Ph.D. Dissertation of Medicine

Single cell transcriptome atlas of human lung airway organoid

인간 폐 오가노이드의 단일세포 전사체 아틀라스

August 2023

College of Medicine
Seoul National University
Biomedical Science Major

이 우 찬

Single cell transcriptome atlas of human lung airway organoid

김 종 일

Submitting a Ph.D. Dissertation of Medicine

August 2023

College of Medicine
Seoul National University
Biomedical Science Major

이 우 찬

Confirming the Ph.D. Dissertation written by
이 우 찬
August 2023

Chair	정 준 호	(Seal)
Vice Chair	김 종 일	(Seal)
Examiner	정 진 행	(Seal)
Examiner	조 영 재	(Seal)
Examiner	이 동 성	(Seal)

Abstract

Keyword : scRNA-seq, organoids, lung, stem cell, differentiation

Student Number : 2016-21999

Despite its importance, the lung is one of the least understood organs in the human body, neither in terms of its development nor in terms of the characteristics of the cells that comprise it. In contrast to the simple cellular composition of the alveoli, which are important for gas exchange, the airways are made up of a wide variety of cells, and new cells are still being discovered. It is also becoming clear that the cells that make up the airways are not just air passages, but also play an important role in the immune system. The main reason why airway cells have been difficult to study in modern cell biology and genomics studies, including single-cell transcriptome analysis, is that airway cells make up a small percentage of the total lung, making it difficult to analyze a sufficient number of cells using single-cell transcriptome analysis.

Recently developed organoid models have the advantage of being able to focus on epithelial cells while simulating the differentiation of cells in the body in an in vitro environment. In particular, the human lung airway organoids can be used to culture all the cells in the airway, starting with the basal cells, through repeated division and differentiation. Establishing

organoids derived from adult stem cells of lung tissue, rather than iPSCs or ESCs, also has the advantage of allowing comparison of lung airway organoids from different individuals.

In this study, we established lung airway adult stem cell-derived organoids from 83 donors and performed single-cell transcriptome analysis of approximately 300,000 cells from 27 donor samples. This has resulted in the world's largest organoid database, which can contribute to the research of a wide range of researchers. In addition, the discovery of Basal-2 and TAB cells revealed for the first time the mechanism of the differentiation process of Tuft, Ionocyte, and PNEC cells, which were difficult to study because they exist in very small numbers. I also discovered that immune-primed cells are also present in lung airway epithelial cells, departing from the previous view that only immune cells are involved.

Through this study, I was able to reveal many precise characteristics of lung airway organoids, which are necessary for various lung studies in the future, and show new possibilities for organoid research. I also made a large database publicly available and created a user-friendly website that allows users to compare various single-cell transcriptome data with my lung airway organoid map. I believe that this large-scale data, which did not exist before, has contributed to our understanding of the overall mechanisms of the lung airways by identifying new cells and conditions involved in immunity and subsequent follow-up studies.

Table of Contents

Chapter 1. Introduction

1.1. Human Lung Airway

The human lung is a complicated and crucial organ with the primary role of facilitating the exchange of gases, mostly oxygen and carbon dioxide, between the blood and the environment. This exchange of gases takes place in the alveoli, which are the air sacs that line the inside of the lung. The airway, which is the channel through which air travels, and the alveolar region, which is responsible for gas exchange with the blood, are the two primary components that make up the lung. In order for the lung to efficiently carry out its job, it is roughly separated into these two main components. Both of these components have highly specialized structures and functions that contribute to the overall efficiency and effectiveness of the respiratory system as a whole.

Not only is the airway responsible for acting as a straightforward channel through which air may go, but it also plays an essential part in the process of removing foreign elements from the air that we breathe. During this stage of the process, the air travels through constricting channels that are lined with specialized cells known as goblet cells. These cells secrete a substance known as mucin. Mucin plays an important role in the process of entrapping foreign particles, such as dust and bacteria, and preventing them

from penetrating the sensitive alveolar region. In addition, the synchronized ciliary motions of ciliated cells that border the airway eliminate any pollutants that have become caught there, thereby ensuring that the air that enters the alveoli is clean.

When one first considers the cells that make up the airway, they may not seem as diverse as they actually are. The following types of cells have been identified: basal, club, goblet, ciliated, deuterosomal, tuft, and ionocyte cells (Hewitt and Lloyd 2021). pulmonary neuroendocrine cells have also been identified. Each of these cell types fulfils a unique job and is responsible for a distinct function in the process of keeping the airway's structure and function intact. In contrast to other organs, which have been subjected to a substantial amount of research, the process of cell differentiation in the airways of the lung has received an insufficient amount of attention. There has been a lot of discussion over how distinct cell types differentiate and how they contribute to the general structure and function of the lung, but new cell

types are still being reported.

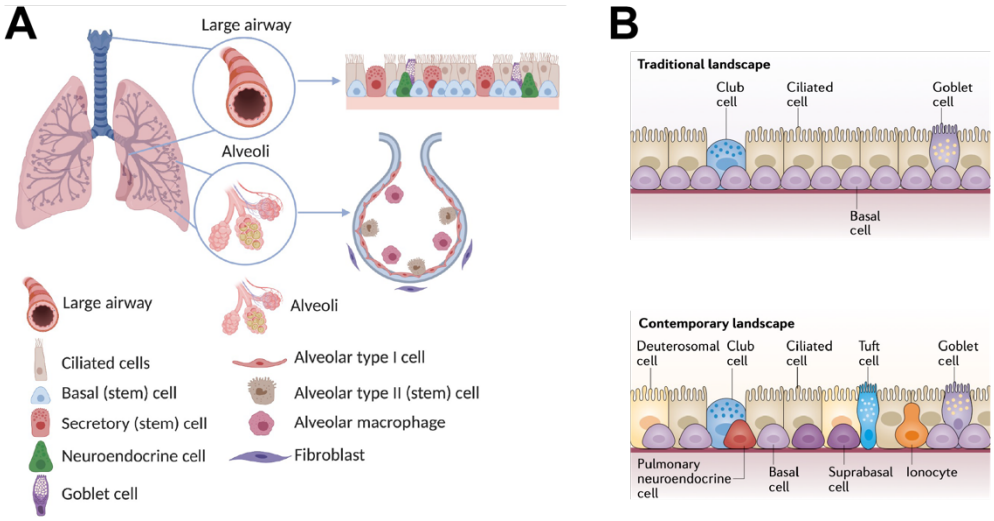

Figure 1. Diverse cell types of human lung airway and alveoli.
(A) Structure of human lung and cell types. **(B)** Cell types which we know in perspectives of traditional (up) and contemporary (down). Modified from (Hewitt and Lloyd 2021)

According to the predominant theory, basal cells in the airway function as stem cells, differentiating into secretory cells such as club or goblet cells, and then further into ciliated cells through the process of deuterosomal cell differentiation. This process of differentiation guarantees that there will always be an adequate supply of specialized cells that are necessary for the airways to carry out their functions in an appropriate manner. In the past, it was thought that the development process of basal cells, which are the stem cells of the airway, was completely distinct from that of alveolar type 2 (AT2) cells, which are the stem cells of the alveolar region. However, recent research has shown that these two types of stem cells actually undergo

the same differentiation process (Kadur Lakshminarasimha Murthy et al. 2022; Basil et al. 2022). On the other hand, new research has demonstrated that transitions between these cells are possible, and that club cells in the post-differentiation stage can revert back to basal cells under homeostatic conditions.

New lines of inquiry have been made possible as a result of recent breakthroughs in our comprehension of the intricate dynamics at play between the various cell types that make up the human lung airway. Scientists have high hopes that they will be able to get a greater understanding of lung development, homeostasis, and the response to damage and disease if they are successful in discovering the exact differentiation lineage of the human lung airway. Asthma, chronic obstructive pulmonary disease (also known as COPD), and lung cancer are only some of the lung conditions that could be helped by the use of new therapeutic approaches, which could be made possible by the knowledge presented here. As our knowledge of the airways in the human lungs continues to expand, so does our capacity to more accurately diagnose, treat, and prevent disorders related to the respiratory system.

Attaining a more profound understanding of the air exchange processes, such as breathing, coughing, and sneezing, is critical for enhancing assessments related to risks affecting respiratory health. The abundance of literature that has been published over the years has enriched our comprehension of airflow dynamics and the transport of particles in the

human lung. It is essential to understand the transport, movement, and deposition of particulate matter in the terminal airways, as this forms the cornerstone for forecasting and potentially mitigating respiratory diseases.

Human activities encompassing the combustion of fossil fuels in automobiles, industrial operations, burning of coal for thermal energy, biomass combustion, among other sources, contribute to the generation of a substantial volume of particulate matter (PM) (Chen et al. 2017). The inhalation of these particles through the nose or mouth sets into motion a complex journey as they traverse through the convoluted, branching structure of the airways. During the course of breathing, inhaled pollutants can either settle within the respiratory tract or escape, influenced by various mechanisms (Laumbach and Kipen 2012). Moreover, there are instances where particles are absorbed through the epithelial layer, instigating a series of inflammatory responses that can culminate in diverse respiratory ailments.

The impact of particulate matter on respiratory health, including its toxic and carcinogenic potentials, is closely linked to the duration for which these particles reside in different sections of the respiratory system. Furthermore, the deleterious effects of both soluble and insoluble aerosols can be magnified in the terminal airways across various lung lobes, depending on their inherent toxic properties. The terminal and alveolar airways, characterized by impaired cilia and the absence of mucociliary clearance mechanisms, are unable to expel these toxic particles effectively.

Research efforts spanning several decades have made significant

strides in improving our understanding of airflow and particle transport within idealized airway models. Furthermore, recent technological advancements have paved the way for the creation of anatomically accurate models based on computed tomography (CT) scans. A plethora of studies has deployed a diverse range of computational techniques and anatomical models to analyze airflow and particle transport in human airways (Bustamante-Marin and Ostrowski 2017).

Lung cancer has emerged as an immense challenge in the realm of public health, claiming around 1.4 million lives annually starting from the year 2008 (Jemal et al. 2011). The gravity of the situation is accentuated by the fact that this particular form of cancer is on the trajectory to become the most common cause of cancer-related fatalities in Western industrialized nations. A stark reminder of lung cancer's severity is its comparatively higher mortality rate when placed alongside other cancers such as breast, prostate, colon, liver, and kidney (Petersen 2008).

Additionally, the spectrum of respiratory illnesses is not limited to lung cancer. There are other debilitating lung diseases, including asthma, emphysema, pulmonary fibrosis, and pneumonia, which have an extensive global impact. Millions of individuals across various continents have had their lives compromised due to these diseases. In the United States, for example, these collective lung ailments hold the dubious distinction of being the third most common cause of death.

In the face of these alarming statistics, the medical community has

been relentlessly seeking effective treatment modalities. Among various approaches, aerosolized therapeutics have gained prominence as the foremost method in combating lung diseases. The utilization of aerosol therapy has seen a substantial upswing in recent years, which can be attributed to the multiple benefits it confers. Notably, it is not just restricted to the treatment of pulmonary diseases; its scope extends to non-pulmonary ailments as well.

One of the fundamental aspects of aerosol therapy is the utilization of the oral pathway for drug administration. This method proves to be highly efficacious as it facilitates a rapid, non-destructive, and non-invasive conduit for drug delivery into the circulatory system. This is made possible owing to the lungs' anatomical attributes, specifically their extensive surface area and proficient blood perfusion. Such attributes are integral in ensuring the prompt and effective dispersion of medication.

The focused delivery of medication to the lungs has its unique set of advantages, especially when treating lung diseases. By employing aerosolized therapeutics, healthcare professionals can administer lower doses of medication directly to the lungs. This focused delivery is highly beneficial as it not only minimizes the quantity of the drug required but also significantly reduces systemic side effects. This is in stark contrast to other drug delivery methods that circulate medication through the entire body and, as a result, may entail a host of side effects.

However, it's critical to understand that for aerosol therapy to be optimally effective in terms of both efficacy and minimizing side-effects,

precise targeting is imperative. In other words, the transfer process must be meticulously calibrated to ensure that the medication is delivered directly to the specific site within the lungs where it is needed most. This necessitates an in-depth understanding of the lung's anatomy and the pathology of the disease being treated, along with the development of innovative technologies to enable targeted delivery.

In summary, lung diseases, with lung cancer at the forefront, present formidable challenges to global health. Aerosolized therapeutics have emerged as a promising approach in addressing these challenges. However, for these therapies to realize their full potential, there must be a continued emphasis on precision and targeting in drug delivery, backed by technological advancements and a thorough understanding of the underlying disease mechanisms.

Pulmonary diseases have been responsible for a staggering number of deaths, making respiratory drug delivery an alluring domain for research over the past several decades. Scholars and researchers have delved into a wide range of topics encompassing pulmonary drug delivery. A wealth of literature covers diverse aspects including the efficacy of drug delivery devices, the development and formulation of micro and nano-particles specifically tailored for drug delivery, and innovative approaches like magnetic drug targeting (Ali, Mazumder, and Martonen 2009; Zhang, Gilbertson, and Finlay 2007; Hasenpusch et al. 2012).

Understanding particle deposition within the human respiratory

airway is a cornerstone for gauging risks associated with exposure to harmful aerosols, as well as for the assessment and creation of drug delivery devices. This knowledge is also vital for advancing treatment protocols for chronic respiratory diseases. However, charting the course for such understanding is laden with challenges. One major hurdle is devising a realistic airway model that can accurately portray particle deposition.

Over the past ten years or so, researchers have made concerted efforts to develop comprehensive models of lung airway geometries. Among the proposed models, Weibel's model stands out as probably the most renowned (Haefeli-Bleuer and Weibel 1988). It is characterized by its simplicity and symmetry. The introduction of Weibel's model marked a watershed moment in the evolution of idealized respiratory airway models and it has since been a mainstay in an array of numerical and experimental studies (Lambert et al. 2011).

In addition to Weibel's model, another significant contribution to the pool of morphometric data is attributed to Raabe et al., which offered invaluable insights regarding the human bronchial tree. The baton was passed on from Weibel and Raabe's models, and further research endeavors were undertaken by various scholars to refine and expand on idealized respiratory airway geometries.

Nonetheless, capturing the complex nuances of the respiratory airway geometry, which is paramount for understanding particle deposition, is beyond the purview of simple models. The complexity of the respiratory

airway's geometry is a determining factor in particle deposition, and basic models are not equipped to elucidate realistic patterns of particle deposition.

Recent strides in medical imaging technologies, such as computed tomography (CT) and magnetic resonance imaging (MRI), have furnished researchers with the means to reconstruct the respiratory airway tree in all its anatomical intricacies (Ismail, Comerford, and Wall 2013). These realistic airway models have become instrumental in computational fluid dynamics (CFD) and have been applied in experimental studies investigating particle deposition within the respiratory system for numerous years (Lambert et al. 2011).

However, this technique is not without its limitations. The capturing of bronchial airway branches is confined to the larger ones, as the resolution of CT images is compromised by factors such as the vibrations stemming from the subject's heartbeat and the inherent resolution constraints of CT-scanners. Given the impracticality of modeling the distal airway generations using CT images, the application of suitable mathematical models is imperative (Ismail, Comerford, and Wall 2013).

In light of this, a plethora of studies employing analytical models have emerged, addressing various dimensions including particle depositions, gas mixing, and flow distributions. These studies have adopted an interdisciplinary approach, bridging the gap between theoretical modeling and practical applications in respiratory drug delivery (Gouinaud et al. 2015; Katz et al. 2011).

In conclusion, respiratory drug delivery research is a multifaceted domain, encompassing the development of drug delivery devices, formulation of particles, and advanced targeting techniques. The creation of realistic airway models and understanding of particle deposition is central to this research, requiring the integration of advanced imaging, computational fluid dynamics, and mathematical modeling. With the rapid progress in technology and collaboration across disciplines, the potential for breakthroughs in the treatment and management of pulmonary diseases is immense.

As has been demonstrated in the preceding discussions, there has been an effusion of scholarly activity over recent decades where myriad researchers have explored particle deposition and pulmonary drug delivery from a variety of modeling angles. It has become unequivocally clear that engaging in research in this domain is not just a matter of academic interest, but an imperative necessity. The primary aim of this study is to present an exhaustive review that takes the reader through the multifaceted dimensions of respiratory airway modeling and the leaps and bounds that have been made in this field over the past decades.

In order to present this information in an organized and coherent manner, the content of this review has been meticulously categorized into two principal segments - the physical domain and the numerical-mathematical domain. Within these two overarching categories, the diverse modeling approaches that have been adopted are further classified in a chronological

manner, painting a vivid picture of the evolution and advancements in this area over time.

In the subsequent sections, we shall embark on a journey to lay down the foundational knowledge concerning the human respiratory tract. It is essential to first build an understanding of the intrinsic anatomical complexities that characterize the respiratory tract as this forms the bedrock upon which respiratory airway modeling is built.

In the section focusing on the physical domain, we shall delve into the world of airway casts. These physical representations have played an instrumental role in elucidating the intricacies of the airway structures and have provided invaluable insights that have shaped the trajectory of research in respiratory drug delivery. The examination of airway casts is critical, as it serves as a stepping stone to the development of more advanced modeling techniques.

Moving on from physical airway casts, we shall venture into the realm of idealized models. These models, while inherently simplistic, have often served as the building blocks for more complex analyses. By approximating the human airways through simplified structures, idealized models have allowed for a conceptual understanding and provided a platform for hypothesis testing in the early stages of research.

Subsequently, we will explore the exciting frontier of image-based models. As technology has advanced, the ability to create highly detailed and anatomically accurate models through medical imaging has revolutionized

the field. These image-based models are the epitome of precision and have allowed for an unprecedented level of insight into the workings of the respiratory tract.

Concluding the physical domain, the baton is passed to the numerical-mathematical domain. In this section, we will immerse ourselves in the sophisticated world of mathematical models and simulations. These models have been particularly instrumental in circumventing some of the limitations inherent in physical models. Through the power of computation and numerical analyses, researchers have been able to simulate complex airflow patterns and particle interactions, which have implications for drug delivery strategies.

In summary, this review aspires to serve as a comprehensive guide through the labyrinth of respiratory airway modeling, unraveling its complexities and highlighting the strides that have been made. Through an amalgamation of physical constructs and numerical wizardry, the field continues to evolve, and it is this very interplay that holds the promise of breakthroughs in the future.

1.2. Differentiation of human lung airway cells

Human Airway Epithelial (HAE) cellular models have garnered extensive attention and are extensively employed in investigative studies focusing on airway physiology and diseases. These models serve as

quintessential tools to shed light on various aspects of airway biology. A specific approach involves utilizing in vitro expanded and differentiated primary HAE cells that have been harvested from patients. This particular approach is lauded for its close approximation to the native human airway. Furthermore, it presents itself as a more expeditious and cost-effective alternative compared to models based on induced pluripotent stem cells (iPSCs).

However, a salient limitation that accompanies primary HAE models is the restricted proliferative lifespan that these cells exhibit when cultured in vitro. This constraint has been the impetus for a plethora of research efforts dedicated to deciphering the determinants governing the proliferation and differentiation of HAE cells. These investigative pursuits aim to gain insights into these cellular processes in both in vivo and in vitro settings.

There is a particular emphasis on conditionally reprogrammed cells (CRC) models, which have been instrumental in facilitating extended in vitro proliferation and differentiation of HAE cells. This is a groundbreaking development as it circumvents the previously mentioned limitation regarding limited proliferative lifespan.

The conducting airways serve as a vital conduit for airflow and encompass the anatomical regions extending from the trachea to the proximal segments of the small bronchioles. This segment of the respiratory system is characterized by a specialized lining known as the pseudostratified respiratory epithelium (PSE) (Berical et al. 2019).

Pertaining to morphogenesis, which is the biological process that gives an organism its shape, the development of the respiratory epithelium that lines the airways is an intricate and gradual process. It is essential to understand that this epithelial formation is the result of a finely orchestrated series of commitments and differentiation stages.

Initially, the definitive endoderm, one of the three germ layers that form during early embryonic development, begins the cascade. This germ layer is critical as it ultimately gives rise to various internal structures, including components of the respiratory system.

As development progresses, the definitive endoderm transitions into what is known as the anterior foregut endoderm. This is a crucial intermediate stage, marking the specialization of cells that will contribute to the formation of the respiratory system among other structures.

Following this transition, there is a further commitment to lung epithelial progenitors. These are the precursor cells that have the capability to differentiate into various cell types that constitute the respiratory epithelium. They are critical players in the development of the lungs and airways.

Ultimately, these lung epithelial progenitors undergo a final differentiation into either distal airway progenitors or alveolar progenitors. The distal airway progenitors contribute to the formation of the smaller airways deeper in the lungs, while the alveolar progenitors are involved in the development of the alveoli, the tiny air sacs where gas exchange occurs.

In summary, the development of the pseudostratified respiratory

epithelium lining the conducting airways is a multi-step process that involves a sequential commitment of cells originating from the definitive endoderm, progressing through stages as anterior foregut endoderm, lung epithelial progenitors, and culminating in the differentiation into distal airway or alveolar progenitors. This coordinated cascade is fundamental to the proper formation and functionality of the respiratory system (Morrisey and Hogan 2010).

The utilization of primary human airway epithelial (HAE) cell culture derived from patients holds a significant advantage, particularly when it comes to studying the pathogenesis of inherited respiratory diseases. The foremost benefit lies in the fact that these cells inherently possess the genetic alterations pertinent to the disease in question. This eliminates the necessity for laboratory-based genetic modification of the cells, thereby offering a more authentic representation of the disease state.

