

UCSF

UC San Francisco Previously Published Works

Title

SARS-CoV-2 infection studies in lung organoids identify TSPAN8 as novel mediator

Permalink

<https://escholarship.org/uc/item/8628c7m9>

Authors

Roose, Jeroen
Hysenaj, Lisiena
Little, Samantha
[et al.](#)

Publication Date

2021

DOI

10.21203/rs.3.rs-581732/v1

Peer reviewed

SARS-CoV-2 infection studies in lung organoids identify TSPAN8 as novel mediator

Jeroen Roose (✉ jeroen.roose@ucsf.edu)

University of California San Francisco <https://orcid.org/0000-0003-4746-2811>

Lisiena Hysenaj

University of California San Francisco

Samantha Little

University of California San Francisco

Kayla Kulhanek

University of California San Francisco

Oghenekevwe Gbenedio

University of California San Francisco

Lauren Rodriguez

University of California, San Francisco

Alan Shen

University of California, San Francisco

Jean-Chrostophe Lone

Essex University

Leonard Lupin-Jimenez

University of California San Francisco

Luke Bonser

University of California San Francisco

Kriti Bahl

University of California San Francisco

Eran Mick

University of California, San Francisco <https://orcid.org/0000-0002-7299-808X>

Nina Serwas

University of California, San Francisco <https://orcid.org/0000-0003-2243-5446>

Dong Vivianne

University of California San Francisco

Jack Li

University of California San Francisco

Mazharul Maishan

University of California San Francisco

Shotaro Matsumoto

University of California San Francisco

Gabriela Fragiadakis

UCSF Data Science CoLab and UCSF Department of Medicine, University of California, San Francisco

Michael Matthay

University of California, San Francisco <https://orcid.org/0000-0003-3039-8155>

David Jablons

University of California San Francisco

Melanie Ott

Gladstone Institutes <https://orcid.org/0000-0002-5697-1274>

Charles Langelier

University of California, San Francisco <https://orcid.org/0000-0002-6708-4646>

Anita Sil

University of California San Francisco

Matthew Krummel

University of California, San Francisco <https://orcid.org/0000-0001-7915-3533>

Alexis Combes

University of California - San Francisco <https://orcid.org/0000-0002-9110-6542>

David Erle

University of California, San Francisco <https://orcid.org/0000-0002-2171-0648>

Johannes Kratz

University of California San Francisco

Jeroen Roose

University of California San Francisco

Biological Sciences - Article

Keywords: SARS coronavirus-2 (SARS-CoV-2), COVID-19, virology, microbiology, biomarkers

DOI: <https://doi.org/10.21203/rs.3.rs-581732/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

**SARS-CoV-2 infection
studies in lung organoids
identify TSPAN8 as novel mediator**

Lisiena Hysenaj¹, Samantha Little¹, Kayla Kulhanek¹, Oghenekevwe M. Gbenedio¹, Lauren Rodriguez^{2, 3}, Alan Shen³, Jean-Christophe Lone⁴, Leonard C. Lupin-Jimenez³, Luke R. Bonser⁵, Nina K. Serwas⁶, Kriti Bahl¹, Eran Mick⁷, Jack Z. Li⁸, Vivianne W. Ding⁸, Shotaro Matsumoto⁹, Mazharul Maishan⁹, Camille Simoneau¹⁰, Gabriela Fragiadakis^{3,11}, David M. Jablons⁸, Charles R. Langelier^{7,12}, Michael Matthay⁹, Melanie Ott¹⁰, Matthew Krummel⁶, Alexis J. Combes³, Anita Sil², David J. Erle^{3,5}, Johannes R. Kratz⁸, and Jeroen P. Roose^{1, #}.

¹ Department of Anatomy, University of California, San Francisco, San Francisco, California 94143, USA.

² Department of Microbiology and Immunology, University of California, San Francisco, San Francisco, California 94143, USA

³ UCSF CoLabs, University of California San Francisco, San Francisco, California 94143, USA

⁴ School of Life Science, University of Essex, Wivenhoe Park, Colchester C04 3SQ, United Kingdom

⁵ Lung Biology Center, Department of Medicine, University of California, San Francisco, San Francisco, California 94143, USA

⁶ Department of Pathology, University of California, San Francisco, San Francisco, California 94143, USA

⁷ Division of Infectious Diseases, University of California, San Francisco, San Francisco, California 94143, USA and Department of Surgery, Division of Cardiothoracic Surgery, University of California, San Francisco, San Francisco, California 94143, USA

⁸ Cardiovascular Research Institute, Departments of Medicine and Anesthesia, University of California, San Francisco, San Francisco, California 94143, USA

⁹ Gladstone Institute of Virology, Department of Medicine, University of California San Francisco, California 94143, USA

¹⁰ Department of Medicine, Division of Rheumatology, University of California, San Francisco, San Francisco, California 94143, USA

¹² Chan Zuckerberg Biohub, San Francisco, California 94158, USA

Correspondence: Jeroen Roose, Department of Anatomy, University of California San Francisco (UCSF) 513 Parnassus Avenue, Room HSW-1326, San Francisco, California 94143-0452, USA; e-mail: jeroen.roose@ucsf.edu.

Abstract

SARS coronavirus-2 (SARS-CoV-2) is causing a global pandemic with large variation in COVID-19 disease spectrum. SARS-CoV-2 infection requires host receptor ACE2 on lung epithelium, but epithelial underpinnings of variation are largely unknown. We capitalized on comprehensive organoid assays to report remarkable variation in SARS-CoV-2 infection rates of lung organoids from different subjects. Tropism is highest for TUBA- and MUC5AC-positive organoid cells, but levels of TUBA-, MUC5A-, or ACE2- positive cells do not predict infection rate. We identify surface molecule Tetraspanin 8 (TSPAN8) as novel mediator of SARS-CoV-2 infection, which is not downregulated by this specific virus. TSPAN8 levels, prior to infection, strongly correlate with infection rate and TSPAN8-blocking antibodies diminish SARS-CoV-2 infection. We propose TSPAN8 as novel functional biomarker and potential therapeutic target for COVID-19.

Introduction

Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) has caused a global pandemic of coronavirus disease (COVID-19) with over 170M cases worldwide (<https://coronavirus.jhu.edu/map.html>) as of May 2021. Clinically, SARS-CoV-2 can lead to lethal acute respiratory distress syndrome (ARDS) ¹. Many SARS-CoV-2 studies have focus on the immune system, currently, we do not understand how the lung epithelium itself may play a role in variation in disease spectrum of COVID-19.

The lung airway epithelium defends against pollutants, allergens, and pathogens and is composed of a variety of cell types including basal cells, ciliated cells, mucus-producing goblet cells, secretory cells, and neuroendocrine cells. The distal lung epithelium also includes alveolar type 1 (AT1) cells, which mediate gas exchange, and AT2 cells, which secrete surfactant ^{2, 3}. In the case of COVID-19, studies thus far suggest that SARS-CoV-2 infects mostly ciliated cells, goblet cells, AT2 cells but also basal stem cells ^{4, 5, 6, 7, 8}. Comprehensive SARS-CoV-2 infection studies focusing on lung epithelial cell subsets in different individuals have not been reported to date.

Coronavirus (CoV) particles are spherical with three viral proteins anchored in the envelope: the triple-spanning membrane (M) protein, the envelope (E) protein, and the spike (S) protein, which form the characteristic trimeric spikes ^{9, 10, 11}. The surface spike glycoprotein of SARS-CoV-2 binds to human ACE2 (Angiotensin-converting enzyme 2) ¹². Binding of the S protein to ACE2 mediates membrane fusion and viral entry. The S protein is cleaved by host cell type II trans-membrane serine proteases resulting in spike protein activation and viral entry ^{13, 14, 15, 16}. As such, ACE2 and Transmembrane protease, serine 2 (TMPRSS2) are critical for SARS-CoV-2 entry into the cell ¹⁷.

Viruses typically co-opt various host proteins to maximize infectious potential ¹⁸. The wide variation in SARS-CoV-2 infection rates and COVID-19 severity suggests that there must be facilitators other than ACE2 and TMPRSS-2 that have yet to be discovered. Epithelial organoids have gained traction as a physiological platform for personal medicine because these organoids retain patient-specific traits¹⁹. A handful of studies have used alveolospheres and organoids to study SARS-CoV-2 epithelial response ^{20, 21, 22, 23, 24, 25} and tropism ^{7, 26, 27, 28, 29, 30}. A comprehensive analysis of

SARS-CoV-2 infection in a large panel of lung organoids has not been reported to date. Here, we performed a comprehensive analysis with the goal of obtaining insights for different subjects as well as discovery of new molecules that play a role in infection.

Results

A comprehensive lung organoid biobank

We generated organoids from biopsies from healthy donors or adjacent normal tissue from lung cancer patients undergoing surgery (**Table 1**). 3D lung organoids from twenty subjects were expanded through passaging and cryopreserved, generating a biobank (**Figs. 1A and 1B, and Suppl. Fig S1A and S1B**).

Spectral flow cytometry analyses with single tube staining of fourteen markers (termed spectral flow here), including a panel of recently developed airway epithelial markers³¹ revealed diversified compositions of cell subsets in fifteen lung organoids. We identified ciliated *like* cells (TUBA^{high} CD271^{neg})³², goblet *like* cells (TUBA^{neg}, MUC5AC⁺)³³, pre-goblet *like* cells (TUBA^{neg}, MUC5AC⁻, TSPAN8⁺)³⁴, three populations of cells expressing basal cell markers CD49⁺CD271⁺, CD49^{f^{neg}}CD271⁺, CD49⁺CD271⁺, and a population of CD49^{f^{neg}}CD271^{neg}TUBA^{neg}MUC5AC^{neg}TSPAN8^{neg} cells (**Figs.1C and 1D**). Lung organoids derived from different donors displayed distinctive cell type compositions (**Fig. 1E**). By contrast, compositions of organoids generated from upper lobe and lower lobe samples from the same patient were very similar (**Suppl. Fig S1B**) and analysis of different passages from distinct organoids demonstrated that our lung organoids are stable and retain patient-specific composition in passaging (**Suppl. Fig S1C**). Spectral Flow Cytometry staining for intra-cellular TMPRSS2 (**Fig. 1F**) and extracellular ACE-2 (**Fig. 1G**) allowed for the assessment of the percent of cells expressing these proteins that play critical roles in SARS-CoV-2 entry. Pearson correlations between age and the proportions of goblet *like*-, ciliated *like*-, and basal-cells in unmanipulated organoids were not significant (**Suppl. Fig S1D-S1F**). We selected a panel of twelve lung organoids that captured the variety in cell composition, age, sex, and ACE-2 and TMPRSS2 expression for further studies.

