

**Title: Three-dimensional human alveolar stem cell culture models reveal infection response to SARS-CoV-2**

Jeonghwan Youk<sup>1,2,14</sup>, Taewoo Kim<sup>1,14</sup>, Kelly V. Evans<sup>3,4,14</sup>, Young-Il Jeong<sup>5,14</sup>, Yongsuk Hur<sup>6,14</sup>, Seon Pyo Hong<sup>7,14</sup>, Je Hyoung Kim<sup>5</sup>, Kijong Yi<sup>1</sup>, Su Yeon Kim<sup>1</sup>, Kwon Joong Na<sup>8</sup>, Thomas Bleazard<sup>9</sup>, Ho Min Kim<sup>1,10</sup>, Mick Fellows<sup>11</sup>, Krishnaa T. Mahubani<sup>12</sup>, Kourosh Saeb-Parsy<sup>12</sup>, Seon Young Kim<sup>13</sup>, Young Tae Kim<sup>8,#</sup>, Gou Young Koh<sup>7,#</sup>, Byeong-Sun Choi<sup>5,#</sup>, Young Seok Ju<sup>1,2,15,#</sup>, and Joo-Hyeon Lee<sup>3,4,#</sup>

**Affiliations**

1 Graduate School of Medical Science and Engineering, Korea Advanced Institute of Science and Technology, Daejeon 34141, Republic of Korea

2 GENOME INSIGHT Inc., Daejeon 34051, Republic of Korea

3 Wellcome-MRC Cambridge Stem Cell Institute, Jeffrey Cheah Biomedical Centre, University of Cambridge, Cambridge, CB2 A0W, United Kingdom

4 Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, CB2 3EL, United Kingdom

5 Division of Viral Disease Research, Center for Infectious Diseases Research, Korea National Institute of Health, Korea Centers for Disease Control and Prevention, Cheongju 28159, Republic of Korea

6 BioMedical Research Center, Korea Advanced institute of Science and Technology, Daejeon 34141, Republic of Korea

7 Center for Vascular Research, Institute for Basic Science, Daejeon 34126, Republic of Korea

8 Department of Thoracic and Cardiovascular Surgery, Seoul National University Hospital, Seoul 03080, Republic of Korea

9 The National Institute for Biological Standards and Control, Blanche Lane, South Mimms, Potters Bar, Hertfordshire EN6 3QG, United Kingdom

10 Center for Biomolecular and Cellular Structure, Institute for Basic Science, Daejeon 34126, Republic of Korea

11 Clinical Pharmacology and Safety Sciences, R&D, AstraZeneca, Cambridge, UK

12 Department of Surgery and Cambridge NIHR Biomedical Research Centre, Biomedical Campus, University of Cambridge, Cambridge, CB2 2QQ, United Kingdom

13 Department of Laboratory Medicine, Chungnam National University College of Medicine, Daejeon 35015, Republic of Korea

14 These authors contributed equally

15 Lead Contact

**#Correspondence to**

Young Seok Ju, ysju@kaist.ac.kr, +82-42-350-4237, Associate Professor, KAIST, Daejeon 34141, Republic of Korea

Joo-Hyeon Lee, jhl62@cam.ac.uk, Group Leader, Wellcome - MRC Cambridge Stem Cell Institute, University of Cambridge, Cambridge CB2 A0W, United Kingdom

Byeong-Sun Choi, byeongsun@korea.kr, Division head, Division of Viral Disease Research, Center for Infectious Disease Research, Korea Centers for Disease Control & Prevention, Cheongju, Republic of Korea

Gou Young Koh, gykoh@kaist.ac.kr, Director, Center for Vascular Research, Institute of Basic Science, Daejeon 34126, Republic of Korea

Young Tae Kim, ytkim@snu.ac.kr, Professor, Department of Thoracic Surgery, Seoul National University Hospital, Seoul 03080, Republic of Korea

## **SUMMARY**

Severe acute respiratory syndrome-coronavirus 2 (SARS-CoV-2), which is the cause of a present global pandemic, infects human lung alveolar type 2 (hAT2) cells. Characterizing pathogenesis is crucial for developing vaccines and therapeutics. However, the lack of models mirroring the cellular physiology and pathology of hAT2 cells limits the study. Here, we develop a feeder-free, long-term three-dimensional (3D) culture technique for hAT2 cells derived from primary human lung tissue, and investigate infection response to SARS-CoV-2. By imaging-based analysis and single-cell transcriptome profiling, we reveal rapid viral replication and the increased expression of interferon-associated genes and pro-inflammatory genes in infected hAT2 cells, indicating robust endogenous innate immune response. Further tracing of viral mutations acquired during transmission identifies full infection of individual cells effectively from a single viral entry. Our study provides deep insights into the pathogenesis of SARS-CoV-2, and the application of defined 3D hAT2 cultures as models for respiratory diseases.

### **Key words**

Human lung alveolar type 2 cells; SARS-CoV-2; COVID-19; Virus; 3D cultures; Transcriptome; Pathogenesis; Single-cell RNA-seq; Interferon

### **Running Title**

Three-dimensional human alveolar culture models reveal infection response to SARS-CoV-2

## Introduction

Currently, Coronavirus disease 2019 (COVID-19), caused by severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), is spreading globally (Wu et al., 2020) and more than 35.8 million confirmed cases with 1.05M deaths have been reported worldwide as of 7<sup>th</sup> Oct 2020. The alveoli in lung tissues are the main target tissues for these emerging viruses especially in SARS-CoV-2-associated pneumonia patients (Muus et al., 2020).

To develop strategies for efficient prevention, diagnosis, and treatment, the characteristics of new viruses, including mechanisms of cell entry and transmission, kinetics in replication and transcription, host reactions and genome evolution, should be accurately understood ideally in its target cell types. Essential features of SARS-CoV-2 virus, for instance ACE2 dependency for cellular entry, patterns of viral transcription and structure of the Spike protein, have been identified (Hoffmann et al., 2020; Kim et al., 2020a; Walls et al., 2020; Wang et al., 2020). However, most findings have been obtained from experiments using non-physiological cell lines (Chu et al., 2020), model animals, such as transgenic mice expressing human angiotensin-converting enzyme 2 (Bao et al., 2020), ferrets (Kim et al., 2020c) and golden hamsters (Sia et al., 2020), or from simple observation in clinical cohorts (Wolfel et al., 2020) and/or inference from in-silico computational methods (Andersen et al., 2020; Forster et al., 2020; Shang et al., 2020). A number of studies utilizing stem cell-based models have been recently established for various tissues (Huang et al., 2020; Jacob et al., 2020; Lamers et al., 2020; Ramani et al., 2020; Yang et al., 2020). However, without competent human alveolar model systems derived from primary tissues, controlled experiments for understanding virus-host interactions, subsequent immune reactions, or for detecting personal genome variants causing susceptibility to the viral infection are challenging. Studies of COVID-19, and respiratory infectious diseases more generally, have been limited by the lack of physiological models that recapitulate normal alveolar physiology and pathology.

Development of organotypic mini-organ models, or organoids, has enabled various physiologic and pathological studies using human-derived tissues *in vitro* (Bartfeld et al., 2015; Fatehullah et al., 2016; Heo et al., 2018). Organoid models established from the human kidney, intestine, and airway have been used to investigate SARS-CoV-2 viral pathogenesis (Elbadawi and Efferth, 2020; Lamers et al., 2020; Monteil et al., 2020).

However, the cellular response of human alveolar type 2 (hereafter referred to as hAT2) cells to SARS-CoV-2 remains elusive, due to difficulty in the long-term expansion of pure hAT2 cells. A recent study utilized a model of hAT2 cells derived from human induced pluripotent stem cells to show the aspects of SARS-CoV-2 infection (Huang et al., 2020). However, inability to differentiate into alveolar type 1 (AT1) cells and to assess potential age- and/or disease-related viral effects limits an understanding of infection response in primary alveolar lung tissues. In this study, we develop a technique for long-term, feeder free human 3D alveolar type 2 cell cultures (hereafter referred to as h3ACs) established from single primary hAT2 cells which serve as stem cells in adult alveolar tissues (Barkauskas et al., 2013). Using our h3AC models, we demonstrate phenotypic changes of hAT2 cells induced by SARS-CoV-2 infection by multi-dimensional methods.

## Results

### Establishing the three-dimensional cultures of human alveolar type 2 cells with chemically defined conditions

We developed chemically defined culture conditions for growing hAT2 cells, which were significantly improved from previous feeder-based systems (Barkauskas et al., 2013; Kathiriya et al., 2020). It allowed the self-organization of hAT2 single-cells into alveolar-like 3D structures with defined factors that support the molecular and functional identity of primary tissue-derived hAT2 cells over multiple passages. Briefly, single-cell dissociated hAT2 cells derived from distal parenchymal regions of healthy donor lungs were isolated by fluorescence-activated cell sorting (FACS) for the hAT2 cell surface marker HTII-280 (CD31<sup>-</sup>CD45<sup>-</sup>EpCAM<sup>+</sup>HTII-280<sup>+</sup>) (**Figure 1A and Figure S1A**) (Barkauskas et al., 2013; Gonzalez et al., 2010). Approximately, HTII-280<sup>+</sup> cells represented 75% of total EPCAM<sup>+</sup> cells (**Figure S1A**). Quantitative PCR (qPCR) analysis confirmed that the hAT2 cell marker *SFTPC* was highly expressed in isolated HTII-280<sup>+</sup> cells, while the basal cell marker *TP63* and secretory cell marker *SCGB1A1* were highly expressed in HTII-280<sup>-</sup> cells (**Figure S1B**). Isolated HTII-280<sup>+</sup> cells were then embedded in Matrigel for 3D culture supplemented with CHIR99021, RSPO1 (R-spondin 1), FGF7, FGF10, EGF, NOG (NOGGIN), and SB431542, factors that have been implicated in lung development and growth of human embryonic lung tip cells (Nikolic et al., 2017). Under this condition, hAT2 single-cells created 3D cellular structures with heterogeneous size and morphology, including both folded and cystic-like structures (**Figures 1B-D**). They consisted of mature hAT2 cells expressing pro-SFTPC, HTII-280, and

ABCA3, as well as exhibiting uptake of LysoTracker, a fluorescent dye that stains acidic organelles such as lamellar bodies, a secretory vesicle containing surfactant proteins (**Figures 1B-D; Figure S1C**) (Barkauskas et al., 2013; Van der Velden et al., 2013). Further, HTII-280<sup>+</sup> hAT2 cells expressing HOPX were also observed in our h3ACs, as well as in alveolar tissue sections (Travaglini et al., 2020) (**Figures S1C and S1D**). Notably, hAT2 cells started to form a small cluster revealing the enriched expressions of HTII-280 and F-actin in the luminal side of h3ACs (**Figure 1E**). Over the growth of h3ACs, although the location of HTII-280 expression became varied (**Figure S1E**), the distribution of F-actin was maintained specifically in the luminal and lateral surfaces of hAT2 cells, consistent with a previous observation in distal embryonic lung epithelium (El-Hashash and Warburton, 2011) (**Figure 1F**). In contrast, SCRIB, which establishes apical-basal polarity (Rodriguez-Boulan and Macara, 2014), was restricted to the basolateral surface of hAT2 cells (**Figure 1F**). We found that CRB3, which has been previously reported to mark an epithelial apical domain including proximal airway epithelium in lung development (Pocha and Knust, 2013; Szymaniak et al., 2015), was expressed broadly within hAT2 cells (**Figures 1E and 1F**).

We found WNT activation as an essential factor for clonal expansion of hAT2 cells, as evidenced by no colony formation in the absence of the WNT activator CHIR99021 in primary culture (**Figure S1F**). HTII-280<sup>-</sup> cells were also cultured under conditions supporting 3D cultures of human bronchial (airway) cells (hereafter referred to as h3BCs) that have previously been reported (Sachs et al., 2019). They grew quickly by day 14 with cystic-like structures consisting of a number of airway cell types, including KRT5<sup>+</sup>TP63<sup>+</sup> basal cells and SCGB1A1<sup>+</sup> secretory cells, as previously reported (**Figures S1G-I**) (Sachs et al., 2019). Significantly, our culture condition allowed the long-term expansion of hAT2 cells up to 10 months, although colony forming efficiency varied between individual donor samples (**Figures 1G and 1H**). Importantly, over serial passaging via single cell dissociation, hAT2s revealed clonal expansion, leading to 3D structures consisting of mature hAT2 cells expressing pro-SFTPC and uptake of LysoTracker following 6 months of continuous culture (**Figure 1I**). In addition, cultured hAT2 cells for 6 months maintained normal karyotypes (**Figure S1J**) suggesting chromosome-level genomic stability for the period. Importantly, upon culture in 2D which promotes AT1 cell phenotype (Dobbs, 1990), clonally expanded hAT2 cells differentiated into hAT1 cells expressing AGER and AQP5, indicating retention of AT1 cell differentiation capacity of hAT2 cells in our culture condition (**Figure 1J**). Over long-term culture, we found that colony forming efficiency reduced at later passages, with colony

size decreasing (**Figure 1H and Figure S1K**). In 8-month-old 3D cultures, some cells lost the expression of pro-SFTPC, although no expression of airway markers such as TP63 was detected (**Figure 1K**).

