Single-cell transcriptomics of lung organoids

Ye Na Song, Man Ryul Lee

Soonchunhyang Institute of Medi-bio Science, Soon Chun Hyang University, Cheonan, Korea

Received: May 16, 2021

Revised: August 23, 2021

Accepted: September 2, 2021

Correspondence to:

Man Ryul Lee, PhD

Soonchunhyang Institute of Medi-bio Science, Soon Chun Hyang University, 25 Bongjeong-ro, Dongnam-gu, Cheonan 31151, Korea

E-mail: leeman@sch.ac.kr

The *in vitro* application of human pluripotent stem cell- or adult stem cell-derived lung organoids has the potential to revolutionize lung disease research, but there are several limitations in the consistent implementation of lung organoids resulting from the structural diversity of the lung tissues and the variety of cell types (more than 40 resident cell types) populating these tissues. However, the evaluation of these complexities using a combination of lung organoids and single-cell transcriptomics has made it possible to identify several key cell types and sub-populations critical to the development of robust *in vitro* organoid models. Recent studies have started to use stem cells to produce these organoids, making it possible to mimic complex 3-dimensional tissues. Furthermore, single-cell mRNA sequencing allows critical comparisons of the transcriptome, which may help focus future research in the field of lung disease.

Keywords: Single cell mRNA sequencing; Organoid; Lung organoid

Introduction

The human body is constantly performing a highly orchestrated process of multiplication and differentiation, taking us from single-cell fertilized eggs to organisms with more than 30 trillion cells in our tissues and organs. Among the various organs that make up the human body, the lung is one of the few to be located on both sides of the body; furthermore, it is found near the heart and in front of the spine. The lungs are mainly responsible for the respiratory function of the organism and are the main site of gas exchange in humans. Lung tissues are made up of more than 40 different resident cell types, including epithelial cells, stromal cells, vascular cells, and immune cells. These cells perform critical functions and contribute to the homeostasis of lung tissues, although more is known about some of these cell types and their functions than others [1]. These cell types are differentially distributed across the various tissues in the lungs, and their density is largely dependent on the part of the lung being evaluated. There is also a distinct difference between these cells and lung stem cells in these tissues [2].

Epithelial cells are present throughout the lung, while mucus-secreting goblet cells, ciliated cells, and club cells, which are differentiated from basal cells and neuroendocrine bodies, are present in the tracheobronchial region. The terminal bronchiolar region has no goblet cells, but contains club cells that can differentiate into ciliated cells and dedifferentiate into basal cells. In the alveolar region, where gas exchange occurs, the blood vessels are highly developed and can differentiate into flat alveolar type 1 (AT1) cells. Other AT1 cells and self-renewed AT2 cells are located next to AT1 cells. The difference between AT1 and AT2 cells were identified using single-cell mRNA sequencing technology, which has led to a more robust analysis of the function of different cell types in the lung, resulting in the development of 3-dimensional (3D) organoids and lung-on-a-chip methods to study lung development, regeneration, and disease [3].

We have been working on methods to simulate the above developmental processes for several years with the aim of producing protocols that can help us understand both normal and pathogenic development *in vitro*. The application of adult stem cells, human

induced pluripotent stem cells (hiPSCs), and pluripotent stem cells has stimulated the development of cell culture-based methods that simulate the development of human lungs and diseases *in vitro*. The advent of 3D culture methods has allowed the expansion of this kind of work. Prior to the development of these technologies, research primarily focused on cell differentiation and development in 2 dimensions (2D) using stem cells. However, 2D models have clear limitations when studying human tissue development, which relies on the intercommunication of several integrated systems. In contrast, 3D culture systems overcome these limitations by facilitating the culture of multiple types of cells in structurally relevant ratios within the same system. Therefore, 3D culture produces a more complex microenvironment and helps to generate a balanced physiological state that more closely mimics *in vivo* tissues, producing what is commonly referred to as an organoid [4]. Stem cell-derived lung organoids are emerging as a new platform for developing novel therapeutics, as they facilitate disease modeling and help to evaluate complex drug interactions in more representative disease models, while also helping to investigate the molecular mechanisms underlying pathogenesis and tissue development [5]. However, it is not clear how accurately lung organoids recreate the lung tissues given the more than 40 resident cell types and their diverse sub-populations present in naturally occurring tissues. That said, several studies have been designed to overcome the limitations of lung organoid systems. These include studies that have analyzed the various characteristics of lung organoids using transcriptome, fluorescence-activated cell sorting (FACS), immunohistochemical analysis, and next-generation sequencing (NGS) [6]. Although studies have shown effectiveness in terms of establishment of the models, these evaluations have not been useful for analyzing the characteristics and diversity of the cells making up specific tissues and their sensitivity to disease. Lung organoids originating from a variety of cells have recently been used to complete single-cell mRNA sequencing analyzes to identify the underlying causes of specific diseases and the molecular mechanisms that underpin organ development to produce more precise and organized lung organoids. However, the lung organoids established so far remain relatively immature compared to *in vivo* tissues, and are often more similar to those found in the fetal environment. This means that these tissues do not reflect the full range of cellular constituents or the cellular ratios found in adult tissues, reducing their overall utility. Thus, it is necessary to establish a cell atlas of current lung organoids and analyze the cellular profiles and ratios using high-resolution single-cell transcriptomics in an effort to identify the missing components needed to model these *in vivo* tissues more accurately.