Viral infections of lung organoids

We utilized H1N1/PR8 influenza virus to optimize viral infection of lung organoids (**Suppl. Fig. S3A-D**). Following infection, we assessed cell-composition changes with spectral flow, adding the cKit marker (**Suppl. Fig S2 and S3F**), as this receptor has been suggested to mark lung regeneration upon injury^{35,36} and tropism of cell types for H1N1 (**Suppl. Fig S3G**). Since viral infections trigger epithelial interferon responses, which help orchestrate immune responses³⁷, we also stained for co-stimulatory molecules CD80 and CD86^{2,38}, or immune-activating molecules CEACAM5 and CEACAM6^{39,40}. ACE2 has recently been described as an interferon-upregulated gene⁴¹. We observed upregulation of these five cell surface molecules after H1N1/PR8 infection (**Suppl. Fig. S3H**).

Having established consistent infection with H1N1/PR8, we switched to SARS-CoV-2 (**Fig 2A**). Whole mount organoids revealed the presence of double stranded RNA (dsRNA) and viral nucleocapsid protein (N) in the infected lung organoids, that were also stained for CD49f and acTUBA (**Fig. 2B**). Five independent SARS-CoV-2 infections of lung organoid 2522UL demonstrated that our infections were robust and consistent (**Suppl. Fig S4A-D**). Spectral flow of dsRNA staining, marking replicating virus, revealed a remarkable variation in the percentage of (replicating) SARS-CoV-2-positive cells in twelve different organoids and 31 SARS-CoV-2 infections with means ranging from fifteen to zero percent (**Figs. 2C and 2D**). The percentage of cell death in the organoids (**Suppl. Fig S5A**) and infection rates (**Fig. 2D**) showed different patterns. We characterized cell compositions (**Fig. 2E and Suppl. Fig S5B**) and plotted changes using PCA (principal component analysis) (**Suppl. Fig S5C**). SARS-CoV-2 infection resulted in increased proportions of cells expressing acTUBA (**Fig. 2F**) and cKit (**Fig. 2G**), but no alterations in the fraction of basal cells (CD49+ or CD271+), MUC5AC-positive, or CD49^{fneg}CD271^{neg} cells (**Fig. 2H ,Suppl. Fig S5B and S5D**).

In terms of functional markers, CD86-positive cells were increased (**Fig. 2I**), but SARS-CoV-2 infection did not cause increases in proportions for CD80-, CEACAM5-, and CEACAM6- expressing cells (**Suppl. Fig S5E**), indicating that there was not a uniform induction of an interferon response program in all cells of infected organoids. Lungs from COVID-19 patients demonstrate elevated ACE-2 protein expression⁴². We observed significant increases in the proportions of ACE2-positive cells (**Fig. 2J**)

upon SARS-CoV-2 infection in the 12 lung organoids. The spectral flow allows us to compare mock-infected to SARS-CoV-2 exposed/infected (dsRNA+), and to SARS-CoV-2 exposed/uninfected (dsRNA-). These comparisons revealed that nearly one hundred per cent of dsRNA-positive, infected cells were ACE2-positive (**Fig. 2K**). Similarly, we observed high proportions of cells positive for CD80, CD86, CEACAM5, and CEACAM6 when analyzing exposed/infected specifically (**Suppl. Fig S4E**).

Finding the rules of variation in SARS-CoV-2 infection

The lung organoids from subject 2525 express low TMPRSS2 (**Fig. 1G**), providing an explanation for the low infection rate in 2525 (**Fig. 2D**). For the other organoids an explanation for the variation was lacking. Our panel of twelve organoids and 31 SARS-CoV-2 infections allowed us to explore if there are rules that help explain the infection variation that ranged from thirteen to zero percent on average. We observed a tropism of SARS-CoV-2 for acTUBA- and MUC5AC- positive cells (**Fig. 3A**). Depicting PCA of many factors in a circle of correlation, we established that there are probable correlations between amount ciliated cells (acTUBA positivity), ACE2 positivity, and CD86 positivity, determined in non-infected organoids and subsequent infection rate. There is also a probable correlation between the age of the donor and infection, since the arrows are of substantial length and point in the same direction (**Fig. 3B**). However, statistical analysis of age (**Fig. 3C**, Pearson 0.45, $P_v = 0.14$), TUBA-positivity (**Fig. 3D**, Pearson 0.48, $P_v = 0.11$), and CD86-, cKit-, or MUC5AC- positivity (**Suppl. Fig S5F**) as single factors revealed that none correlate with SARS-CoV-2 infection rate in organoids.

The percentage of ACE2 positive cells was significantly upregulated upon infection of lung organoids (**Fig. 2J**) and nearly all dsRNA-positive cells are also ACE2-positive (**Fig. 2K**). However, the proportion of ACE2-positive cells in the organoids prior to infection, either as mean infection or as single datapoints of all 31 infections, did not correlate with infection rate (**Fig. 3E and 3F**). These findings motivated us to search for novel molecules that show strong correlations with rate of infection and may function as host proteins co-opted by SARS-CoV-2 in infecting the lung.

Single cell gene expression points to TSPAN8

In search of new mediators of SARS-CoV-2 infection, we infected four organoids with SARS-CoV-2 and performed single-cell RNA sequencing (sc-RNAseq) of organoids. Unsupervised clustering analysis based on most variable gene expression across all cells ⁴³, regardless of infection status, identified sixteen unique cell subsets represented in a UMAP plot (**Fig. 4A**). Based on the top five most differentially expressed genes by cluster we assigned relative identities to these sixteen populations (**Suppl. Fig S6**), though it should be noted that assignment of identity of single cells of different lung organoids was not our objective here. Furthermore, while clustering analysis of individual organoids showed significant differences (**Suppl. Fig S7A**), read counts were relatively low for organoids L7 and 2524. For this reason, we treated the four organoids as one collective dataset, while capitalizing on the resolution of the sc-RNAseq for discovery.

We next focused on cells expressing SARS-CoV-2 transcript to identify specific transcriptomic signatures in infected cells (**Table 2**). Metascape pathway analysis ⁴⁴ confirmed a signature of inflammatory response, in agreement with the identity of virus-infected epithelial cells (**Suppl. Fig S7C**). We plotted a cluster of sixty genes that were differentially expressed in single cells positive for SARS-CoV-2 viral read compared to false identities (**Suppl. Fig. S7B** and **Table 3A**). We mined our dataset in search of possible novel receptors for SARS-CoV-2 and noted that the surface molecules TSPAN8, AREG, and CD24 (**Suppl. Fig. S7B** and **Table 3B**) were positively expressed in single cells containing SARS-CoV-2 reads. We were particularly intrigued by TSPAN8 (Tetraspanin 8), since members of the TSPAN family are believed to promote cell entry of different viruses ⁴⁵, but TSPANs have not been linked to SARS-CoV-2. Single cell *TSPAN8* mRNA reads were present in 64% of cells with SARS-CoV-2 reads (**Suppl. Fig. S7B**) and *TSPAN8* expression was not limited to a unique cell subset (**Fig. 4B**).

TSPAN8 facilitates SARS-CoV-2 infection and can be inhibited

TSPAN8 mRNA expression had been documented in the supplemental dataset of a single SARS-CoV-2-infected lung organoid⁸, but the role of *TSPAN8* in SARS-CoV-2 infection has not been investigated. Fortuitously, we had included anti-*TSPAN8* antibody in our spectral flow to distinguish goblet like cells from pre-goblet cells (**Suppl. Fig. S2**), allowing for surface protein expression to confirm the sc-RNAseq discovery. The proportion *TSPAN8*-positive cells in the organoids increased upon SARS-CoV-2 infection (**Fig. 4C**), but decreased with H1N1 infections (**Fig. 4D**). Just as for ACE2 (**Fig 2K**), most infected cells (dsRNA positive) expressed *TSPAN8* (**Fig. 4E**), a pattern not observed for H1N1 (**Fig. 4F**). We further investigated *TSPAN8* in the context of viral infections. Nasal swabs of adult patients with acute respiratory illness (ARI)⁴⁶ due to non-COVID-19, respiratory viral infection revealed downregulation of *TSPAN8* (**Fig. 4G**). However, such airway brushes from COVID patients revealed that *TSPAN8* expression remained of the same level compared to comparison to patients with ARI due to non-viral etiology **Fig. 4G**.

Mining published scRNAseq data of human lungs³ that predominantly reported on immune cell composition, we noted that of 61,000 cells sequenced from lung tissue (as opposed to blood), only 48 cells had expression of both *TSPAN8* and ACE2 and all of these 48 cells were of epithelial origin (**Fig. 4H**). Furthermore, back-gating on dsRNA-positive cells and overlaying this onto a ACE2/*TSPAN8* contour plot, our spectral flow revealed that a large proportion of SARS-CoV-2 infected cells simultaneously expresses both ACE2 and *TSPAN8* (**Fig. 4J** and **Suppl. Fig S7F**). Collectively, these data suggest that the two surface proteins may cooperate in SARS-CoV-2 infection (**Fig. 4I**).

Given the *TSPAN8* data and the report that *TSPAN* family member CD9 facilitates MERS-CoV infection in mice⁴⁷, we considered whether *TSPAN8* is a facilitator of SARS-CoV-2 infection (**Fig. 4I**). First, we explored *TSPAN8* as biomarker. Different from our studies on ACE2 expression (**Figs. 3E** and **3F**), high percentages of *TSPAN8*-positive cells prior to infection did correlate very strongly with subsequent SARS-CoV-2 infection rates in the 31 infections (**Fig. 4K**, **Suppl. Fig S7G**). Lastly, we explored *TSPAN8* as potential therapeutic target. We chose organoid 2450UL that

displays high baseline TSPAN8 expression (**Suppl. Fig. S7F**), infected with SARS-CoV-2, and subjected these to blocking antibody treatment. Addition of a blocking antibody that binds to one of TSPAN8's extracellular loops reduced SARS-CoV-2 infection by 60 per cent, roughly in the same range of inhibition of an ACE2-blocking antibody (**Figs. 4L and 4M, Supplemental Fig. S7H and I**), which had been reported previously ¹⁴.