Additionally, not only do these primary HAE cells harbor genetic modifications, but they also embody relevant epigenetic changes. Epigenetic modifications are alterations in gene expression that do not involve changes to the underlying DNA sequence. Such changes are critical in understanding various disease processes. For instance, it has been demonstrated that when primary airway cells are cultured from patients suffering from Chronic Obstructive Pulmonary Disease (COPD), these cells exhibit epigenetic changes in their DNA. Furthermore, these cells have been observed to re-enact epithelial reactions to IL-13, a cytokine involved in the inflammatory

response, which is indicative of the disease state (Song et al. 2017).

Another important aspect is that cells cultured in vitro from patients with asthma maintain certain characteristics of abnormal airway epithelium even while in culture. This retention of disease-specific traits is exceedingly valuable for studies that focus on airway disorders influenced by environmental factors. Given that asthma is a condition where environmental triggers play a significant role, having a model that retains characteristics of the distorted airway epithelium is crucial for understanding the interactions between genetic predispositions and environmental triggers (Hiemstra et al. 2018; Gras et al. 2012).

In summary, primary HAE cell cultures obtained from patients are an invaluable resource for studying the pathogenesis of inherited respiratory diseases. Their inherent genetic and epigenetic modifications, coupled with the ability to mimic in vivo responses, make them an authentic and reliable model for understanding disease mechanisms and exploring potential therapeutic interventions.

Conversely, induced pluripotent stem cells (iPSCs) exhibit certain limitations that make them less desirable for certain applications as compared to primary human airway epithelial (HAE) cells (Gomes et al. 2017; Calvert and Ryan Firth 2020). One of the notable shortcomings of iPSCs is the loss of epigenetic marks. Epigenetic marks are vital in the regulation of gene expression and are essential in understanding various biological processes and diseases. The loss of these marks in iPSCs may lead to an inaccurate

representation of the disease state, and this can be a significant limitation in the study of pathogenesis.

Another major concern regarding the use of iPSCs is their propensity for tumorigenesis and genetic instability (Awatade et al. 2018; Hynds, Bonfanti, and Janes 2018; Hiemstra et al. 2018). These cells have the potential to form tumors, and their genetic material can be unstable. This poses a substantial hurdle, especially in the fields of gene therapy and tissue regeneration, where the goal is often to replace or repair damaged tissues without introducing further complications such as tumor formation.

Furthermore, the financial aspect and time investment associated with the use of iPSCs should not be overlooked. The costs involved in iPSC generation and differentiation are considerably high. In addition to the financial burden, the process of directing iPSCs to differentiate into airway progenitors is time-consuming. This long-term derivation process can be a drawback, especially in research settings where time is a critical factor.

In light of these limitations - the loss of epigenetic marks, the potential for tumorigenesis and genetic instability, coupled with the high costs and extended time required for derivation - iPSCs may be considered less favorable compared to primary HAE models for certain applications. Primary HAE models, on the other hand, offer a more accurate representation of the in vivo state, without the concerns of tumorigenesis, and are both quicker and more cost-effective to establish, making them a more appealing choice for studies related to airway physiology and diseases.

There is a variety of culture platforms available for the cultivation and differentiation of primary human airway epithelial (HAE) cells. Among these, the adherent two-dimensional (2D) culture is one of the most established and traditional methods. In this approach, HAE cells are cultivated in a monolayer, where they are entirely submerged in a growth medium. It's important to note that this environment predominantly promotes an undifferentiated HAE phenotype. As a result, adherent 2D cultures are primarily utilized for the proliferation of HAE basal cells.

One of the challenges in 2D culture is ensuring that the HAE cells properly adhere to the culture vessels, which is critical for their growth and proliferation. To enhance adhesion, researchers have developed strategies that include the use of a collagen coating. By applying a thin layer of collagen, one can improve the adherence of the cells to the surface of the culture vessels. Additionally, another alternative is to use a thicker layer of collagen gel. The utilization of collagen not only aids in cell adhesion but can also provide a more supportive environment for the cells, somewhat mimicking the extracellular matrix found in vivo.

Despite the widespread use of adherent 2D cultures, it's worth mentioning that this method has limitations when it comes to simulating the complex three-dimensional structure of human airway tissues. Consequently, researchers have been exploring alternative culture methods that more closely resemble the natural environment of HAE cells. This includes the development of three-dimensional (3D) culture systems which allow cells to

grow in multiple layers, offering a more physiologically relevant model for studying airway epithelial biology and disease mechanisms. However, the adherent 2D culture remains an essential tool, particularly for the initial expansion of HAE basal cells before they may be further differentiated in more complex culture systems (Fulcher et al. 2005; Jorissen and Willems 2004).

The advent of the Conditionally Reprogrammed Cells (CRC) method for culturing Human Airway Epithelial (HAE) cells has marked a significant advancement in the field. This novel approach has remarkably enhanced the proliferative efficiency of primary HAE cell cultures. Consequently, it is now feasible to obtain substantial quantities of HAE cells, which can then be subjected to genetic modifications, cryopreservation, and even employed in cutting-edge tissue engineering applications.

Nevertheless, it is crucial to acknowledge that the CRC HAE model, despite its groundbreaking contributions, is not without flaws. There is still ample room for refinement to enhance both its robustness and the reproducibility of the results it yields. One of the aspects that necessitates further investigation is a meticulous comparison of the existing CRC culturing variations, specifically the CRC HAE mono-culture and co-culture regimens. By examining these methodologies, it is imperative to evaluate how closely they emulate the native HAE epithelium in terms of physiological attributes. Such evaluation is instrumental in determining the suitability of each variation for different kinds of experimental undertakings.

Moreover, delving deeper into the factors that influence the differentiation and functioning of HAE cells is fundamental to the advancement of the CRC HAE model. A pivotal component in this regard is understanding the interplay between the extracellular matrix (ECM) and various tissue types, such as fibroblasts. Understanding the nuances of this interaction can provide critical insights into how to optimize the culture environment to promote HAE cell growth and differentiation.

Another area that warrants attention is the identification of factors that contribute to variability in culture outcomes from one donor to another. By pinpointing these factors, researchers can better predict the success of cultures derived from different individuals. This, in turn, can pave the way for enhancing the reproducibility and reliability of the CRC culture approach.

In conclusion, the CRC HAE model represents a significant step forward in the culturing of HAE cells. However, continual refinements and a deeper understanding of the underlying biological processes are essential for maximizing the potential of this innovative approach. Through such efforts, the CRC HAE model could become an even more powerful tool for research and therapeutic applications.

1.3. Organoids

In 2009, Hans Clevers was the first person to present organoids to the scientific community as a revolutionary in vitro model system (Sato et al. 2009). With the help of this cutting-edge method, researchers are able to keep

and examine a wide variety of epithelial cells derived from real organs in an in vitro environment for extended periods of time. In terms of gaining an understanding of organ development, tissue regeneration, and disease modelling, organoids have shown to be extremely useful. Even though organoids are not a perfect system, they have many advantages over the more straightforward 2D cell line culture methods. One of these advantages is that organoids are able to mimic a much more heterogeneous system in vitro, which results in the collection of data that is more physiologically relevant.

Adult stem cell-derived organoids (also referred to as ASC organoids) and pluripotent stem cell-derived organoids (also referred to as PSC organoids), which include both embryonic stem cells and induced pluripotent stem cells, are the two primary classifications that can be used to classify organoids (Kim, Koo, and Knoblich 2020). PSC organoids have the advantage of being able to construct organoids for practically all organs, which provides a broad platform for researching organ development and the disease pathways that underlie diseases. Nevertheless, during the process of proliferation, they cannot be continuously maintained, which is a significant limitation. Therefore, in a real experimental setting, the laborious process of differentiating PSC cultures needs to be carried out whenever it is required, which can cause potential delays in the experiment and in the collection of

data.

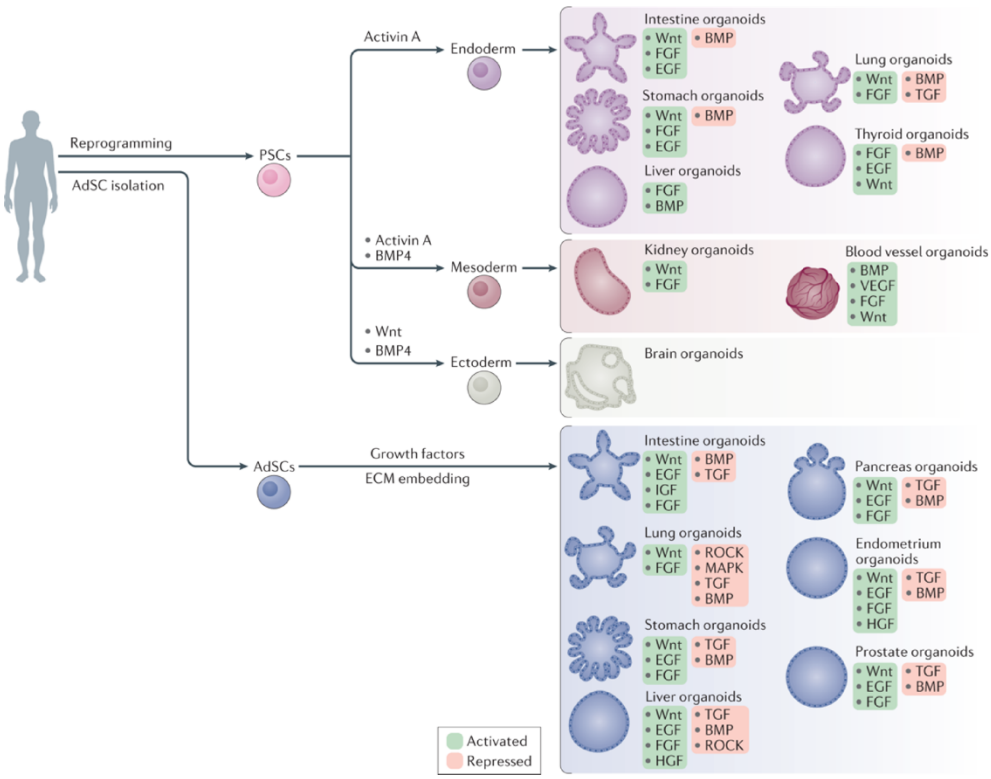

Figure 2. Establishment of human PSC organoids and ASC organoids. Schematic figure from (Kim, Koo, and Knoblich 2020)

Organoids derived from human pluripotent stem cells (PSCs) are produced through a series of guided differentiation protocols which are designed to emulate the developmental processes that have been identified in earlier research, both in laboratory conditions and in living organisms. Considering that our understanding of human development is rather scant, initial studies that aspired to create organoids bearing resemblance to human tissues principally relied on drawing analogies from the development of mice.

Theoretically, for the generation of an organoid that closely mimics

a human organ, it is imperative that the entire gamut of the organ's developmental process, commencing from PSCs, be replicated with utmost fidelity. However, in practical terms, achieving this level of precision in a laboratory setting is exceedingly challenging. The principal reason behind this is the sheer complexity and dynamism of embryonic development, wherein a myriad of biochemical signals are responsible for driving cell differentiation, as well as the three-dimensional assembly and organization of tissues. These signals need to be delivered at exact locations, times, and concentrations, an orchestration that is incredibly intricate.

Interestingly, cells have an inherent tendency to follow what can be termed as a semi-autonomous differentiation pathway, a phenomenon observed both in vivo and in vitro. This means that, to some extent, cells are able to navigate the labyrinthine journey of differentiation and assembly into complex structures with minimal external guidance. Capitalizing on this intrinsic propensity of cells, researchers employ three primary categories of protocols to cultivate functional organoids in vitro.

Firstly, the 'guided differentiation' approach is employed where scientists attempt to approximate the natural biochemical cues by adding specific growth factors and signaling molecules to the culture at different stages of development. Secondly, some protocols rely on the 'self-organization' principle, where minimal external guidance is provided, and the cells are largely left to their own devices to spontaneously differentiate and organize into complex structures. Lastly, a hybrid approach can be used,

which combines elements of guided differentiation and self-organization, often yielding organoids that are more representative of the in vivo tissue structures.

The use of organoids in biomedical research heralds a new era of possibilities. These miniature organ-like structures offer unprecedented insights into human development, disease progression, and drug response. Despite the challenges, the field is rapidly advancing and holds immense promise in revolutionizing our understanding of biology and improving healthcare outcomes. The continuous refinement of protocols and incorporation of new knowledge in human development will play a crucial role in the maturation of this exciting area of research.

ASC organoids, on the other hand, are created by cultivating genuine organ stem cells outside of the body in a medium that contains the relevant niche elements. This allows them to keep a stable balance between stem cell division and differentiation. Even though this method necessitates the locating of stem cells as well as the establishment of suitable culture conditions, once these have been determined, the experiments can be carried out much more quickly, which makes the process of data collection and analysis much more effective. Despite this, all that has been accomplished with ASC organoids is the establishment of organoids that are predominantly derived from endoderm organs, which restricts their applicability to a more limited variety of organ systems.

ASC-based human airway organoids were described for the first time

in 2019 by the group led by Hans Clevers (Sachs et al. 2019). They were able to maintain airway stem cells (basal cells) for more than six months and differentiate them into practically all types of airway cells by modulating cell signalling pathways including as Wnt, EGF, BMP, and TFGb. This allowed them to keep airway stem cells as basal cells. The research of airway illnesses, the discovery of new drugs, and the creation of personalized medicine have all seen substantial advances as a result of this recent breakthrough. To fully recapitulate the in vivo airway cell diversity, however, the proportions of pulmonary neuroendocrine, tuft, and ionocyte cells are quite limited, which presents a problem.

Rooted in the core principle that biological processes and mechanisms are preserved across evolutionary lines, biomedical research has conventionally concentrated on employing a select group of model organisms. The favored choices are usually resilient, characterized by rapid growth, and have the capacity to yield large litters in brief time spans, all while being cost-effective for lab cultivation. The utilization of a well-established model organism offers researchers a plethora of advantages, encompassing a wealth of information regarding its development and physiology, an extensive repertoire of experimental methodologies, and a host of community-driven resources such as reagents and databases tailored specifically for the species in question.

Among the model organisms, certain ones like yeast (Saccharomyces cerevisiae), fruit flies (Drosophila melanogaster), and mice (Mus musculus),

boast a storied legacy in the realm of scientific research. However, there are also more recent additions to the roster of model organisms which, despite their relatively shorter history, have been instrumental in widening the scope of our understanding by providing us with an innovative arsenal of experimental instruments. For instance, the nematode worm, Caenorhabditis elegans, stands out due to its highly reproducible cell lineage, which grants researchers an unparalleled opportunity to scrutinize cell fate decisions in great detail (Sulston et al. 1983). Another noteworthy example is the zebrafish, Danio rerio, which holds the distinction of being the pioneering vertebrate system wherein comprehensive loss-of-function mutagenesis screenings were conducted (Haffter and Nusslein-Volhard 1996; Mullins et al. 1994). The worm, Caenorhabditis elegans, gained prominence as a frequently employed experimental model from the mid-1970s onwards, while zebrafish became an integral part of large-scale genetic studies starting in the 1990s (Nusslein-Volhard 2012).

It is imperative, however, to recognize that despite the wealth of information that can be gleaned from animal models, they do not wholly emulate human physiology. This caveat underscores the need for alternative experimental models that can better replicate human biological processes. This is where human organoids, which surfaced in the early 2010s, assume significance (Takebe et al. 2013; Lancaster et al. 2013; Sato et al. 2011). Human organoids represent a cutting-edge experimental model that strives to bridge the chasm between animal models and human beings.

Composed of stem cells, human organoids have the ability to mimic the architecture and functionality of human organs on a diminutive scale. This enables scientists to carry out intricate studies that simulate human physiological conditions more closely than conventional animal models. They have become particularly valuable for drug testing, understanding disease mechanisms, and even exploring possibilities in regenerative medicine.

In summary, while traditional model organisms have and continue to play a vital role in biomedical research, the advent of human organoids signifies a momentous leap forward, offering a more representative and human-centric approach to studying biology and diseases. The continuous evolution of experimental models, coupled with an integration of classic and novel approaches, holds the key to unlocking a deeper understanding of the complex tapestry of life.

After the successful development of organoids based on human stem cells, these miniature organ-like structures have been employed extensively to model a plethora of human diseases. Human organoids have proven to be invaluable in the study of a wide range of diseases, including but not limited to, infectious diseases, hereditary genetic disorders, and various forms of cancer.

In the realm of infectious diseases, organoids have provided crucial insights into the complex interactions between pathogens and human cells. They serve as miniature replicas of human organs that can be infected under

controlled conditions, allowing for the detailed study of the mechanisms by which pathogens invade and spread within human tissues.

When it comes to inheritable genetic disorders, human organoids have been instrumental in understanding the underlying genetic factors and their manifestations. They allow for the simulation of the disease in a controlled environment, which helps in unraveling the mechanisms that contribute to the progression of these genetic disorders.

In the context of cancer research, organoids offer a unique platform for modeling the various stages of cancer development and progression. Through the use of cancer-derived organoids, researchers can study tumorigenesis and metastasis in a system that closely mimics the in vivo environment. This has far-reaching implications for the development of new therapeutic strategies and drugs.

Furthermore, the breakthroughs in genetic engineering technologies have significantly augmented the potential applications of organoids in disease modeling. With tools such as CRISPR/Cas9, it is now possible to engineer organoids with specific genes or mutations that are associated with particular diseases. One of the most promising approaches involves deriving organoids from cells isolated from healthy donors and then introducing pathogenic genes and mutations into these organoids. This enables researchers to conduct human genetic studies within controlled genetic backgrounds, which is critical for isolating the effects of specific genes or mutations.

Additionally, this combination of organoid technology with genetic engineering opens up the possibility of personalized medicine. By creating organoids derived from a patient's own cells and introducing specific mutations, it is feasible to study how an individual's unique genetic makeup can affect disease progression and response to treatment.

There are various approaches to the cultivation of organoids in the laboratory, each of which has its own set of benefits and downsides; the two primary approaches that are utilized for airway organoids are described below. The initial step is to develop a three-dimensional culture by means of Matrigel domes, which is the approach that is utilized for the majority of ASC organoids (Sachs et al. 2019). This method permits the growth of cells in a three-dimensional environment that is highly reminiscent of the conditions seen in living organisms, and as a result, it fosters the development and differentiation of a wide variety of cell types. Creating an air-liquid interface culture (also known as an ALI culture) can also be done with the assistance of a specialized culture device known as Transwell. This apparatus is made up of a porous membrane. This method has the advantage of producing a setting in which cells can have direct contact with air, which is analogous to the conditions that exist within the lungs of living organisms. In addition, because the apical side of the cells faces upward, it is possible to establish an environment that is more similar to that found in vivo for research such as viral infection. This capability is very helpful for research into the mechanics of viral entrance, replication, and host response in the context of a

physiologically appropriate setting.

In the history of genetic engineering, a significant milestone was reached in 1987 when Mario Capecchi demonstrated that homologous recombination events occur with high frequency in mouse Embryonic Stem Cells (ESCs) (Thomas and Capecchi 1987). Homologous recombination is a process where genetic information is exchanged between two similar or identical strands of DNA, which was observed to be an extremely rare event in human Pluripotent Stem Cells (PSCs) (Liu and Rao 2011; Porteus and Baltimore 2003).

Subsequent research discovered that double-strand breaks (DSBs), which are ruptures that span both strands of the DNA double helix, could facilitate homologous recombination events as part of the cell's intrinsic DNA repair mechanism. This finding ignited the quest for developing a highly efficient method to generate DSBs at specific genomic locations, with the aim of harnessing the cell's repair machinery to introduce desired genetic modifications such as repairing a sequence or introducing a pathogenic mutation into the targeted site.

Initial attempts to achieve this goal involved the use of meganucleases, which are endonucleases known for their long recognition site (12-40 base pairs) (Bibikova et al. 2003). Following this, researchers experimented with zinc finger nucleases and transcription activator-like effector nucleases, with varying degrees of target specificity and activity (Miller et al. 2011). While these technologies did pave the way for

advancements in genetic engineering, they were not without their limitations.

The advent of Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) and CRISPR-associated protein 9 (Cas9) endonuclease technology heralded a new era in genetic engineering, making a wide array of genetic manipulation techniques accessible to researchers worldwide (Cho et al. 2013; Mali et al. 2013; Cong et al. 2013; Wiedenheft, Sternberg, and Doudna 2012).

Unlike previous technologies, the Cas9 endonuclease operates in conjunction with a guide RNA sequence (gRNA) that directs it to the desired genomic sequence to generate a DSB. This element of the CRISPR-Cas9 system adds a high degree of versatility and ease of application, given that it allows researchers to specify the exact genomic location that the system will target (Pickar-Oliver and Gersbach 2019).

Since its inception, the CRISPR-Cas9 technology has revolutionized genetic engineering, allowing for unprecedented precision and efficiency in introducing genetic modifications. It has opened new doors in the realm of biomedical research, offering potential solutions for treating genetic disorders and advancing our understanding of human genetics. The system's simplicity and flexibility have made it an indispensable tool for scientists in the field, contributing greatly to rapid advancements in genetic research and therapy.

1.4. Single Cell RNA sequencing

The application of the next-generation sequencing (NGS) technology is required for the cutting-edge transcriptome analysis approach known as single-cell RNA sequencing (scRNA-seq). This method is found in the field of genomics (Papalexi and Satija 2018). Our understanding of cellular heterogeneity, gene expression patterns, and cell lineage relationships within complex tissues has been completely transformed as a result of this discovery. The primary distinction between traditional bulk RNA sequencing and single-cell RNA sequencing, also known as single-cell RNA sequencing (scRNA-seq), is that the former method extracts RNA from whole cells or tissues, whereas the latter method analyses the transcriptome of individual cells and enables the simultaneous analysis of hundreds of thousands to millions of cells at a single-cell resolution.