### **SARS-CoV-2 infection in h3ACs**

We next infected our established h3ACs (and h3BCs) with SARS-CoV-2 (**Figure 2A**). To this end, we physically and chemically broke h3ACs and h3BCs into pieces to enhance the access of SARS-CoV-2 to the apical cell surfaces, followed by viral incubation at a multiplicity of infection (MOI) of 1.0 and 0.1. The viral particles were collected and prepared from a Korean patient (known as KCDC03) who was diagnosed with COVID-19 on January 26, 2020, after traveling to Wuhan, China (Kim et al., 2020b). Vero cells, an interferon-deficient 2D cell line conventionally used as viral host, were also experimentally infected as a positive control, although this was not directly comparable to our 3D infection models due to different technical procedures.

In the plaque forming assay with MOI of 1.0, infectious viral particles increased to significant titers in h3ACs within the 1st day post infection (dpi), suggesting that exponential viral replication occurs during <24 hrs from viral entry in hAT2 cells (**Figures 2B and 2C**). The increment of viral particles was observed in h3BCs as reported (Suzuki et al., 2020), but their titers were ~100 times lower than h3ACs (**Figures 2B and 2C**). Infected Vero cells exhibited an even higher SARS-CoV-2 viral burden than h3ACs (**Figures 2B and 2C**). In line with viral particles, the viral RNA levels in the infected h3ACs also highly increased at 1 dpi (**Figure 2D**). Further, viral transcripts were also detected in the supernatant of h3ACs (**Figure 2D**), suggesting active secretion of viral transcripts (or particles) from the infected hAT2 cells. In the lower infection experiments with MOI of 0.1, a substantially lower viral burden was observed for the three days, suggesting that a fraction of hAT2 cells were likely left uninfected by 3 dpi (**Figure 2D**). Notably, infected h3ACs did not show substantial cytopathies, such as cell degeneration and syncytium formation (Chu et al., 2020), and maintained cell viability, unlike Vero cells (**Figure 2E and Figure S2**).

### **Visualization of SARS-CoV-2 infection in h3ACs**

Immunofluorescent (IF) staining of the infected h3ACs revealed widespread expression of ACE2 and TMPRSS2 proteins, which are necessary for SARS-CoV-2 infection (**Figures 3A and 3B**)(Hoffmann et al., 2020). Despite conventional wisdom, these proteins were not

present exclusively in membrane but also in the cytoplasm as reported (Hamming et al., 2004; Manna and Caradonna, 2020; Zhou et al., 2020). In the infected h3ACs (MOI of 1.0 at 1 dpi), viral components were clearly observed by IF staining of nucleocapsid protein (NP), and double-stranded viral RNA (dsRNA) of SARS-CoV-2. These two viral components were robustly detected in the cytoplasm of hAT2 cells with expression of pro-SFTPC, ACE2 and TMPRSS2 (**Figures 3C and 3D; Figures S3A and S3B**). IF staining of infected h3ACs revealed that approximately 94% of h3ACs (n=61 out of 65 h3ACs) harbored cells with detectable levels of the viral components (**Figure S3C**). At the level of individual cells, approximately 61% of the cells within individual h3ACs showed evidence of the viral components (**Figure S3C**). However, more widespread infection may have been present, as cells with minimal viral components may have fallen below our detection threshold (**Discussion**). When h3ACs were treated by camostat mesylate, a serine protease inhibitor active against TMPRSS2, SARS-CoV-2 entry and progression was partially blocked, corroborating that the activity of TMPRSS2 is required for viral infection of hAT2 cells (**Figure S3D**).

To further determine subcellular consequences of the SARS-CoV-2 infection at a higher resolution, we performed a transmission electron microscopic (**TEM**) analysis. Compared with uninfected h3ACs (**Figure S4**), infected h3ACs at 2 dpi with MOI of 1.0 showed viral particles and aggregated proteins which appeared as electron-dense objects (Knoops et al., 2008) in ~51% of cells within four h3ACs explored (**Figures 4A-J; Discussion**). Infected hAT2 cells exhibited pathogenic subcellular alterations, such as massive vacuoles as seen in Zika virus-infected epithelial cells (**Figures 4A and 4F**)(Arbour et al., 1999; Monel et al., 2017). Viral particles were either dispersed or encapsulated in a vesicular structure in cytoplasm (**Figures 4C, 4D, and 4G**). Double-membrane vesicles (**DMVs**), subcellular structures known as sites for early viral replication, were rarely observed in the vicinity of zippered endoplasmic reticulum in an infected cell (**Figure 4I**)(Maier et al., 2013; Ogando et al., 2020). Virus secretion was also observed in the luminal spaces of h3ACs (**Figures 4E and 4J**).

Intriguingly, a fraction of cells in h3ACs (~19% of cells within the four h3ACs explored) showed much higher viral burdens than other cells (**Figure 4C**), with as many as 500 copies in the 100 nm section, implying >10,000 SARS-CoV-2 particles in the cell under an assumption of uniform intracellular viral distribution. It remains to be studied whether these

viral particles are to be secreted, or are competent for infection of other cells. However, if 10% of them are actively released, the burst size of SARS-CoV-2 from hAT2 cells would be larger than previously suggested for murine hepatitis virus, a viral species in beta coronavirus (Bar-On et al., 2020).

### **Transcriptome changes of h3ACs after SARS-CoV-2 infection**

We next explored the gene expression changes after SARS-CoV-2 infection in h3ACs. To this end, strand-specific Illumina mRNA libraries were constructed from infected h3ACs with MOI of 1.0 and sequenced (overall 7 samples; at 0, 1, and 3 dpi with 2, 3, and 2 biological replicates, respectively). For comparison, the transcriptome of infected Vero cells was also sequenced.

A set of human genes were differentially expressed among the three time points (**Figure 5A**), although most genes showed good correlations in gene expression levels (**Table S1, Figure S5A**). Cytokeratin genes (including KRT16, KRT6A, KRT6B, and KRT6C), genes involved in keratinization (including SPRR1A), cytoskeleton (including S100A2) and cell-cell adhesion genes (including DSG3), were significantly reduced to ~2-3% in h3ACs at 3 dpi compared to the levels at 0 dpi (**Figures 5A and 5B**). Many more genes were upregulated in the infected h3ACs specifically at 3 dpi. In particular, transcription of a broad range of interferon-stimulated genes (ISGs), known to be typically activated by type I and III interferons (Iwasaki et al., 2017), was remarkably increased. These genes include interferon induced protein genes (such as *IFI6*, *IFI27*, *IFI44*, *IFI44L*), interferon induced transmembrane protein genes (such as *IFITM1*), interferon induced transmembrane proteins with tetratricopeptide repeats genes (*IFIT1*, *IFIT2*, *IFIT3*), 2'-5'-oligoadenylate synthetase genes (*OAS1*, *OAS2*), and miscellaneous genes known to be involved in innate cellular immunity (*MX1*, *MX2*, *RSAD2*, *ISG15*). These genes were expressed to >20 times higher levels in h3ACs at 3 dpi than at 0 dpi. Many other known ISGs also showed moderate inductions (2-20 times), including *BTS2*, *OAS3*, *HERC5*, *HERC6* and *USP18*. Antiviral functions are known for these ISGs (Schneider et al., 2014) including (1) inhibition of virus entry (*MX* genes, *IFITM* genes), (2) inhibition of viral replication and translation (*IFIT* genes, *OAS* genes, *ISG15*, *HERC5*, *HERC6*, *USP18*) and (3) inhibition of viral egress (*RSAD2* and *BST2*). Of the 20 interferon genes in the human genome, an interferon beta (*IFNB1*) and three interferon lambda genes (*IFNL1*, *IFNL2*, and *IFNL3*) showed significant transcriptional induction, although their absolute changes were not substantial (**Figure 5C**). The surface

receptors of interferons were stably expressed in h3AC cells without reference to viral infection (**Figure 5C**). Downstream signaling genes of the receptors were appreciably upregulated (2-7 times), for example, *STAT1*, *STAT2* and their associated genes, such as *IRF1* and *IRF9*. Of note, *IRF1* is known to be specific to type I interferon responses (Forero et al., 2019), while type I and type III ISGs are generally overlapping (Park and Iwasaki, 2020).

In addition to ISGs, genes in the viral sensing pathway in cytosol showed increased expression in the infected h3ACs at 3 dpi, for example, *DDX58* (official gene name of RIG-1, ~13 times), *IFIH1* (also known as MDA5, ~5 times), *TLR3* (Toll-like receptor 3, ~3 times), *IRF7* (Interferon regulatory factor 7, ~7 times), and *IL6* (~4 times at 1 dpi; proinflammatory factor). Of note, the innate immune response observed was completely autologous because immune cells are absent in our h3AC system, thus mimicking the very early phase of SARS-CoV-2 alveolar infection. As a validation study, we performed the analyses of qPCR and IF staining in the infected h3ACs (**Figure 5D and Figure S5B**), which supported the transcriptome changes.

To understand the biological reliability of these upregulated genes, we compared the gene list with transcriptomes produced from the lung tissue of COVID-19 patients (obtained from 8 tissues from two COVID-19 pneumonia patients)(Blanco-Melo et al., 2020). Unlike our models, tissue transcriptome showed a substantially higher level of variance presumably caused from the diverse cell-type composition in the tissue (i.e. airway epithelial, immune, mesenchymal and endothelial cells; **Figure S5C**). Of the 97 DEGs in our h3AC models, 26.8% were confirmed in the infected human tissues (**Figure S5D; Table S2**).

We then further explored the utility of our h3AC models by comparison with other cellular models. In h3BCs, for example, most of the DEGs observed in h3ACs were not substantially altered (**Figures S5E; Table S3**). Indeed, only a few genes were transcriptionally altered in h3BCs at 3 dpi (**Figure S5F**), presumably due to the insufficient SARS-CoV-2 infection as mentioned above (**Figure 2C**). Similarly, 2D cell-lines (Emanuel et al., 2020b) established from lung and colon cancers (NCI-H1299 and Caco-2, respectively) showed no evidence of sufficient SARS-CoV-2 infection (**Figure S5G**). Vero cells (sequenced in this study) showed expression changes of many genes, but transcription of ISGs was not altered presumably due to the absence of interferon genes in the genome (**Figure S5G**). Calu-3, established

from lung cancer, showed quite similar gene expression changes with our h3ACs (**Figure S5G**). Of these models, our h3ACs showed the highest concordance with human lung tissues (**Figure S5H**). In addition, a prolonged observation (>1 dpi) of viral infection is not feasible with these cell-lines, because infected 2D cells are easily suspended from the culture plate.

### Expression of viral genes

We further analyzed the viral RNAs from the transcriptomes of infected h3AC and h3BC models. In agreement with the plaque assay (**Figures 2B and 2C**), the proportion of viral reads in whole-transcriptome sequences plateaued by 1 dpi in the h3ACs (**Figure 5E**). Approximately 50% of the sequencing reads were mappable to the SARS-CoV-2 genome in h3ACs from 1 dpi, indicating prevailing viral gene expression in infected AT2 cells as observed in Vero cells (Kim et al., 2020a). In contrast, the proportion of viral reads was less than 0.3% in the infected h3BCs until 3 dpi (**Figure 5E**).

Viral transcripts were not mapped uniformly to the viral genome sequence, but 3' genomic regions where canonical subgenomic RNAs (sgRNAs) showed much higher read-depth in all samples, consistent with the previous report (**Figure S5I**)(Kim et al., 2020a). The vast majority of viral RNA sequences were in the orientation of positive-sense RNA strands (**Figure S5I**; for example, 99.98% vs 0.02% for positive- and negative-sense RNAs, respectively, from h3ACs at 1 dpi). This is in good agreement with the nature of SARS-CoV-2, which is an enveloped, non-segmented, and positive-sense RNA virus. Therefore, its genomic RNA in viral particles and sgRNA transcripts being translated in the cytoplasm should be positive-sense RNAs. Negative-sense RNAs are intermediate templates for replication and/or transcription of positive-sense RNAs.

Cross-comparison of viral RNA sequences obtained from the 11 infected h3ACs (n=5) and h3BCs (n=6) revealed 20 viral base substitutions (**Table S4**). No mutation was at 100% variant allele fraction (VAF) and exclusive to an infection experiment. Instead, sequence alterations showed a broad range of quasispecies heterogeneity in each culture (VAF ranges from 0.1% to 73.1%; **Figure 5F**), and a large proportion of the mutations (n=16; 80%) were shared by two or more infected models (by the cut-off threshold of 0.1%). Therefore, we speculate that most of the base changes were originally present in the pool of viral particles before their inoculation to our models. Given the fact that these viral particles were prepared

from one of the earliest COVID-19 patients, our finding suggests that mutations can accumulate in the pool of viral genomes in a small number of rounds of viral transmissions, and appear with dramatic changes in quasispecies abundance. A substantially higher proportion of specific mutations in a sample may suggest a bottleneck in viral entry or stochasticity in viral replication.