Discussion

Here we characterized the composition of 15 lung organoids derived from different donors. Multi-parameter spectral flow of cell sub-sets shows that the organoids are stable, but unique from subject to subject, generating a diverse lung organoid biobank. Subject-specific traits can be investigated, such as the pattern of decreasing amounts of mucus-producing cells as the age of the donor subject increases (**Suppl. Fig S1F**). H1N1/PR8 influenza and SARS-CoV-2 infection and spectral flow allowed for discovery of cell responses. Fifteen antibody spectral flow generates gating strategies where cell populations are embedded in each other, providing quantitative, linked data that can be analyzed in many ways. We were intrigued by the increase in cKit-positive cells after infection, as cKit has been suggested as marker of lung regeneration upon injury^{35, 36}. Possibly infected organoids sense damage and attempt to repair injury. Spectral flow-assisted gating on SARS-CoV-2-infected cells revealed a clear tropism for ciliated and MUC5AC-positive cells, but other cell types could also get infected. It should be noted here that our lung organoids are all embedded in Matrigel as we have not yet been able to obtain high throughput for organoid with air-liquid interface⁴⁸.

Neutralizing antibodies from COVID-19 patients have multiple targets^{49, 50}, suggesting that protective immune responses occur that may block interactions of molecules other than the S protein-ACE2 receptor pair. In our scRNA data *RALA* and *CD24* reads were present in infected cells. *RALA* (RAS like Proto-Oncogene A) was identified in the host-coronavirus protein network⁵¹ and infection levels of a hepatoma cell line is reduced when deleted by CRISPR¹⁷. *CD24* is a glycosyl-phosphatidyl-inositol (GPI)-anchored membrane protein that can repress the host response to DAMPS⁵² and EXO-*CD24* exosomes are explored in clinical trials as therapy for COVID19 patients (Clinical Trial NCT047477574). Studies using cell lines had also implicated additional mediators of infection, including *AXL*⁵³ *CD147*^{54, 55} and neuropilin-1 (*NRP1*)^{56, 57}.

We discovered that *TSPAN8* strongly correlates with infected cells and infection rate. Tetraspanin proteins have four transmembrane domains (**Fig. 4I**) that can form lateral associations with multiple molecular partners and with each other, organizing the surface membrane proteins in dynamic microdomains⁴⁵. Tetraspanins promote the entry of multiple viruses, including influenza A virus (IAV), human cytomegalovirus

(HCMV), human papillomavirus, hepatitis C virus, Lujo virus, and several alphaviruses⁴⁵. Depleting the tetraspanin CD9 reduced MERS-CoV lung titers by ~90% in the infected mice⁴⁷. Our analyses of ACE2 and TSPAN8 protein expression in the context of SARS-CoV-2 infection of a large panel of lung organoids strongly suggested that TSPAN8 may play a role in SARS-CoV-2 infections. SARS-CoV-2 infection does not lead to TSPAN8 downregulation, whereas other viral infections do. Bioinformatic Analysis using published data performed by the GeneMania⁵⁸ speculates that ACE2 interacts with TSPAN8⁵⁹ and data from Traveglini *et al.*³ showed that epithelial cells are the only cells in the lung that can express both TSPAN8 and ACE2. Alternatively, TSPAN8 may cooperate in the lung organoid infections with other mediators, such as neuropilin-1 (NRP1) described in cell line studies^{56, 57}. Future studies are required to understand the mechanistic details of TSPAN8 and ACE2-mediated SARS-CoV-2 infection.

Clonal cell lines are powerful discovery tools but lack variations in genetic and proteomic traits. Lung epithelial organoids are an attractive platform to study responses to airway infection in the context of diverse cell types⁴⁸, but every organoid from a human subject is distinct. Our study with a comprehensive panel of lung organoids capitalized on the unique diversity in lung organoids derived from different subjects to discover TSPAN8 as novel biomarker. Furthermore, blocking TSPAN8 showed a reduction of SARS-CoV-2 viral load and TSPAN8 may be a novel potential therapeutic target in SARS-CoV-2 infection.

Acknowledgements

We thank Drs. Melia Magnen, Leonardo Ferreira, Vilma Arce-Gorvel, Jean-Pierre Gorvel, Jean Louis Mege and the entire Roose lab for stimulating discussions. We thank Drs. Hans Clevers, Rob Vries, Bahar Ramezanzpour, and Sylvia Boj for helpful discussion on organoid technology. We thank the Aurora Cytex team for technical support on the spectral flow and Arjun Rao for help with scRNAseq analysis.

We thank Garcia lab at Standford for the R-spondin surrogates. Most of the work was supported by an Administrative COVID-19 Supplement 3P01AI091580-09S1 (to JPR) on the parent NIH/NIAID P01-AI091580 (Weiss). Organoids were generated in the Roose Organoid D2B unit, started through a UCSF PBBR TMC (Technologies, methodologies, and Cores) grant in 2018 and a gift from the Pathology department and now, in part, funded through a Mark Foundation for Cancer Research Endeavor Program grant (all to JPR). KB is a Mark Foundation Momentum fellow. JRK is funded by the UCSF Bakar ImmunoX Initiative, the UCSF Helen Diller Family Cancer Center, and the American Association for Thoracic Surgery Foundation. DMJ is funded by a private endowment fund. Human Frontier Science Program Fellowship LT000061/2018-L to NKS Additional funds came a grant from the Innovative Genomics Institute and UCSF PBBR (M.O.).

Disclosure statement

The authors have no potential conflicts of interest.

Author contributions

LH and JPR: conceived and designed study. LH, SL and KK: performed spectral flow and analyses. LH, KK, SL, OMG: performed organoid experiments. LH and LR: BSL3 work. AS and AJC: scRNAseq experiments and advice. JCL: statistical analyses. LCL: computational analysis of scRNAseq data. LB and DE: advice on antibody staining and lung populations. NKS and MK: H1N1 virus. LH, OMG and KB: microscopy. JZL, VD, SM, MM: patient samples for organoids. GK, DMJ, MM, AJC: (funding) support of their team members. MO: SARS-CoV-2 virus. JRK: surgery lung samples, clinical data discussion. LH, SL, LCL, JPR: figure panels. LH, SL, JPR: manuscript writing. JPR: funding for the study. F, DMJ, MM, MO, MM, AJC, DE, ANS, and JRK: edits on draft manuscript.

References

1. Zhu, N., et al., *A novel coronavirus from patients with pneumonia in China, 2019*. New England journal of medicine, 2020.
2. Zhou, P., et al., *A pneumonia outbreak associated with a new coronavirus of probable bat origin*. nature, 2020. **579**(7798): p. 270-273.
3. Corbiere, V., et al., *Phenotypic characteristics of human type II alveolar epithelial cells suitable for antigen presentation to T lymphocytes*. Respir Res, 2011. **12**: p. 15.
4. Travaglini, K.J., et al., *A molecular cell atlas of the human lung from single-cell RNA sequencing*. Nature, 2020. **587**(7835): p. 619-625.
5. Mason, R.J., *Pathogenesis of COVID-19 from a cell biology perspective*. Eur Respir J, 2020. **55**(4).
6. Chua, R.L., et al., *COVID-19 severity correlates with airway epithelium-immune cell interactions identified by single-cell analysis*. Nat Biotechnol, 2020. **38**(8): p. 970-979.
7. Ravindra, N.G., et al., *Single-cell longitudinal analysis of SARS-CoV-2 infection in human bronchial epithelial cells*. bioRxiv, 2020.
8. Fiege, J.K., et al., *Single cell resolution of SARS-CoV-2 tropism, antiviral responses, and susceptibility to therapies in primary human airway epithelium*. PLoS Pathog, 2021. **17**(1): p. e1009292.
9. Lamers, M.M., et al., *SARS-CoV-2 productively infects human gut enterocytes*. Science, 2020.
10. Li, F., et al., *Structure of SARS coronavirus spike receptor-binding domain complexed with receptor*. Science, 2005. **309**(5742): p. 1864-1868.
11. Walls, A.C., et al., *Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein*. Cell, 2020. **181**(2): p. 281-292 e6.
12. Walls, A.C., et al., *Cryo-electron microscopy structure of a coronavirus spike glycoprotein trimer*. Nature, 2016. **531**(7592): p. 114-117.
13. Yan, R., et al., *Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2*. Science, 2020. **367**(6485): p. 1444-1448.
14. Hoffmann, M., H. Kleine-Weber, and S. Pöhlmann, *A multibasic cleavage site in the spike protein of SARS-CoV-2 is essential for infection of human lung cells*. Molecular Cell, 2020.
15. Hoffmann, M., et al., *SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor*. Cell, 2020. **181**(2): p. 271-280 e8.
16. Firas A. Rabi 1, Mazhar S. Al Zoubi 2, et al., *SARS-CoV-2 and Coronavirus Disease 2019: What We Know So Far*. <https://doi.org/10.3390/pathogens9030231>.
17. Kim, I.S., et al., *Mechanism of membrane fusion induced by vesicular stomatitis virus G protein*. Proc Natl Acad Sci U S A, 2017. **114**(1): p. E28-E36.
18. Wang, R., et al., *Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses*. Cell, 2021. **184**(1): p. 106-119. e14.
19. Thorley, J.A., J.A. McKeating, and J.Z. Rappoport, *Mechanisms of viral entry: sneaking in the front door*. Protoplasma, 2010. **244**(1): p. 15-24.

20. Kim, M., et al., *Patient-derived lung cancer organoids as in vitro cancer models for therapeutic screening*. Nat Commun, 2019. **10**(1): p. 3991.
21. Katsura, H., et al., *Human Lung Stem Cell-Based Alveolospheres Provide Insights into SARS-CoV-2-Mediated Interferon Responses and Pneumocyte Dysfunction*. Cell Stem Cell, 2020. **27**(6): p. 890-904 e8.
22. Youk, J., et al., *Three-Dimensional Human Alveolar Stem Cell Culture Models Reveal Infection Response to SARS-CoV-2*. Cell Stem Cell, 2020. **27**(6): p. 905-919 e10.
23. Lamers, M.M., et al., *SARS-CoV-2 productively infects human gut enterocytes*. Science, 2020. **369**(6499): p. 50-54.
24. Lamers, M.M., et al., *Human airway cells prevent SARS-CoV-2 multibasic cleavage site cell culture adaptation*. bioRxiv, 2021.
25. Salahudeen, A.A., et al., *Progenitor identification and SARS-CoV-2 infection in long-term human distal lung organoid cultures*. bioRxiv, 2020.
26. Yang, L., et al., *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, 2020. **27**(1): p. 125-136. e7.
27. Zhang, B.Z., et al., *SARS-CoV-2 infects human neural progenitor cells and brain organoids*. Cell Res, 2020. **30**(10): p. 928-931.
28. Clevers, H., *COVID-19: organoids go viral*. Nat Rev Mol Cell Biol, 2020. **21**(7): p. 355-356.
29. Bonser, L.R., et al., *Flow-Cytometric Analysis and Purification of Airway Epithelial-Cell Subsets*. American Journal of Respiratory Cell and Molecular Biology, 2021. **64**(3): p. 308-317.
30. Griggs, T.F., et al., *Rhinovirus C targets ciliated airway epithelial cells*. Respiratory research, 2017. **18**(1): p. 1-11.
31. Gray, T., et al., *Regulation of MUC5AC mucin secretion and airway surface liquid metabolism by IL-1 β in human bronchial epithelia*. American Journal of Physiology-Lung Cellular and Molecular Physiology, 2004. **286**(2): p. L320-L330.
32. Duclos, G.E., et al., *Characterizing smoking-induced transcriptional heterogeneity in the human bronchial epithelium at single-cell resolution*. Science advances, 2019. **5**(12): p. eaaw3413.
33. Fang, S., et al., *Generation of functional blood vessels from a single c-kit⁺ adult vascular endothelial stem cell*. PLoS Biol, 2012. **10**(10): p. e1001407.
34. Lopez-Giraldo, A., et al., *Characterization, localization and comparison of c-Kit⁺ lung cells in never smokers and smokers with and without COPD*. BMC Pulm Med, 2018. **18**(1): p. 123.
35. Schleimer, R.P., et al., *Epithelium: at the interface of innate and adaptive immune responses*. J Allergy Clin Immunol, 2007. **120**(6): p. 1279-84.
36. Kaneko, Y., et al., *B7-1, B7-2 and class II MHC molecules in idiopathic pulmonary fibrosis and bronchiolitis obliterans-organizing pneumonia*. European Respiratory Journal, 2000. **15**(1): p. 49-55.
37. Lambrecht, B.N. and H. Hammad, *The role of dendritic and epithelial cells as master regulators of allergic airway inflammation*. The Lancet, 2010. **376**(9743): p. 835-843.
38. Zissel, G., et al., *Human alveolar epithelial cells type II are capable of regulating T-cell activity*. Journal of investigative medicine: the official publication of the American Federation for Clinical Research, 2000. **48**(1): p. 66-75.