Plate-based and droplet-based procedures are the two most common ways to conduct scRNA-seq experiments, respectively. Creating a library from each well of a plate requires the use of the plate-based technique, which entails inserting individual cells into the small wells of a plate. This method has the potential to provide more sensitivity and depth of coverage for each cell, but it is restricted in the number of cells that it can analyze in a single experiment. As a result, it is best suited for investigations of a more limited

scope.

The droplet-based approach, on the other hand, makes use of microfluidics to produce extremely small droplets, each of which encapsulates a single cell, and then builds libraries from the cells that are encased within the droplets. In comparison to plate-based methods, this technique enables the examination of a significantly greater number of individual cells, albeit at the expense of a slightly lower data quality for each of those cells. Despite this trade-off, droplet-based scRNA-seq is more cost-effective and has gained widespread adoption due to its high throughput capacity and lower associated costs.

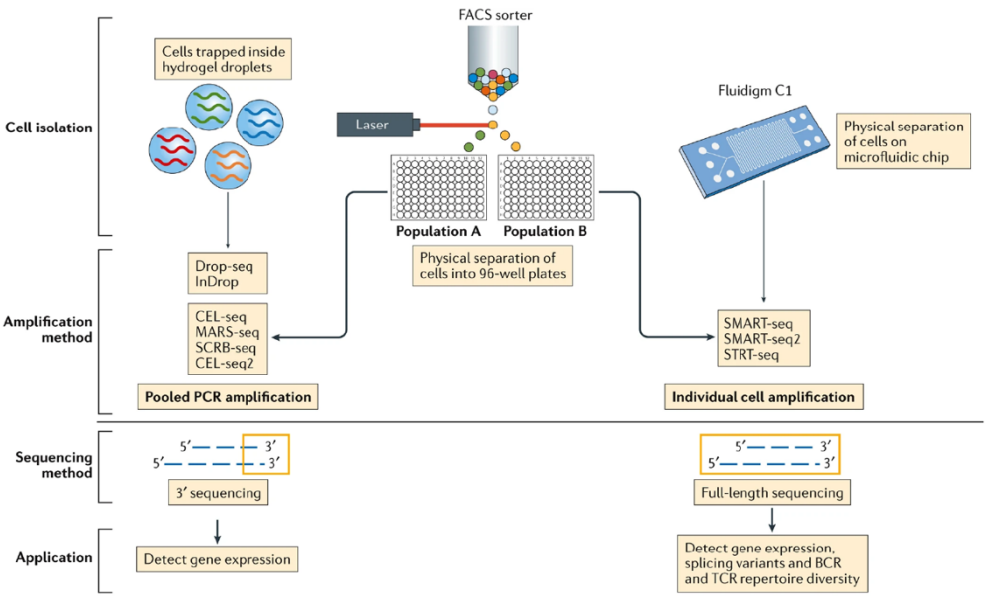

Figure 3. Overview of scRNA-seq technologies.
Two major approaches of scRNA-seq; droplet-based (left) and plate-based (right). Images from (Papalexi and Satija 2018)

Recent developments in scRNA-seq technology and computational

methodologies have broadened the range of applications that are potentially feasible for this methodology. These include the identification of rare cell populations, the investigation of cellular differentiation and development, the analysis of transcriptional networks underlying cellular responses to various stimuli, and the dissection of complex cellular interactions within tissues.

Gaining insights into cellular responses necessitates evaluations of either gene or protein expression. With regards to protein expression, multicolor flow cytometry in conjunction with fluorescently conjugated monoclonal antibodies has revolutionized the field by enabling the simultaneous analysis of a small set of proteins across a large number of individual cells. This has become a standard practice in both experimental and clinical research.

An advanced development in this realm is mass cytometry. This technique employs staining of cells with antibodies that are tagged with heavy metal ions, followed by quantitative measurements utilizing time-of-flight detectors. This has exponentially increased the number of proteins that can be assessed, typically ranging from five to ten times more than what was previously possible using flow cytometry. Moreover, mass cytometry has unveiled remarkable levels of diversity and complexity within cell populations that seemed homogeneous at first glance, particularly among immune cells (Giesen et al. 2014).

Nevertheless, a significant challenge remains - the comprehensive analysis of the entire collection of proteins, referred to as the 'proteome,'

expressed by the genome in a single cell.

In light of this challenge, a great deal of research has shifted towards studying protein-encoding mRNA molecules, collectively known as the 'transcriptome'. The expression levels of these mRNA molecules often correlate closely with cellular characteristics and alterations in cellular states. Transcriptomics began with studies on large collections of cells, initially employing hybridization-based microarrays and later transitioning to next-generation sequencing techniques, notably RNA sequencing (RNA-seq). RNA-seq has been invaluable, producing an abundance of data that has continued to propel advances and discoveries in biomedicine. For instance, RNA-seq has been applied to hematopoietic stem cells in order to categorize patients with acute myeloid leukemia into different groups, each necessitating distinct treatment strategies (Ng et al. 2016).

However, one inherent limitation of traditional RNA-seq is that it involves pooling large numbers of cells, which means that the data is averaged across these cells. This pooling prevents a detailed analysis of individual cells or nuclei, which are the fundamental biological units containing the genome.

Single-cell RNA sequencing (scRNA-seq) has emerged as a transformative approach to overcome this limitation. Since the publication of the first study employing scRNA-seq in 2009, there has been a surge in interest and application of this technique (Tang et al. 2009). One of the most notable aspects of scRNA-seq is its ability to characterize RNA molecules within individual cells at a high resolution and on a genomic scale.

Initially, scRNA-seq was primarily conducted by specialized research groups, but it has become increasingly evident that biomedical researchers and clinicians can achieve significant breakthroughs using this potent methodology (Kolodziejczyk et al. 2015; Zeisel et al. 2015; Patel et al. 2014; Shalek et al. 2014; Shalek et al. 2013; Treutlein et al. 2014; Deng et al. 2014; Mahata et al. 2014; Brennecke et al. 2013; Sasagawa et al. 2013). This is in part due to the fact that the technologies and tools needed for conducting scRNA-seq studies have become more readily available and accessible.

In addition, scRNA-seq has been applied in a wide variety of domains, including as cancer research, immunology, neuroscience, and developmental biology, amongst others. scRNA-seq is paving the way for the creation of more targeted and individualized therapy methods for a wide variety of diseases by offering a more in-depth understanding of cellular variability.

While single-cell RNA sequencing (scRNA-seq) has become more accessible to newcomers in the field, thanks to commercial reagents and platforms, the bioinformatic and computational requirements of scRNA-seq still pose significant challenges. As of now, there is a scarcity of user-friendly software packages that allow researchers to efficiently conduct quality control, analysis, and interpretation of scRNA-seq data without a steep learning curve.

However, companies that supply laboratory equipment and reagents for scRNA-seq are gradually introducing free software tools. Examples of such software include Loupe by 10x Genomics and Singular by Fluidigm. These software options are designed to be user-friendly and cater to

researchers who may not have extensive experience in bioinformatics. One limitation of these tools is that they often operate as a “black box,” meaning that users may not have clear insights into the specific algorithms and parameters that are being used in the background. This lack of transparency can be an obstacle for researchers who need a deeper understanding of the computational processes behind the data analysis.

It’s important to recognize that the field of scRNA-seq data analysis is rapidly evolving. Recent trends suggest that web-browser-based interfaces, which are typically more user-friendly, are likely to be more widely available in the near future (Zhu et al. 2017). This development could significantly lower the barriers to entry for researchers and clinicians who want to employ scRNA-seq in their work.

Nonetheless, identifying the precise functionalities that should be integrated into these platforms continues to be an area of active research and development. The requirements for analyzing scRNA-seq data are complex and varied, and an optimal platform would need to be both powerful and flexible.

In summary, a solid grasp of the bioinformatic and computational aspects of scRNA-seq studies is essential. As the field continues to evolve, it is likely that more accessible tools will be developed. However, in the meantime, it is highly beneficial for biomedical researchers and clinicians to seek collaboration with bioinformaticians who have expertise in handling scRNA-seq datasets. Such collaborations can help to ensure that the data is

analyzed accurately and that the results are interpreted with the depth and nuance required for scientific rigor.

In conclusion, single-cell RNA sequencing has developed into an essential tool for investigating cellular heterogeneity, the dynamics of gene expression, and the complex interactions that exist between cells within complex tissues. Our understanding of the molecular basis of biological processes and diseases will continue to expand as a direct result of continuous developments in scRNA-seq technologies and computational techniques. This, in turn, will ultimately lead to the development of new therapeutic options and personalized medicine.

The field of single-cell transcriptomics has undergone significant growth and maturation throughout the current decade. A burgeoning selection of commercial solutions has been made available, many of which are complemented by specialized bioinformatics tools designed to analyze the data. In addition, cutting-edge innovations in microfluidics and cellular barcoding have substantially bolstered the throughput of single-cell RNA sequencing (scRNA-seq) experiments, meaning that a larger number of cells can be studied in parallel.

Concurrently, new protocols that are compatible with cell fixation and freezing have started to surface. This is a critical advancement as it enhances the practicality of scRNA-seq for applications in biomedical research and clinical settings. For instance, the capacity to examine thousands of cells in a single experiment is particularly advantageous for forward-

looking studies that analyze clinical samples, which are often highly diverse in nature. It is anticipated that these developments will have a far-reaching impact, not only in applications that bridge research and clinical practice but also in enriching our comprehension of fundamental tissue structure and physiology.

As the scope for single-cell transcriptome analysis broadens, there has been a notable proliferation of experimental protocols. Each protocol presents its own unique set of advantages and limitations. Consequently, researchers are tasked with making key decisions, such as prioritizing the number of cells analyzed versus the depth of sequencing, determining if information on the full-length of transcripts is necessary, and deciding whether measurements at the protein level or epigenomic data are needed from the same set of cells.

In the face of these choices, it is essential for researchers to have a well-defined set of biological objectives and a thoughtfully constructed experimental design. This foundation is often crucial in guiding informed decisions regarding the selection of the most appropriate approach for a given study. Moreover, as the field continues to evolve at a rapid pace, staying abreast of the latest technological advancements and analytical tools is essential for researchers to harness the full potential of single-cell transcriptomics in answering complex biological questions and addressing clinical challenges.

1.5. Epithelial cell response to SARS-CoV-2 virus

COVID-19 presents itself as a highly intricate disease characterized by an extensive range of clinical responses. These responses can be as mild as showing no symptoms or as severe as experiencing overwhelming respiratory failure that can be fatal. Interestingly, the efficacy of certain treatments varies at different stages of the disease. Some treatments prove to be beneficial when administered early on in the disease, but lose their effectiveness when the patient is critically ill. Conversely, there are treatments that are not recommended at the onset of the disease but have been found to be highly effective in the later stages.

To truly comprehend the diverse nature of COVID-19 and how it responds to different therapies, it is essential to consider the progression of the infection in terms of its location and the types of cells it infects over the course of the disease.

Initially, the infection takes hold in the sinonasal airway epithelium, which is an interwoven mesh of luminal cells responsible for moving mucus (ciliated cells) and mucus-secreting cells that work in tandem for mucociliary clearance, a critical process in maintaining respiratory health. It has been observed that the ciliated cells are particularly susceptible to the SARS-CoV-2 virus and serve as the primary site for initial viral replication and release into the luminal area (Ravindra et al. 2021). During this initial phase, the key objective of treatment should be to thwart the virus from infecting cells and

to curtail the propagation of the virus.

As the disease progresses, the infection begins to migrate down the respiratory tract and eventually reaches the alveoli, tiny air sacs in the lungs crucial for oxygen exchange. At this juncture, the cells targeted by the virus and the corresponding responses undergo a change, necessitating a shift in treatment strategies. The alveolar region is particularly critical because this is where the infection can take a turn for the worse and become life-threatening.

Within the alveoli, the primary cell that becomes infected is the alveolar type II cell. These cells are multifunctional; they produce pulmonary surfactant, which is essential for efficient gas exchange, and also serve as progenitor cells for type II and type I epithelial cells. Notably, type I epithelial cells cover a whopping 95% of the alveolar surface.

Given the paramount role of these cells in respiratory function, the treatment for SARS-CoV-2 in the alveoli should encompass multiple objectives. This includes not only blocking the virus from entering and replicating within the cells but also mitigating damage to the alveolar epithelial cells and the endothelial cells lining the blood vessels in the region. Moreover, treatment strategies should aim to restrain the influx of inflammatory cells that can exacerbate the condition and to foster the repair of epithelial cells. In summary, a comprehensive understanding of the progression of COVID-19 infection and its impact on different respiratory cells is pivotal in devising effective, stage-specific treatment strategies (V'Kovski et al. 2021; Wiersinga et al. 2020).

When viral particles are inhaled, they predominantly settle on the nasal mucosa, which is the mucus-producing tissue lining the inside of the nose (Gengler et al. 2020). It is here that the virus invades, multiplies within, and is subsequently released from specific target cells within the epithelium. Such infections might not show any symptoms or, in some cases, may give rise to localized symptoms.

To gain initial insights into which cells might be the likely targets of the SARS-CoV-2 virus within any tissue, researchers looked at the presence of the transcripts for angiotensin-converting enzyme 2 (ACE2) and transmembrane serine protease 2 (TMPRSS2) in data sets obtained from single-cell RNA sequencing. ACE2 functions as the cell surface receptor that is identified by the spike (S) protein of the SARS-CoV-2 virus. In contrast, TMPRSS2 is an enzyme belonging to the serine protease family, which plays a role in enabling the virus to enter cells by cleaving the S protein (Muus et al. 2021; Sungnak et al. 2020).

However, it's important to acknowledge that the expression of ACE2 is not uniform among cells and is generally present in low levels. This means that not all cells express ACE2 to the same degree, and in many cases, the expression might be relatively scant. This variability and overall low level of expression add a layer of complexity in understanding the cells that are most vulnerable to infection by the SARS-CoV-2 virus within a particular tissue.

As we consider the role of ACE2 and TMPRSS2 in the context of SARS-CoV-2 infection, it becomes apparent that these molecules are

instrumental in the virus's ability to invade cells. Understanding the expression patterns and roles of these molecules can provide valuable insights into the initial stages of infection and how the virus establishes itself within the host. This knowledge is vital for the development of therapeutic strategies aimed at curtailing the virus's ability to enter and replicate within cells.

In addition to patient samples, primary airway epithelial cell cultures obtained from nasal scrapings have been instrumental in enriching our understanding of the mechanisms of SARS-CoV-2 infection and the cellular responses that follow. These primary cultured epithelial cells are susceptible to SARS-CoV-2 infection and, once infected, release a considerable volume of the virus from the surface of the apical cells (Hou et al. 2020; Gamage et al. 2020).

Within the nasal epithelium, both the ciliated cells, which have hair-like structures that move mucus and particles out of the respiratory system, and the goblet cells, which are responsible for secreting mucus, have been found to express ACE2 and TMPRSS2 (Ziegler et al. 2021). Interestingly, the ACE2 protein has only been observed on the cilia and the apical surface of ciliated cells. This suggests that the virus's tropism, meaning the cells it prefers to infect, and its entry into cells could also be influenced by additional molecules and cellular structures.

In this context, it has been suggested that other enzymes, such as cathepsin B, cathepsin L, and furin, might also play a role in the infection process. Moreover, there are indications that SARS-CoV-2 attaches itself to

the microvilli, which are tiny, finger-like projections on the surface of ciliated cells and are abundant in actin. Contrary to what one might expect, the virus does not attach to the cilia themselves. This warrants a more detailed examination of ACE2 localization and a more thorough understanding of the roles of various structures within ciliated cells.

Another crucial aspect to consider is that the airway epithelium has an altered cellular composition in individuals suffering from chronic inflammatory airway diseases, such as asthma, cystic fibrosis, chronic rhinosinusitis, and chronic obstructive pulmonary disease (Vladar et al. 2016; Thomas et al. 2010; Leung et al. 2019). This altered state includes ciliated cells that are structurally and functionally abnormal. Adding to the complexity, it has been observed that the expression of ACE2 is influenced by different factors, being increased by IL-1β and IFN-β, and decreased by IL-13 in bronchial epithelial cells (Hou et al. 2020).

Understanding how these factors influence SARS-CoV-2 infection is critical. Particularly, it's essential to delve into whether these altered states of the airway epithelium and the modulation of ACE2 expression have any bearing on the susceptibility to infection and the progression of COVID-19 in patients with chronic airway diseases. The interplay between these factors could provide valuable insights into designing targeted therapeutic strategies and managing patient care more effectively.

The innate immune response to SARS-CoV-2 within the nasal passages is intricate and multifaceted (Gallo et al. 2021). There is

accumulating evidence that suggests the interferon response, which is a crucial aspect of the body's natural defense mechanisms against viral infections, is subdued in the case of SARS-CoV-2. Interestingly, this dampening of the interferon response might have connections to the severity of the disease.

In one study, single-cell RNA sequencing (scRNAseq) of cells gathered through nasal scrapings from patients infected with COVID-19 revealed some intriguing aspects. It was observed that there was no detection of type I interferons (such as interferon-alpha or interferon-beta) or type III interferons (interferon lambda) in the transcripts. However, cells did exhibit the expression of genes responsive to interferons, such as MX1 and IFITM3, which have antiviral properties. What’s noteworthy is that the expression of interferon genes was considerably lower in patients suffering from severe COVID-19 infection compared to those with mild disease (Fiege et al. 2021).

Another study that employed scRNAseq, but on cultures of human tracheal cells, not only corroborated the primary infection in ciliated cells and a subdued interferon response but also brought to light an intriguing finding (Ouchetto, Moulin, and Roche 2021). It was observed that when the cultures were treated with interferon-beta, there was a significant reduction in the rate of infection.

Furthermore, an additional study investigated both nasal and bronchial primary airway epithelial cultures. It was found that in the first 24 hours following SARS-CoV-2 infection, there was minimal production of

interferon and expression of genes regulated by interferon. However, as time progressed, this expression started to increase.

Another study focused on the array of cytokines, which are signaling proteins in the immune system, secreted after SARS-CoV-2 infection of nasal epithelial cells. With the exception of CXCL10, this study also observed a subdued cytokine secretion. Moreover, when comparing the innate immune response to SARS-CoV-2 with that of influenza A infection, it was observed that the innate immune response to SARS-CoV-2 was significantly reduced.

These insights underline the complexity of the immune response to SARS-CoV-2, especially regarding interferons. It hints at possible avenues for therapeutic intervention by modulating the interferon response, but also demands further investigation into the intricacies of how the immune system interacts with the virus. Understanding these interactions in-depth can be fundamental in designing targeted interventions and treatments.

The treatment strategies for COVID-19 are highly contingent on the progression stage of the disease and the specific cell types that have been infected. It is imperative to recognize that the extent of damage inflicted upon the alveolar epithelium is a critical determinant in the mortality and long-term consequences ensuing from a SARS-CoV-2 infection.

A distinctive feature observed in the long-term ramifications of SARS-CoV-2 infections is the damage sustained by the progenitor cells, which are essential for the regeneration of the epithelium within the conducting airways and the alveoli. Within the conducting airways, basal

stem cells appear to escape significant damage. However, within the alveoli, a contrasting picture emerges where type II cells are subjected to considerable damage.

Several pressing questions emerge from these observations that require urgent attention. Firstly, it is crucial to decipher the underlying mechanism that leads to the observed dampened immune response in the nasal and bronchial regions, especially in contrast to what is witnessed in the lower respiratory tract. Understanding this could unveil new avenues for therapeutic interventions.

Another key question pertains to the timing and modulation of the innate immune system during the disease's progression. Specifically, at what junctures in the disease's progression would bolstering or suppressing the innate immune response prove to be advantageous? This is critical for the development of targeted and effective treatment strategies.

Furthermore, another area that necessitates focused inquiry is identifying potential treatments that can not only protect the type II cells in the lungs but also bolster their resilience and functionality. This is of paramount importance, as type II cells are instrumental in restoring the alveolar epithelium after it has sustained damage. Ensuring their survival and functionality could be central to improving patient outcomes.

Additionally, a deeper understanding of the cellular and molecular underpinnings behind the observed impaired response to the virus among the elderly population is critical. As this demographic is at a higher risk for severe

outcomes, elucidating these mechanisms can have profound implications for the development of targeted therapeutics.

In conclusion, the pathogenesis of COVID-19 harbors many uncertainties and unknowns. Rigorous research endeavors are underway to unravel these mysteries, and it is anticipated that significant insights will emerge in the near future. These findings have the potential to considerably enhance our understanding and management of this complex disease.

1.6. Purpose of this study

The primary goal of this study is to advance our understanding of human lung biology and disease modeling by generating biobanks of 83 human primary airway organoid lines and establishing a comprehensive single-cell atlas of these lines, derived from a diverse subset of 27 individuals. To achieve this, we will generate biobanks of 83 human primary airway organoid lines. This work represents a significant expansion of previously published lung cell atlases and provides a more in-depth investigation of the cellular makeup and molecular properties of healthy and diseased human lung tissues.

To achieve this, I integrated public datasets from the Human Lung Cell Atlas and directly compared the transcriptomes of 2D and 3D human airway in vitro models with those of in vivo human lung tissues to determine any correlation. This comparative approach allows us to evaluate the fidelity

and robustness of in vitro models in recreating the complex cellular and molecular landscape of human lungs in vivo.