### **Transcriptome changes at a single-cell resolution**

To understand transcriptional changes of the infected h3ACs at a single-cell resolution, we employed a 10X Genomics single-cell RNA-seq platform for h3ACs at five different conditions (uninfected, MOI of 0.1 at 4, 16 and 60 hpi (hour post infection), and MOI of 1.0 at 60 hpi). From the five conditions, we overall captured a total of 14,174 single-cells, with 3,266 detectable genes per cell by 13,500 unique molecular identifiers (UMIs), on average. Using UMIs from the host transcripts, we conducted integration (count normalization, feature selection, dimension reduction, and visualization through UMAP) (**Figure 6A**) and unsupervised clustering of the single-cells (**Figure 6B** and **Figure S6A**) through Seurat and SC3 packages (Kiselev et al., 2019; Stuart et al., 2019). With the exception of Cluster 7 (from Seurat), showing airway-like cells in h3ACs (778 cells from the 5 conditions; **Figure S6B; Discussion**), the other six clusters exhibited hAT2 cell features in the course of the viral infection. The Seurat clusters consisted of: Cluster 1, mostly with uninfected cells (3,146 cells); Clusters 2-4, mostly with cells from lower infection experiments at different time points (2,048, 1,932, and 2,092 cells for Cluster 2,3, and 4, respectively); Cluster 5, mostly with cells with moderate to high levels of viral transcripts (2,111 cells); and Cluster 6, cells disintegrating and likely close to cell death (2,067 cells). Generally, cells in MOI of 1.0 and/or in the later stage (60 hpi) showed a higher proportion of single-cells harboring viral UMIs (**Figure 6C**). In particular, 99.9% of cells in h3ACs with MOI of 1.0 at 60 hpi showed viral UMIs, confirming our previous imaging analyses that a substantial number of cells in the infected h3ACs were infected at 1-3 dpi (**Figures 3C, 3D, and 4; Figures S6C and S6D; Discussion**). The number of viral transcripts, however, was not uniformly distributed, but enriched in a fraction of cells in each condition (**Figure 6C**). For example, infected cells in Cluster 6 exhibited on average 10.7 times more viral UMI counts than cells in Cluster 5 (1,280 vs. 119 UMIs, respectively), despite cells in Cluster 6 containing relatively lower total UMI counts than ones in the other clusters (4,534 vs. 17,923 UMIs, respectively; **Figures 6D and 6E**). When normalized with cellular UMI counts, the proportion of highly infected cells ( $\geq 2^9$  viral reads per cell) in Cluster 6 was 28 times higher than that in Cluster 5 (**Figure 6D**).

Interestingly, the highly infected cells in Cluster 6 showed reduced expression of canonical hAT2 marker genes, including *SFTPB* (Surfactant Protein B) and *NKX2-1* (NK2 homeobox 1) (**Figure 6E**). Further, we found 8.2% of actively proliferating hAT2 cells (*MKI67*<sup>+</sup> h3ACs) in uninfected h3ACs (Cluster 1), which was more-or-less maintained across the infection progression in infected h3ACs (Cluster 2, 10.6%; Cluster 3, 12.8%; Cluster 4, 3.6%; Cluster 5, 30.3%; Cluster 6, 5.1%) (**Figure S6F**).

Of note, we found 52 cells (0.4%) showing *ACE2* transcripts, 2,661 cells (18.8%) expressing *TMPRSS2* transcripts, and 17 cells (0.1%) expressing both in single-cells (**Figure 6E**), despite the widespread protein expression (**Figures 3A and 3B**). These proportions are low at face value, but are consistent with a previous observation (Ziegler et al., 2020). Although the previous report also suggested that *ACE2* RNA expression can be stimulated as an infection-mediated response, particularly in human airway cells, such a trend was not observed in our dataset.

Our single-cell clusters illustrated the characteristic transcriptional changes in the course of SARS-CoV-2 infection (**Figures 7A and 7B; Table S5**). For example, *AQP5* transcription, highly expressed in a subset of hAT2 cells that selectively express Wnt pathway genes and were proposed as alveolar stem cells (Travaglini et al., 2020), was robust in uninfected h3AC cells (Cluster 1), yet downregulated in the other clusters of infected cells. In the acute phase of viral infection (4 hpi; Cluster 2), genes responsive to the ER stress, such as *HSPA1A*, *HSPA5* and *HERPUD1*, were specifically upregulated suggesting that viral infection affects cellular phenotypes from the early hours after cellular entry. In Clusters 4 and 5, hAT2 cells at 60 hpi upregulated ISGs, such as *IFI27* and *IFI6*. In the cluster of disintegrating cells (Cluster 6), expression levels of ISGs were mostly reduced. Instead, these cells showed transcriptional induction of apoptosis mediators, such as *PPP1R15A* and *GADD45B*, suggesting an active catastrophic cellular pathway due to the extreme viral burdens. Gene ontology analyses with the variable genes in single-cell transcriptome sequencing mostly highlighted infection-related pathways (**Figure S6E**).

#### **Effective number of viral entry for infection of an alveolar cell**

Finally, we statistically inferred the number of viral particles effectively entering each alveolar cell for infection. Although thousands of viral copies were often observed in the infected cells (**Figure 4**), it is generally not known how many viral particles are necessary for effective

infection of an alveolar cell. In an extreme scenario, one viral particle is sufficient. Alternatively, infection may be initiated with the entry of multiple viruses. We tracked the effective viral number of cellular entries using a mutation (NC\_045512.2: 23,707C>U; a silent mutation in Spike protein) as a viral barcode in the viral quasispecies (**Figure 7C**). From our sequencing, the mutation was estimated to be present at 4.3% VAF in the initial viral pool for inoculation. If the first scenario dominantly applies, the infected alveolar cells will be dichotomized, i.e., 95.7% cells with solely wild-type virus and the rest of the cells with solely mutant virus, exclusively. Alternatively, if multiple viruses are effectively entering a cell, a large proportion of alveolar cells harboring detectable mutant virus will have intermediate VAF for the mutation. In our single-cell transcriptomes of h3ACs with MOI of 1 infection, 212 cells had at least two independent transcripts for the mutant locus. The majority of the cells harboring the mutant allele did not have the wild-type allele present (n=8 vs 4 for cells showing solely mutant allele vs mixture of the quasispecies, respectively; **Figure 7C**). In a more sophisticated statistical analysis, infection by single viral entry is estimated as >2 times more frequent than by multiple viral entry (69% vs. 31%, respectively; **Figure 7D**). Our calculation indicates that a single viral particle is mainly responsible for SARS-CoV-2 infection in most hAT2 cells, although multiple viral entry is also possible. It may also reflect viral interference in SARS-CoV-2 alveolar infection.

## Discussion

In this study, we have established the experimental condition for feeder-free, long-term 3D cultures of adult hAT2 cells. We demonstrated that WNT activity is crucial for hAT2 maintenance, as consistent with previous studies of murine AT2 cells (Frank et al., 2016; Nabhan et al., 2018; Zacharias et al., 2018). Although our h3ACs may not produce the full cellular spectrum present in the adult alveoli, h3ACs established cellular polarity and maintained functional mature and genomic stable hAT2 cells over multiple passages.

SARS-CoV-2 infected h3ACs showed remarkable cellular and transcriptional changes far more clearly than other models, including h3BCs and 2D Vero cell lines, showing cellular tropism in the viral replication and transcription as well as the resultant reaction from the host cell.

SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 (Hoffmann et al., 2020). Although their protein levels are widespread in hAT2 cells (Hamming et al., 2004; Ren et al., 2006), their transcription is not very active in our single-cell transcriptome analysis of h3ACs, in agreement with recent reports (Zhao et al., 2020; Ziegler et al., 2020). Estimating the proportion of cells harboring low abundance transcripts with single-cell RNAseq is generally challenging, frequently resulting in underestimation.

On the contrary, we observed that single-cell transcriptome sequencing enables more sensitive and scalable detection of the infected cells compared to immunostaining and TEM analyses (**Figure S6D**). We believe that newly infected cells, which include abundant viral transcripts prior to the production of viral protein components, are more likely to be detected by single-cell transcriptome techniques. However, the detection sensitivity is dependent on the burden of viral transcripts in cells as well as the depth of single-cell sequencing.

Through the multi-dimensional analyses of the infected h3ACs, we found that (1) viral replication is active during the first day after cellular infection, (2) expression of type I and III interferons begins at 1 dpi, and (3) expression of ISGs follows at 3 dpi. The timing of ISG induction in alveolar cells may be more rapid *in vivo* in concert with exogenous interferons from immune cells. For more physiological understanding, co-culturing SARS-CoV-2 infected h3AC models with immune cells obtained from the same donor would be helpful.

Notably, we identified five states of infected cells across the infection progression with MOI of 0.1 and 1.0 (Clusters 2-6). Given the hAT2 cells are more-or-less homogeneous before being infected, this indicates a distinct switch between cell states occurring during infection. If *in vivo* SARS-CoV-2 infection transforms hAT2 cells into Cluster 6-like cells showing the loss of hAT2 identity, the integrity and function of alveoli will be highly compromised. The route towards the quantum cellular change is of significant interest to be answered in future studies.

Although we observe in this study that physiological and molecular changes of alveolar cells appear within 3 days post infection, clinical symptoms of COVID-19 sometimes occur after >10 days of viral exposure (Lauer et al., 2020). Several possible explanations could underpin the temporal difference. Viral infiltration from the upper respiratory tract to terminal alveoli may take several days; a substantial proportion of alveolar cells may need to be

infected before displaying symptoms; or, further interactions with immune cells resulting in inflammation may be necessary for developing symptoms.

From single-cell transcriptome sequences, we observed a subset of cells with distinct transcriptional features (Cluster 7), which express lower levels of canonical hAT2 marker genes, including *SFTPB* (**Figure 6E**) but detectable levels of airway marker genes, including *SOX2*, *KRT5*, and *TP63* (**Figure S6B**). Some of these cells harbored SARS-CoV-2 transcripts, but expression of the marker genes was not highly affected. We speculate that these cells are hAT2 cells with pathologic phenotypes of alveolar bronchiolization and/or basal-like hAT2 cells, which are stochastically observed in chronic lung diseases (Jensen-Taubman et al., 1998; Xu et al., 2016). Given the fact that h3ACs used for SARS-CoV-2 infection were established from hAT2 cells isolated from adjacent normal counterparts of lung cancer and/or IPF, it is likely that this transcriptional feature reflects the cellular status of original tissues rather than virus-associated phenotype.

In summary, our study highlights the advantages of h3AC models to elucidate the pathogenesis of SARS-CoV-2 infection in alveolar stem cells. Our data will be a great resource for the biomedical community for deeper characterization of the viral disease. We believe that our models will enable more accurate and sophisticated analyses in the near future, especially for studying the response to viral infection within vulnerable groups with aged or diseased lungs, providing the opportunity to elucidate individual patient responses to viral infection. Furthermore, our models can be combined with other techniques, such as co-culture experiments with immune cells and *in vitro* screening of antiviral agents. We believe that our models are also applicable for the study of basic biology and other diseases in hAT2 cells.

### **Limitations of Study**

Our h3AC models were established from lung tissue specimen collected from patients with heterogeneous clinical history and genetic background. These confounding factors have not been thoroughly controlled in the study. Due to the lack of immune cells in our system, molecular interaction between hAT2 and immune cells could not be investigated in this study. With h3AC models from additional donors as well as introducing immune-cells in culture system will likely advance our understanding of the pathogenesis of SARS-CoV-2 infection.

## **Acknowledgements**

We thank Jinwook Choi (Wellcome - MRC Cambridge Stem Cell Institute), Sam M Janes, Rob Hynds, Sarah Clarke (University College London), Yong Man Han, Eui-Cheol Shin, Heung Kyu Lee, Su-Hyung Park, Jeong Seok Lee, Myungsuk Choi (KAIST), Jung-Ki Yoon (Seoul National University Hospital), Yoon Ho Kim (Seoul National University Cancer Research Institute), and Jong-Yeon Shin (Macrogen Inc.) for valuable discussion, productive comments, and technical help; Irina Pshenichnaya (Histology), Peter Humphreys (Imaging), Simon McCallum (Flow cytometry, Cambridge NIHR BRC Cell Phenotyping Hub), and Cambridge Stem Cell Institute core facilities for technical assistance. This work was supported by Suh Kyungbae Foundation (SUHF-18010082 to Y.S.J.); National Research Foundation of Korea (for Brain Pool Program NRF-2019H1D3A2A02061168 to Y.S.J and S.Y.K; Leading Researcher Program NRF-2020R1A3B2078973 to Y.S.J); Research of Korea Centers for Disease Control and Prevention (2019-NI069-00 to Y.-I.J.); Institute for Basic Science (IBS-R025-D1, G.Y.K) funded by Ministry of Science and ICT of Korea; Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (HI14C1277 to Y.S.J. and Y.T.K.); Research Program 2020 funded by Seoul National University College of Medicine Research Foundation (800-20200516); and European Research Council Starting Grant (679411). J.-H.L is supported by Wellcome and the Royal Society (107633/Z/15/Z). K.V.E is supported by Biotechnology and Biological Sciences Research Council Industrial CASE (BBSRC iCASE) studentships (BB/R505328/1). Y.S.J. is supported by Human Frontier Science Program Research Grant (RGY0071/2018).