39. Ziegler, C.G.K., et al., *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*, 2020. **181**(5): p. 1016-1035 e19.
40. Ackermann, M., et al., *Pulmonary vascular endothelialitis, thrombosis, and angiogenesis in Covid-19*. *New England Journal of Medicine*, 2020. **383**(2): p. 120-128.
41. Becht, E., et al., *Dimensionality reduction for visualizing single-cell data using UMAP*. *Nat Biotechnol*, 2018.
42. Zhou, Y., et al., *Metascape provides a biologist-oriented resource for the analysis of systems-level datasets*. *Nature communications*, 2019. **10**(1): p. 1-10.
43. Hantak, M.P., et al., *Tetraspanins: architects of viral entry and exit platforms*. *Journal of virology*, 2019. **93**(6).
44. Earnest, J.T., et al., *The tetraspanin CD9 facilitates MERS-coronavirus entry by scaffolding host cell receptors and proteases*. *PLoS pathogens*, 2017. **13**(7): p. e1006546.
45. Sachs, N., et al., *Long-term expanding human airway organoids for disease modeling*. *EMBO J*, 2019. **38**(4).
46. Chi, X., et al., *A neutralizing human antibody binds to the N-terminal domain of the Spike protein of SARS-CoV-2*. *Science*, 2020. **369**(6504): p. 650-655.
47. Brouwer, P.J., et al., *Potent neutralizing antibodies from COVID-19 patients define multiple targets of vulnerability*. *Science*, 2020. **369**(6504): p. 643-650.
48. Gordon, D.E., et al., *A SARS-CoV-2 protein interaction map reveals targets for drug repurposing*. *Nature*, 2020. **583**(7816): p. 459-468.
49. Chen, G.-Y., et al., *CD24 and Siglec-10 selectively repress tissue damage-induced immune responses*. *Science*, 2009. **323**(5922): p. 1722-1725.
50. Wang, S., et al., *AXL is a candidate receptor for SARS-CoV-2 that promotes infection of pulmonary and bronchial epithelial cells*. *Cell Res*, 2021.
51. Wang, K., et al., *SARS-CoV-2 invades host cells via a novel route: CD147-spike protein*. *BioRxiv*, 2020.
52. Shilts, J., et al., *No evidence for basigin/CD147 as a direct SARS-CoV-2 spike binding receptor*. *Scientific reports*, 2021. **11**(1): p. 1-10.
53. Daly, J.L., et al., *Neuropilin-1 is a host factor for SARS-CoV-2 infection*. *Science*, 2020. **370**(6518): p. 861-865.
54. Cantuti-Castelvetri, L., et al., *Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity*. *Science*, 2020. **370**(6518): p. 856-860.
55. Warde-Farley, D., et al., *The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function*. *Nucleic acids research*, 2010. **38**(suppl_2): p. W214-W220.
56. Samad, A., T. Jafar, and J.H. Rafi, *Identification of angiotensin-converting enzyme 2 (ACE2) protein as the potential biomarker in SARS-CoV-2 infection-related lung cancer using computational analyses*. *Genomics*, 2020. **112**(6): p. 4912-4923.

Figure Legends

Figure 1: Donor-derived lung organoids are stable and distinctive.

A) Workflow of lung organoid generation.

B) Brightfield images of lung organoids derived from different donors.

C) Spectral Flow Cytometry gating strategy to define seven cell populations: (1) TUBA+ CD271- as ciliated cells. (2) MUC5AC+ TUBA- as goblet-like cells. (3) TSPAN8+ MUC5AC- TUBA- as pre-goblet cells. (4) CD49f+ CD271+, (5) CD49f- CD271+, (6) CD49f+ CD271+ as basal cells and (7) CD49f- CD271- as undefined cells.

D) Scheme of cell types found in the 3D lung organoids.

E) Pie charts representing spectral flow cytometry analyses of lung organoids from indicated donors.

F-G) Percentage of F) ACE2+ and G) TMPRSS2+ cells in lung organoids based on spectral flow cytometry analyses. Bars represent mean, error bars are SEM, n=3. Paired t-test, *p<0.05; ns, non-significant.

Figure 2: Patient derived organoids show distinct responses to SARS-CoV-2 infection.

A) Experimental scheme SARS-CoV-2 (MOI=0.3) infection.

B) Representative, confocal images (Z-stack) of SARS-CoV-2 infected (top and middle panel) and mock infected (bottom panel) whole mounted organoids. Scale bars are 50uM.

C) Spectral flow for dsRNA, marking cells with replicating SARS-CoV-2.

D) Percentage of dsRNA-positive cells in SARS-CoV-2 infected lung organoids

E) Composition of Mock and SARS-CoV-2 infected organoids.

F-H) Comprehensive analysis of % of cell populations (F) Ciliated like cells (TUBA+), (G) cKit+, (H) Basal cells) in Mock and SARS-CoV-2 infected organoids for 12 different donor derived lung organoids. Each point represents the mean of % cell type for the distinct donor for Mock or SARS-CoV-2 conditions. Wilcoxon signed-rank paired test, * $p < 0.05$, $N=12$, $n=3$.

I-J) Percentages of ACE2+ and CD86+ cells. Each point represents the mean of % cell type for the distinct donor for Mock or SARS-CoV-2 conditions. Wilcoxon signed-rank paired test, * $p < 0.05$, $N=12$, $n=3$.

K) The percentage of ACE2+ cells in live cells of non infected organoids (Mock), and the percentage of ACE2 positive cells in live dsRNA- and in dsRNA+ live cells of SARS-CoV-2 infected organoids at 72h p.i.. Each point represents the mean of % cell type for the distinct donor for Mock or SARS-CoV-2 conditions. Friedman Test, ** $p < 0.01$; *** $p < 0.001$. ns, non-significant.

Figure 3: Susceptibility of lung organoids to SARS-CoV-2 infection is not predicted by ACE2.

- A) Percentage of dsRNA+ cells (left) and the cell types in dsRNA+ cells (right) in SARS-CoV-2 infected organoids (MOI=0.3, 72h p.i.). Mean values of 31 infections are depicted and plotted.
- B) PCA of different variables impacting infection. Each arrow corresponds to one biological descriptor; the longer the arrow, the better the representation (the color displays the \cos^2).
- C) Linear regression modeling the relationship between age of the donor and infection rate.
- D) Linear regression modeling the relationship between percentage of TUBA+ CD271+ cells and infection rate. Each point represents the mean of % cell type for the distinct donor.
- E) Linear regression modeling the relationship between Infection rate and % ACE2+ cells prior to infection. Each point represents the mean of % cell type for the distinct donor.
- F) Infection rate and % ACE2+ cells prior to infection for 31 single infections. Statistical significance is displayed by P values of Pearson Correlation for each graph. Mean values, N=12, n=3.

Figure 4: TSPAN8 is a facilitator, risk factor, and target for SARS-CoV-2.

A) UMAP reduction on the merged cell data with overlaid clusters and identified cell types.

B) TSPAN8 expression overlay on the UMAP seen in (A).

C-D) Percentage of TSPAN8 positive cells for 12 SARS-CoV-2 (MOI=0.3) infected or 3 H1N1(MOI=0.15) infected lung organoids for Mock and infected organoids at 72h p.i. . Every point represents a single experiment. Each experiment was performed independently at least 2 times using three replicates per experiment. Friedman Test, * $p < 0.05$; **** $p < 0.0001$.

E-F) The percentage of TSPAN8+ cells in live cells of non infected organoids (Mock), and the percentage of ACE2 positive cells in (E,F) live dsRNA- and in dsRNA+ live cells of SARS-CoV-2 infected organoids or mCherry+ and mCherry- live cells of H1N1 infected organoids at 72h p.i.. Each point represents the mean of % cell type for the distinct donor for Mock or viral infected conditions. Friedman Test, ** $p < 0.01$; *** $p < 0.001$. ns, non-significant.

H) Analyses of ACE2 and TSPAN8 co-expressing cells in vivo in normal human lung, from 10X single-cell sequencing data from Travaglini et al.

I) SARS-CoV-2 virus, ACE2, and TSPAN8 in a representation of lung epithelial cell

G) Differential expression of TSPAN8 in nasal swabs of adult patients with acute respiratory illness (ARI) due to COVID-19 (n=93) or other viral infection (n=41), in comparison to patients with ARI due to non-viral etiology (n=100)

J) Spectral flow cytometry plots showing live dsRNA- cells (in gray) and live dsRNA+ cells (overlaid in red) in four different SARS-CoV-2 infected organoids at 72h p.i.