Through thorough single-cell profiling, I can trace the differentiation trajectories of rare airway cell types originating from basal cells over pseudotime. This provides invaluable insights into lineage linkages and developmental processes driving lung cell differentiation. Additionally, I discover unique immune-primed subsets in in vitro models taken from the lungs of individuals with chronic respiratory disorders, demonstrating organoid models' potential for studying clinical conditions' influence on lung cell states.

I illustrate the practical applications of these in vitro models and our organoid single-cell atlas in investigating cellular states and immune responses at the single-cell level following infection with various viruses, in addition to characterizing cellular and molecular features of human primary airway organoids. This approach has significant implications for studying viral respiratory disease pathogenesis, including the ongoing COVID-19 pandemic, and informing the development of novel therapeutic strategies.

In conclusion, I offer our organoid single-cell atlas to the scientific community as an explorable and usable interactive online resource (OSCA: http://osca.snu.ac.kr) that can be incorporated into their research endeavors. I hope that by providing this valuable resource, I can encourage collaboration and accelerate breakthroughs in lung biology and disease modeling, ultimately leading to an enhanced ability to understand, diagnose, and treat

lung disorders.

Chapter 2. Materials and Methods

2.1. Human tissue samples

To establish human primary distal airway organoid cultures, distal lung tissues were obtained from lung cancer patients undergoing lung resection surgery at Seoul National University Bundang Hospital (SNUH). Informed consent was received and the study was approved by the ethical committee (IRB No. 2008-065-1148). Organoids were derived from lung tissues that were either farthest from a tumor lesion (normal lung tissues) or had a history of chronic respiratory diseases (diseased lung tissues). Fresh lung tissues were stored in media (Advanced DMEM/F12 (Gibco), 1% HEPES (Gibco), 1% Glutamax (Gibco), and 0.125 ug/ml Amphotericin B (Gibco)) at 4°C for less than 12 hours before processing.

2.2. Human primary distal airway organoids

Lung tissues were cut into <5 mm pieces and washed in ice-cold PBS to remove residue mucus and blood cells. Before snap freezing for DNA extraction, 2-3 pieces of lung tissues were kept in RNAlater solution (Invitrogen) overnight at 4°C. The remaining tissue pieces were dissociated into single cells using a gentleMACS Octo Dissociator with Heater (Miltenyi Biotec) following the manufacturer's protocol. Tissue suspensions were

filtered through a 70 µm cell strainer (Miltenyi Biotec) and washed with 5 ml of Advanced DMEM/F12. Cells were centrifuged at 400g for 10 min at 4°C, and the supernatant was removed. Cell pellets were resuspended in 1 ml of red blood cell lysis buffer (Miltenyi Biotec) for 10 min at room temperature (RT). The reaction was stopped by adding 9 ml of Advanced DMEM/F12 with 1% HEPES, 1% Glutamax, and 1% Penincillin-Streptomycin (Gibco) (hereafter ADF+++), followed by centrifugation at 400 g for 10 min at 4°C. Cell pellets were then resuspended in growth factor reduced Matrigel (GFR-Matrigel; Corning) and plated as 40 µl droplets in a pre-warmed 24-well tissue culture plate. Plates were incubated at 37°C for 15 mins, followed by submersion in 500 µl of pre-warmed airway organoid medium (AO medium) with 10 uM Y-27632 (Tocris) for the first 3 days (Sachs et al. 2019). The medium was changed every 3 days. Established three-dimensional (3D) airway organoids (3D-AOs) were passaged at 1:4 ratios every 1-2 weeks. For passaging, TrypLE express (Gibco) was added to each well, and GFR-Matrigel was mechanically disrupted. Organoids were incubated at 37°C for 5 min. The reaction was quenched by adding ice-cold ADF+++, and cells were centrifuged at 400 g for 5 min at 4°C. Cells were resuspended in GFR-Matrigel and plated as described above. For cytokine treatment, 1 ng/ml IL-13 (Biolegend), 10 ng/ml IL-1β (Biolegend), or 10 ng/ml IL-6 (R&D systems) were added to 3D-AOs for 3 weeks.

2.3. Air-liquid interface (ALI) culture of organoids

Established 3D-AOs were dissociated into single cells as previously described. Transparent 24-well transwell inserts (Corning) were precoated with 1% v/v GFR-Matrigel for a minimum of 1 hour at 37°C. Subsequently, 200k dissociated airway cells were resuspended in 200 μL of AO medium or PneumaCultTM-Ex Plus medium (STEMCELL Technologies) and seeded in each transwell insert for 2D ALI cultures (ALI-Om or ALI-Pm, respectively). Then, 500 μL of medium was added to the lower compartment of transwell plates. After 5-8 days, when monolayers reached 100% confluency, growth medium was removed from the inserts to induce differentiation for another 21 days. For ALI-Pm, PneumaCultTM-ALI Medium (STEMCELL Technologies) was used for inducing differentiation.

2.4. scRNA-seq of in vitro human primary airway cultures

For sequencing, established 3D-AOs between passages 3 and 10 from individual donors were used. 3D-AOs were dissociated into single cells and plated for establishing ALI-Om and ALI-Pm, which were used for sequencing after 3 weeks of differentiation. Dissociated cells were resuspended in 1 ml ADF+++ and filtered through a 30 μm SmartStrainer (Miltenyi Biotec). Each sample was then analyzed using ReadyCount Green/Red Viability Stain (Invitrogen) for counting the number of cells and measuring cell viability with an automated cell counter, Countess 3 FL (Thermo Fisher). Cell suspensions from in vitro cultures established from different donors were pooled to obtain

a cell suspension mixture with an equal proportion of each sample. The pooled cell suspension was centrifuged at 400 g for 5 min at 4°C, and then resuspended in a 0.04% bovine serum albumin (BSA) solution at an appropriate volume for microfluidic chip loading. Typically, 6 to 8 samples were loaded per lane of a 10x microfluidic chip device, and demultiplexed based on single-nucleotide polymorphisms. For 10x multiome analysis, one sample was loaded on a Chromium Next GEM Chip J. All libraries were sequenced in a NovaSeq 6000 system (Illumina) in paired-end mode.

2.5. Immunofluorescent staining of 2D ALI and 3D cultures

For immunofluorescent (IF) staining of 3D-AOs, the Matrigel was dissolved using ice-cold Cell Recovery Solution (Corning) at 60 rpm for 30 to 60 minutes at 4°C on a horizontal shaker (CRYSTE). The recovered organoids were fixed in 4% paraformaldehyde for 30 minutes at 4°C. For IF staining of ALI-Oms, the insert was removed from the plate and placed upside down on a clean petri dish. Using a surgical blade, the membrane with cells was detached from the insert and fixed in 4% paraformaldehyde for 30 minutes at 4°C. The samples were then washed with 0.1% TritonX-100/PBS (PBS-T). Next, they were permeabilized in 0.2% Triton X-100/PBS for 5 minutes and incubated in 5% donkey serum (Jackson ImmunoResearch)/PBS-T blocking solution for 2 hours at room temperature (RT). Primary antibodies were incubated overnight at 4°C at the indicated

dilutions: Chicken-anti-CK5 (1:500, Biolegend, 905904), Mouse-anti-TP63 (1:100, Thermo Fisher, MA1-21871), Rabbit-anti-MUC5AC (1:200, Cell Signaling Technology, 61193), Rat-anti-SCGB1A1 (1:200, R&D Systems, MAB4218), Rabbit-anti-Acetyl-α-Tubulin (1:200, Cell Signaling Technology, 5335), Mouse-anti-SARS-CoV NP (1:150, Sino Biological, 40143-MM05), Mouse-anti-Influenza A NP (1:100, Meridian, C87050M), and Rabbit-anti-SARS-CoV NP (1:1000, Sino biological, 40143-T62). Alexa Fluor-coupled secondary antibodies (1:500, Invitrogen) were incubated overnight at 4°C. After antibody staining, nuclei were stained with DAPI (1:1000, Sigma) and sections were embedded in Vectashield Plus antifade mounting medium (H-1900, Vector laboratories). Fluorescence images were acquired using either Leica SP8 or LSM700 confocal microscopes. LAS X (Leica) or ZEN software (Zeiss) was used for processing fluorescent images.

2.6. Quantitative PCR

Total RNA was extracted from 3D-AO and ALI-Om using the RNeasy mini kit (Qiagen). Complementary DNA (cDNA) was synthesized from the extracted RNA using the 1st strand cDNA synthesis kit (Takara) according to the manufacturer's protocols. Quantitative PCR (q-PCR) was performed with the Viia7 Real-time PCR system (Thermofisher) using TB green premix ex taq II (Takara) according to the manufacturer's instructions. Relative expression (2-ddCt) was normalized with the Ct of GAPDH and the

dCt of 3D-AOs.

2.7. Virus preparation

SARS-CoV-2 alpha and delta strain was obtained from the Korean National Culture Collection for Pathogen (NCCP no.43326 and 43390, respectively). Viruses included in the beta and omicron lineage were isolated from nasopharyngeal swabs taken from COVID-19 patients (GenBank accession no. OP349649.1 and 349650.1, respectively). Wild type MERS-CoV (GenBank accession no. KT029139.1) was provided by the Korean Centers for Disease Control (KCDC). MERS-CoV and SARS-CoV-2 were propagated on VeroE6 cells (CRL-1586, ATCC) in DMEM (Welgene) supplemented with 2% fetal bovine serum (FBS, Gibco) and 100 IU/ml penicillin-streptomycin at 37°C in a humidified CO2 incubator for 3 days and titrated under overlay medium containing 0.8% Methylcellulose (Sigma) and 2% FBS in DMEM. Influenza A (VR-95™, ATCC) was propagated in embryonated chicken eggs and titrated in MDCK cells (CCL-34™, ATCC) by plaque assay. The culture supernatant was cleared by centrifugation and stored in aliquots at −80°C until use.

2.8. Viral infection procedure

All experiments involving infectious viruses were conducted in a Class II Biosafety Cabinet under BSL-3 (MERS-CoV, SARS-CoV-2) or BSL-

2 (Influenza A) conditions at Seoul National University. For the viral infection of 2D differentiated cultures (SARS-CoV-2, MERS-CoV, and Influenza A), samples were pre-washed and infected from the apical side of the ALI culture. Cell count in each culture was determined, and cultures were incubated before washing and culturing under normal conditions for 3 days prior to scRNA-seq library generation, IF staining, or TEM.

2.9. Transmission electron microscopy (TEM)

To visualize SARS-CoV-2 virus particles using TEM, infected ALI-Om samples were fixed and processed through a series of steps before ultrathin sections were prepared and stained. TEM imaging and analysis were conducted using a transmission electron microscope (JEM-1400) at 80 kV in the Department of Research & Experiment at Seoul National University Hospital.

2.10. K-chip genotyping array

DNA was extracted from frozen tissue or organoids and genotyped using the Korea Biobank Array V1.1 (ThermoFisher). Data preprocessing followed the Korea Biobank Array Project analytic protocol, and specific criteria were used to select variants. Genotype principal component analysis was performed using Plink v1.0.9, and imputation was conducted with the East Asian-specific WGS imputation reference panel (Northeast Asian

Reference Database V2, unpublished). Phasing and imputation were executed using Beagle5 and Minimac4, respectively.

2.11. scRNA-seq data analysis

We used Cell Ranger (v6.0.1) to align sequenced reads to the GRCh38 human reference genome. Demultiplexing of individuals was done using souporcell (v2.0) with donors' genotype array data from K-chip. Samples underwent ambient RNA removal using cellbender (v0.2.1) for background noise expression and doublet removal. Only cells meeting specific criteria were used for further analysis. Normalization was done using SCTransformation in the Seurat package (v4.0.3). We carried out canonical correlation analysis (CCA) to correct batch effects between sequencing libraries. UMAP plots and cell clustering were visualized using Seurat package functions.

2.12. Cell cycle scoring

To compare cell cycling phase proportions between culture methods, we employed CellCycleScoring in the Seurat package. Specific marker genes were used for S phase and G2M phase. TUBB4B gene was removed from default Seurat G2M marker genes due to its role as a ciliated cell marker. The ratio of cycling cells for each culture method of a cell type was calculated based on average ratios of S or G2M phase cells for individuals.

2.13. Integration of OSCA and tissue scRNA-seq data

To compare in vitro human lung cell transcriptomes with in vivo human lung tissues, we integrated our in vitro single cell datasets with in vivo tissue datasets from the European Genome-phenome Archive (EGA) under accession EGAS00001004344 (Travaglini et al. 2020). We used the FindTransferAnchors function for transferring anchors between in vitro and tissue data. After performing MapQuery, we regenerated UMAP and Louvain clusters. Cell types were identified using scHCL (v0.1.1), Azimuth, and manual curation with known marker gene expressions.

2.14. Cell type annotation prediction using CellTypist

We employed the Python package CellTypist (v.1.2.0) for cell type annotation prediction using logistic regression models. All models provided by CellTypist, trained with lung single cell data, were used on our data for annotation prediction and validation. Additionally, we used lung single cell data and annotated metadata from P.K.L. Murphy et al. to create a CellTypist model. Default parameters were used for model building, and all predictions were made with majority voting.

2.15. Rare cell score

For cell type identification, scHCL was employed to match each cell

to annotated scRNA-seq data. Within the 23 clusters divided by the Louvain clustering algorithm, cluster 22 exclusively expressed canonical marker genes of rare airway epithelial cell populations, such as tuft cells, ionocytes, and pulmonary neuroendocrine cells (PNECs). This cluster was primarily labeled as 'Basal.cell.Airway.Epithelium_Plasschaert' or 'Ionocyte.Airway.Epithelium_Plasschaert'. We defined the scHCL_score of these labels as the 'rare cell score'.

2.16. Rare cell population lineage trajectory

Single cell RNA-seq data from different sample origins were reintegrated using Seurat's CCA algorithm. Monocle (v3.1.0.0) was used to determine the developmental paths of cells along the UMAP via the learnGraph function. Cells were ordered along the identified pseudotime trajectory using the orderCell function, with the cluster predominantly composed of basal cells serving as the initial root group. The prepPseudotimePlotDatasets function, adopted from Goldfarbmuren et al., was used to view key marker gene expression along pseudotime on a shared scale, generating smoothed expression curves normalized between zero and one (Goldfarbmuren et al. 2020).

2.17. Single cell multiome data analysis

Single cell multiome data were aligned to the GRCh38 reference

genome using cellranger-arc (v2.0.0). Gene expression data from each single cell followed the same pipeline as 3' scRNA-seq data with Seurat (v4.0.3). The Signac (v1.6.0) package was employed for filtering, clustering, and visualizing Assay for Transposase-Accessible Chromatin (ATAC) data. Cells with specific criteria were used for further analysis. After normalization using RunTFIDF, FindTopFeatures, and RunSVD, we identified TAB cell-specific overrepresented motifs by comparing TAB cells and basal-2 cells using FindMotifs. The most significantly overrepresented motif profiles were represented using the JASPAR database. Motif activity scores were calculated with chromVAR using the bSgenome.Hsapiens.UCSC.hg38 reference.

2.18. Identification of expression profiles

Normalized expression profiles for each culture method or subcluster were compared using the Wilcoxon rank sum test with the FindMarkers function. Genes with an adjusted p-value < 0.05 were considered differentially expressed genes (DEGs). Gene set enrichment analysis (GSEA) was conducted using GSEAPreranked with DEGs ranked by log2FC. Hallmark and C5: Biological process were tested for enrichment. Gene sets with an FDR q-value < 0.01 were deemed significant. Various scRNA-seq data with lung disease patient tissues and organoids with cytokine treatment were reference mapped to OSCA integrated with tissue to predict cell types and subclusters for each cell. Idiopathic pulmonary fibrosis (IPF), chronic

obstructive pulmonary disease (COPD), cystic fibrosis (CF), healthy control, and fetal lung tissue data from GEO using GSE135893, GSE150674, and EMBL-EBI ArrayExpress database (E-MTAB-8221) were utilized (Miller et al. 2020; Habermann et al. 2020; Carraro et al. 2021). Prior to reference mapping, each query data was processed with SCTransform, RunPCA, and RunUMAP in the same manner as reference OSCA data. After transferring anchors, the ratio of predicted subclusters by MapQuery was calculated for each sample.

2.19. scRNA-seq data analysis with virus-infected samples

We used a chimeric reference genome comprising GRCh38, NC_045512.2, NC_019843.3, and NC_002016.1~NC_002023.1 for SARS-CoV-2, MERS-CoV, and Influenza A, respectively (Wu et al. 2020; Zaki et al. 2012; Winter and Fields 1980). All processes, including alignment, normalization, and filtering, were identical to the pipeline used to construct OSCA. Infected cells were defined as cells expressing viral genes. Interferon (IFN) and Entry scores were measured using AddModuleScore in the Seurat package with genes related to interferon signaling and viral entry factors.

Chapter 3. Body

3.1. Archiving human airway organoids

I created human primary airway organoid biobanks, which resulted in the generation of a complete collection that included 83 cryopreserved organoid lines, matched snap-frozen tissue samples, and genotyping information. The human distal lung tissues used in this investigation were first removed surgically, then carefully processed, and then fragmented enzymatically into single cells so that they could be used later on in the production of 3D airway organoids. A success rate of 70% (83/118) was achieved when attempting to keep established organoids in culture by passaging them up to three times. After they had been successfully cultured, these organoids were carefully stored in biobanks so that they could be used as useful resources in subsequent research.

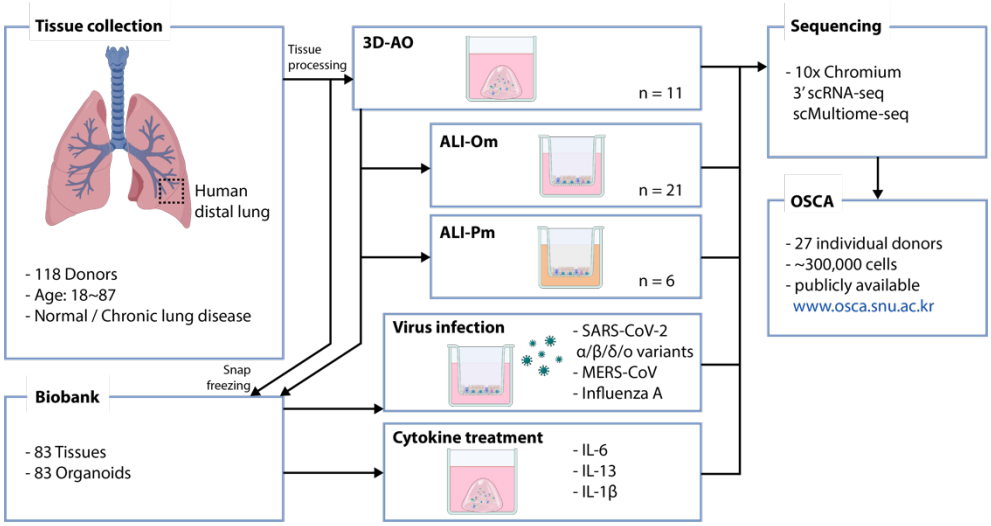

Figure 4. Schematics of data generation.
Schematic of how to make organoid biobanks, analyze the transcripts of a single cell, and use organoids to treat respiratory viruses and cytokines. 3D-AO (3D airway organoid), ALI-Om (Air-liquid interface culture with organoid media), ALI-Pm (Air-liquid interface culture with Pneumacult media), and OSCA (Organoid single cell atlas). n is the number of samples.

I directly examined the two most common types of in vitro human airway models, which are known as the 2D air-liquid interface (ALI) and the 3D organoid cultures, in order to acquire insights into the influence that culture conditions have on the behaviors of cells and the gene expression programs that are active in vitro. These models include the two-dimensional ALI and the three-dimensional organoid culture. Each of these models simulates human airway physiology in their own unique way and has both advantages and downsides. Airway organoids were dissociated into single cells and then seeded onto porous membrane inserts in order to establish a submerged 2D ALI culture. Both the 3D airway organoid, which will be

referred to henceforth as the 3D-AO, and the 2D ALI, which will be referred to henceforth as the ALI-Om, were cultivated using the same medium, as detailed in the Methods section, in order to minimize any potential bias that could be introduced by using separate culture media.

In addition, I grew the cells in a typical ALI culture system, which encourages mucociliary differentiation, by making use of the commercially available PneumaCult medium (referred to as ALI-Pm) (See Methods)(Tata et al. 2013; He et al. 2022). This culture system was performed using the commercially available PneumaCult media. Taking this approach allowed for a comparison with the other culture conditions that was more closely controlled. After that, I labeled airway lineage markers using immunofluorescent staining (IF), which included basal, goblet, club, and ciliated cells. This allowed me to evaluate the spatial distribution of differentiated cells within the pseudostratified epithelium of both the 3D-AO and ALI-Om models.