## **Author contributions**

Conceptualization, Y.S.J. and J.-H.L.; hAT2 3D culture optimization, K.V.E., J.-H.L.; 3D culture establishments, K.V.E., J.Y., T.K., J.-H.L.; Specimen preparation; K.J.N, M.F., K.T.M., K.S.-P., Y.T.K.; Infection, J.K., J.Y., T.K., Y.-I.J., B.-S.C; Confocal microscope, K.V.E., S.P.H., G.Y.K.; Electron microscope, Y.H., T.K., H.M.K.; Karyotyping, Seon.Y.K; Transcriptome analysis, J.Y., T.K., K.Y., Su.Y.K., Y.S.J.; Statistics, Su.Y.K.; Writing, J.Y., T.K., K.V.E., T.B., S.-H.K., G.Y.K., Y.S.J., J.-H.L.; Manuscript finalization, all authors; Supervision Y.S.J. and J.-H.L.

## **Declaration of interests**

Y.S.J. is a founder and CEO of GENOME INSIGHT Inc..

## Figure Legends

### Figure 1. Long-term, three-dimensional cultures of human alveolar type 2 cells in chemically-defined conditions

(A) Schematic diagram outlining method for h3ACs.

(B) A representative image of primary hAT2-derived 3D structures (h3ACs) from freshly isolated HTII-280<sup>+</sup> cells at day 27 in culture (top) and with LysoTracker (bottom; red). P0; passage 0. Scale bar, 2000  $\mu\text{m}$  (top), 1000  $\mu\text{m}$  (bottom).

(C) Morphological heterotypic colony formation from isolated hAT2 cells in primary h3ACs (P0). Hematoxylin and eosin (H&E, left), and IF staining (right) for HTII-280 (green), pro-SFTPC (red) and DAPI (blue). Scale bar, 50  $\mu\text{m}$ .

(D) Quantification of the folded and cystic 3D structures in primary h3ACs (P0). Data presented are the mean  $\pm$  SEM for 3 individual donor samples (n = 67 for donor 1, n = 50 for donor 2, n = 50 for donor 3). n=total number of colonies scored.

(E) IF images of primary h3ACs expressing HTII-280 (green), F-actin (white), CRB3 (red), and DAPI (blue). Scale bar, 50  $\mu\text{m}$ .

(F) IF images of primary h3ACs expressing HTII-280 (green), F-actin (white), CRB3 (red), SCRIB (red), and DAPI (blue). Scale bar, 50  $\mu\text{m}$ .

(G) Serial passage of h3ACs via single cell dissociation at various time points depending on growth from 3 individual donors. Each point represents a single passage for n = 3 individual donor samples with >3 technical replicates.

(H) Quantification of colony forming efficiency for h3ACs at day 14 of culture up to 9 total passages. Colony forming efficiency: number of colonies formed/number of cells plated per well as a percentage. Data are presented as mean  $\pm$  SEM for 3 individual donor samples (n = 3 biological samples). Each point represents the average of 3 technical replicates calculated for each biological sample (except P8, which is n = 2).

(I) Flow cytometry analysis of LysoTracker (deep red) uptake in freshly isolated HTII-280<sup>+</sup> cells and dissociated h3ACs at passage 6 (P6, 6 months in culture).

(J) IF images of hAT1 cells, induced by 2D-plating of cultured h3ACs at passage 2. Pro-SFTPC (hAT2, white), AGER (hAT1, red), AQP5 (hAT1, green), and DAPI (blue). Scale bar, 50  $\mu\text{m}$ .

(K) Quantification of h3ACs expressing pro-SFTPC and TP63 over multiple passages (passage 0, P0; passage 1, P1; passage 8, P8). h3ACs were classified based on IF staining for pro-SFTPC (hAT2), TP63 (basal), or lack of both markers (pro-SFTPC<sup>-</sup> TP63<sup>-</sup>). The mean

percentage of total colonies per well represented by each class of colony is shown. Data is presented as mean  $\pm$  SEM at each passage (n = 131 for P0, n = 55 for P1, and n = 25 for P8). Three donor samples were used, except P8 (one donor sample was used). n=total number of colonies scored across passages.

### **Figure 2. SARS-CoV-2 infectivity in human 3D alveolar and bronchial cultures**

**(A)** Schematic diagram outlining the method for SARS-CoV-2 infection in h3ACs and human 3D bronchial cultures (h3BCs). To infect these models, h3ACs and h3BCs were collected and broken to expose the luminal space. Then, the pieces of h3ACs and h3BCs were incubated with SARS-CoV-2 at MOI of 0.1 or 1.0 for 2 hours.

**(B)** Representative images for plaque assay using SARS-CoV-2 infected cells at 3 dpi. Dilution factors are shown in numbers on the right upper corner. Scale bar, 1 cm.

**(C)** Plaque assay showing that SARS-CoV-2 actively replicates in h3ACs at 1dpi. Mean  $\pm$  SEM (n = 2, two plaque assays at each time point). h3ACs at passage 2-3 from 2 donors and h3BCs at passage 2 from 2 donors were used in the plaque assays.

**(D)** Quantitative PCR (qPCR) analysis measuring the viral N gene transcripts of SARS-CoV-2 in lysed h3ACs with MOI of 0.1 and 1.0 and in supernatant of h3ACs with MOI of 1.0. Mean  $\pm$  SEM (n = 3, three qPCR assays at each time point). h3ACs at passage 1-3 from 3 donors and h3BCs at passage 2 from 2 donors were used in the experiment.

**(E)** The viability of Vero cells remarkably decreases at 2 days after SARS-CoV-2 infection, whereas that of h3ACs does not significantly change. Mean  $\pm$  SEM (n = 2, three measurements of luminescence at each time point). h3ACs at passage 0 from 1 donor were used in the experiment.

### **Figure 3. Confocal imaging of SARS-CoV-2 infected h3ACs**

**(A)** A representative image of IF staining of ACE2 (green) and HTII-280 (red) in h3ACs. All replicates (n = 5 at passage 1-2 from 3 donors) of control h3ACs express ACE2 and HTII-280. Scale bar, 50  $\mu$ m.

**(B)** A representative image of IF staining of TMPRSS2 (green) and pro-SFTPC (red) in h3ACs. All replicates (n = 5 at passage 1-2 from 3 donors) of control h3ACs express TMPRSS2 and pro-SFTPC. Scale bar, 50  $\mu$ m.

**(C)** Representative images of an infected h3AC at 1 dpi. Viral components (NP or dsRNA; green) are co-stained with pro-SFTPC (red). At 1 dpi, SARS-CoV-2 highly infects h3ACs

which show a punctated pattern of pro-SFTPC. N ≥ 5 replicates of infected h3ACs; P2-3 from 3 donors. Scale bar, 50 μm.

**(D)** Representative images of an infected h3AC expressing ACE2 (red). SARS-CoV-2 are identified by viral dsRNA (green) or NP (green). Viral dsRNA appears punctated. N ≥ 5 replicates of infected h3ACs; P2-3 from 3 donors. Scale bar, 50 μm.

#### **Figure 4. Transmission electron microscopic imaging analysis of SARS-CoV-2 infected h3ACs at 2 days post infection**

**(A)** A low magnification representative image of an infected h3ACs (n=10 h3ACs from 3 donors; P2-3). Red asterisks: alveolar space; white arrowheads: aggregated viral particles; Vc: large pathologic vacuoles; white dashed line: hAT2 cell membrane; Nu: nucleus.

**(B)** Lamellar bodies (skyblue arrow) are observed.

**(C)** An hAT2 cells with a high density of SARS-CoV-2 particles. Viral particles are dispersed in a cytoplasm of the cell (red arrow).

**(D)** Multiple viral particles included in the vesicular structures (red arrow).

**(E)** Virus particles secreted to the lumen of the h3AC.

**(F)** A partial image of another infected h3AC.

**(G)** Virus particles encapsulated in vesicular structures (red arrow).

**(H)** Virus-containing vesicles aggregated in the vicinity of ER (yellow arrow).

**(I)** Double membrane vesicles (red arrowhead) located near zippered ER (yellow arrowhead).

**(J)** Viral particles secreted into the lumen of a h3AC (red arrows). Microvilli (black arrow) are shown at the apical side of a hAT2 cell. All Scale bars, 1 μm.

#### **Figure 5. RNA-sequencing analyses of infected h3ACs and h3BCs**

**(A)** Heatmap of the most variable 100 genes among three groups of h3ACs at 0, 1 and 3 dpi.

**(B)** A volcano plot showing differentially expressed genes between h3ACs at 0 dpi and 3 dpi.

**(C)** Transcriptional changes of interferon genes in the infected h3ACs.

**(D)** IF imaging for upregulated MX1 (green). The intensity of MX1 significantly increases in infected h3ACs (*p-value* < 0.001). n = 11 for control and n = 13 for infected h3ACs.

**(E)** Proportion of viral RNA reads in the h3ACs and h3BCs transcriptomes.

**(F)** An example of missense mutation (NC\_045512.2: 3,177C>U) detected from a h3BC transcriptome at 3 dpi.

**Figure 6. Single-cell transcriptome analysis of uninfected and infected h3ACs**

- (A) Five experimental conditions of the dataset in the UMAP plot. MOI = multiplicity of infection. hpi = hours post infection.
- (B) Unsupervised UMAP clustering of uninfected and infected h3ACs.
- (C) Normalized levels of SARS-CoV-2 viral UMI counts in each experimental condition.
- (D) Normalized levels of SARS-CoV-2 viral UMI counts in the UMAP plot. Highly infected cells (viral UMI counts per 10k human UMI counts  $\geq 2^9$ ) are mostly enriched in Cluster 6.
- (E) Expression of genes showing general features of h3ACs in the dataset.

**Figure 7. Gene markers for each of the cluster and estimation of the number of viruses entering a host cell**

- (A) Heatmap showing highly expressed variable genes in each cluster. The most significantly variable genes ( $\leq 10$ ) in each cluster are demonstrated. Level\_CoV2:  $\log_2$  level of SARS-CoV-2 UMI counts per 10k human UMI counts. Normalized read counts ( $\log_2$ ): UMI counts of a certain gene per 10k human UMI counts.
- (B) Expression of genes which are represent in Uninfected hAT2s cluster (*AQP5*), Acutely infected hAT2s cluster (*HSPA1A*), LowIFN\_lowCoV2 and LowIFN\_mod.CoV2 clusters (*IFI27*), and Dying hAT2s\_highCoV2 (*PPP1R15A*)
- (C) Distribution of single base substitution (variant allele fraction (VAF) = 4.3% in original viral particles) in each single cell.
- (D) Maximum likelihood estimation for the proportion of cells infected with a single virus.

## **STAR Methods**

### **KEY RESOURCES TABLE**

#### **RESOURCE AVAILABILITY**

Lead Contact  
Materials Availability  
Data and Code Availability

#### **EXPERIMENT MODEL AND SUBJECT DETAILS**

Human Tissues  
Virus particle preparation for infection

#### **METHODS DETAILS**

Human lung tissue dissociation and flow cytometry  
In vitro 3D culture and passage  
Karyotype of *in vitro* 6 month grown h3AC  
Assessment of lung lineage transcripts by qRT-PCR  
Differentiation of hAT2 cells into hAT1 cells in 2D culture  
SARS-CoV-2 infection of 3D culture cells and Vero cell  
Viral RNA extraction  
Cellular total RNA extraction  
Viral RNA copy number calculation with qRT-PCR  
Validation of upregulated genes with qRT-PCR  
Live virus titer calculation with plaque assay  
Viability and cytotoxicity test  
Immunofluorescence staining of paraffin-embedded h3AC and LysoTracker  
Immunofluorescence staining of infected h3AC with cryo-section  
Immunofluorescence staining of whole-mount h3ACs  
Immunofluorescence staining of hAT1-like cell  
Transmission electron microscopy  
Bulk RNA sequencing and data processing  
Single cell RNA sequencing (scRNAseq) and data processing  
Statistical inference on the effective number of viral entry

#### **QUANTIFICATION AND STATISTICAL ANALYSIS**

## **RESOURCE AVAILABILITY**

### **Lead Contact**

Further information and requests for resources and reagents should be directed to and will be fulfilled by the Lead Contact, Young Seok Ju (ysju@kaist.ac.kr).

### **Materials Availability**

All 3D models generated in this study are available from the Lead Contact with a completed Materials Transfer Agreement.

### **Data and Code Availability**

Processed single cell RNA sequencing data including Seurat objects, cellular metadata, and counts/UMI tables are available on Synapse (accession syn22146555). Scripts reproducing the single cell RNA sequencing analysis are deposited on GitHub ([https://github.com/ju-lab/SARS-CoV-2\\_alveolar\\_organoids](https://github.com/ju-lab/SARS-CoV-2_alveolar_organoids)). Bulk RNA and single cell RNA sequencing datasets are uploaded on the European Genome-Phenome Archive (EGA) with accession ID EGAS00001004508 for human-derived data and the NCBI Gene Expression Omnibus (GEO) GSE159316 for Vero cell data. Transmission electron microscopy images (n>300) are uploaded in EM Public Image Archive (EMPIAR) with an accession ID EMPIAR-10533. Supplemental information is uploaded in Mendeley Data with a DOI 10.17632/hzbrzdkcfr.1.