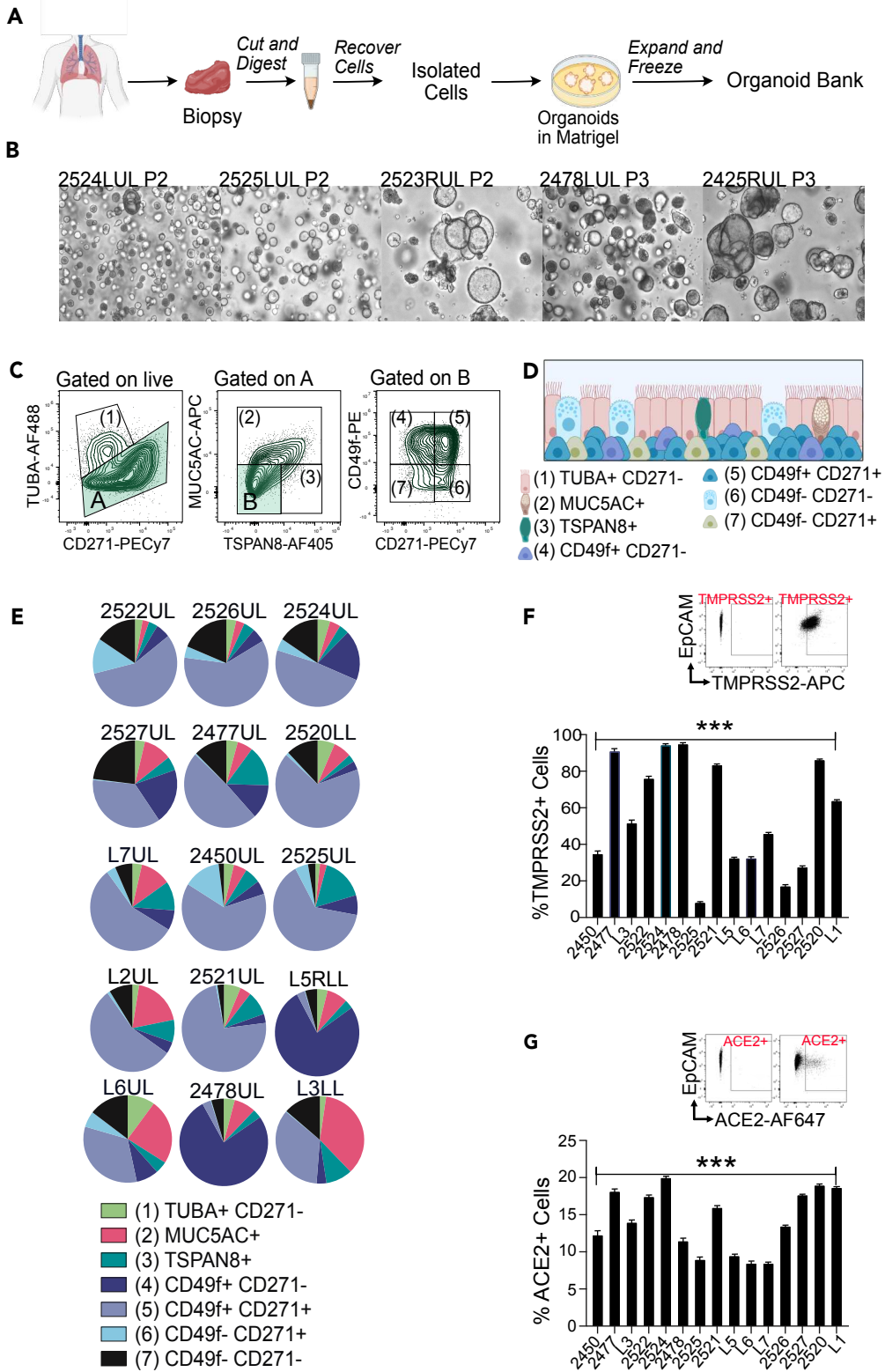
K) Correlation between percentage of TSPAN8 positive cells prior to infection and SARS-CoV-2 infection rate. 31 individual infections. P values indicate statistical significance.

L) Spectral Flow Cytometry Plots of dsRNA-positive cells in Mock, SARS-CoV-2 infected organoids in presence or absence of TSPAN8 blocking antibody.

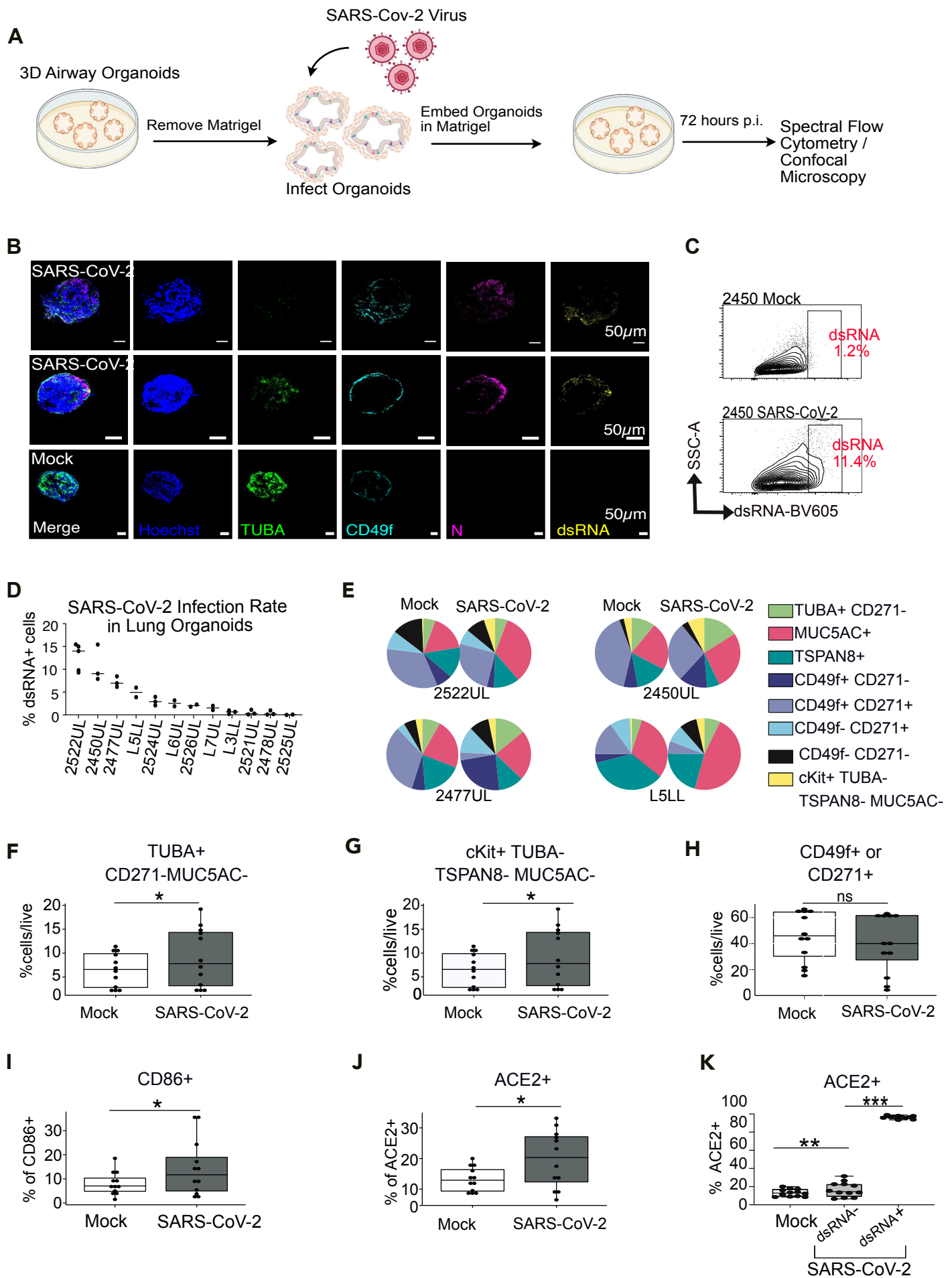
M) Percentage of dsRNA+ cells in Mock, SARS-CoV-2 infected organoids in presence or absence of TSPAN8 or ACE2 blocking antibody. SARS-CoV-2 infection in the indicated blocking conditions (N=1, n=3).

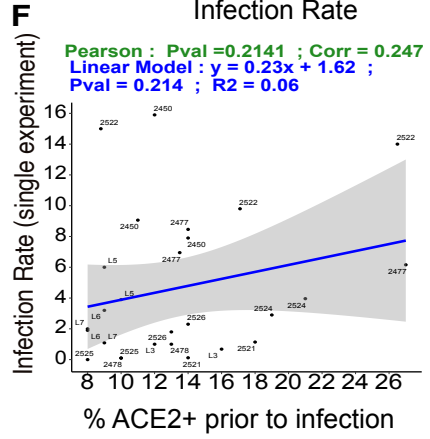
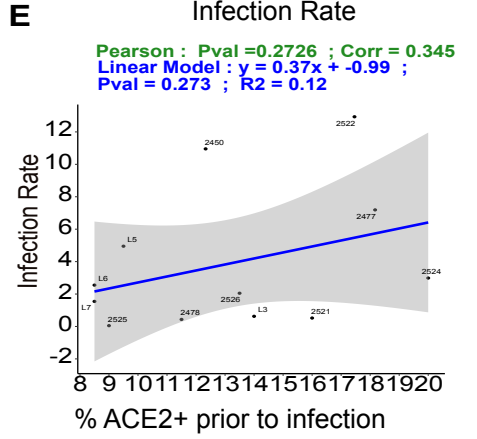
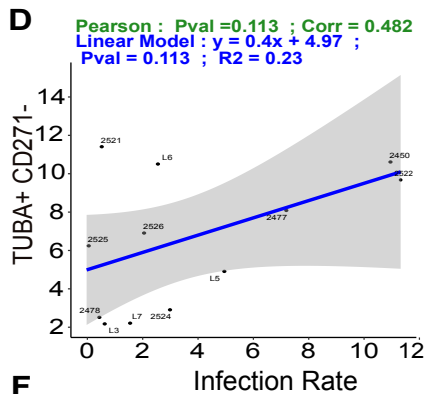
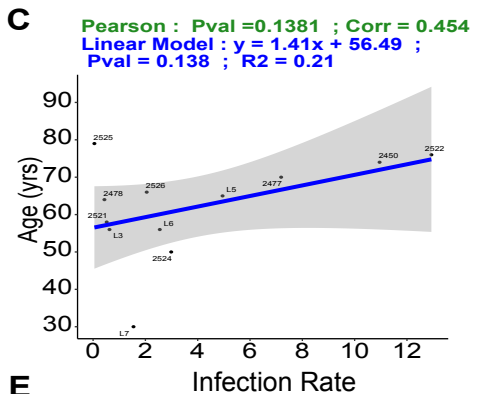
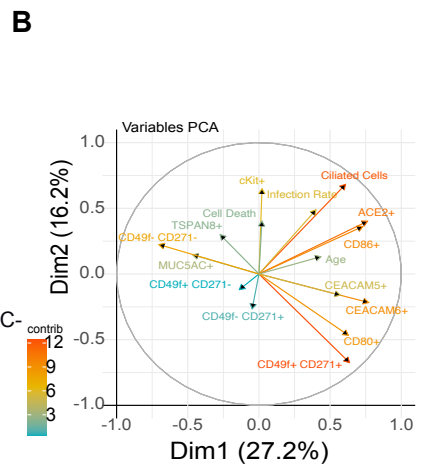
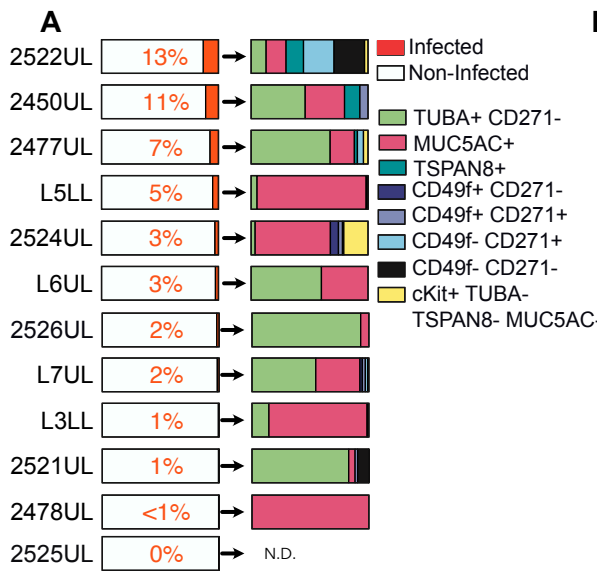
M) Percentage of dsRNA-positive cells in Mock, SARS-CoV-2 infected organoids in presence or absence of TSPAN8 or ACE2 blocking antibody. SARS-CoV-2 infection