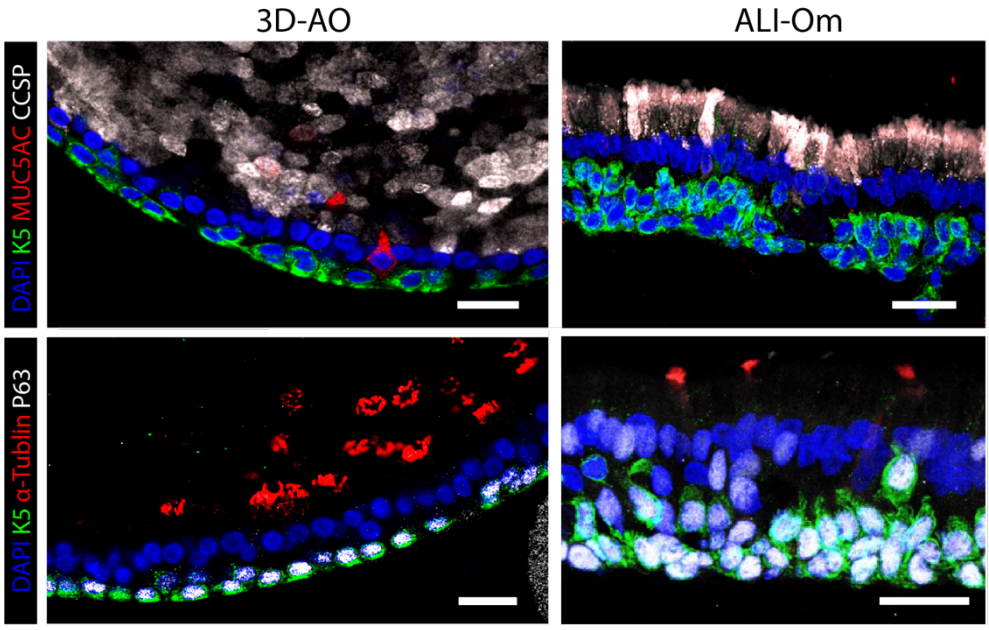

Figure 5. Immunofluorescent staining for marker genes.
Immunofluorescent marking of airway lineage markers in 3D-AO and ALI-Om: nuclei (DAPI), basal cell (KERATIN-5, TP63), goblet cell (MUC5AC), club cell (CCSP), and ciliated cell (ACETYLATED-TUBLIN). Bars of scale are 30 m.

3.2. Single-cell atlas of cellular diversity in airway cells

Following that, my objective was to create a comprehensive single-cell transcriptional atlas of human primary distal airway organoids that had been cultured under the three distinct conditions that were described earlier. I evaluated the single-cell transcriptome of 3D-AO, ALI-Om, and ALI-Pm acquired from 27 individuals by utilizing the 10X genomics platform. Of the 27 individuals, 19 had normal lungs, and 8 had sick lungs. After careful checks to ensure the quality of the raw data, the separate datasets were combined and harmonised with the help of Canonical Correlation Analysis (CCA) for batch correction (see Methods). The overall quality control tests

were successfully completed by 95,170 cells out of the entire population.

Methods	# samples	# cells
3D-AO	11	20,576
ALI-AO	21	58,765
ALI-Pm	6	15,829
	38	95,170

Table 1. scRNA-seq summary
The number of sequenced samples and cells for each culture methods.

Using standard canonical cell type marker genes, I was able to identify and annotate a variety of primary categories of airway epithelial cells as I progressed through the investigation. Interestingly, the findings demonstrated that each of the 27 organoid lines had all of the primary cell types, despite the various growing techniques that were used. Basal cells (TP63, KRT5, KRT15), secretory club cells (SCGB1A1, BPIFB1), goblet cells (MUC1, MUC5B, MSMB), and ciliated cells (FOXJ1, PIFO, TPPP3, SNTN) displayed profiles that were consistent with their in vivo counterparts across the in vitro multi-culture systems.

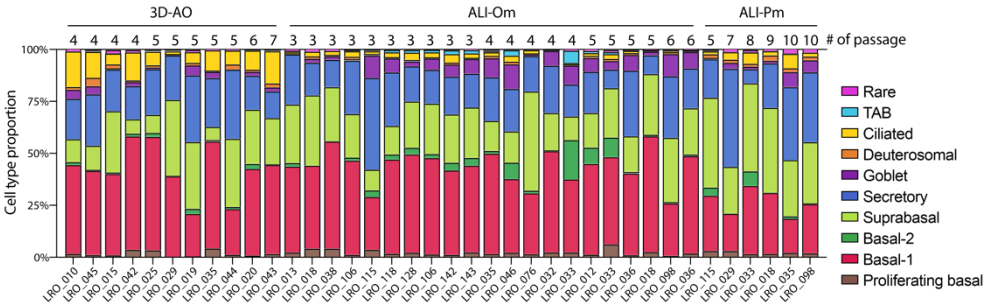

Figure 6. Proportion of various airway cell types.

A bar chart illustrating the percentages of different airway cell types present in each of the 38 samples, arranged according to culture techniques and passage numbers.

In addition, I identified two different subsets of basal cells (basal-1 and basal-2), with basal-2 expressing higher levels of CXCL14, RNF43, LGR6, and LRP4 than basal-1. Particularly noteworthy is the fact that my comprehensive single-cell profiling made it possible to locate SCGB3A2+ cells, which were just recently found in the terminal bronchioles of human lungs.

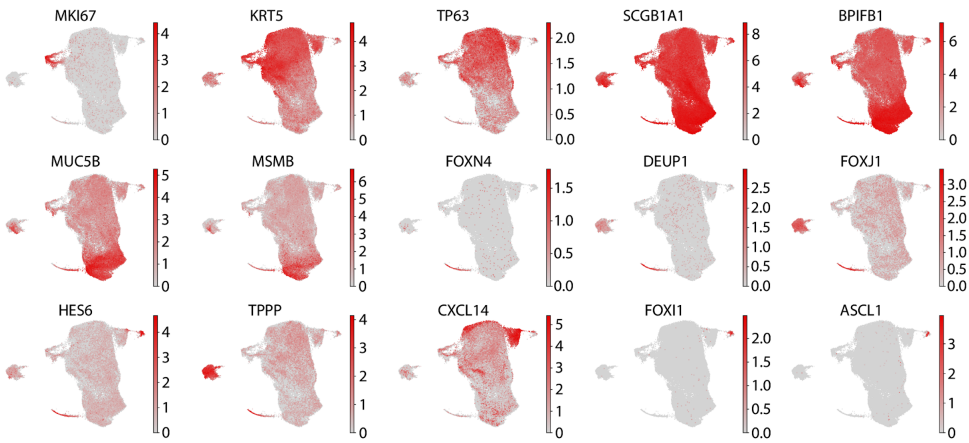

Figure 7. Marker genes of each lung airway cell types

It is important to note that because organoids are actively growing, I was able to capture cellular transitional stages that are not easily observable in the lung that is in its homeostatic condition. For instance, one population of organoid cells co-expresses basal and secretory cell marker genes (KRT5, SCGB1A1), which indicates that suprabasal cells are in the process of

differentiating from the basal secretory cell stage to the luminal secretory cell stage. In addition, I found a distinct cluster of deuterosomal cells that expressed both secretory and ciliated cell marker genes (SCGB1A1, FOXJ1), which is suggestive of a transition from the secretory state to the ciliated state. As was to be anticipated, a subset of basal cells, which are referred to as proliferating basal cells, displayed active proliferation, whereas differentiated cells displayed modest levels of cycle activity. In both organoids and lung tissues, deuterosomal cells, which are the progenitors of ciliated cells, displayed a significant amount of cycling activity.

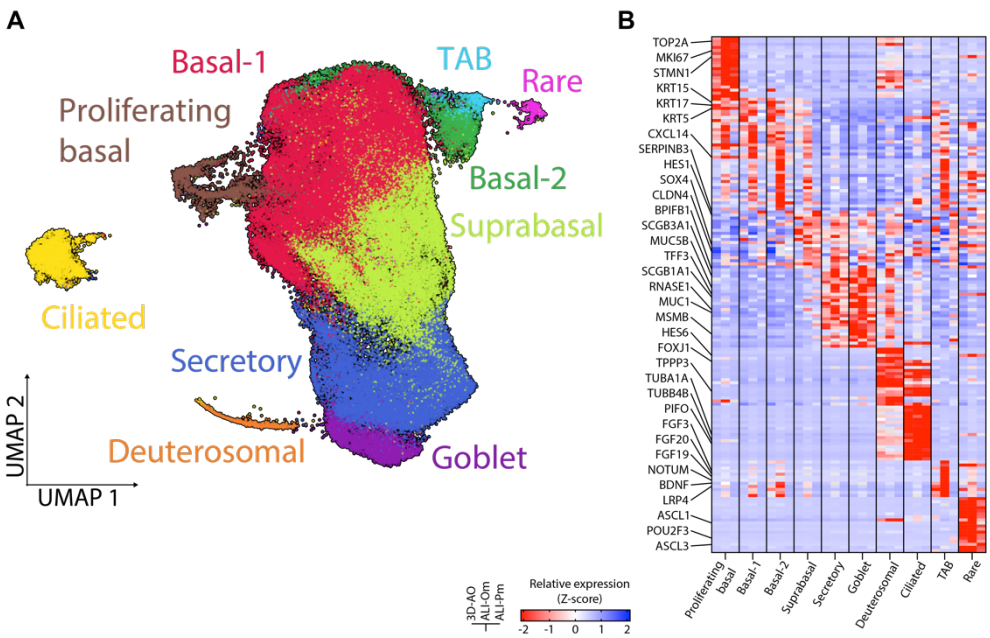

Figure 8. Single cell transcriptomic profiles of organoids.
(A) UMAP displaying the scRNA-seq information of in vitro human primary airway cells. Ten distinct airway epithelial cell types are represented in various colors. TAB refers to transitioning airway basal cells. **(B)** A heatmap that demonstrates the differential expression of marker genes for each epithelial cell type. Each cell type column is ordered as follows: 3D-AO (left), ALI-Om (middle), ALI-Pm (right).

It is interesting to note that regardless of the culture conditions or the number of passages, the cellular diversity and state of all 27 in vitro lines remained practically the same, with only modest alterations in their proportions. In addition, I could not find any distinguishable differences between the groups with regard to age, gender, smoking status, or previous disease history. Using the Organoid Single Cell Atlas (OSCA: http://osca.snu.ac.kr), the datasets can be investigated in an interactive manner as well as analysed.

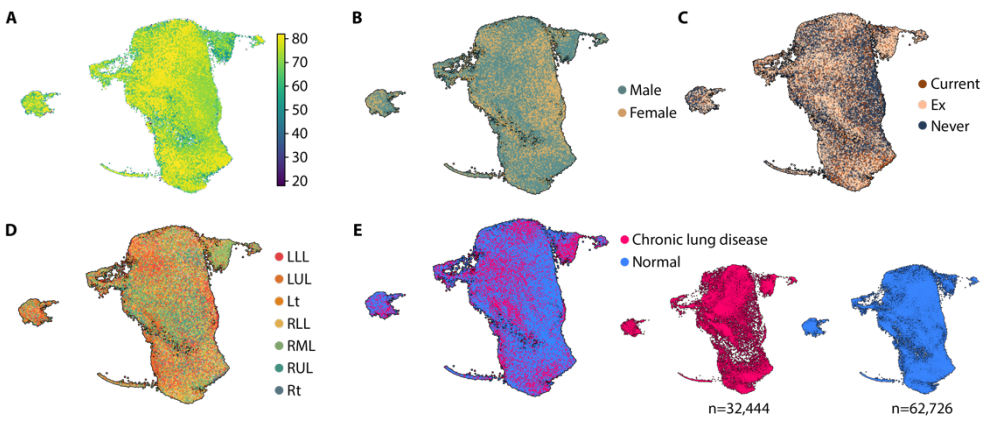

Figure 9. Group comparison
UMAP labeled by age **(A)**, sex **(B)**, smoking status **(C)**, specimen location **(D)** and comorbidity **(E)**.

This single-cell transcriptional atlas is a significant resource for gaining a knowledge of the complexity and dynamics of human primary airway cells in a variety of in vitro multi-culture systems. My findings contribute to a more comprehensive understanding of the formation,

differentiation, and function of airway cells by capturing the cellular heterogeneity that exists within these systems as well as the transitional phases that occur within them.

3.3. Single-cell multiomics of rare airway lineages

Recent developments in single-cell transcriptome analysis of human primary lung tissues have made it possible to map, at the single-cell level, cellular composition as well as the lineage hierarchy (Papalexi and Satija 2018). However, due to the rarity of certain epithelial cell types, such as tuft cells, ionocytes, and pulmonary neuroendocrine cells (PNECs), it is challenging to conduct in-depth research on these cell types. As a result, in vivo lung atlases only provide a limited understanding of the lineage trajectories of these cell types.

It is remarkable that my large-scale single-cell transcriptome analysis was able to identify all of these uncommon airway epithelial cell types in both the 2D ALI and the 3D organoid models. Ionocytes (FOXI1), tuft cells (POU2F3), and PNECs (ASCL1) are some of the cell types that fall within this category. Validating the in vitro lineage link between basal and secretory cells, which corresponds to the well-established in vivo lineage hierarchy from basal to secretory cells in the lung, was the first step in the process of constructing a framework for tracking the lineage trajectories of these rare epithelial cell types. This allowed me to design a framework for tracing the

lineage trajectories of these rare epithelial cell types.

I next developed a system for monitoring the lineage trajectories of these uncommon epithelial cell types, and I reintegrated these cells into the datasets by mining my single-cell datasets for clusters of proliferating basal, basal-1, suprabasal, and secretory cells. Pseudotime trajectory analysis did what was predicted and indicated the differentiation pathway from basal cells to secretory cells. This ensured the fidelity of in vitro differentiation programmes in comparison to in vivo lungs.

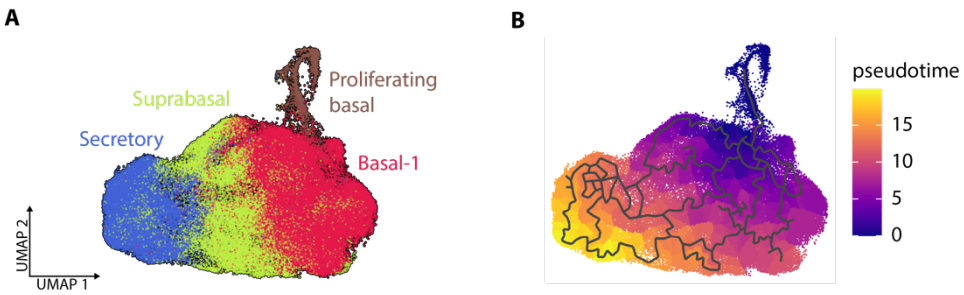

Figure 10. In vitro human airway epithelial cell differentiation
(A) Reorganized UMAP including proliferating basal, basal-1, suprabasal, and secretory cells. **(B)** Pseudotime trajectory assessment conducted using Monocle3.

Following this, I looked into the lineage differentiation of unusual cell types by first generating a rare cell score. This score was based on the degree to which each cell's transcriptome was similar to that of known reference rare cells. It was interesting that I found that basal-2 cells (cluster 13) had the highest rare cell score among non-rare lung epithelial cells; this leads me to hypothesize that basal-2 cells might be the precursor for rare cell

development. Further investigation by my team uncovered transitioning airway basal cells (TAB), which had a reduced expression of basal cell marker genes and were tightly connected with basal-2 and uncommon cell types. The expression of NOTUM, FGF3, and FGF19 was what ultimately provided a definitive understanding of the molecular program that modifies TAB cells during this transition.

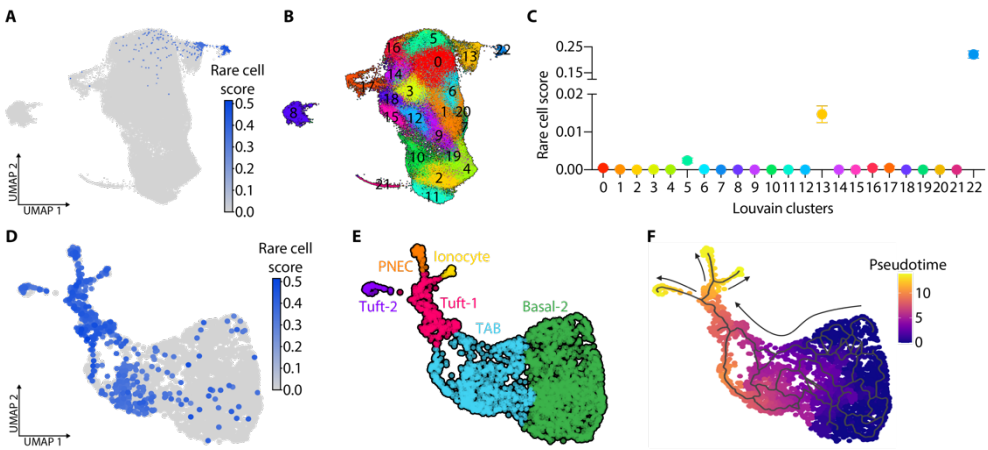

Figure 11. Transitioning airway basal cells.
(A) UMAP displaying rare cell scores based on transcriptomic similarity to rare cells. **(B)** UMAP with Louvain clusters, featuring 23 distinct clusters. **(C)** Mean scatter plot with 95% CI illustrating average rare cell scores for each Louvain cluster. **(D)** Re-clustered UMAP focusing on Louvain clusters 13 and 22, colored by rare cell score (n = 2,962 cells). **(E)** Re-clustered UMAP identifying expected rare epithelial cell types, such as ionocytes, tuft cells, PNECs, and a novel transitioning airway basal cell state (TAB). **(F)** Pseudotime trajectory analysis conducted with Monocle3.

In order to get a more in-depth comprehension of the connection that exists between TAB cells and unusual airway lineages, I carried out the 10x single-cell multiomics approach on 3D-AO. It is important to note that the single-cell Assay for Transposase Accessible Chromatin utilizing sequencing

(scATAC-seq) indicated that TAB regions characterized by gene expression profiles were very closely connected with rare cell populations. In addition, I found transcription factor (TF) binding sites that were opened up in a manner that was exclusive to TAB cells. This led me to hypothesize that TAB cells would reflect primed cell states that lead to the differentiation of basal cells into PNECs. In addition, I separated tuft cells into tuft-1 (ASCL2) and tuft-2 (TRPM5) cells on the basis of the different signatures that each of these cell types possessed.

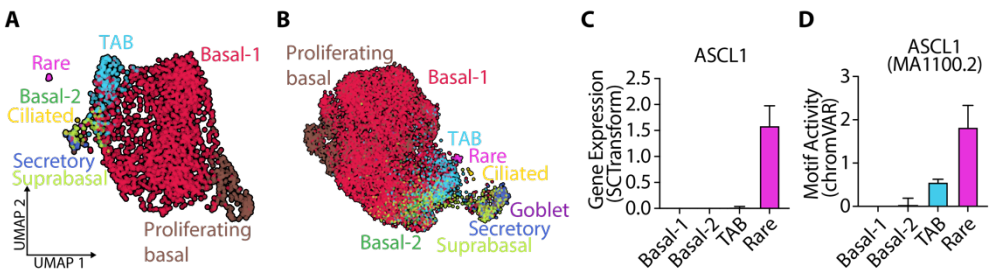

Figure 12. Single cell multiomics on TAB cells.
(A and B) UMAP illustrating gene expression data **(A)** and ATAC peak data **(B)** acquired from 10x multiome sequencing. **(C and D)** Bar plot with 95% CI displays normalized expressions determined by SCTransform **(C)** and motif activity assessed by ChromVar **(D)** for ASCL1.

I also made 3 samples of a lung organoid containing AT2 cells. This organoid is a model that contains both airway cells and alveolar cells. This organoid was developed very recently and may not be fully validated. I was able to get data that included the entire lung epithelial cell, including AT2 and AT1 cells, and also found new cell types. The SCGB3A2+ club cells that differentiate into AT2 cells seem to be slightly clustered to one side, but the difference is

not significant.

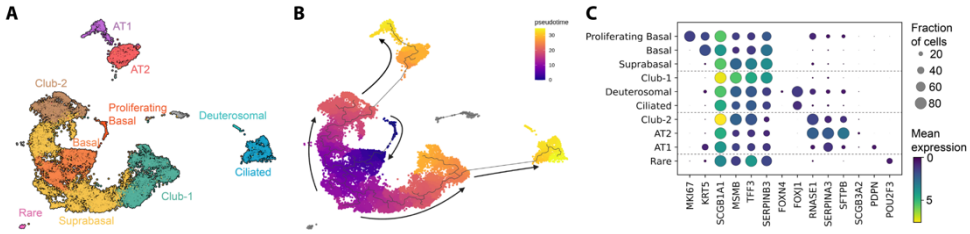

Figure 13. Lung organoids containing alveolar cells
(A) UMAP plot of lung organoid with three samples. **(B)** Epithelial cell differentiation trajectory including two pathways. **(C)** Dot plot representing marker genes of each cell type.

3.4. Immune response signatures in lung tissues

I merged my dataset with the well-annotated data of epithelial cell clusters from the Human Lung Cell Atlas in order to determine how closely the cellular and molecular features of in vitro human airway cultures resemble those of in vivo human lung tissue (Travaglini et al. 2020). This allowed me to determine the degree to which in vitro human airway cultures mimic in vivo human lung tissue. Although the cellular composition of in vitro models is comparable to that of lung tissues, the proportions of each cell type are different. For example, the number of basal and intermediate cells present in in vitro airway cells is significantly higher.