This review aims to provide a general explanation of the workflow underlying single-cell transcriptomics and discuss the advanc-

es in this field over the last 10 years, while also evaluating the data from those single-cell transcriptome studies evaluating lung organoids. This review was devised to act as a reference point for novel investigations designed to improve lung organoid models.

The single-cell transcriptome analysis workflow

Successful single-cell transcriptome analysis is performed in 5 steps: (1) tissue/cell preparation, (2) single-cell capture, (3) library production, (4) transcriptome sequencing, and (5) data analysis and visualization.

1. Tissue/cell preparation

The first step in the analysis of single-cell transcriptomes is cellular dissociation. This process produces a living monodispersed cell population from cultured cells or tissues. It is important to minimize cell death during the cell dissociation process, as cell death will reduce the variability in the downstream single-cell transcriptome data. Because the analysis of single-cell transcriptomes is both time- and resource-intensive, single-cell analyzes should be subject to various quality control (QC) evaluations at every step. It is relatively easy to obtain living single cells from cultured cells or human blood samples, but obtaining living single cells from mass tissues, such as organoids, requires considerable effort. The platform provided by Miltenyi Biotec (Bergisch Gladbach, Germany) has been shown to be superior in these kinds of applications, as it provides different cell dissociation methods for different tissues, which makes it possible to secure single cells with relatively little damage. After dissociation, it is advisable that researchers evaluate the live:dead cell ratio using DNA dyes, such as PI and 7AAD, or that they increase the ratio of living cells via FACS prior to moving to the next step of the experiment. It is also common for users to evaluate the cell size, cell concentration, and number of living cells using a commercial automated cell counter to identify the number of cells required for the next step in the analysis workflow.

2. Single-cell capture

Efforts to analyze single-cell transcriptomes began in earnest with research into oocytes in 2009 [7]. The subsequent rapid development of NGS analysis technology and the creation of various single-cell isolation methods have resulted in an exponential increase in this type of analysis, which has fostered further development. In the past, it was only possible to separate cells into single cells using the serial dilution method or a labeled protein FACS-based method. However, these methods have limited efficiency because the number of cells that can be obtained per experiment is limit-

ed. Recently, the development and dissemination of commercial microfluidic platforms have enabled high-throughput single-cell transcriptome analysis. The 2 most widely used methods are microwell- and droplet-based methods. Microwell-based single-cell capture relies on introducing individual cells onto a chip one by one using a very fine well, followed by reverse transcription and amplification [8,9]. Unlike other methods, this approach has the advantage of being applicable to a variety of protocols [10–12], and the Fluidigm C1 platform is largely considered to be the representative commercial platform for microwell-based single-cell capture (<https://www.fluidigm.com>). The second method is the microdroplet method, which is widely used in the analysis of single-cell transcriptomes. In this method, users synthesize cDNA libraries by making very small oil droplets that contain a single cell and an oligo-dT labeled gel bead designed to capture each cell and facilitate its lysis within the droplet [13,14]. The largest advantage of this platform is its high capacity; researchers can process thousands of cells concomitantly, and 10x Genomics Chromium is the representative commercial platform for this type of single-cell capture (<https://www.10xgenomics.com>).