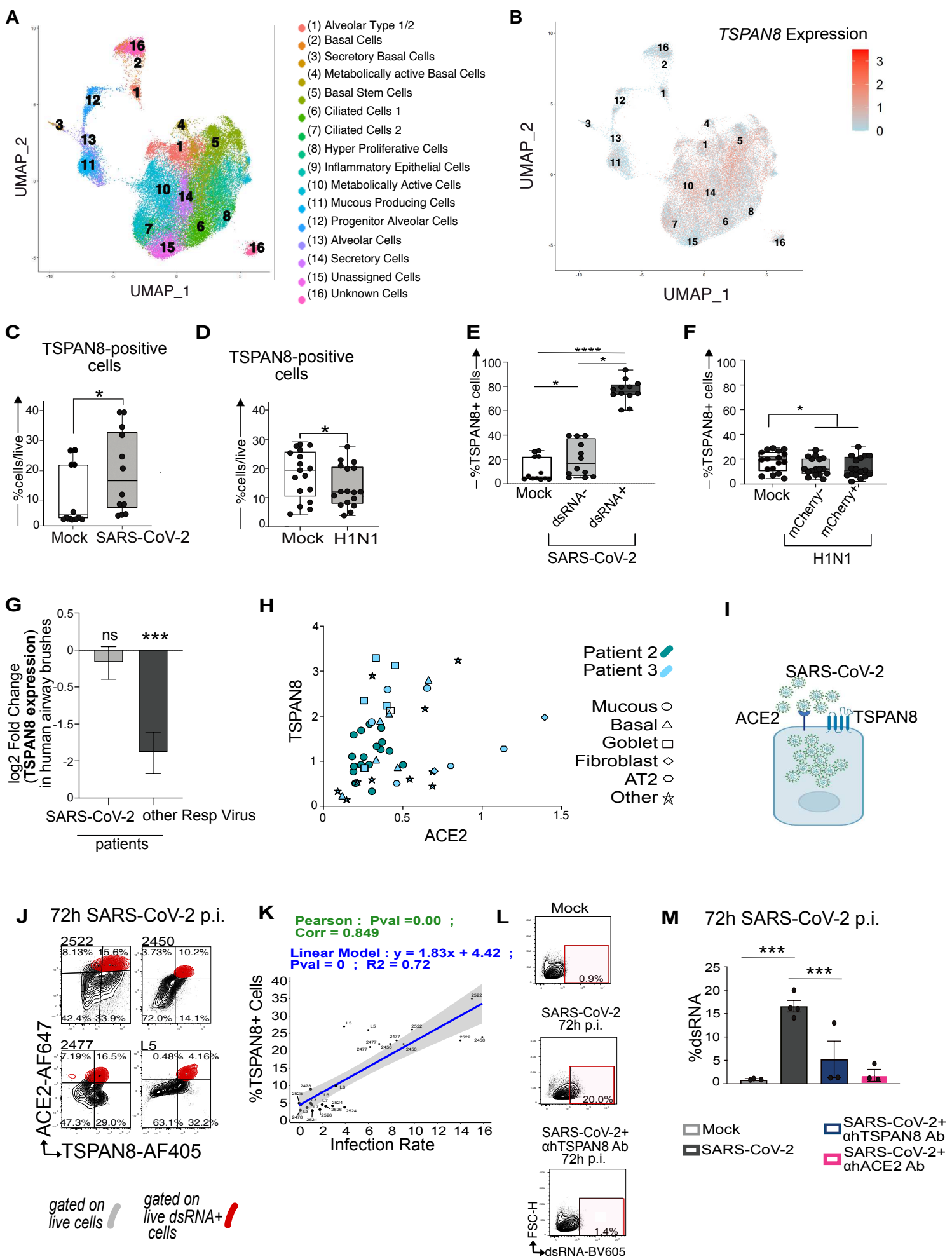
in the indicated blocking conditions (N=1, n=3). Bars represent mean, error bars are SEM, n=3. Paired t-test, * $p < 0.05$; ns, non-significant.



Hysenaj et al., Fig.1







Hysenaj et al., Fig.4

Methods and Materials

Viruses and Cell Lines

Vero E6 cells were cultured in Dulbecco's Modified Eagle Medium (UCSF Media Production) supplemented with 10% fetal bovine serum (Corning), penicillin/streptomycin (UCSF Media Production) and L-glutamine (Corning) in a humidified incubator at 37C and 5% CO₂. SARS-CoV-2 virus (USA-WA1/2020 strain) was propagated in Vero E6 cells. The Vero cells were infected with the SARS-CoV-2 virus, incubated at 37C, 5% CO₂ and after 72h the supernatant was collected. The virus was aliquoted and stored at -80C. This work was done under Biosafety Level 3 (BSL-3) conditions.

Determination of Virus Titers using Plaque Assay

Viral titer was quantified using a plaque assay in Vero cells. 10-fold dilutions of the virus stock were added to Vero cells in a 12-well plate for 1 hour, after which an overlay of 1.25% Avicel RC-591 in DMEM (UCSF Media Production) was added. The cells were incubated at 37C, 5% CO₂ for 72 hours. The cells were fixed with 10% formalin, stained with crystal violet, and washed with water. The plaques were counted to determine the titer of the virus stock. All work was done under BSL-3 conditions.

H1N1/PR8¹ (gift from Dr. Yoshihiro Kawaoka) was propagated in serum pathogen-free fertilized chicken eggs (Charles River) as previously described². In brief, freshly fertilized eggs were kept in an automatic egg turner for 10 days before injection of the virus into the allantoic cavity. Embryos were incubated with the virus for 2 days. Allantoic fluid was harvested and snap-frozen in liquid nitrogen. Titers were determined with a hemagglutination assay. This work was done under BSL-2 conditions.

Patient samples

Human cancer patient samples, 2425LL, 2450UL, 2477UL, 2478UL, 2520LL, 2521UL, 2522UL, 2524UL, 2525UL, 2526UL, 2527UL, 2531UL, 2547UL, 2551ML, were collected under UCSF study CC#00654 (IRB 11-06107) from patients undergoing thoracic surgery.

Donor lungs rejected for transplantation were received from organ procurement organization (Donor Network West) as previously described³. Explicit approval for the use of donor lungs for research was sought from each donor's family by Donor Network West as part of the standard organ donation process. Local institutional review board approval is not required because research on tissues from deceased organ donors is not considered human subject research, although institutional biosafety approval was

obtained from University of California, San Francisco (UCSF) Institutional Review Board.

Lung Organoid Culture

Adult human lung stem cells were derived from non-tumor lung tissue obtained from patients undergoing lung resection or from donor organs. The Roose lab has IRB Exempt certification (IRB #: 12-09467). Isolation of lung stem cells was performed using a protocol adapted from Sachs and colleagues (2019). Lung tissue was washed 2X with DPBS, placed in sterile petri dishes (Corning® Gosselin™) and filled with 10 ml of digestion buffer (DMEM/F12, Collagenase I 1.5mg/ml (Thermo Fischer Scientific, cat#17100017), HEPES (UCSF cell culture facility), penicillin 10,000 IU/mL (Thermo Fischer Scientific Cat#15140122). The tissue chunk was minced in 1-3mm pieces using a scalpel blade and transferred with the digestion buffer into a 50 ml Falcon tube. After 1 hour of incubation at 37Celsius with shaking at 125 rpm, the digested tissue was filtered through a 100µM filter and transferred into a 50 ml Falcon tube. Cells were pelleted (600g, 5 minutes at 4Celsius) and washed twice with DPBS (UCSF Cell Culture Facility). Lung cell pellet was then resuspended in 1300 µL Matrigel (Corning Cat#356230) and plated in ~50 µL droplets in a 24 well tissue culture plate. Plates were placed at 37Celsius with 5% CO2 for 20 min to solidify the Matrigel droplets upon which 550uL of lung organoid media was added to each well. Plates were incubated in the standard tissue culture incubator at 37Celsius. Images were taken on Days 10 to 21 using a BZ-X700 inverted microscope with a CCD cooling camera and BZ-X analysis software (KEYENCE).

Lung Organoid Biobank

After 1 to 3 weeks of growth, organoids were dissociated into single cells using TrypLE express. Cells were washed twice with DPBS, pelleted and resuspended in freezing media (Gibco™ Recovery™ Cell Culture Freezing Medium -1560446) and transferred immediately in the Mr. Frosty™ Freezing Container (Cat 51100-001) in the -80C Freezer. Frozen vials were then transferred to the Liquid Nitrogen freezer. When needed, cells were quickly thawed using the 37C water bath, and transferred to 15ml tubes filled with DPBS. Cells were pelleted (600g, 5 minutes at 4Celsius) and washed twice with DPBS. Lung cell pellets were then resuspended in 1300 µL Matrigel (Corning) and plated in ~50 µL droplets in a 24 well tissue culture plate. Plates were placed at 37 degrees Celsius with 5% CO2 for 20 min to solidify the Matrigel droplets upon which 550uL of lung organoid media was added to each well. Plates were incubated in a standard tissue culture incubator at 37 degrees Celsius.