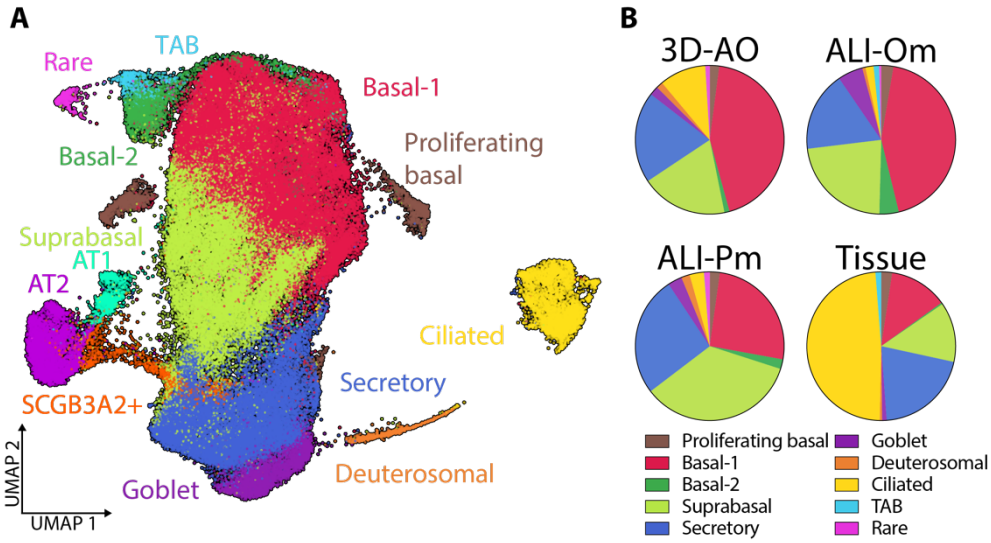

Figure 14. Different cell type compositions of organoids and tissues. **(A)** Combined UMAP representation of cells from in vitro airway models and in vivo human lung tissues (n = 104,577 cells in total). Note that alveolar lineage cells, such as alveolar type 1 (AT1) and type 2 (AT2) cells, are identified only in in vivo lungs and not in in vitro cultures. **(B)** Pie charts displaying the percentage of various cell types in each culture method and tissue. Only airway cells are considered when calculating the proportion of cell types in in vivo lungs.

I found that the Gene Set Enrichment Analysis (GSEA) of Differentially Expressed Genes (DEGs) indicated enhanced gene expression patterns relevant to immune responses in in vivo lungs by comparing the cellular molecular programs of in vitro airway cells with those of lung tissues. This observation is most likely attributable to a distinctive property of in vitro airway models, which are able to correctly mimic in vivo tissue features despite being made up entirely of epithelial cells rather than immune cells and being the single component of the model. It is interesting to note that only certain fractions of basal and ciliated cells displayed increased expressions of

immune response genes, which suggests the existence of "immune primed cell states."

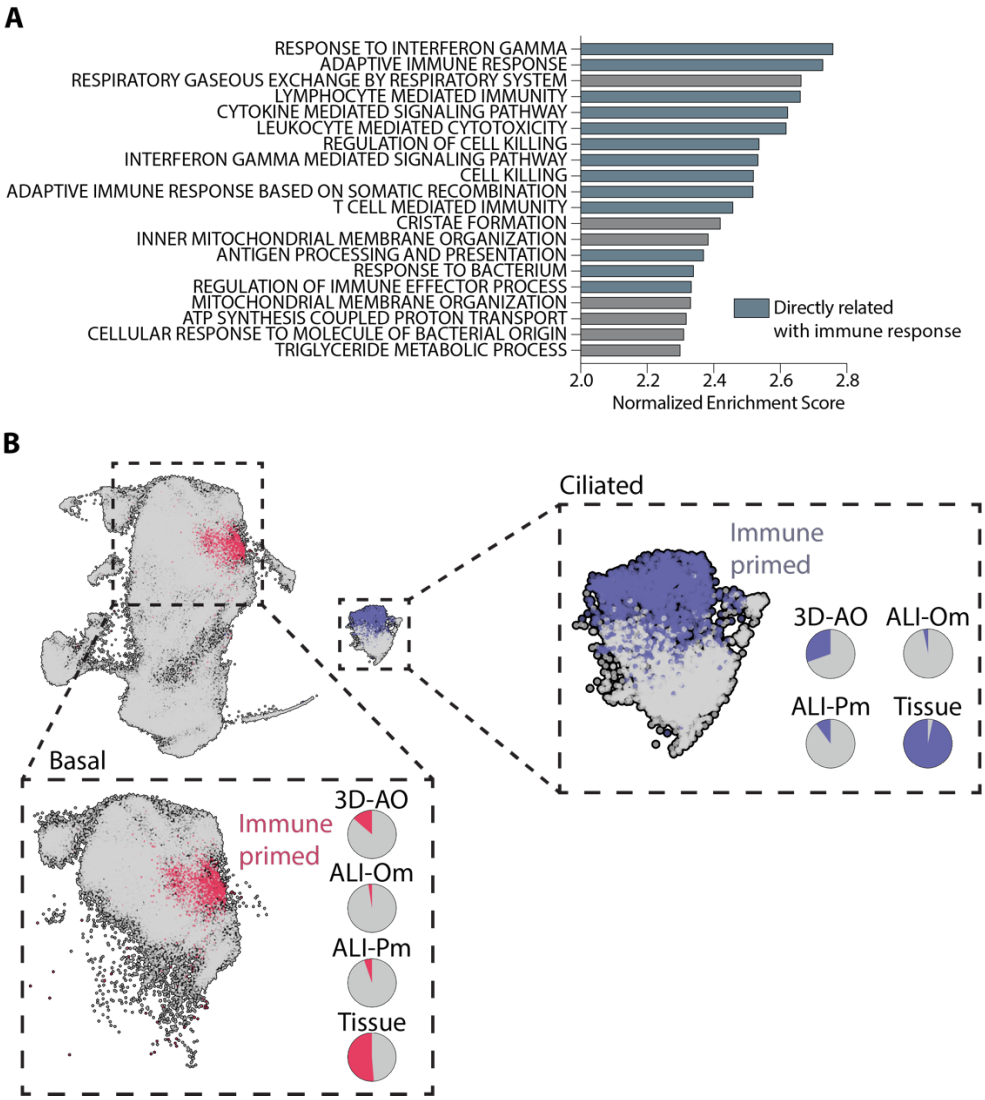

Figure 15. Immune primed cells in human tissues.
(A) Bar plot displaying pathways more enriched in in vivo lungs compared to in vitro airway models, sorted by normalized enrichment scores computed through GSEA of DEGs. In vivo lungs exhibit activation of biological processes. **(B)** UMAP illustrates unique immune-primed cell states present solely in basal and ciliated cells. Pie charts demonstrate a higher proportion of immune-primed basal and ciliated cells in in vivo lungs compared to in vitro lung cells.

After that, I investigated the therapeutic significance of immune primed cells in chronic lung diseases such as idiopathic pulmonary fibrosis (IPF), chronic obstructive pulmonary disease (COPD), and cystic fibrosis (CF) by integrating my single-cell RNA sequencing (scRNA-seq) data with the findings of previous research (Carraro et al. 2021; Adams et al. 2020; Sauler et al. 2022). In sick lungs, I discovered a much higher frequency of immune primed subsets in basal cells as compared to healthy controls; however, immune primed cell states were extremely uncommon in fetal lung tissues (Miller et al. 2020).

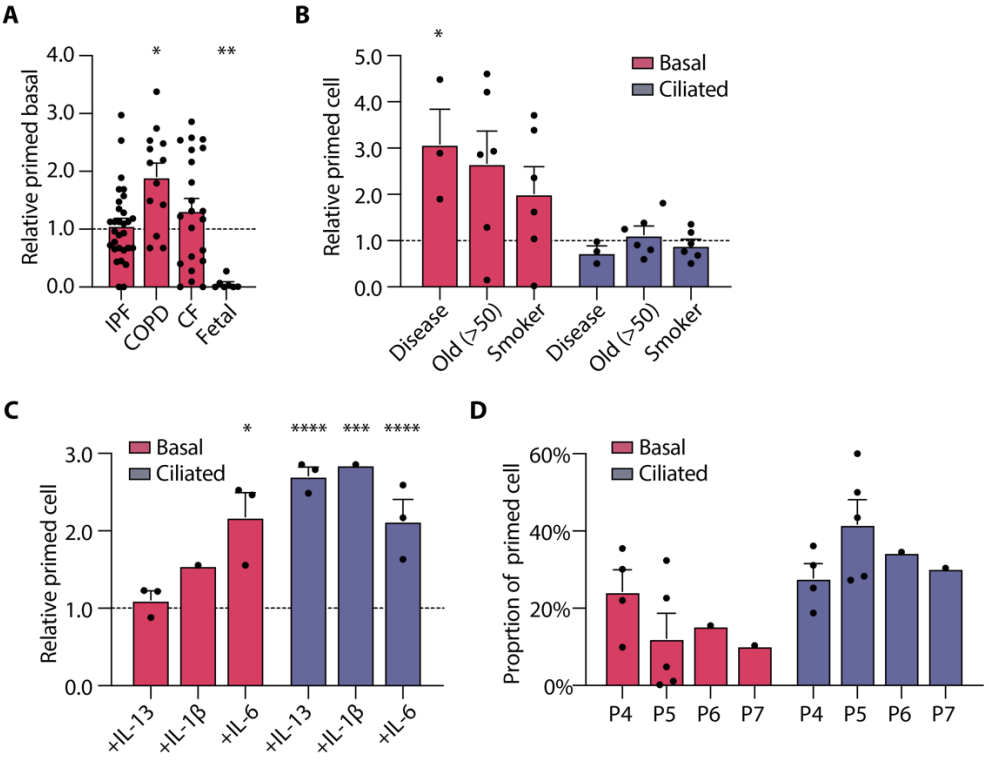

Figure 16. Proportions of immune primed cells.
(A) The relative percentages of immune-primed basal cells in in vivo human

lung single-cell transcriptome data with chronic respiratory diseases, normalized against their healthy controls. **(B and C)** The relative percentages of immune-primed basal and ciliated cells in 3D-AO, considering donor characteristics **(B)** and cytokine treatments **(C)**. Each group is normalized against healthy **(B)** and vehicle-treated **(C)** samples. **(D)** The proportion of immune-primed basal and ciliated cells over multiple passages in 3D-AO. *P < 0.05, **P <0.01, ***P < 0.001, ****P < 0.0001, unpaired t-test.

Additionally, I investigated immune cells that were primed in airway organoids that were derived from the lung tissue of healthy individuals as well as individuals who suffered from chronic respiratory diseases. When compared to organoids derived from healthy lung tissues, those derived from diseased lung tissues displayed a statistically significant increase in the number of immune primed basal cells. Additionally, organoids obtained from older donors or smokers displayed a higher number of immune primed basal cells in comparison to organoids obtained from younger donors or donors who did not smoke.

This demonstrates that human primary airway organoids can imitate in vivo immune responses by modifying the characteristics of the culture microenvironment. The 3D-AOs were created from healthy lungs and treated with pro-inflammatory cytokines such as IL-1 and IL-6, which induced immune primed basal and ciliated cells. It is important to note that ciliated cells displayed a greater responsiveness to inflammatory stimuli than basal cells did. I also found a steady drop in the number of immune primed basal cells in 3D-AO across numerous passages, which suggests that cell states that were generated in response to signals from local microenvironments may be

lost during long-term in vitro cultures.

In a nutshell, my findings highlight the distinctive transcriptome signature of immune response in human lung tissues cultured in vivo in comparison to organoid models cultured in vitro. These findings provide invaluable insights into the intricate interplay between epithelial and immune cells in lung diseases, as well as the potential for the development of more physiologically relevant models for the study of airway diseases and the identification of therapeutic targets.

3.5. Airway organoid responses to viral infections

For the purpose of this investigation, I made use of my platform to mimic a number of respiratory viral infections and then compared the transcriptome signatures of infected human primary airway epithelium. When testing for infection with SARS-CoV-2 variants (Alpha, Beta, Delta, and Omicron), MERS-CoV, and Influenza A viruses, I used established ALI-Om that was derived from a total of 20 individual tissue donors. Initially, I investigated the levels of expression of the ACE2 and TMPRSS2 genes, both of which are necessary for SARS-CoV-2 viral entry, in ALI-Om while it was undergoing the process of air-liquid differentiation. When the virus was delivered three weeks after differentiation induction, the levels of expression for both ACE2 and TMPRSS2 rose.

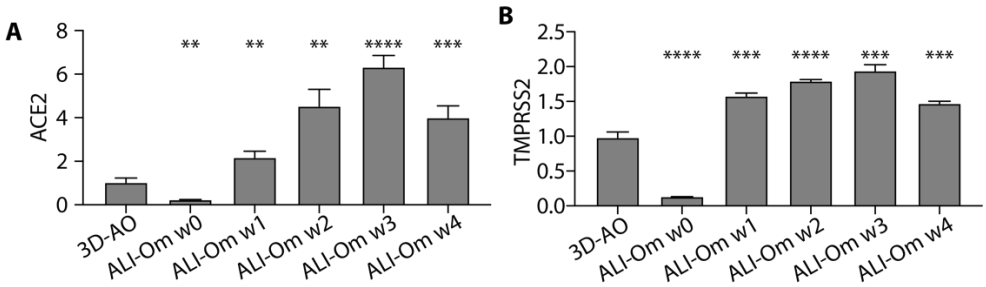

Figure 17. Expression levels or viral entry genes.
(A and B) Relative expression levels (2-ddCt) measured by quantitative PCR of two viral entry genes. **P < 0.01, ***P < 0.001, ****P < 0.0001, unpaired t test between 3D-AO and each ALI-Om sample.

At 72 hours after infection, a transmission electron microscopy (TEM) examination revealed a widespread presence of SARS-CoV-2 alpha virus particles in ALI-Om. By using immunofluorescence labelling for the nucleocapsid protein (NP), I was able to demonstrate that these cells had been successfully infected with the SARS-CoV-2 alpha and Influenza A viruses.

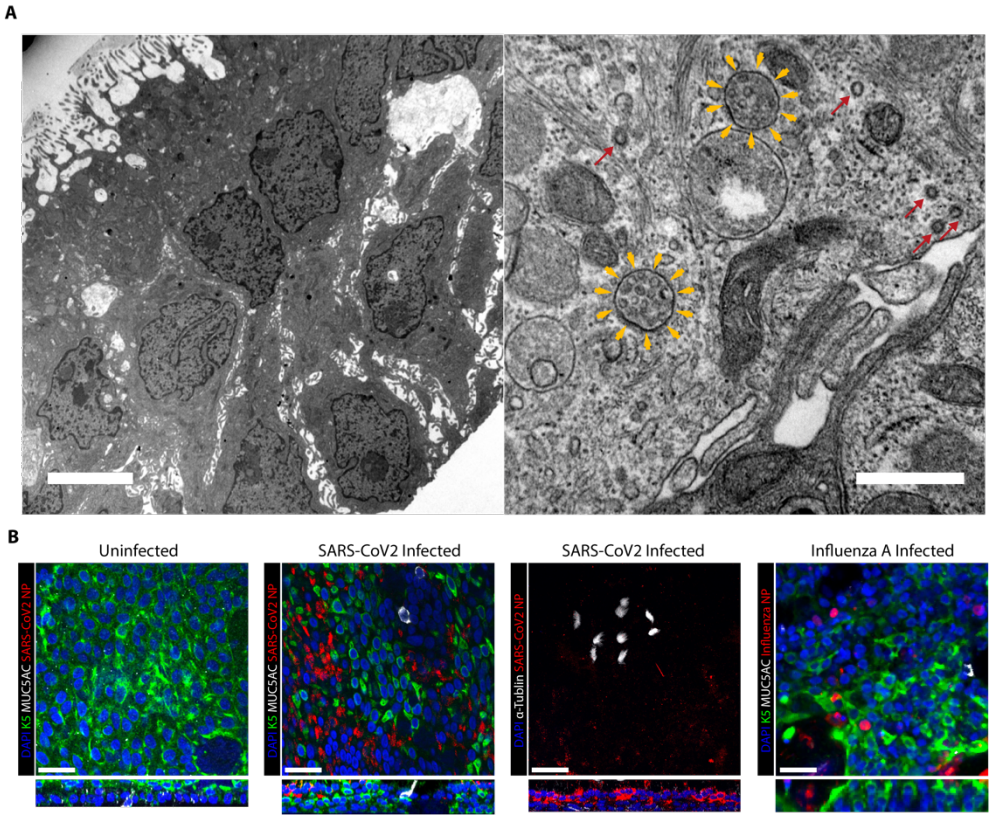

Figure 18. Viral infection on ALI-Om.
(A) This image depicts an SARS-CoV-2 alpha infection in ALI-Om, captured using transmission electron microscopy (TEM). Virus particles are highlighted by red arrows, while virus vesicles are indicated by yellow arrows. The scale bars on the left and right represent 5 μm and 500 nm, respectively. **(B)** Immunofluorescence (IF) images show three types of ALI-Oms: uninfected (left), SARS-CoV-2 infected (middle), and Influenza A infected (right). The scale bar on the image is 30 μm.

On ALI-Om that was infected with SARS-CoV-2 variants, MERS-CoV, and Influenza A viruses, I carried out an analysis using scRNA-seq. Through analysis of aligned reads in relation to the genome of the viral reference strain, I was able to calculate the infection rate. The infection rate of ALI-Om that was isolated from diseased lung tissues exhibited a

considerable increase when compared to the infection rate of ALI-Om that was isolated from healthy lung tissues. On the other hand, neither the age nor the gender of the people who donated lung tissue had a discernible impact on the infection rate.

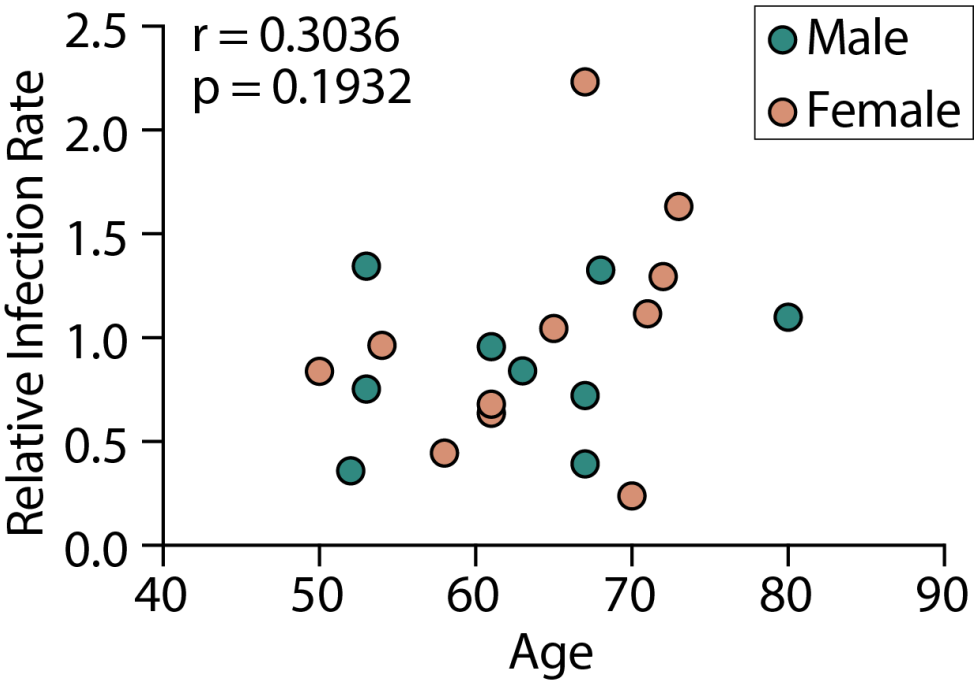

Figure 19. Relative infection rate.
The correlation between age, sex, and the infection rate of SARS-CoV-2 viruses in each sample was assessed using a two-tailed Pearson correlation analysis.

I used my single-cell transcriptome platform to analyze the single-cell data obtained from infected samples as a reference dataset for controls that were not subjected to any treatment. The percentage of different cell types was not significantly altered by viral infection, with the exception of the omicron variant, which caused a reduction in the number of basal-1 cell

populations. Notably, ALI-Om that had been infected with SARS-CoV-2 Beta, Omicron, and Influenza A demonstrated significant increases in immune primed basal subsets. On the immune primed cell states that were generated by viral infection in ALI-Om, I did not find any significant influence of donor age, sex, or disease history. My infected cells, in particular those infected with SARS-CoV-2 Omicron, MERS-CoV, and Influenza A viruses, demonstrated increased levels of interferon as well as increased viral entrance scores.

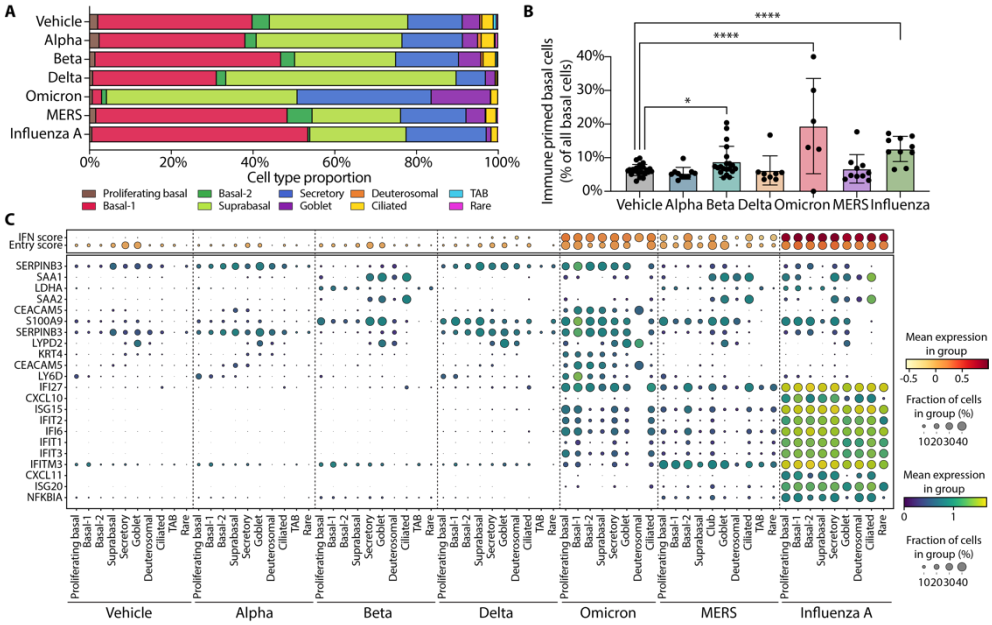

Figure 20. Transcriptomic responses of organoids to viruses.
(A) The distribution of cell types in each infection group is presented in stacked bar charts. **(B)** The ratio of immune primed basal cells in each infection group is represented by a bar plot, accompanied by a 95% confidence interval (CI). Statistical significance (*P < 0.05, ****P < 0.0001) was determined using an unpaired t-test between the vehicle group and each infection group. **(C)** The expression levels of immune-responsive marker genes and the proportion of cells expressing these genes are visualized in a dot plot, categorized by cell type and infection groups.