3. Producing cDNA libraries

A variety of methods are available for synthesizing single-cell cDNA libraries, but the methods that use oligo-dT primers, which specifically select for mRNA with a poly(A) tail, and subsequent reverse transcription, tend to be the most popular [15]. This method facilitates the addition of a barcode sequence to the transcripts of the library. These barcodes include a unique base sequence consisting of 12 base pairs that can then be used to identify specific transcripts from specific cells during data analysis [16]. When a cDNA library is established, noise is inevitably present; the noise can be mitigated by adding a short unique molecular identifier (UMI) to the primer [17]. These UMIs make it possible to quantify the number of specific transcripts more accurately from each cell by correcting for the noise generated during the amplification process [18]. The quality and concentration of these cDNA libraries are then evaluated using a fragment analyzer to reduce the likelihood of sequencing failure. While these steps help, only about 20% of all transcripts appear in the cDNA library [19], suggesting that there is room for improvement in the mRNA identification process.

4. cDNA sequencing

A small amount of synthesized cDNA can be amplified by polymerase chain reaction and sequenced using a commercial sequencing platform [19]. The system provided by Illumina (San Diego, CA, USA), which is the most representative commercial sequenc-

ing platform, is capable of sequencing single-cell transcriptomes. It uses a sequencing by synthesis technique that allows the massive parallel addition of a single base to every position in the DNA strand. Determining the read depth in single-cell transcriptome analyses may vary depending on the purpose of the experiment, and the read depth is determined based on the number of reads per cell without the number of reads per base. Therefore, in single-cell transcriptome analysis, unbiased sub-populations can be distinguished from 10,000 to 50,000 read depths per cell. While read-depth requirements are dependent on the purpose of the experiment, paired-end sequencing, which increases the reliability of the sequencing results and allows mutation identification, remains the most popular read method. The single-read sequencing method produces limited sequence information because 3' sequencing produces only approximately 100 bp of information per read close to the poly-A tail [20]. This means that most studies use 3'- and 5'-paired-end sequencing in the hope of improving the data produced in these single-cell experiments [21].

5. Data analysis and visualization

The final step in single-cell transcriptome analysis is the visualization and interpretation of the sequencing results. The sequencing results are converted to the FASTQ format for analysis, and the cells are divided into groups using their barcode and UMI information. Each read is then mapped and aligned to a reference genome following QC to improve the reproducibility and reliability of the datasets. This QC process was designed to check the ratio of genes in the non-cellular barcodes and the mitochondrial genome or remove doublet results by evaluating the number of genes per cell [22,23]. It is important to complete a thorough QC evaluation before moving on to data analysis and visualization, as QC improves the quality of the results produced in the study. Data that have undergone QC and normalization can, then, be used to produce new insights into the transcriptome by undergoing dimension reduction, clustering, visualization, and trajectory analysis. Seurat (<https://www.10xgenomics.com>), an open-source software program, is widely used to visualize single-cell transcriptome analysis results. Seurat is an R-based program (<https://satijalab.org/seurat>) that can be used to evaluate and visualize the normalization, dimensionality reduction, and heatmapping of these datasets. Single-cell RNA-sequencing results are considered big data and produce high-level information because thousands to tens of thousands of cells are captured and the expression of 30,000 genes is evaluated in each one. As it is impossible to visualize such high-level information directly, it is essential to visualize groups of cells with similar gene expression profiles. This is accomplished via the process of dimensional reduction, which allows us to project

them into 2D space and express these groups of cells as dots [24]. The most important aspect of dimensional reduction is to avoid the loss of important biological information while excluding noise as much as possible. This process facilitates the evaluation of population heterogeneity, subpopulation identities, and cell trajectories for each group of cells [24].

Single-cell transcriptomes from lung organoids

Lung organoids are frequently established using adult stem cells present in the lungs of human tissues, embryonic stem cells, or dedifferentiated hiPSCs. This review focuses on the application of single-cell transcriptome analysis in the evaluation of lung organoids established using these 3 types of stem cells. Adult stem cells are multipotent and cannot differentiate into cells from other organs, as they are already programmed to differentiate into lung-tissue cells. This means that lung lineage analysis following lung organoid differentiation is not required for these cells. However, only a very small number of cells can be obtained with the consent of the patient, and given the invasive nature of the procedure, it is often difficult to obtain these stem cells from healthy volunteers. Embryonic stem cells are pluripotent cells that can differentiate into any lineage, which means that lung lineage analysis is required after lung organoid differentiation. These evaluations often need to be extensive to clearly demonstrate that they have not become other tissues or organs, but these cells are easier to obtain than adult stem cells. hiPSCs are pluripotent stem cells, such as embryonic stem cells, and display similar disadvantages; however, hiPSCs are more amenable to disease modeling as they are produced following the dedifferentiation of patient-derived somatic cells. This means that each of the 3 different stem cell lineages have different strengths and weaknesses, but they are all more useful for the *in vitro* modeling of human lung diseases than their animal model counterparts.