## **EXPERIMENT MODEL AND SUBJECT DETAILS**

### **Human Tissues**

For the establishment of human lung 3D-culture models, human distal lung parenchymal tissues from deidentified lungs not required for transplantation were obtained from adult donors with no background lung pathologies from Papworth Hospital Research Tissue Bank (T02233), and Addenbrookes Hospital (Cambridge University NHS foundations trust) under the collaboration of Cambridge Biorepository for Translational Medicine (CBTM) project. Appropriate Human Tissue Act (HTA) guidance was followed; For the viral infection and following analysis, human lung tissues were obtained from patients undergoing lobectomy surgery at Seoul National University Hospital (SNUH) with written informed consent from approval of the ethical committee (approval no. C-1809-137-975). hAT2 3D cultures for infection studies were established from normal counterpart tissues in lung cancer patients or an idiopathic pulmonary fibrosis (IPF) patient. Overall, h3ACs at passage 0-3 from a total of 11 donors and h3BCs at passage 2 from 3 donors were used in our infection experiments (**Table S6**).

### **Virus particle preparation for infection**

SARS-CoV-2 virus strain is BetaCov/Korea/KCDC03/2020 (Kim et al., 2020b). The patient (KCDC03) was diagnosed with COVID-19 on January 26, 2020, after traveling to Wuhan, China. The virus was also sequenced in the previous paper (Kim et al., 2020a). For virus replication, Vero cells were infected with MOI of 0.01 and grown under DMEM with 2% FBS, 1% P/S for 48 hours at 37°C 5% CO<sub>2</sub> (hereafter, the media is termed as infection media). Media was centrifuged with 2500rpm for 25min, and supernatant without cell debris was stocked at -80°C with 4 x 10<sup>6</sup> pfu / ml.

## **METHOD DETAILS**

### **Human lung tissue dissociation and flow cytometry**

Distal lung parenchymal tissue was processed as soon as possible in order to minimize cell yield loss and maintain cell viability. Briefly, fresh tissue was washed in cold PBS and minced into small (1 mm) pieces with a scalpel, followed by further dissociation using pre-warmed digestion buffer containing 2U/ mL Dispase II (Sigma-Aldrich, CORNING), 1 mg/mL Dispase/Collagenase (Sigma-Aldrich) and 0.1 mg/mL DNase I (Sigma-Aldrich) in PBS at 37 °C for 1 hr with agitation. Tissue cell suspensions were filtered through a 100 µm cell strainer into a 50 mL falcon tube to remove cell debris, and washed with 10 mL of DMEM (Gibco, Thermo Fisher Scientific). Cells were centrifuged at 350 g for 10 min, supernatant carefully aspirated, and cell pellet resuspended in 5 mL red blood cell lysis buffer for 5 min at room temperature (RT). The reaction was quenched using 5 mL of DMEM, and the entire 10 mL of cell suspension was transferred to a 15 mL falcon tube, followed by 10 min centrifugation at 350 g. Supernatant was removed, and the cell pellet was resuspended in 10% FBS in PBS (PF10 buffer) for counting. Cells were prepared for flow cytometry with primary antibodies CD31-APC (Biolegend, 303116), CD45-APC (Biolegend, 368512), EpCAM-FITC (Biolegend, 324204) and HTII-280-IgM (Terrace Biotech, TB-27AHT2-280) at 1:40 per 4 million cells for 30 min on ice. Following two washes with cold PF10 buffer and centrifugation at 350 g for 5 min, cells were stained with secondary PE goat anti-mouse IgM (eBioscience, 12-5790-81) for HTII-280 at 1:100. Stained cells were washed with PF10 buffer, and counted using a hemocytometer to assess dilution required for final volume. Cells were diluted at a concentration of 30 million cells/ mL and filtered through a 35 µm cell strainer into polypropylene FACS tubes. Cell sorting was performed on an Aria III fusion (UK; BD Biosciences) or an Aria II (South Korea; BD Biosciences) using a 100 µm nozzle (UK) or 85 µm nozzle, and data were analyzed with FlowJo software (UK; Tree Star, Inc.) or FACS DIVA (South Korea, BD Biosciences).

### ***In vitro* 3D culture and passage**

Freshly isolated HTII-280<sup>+</sup> and HTII-280<sup>-</sup> cells derived from CD31<sup>-</sup>CD45<sup>-</sup>EpCAM<sup>+</sup> cells of human distal lungs were resuspended in base medium (Advanced DMEM/F12 (Thermo Fisher Scientific) supplemented with 10 mM HEPES (Gibco), 1 U/mL Penicillin/Streptomycin (Gibco), 1 mM *N*-Acetylcysteine (Sigma-Aldrich), and 10 mM Nicotinamide (Sigma-Aldrich)). Growth factor-reduced (GFR) Matrigel (CORNING) was added to the cell suspension at a ratio of 1:1, and 100 µL of suspension was added to a 24-well transwell insert with a 0.4 µm pore

(CORNING) so that there were approximately  $10 \times 10^4$  cells per insert. GFR-Matrigel was allowed to solidify for 1 hr at 37 °C, after which 500  $\mu$ L of pre-warmed alveolar media (base medium supplemented with 1 x B27 (Thermo Fisher Scientific), 10% R-SPONDIN-1 (Cambridge Stem Cell Institute tissue culture core facility, manually produced), 50 ng/ml human EGF (PEPROTECH), 100 ng/ml human FGF7/KGF (PEPROTECH), 100 ng/ml human FGF10 (PEPROTECH), 100 ng/ml NOGGIN (PEPROTECH), 10  $\mu$ M SB431542 (TOCRIS) and 3  $\mu$ M CHIR99021 (TOCRIS)) was added to each lower chamber. For assessment of the effect of Wnt activity on hAT2 culture ability, primary cultures were also established without CHIR99021. Cultures were maintained under standard cell culture conditions (37 °C, 5% CO<sub>2</sub>), with media changes every 2-3 days. Y-27632 (10  $\mu$ M, Sigma-Aldrich) was added for the first 48 hr of culture to promote cell survival. To avoid the growth of fungal and bacterial infection, 250 ng/mL Amphotericin B and 50  $\mu$ g/mL gentamicin were added to culture medium for 5 days. For culture in 48-well plates,  $5 \times 10^3$  cells were resuspended in 100% GFR-Matrigel, and allowed to solidify in a 20  $\mu$ L droplet per well at 37 °C for 20 min, followed by submersion in 250  $\mu$ L of pre-warmed medium. Cellular structures with a surface area of more than 500  $\mu$ m at day 14 of culture were counted at day 14 to assess colony forming efficiency, and were either fixed and stained at day 21 for analysis, or enzymatically dissociated into single cells for further culture without sorting. Cells in 3D cultures were passaged at different days depending on size, with culture days varying from 21-35 days. For passaging, Matrigel was disrupted by incubation with dispase (Sigma-Aldrich) at 37 °C for 45 min, followed by single cell-dissociation through addition of TrypLE (Gibco) for 5 min at 37 °C. The reaction was quenched with base medium, and cells were centrifuged at 350 x g for 5 min. Cells were resuspended in fresh GFR-Matrigel at a ratio of  $5 \times 10^3$  (48-well plates) or  $10 \times 10^4$  (24-well transwell inserts) cells as before. Bronchial cells were passaged every 21-28 days due to accelerated growth compared with hAT2 cells, and were cultured in previously reported medium conditions (Sachs et al., 2019) with the following concentration/factor edits; 100 ng/ml human FGF10, 10% R-SPONDIN-1, 10  $\mu$ M SB431542 (instead of A83-01).

### **Karyotype of *in vitro* 6 months grown h3ACs**

Cultured h3ACs for 6 months were retrieved from Matrigel using Cell Recovery Solution (Corning). h3ACs were dissociated into single cells using Accustase (STEMCELL Technologies) at 37°C for 5-10 minutes. After washing with PBS once, h3ACs were incubated with colcemide (Gibco) at 37°C for 45 minutes. After washing with PBS once, h3AC cells were incubated with KCl to enlarge the cells. After fixation with acetone, cells were stained with

Giemsa. Karyotypes were observed using Cytovision (Leica).

### **Assessment of lung lineage transcripts by qRT-PCR**

Freshly sorted HTII-280<sup>+</sup> and HTII-280<sup>-</sup> cells were lysed with TRIzol, and RNA was extracted. RNA was reverse transcribed using SuperScript IV (Thermo Fisher Scientific), and were assessed using the following TaqMan probes; SFTPC ([Hs00951326\\_g1](#)), TP63 ([Hs01114115\\_m1](#)), SCGB1A1 ([Hs00171092\\_m1](#)) (**Table S7**).

### **Differentiation of hAT2 cells into hAT1 cells in 2D culture**

h3ACs were dissociated into single cells with Accutase (STEMCELL Technologies) at 37°C for 5-10 minutes. Cells were washed with PBS and then seeded in Lab-tek 8 well slide glass (Thermo Fisher Scientific) with 25,000 cells per well or seeded in Collagen I coated 8 well slide glass (SPL). The cells were grown under 10% human serum (Sigma), 1% P/S in DMEM. After 4 days incubation, cells were allowed to attach to slides and differentiate.

### **SARS-CoV-2 infection of 3D culture cells and Vero cells**

For 3D culture cells, cells were recovered from Matrigel with Cell Recovery Solution (CORNING). 3D culture cells were sheared with 1000 p pipette tips and incubated with Accutase (STEMCELL Technologies) at 37 °C for 5-10 minutes. After washing, the number of 3D culture cells was calculated by hemocytometer C-chip (iNCYTO), and the appropriate amount of SARS-CoV-2 virus particles was prepared from a stock vial. 3D culture cells were resuspended with each media and were infected with virus multiplicity of infection (MOI) of 1.0 or 0.1 for 2 hrs at 37 °C 5% CO<sub>2</sub>.

After 2 hrs incubation, 3D culture cells were washed twice with Advanced DMEM/F12 with 1 U/ml Penicillin/Streptomycin, 10 mM HEPES, and 1% Glutamax (v/v) (hereafter referred to as ADF+++ ) and embedded in 50 µl of GFR-Matrigel (CORNING) in 24-well plate. Each well contained at least 10,000 cells. To release viral particles from the infected cells, Matrigel was physically disrupted, and cells were lysed by freezing at -80 °C and thawing 3 times. Live virus titers were determined by plaque assay and viral RNA titer was calculated using qRT-PCR.

For Vero cell infection, Vero cells were prepared in DMEM with 10% FBS and 1% Penicillin/Streptomycin (Gibco). At the day of SARS-CoV-2 infection, media was aspirated from 6-well plates, and SARS-CoV-2 viral particles with MOI of 0.01 in 500ul of DMEM with 2% FBS and 1% Penicillin/Streptomycin (Gibco) were added to Vero cells on the 6-well plates for 1 hr at 37 °C 5% CO<sub>2</sub>. After 1 hr, infection media was aspirated, and cells were washed with

cold PBS twice. Then, 2 ml of DMEM with 2% FBS and 1% p/s solution was added to each 6-well. All work was performed in a Class II Biosafety Cabinet under BSL-3 conditions at Korea Center for Disease Control (KCDC).

### **Viral RNA extraction**

For viral RNA extraction in the infected cells, cells were retrieved from Matrigel using 500 ul of Cell Recovery Solution (Corning), and were lysed by repetitive freezing at -80 °C and thawing 3 times. Then, QIAamp Viral RNA Mini Kit (Qiagen) was used to obtain SARS-CoV-2 viral RNA according to the manufacturer's instructions. For collection of viral RNAs from culture media, 140 ul of the culture media was obtained followed by QIAamp Viral RNA Mini Kit (Qiagen).

### **Cellular total RNA extraction**

For h3ACs and h3BCs, cells were retrieved from Matrigel using Cell Recovery Solution (45 minutes incubation at 4 °C; Corning). For Vero cells, cells were detached using TrypLE Select (5 minutes incubation at 37 °C; Gibco). Cell pellets were obtained after centrifugation of retrieved cells at 300g for 3-5 minutes at 4 °C. Total RNA was extracted using RNeasy Plus Mini Kit (Qiagen).

### **Viral RNA copy number calculation with qRT-PCR**

Purified viral RNAs were reverse-transcribed using SuperScript IV (Thermo Fisher Scientific). Viral *N* gene with 2019-nCoV\_N3-F and 2019-nCoV\_N3-R probes (CDC) was targeted for qRT-PCR (**Table S7**). Each RNA sample was measured three times. Viral RNA copy number was estimated using the standard curve of virus copy number estimated by Ct value of the qRT-PCR results.

The viral copy number standard curve was generated as described below. Positive viral RNA template was reverse-transcribed, and cDNA was amplified with the CDC designed *N* gene primers and cloned into pGEM-T Easy vector (Promega, USA). The resultant plasmid DNA was linearized with PstI restriction enzyme and purified with a QIAquick PCR Purification Kit (QIAGEN). Purified template was in vitro transcribed by RiboMAX™ Large Scale RNA Production System with T7 RNA polymerase (Promega). RNA transcript was further purified with the NucleoSpin RNA Mini kit (MACHEREY-NAGEL) and quantified with spectrophotometer at 260 nm. The quantified RNA was serially diluted and each diluted

sample was reverse transcribed to measure Ct value. Then, the standard curve for viral copy number with Ct values was depicted using the Ct values with corresponding RNA doses.