Lung Organoid Infection

Lung organoids were removed from Matrigel following 1 minute incubation with Dispase 0.5U/ml (StemCell Technologies, Cat #07913), carefully transferred into a 15 ml tube,

washed with DPBS (UCSF Cell Culture Facility) supplemented with 5 mM EDTA (Cellgro, Cat#46-034-CI) and pelleted (200g, 3 minutes at Room Temperature). Organoid pellets were resuspended in 500 µl of lung organoid media. Virus was added at MOI=0.15 for H1N1/PR8 mCherry and MOI=0.3 for SARS-CoV-2. Every organoid well contained 10,000 cells. After 2 hours of incubation, lung organoids were washed twice with PBS and resuspended in 300µL Matrigel (Corning) and plated in ~50 µL droplets in a 24 well tissue culture plate. Plates were placed at 37 degrees Celsius with 5% CO₂ for 20 min to solidify the Matrigel droplets upon which 550uL of lung organoid media was added to each well. Plates were incubated in the standard tissue culture incubator at 37 degrees Celsius for 92 hours. After 72 hours spectral flow cytometry analyses (CYTEK Aurora) and Confocal Analyses (Zeiss, SP8) were performed.

TSPAN8- and ACE2- blocking antibody assays

Lung organoids were removed from Matrigel following a 1 minute incubation with Dispase 0.5U/ml (StemCell Technologies, Cat #07913), carefully transferred into a 15 ml tube, washed with DPBS (UCSF Cell Culture Facility) supplemented with 5 mM EDTA (Cellgro, Cat#46-034-CI) and pelleted (200g, 3 minutes at Room Temperature). Organoid pellets were resuspended in 500 µl of lung organoid media with human TSPAN8 blocking antibody (R&D Systems, #MAB4734-SP) and/or human ACE2 (R&D Systems, #AF933-SP) at 50µg/ml. After 1 hour incubation with the antibody, SARS-CoV-2 was added. After 2 hours of incubation with the virus, lung organoids were washed twice with PBS and resuspended in 300µL Matrigel (Corning) containing human TSPAN8 blocking antibody (R&D Systems, #MAB4734-SP) and/or human ACE2 (R&D Systems, #AF933-SP) at 50µg/ml. and plated in ~50 µL droplets in a 24 well tissue culture plate. Plates were placed at 37 degrees Celsius with 5% CO₂ for 20 min to solidify the Matrigel droplets upon which 550uL of lung organoid media was added to each well. Plates were incubated in the standard tissue culture incubator at 37 degrees Celsius for 24 to 72 hours. Spectral flow cytometry analyses (CYTEK Aurora) and Confocal Analyses (Zeiss, SP8) were performed.

Flow Cytometry

Organoids were dissociated into single cells using TrypLE express (ThermoFisher Scientific Cat#12604012). Single Cell suspensions were transferred to 2 ml cryovial tubes and washed twice with FACS buffer (2% FBS, 0,5% BSA, DPBS,0.01% Azide, 0.1mg DNase + Y-27632 5mM (Abmole) before incubating for 10 minutes with Ig block (True-Stain Monocyte Blocker™ Innovex Bioscience). Cell pellets were incubated with Zombie NiR (Biolegend) and the antibody mix for 35 minutes followed by two washes with FACS Buffer. Cells were fixed for 30 minutes with Permeabilization Fixation Buffer (Ebioscience). After washing twice with washing buffer cells were incubated for 45

minutes with antibodies staining for intracellular targets followed by a 20-minute incubation with secondary conjugated antibodies. Cells were washed with DPBS before acquiring in the CYTEK Aurora Machine.

Antibodies used for spectral flow cytometry: TSPAN8 AF405, clone 45811 (R&D systems), CD66c (CEACAM6) BV510, clone B6.2 (BD Bioscience), EPCAM BV711, clone 9C4, BioLegend; CD86 BV650, clone IT2.2 (BioLegend), CD117 BV785, clone 104D2, BioLegend, CD49f PE, clone GoH3 (BioLegend), CD80 PE-Cy5, clone 2D10 (BioLegend), CD271 PE-Cy7, clone ME20.4, BioLegend, ACE2 AF647, clone Q9BYF1 (R&D systems), CEACAM5 APC-Fire750, clone 487609 (R&D systems), MUC5AC Biotin, clone 45M1 (ThermoFischer Scientific), Streptavidine APC (eBioscience), anti-TMPRSS2, clone EPR3862 (Abcam), acetylated Tubulin AF488, clone 6-11B-1 (SantaCruz Biotechnologies), anti-dsRNA (Scicons), IgG BV605, clone4053 (Biolegend)

Confocal Imaging

Lung organoids were removed from Matrigel following 1 minute incubation with Dispase 0.5U/ml (StemCell Technologies, Cat #07913), carefully transferred into a 15 ml tube, washed with DPBS (UCSF Facility) and pelleted (200g, 2 minutes at Room Temperature). The organoid pellet was incubated for 30 minutes in 4% PFA 10% FBS Triton 0.1X and TrueStain. Organoids were washed twice with PBS-Triton 0.1X and stained overnight at 4C with primary antibodies dsRNA Ab (clone J2, Scicons), SARS-CoV-2 Nucleocapsid Ab (Novus Biologicals, Cat#NB100-56576SS). Organoids were then washed and incubated for two hours with secondary antibodies Goat anti mouse IgG AF555, clone4053 (Biolegend), CD49f PE, clone GoH3 (BioLegend), ACE2 AF647, clone Q9BYF1 (R&D systems), and acetylated Tubulin AF488, clone 6-11B-1 (SantaCruz Biotechnologies). After staining, organoids were washed with DPBS- 0.1X Triton and resuspended in Fructose-Glycerol Clearing Solution 60% (vol/vol) glycerol and 2.5 M fructose. Organoids were mounted on coverslips and imaged with Leica SP8 Confocal Microscopy. Confocal Z-stack images were generated using the staining maxima.

Statistics

Statistical analyses were run in R (version 4.0.2) (quote 1). Paired Samples Wilcoxon Test were performed using Wilcox.Test (stat v4.0.2) (quote 1). Outliers were removed if the value was over $Q3+1.5 IQR$. We used PCA (FactoMineR v2.4) to perform Principal Component Analysis and fviz_pca_ind or fviz_pca_var (factoextra v1.0.7) (quote 3) for visualization. Community distances were evaluated based on relative abundance of cell populations by PERMANOVA (Bray, $*p<0.05$). For spearman's and pearson's correlations, we used cor function (stat v4.0.2) to compute the correlation coefficients and lm (stat v4.0.2) to fit linear models. Data management was done using tidyverse (v1.3.0) (quote 4). All graphs were built using ggplot2 (quote 5) ^{4, 5, 6, 7, 8}.

UMAP dimensionality reduction, discovery of upregulated/downregulated genes, and gene expression related plots were constructed in R v 4.0.3 (1) via Seurat v 4.0 provided by the Satija Lab⁹ and ggplot2 v 3.3.3. Clusters were separated using the Louvain clustering method with a resolution of 0.6, and upregulated differential expression gene scores between clusters were used to establish cell type identities⁸. Bar graph of enrichment analysis up regulated pathway and processes based in cells positive for SARS-CoV-2 reads were generated using Metascape¹⁰.

Software

GraphPad Prism 9 was used for DATA visualization. Fiji/ImageJ (V) software 2.1.0/1.53c was used for confocal microscopy analysis. FlowJO and CYTEK Aurora software 10.7.1 was used for spectral flow analyses.

Single cell and library preparation for scRNA-sequencing

For the single-cell RNA sequencing experiments each organoid was generated from 4 different donors. After organoid dissociation into single-cell suspension using TrypLE express (ThermoFisher Scientific Cat#12604012), cells were counted, and pooled together (78,000 total cells). Pooled cells were then loaded evenly across two lanes in the Chromium Controller for generating single-cell contained in lipid droplets following the manufacturer's instructions (10X Genomics). A Chromium Single cell 3' Reagent Kit (v3.1) (10X Genomics) was used for reverse transcription, cDNA amplification and library construction of the gene expression libraries following the manufacturer's instructions. All samples were encapsulated, and cDNA was generated within 6 hours after organoid processing. Finally, Pooled libraries were sequenced on an Illumina NovaSeq 6000.

Single cell RNAseq analysis

Data pre-processing of 10x Genomics Chromium scRNA-seq data:

Data pre-processing was performed as previously described (CITE: Combes et al, Nature, 2021). Briefly, sequencer-obtained bcl files were demultiplexed into individual samples using the Cellranger 3.0.2 suite of tools (<https://support.10xgenomics.com>). Feature-barcode matrices were obtained for each sample by aligning the raw fastqs to GRCh38 reference genome (annotated with Ensembl v85) using the Cellranger count. Raw feature-barcode matrices were loaded into Seurat and genes with fewer than 3 UMIs were dropped from the analyses. Matrices were further filtered to remove events with greater than 30% percent mitochondrial content, events with greater than 50% ribosomal

content, or events with fewer than 250 total genes. The cell cycle state of each cell was assessed using a published set of genes associated with various stages of human mitosis.

Inter-sample doublet detection:

Inter-sample doublet detection was performed as previously described (CITE: Combes). Libraries containing samples pooled prior to loading were processed using Freemuxlet (<https://github.com/statgen/popsicle>) to identify clusters of cells belonging to the same patient via SNP concordance. Cells are classified as singlets arising from a single library, doublets arising from two or more libraries, or as ambiguous cells that cannot be accurately assigned to any existing cluster (due to a lack of sufficient genetic information).

Data quality control and Normalization:

The filtered count matrices were normalized, and variance stabilized using negative binomial regression via the scTransform method offered by Seurat ¹¹. The effects of mitochondrial content, ribosomal content, and cell cycle state were regressed out of the normalized data to prevent any confounding signal. The scTransformed data from different sequencing libraries were combined and normalized using Harmony sc integration software.

Intra-sample heterotypic doublet detection:

All libraries were further processed to identify heterotypic doublets arising from the 10X sample loading. Processed, annotated Seurat objects were processed using the DoubletFinder package ¹². Briefly, the cells from the object are modified to generate artificial duplicates, and true doublets in the dataset are identified based on similarity to the artificial doublets in the modified gene space. The prior doublet rate per library was approximated using the information provided in the 10x knowledgebase (<https://kb.10xgenomics.com/hc/en-us/articles/360001378811>) and this was corrected to account for homotypic doublets using the per-cluster numbers in each dataset.