1. Single-cell RNA sequencing for lung organoids produced using adult stem cells

Many studies have focused on identifying and elucidating the infection pathway for severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) and the subsequent identification of novel therapeutic agents for treating coronavirus disease 2019 (COVID-19). This has led to the use of lung organoids, which overcome the limitations of COVID-19 animal models, into the spotlight [25]. SARS-CoV-2 is replicated and released into the proximal airway [26]. The immune response is active in the distal airway [27]. However, a lung organoid protocol incorporating both cell types has been developed. These lung organoids were then subjected to

single-cell RNA sequencing to determine the significance of their expression in infected patients, thereby narrowing the infection route and cellular target, which are important for the development of novel therapeutic agents. Adult stem cells isolated from lung biopsies produced lung organoids with complete proximal and distal airways *in vitro*, allowing the simulation of SARS-CoV-2 infection without additional lung lineage analysis. Single-cell RNA sequencing analysis of these lung organoids enabled accurate disease modeling and showed that these techniques are suitable for the development of therapeutic agents as they facilitate the identification of specific cell types associated with specific lung-related diseases and can be used to develop novel protocols for producing appropriate organoids and exploring the pathology and physiology of the disease [28].

Unlike in rats, the basal stem cells of the human lung can be expanded and differentiated into the bronchial tree, indicating that lung organoids can be used to study human lung development *in vitro* [29,30]. During lung development, basal stem cells play an important role in maintaining lung homeostasis and facilitating regeneration [31]. This process can be explored using single-cell RNA sequencing and bud tip progenitor organoids [32] made from bud tip progenitor cells (such as adult stem cells) extracted from the human fetus. These evaluations allow researchers to determine how similar the differentiation patterns are for specific cell populations from bud tip progenitor cells *in vitro* and *in vivo*. This was done by using single-cell RNA sequencing to evaluate which cell types from the bud tip progenitor produced a differentiation marker (TP63), and then comparing these to the expression patterns of this marker in bud tip organoids during differentiation. In addition, specific transient cell types present only during lung differentiation, such as the bud tip adjacent and secretory progenitor cell types, can be observed using lung organoids and single-cell RNA sequencing [33].

Next, a single-cell mRNA analysis confirmed the impact of SARS-CoV-2 on human lung alveolar type 2 (hAT2) cells using a 3D culture model. Youk et al. [34] created a 3D culture model with the folded and cystic-like structure of hAT2 cells derived from substantial areas of the lungs from a healthy donor. The analysis was conducted using the 10x Genomics single-cell sequence platform after SARS-CoV-2 infection in the established 3-dimensional culture medium. The analysis captured a total of 14,174 cells with 3,266 detectable genes per cell by an average of 13,500 UMIs, while 6 other clusters characterized hAT2 cells during virus infection. The authors analyzed the single-cell transcription results and confirmed that during SARS-COV-2 infection, no viral infection was found in the cell cluster defined as containing lung stem cells. Genes that respond to endoplasmic reticulum stress,

such as *HSPA1A*, *HSPAS*, and *HERPUD1*, were particularly up-regulated during the acute period of viral infection. As a result of viral infection, hAT2 cells showed significant upregulation of interferon-stimulated genes (ISGs), such as *IFI27* and *IFI6*. In clusters where cells die out due to viral infections, the level of ISG expression mostly decreased. Instead, these cells induced transcription of cell apoptosis mediators, such as *PPP1R15A* and *GADD45B*, suggesting active catastrophic cell pathways due to extreme viral strain. A Gene Ontology analysis using the variable genes obtained through sequencing single-cell transcriptomes primarily highlighted infection-related pathways.

2. Single-cell RNA sequencing for lung organoids using embryonic stem cells

In addition to adult stem cells, which are difficult to extract, embryonic stem cells can be used to produce lung organoids for modeling lung-related diseases. Embryonic stem cells are generally easier to obtain, but lung organoids created using this method need to be evaluated for the lung lineage [35,36]. In particular, for analyses related to SARS-CoV-2, it was important to verify the expression of *ACE2* (the viral receptor), *TMPRSS2* (a protease involved in viral entry) [37], and *FURIN* (a pre-protein convertase that pre-activates the virus) [38], which are essential for SARS-CoV-2 infection. Single-cell RNA sequencing was then able to confirm the cell type in these organoids, indicating that they can be used to model SARS-CoV-2 disease. This system can now be used to rapidly evaluate potential therapeutic interventions, such as imatinib, mycophenolic acid, and quinacrine dihydrochloride (approved by the US Food and Drug Administration [FDA]) for efficacy against COVID-19. In addition, the analysis of lung organoid-modeled SARS-CoV-2 disease using single-cell RNA sequencing made it possible to identify multiple novel colonic cell types [39].