### **Validation of upregulated genes with qRT-PCR**

To validate upregulated interferon genes (*INFB1*, *INFL1*, and *INFL2/3*) and ISGs (*IFI44L*, *IFI6*, *IFIT1*, and *MX1*), infected h3ACs were harvested at 0 and 3 dpi. Total cellular RNA was extracted as described above. From equal quantity of RNA (50 ng), cDNA was generated using SuperScript™ III Reverse Transcriptase kit (Invitrogen) as described previously (Pfaffl, 2001) and stored at -20°C until use. Real time PCR assay (qRT-PCR) was conducted by the ABI applied biosystems using Power SYBR Green PCR Master Mix (Applied Biosystems) (**Table S7**). The relative amounts of cytokine mRNA present (normalized with GAPDH) was determined by  $2^{-\Delta\Delta Ct}$  method.

### **Live virus titer calculation with plaque assay**

To prepare viral specimens released from infected h3AC and h3BC cells, Matrigel was disrupted by repetitive pipetting and cells were lysed by -80°C freezing and thawing once. In the case of Vero cells, culture media were directly used as viral specimens to be used in plaque assays. To perform plaque assays, newly seeded Vero cells in 12-well plates were prepared. Test specimens derived from h3ACs, h3BCs, and Vero cells were serially diluted by a scale of  $10^0 - 10^5$ . Then, 250ul of each diluted specimen was dispensed into the designated well, and incubated for 1hr at 37 °C, 5% CO<sub>2</sub>. After 1 hr, the diluted specimens were aspirated. Vero cells were washed with PBS two times, then agar and Modified Eagle's Medium (Thermo Fisher Scientific) were poured into each well. After agar mixture was solidified, the mixtures were fixed with 4% PFA for 3 days, and stained with crystal violet (Sigma-Aldrich). From the number of colonies, live virus titers were calculated.

### **Viability and cytotoxicity test**

To measure the viability and cytotoxicity in h3ACs and Vero cells after SARS-CoV-2 infection with MOI of 1.0 and 0.01, respectively, infected h3ACs were embedded in Matrigel with 10,000 cells per a well. CellTiter-Glo 3D Cell Viability Assay (Promega) was used to test viability according to the manufacturer's instructions. Briefly, 250 ul of CellTiter-Glo 3D reagent was added to each well in 24-well plates (200ul media with 50ul Matrigel, and Matrigel was disrupted. The mixture of the reagent and media was stirred up, then dispensed into

luminescence plates. After 25 mins of incubation, we recorded luminescent intensity using Spectramax L (Molecular Devices). Three biological replicates were performed for each time point.

To measure lactate dehydrogenase levels released from ruptured cells, LDH Glo Cytotoxicity Assay (Promega) was used according to the manufacturer's instructions. Briefly, triton X-100 (Sigma-Aldrich) was added to 100ul of culture media (0.2% triton X-100) to inactivate live SARS-CoV-2 virus. Then, storage buffer (Tris-HCl pH 200mM pH7.3, 10% glycerol, and 1% BSA) was added to dilute the media as indicated by the manufacturer's protocol. Diluted media was incubated with the same amount of LDH Detection Reagent (Promega) for 1 hr. After 1 hr incubation, the intensity of luminescence was measured by Spectramax L (Molecular Devices).

### **Immunofluorescence staining of paraffin-embedded h3ACs and LysoTracker**

Control h3ACs were fixed and embedded in a paraffin block (Lee et al., 2014). Pre-cut 7  $\mu$ M paraffin sections were de-waxed and rehydrated (sequential immersion in xylene, 100% EtOH, 90% EtOH, 75% EtOH, distilled water) and either stained with hematoxylin and eosin (H&E) or immunostained. For antigen retrieval, slides were submerged into pre-heated citrate antigen retrieval buffer (10 mM sodium citrate, pH 6.0) and allowed to boil for 15 min. Slides were cooled in a buffer for 20 min, washed in running water for 3 min, and permeabilized with 0.3% Triton-X in PBS for 15 min. Following permeabilization, sections of 3D models were blocked for 1 hr in 5% normal donkey serum in PBS at RT, and incubated with primary antibody mixtures overnight at 4 °C at the following dilutions; rabbit pro-SFTPC (1:500, Merck Millipore, Ab3786), mouse anti-HTII-280 (1:500, Terrace Biotech, TB-27AHT2-280), rat anti-SCGB1A1 (1:200, R&D systems, MAB4218), rabbit anti-KRT5 (1:500, Biolegend, 905501), Rabbit anti-ABCA3 (1:300, Seven Hills Bioreagents, WRAB-ABCA3), mouse anti-TP63 (1:500, abcam, ab735), and rabbit anti-HOPX (1:200, Santa Cruz Biotechnologies, sc-30216), and rabbit anti-SCRIB (1:100, GeneTex, GTX107692). Antibodies were removed with three PBS washes, and samples were incubated with Alexa Fluor-couple secondary antibodies (1:1000, Jackson ImmunoResearch Laboratories) for 1 hr at RT. Following PBS washes, nuclei were stained with DAPI for 5 min, slides mounted with Rapiclear (SUNjin lab), and sealed with clear nail polish. For LysoTracker™ staining of lysosomes, live h3ACs in 48-well plates were incubated *in situ* with 50 ng/ $\mu$ L of LysoTracker™ (Invitrogen, L12492), diluted in pre-warmed expansion medium, for 30 min at 37 °C. LysoTracker was removed, and Matrigel suspension was carefully washed for 5 min in PBS, followed by addition of fresh, pre-warmed expansion medium. Cells were protected from light and either imaged immediately using an EVOS cell imaging system,

or analyzed by FACS on an Aria III fusion (BD Biosciences) and FlowJo software (Tree Star, Inc.)

### **Immunofluorescence staining of infected h3ACs with cryo-section**

Both uninfected and infected h3ACs were fixed in 4% paraformaldehyde (PFA) for 3 hrs at 4 °C, and then dehydrated in PBS with 30% sucrose (v/v) (Sigma-Aldrich). The h3ACs were embedded with optimal cutting temperature (OCT) compound (Leica) and cut into 7-10 µm thick sections. The sections were blocked with PBS with 5% normal donkey serum and 1% triton X-100 (Sigma-Aldrich). The sections were incubated with primary antibodies overnight at 4 °C, then washed, and incubated with host matched Alexa Fluor-couple secondary antibodies (1:1000, Jackson ImmunoResearch Laboratories) for 1.5 hr at RT. Following DAPI incubation, slides were mounted. E-cadherin (1:300, R&D SYSTEMS, AF748), NP (1:200, Sino Biological, 40143-MM05), NP (1:200, Sino Biological, 40143-T62), dsRNA (1:2, SCICONS, 10030005), ACE2 (1:400, abcam, ab15348), TMPRSS2 (1:400, Santa Cruz, 515727), pro-SFTPC (1:400, abcam, ab90716), MX1 (1:400, GeneTex, GTX110256), and HTII-280 (1:50, Terrace Biotech, TB-27AHT2) were used.

### **Immunofluorescence staining of whole-mount h3ACs**

To check polarity of h3ACs, we conducted Z-stack whole mount immunofluorescence. H3ACs were seeded with the Phenol-free Matrigel (Corning) in the 8-well chamber slides (Thermo Fisher Scientific). Each well was fixed with 4% PFA (Biosesang) at 4°C for 45 min and washed with PBS for 5 min 2 times. Organoids were blocked with PBS with 1% BSA (Sigma-Aldrich) and 0.1% triton X-100 (Sigma-Aldrich). The sections were incubated with primary antibodies overnight at 4°C, then washed, and incubated with host matched Alexa Fluor-couple secondary antibodies (1:1000, Jackson ImmunoResearch Laboratories) for 1.5 hr at RT. Alexa 647 Phalloidin (1:100) were incubated for 1hr at RT. Following DAPI incubation, slides were mounted. Crb3 (1:400, NOVUS BIOLOGICALS, NBP1-81185), HTII-280 (1:200, Terrace Biotech, TB-27AHT2) were used.

### **Immunofluorescence staining of hAT1-like cells**

Attached hAT1-like cells were fixed with 4% PFA at 4°C for 3 hrs. After cells were blocked in PBS with 5% normal donkey serum and 1% triton X-100 (Sigma-Aldrich), sections were incubated with primary antibodies overnight at 4 °C, then washed, and incubated with host matched Alexa Fluor-couple secondary antibodies (1:1000, Jackson ImmunoResearch

Laboratories) (1:1000) for 1.5 hr at RT. Following DAPI incubation, slides were mounted. AGER (1:400, R&D SYSTEMS, AF1145), Aquaporin5 (1:200, abcam, ab92320) pro-SFTPC (1:400, abcam, ab90716) were used.

### **Transmission electron microscopy**

Uninfected h3ACs and infected h3ACs at 2 dpi were fixed with 2.5 % glutaraldehyde in 0.1 M PBS for overnight incubation at 4 °C (Fujii et al., 2018). The h3ACs were washed with PBS and post-fixed with 2% osmium tetroxide for 1.5 hr. The fixed samples were dehydrated in graded ethanol, substituted with propylene oxide, and finally embedded in EMbed-812 resin (EMS). Polymerization was performed at 60 °C for 24 hrs. Ultrathin (100 nm) sections were prepared using an ultramicrotome (Leica, EM UC7). Images were captured with a transmission electron microscope (FEI Tecnai G2 spirit TWIN, eagle 4K CCD camera) at 120kV acceleration voltage. All work was carried out in EM & Histology Core Facility, at BioMedical Research Center, KAIST. For control EM image, data were stitched manually (**Figure S4A**).

### **Bulk RNA sequencing and data processing**

Total RNA sequencing library was constructed using Truseq Stranded Total RNA Gold kit (Illumina) according to the manufacturer's protocol followed by sequencing with 2 x 100 bp using Hiseq 2500. Fastq files were aligned to GRCh38 (human cells) with SARS-CoV-2 virus sequence (NC 045512.2 from NCBI) or ChISab1.1 (Vero cells) with SARS-CoV-2 virus sequence, yeast *ENO2* cDNA sequence (SGD: YHR174W), and human ribosomal DNA complete repeat unit sequence (U13369.1)(Kim et al., 2020a) using STAR v2.6.1d (Dobin et al., 2013) and normalized counts of total RNA expression were calculated using RSEM v1.3.1 (Li and Dewey, 2011).

The top 100 genes with highest variations (standard deviations / mean of TPM values of each gene) in h3ACs were demonstrated as heatmap (**Figure 5A and Figure S5E**) using R package "ComplexHeatmap". Enriched pathways were acquired from Enrichr website with input of the top 100 variable genes (Kuleshov et al., 2016).

Previously published bulk RNA sequencing data (fastq files) were obtained from the NCBI Gene Expression Omnibus (GEO) server under the accession number GSE147507 for COVID-19 patients and healthy donors and GSE148729 for 2D cell lines including mock and

SARS-CoV-2 treatment of Calu-3, Caco-2, and NCI-H1299 at 24 hpi (Blanco-Melo et al., 2020; Emanuel et al., 2020a). The fastq files were processed as mentioned above.

Fold changes and adjusted  $p$ -values of DEGs were calculated using R package “DEseq2” (Love et al., 2014). The results were demonstrated as volcano plots using R package “EnhancedVolcano” (Blighe et al., 2018). Differentially upregulated genes were selected when fold change and adjusted  $p$ -value of each DEG were more than 1.5 and less than  $10^{-6}$ , respectively, in SARS-CoV-2 infected samples compared to uninfected samples (**Figures S5D, S5H, and Table S2**). To investigate single base substitutions and small indels, we used union sets of variants called by Strelka2 (Saunders et al., 2012) and VarScan2 (Koboldt et al., 2012) Then, in-house script was used to filter false-positive calls, and we manually curated the variant positions using Integrative Genomics Viewer (Robinson et al., 2011).

### **Single-cell RNA sequencing and data processing**

Fastq files were aligned and UMI counts were calculated using Cell Ranger software (3.1.0) provided by the manufacturer (10X Genomics). The 10X gene count data includes 19,941 human genes and SARS-CoV-2 genome as an extra gene.

Normalization was done using a log-normalization method (same as ‘NormalizeData from ‘Seurat’ R package) controlling the total human gene UMI count (after scaling) at 10k. After scaling, each count is log2-transformed after added by 1.

From the initial 30,027 cells with reported UMI counts across five experimental conditions, we filtered out cells with >25% mitochondria RNA percent ( $n=2,479$ ), pathologic cell clusters characterized with KRT5-/KRT17+ ( $n=13,175$ ) (Adams et al., 2020; Habermann et al., 2020), and contaminated cells ( $n=199$ ). The remaining 14,174 cells constitutes the core dataset, and used for graphical visualization as well as clustering analyses.