In conclusion, my research shows that human airway organoids are able to respond to respiratory virus infections at the transcriptome level. This finding provides important insights into the cellular and molecular mechanisms that underlie host-pathogen interactions as well as potential therapeutic targets for the treatment of respiratory viral diseases.

Chapter 4. Conclusion

The use of human primary organoids has a tremendous amount of untapped potential as a research platform for human biology, particularly for the purpose of better comprehending the intricate biological processes that take place in human tissues in vivo. During the course of my research, I created a comprehensive single-cell atlas of three typical in vitro human airway models. Additionally, I devised a method for comparing the datasets obtained from these in vitro human airway models to those derived from in vivo human lungs. Both of these contributions were made as part of my investigation.

My findings provide compelling evidence that, despite differences in the percentage of cellular composition, the cellular diversity as well as the transcriptional states of the cells are significantly maintained across the board in in vitro models. This is a significant discovery because it lends credence to the concept that it is possible to successfully carry out comparisons between the various models. In addition, the maintenance of cellular diversity even after numerous passages, up to ten of them, hints at the durability and dependability of in vitro human primary airway models for the purposes of further research and application.

I found a separate population of basal cells that have unique stem cell

characteristics; hence, additional validation of this population is required. In addition, I found a tiny percentage of secretory cells that expressed SCGB3A2. These cells could be potential progenitor cells in human distal terminal bronchioles, as they were just recently recognized as having these properties. In addition, I was able to identify intermediate cell states that were associated with two distinct lineages. These findings shed light on the creation of intermediate cell states throughout the process of lung regeneration as well as their possible therapeutic applications in the treatment of human lung disorders such as pulmonary fibrosis and lung cancer.

I was able to capture the differentiation processes of rare epithelial cell types, such as PNECs, ionocytes, and tuft cells, from basal cells as a result of my large-scale single-cell profiling of in vitro multi-culture models. These findings were comparable to those that have been described in the scRNA-seq analysis of raw airway tissues earlier. This study indicates the capability of my technology to mimic the lineage linkages of human lung tissues in vivo, and it also demonstrates the potential to gain unique insights into the biology of the lung and the disease processes that affect it.

Because in vitro models cannot fully replicate the dynamics of the lung epithelium in vivo, which is continuously subjected to a variety of environmental stresses, one of the most important aspects of my study was the analysis of the impact that immune cells have on the processes that occur within the lung epithelium. I was able to identify unique immune primed subsets of basal and ciliated cells that were enriched in human lung tissues

compared to in vitro airway preparations. It is interesting to note that I found a greater number of immune primed cell states in 3D-AO in comparison to ALI-Om and ALI-Pm. This leads me to hypothesize that the dimensionality of the extracellular matrix may play a part in the in vitro maintenance of immune responsive cells. Donor features, such as smoking history and lung illness, as well as age, were found to alter the status of immune primed cells in organoids. This suggests that donor characteristics play an important role in the behavior of these cells in vitro, and that cellular states can be modified by elements in the microenvironment.

I evaluated the transcriptome responses of human primary airway epithelial cells to diverse respiratory viral infections using in vitro organoids. These respiratory virus infections included SARS-CoV-2 variants, MERS-CoV, and Influenza A viruses. I found that demographic characteristics such as age, gender, or a history of lung disease did not have a significant impact on the infection rate. However, one of the things that I discovered was that infected cells displayed increased levels of interferon responses. This was especially the case in cells associated with less severe symptoms. This result may have crucial ramifications for understanding the severity of the disease as well as the potential treatment options, and it also offers up new possibilities for further inquiry into the role that interferon signaling plays in the setting of respiratory viral infections.

In conclusion, I have established a reliable database that has the potential to be used to get a deeper comprehension of human lung stem cell

activities and lineage trajectories. This was accomplished by combining single-cell multiomics data with in vitro human primary organoids obtained from numerous individuals. My method has the potential to serve as a pre-clinical model for directing optimized personalized therapeutics based on molecular profiling, which would ultimately help to hasten the deployment of effective treatment regimens for both known and new viral infections. In addition, my research demonstrates the potential of in vitro organoid models as strong tools for examining the effects of environmental factors and a variety of pathological situations on lung cell states as well as the responses of lung cells to infections or inflammations.

The use of organoids as a base for the testing of drugs and the development of personalized medical treatments has the potential to fundamentally alter how we approach the treatment of lung illnesses. For instance, organoids can be utilized in the context of chronic respiratory diseases such as asthma or chronic obstructive pulmonary disease (COPD) in order to evaluate the efficacy and safety of novel therapeutic interventions, as well as to identify patient-specific factors that influence treatment response. This would make it possible to take a more targeted and individualized approach to therapy, which would eventually result in better outcomes for patients.

In addition, the findings that I have obtained highlight the significance of including immune cells and inflammatory stimuli inside in vitro models in order to more accurately replicate the conditions that exist

within the lungs of living animals. To further understand the interactions that occur between lung epithelial cells and immune cells during homeostasis and disease, future research may investigate the use of co-culture methods, which involve the introduction of immune cells into an organoid model. This technique would provide significant insights into the intricate interplay between the immune system and lung tissue, which is critical for understanding the pathogenesis of many lung diseases and identifying novel therapeutic targets.

Investigating the possibility for distinct signaling pathways to communicate with one another in the setting of lung cell differentiation and regeneration is an additional fascinating direction that future research may take. For instance, I found evidence that the Wnt signaling pathway may have a part to play in the development of Basal-2, TAB, and rare cells. This suggests that these cell types may have a function in the body. Additional research into these signaling pathways and the ways in which they interact with one another may yield important insights into the molecular mechanisms that underlie the differentiation of lung cells and the homeostasis of lung tissue.

In a nutshell, the findings of my research have laid a rock-solid groundwork for the analysis of human lung biology and the modeling of lung illnesses that may be carried out with the assistance of in vitro human primary organoids as a platform. We can unlock their full potential as instruments for improving lung research and establishing more effective, individualized

treatments for patients who are afflicted with lung disorders if we further refine and broaden our understanding of these models.

Abstract

인체의 다양한 장기 중 폐는 그 중요성에 비해 발달 과정이나 구성하는 세포들의 특성이 명확히 규명되지 않은 장기 중 하나이다. 기체교환에 중요하게 기능하는 폐포의 단순한 세포 구성과는 다르게 기도는 아주 다양한 세포들로 구성 되어있고 아직까지도 새로운 세포가 발견되고 있다. 또한 기도를 구성하는 세포들은 단순히 공기의 통로로서 기능하는 것이 아니라 면역과정에 중요한 역할을 한다는 것 또한 밝혀지고 있다. 단일세포 전사체 분석을 필두로 하는 현대의 세포 생물학, 유전체학 연구에서 기도의 세포들에 대한 연구가 어려운 주요한 이유는 폐 전체 중 기도 세포들이 차지하는 비중이 적어서 단일세포 전사체 분석으로 충분한 수의 세포를 분석하기 어렵다는 점이다.

최근 개발된 오가노이드 모델은 in vitro 환경에서 체내의 세포들의 분화를 모사하면서 상피세포만을 집중하여 연구할 수 있다는 장점을 가지고 있다. 특히 인간 폐 기도 오가노이드는 basal 세포부터 시작하여 기도에 존재하는 모든 세포를 분열과 분화를 반복하며 배양할 수 있다. 또한 iPSC 또는 ESC이 아닌 폐 조직의 성체줄기세포에서 유래한 오가노이드를 수립하면 다양한 사람들의 폐 기도 오가노이드를 비교할 수 있다는 장점을 가진다.

본 연구에서는 83명의 공여자로부터 폐 기도 성체줄기세포유래 오가노이드를 수립하였고 27명의 공여자 샘플로부터 약 30만 세포의

단일세포 전사체 분석을 수행하였다. 이를 통해 세계 최대 규모의 오가노이드 데이터베이스를 구축하여 다양한 연구자들의 연구에 기여할 수 있게 되었다. 그뿐 아니라 Basal-2와 TAB 세포를 새롭게 발견하여 극소수만 존재하여 연구가 어렵던 Tuft, Ionocyte, PNEC 세포의 분화 과정에 대한 기전을 처음으로 밝혔다. 또한 면역세포만이 관여한다고 생각되던 기존의 관점에서 벗어나 폐 기도 상피세포에도 immune-primed 세포가 있다는 것을 알게 되었다.

이 연구를 통하여 나는 앞으로 다양한 폐 연구에 필요한 폐 기도 오가노이드의 정확한 특성들을 다방면에서 밝혀냈고, 오가노이드 연구의 새로운 가능성을 보여줄 수 있었다. 또한 대규모 데이터베이스를 공개하고 사용자 친화적인 웹사이트를 만들어 다양한 단일세포 전사체 데이터를 나의 폐 기도 오가노이드 지도와 비교해 볼 수 있게 하였다. 기존에 존재하지 않던 대규모의 데이터를 통해 새로운 세포를 찾고 면역에 관여하는 상태를 찾아서 이후의 후속연구를 통해 폐 기도의 전반적인 기전을 이해하는데 기여하였다고 생각한다.

Bibliography

Adams, T. S., J. C. Schupp, S. Poli, E. A. Ayaub, N. Neumark, F. Ahangari, S. G. Chu, B. A. Raby, G. DeIuliis, M. Januszyk, Q. Duan, H. A. Arnett, A. Siddiqui, G. R. Washko, R. Homer, X. Yan, I. O. Rosas, and N. Kaminski. 2020. 'Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis', *Sci Adv*, 6: eaba1983.

Ali, M., M. K. Mazumder, and T. B. Martonen. 2009. 'Measurements of electrodynamic effects on the deposition of MDI and DPI aerosols in a replica cast of human oral-pharyngeal-laryngeal airways', *J Aerosol Med Pulm Drug Deliv*, 22: 35-44.

Awatade, N. T., S. L. Wong, C. K. Hewson, L. K. Fawcett, A. Kicic, A. Jaffe, and S. A. Waters. 2018. 'Human Primary Epithelial Cell Models: Promising Tools in the Era of Cystic Fibrosis Personalized Medicine', *Front Pharmacol*, 9: 1429.

Basil, M. C., F. L. Cardenas-Diaz, J. J. Kathiriya, M. P. Morley, J. Carl, A. N. Brumwell, J. Katzen, K. J. Slovik, A. Babu, S. Zhou, M. M. Kremp, K. B. McCauley, S. Li, J. D. Planer, S. S. Hussain, X. Liu, R. Windmueller, Y. Ying, K. M. Stewart, M. Oyster, J. D. Christie, J. M. Diamond, J. F. Engelhardt, E. Cantu, S. M. Rowe, D. N. Kotton, H. A. Chapman, and E. E. Morrisey. 2022. 'Human distal airways contain a multipotent secretory cell that can regenerate alveoli', *Nature*, 604: 120-26.

Berical, A., R. E. Lee, S. H. Randell, and F. Hawkins. 2019. 'Challenges Facing Airway Epithelial Cell-Based Therapy for Cystic Fibrosis', *Front Pharmacol*, 10: 74.

Bibikova, M., K. Beumer, J. K. Trautman, and D. Carroll. 2003. 'Enhancing gene targeting with designed zinc finger nucleases', *Science*, 300: 764.

Brennecke, P., S. Anders, J. K. Kim, A. A. Kolodziejczyk, X. Zhang, V. Proserpio, B. Baying, V. Benes, S. A. Teichmann, J. C. Marioni, and M. G. Heisler. 2013. 'Accounting for technical noise in single-cell RNA-seq experiments', *Nat Methods*, 10: 1093-5.

Bustamante-Marin, X. M., and L. E. Ostrowski. 2017. 'Cilia and Mucociliary Clearance', *Cold Spring Harb Perspect Biol*, 9.

Calvert, B. A., and A. L. Ryan Firth. 2020. 'Application of iPSC to Modelling of Respiratory Diseases', *Adv Exp Med Biol*, 1237: 1-16.

Carraro, G., J. Langerman, S. Sabri, Z. Lorenzana, A. Purkayastha, G. Zhang, B. Konda, C. J. Aros, B. A. Calvert, A. Szymaniak, E. Wilson, M. Mulligan, P. Bhatt, J. Lu, P. Vijayaraj, C. Yao, D. W. Shia, A. J. Lund, E. Israely, T. M. Rickabaugh, J. Ernst, M. Mense, S. H. Randell, E. K. Vladar, A. L. Ryan, K. Plath, J. E. Mahoney, B. R. Stripp, and B. N. Gomperts. 2021. 'Transcriptional analysis of cystic fibrosis airways at

single-cell resolution reveals altered epithelial cell states and composition', *Nat Med*, 27: 806-14.

Chen, J., C. Li, Z. Ristovski, A. Milic, Y. Gu, M. S. Islam, S. Wang, J. Hao, H. Zhang, C. He, H. Guo, H. Fu, B. Miljevic, L. Morawska, P. Thai, Y. F. Lam, G. Pereira, A. Ding, X. Huang, and U. C. Dumka. 2017. 'A review of biomass burning: Emissions and impacts on air quality, health and climate in China', *Sci Total Environ*, 579: 1000-34.

Cho, S. W., S. Kim, J. M. Kim, and J. S. Kim. 2013. 'Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease', *Nat Biotechnol*, 31: 230-2.

Cong, L., F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini, and F. Zhang. 2013. 'Multiplex genome engineering using CRISPR/Cas systems', *Science*, 339: 819-23.

Deng, Q., D. Ramskold, B. Reinius, and R. Sandberg. 2014. 'Single-cell RNA-seq reveals dynamic, random monoallelic gene expression in mammalian cells', *Science*, 343: 193-6.

Fiege, J. K., J. M. Thiede, H. A. Nanda, W. E. Matchett, P. J. Moore, N. R. Montanari, B. K. Thielen, J. Daniel, E. Stanley, R. C. Hunter, V. D. Menachery, S. S. Shen, T. D. Bold, and R. A. Langlois. 2021. 'Single cell resolution of SARS-CoV-2 tropism, antiviral responses, and susceptibility to therapies in primary human airway epithelium', *PLoS Pathog*, 17: e1009292.

Fulcher, M. L., S. Gabriel, K. A. Burns, J. R. Yankaskas, and S. H. Randell. 2005. 'Well-differentiated human airway epithelial cell cultures', *Methods Mol Med*, 107: 183-206.

Gallo, O., L. G. Locatello, A. Mazzoni, L. Novelli, and F. Annunziato. 2021. 'The central role of the nasal microenvironment in the transmission, modulation, and clinical progression of SARS-CoV-2 infection', *Mucosal Immunol*, 14: 305-16.

Gamage, A. M., K. S. Tan, W. O. Y. Chan, J. Liu, C. W. Tan, Y. K. Ong, M. Thong, A. K. Andiappan, D. E. Anderson, Y. Wang, and L. F. Wang. 2020. 'Infection of human Nasal Epithelial Cells with SARS-CoV-2 and a 382-nt deletion isolate lacking ORF8 reveals similar viral kinetics and host transcriptional profiles', *PLoS Pathog*, 16: e1009130.

Gengler, I., J. C. Wang, M. M. Speth, and A. R. Sedaghat. 2020. 'Sinonasal pathophysiology of SARS-CoV-2 and COVID-19: A systematic review of the current evidence', *Laryngoscope Investig Otolaryngol*, 5: 354-59.

Giesen, C., H. A. Wang, D. Schapiro, N. Zivanovic, A. Jacobs, B. Hattendorf, P. J. Schuffler, D. Grolimund, J. M. Buhmann, S. Brandt, Z. Varga, P. J. Wild, D. Gunther, and B. Bodenmiller. 2014. 'Highly multiplexed imaging of tumor tissues with subcellular resolution by mass cytometry', *Nat Methods*, 11: 417-22.

Goldfarbmuren, K. C., N. D. Jackson, S. P. Sajuthi, N. Dyjack, K. S. Li, C. L. Rios, E. G. Plender, M. T. Montgomery, J. L. Everman, P. E. Bratcher, E. K. Vladar, and M. A. Seibold. 2020. 'Dissecting the cellular specificity of smoking effects and reconstructing lineages in the human airway epithelium', *Nat Commun*, 11: 2485.

Gomes, K. M., I. C. Costa, J. F. Santos, P. M. Dourado, M. F. Forni, and J. C. Ferreira. 2017. 'Induced pluripotent stem cells reprogramming: Epigenetics and applications in the regenerative medicine', *Rev Assoc Med Bras (1992)*, 63: 180-89.

Gouinaud, L., I. Katz, A. Martin, J. Hazebroucq, J. Texereau, and G. Caillibotte. 2015. 'Inhalation pressure distributions for medical gas mixtures calculated in an infant airway morphology model', *Comput Methods Biomech Biomed Engin*, 18: 1358-66.

Gras, D., A. Bourdin, I. Vachier, L. de Senneville, C. Bonnans, and P. Chanez. 2012. 'An ex vivo model of severe asthma using reconstituted human bronchial epithelium', *J Allergy Clin Immunol*, 129: 1259-66 e1.

Habermann, A. C., A. J. Gutierrez, L. T. Bui, S. L. Yahn, N. I. Winters, C. L. Calvi, L. Peter, M. I. Chung, C. J. Taylor, C. Jetter, L. Raju, J. Roberson, G. Ding, L. Wood, J. M. S. Sucre, B. W. Richmond, A. P. Serezani, W. J. McDonnell, S. B. Mallal, M. J. Bacchetta, J. E. Loyd, C. M. Shaver, L. B. Ware, R. Bremner, R. Walia, T. S. Blackwell, N. E. Banovich, and J. A. Kropski. 2020. 'Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis', *Sci Adv*, 6: eaba1972.

Haefeli-Bleuer, B., and E. R. Weibel. 1988. 'Morphometry of the human pulmonary acinus', *Anat Rec*, 220: 401-14.

Haffter, P., and C. Nusslein-Volhard. 1996. 'Large scale genetics in a small vertebrate, the zebrafish', *Int J Dev Biol*, 40: 221-7.

Hasenpusch, G., J. Geiger, K. Wagner, O. Mykhaylyk, F. Wiekhorst, L. Trahms, A. Heidsieck, B. Gleich, C. Bergemann, M. K. Aneja, and C. Rudolph. 2012. 'Magnetized aerosols comprising superparamagnetic iron oxide nanoparticles improve targeted drug and gene delivery to the lung', *Pharm Res*, 29: 1308-18.

He, P., K. Lim, D. Sun, J. P. Pett, Q. Jeng, K. Polanski, Z. Dong, L. Bolt, L. Richardson, L. Mamanova, M. Dabrowska, A. Wilbrey-Clark, E. Madissoon, Z. K. Tuong, E. Dann, C. Suo, I. Goh, M. Yoshida, M. Z. Nikolic, S. M. Janes, X. He, R. A. Barker, S. A. Teichmann, J. C. Marioni, K. B. Meyer, and E. L. Rawlins. 2022. 'A human fetal lung cell atlas uncovers proximal-distal gradients of differentiation and key regulators of epithelial fates', *Cell*, 185: 4841-60 e25.

Hewitt, R. J., and C. M. Lloyd. 2021. 'Regulation of immune responses by the airway epithelial cell landscape', *Nat Rev Immunol*, 21: 347-62.

Hiemstra, P. S., G. Grootaers, A. M. van der Does, C. A. M. Krul, and I. M. Kooter. 2018. 'Human lung epithelial cell cultures for analysis of inhaled toxicants: Lessons learned and future directions', *Toxicol In Vitro*, 47: 137-46.

Hou, Y. J., K. Okuda, C. E. Edwards, D. R. Martinez, T. Asakura, K. H. Dinnon, 3rd, T. Kato, R. E. Lee, B. L. Yount, T. M. Mascenik, G. Chen, K. N. Olivier, A. Ghio, L. V. Tse, S. R. Leist, L. E. Gralinski, A. Schafer, H. Dang, R. Gilmore, S. Nakano, L. Sun, M. L. Fulcher, A. Livraghi-Butrico, N. I. Nicely, M. Cameron, C. Cameron, D. J. Kelvin, A. de Silva, D. M. Margolis, A. Markmann, L. Bartelt, R. Zumwalt, F. J. Martinez, S. P. Salvatore, A. Borczuk, P. R. Tata, V. Sontake, A. Kimple, I. Jaspers,

W. K. O'Neal, S. H. Randell, R. C. Boucher, and R. S. Baric. 2020. 'SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract', *Cell*, 182: 429-46 e14.