The fact that SARS-CoV-2 recognizes ACE (a viral receptor) and that infection begins in lung cells expressing this receptor was first reported in a study using primary epithelial cell lines [40]. However, Vero cells, which are commonly used for drug screening, have a significantly different ACE expression pattern from patient tissues, making it impossible to screen drugs for SARS-CoV-2 in these cells. Various analytical methods have made it possible to identify FDA-approved drugs that inhibit the process of ACE2 activation by 5 α -dihydrotestosterone during the process of ACE2-mediated SARS-CoV-2 internalization. These drugs were then used to treat human lung organoids composed of AT2 cells expressing *ACE2*, *TMPRSS2*, *AR*, *SRD5A1*, and *SRD5A2* that acted as an *in vitro* SARS-CoV-2 disease model, and researchers were able to evaluate drug efficiency in these settings as a primary screening for clinical evaluation. These researchers, then, used single-cell RNA-sequenc-

ing analysis to identify the factors required for the disease model to produce fully differentiated human lung organoids and to assess whether the percentage of AT2 cells in these organoids was excessive [41].

As primary SARS-CoV-2 infection is respiratory-based, we developed a lung organoid model using human pluripotent stem cells that could be adapted for drug screening. Lung organoids, particularly AT2 cells, express *ACE2* and are permissive to SARS-CoV-2 infection. Transcriptomic analysis following SARS-CoV-2 infection revealed a robust induction of chemokines and cytokines with little type I/III interferon signaling, similar to that observed in human COVID-19 pulmonary infections. We performed high-throughput screening using hPSC-derived lung organoids and identified FDA-approved drug candidates, including imatinib and mycophenolic acid, as inhibitors of SARS-CoV-2 entry. Pre- or post-treatment with these drugs at physiologically relevant levels decreased SARS-CoV-2 infection of hPSC-derived lung organoids. Together, these data demonstrate that hPSC-derived lung cells infected by SARS-CoV-2 can mimic human COVID-19 disease and provide a robust model to screen for FDA-approved drugs that might be repurposed and should be considered for COVID-19 clinical trials [42].

3. Single-cell RNA sequencing for lung organoids produced using dedifferentiated stem cells

Alveologenesis during human lung development is established between 36 weeks of gestation and 3 years of age [3]. This developmental process accompanies the expansion of AT1 cells and their differentiation from AT2 cells, which are essential for gas exchange [43–45]. However, this differentiation process is not well characterized because of the difficulties associated with the isolation and culture of primary AT1 cells. However, producing fibroblast-dependent alveolar organoids (FD-AOs) using dedifferentiated stem cells [46–48] makes it possible to obtain AT1 cells more easily and may facilitate patient-specific cell therapy. Single-cell RNA sequencing analysis showed that induced AT1 cells made from dedifferentiated stem cells and induced AT1 cells obtained from hiPSC-derived AOs and primary AT1 cells were all very similar. In addition, single-cell RNA sequencing analysis revealed that substances such as XAV-939 increase AT1 cell differentiation in FD-AOs via inhibition of canonical Wnt signaling [49].

AT2 cells act as tissue stem cells within the alveolar region of the lung and can differentiate into AT1 cells, which occupy more than 90% of the total area of the lung. AT2 cells can also differentiate into themselves, thereby facilitating their self-renewal. This means that if these AT2 cells can be obtained more stably and efficiently, they may provide a useful tool for evaluating the treatment of ir-

reversible lung diseases. Stable AT2 cell production usually relies on the production and separation of lung AOs, particularly those produced from hiPSCs. These organoids also have a specific advantage for the therapeutic application of AT2 cells, as these cells are less likely to experience immune rejection, making them more efficient for patient-specific disease treatment. Single-cell RNA sequencing analysis following lung AO culture is widely used to develop novel methods to efficiently differentiate and expand hiPSC-derived alveolar stem cells to facilitate long-term evaluation and promote therapeutic efficacy. The efficiency of the differentiation protocol can be evaluated by analyzing lung AOs at each stage, including ventralized anterior foregut endoderm, distal lung progenitor, and AOs using single-cell RNA units during the differentiation process and comparing the expression levels of key markers at each stage [47].