Using ‘Seurat V3’ package (Stuart et al., 2019), 3,000 variable features were selected, principal components (PCs) were constructed, and UMAP dimension reduction was performed based on 30 PCs. Then, unsupervised clustering analysis was carried out by constructing a shared nearest neighbor (SNN) graph (using ‘FindNeighbors’ function with parameters of ‘dims=1:20’) and identifying clusters of cells (using ‘FindClusters’ function with ‘resolution’ parameter of 0.1). Using ‘SC3’ package (Kiselev et al., 2019), additional

clustering exercise was carried out and marker genes were highlighted. Specifically, ‘SC3’ clusters were constructed at ‘k=9’ parameter setting. Then, marker genes most representing each cluster were found and ranked by internal ‘auroc’ metric. The marker genes with the associated auroc > 0.85 (default) were visualized. Using 150 variable genes calculated by ‘SC3’ package, enriched pathways in SARS-CoV-2 infected cells were obtained from Enrichr website (Kuleshov et al., 2016).

For investigating C>U at the 23,707 position of SARS-CoV-2 transcripts in each cell, subset-bam tool (10X Genomics) was used to get an individual bam file per each cell. Samtools (Li et al., 2009) was used to call the substitution. In house python and R scripts were used for more downstream analyses.

### Statistical inference on the effective number of viral entry

To assess whether hAT2 cells tend to be infected by a single viral particle or multiple particles, we employed a likelihood approach. As a proof-of-concept, we assumed only two scenarios exist, one supporting a single viral entry and the other supporting double viral entry, then aimed to estimate the proportion of cells with a single viral particle ( $w$ ). Our data consists of the observed reference ( $n_{cell}^r$ ) and alternative ( $n_{cell}^v$ ) read counts for each of the 547 cells reporting at least one read at the mutation site of NC\_045512.2:23,707. Assuming a sequencing error rate ( $\epsilon$ ) of 0.1%, which will cover any Illumina sequencing errors or misalignment, the likelihood of data given the weight supporting a single virus scenario ( $w$ ) was computed as follows.

$$L(w) = \prod_{cell=1}^{547} L_{cell}(w)$$

$$L_{cell}(w) = w \times P(Data_{cell} | \text{single entry}) + (1 - w) \times P(Data_{cell} | \text{double entry})$$

$$P(Data_{cell} | \text{single entry}) = P(G = R) \times P(n_{cell}^r, n_{cell}^v | vaf_{cell} = \epsilon)$$

$$+ P(G = V) \times P(n_{cell}^r, n_{cell}^v | vaf_{cell} = 1 - \epsilon)$$

$$P(Data_{cell} | \text{double entry}) = P(G = RR) \times P(n_{cell}^r, n_{cell}^v | vaf_{cell} = \epsilon)$$

$$+ P(G = RV) \times P(n_{cell}^r, n_{cell}^v | vaf_{cell} = 0.5)$$

$$+ P(G = VV) \times P(n_{cell}^r, n_{cell}^v | vaf_{cell} = 1 - \epsilon)$$

$G$  stands for the viral genotype for each cell, and  $R$  and  $V$  denote for the reference (wildtype) and variant (mutant) alleles in the genotype, respectively.  $vaf_{cell}$  refers to the true mutant allele frequency within an infected cell given the genotype after accounting for the sequencing error rate.

For both single- and double-entry scenarios, the prior for each genotype was computed using the mutant allele frequency estimated from the whole cell population ( $vaf_{pop}$ ; 4.3%=42 variant reads out of 986 reads). For example, the prior of having the genotype with two mutant alleles in a double-entry scenario is computed as follows:

$$P(G = VV) = vaf_{pop}^2$$

Within a cell, the probability of observing the number of reference and variant reads given the mutant allele frequency,  $P(n_{cell}^r, n_{cell}^v | vaf_{cell})$ , was computed using a binomial distribution.

## QUANTIFICATION AND STATISTICAL ANALYSIS

To quantify folded and cystic 3D structures of h3ACs, we at least counted 50 organoids for each donor, and the total number of donors is 3. For quantification of colony forming efficiency (CFE) for h3AC, CFE was calculated from total 3 donors at each passage (Data is represented as mean  $\pm$  SEM). In Plaque assay for live virus titer, the number of viruses was counted from two technical replicates for each designated timepoints (Data is represented as mean  $\pm$  SEM). In qPCR for virus RNA titer and interferon induced gene validation, 3 technical replicates were used (Data is represented as mean  $\pm$  SEM). To compare AT2 marker gene's expression, qPCR was conducted with 2 technical replicates (Data is represented as mean  $\pm$  SEM). Viability and LDH cytotoxicity assay were replicated with 3 biological replicates (Data is represented as mean  $\pm$  SEM). All the immunofluorescence images were taken at least 5 different biological replicates and representative image was used. To compare MX1 intensity in control (n = 11) and infected (n = 13) h3ACs, Wilcoxon rank sum test was used. For transmission electron microscopy, more than 10 biological replicates were taken. For karyotyping, 6 normal cell karyotypes were observed and representative image was used. RNA sequencing statistical analysis was written in "RNA sequencing and data processing" session in METHOD DETAILS. Single cell RNA sequencing statistical analysis was written in "Single cell RNA sequencing and data processing" session in METHOD DETAILS.

## **SUPPLEMENTAL EXCEL TABLE TITLES AND LEGENDS**

**Table S1.** RNA expression levels (TPM) of all genes in seven human 3D alveolar culture samples. Related to Figure 5.

**Table S2.** List of differentially upregulated genes in SARS-CoV-2 infected human lung, h3AC, Vero cell line, Calu-3 cell line, and Caco-2 cell lines. Related to Figures 5 and S5.

**Table S3.** RNA expression levels (TPM) of all genes in seven human 3D bronchial (airway) culture samples. Related to Figure 5.

**Table S4.** Twenty single base substitutions of SARS-CoV-2 which were detected in infected human 3D alveolar cultures with MOI of 1 at 3 dpi. Related to Figure 7.

**Table S5.** List of differentially expressed genes in single cell transcriptome data, calculated by SC3 packages. Related to Figure 7.

## References

- Adams, T.S., Schupp, J.C., Poli, S., Ayaub, E.A., Neumark, N., Ahangari, F., Chu, S.G., Raby, B.A., Deluiliis, G., Januszyk, M., *et al.* (2020). Single-cell RNA-seq reveals ectopic and aberrant lung-resident cell populations in idiopathic pulmonary fibrosis. *Science Advances* 6, eaba1983.
- Andersen, K.G., Rambaut, A., Lipkin, W.I., Holmes, E.C., and Garry, R.F. (2020). The proximal origin of SARS-CoV-2. *Nat Med* 26, 450-452.
- Arbour, N., Cote, G., Lachance, C., Tardieu, M., Cashman, N.R., and Talbot, P.J. (1999). Acute and persistent infection of human neural cell lines by human coronavirus OC43. *J Virol* 73, 3338-3350.
- Bao, L., Deng, W., Huang, B., Gao, H., Liu, J., Ren, L., Wei, Q., Yu, P., Xu, Y., Qi, F., *et al.* (2020). The pathogenicity of SARS-CoV-2 in hACE2 transgenic mice. *Nature*.
- Bar-On, Y.M., Flamholz, A., Phillips, R., and Milo, R. (2020). SARS-CoV-2 (COVID-19) by the numbers. *Elife* 9.
- Barkauskas, C.E., Crouce, M.J., Rackley, C.R., Bowie, E.J., Keene, D.R., Stripp, B.R., Randell, S.H., Noble, P.W., and Hogan, B.L.M. (2013). Type 2 alveolar cells are stem cells in adult lung. *J Clin Invest* 123, 3025-3036.
- Bartfeld, S., Bayram, T., van de Wetering, M., Huch, M., Begthel, H., Kujala, P., Vries, R., Peters, P.J., and Clevers, H. (2015). In vitro expansion of human gastric epithelial stem cells and their responses to bacterial infection. *Gastroenterology* 148, 126-136 e126.
- Blanco-Melo, D., Nilsson-Payant, B.E., Liu, W.C., Uhl, S., Hoagland, D., Moller, R., Jordan, T.X., Oishi, K., Panis, M., Sachs, D., *et al.* (2020). Imbalanced Host Response to SARS-CoV-2 Drives Development of COVID-19. *Cell* 181, 1036-1045 e1039.
- Blighe, K., Rana, S., and Lewis, M. (2018). EnhancedVolcano: Publication-ready volcano plots with enhanced colouring and labeling.
- Chu, H., Chan, J.F.-W., Yuen, T.T.-T., Shuai, H., Yuan, S., Wang, Y., Hu, B., Yip, C.C.-Y., Tsang, J.O.-L., Huang, X., *et al.* (2020). Comparative tropism, replication kinetics, and cell damage profiling of SARS-CoV-2 and SARS-CoV with implications for clinical manifestations, transmissibility, and laboratory studies of COVID-19: an observational study. *The Lancet Microbe* 1, e14-e23.
- Dobbs, L.G. (1990). Isolation and culture of alveolar type II cells. *Am J Physiol* 258, L134-147.
- Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15-21.
- El-Hashash, A.H., and Warburton, D. (2011). Cell polarity and spindle orientation in the distal epithelium of embryonic lung. *Dev Dyn* 240, 441-445.
- Elbadawi, M., and Efferth, T. (2020). Organoids of human airways to study infectivity and cytopathy of SARS-CoV-2. *Lancet Respir Med*.
- Emanuel, W., Kirstin, M., Vedran, F., Asija, D., Theresa, G.L., Roberto, A., Filippos, K., David, K., Salah, A., Christopher, B., *et al.* (2020a). Bulk and single-cell gene expression profiling of SARS-CoV-2 infected human cell lines identifies molecular targets for therapeutic intervention. *bioRxiv*, 2020.2005.2005.079194.
- Emanuel, W., Kirstin, M., Vedran, F., Asija, D., Theresa, G.L., Roberto, A., Filippos, K., David, K., Salah, A., Christopher, B., *et al.* (2020b). Bulk and single-cell gene expression profiling of SARS-CoV-2

infected human cell lines identifies molecular targets for therapeutic intervention. bioRxiv.

Fatehullah, A., Tan, S.H., and Barker, N. (2016). Organoids as an in vitro model of human development and disease. *Nat Cell Biol* 18, 246-254.

Forero, A., Ozarkar, S., Li, H., Lee, C.H., Hemann, E.A., Nadjombati, M.S., Hendricks, M.R., So, L., Green, R., Roy, C.N., *et al.* (2019). Differential Activation of the Transcription Factor IRF1 Underlies the Distinct Immune Responses Elicited by Type I and Type III Interferons. *Immunity* 51, 451-464 e456.

Forster, P., Forster, L., Renfrew, C., and Forster, M. (2020). Phylogenetic network analysis of SARS-CoV-2 genomes. *Proc Natl Acad Sci U S A* 117, 9241-9243.

Frank, D.B., Peng, T., Zepp, J.A., Snitow, M., Vincent, T.L., Penkala, I.J., Cui, Z., Herriges, M.J., Morley, M.P., Zhou, S., *et al.* (2016). Emergence of a Wave of Wnt Signaling that Regulates Lung Alveologenesis by Controlling Epithelial Self-Renewal and Differentiation. *Cell Rep* 17, 2312-2325.

Fujii, M., Matano, M., Toshimitsu, K., Takano, A., Mikami, Y., Nishikori, S., Sugimoto, S., and Sato, T. (2018). Human Intestinal Organoids Maintain Self-Renewal Capacity and Cellular Diversity in Niche-Inspired Culture Condition. *Cell Stem Cell* 23, 787-793 e786.

Gonzalez, R.F., Allen, L., Gonzales, L., Ballard, P.L., and Dobbs, L.G. (2010). HTII-280, a biomarker specific to the apical plasma membrane of human lung alveolar type II cells. *J Histochem Cytochem* 58, 891-901.

Habermann, A.C., Gutierrez, A.J., Bui, L.T., Yahn, S.L., Winters, N.I., Calvi, C.L., Peter, L., Chung, M.-I., Taylor, C.J., Jetter, C., *et al.* (2020). Single-cell RNA sequencing reveals profibrotic roles of distinct epithelial and mesenchymal lineages in pulmonary fibrosis. *Science Advances* 6, eaba1972.

Hamming, I., Timens, W., Bulthuis, M.L., Lely, A.T., Navis, G., and van Goor, H. (2004). Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J Pathol* 203, 631-637.

Heo, I., Dutta, D., Schaefer, D.A., Iakobachvili, N., Artegiani, B., Sachs, N., Boonekamp, K.E., Bowden, G., Hendrickx, A.P.A., Willems, R.J.L., *et al.* (2018). Modelling Cryptosporidium infection in human small intestinal and lung organoids. *Nat Microbiol* 3, 814-823.

Hoffmann, M., Kleine-Weber, H., Schroeder, S., Kruger, N., Herrler, T., Erichsen, S., Schiergens, T.S., Herrler, G., Wu, N.H., Nitsche, A., *et al.* (2020). SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 181, 271-280 e278.

Huang, J., Hume, A.J., Abo, K.M., Werder, R.B., Villacorta-Martin, C., Alysandratos, K.D., Beermann, M.L., Simone-Roach, C., Lindstrom-Vautrin, J., Olejnik, J., *et al.* (2020). SARS-CoV-2 Infection of Pluripotent Stem Cell-Derived Human Lung Alveolar Type 2 Cells Elicits a Rapid Epithelial-Intrinsic Inflammatory Response. *Cell Stem Cell*.

Iwasaki, A., Foxman, E.F., and Molony, R.D. (2017). Early local immune defences in the respiratory tract. *Nat Rev Immunol* 17, 7-20.