Hynds, R. E., P. Bonfanti, and S. M. Janes. 2018. 'Regenerating human epithelia with cultured stem cells: feeder cells, organoids and beyond', *EMBO Mol Med*, 10: 139-50.

Ismail, M., A. Comerford, and W. A. Wall. 2013. 'Coupled and reduced dimensional modeling of respiratory mechanics during spontaneous breathing', *Int J Numer Method Biomed Eng*, 29: 1285-305.

Jemal, A., F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman. 2011. 'Global cancer statistics', *CA Cancer J Clin*, 61: 69-90.

Jorissen, M., and T. Willems. 2004. 'The secondary nature of ciliary (dis)orientation in secondary and primary ciliary dyskinesia', *Acta Otolaryngol*, 124: 527-31.

Kadur Lakshminarasimha Murthy, P., V. Sontake, A. Tata, Y. Kobayashi, L. Macadlo, K. Okuda, A. S. Conchola, S. Nakano, S. Gregory, L. A. Miller, J. R. Spence, J. F. Engelhardt, R. C. Boucher, J. R. Rock, S. H. Randell, and P. R. Tata. 2022. 'Human distal lung maps and lineage hierarchies reveal a bipotent progenitor', *Nature*, 604: 111-19.

Katz, I. M., A. R. Martin, P. A. Muller, K. Terzibachi, C. H. Feng, G. Caillibotte, J. Sandeau, and J. Texereau. 2011. 'The ventilation distribution of helium-oxygen mixtures and the role of inertial losses in the presence of heterogeneous airway obstructions', *J Biomech*, 44: 1137-43.

Kim, J., B. K. Koo, and J. A. Knoblich. 2020. 'Human organoids: model systems for human biology and medicine', *Nat Rev Mol Cell Biol*, 21: 571-84.

Kolodziejczyk, A. A., J. K. Kim, J. C. Tsang, T. Ilicic, J. Henriksson, K. N. Natarajan, A. C. Tuck, X. Gao, M. Buhler, P. Liu, J. C. Marioni, and S. A. Teichmann. 2015. 'Single Cell RNA-Sequencing of Pluripotent States Unlocks Modular Transcriptional Variation', *Cell Stem Cell*, 17: 471-85.

Lambert, A. R., P. O'Shaughnessy, M. H. Tawhai, E. A. Hoffman, and C. L. Lin. 2011. 'Regional deposition of particles in an image-based airway model: large-eddy simulation and left-right lung ventilation asymmetry', *Aerosol Sci Technol*, 45: 11-25.

Lancaster, M. A., M. Renner, C. A. Martin, D. Wenzel, L. S. Bicknell, M. E. Hurles, T. Homfray, J. M. Penninger, A. P. Jackson, and J. A. Knoblich. 2013. 'Cerebral organoids model human brain development and microcephaly', *Nature*, 501: 373-9.

Laumbach, R. J., and H. M. Kipen. 2012. 'Respiratory health effects of air pollution: update on biomass smoke and traffic pollution', *J Allergy Clin Immunol*, 129: 3-11; quiz 12-3.

Leung, H. M., S. E. Birket, C. Hyun, T. N. Ford, D. Cui, G. M. Solomon, R. J. Shei, A. T. Adewale, A. R. Lenzie, C. M. Fernandez-Petty, H. Zheng, J. H. Palermo, D. Y. Cho, B. A. Woodworth, L. M. Yonker, B. P. Hurley, S. M. Rowe, and G. J. Tearney. 2019. 'Intranasal micro-optical coherence tomography imaging for cystic fibrosis studies', *Sci Transl Med*, 11.

Liu, Y., and M. Rao. 2011. 'Gene targeting in human pluripotent stem cells', *Methods Mol Biol*, 767: 355-67.

Mahata, B., X. Zhang, A. A. Kolodziejczyk, V. Proserpio, L. Haim-Vilmovsky, A. E. Taylor, D. Hebenstreit, F. A. Dingler, V. Moignard, B. Gottgens, W. Arlt, A. N. McKenzie, and S. A. Teichmann. 2014. 'Single-cell RNA sequencing reveals T helper cells synthesizing steroids de novo to contribute to immune homeostasis', *Cell Rep*, 7: 1130-42.

Mali, P., L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, and G. M. Church. 2013. 'RNA-guided human genome engineering via Cas9', *Science*, 339: 823-6.

Miller, A. J., Q. Yu, M. Czerwinski, Y. H. Tsai, R. F. Conway, A. Wu, E. M. Holloway, T. Walker, I. A. Glass, B. Treutlein, J. G. Camp, and J. R. Spence. 2020. 'In Vitro and In Vivo Development of the Human Airway at Single-Cell Resolution', *Dev Cell*, 53: 117-28 e6.

Miller, J. C., S. Tan, G. Qiao, K. A. Barlow, J. Wang, D. F. Xia, X. Meng, D. E. Paschon, E. Leung, S. J. Hinkley, G. P. Dulay, K. L. Hua, I. Ankoudinova, G. J. Cost, F. D. Urnov, H. S. Zhang, M. C. Holmes, L. Zhang, P. D. Gregory, and E. J. Rebar. 2011. 'A TALE nuclease architecture for efficient genome editing', *Nat Biotechnol*, 29: 143-8.

Morrisey, E. E., and B. L. Hogan. 2010. 'Preparing for the first breath: genetic and cellular mechanisms in lung development', *Dev Cell*, 18: 8-23.

Mullins, M. C., M. Hammerschmidt, P. Haffter, and C. Nusslein-Volhard. 1994. 'Large-scale mutagenesis in the zebrafish: in search of genes controlling development in a vertebrate', *Curr Biol*, 4: 189-202.

Muus, C., M. D. Luecken, G. Eraslan, L. Sikkema, A. Waghray, G. Heimberg, Y. Kobayashi, E. D. Vaishnav, A. Subramanian, C. Smillie, K. A. Jagadeesh, E. T. Duong, E. Fiskin, E. Torlai Triglia, M. Ansari, P. Cai, B. Lin, J. Buchanan, S. Chen, J. Shu, A. L. Haber, H. Chung, D. T. Montoro, T. Adams, H. Aliee, S. J. Allon, Z. Andrusivova, I. Angelidis, O. Ashenberg, K. Bassler, C. Becavin, I. Benhar, J. Bergenstrahle, L. Bergenstrahle, L. Bolt, E. Braun, L. T. Bui, S. Callori, M. Chaffin, E. Chichelnitskiy, J. Chiou, T. M. Conlon, M. S. Cuoco, A. S. E. Cuomo, M. Deprez, G. Duclos, D. Fine, D. S. Fischer, S. Ghazanfar, A. Gillich, B. Giotti, J. Gould, M. Guo, A. J. Gutierrez, A. C. Habermann, T. Harvey, P. He, X. Hou, L. Hu, Y. Hu, A. Jaiswal, L. Ji, P. Jiang, T. S. Kapellos, C. S. Kuo, L. Larsson, M. A. Leney-Greene, K. Lim, M. Litvinukova, L. S. Ludwig, S. Lukassen, W. Luo, H. Maatz, E. Madissoon, L. Mamanova, K. Manakongtreecheep, S. Leroy, C. H. Mayr, I. M. Mbano, A. M. McAdams, A. N. Nabhan, S. K. Nyquist, L. Penland, O. B. Poirion, S. Poli, C. Qi, R. Queen, D. Reichart, I. Rosas, J. C. Schupp, C. V. Shea, X. Shi, R. Sinha, R. V. Sit, K. Slowikowski, M. Slyper, N. P. Smith, A. Sountoulidis, M. Strunz, T. B. Sullivan, D. Sun, C. Talavera-Lopez, P. Tan, J. Tantivit, K. J. Travaglini, N. R. Tucker, K. A. Vernon, M. H. Wadsworth, J. Waldman, X. Wang, K. Xu, W. Yan, W. Zhao, C. G. K. Ziegler, Nhlbi LungMap Consortium, and Network Human Cell Atlas Lung Biological. 2021. 'Single-cell meta-analysis of SARS-CoV-2 entry genes across tissues and demographics', *Nat Med*, 27: 546-59.

Ng, S. W., A. Mitchell, J. A. Kennedy, W. C. Chen, J. McLeod, N. Ibrahimova, A. Arruda, A. Popescu, V. Gupta, A. D. Schimmer, A. C. Schuh, K. W. Yee, L. Bullinger, T. Herold, D. Gorlich, T. Buchner, W. Hiddemann, W. E. Berdel, B. Wormann, M. Cheok, C. Preudhomme, H. Dombret, K. Metzeler, C. Buske, B. Lowenberg, P. J. Valk, P. W. Zandstra, M. D. Minden, J. E. Dick, and J. C. Wang. 2016. 'A 17-gene stemness score for rapid determination of risk in acute leukaemia', *Nature*, 540: 433-37.

Nusslein-Volhard, C. 2012. 'The zebrafish issue of Development', *Development*, 139: 4099-103.

Ouchetto, W., C. Moulin, and P. G. Roche. 2021. '[Characterization and treatment of SARS-CoV-2 in nasal and bronchial human airway epithelia]', *Med Sci (Paris)*, 37: 1073-76.

Papalexi, E., and R. Satija. 2018. 'Single-cell RNA sequencing to explore immune cell heterogeneity', *Nat Rev Immunol*, 18: 35-45.

Patel, A. P., I. Tirosh, J. J. Trombetta, A. K. Shalek, S. M. Gillespie, H. Wakimoto, D. P. Cahill, B. V. Nahed, W. T. Curry, R. L. Martuza, D. N. Louis, O. Rozenblatt-Rosen, M. L. Suva, A. Regev, and B. E. Bernstein. 2014. 'Single-cell RNA-seq highlights intratumoral heterogeneity in primary glioblastoma', *Science*, 344: 1396-401.

Petersen, P. E. 2008. 'World Health Organization global policy for improvement of oral health--World Health Assembly 2007', *Int Dent J*, 58: 115-21.

Pickar-Oliver, A., and C. A. Gersbach. 2019. 'The next generation of CRISPR-Cas technologies and applications', *Nat Rev Mol Cell Biol*, 20: 490-507.

Porteus, M. H., and D. Baltimore. 2003. 'Chimeric nucleases stimulate gene targeting in human cells', *Science*, 300: 763.

Ravindra, N. G., M. M. Alfajaro, V. Gasque, N. C. Huston, H. Wan, K. Szigeti-Buck, Y. Yasumoto, A. M. Greaney, V. Habet, R. D. Chow, J. S. Chen, J. Wei, R. B. Filler, B. Wang, G. Wang, L. E. Niklason, R. R. Montgomery, S. C. Eisenbarth, S. Chen, A. Williams, A. Iwasaki, T. L. Horvath, E. F. Foxman, R. W. Pierce, A. M. Pyle, D. van Dijk, and C. B. Wilen. 2021. 'Single-cell longitudinal analysis of SARS-CoV-2 infection in human airway epithelium identifies target cells, alterations in gene expression, and cell state changes', *PLoS Biol*, 19: e3001143.

Sachs, N., A. Papaspyropoulos, D. D. Zomer-van Ommen, I. Heo, L. Bottinger, D. Klay, F. Weeber, G. Huelsz-Prince, N. Iakobachvili, G. D. Amatngalim, J. de Ligt, A. van Hoeck, N. Proost, M. C. Viveen, A. Lyubimova, L. Teeven, S. Derakhshan, J. Korving, H. Begthel, J. F. Dekkers, K. Kumawat, E. Ramos, M. F. van Oosterhout, G. J. Offerhaus, D. J. Wiener, E. P. Olimpio, K. K. Dijkstra, E. F. Smit, M. van der Linden, S. Jaksani, M. van de Ven, J. Jonkers, A. C. Rios, E. E. Voest, C. H. van Moorsel, C. K. van der Ent, E. Cuppen, A. van Oudenaarden, F. E. Coenjaerts, L. Meyaard, L. J. Bont, P. J. Peters, S. J. Tans, J. S. van Zon, S. F. Boj, R. G. Vries, J. M. Beekman, and H. Clevers. 2019. 'Long-term expanding human airway organoids for disease modeling', *EMBO J*, 38.

Sasagawa, Y., I. Nikaido, T. Hayashi, H. Danno, K. D. Uno, T. Imai, and H. R. Ueda. 2013. 'Quartz-Seq: a highly reproducible and sensitive single-cell RNA sequencing method, reveals non-genetic gene-expression heterogeneity', *Genome Biol*, 14: R31.

Sato, T., D. E. Stange, M. Ferrante, R. G. Vries, J. H. Van Es, S. Van den Brink, W. J. Van Houdt, A. Pronk, J. Van Gorp, P. D. Siersema, and H. Clevers. 2011. 'Long-term expansion of epithelial organoids from human colon, adenoma, adenocarcinoma, and Barrett's epithelium', *Gastroenterology*, 141: 1762-72.

Sato, T., R. G. Vries, H. J. Snippert, M. van de Wetering, N. Barker, D. E. Stange, J. H. van Es, A. Abo, P. Kujala, P. J. Peters, and H. Clevers. 2009. 'Single Lgr5 stem cells build crypt-villus structures in vitro without a mesenchymal niche', *Nature*, 459: 262-5.

Sauler, M., J. E. McDonough, T. S. Adams, N. Kothapalli, T. Barnthaler, R. B. Werder, J. C. Schupp, J. Nouws, M. J. Robertson, C. Coarfa, T. Yang, M. Chioccioli, N. Omote, C. Cosme, Jr., S. Poli, E. A. Ayaub, S. G. Chu, K. H. Jensen, J. L. Gomez, C. J. Britto, M. S. B. Raredon, L. E. Niklason, A. A. Wilson, P. N. Timshel, N. Kaminski, and I. O. Rosas. 2022. 'Characterization of the COPD alveolar niche using single-cell RNA sequencing', *Nat Commun*, 13: 494.

Shalek, A. K., R. Satija, X. Adiconis, R. S. Gertner, J. T. Gaublomme, R. Raychowdhury, S. Schwartz, N. Yosef, C. Malboeuf, D. Lu, J. J. Trombetta, D. Gennert, A. Gnirke, A. Goren, N. Hacohen, J. Z. Levin, H. Park, and A. Regev. 2013. 'Single-cell transcriptomics reveals bimodality in expression and splicing in immune cells', *Nature*, 498: 236-40.

Shalek, A. K., R. Satija, J. Shuga, J. J. Trombetta, D. Gennert, D. Lu, P. Chen, R. S. Gertner, J. T. Gaublomme, N. Yosef, S. Schwartz, B. Fowler, S. Weaver, J. Wang, X. Wang, R. Ding, R. Raychowdhury, N. Friedman, N. Hacohen, H. Park, A. P. May, and A. Regev. 2014. 'Single-cell RNA-seq reveals dynamic paracrine control of cellular variation', *Nature*, 510: 363-9.

Song, J., I. H. Heijink, L. E. M. Kistemaker, M. Reinders-Luinge, W. Kooistra, J. A. Noordhoek, R. Gosens, C. A. Brandsma, W. Timens, P. S. Hiemstra, M. G. Rots, and M. N. Hylkema. 2017. 'Aberrant DNA methylation and expression of SPDEF and FOXA2 in airway epithelium of patients with COPD', *Clin Epigenetics*, 9: 42.

Sulston, J. E., E. Schierenberg, J. G. White, and J. N. Thomson. 1983. 'The embryonic cell lineage of the nematode Caenorhabditis elegans', *Dev Biol*, 100: 64-119.

Sungnak, W., N. Huang, C. Becavin, M. Berg, R. Queen, M. Litvinukova, C. Talavera-Lopez, H. Maatz, D. Reichart, F. Sampaziotis, K. B. Worlock, M. Yoshida, J. L. Barnes, and H. C. A. Lung Biological Network. 2020. 'SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes', *Nat Med*, 26: 681-87.

Takebe, T., K. Sekine, M. Enomura, H. Koike, M. Kimura, T. Ogaeri, R. R. Zhang, Y. Ueno, Y. W. Zheng, N. Koike, S. Aoyama, Y. Adachi, and H.

Taniguchi. 2013. 'Vascularized and functional human liver from an iPSC-derived organ bud transplant', *Nature*, 499: 481-4.

Tang, F., C. Barbacioru, Y. Wang, E. Nordman, C. Lee, N. Xu, X. Wang, J. Bodeau, B. B. Tuch, A. Siddiqui, K. Lao, and M. A. Surani. 2009. 'mRNA-Seq whole-transcriptome analysis of a single cell', *Nat Methods*, 6: 377-82.

Tata, P. R., H. Mou, A. Pardo-Saganta, R. Zhao, M. Prabhu, B. M. Law, V. Vinarsky, J. L. Cho, S. Breton, A. Sahay, B. D. Medoff, and J. Rajagopal. 2013. 'Dedifferentiation of committed epithelial cells into stem cells in vivo', *Nature*, 503: 218-23.

Thomas, B., A. Rutman, R. A. Hirst, P. Haldar, A. J. Wardlaw, J. Bankart, C. E. Brightling, and C. O'Callaghan. 2010. 'Ciliary dysfunction and ultrastructural abnormalities are features of severe asthma', *J Allergy Clin Immunol*, 126: 722-29 e2.

Thomas, K. R., and M. R. Capecchi. 1987. 'Site-directed mutagenesis by gene targeting in mouse embryo-derived stem cells', *Cell*, 51: 503-12.

Travaglini, K. J., A. N. Nabhan, L. Penland, R. Sinha, A. Gillich, R. V. Sit, S. Chang, S. D. Conley, Y. Mori, J. Seita, G. J. Berry, J. B. Shrager, R. J. Metzger, C. S. Kuo, N. Neff, I. L. Weissman, S. R. Quake, and M. A. Krasnow. 2020. 'A molecular cell atlas of the human lung from single-cell RNA sequencing', *Nature*, 587: 619-25.

Treutlein, B., D. G. Brownfield, A. R. Wu, N. F. Neff, G. L. Mantalas, F. H. Espinoza, T. J. Desai, M. A. Krasnow, and S. R. Quake. 2014. 'Reconstructing lineage hierarchies of the distal lung epithelium using single-cell RNA-seq', *Nature*, 509: 371-5.

V'Kovski, P., A. Kratzel, S. Steiner, H. Stalder, and V. Thiel. 2021. 'Coronavirus biology and replication: implications for SARS-CoV-2', *Nat Rev Microbiol*, 19: 155-70.

Vladar, E. K., J. V. Nayak, C. E. Milla, and J. D. Axelrod. 2016. 'Airway epithelial homeostasis and planar cell polarity signaling depend on multiciliated cell differentiation', *JCI Insight*, 1.

Wiedenheft, B., S. H. Sternberg, and J. A. Doudna. 2012. 'RNA-guided genetic silencing systems in bacteria and archaea', *Nature*, 482: 331-8.

Wiersinga, W. J., A. Rhodes, A. C. Cheng, S. J. Peacock, and H. C. Prescott. 2020. 'Pathophysiology, Transmission, Diagnosis, and Treatment of Coronavirus Disease 2019 (COVID-19): A Review', *JAMA*, 324: 782-93.

Winter, G., and S. Fields. 1980. 'Cloning of influenza cDNA ino M13: the sequence of the RNA segment encoding the A/PR/8/34 matrix protein', *Nucleic Acids Res*, 8: 1965-74.

Wu, F., S. Zhao, B. Yu, Y. M. Chen, W. Wang, Z. G. Song, Y. Hu, Z. W. Tao, J. H. Tian, Y. Y. Pei, M. L. Yuan, Y. L. Zhang, F. H. Dai, Y. Liu, Q. M. Wang, J. J. Zheng, L. Xu, E. C. Holmes, and Y. Z. Zhang. 2020. 'A new coronavirus associated with human respiratory disease in China', *Nature*, 579: 265-69.

Zaki, A. M., S. van Boheemen, T. M. Bestebroer, A. D. Osterhaus, and R. A. Fouchier. 2012. 'Isolation of a novel coronavirus from a man with pneumonia in Saudi Arabia', *N Engl J Med*, 367: 1814-20.

Zeisel, A., A. B. Munoz-Manchado, S. Codeluppi, P. Lonnerberg, G. La Manno, A. Jureus, S. Marques, H. Munguba, L. He, C. Betsholtz, C. Rolny, G. Castelo-Branco, J. Hjerling-Leffler, and S. Linnarsson. 2015. 'Brain structure. Cell types in the mouse cortex and hippocampus revealed by single-cell RNA-seq', *Science*, 347: 1138-42.

Zhang, Y., K. Gilbertson, and W. H. Finlay. 2007. 'In vivo-in vitro comparison of deposition in three mouth-throat models with Qvar and Turbuhaler inhalers', *J Aerosol Med*, 20: 227-35.

Zhu, X., T. K. Wolfgruber, A. Tasato, C. Arisdakessian, D. G. Garmire, and L. X. Garmire. 2017. 'Granatum: a graphical single-cell RNA-Seq analysis pipeline for genomics scientists', *Genome Med*, 9: 108.

Ziegler, C. G. K., V. N. Miao, A. H. Owings, A. W. Navia, Y. Tang, J. D. Bromley, P. Lotfy, M. Sloan, H. Laird, H. B. Williams, M. George, R. S. Drake, T. Christian, A. Parker, C. B. Sindel, M. W. Burger, Y. Pride, M. Hasan, G. E. Abraham, 3rd, M. Senitko, T. O. Robinson, A. K. Shalek, S. C. Glover, B. H. Horwitz, and J. Ordovas-Montanes. 2021. 'Impaired local intrinsic immunity to SARS-CoV-2 infection in severe COVID-19', *Cell*, 184: 4713-33 e22.