Conclusion

In vitro models of lung disease have recently been thrust into the limelight, as research focus has shifted to the impact of air pollution on humans and lung-targeting diseases such as COVID-19, the prevalence of which has increased. Lung disease modeling relied heavily on both 2D cell culture and animal models until the advent of organoid fabrication techniques; this means that previous work suffered from numerous drawbacks, and most disease models were incomplete. For example, rats, which are the representative animal model for most lung studies, have relatively simple lungs compared to humans, who have over 40 types of cells with varying functions depending on where they are in the lung, making rat models less representative of both disease and development processes, ultimately limiting our ability to understand the full complexity of the lungs (Fig. 1). Therefore, the need for disease modeling using human lung organoids has increased. Recent advances in the production of lung organoids from various stem cells, such as patient-derived adult stem cells, embryonic stem cells, and iPSC dedifferentiated stem cells, have expanded the scope of *in vitro* modeling. However, given that these lung organoids struggle to mature beyond the level of fetal lung tissues, there remains a critical need to advance organoid technologies and better understand the lung tissue and development environments. Therefore, it is essential to verify the cell populations in human adult lungs and lung organoids using various technologies, including single-cell RNA sequencing, and to identify whether mature lung cells can be produced using lung organoids. In addition, various cells interact during gas exchange and at immune barriers and form a complex signaling network, which is prone to dynamic changes that cannot be recapitulated within the organoid

Fig. 1. Lung cell types found in the conducting airways and alveoli. The epithelia of adult lungs can be divided into 4 major, biologically distinct, trophic units (trachea, bronchi, bronchioles, and alveoli), each of which comprises unique types of airway epithelial cells.

environment. This means that it is critical to identify novel technologies that will help define these interactions, including those that would help define the transcriptional profile of these complex cellular environments. Taken together, the data in this review suggest that single-cell transcriptome analysis can help to identify fate determinants during lung development via the production of a pseudo-time trajectory in a reduced data space and help to define cell populations essential to both development and disease in human organs.

Notes

Conflict of interest

No potential conflict of interest relevant to this article was reported.

Funding

This research was funded by The National Research Foundation of Korea, grant number NRF-2020R1A2C1012294 and supported by the Soon Chun Hyang University Research Fund.

Additional contributions

We would like to thank the Soonchunhyang Biomedical Research Core Facility of the Korea Basic Science Institute (KBSI), especially STEMOPIA and CREKA, for the illustration.