Jacob, F., Pather, S.R., Huang, W.K., Zhang, F., Wong, S.Z.H., Zhou, H., Cubitt, B., Fan, W., Chen, C.Z., Xu, M., *et al.* (2020). Human Pluripotent Stem Cell-Derived Neural Cells and Brain Organoids Reveal SARS-CoV-2 Neurotropism Predominates in Choroid Plexus Epithelium. *Cell Stem Cell*.

Jensen-Taubman, S.M., Steinberg, S.M., and Linnoila, R.I. (1998). Bronchiolization of the alveoli in lung cancer: pathology, patterns of differentiation and oncogene expression. *Int J Cancer* 75, 489-496.

Kathiriyai, J.J., Wang, C., Brumwell, A., Cassandras, M., Le Saux, C., Wolters, P., Matthay, M., Chapman,

H.A., and Peng, T. (2020). Human alveolar Type 2 epithelium transdifferentiates into metaplastic KRT5+ basal cells during alveolar repair. *bioRxiv*.

Kim, D., Lee, J.Y., Yang, J.S., Kim, J.W., Kim, V.N., and Chang, H. (2020a). The Architecture of SARS-CoV-2 Transcriptome. *Cell* 181, 914-921 e910.

Kim, J.M., Chung, Y.S., Jo, H.J., Lee, N.J., Kim, M.S., Woo, S.H., Park, S., Kim, J.W., Kim, H.M., and Han, M.G. (2020b). Identification of Coronavirus Isolated from a Patient in Korea with COVID-19. *Osong Public Health Res Perspect* 11, 3-7.

Kim, Y.I., Kim, S.G., Kim, S.M., Kim, E.H., Park, S.J., Yu, K.M., Chang, J.H., Kim, E.J., Lee, S., Casel, M.A.B., *et al.* (2020c). Infection and Rapid Transmission of SARS-CoV-2 in Ferrets. *Cell Host Microbe* 27, 704-709 e702.

Kiselev, V.Y., Andrews, T.S., and Hemberg, M. (2019). Challenges in unsupervised clustering of single-cell RNA-seq data. *Nat Rev Genet* 20, 273-282.

Koboldt, D.C., Zhang, Q., Larson, D.E., Shen, D., McLellan, M.D., Lin, L., Miller, C.A., Mardis, E.R., Ding, L., and Wilson, R.K. (2012). VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* 22, 568-576.

Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., *et al.* (2016). Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. *Nucleic Acids Res* 44, W90-97.

Lamers, M.M., Beumer, J., van der Vaart, J., Knoop, K., Puschhof, J., Breugem, T.I., Ravelli, R.B.G., Paul van Schayck, J., Mykytyn, A.Z., Duimel, H.Q., *et al.* (2020). SARS-CoV-2 productively infects human gut enterocytes. *Science*.

Lauer, S.A., Grantz, K.H., Bi, Q., Jones, F.K., Zheng, Q., Meredith, H.R., Azman, A.S., Reich, N.G., and Lessler, J. (2020). The Incubation Period of Coronavirus Disease 2019 (COVID-19) From Publicly Reported Confirmed Cases: Estimation and Application. *Ann Intern Med* 172, 577-582.

Lee, J.H., Bhang, D.H., Beede, A., Huang, T.L., Stripp, B.R., Bloch, K.D., Wagers, A.J., Tseng, Y.H., Ryeom, S., and Kim, C.F. (2014). Lung stem cell differentiation in mice directed by endothelial cells via a BMP4-NFATc1-thrombospondin-1 axis. *Cell* 156, 440-455.

Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12, 323.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550.

Maier, H.J., Hawes, P.C., Cottam, E.M., Mantell, J., Verkade, P., Monaghan, P., Wileman, T., and Britton, P. (2013). Infectious bronchitis virus generates spherules from zippered endoplasmic reticulum membranes. *mBio* 4, e00801-00813.

Manna, V.J., and Caradonna, S.J. (2020). The heterogeneous nature of the Coronavirus receptor, angiotensin-converting enzyme 2 (ACE2) in differentiating airway epithelia. *bioRxiv*.

Monel, B., Compton, A.A., Bruel, T., Amraoui, S., Burlaud-Gaillard, J., Roy, N., Guivel-Benhassine, F., Porrot, F., Genin, P., Meertens, L., *et al.* (2017). Zika virus induces massive cytoplasmic vacuolization

and paraptosis-like death in infected cells. *EMBO J* 36, 1653-1668.

Monteil, V., Kwon, H., Prado, P., Hagelkruys, A., Wimmer, R.A., Stahl, M., Leopoldi, A., Garreta, E., Hurtado Del Pozo, C., Prosper, F., *et al.* (2020). Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues Using Clinical-Grade Soluble Human ACE2. *Cell* 181, 905-913 e907.

Muus, C., Luecken, M.D., Eraslan, G., Waghray, A., Heimberg, G., Sikkema, L., Kobayashi, Y., Vaishnav, E.D., Subramanian, A., Smilie, C., *et al.* (2020). Integrated analyses of single-cell atlases reveal age, gender, and smoking status associations with cell type-specific expression of mediators of SARS-CoV-2 viral entry and highlights inflammatory programs in putative target cells. *bioRxiv*.

Nabhan, A.N., Brownfield, D.G., Harbury, P.B., Krasnow, M.A., and Desai, T.J. (2018). Single-cell Wnt signaling niches maintain stemness of alveolar type 2 cells. *Science* 359, 1118-1123.

Nikolic, M.Z., Caritg, O., Jeng, Q., Johnson, J.A., Sun, D., Howell, K.J., Brady, J.L., Laresgoiti, U., Allen, G., Butler, R., *et al.* (2017). Human embryonic lung epithelial tips are multipotent progenitors that can be expanded in vitro as long-term self-renewing organoids. *Elife* 6.

Ogando, N.S., Dalebout, T.J., Zevenhoven-Dobbe, J.C., Limpens, R.W., van der Meer, Y., Caly, L., Druce, J., de Vries, J.J.C., Kikkert, M., Bárcena, M., *et al.* (2020). SARS-coronavirus-2 replication in Vero E6 cells: replication kinetics, rapid adaptation and cytopathology. *bioRxiv*.

Park, A., and Iwasaki, A. (2020). Type I and Type III Interferons - Induction, Signaling, Evasion, and Application to Combat COVID-19. *Cell Host Microbe*.

Pfaffl, M.W. (2001). A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29, e45.

Pocha, S.M., and Knust, E. (2013). Complexities of Crumbs function and regulation in tissue morphogenesis. *Curr Biol* 23, R289-293.

Ramani, A., Muller, L., Ostermann, P.N., Gabriel, E., Abida-Islam, P., Muller-Schiffmann, A., Mariappan, A., Goureau, O., Gruell, H., Walker, A., *et al.* (2020). SARS-CoV-2 targets neurons of 3D human brain organoids. *EMBO J*, e106230.

Ren, X., Glende, J., Al-Falah, M., de Vries, V., Schwegmann-Wessels, C., Qu, X., Tan, L., Tschernig, T., Deng, H., Naim, H.Y., *et al.* (2006). Analysis of ACE2 in polarized epithelial cells: surface expression and function as receptor for severe acute respiratory syndrome-associated coronavirus. *J Gen Virol* 87, 1691-1695.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nat Biotechnol* 29, 24-26.

Rodriguez-Boulan, E., and Macara, I.G. (2014). Organization and execution of the epithelial polarity programme. *Nat Rev Mol Cell Biol* 15, 225-242.

Sachs, N., Papaspyropoulos, A., Zomer-van Ommen, D.D., Heo, I., Bottinger, L., Klay, D., Weeber, F., Huelsz-Prince, G., Iakobachvili, N., Amatngalim, G.D., *et al.* (2019). Long-term expanding human airway organoids for disease modeling. *EMBO J* 38.

Saunders, C.T., Wong, W.S., Swamy, S., Becq, J., Murray, L.J., and Cheetham, R.K. (2012). Strelka: accurate somatic small-variant calling from sequenced tumor-normal sample pairs. *Bioinformatics* 28, 1811-1817.

Schneider, W.M., Chevillotte, M.D., and Rice, C.M. (2014). Interferon-stimulated genes: a complex web of host defenses. *Annu Rev Immunol* 32, 513-545.

- Shang, J., Ye, G., Shi, K., Wan, Y., Luo, C., Aihara, H., Geng, Q., Auerbach, A., and Li, F. (2020). Structural basis of receptor recognition by SARS-CoV-2. *Nature* 581, 221-224.
- Sia, S.F., Yan, L.M., Chin, A.W.H., Fung, K., Choy, K.T., Wong, A.Y.L., Kaewpreedee, P., Perera, R., Poon, L.L.M., Nicholls, J.M., *et al.* (2020). Pathogenesis and transmission of SARS-CoV-2 in golden hamsters. *Nature*.
- Stuart, T., Butler, A., Hoffman, P., Hafemeister, C., Papalexi, E., Mauck, W.M., 3rd, Hao, Y., Stoeckius, M., Smibert, P., and Satija, R. (2019). Comprehensive Integration of Single-Cell Data. *Cell* 177, 1888-1902 e1821.
- Suzuki, T., Itoh, Y., Sakai, Y., Saito, A., Okuzaki, D., Motooka, D., Minami, S., Kobayashi, T., Yamamoto, T., Okamoto, T., *et al.* (2020). Generation of human bronchial organoids for SARS-CoV-2 research. *bioRxiv*.
- Szymaniak, A.D., Mahoney, J.E., Cardoso, W.V., and Varelas, X. (2015). Crumbs3-Mediated Polarity Directs Airway Epithelial Cell Fate through the Hippo Pathway Effector Yap. *Dev Cell* 34, 283-296.
- Travaglini, K.J., Nabhan, A.N., Penland, L., Sinha, R., Gillich, A., Sit, R.V., Chang, S., Conley, S.D., Mori, Y., Seita, J., *et al.* (2020). A molecular cell atlas of the human lung from single cell RNA sequencing. *bioRxiv*.
- Van der Velden, J.L., Bertoncello, I., and McQualter, J.L. (2013). LysoTracker is a marker of differentiated alveolar type II cells. *Respir Res* 14, 123.
- Walls, A.C., Park, Y.J., Tortorici, M.A., Wall, A., McGuire, A.T., and Velesler, D. (2020). Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* 181, 281-292 e286.
- Wang, Q., Zhang, Y., Wu, L., Niu, S., Song, C., Zhang, Z., Lu, G., Qiao, C., Hu, Y., Yuen, K.Y., *et al.* (2020). Structural and Functional Basis of SARS-CoV-2 Entry by Using Human ACE2. *Cell* 181, 894-904 e899.
- Wolfel, R., Corman, V.M., Guggemos, W., Seilmaier, M., Zange, S., Muller, M.A., Niemeyer, D., Jones, T.C., Vollmar, P., Rothe, C., *et al.* (2020). Virological assessment of hospitalized patients with COVID-2019. *Nature* 581, 465-469.
- Wu, F., Zhao, S., Yu, B., Chen, Y.M., Wang, W., Song, Z.G., Hu, Y., Tao, Z.W., Tian, J.H., Pei, Y.Y., *et al.* (2020). A new coronavirus associated with human respiratory disease in China. *Nature* 579, 265-269.
- Xu, Y., Mizuno, T., Sridharan, A., Du, Y., Guo, M., Tang, J., Wikenheiser-Brokamp, K.A., Perl, A.T., Funari, V.A., Gokey, J.J., *et al.* (2016). Single-cell RNA sequencing identifies diverse roles of epithelial cells in idiopathic pulmonary fibrosis. *JCI Insight* 1, e90558.
- Yang, L., Han, Y., Nilsson-Payant, B.E., Gupta, V., Wang, P., Duan, X., Tang, X., Zhu, J., Zhao, Z., Jaffre, F., *et al.* (2020). A Human Pluripotent Stem Cell-based Platform to Study SARS-CoV-2 Tropism and Model Virus Infection in Human Cells and Organoids. *Cell Stem Cell* 27, 125-136 e127.
- Zacharias, W.J., Frank, D.B., Zepp, J.A., Morley, M.P., Alkhaleel, F.A., Kong, J., Zhou, S., Cantu, E., and Morrissey, E.E. (2018). Regeneration of the lung alveolus by an evolutionarily conserved epithelial progenitor. *Nature* 555, 251-255.
- Zhao, Y., Zhao, Z., Wang, Y., Zhou, Y., Ma, Y., and Zuo, W. (2020). Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCov. *bioRxiv*.
- Zhou, L., Niu, Z., Jiang, X., Zhang, Z., Zheng, Y., Wang, Z., Zhu, Y., Gao, L., Wang, X., and Sun, Q.

(2020). Systemic analysis of tissue cells potentially vulnerable to SARS-CoV-2 infection by the protein-proofed single-cell RNA profiling of ACE2, TMPRSS2 and Furin proteases. bioRxiv.

Ziegler, C.G.K., Allon, S.J., Nyquist, S.K., Mbanjo, I.M., Miao, V.N., Tzouanas, C.N., Cao, Y., Yousif, A.S., Bals, J., Hauser, B.M., *et al.* (2020). SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. *Cell* *181*, 1016-1035 e1019.