ORCID

Ye Na Song, <https://orcid.org/0000-0002-8451-4847>

Man Ryul Lee, <https://orcid.org/0000-0002-2351-9679>

References

- Choi J, Ilich E, Lee JH. Organogenesis of adult lung in a dish: differentiation, disease and therapy. *Dev Biol* 2016;420:278–86.
- Wang J, Li X, Chen H. Organoid models in lung regeneration and cancer. *Cancer Lett* 2020;475:129–35.
- Gkatzis K, Taghizadeh S, Huh D, Stainier DY, Bellusci S. Use of three-dimensional organoids and lung-on-a-chip methods to study lung development, regeneration and disease. *Eur Respir J* 2018;52:1800876.
- Lu T, Cao Y, Zhao P, Shen S, Xi Y. Organoid: a powerful tool to study lung regeneration and disease. *Cell Regen* 2021;10:21.
- Huang SX, Green MD, de Carvalho AT, Mumau M, Chen YW, D'Souza SL, et al. The in vitro generation of lung and airway progenitor cells from human pluripotent stem cells. *Nat Protoc* 2015;10:413–25.
- Goodwin S, McPherson JD, McCombie WR. Coming of age: ten years of next-generation sequencing technologies. *Nat Rev Genet* 2016;17:333–51.
- Tang F, Barbacioru C, Wang Y, Nordman E, Lee C, Xu N, et al. mRNA-Seq whole-transcriptome analysis of a single cell. *Nat Methods* 2009;6:377–82.
- Yuan J, Sims PA. An automated microwell platform for large-scale single cell RNA-Seq. *Sci Rep* 2016;6:33883.
- Zhang X, Li T, Liu F, Chen Y, Yao J, Li Z, et al. Comparative analysis of droplet-based ultra-high-throughput single-cell RNA-Seq systems. *Mol Cell* 2019;73:130–42.
- Der E, Suryawanshi H, Morozov P, Kustagi M, Goilav B, Ranabothu S, et al. Tubular cell and keratinocyte single-cell transcriptomics applied to lupus nephritis reveal type I IFN and fibrosis relevant pathways. *Nat Immunol* 2019;20:915–27.
- Gong H, Do D, Ramakrishnan R. Single-cell mRNA-Seq using the Fluidigm C1 system and integrated fluidics circuits. *Methods Mol Biol* 2018;1783:193–207.
- Nelson L, Tighe A, Golder A, Littler S, Bakker B, Moralli D, et al. A living biobank of ovarian cancer ex vivo models reveals profound mitotic heterogeneity. *Nat Commun* 2020;11:822.
- Matula K, Rivello F, Huck WTS. Single-cell analysis using droplet microfluidics. *Adv Biosyst* 2020;4:e1900188.
- Mazutis L, Gilbert J, Ung WL, Weitz DA, Griffiths AD, Heyman JA. Single-cell analysis and sorting using droplet-based microfluidics. *Nat Protoc* 2013;8:870–91.
- van Dijk EL, Auger H, Jaszczyszyn Y, Thermes C. Ten years of next-generation sequencing technology. *Trends Genet* 2014;30:418–26.
- Tambe A, Pachter L. Barcode identification for single cell genomics. *BMC Bioinformatics* 2019;20:32.
- Chen W, Li Y, Easton J, Finkelstein D, Wu G, Chen X. UMI-count modeling and differential expression analysis for single-cell RNA sequencing. *Genome Biol* 2018;19:70.
- Kolodziejczyk AA, Kim JK, Svensson V, Marioni JC, Teichmann SA. The technology and biology of single-cell RNA sequencing. *Mol Cell* 2015;58:610–20.
- Hwang B, Lee JH, Bang D. Single-cell RNA sequencing technologies and bioinformatics pipelines. *Exp Mol Med* 2018;50:1–14.
- Patrick R, Humphreys DT, Janbandhu V, Oshlack A, Ho JW, Harvey RP, et al. Sierra: discovery of differential transcript usage from polyA-captured single-cell RNA-seq data. *Genome Biol* 2020;21:167.
- Cole C, Byrne A, Beaudin AE, Forsberg EC, Vollmers C. Tn-5Prime, a Tn5 based 5' capture method for single cell RNA-seq. *Nucleic Acids Res* 2018;46:e62.
- Ding B, Zheng L, Zhu Y, Li N, Jia H, Ai R, et al. Normalization and noise reduction for single cell RNA-seq experiments. *Bioinformatics* 2015;31:2225–7.
- Hafemeister C, Satija R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome Biol* 2019;20:296.
- Cakir B, Prete M, Huang N, van Dongen S, Pir P, Kiselev VY. Comparison of visualization tools for single-cell RNAseq data. *NAR Genom Bioinform* 2020;2:lqaa052.
- Salahudeen AA, Choi SS, Rustagi A, Zhu J, van Unen V, de la O SM, et al. Progenitor identification and SARS-CoV-2 infection in human distal lung organoids. *Nature* 2020;588:670–5.
- Hou YJ, Okuda K, Edwards CE, Martinez DR, Asakura T, Dinno KH 3rd, et al. SARS-CoV-2 reverse genetics reveals a variable infection gradient in the respiratory tract. *Cell* 2020;182:429–46.
- Chen Y, Li L. SARS-CoV-2: virus dynamics and host response. *Lancet Infect Dis* 2020;20:515–6.
- Tindle C, Fuller M, Fonseca A, Taheri S, Ibeawuchi SR, Beutler N, et al. Adult stem cell-derived complete lung organoid models emulate lung disease in COVID-19. *Elife* 2021;10:e66417.
- Hong KU, Reynolds SD, Watkins S, Fuchs E, Stripp BR. Basal cells are a multipotent progenitor capable of renewing the bronchial epithelium. *Am J Pathol* 2004;164:577–88.
- Rock JR, Onaitis MW, Rawlins EL, Lu Y, Clark CP, Xue Y, et al. Basal cells as stem cells of the mouse trachea and human airway epithelium. *Proc Natl Acad Sci U S A* 2009;106:12771–5.
- Murata K, Ota S, Niki T, Goto A, Li CP, Ruriko UM, et al. p63 - Key molecule in the early phase of epithelial abnormality in idiopathic pulmonary fibrosis. *Exp Mol Pathol* 2007;83:367–76.

32. Miller AJ, Hill DR, Nagy MS, Aoki Y, Dye BR, Chin AM, et al. In vitro induction and in vivo engraftment of lung bud tip progenitor cells derived from human pluripotent stem cells. *Stem Cell Reports* 2018;10:101–19.
33. Miller AJ, Yu Q, Czerwinski M, Tsai YH, Conway RF, Wu A, et al. In vitro and in vivo development of the human airway at single-cell resolution. *Dev Cell* 2020;53:117–28.
34. Youk J, Kim T, Evans KV, Jeong YI, Hur Y, Hong SP, et al. Three-dimensional human alveolar stem cell culture models reveal infection response to SARS-CoV-2. *Cell Stem Cell* 2020;27:905–19.
35. Duan F, Guo L, Yang L, Han Y, Thakur A, Nilsson-Payant BE, et al. Modeling COVID-19 with human pluripotent stem cell-derived cells reveals synergistic effects of anti-inflammatory macrophages with ACE2 inhibition against SARS-CoV-2. *Res Sq* [Preprint]. Posted 2020 Aug 20:rs.3.rs-62758. <https://doi.org/10.21203/rs.3.rs-62758/v1>.
36. Mulay A, Konda B, Garcia G Jr, Yao C, Beil S, Villalba JM, et al. SARS-CoV-2 infection of primary human lung epithelium for COVID-19 modeling and drug discovery. *Cell Rep* 2021;35:109055.
37. Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, et al. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* 2020;181:271–80.
38. Shang J, Wan Y, Luo C, Ye G, Geng Q, Auerbach A, et al. Cell entry mechanisms of SARS-CoV-2. *Proc Natl Acad Sci U S A* 2020;117:11727–34.
39. Han Y, Duan X, Yang L, Nilsson-Payant BE, Wang P, Duan F, et al. Identification of SARS-CoV-2 inhibitors using lung and colonic organoids. *Nature* 2021;589:270–5.
40. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. *Nature* 2020;581:215–20.
41. Samuel RM, Majd H, Richter MN, Ghazizadeh Z, Zekavat SM, Navickas A, et al. Androgen signaling regulates SARS-CoV-2 receptor levels and is associated with severe COVID-19 symptoms in men. *Cell Stem Cell* 2020;27:876–89.e12.
42. Han Y, Yang L, Duan X, Duan F, Nilsson-Payant BE, Yaron TM, et al. Identification of candidate COVID-19 therapeutics using hPSC-derived lung organoids. *bioRxiv* [Preprint]. Posted 2020 May 5. *bioRxiv* 2020.05.05.079095. <https://doi.org/10.1101/2020.05.05.079095>.
43. Borok Z, Danto SI, Lubman RL, Cao Y, Williams MC, Crandall ED. Modulation of α 1 expression with alveolar epithelial cell phenotype in vitro. *Am J Physiol* 1998;275:L155–64.
44. Barkauskas CE, Cronce MJ, Rackley CR, Bowie EJ, Keene DR, Stripp BR, et al. Type 2 alveolar cells are stem cells in adult lung. *J Clin Invest* 2013;123:3025–36.
45. Zacharias WJ, Frank DB, Zepp JA, Morley MP, Alkhaleel FA, Kong J, et al. Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. *Nature* 2018;555:251–5.
46. Jacob A, Morley M, Hawkins F, McCauley KB, Jean JC, Heins H, et al. Differentiation of human pluripotent stem cells into functional lung alveolar epithelial cells. *Cell Stem Cell* 2017;21:472–88.
47. Yamamoto Y, Gotoh S, Korogi Y, Seki M, Konishi S, Ikeo S, et al. Long-term expansion of alveolar stem cells derived from human iPS cells in organoids. *Nat Methods* 2017;14:1097–106.
48. Gotoh S, Ito I, Nagasaki T, Yamamoto Y, Konishi S, Korogi Y, et al. Generation of alveolar epithelial spheroids via isolated progenitor cells from human pluripotent stem cells. *Stem Cell Reports* 2014;3:394–403.
49. Kanagaki S, Ikeo S, Suezawa T, Yamamoto Y, Seki M, Hirai T, et al. Directed induction of alveolar type I cells derived from pluripotent stem cells via Wnt signaling inhibition. *Stem Cells* 2021;39:156–69.