Differential expression tests and cluster marker genes, cluster annotation:

Differential gene expression (DGE) tests were performed on log-normalized gene counts using the Poisson test (with a latent batch variable to account for multiple library preparations) provided by the FindMarkers/FindAllMarkers functions in Seurat. Cluster marker genes were identified by applying the DE tests for upregulated genes between cells in one cluster to all other clusters in the dataset. Top ranked genes (by log-fold changes) from each cluster and between clusters are used for further analysis.

Data Mining

To examine presence of *ACE2* and *TSPAN8* co-expressing cells in vivo in normal human lung, we analyzed the processed 10X single-cell sequencing data from Travaglini et al¹³. Of the 60,993 cells derived from lung tissue of 3 patient donors in this dataset, 48 cells were found with at least 1 UMI (unique molecular identifier) for both genes. Only 1 cell derived from patient 1, which had fewer cells sequenced overall and so we excluded it. The expression values represent $\ln(\text{UMI-per-10K} + 1)$ in each of the 47 cells from patients 2 and 3. Cell type designations were determined by Travaglini et al.

Differential expression of *TSPAN8* in nasal swabs of adult patients with acute respiratory illness (ARI) due to COVID-19 (n=93) or other viral infection (n=41), in comparison to patients with ARI due to non-viral etiology (n=100), was derived from Mick et al¹⁴. The differential expression analysis between the 3 viral status groups was performed with the R package limma while controlling for gender and age.

References

1. Fukuyama, S. et al. Multi-spectral fluorescent reporter influenza viruses (Color-flu) as powerful tools for in vivo studies. *Nature communications* **6**, 1-8 (2015).
2. Brauer, R. & Chen, P. Influenza virus propagation in embryonated chicken eggs. *JoVE (Journal of Visualized Experiments)*, e52421 (2015).
3. Ross, J.T., Nesseler, N., Lee, J.W., Ware, L.B. & Matthay, M.A. The ex vivo human lung: research value for translational science. *JCI Insight* **4** (2019).
4. Team, R.C. R: A language and environment for statistical computing. (2013).
5. Lê, S., Josse, J. & Mazet, F. Package 'FactoMineR'. *J Stat Softw [Internet]* **25**, 1-18 (2008).
6. Kassambara, A. & Mundt, F. Package 'factoextra'. *Extract and visualize the results of multivariate data analyses* **76** (2017).
7. Wickham, H., Francois, R., Henry, L. & Müller, K. others. 2015. ". *Dplyr: A Grammar of Data Manipulation.*" *R Package Version 0.4* **3**.
8. Wickham, H. et al. Springer-Verlag. *New York* (2016).
9. Hao, Y. et al. Integrated analysis of multimodal single-cell data. *bioRxiv* (2020).
10. Zhou, Y. et al. Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nature communications* **10**, 1-10 (2019).

11. Hafemeister, C. & Satija, R. Normalization and variance stabilization of single-cell RNA-seq data using regularized negative binomial regression. *Genome biology* **20**, 296 (2019).
12. McGinnis, C.S., Murrow, L.M. & Gartner, Z.J. DoubletFinder: Doublet Detection in Single-Cell RNA Sequencing Data Using Artificial Nearest Neighbors. *Cell Syst* **8**, 329-337 e324 (2019).
13. Travaglini, K.J. *et al.* A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619-625 (2020).
14. Mick, E. *et al.* Upper airway gene expression reveals suppressed immune responses to SARS-CoV-2 compared with other respiratory viruses. *Nature communications* **11**, 1-7 (2020).

Legends Supplementary Figures

Supplementary 1: Donor derived lung organoids display distinct phenotypes.

- A) Brightfield images of twenty lung organoids derived from different donors.
- B) Brightfield images of organoids derived from the upper and lower region of the left lung of individual L2. (Right) The composition of lung organoids based on spectral flow cytometry analyses derived from the upper and lower lobes of subject L2 and L7 are represented in pie charts. Each part of the pie chart represents the mean value of the percentage of the cells population in distinct donor derived lung organoids.
- C) Lung organoids compositions derived from spectral flow cytometry analysis of different passages of organoids from multiple donors. Each section of the pie chart represents the cell population mean from 3 independent experiments.
- (D-F) Linear regression showing the relationship between age of the donor and cell types found in the donor derived lung organoids (N=3, n=3). Each dot represents the mean value of the percentage of the cells population in distinct donor derived lung organoids. Values for Pearson correlations and corresponding P values are depicted. Functions of the positive or negative correlations are depicted by the Linear Model.
- (D) Age and % of ciliated like cells (TUBA+ CD271-). (E) Age and % of basal stem cells (CD271+ or CD49f+). (F) Correlation between age of the donors and proportion of mucus producing cells.

Supplementary Figure 2: Spectral flow cytometry analyses of mock and infected lung organoids.

A) Spectral flow cytometry gating strategy. Cell populations are defined after excluding ZombieNIR-positive, dead cells and doublets in steps a-c. (1) TUBA + CD271- are considered as ciliated cells. Fraction (d) is analyzed for MUC5AS and TSPAN8 and (2) MUC5AC+ TUBA- cells are considered goblet-like cells. (3) TSPAN8+ MUC5AC- TUBA- are defined as pre-goblet cells. Fraction (e) is analyzed for cKit and (g) are considered cKit-positive cells. Fraction (f) is analyzed for CD49f and CD271, (4) CD49f+ CD271-, (5) CD49f+ CD271+ and (6) CD49f- CD271+ are considered basal stem cells and (7) CD49f- CD271- are undefined cells.

B) Spectral flow cytometry gating strategy for mCherry+ cell populations. The same gates as in Supplementary 1 A) were drawn on live, single cell and mCherry+ cells (h).

Supplementary Figure 3: H1N1 viral infection of lung organoids.

- A) Experimental scheme of lung organoid infection with H1N1/PR8 mCherry virus.
- B) Confocal images (Z-stack) of H1N1 infected (top panel) and mock infected (bottom panel) whole mounted lung organoids. Scale bars are 50uM.
- C) Spectral flow cytometry plot of live cells and gating for mCherry-positive cells.
- D) Percentage of H1N1/PR8 mCherry-positive lung organoids.
- E) Percentage of cell death in indicated organoids.
- F) Cell composition in mock and infected organoids. Pie charts represent mean percentages of n=3 independent experiments for each donor derived organoids.
- G) Percentage of H1N1/PR9 mCherry+ cells (left) and H1N1/PR8 mCherry-infected cell types (right). Colors as in 2F. Mean values of n=3 experiments shown in bars.
- H) Fold change in median fluorescence intensity (MFI) of CEACAM6, CEACAM5, CD80, CD86 and ACE2 during infection at 36h H1N1 p.i. (MOI=0.15). Bars represent mean, error bars are SEM, n=3. Paired t-test, *p<0.05; ns, non-significant.

Supplementary Figure 4. Reproducibility of SARS-CoV-2 infection in lung organoids.

- A) dsRNA staining for spectral flow cytometry analyses of SARS-CoV-2 infected organoids at 72h p.i.
- B) Percentage of dsRNA+ in replicate experiments infecting the same donor derived organoid with SARS-CoV-2. Each point represents an independent experiment.
- C) Cell composition of organoid 2522UL in Mock and infected organoids at 72h p.i. (MOI=0.3) in five separate experiments.
- D) The percentage of dsRNA+ cells (left) and the cell types in dsRNA+ cells (right) based on spectral flow cytometry analyses.
- E) The percentage of cells positive for CEACAM6, CEACAM5, CD80, CD86 and ACE2 at 72h p.i. in mock infected, SARS-CoV-2 exposed (dsRNA-) and SARS-CoV-2 infected (dsRNA+) cells from organoid 2522UL. Every point represents data for a single experiment. Bars represent mean, error bars are SEM. Paired t-test, * $p < 0.05$, **** $p < 0.0001$; ns, non-significant.

Supplementary Figure 5: Donor derived lung organoid response to SARS-CoV-2.

- A) The percent of cell death in mock and SARS-CoV-2 infected organoids (MOI=0.3, 72h p.i.). Each DOT represents the mean value of the percentage of the cells population in distinct donor derived lung organoids (N=3, n=3).
- B) Cell composition based on spectral flow cytometry analyses of mock and SARS-CoV-2 infected organoids (MOI=0.3, 72h p.i.). Each part of the pie chart represents the mean value of the percentage of the cells population in distinct donor derived lung organoids (N=3, n=3).
- C) Principal component analysis using cell populations and infection conditions. PCA reduced the descriptors into two dimensions as shown in the individual plot. Each point corresponds to an observation in the dataset and were colored according to their mock or infection state. Community distances were evaluated based on relative abundance of cell populations by PERMANOVA (Bray, *p<0.05).
- D) Comprehensive analyses of the percentage of MUC5AC+ and CD49f- CD271- cells in mock and infected organoids. Each point represents the mean of % cell type for the distinct donor for Mock or SARS-CoV-2 conditions. Wilcoxon signed-rank test, *p<0.05, N=12, n=3. Outliers were removed if the value was over Q3+1.5 IQR.
- E) Comprehensive analyses of the percentage of cells positive for activation markers in mock and infected organoids. Wilcoxon signed-rank test, *p<0.05, N=12, n=3. Outliers were removed if the value was over Q3+1.5 IQR.
- F) Linear regression showing the relationship between SARS-CoV-2 infection rate and fraction CD86+ cells, TUBA+ cells and MUC5AC+ cells prior to infection. Each dot represents the mean of % cell type in 3 independent experiment (n=3) for the distinct donor. Linear regression, R2 value and statistical significance stated on graph. Mean values, N=2 to 3, n=3 for every donor derived organoid.

Supplementary Figure 6: Differentially expressed genes donor derived organoids.

Clustermap of top differentially regulated genes between clusters established in Fig 5A. Hierarchical clustering is used on a concatenated set of genes constructed from the top 5 statistically significant expressed genes from each cluster.

Supplementary Figure 7: TSPAN8 a potential biomarker in COVID-19 severity.

- A) UMAP's constructed for individual organoids from four different subjects. cell subsets from the merged data, with overlaid clusterings from Fig 5A.
- B) Top statistically significant genes when computing differential expression between cells that had any genes mapping to the covid genome (TRUE) versus those that had no genes mapped (FALSE). The size of the dot shows the percentage of cells expressing the gene.
- C) Bar graph of enrichment analysis up regulated pathway and processes based on cells positive for SARS-CoV-2 reads, colored by p-values. The darker the color, the lower the p-value. Figure created in Metascape.
- D) Analysis of cell populations in mock and infected organoids. Percentages of TSPAN8+ cells. Wilcoxon signed-rank test, * $p < 0.05$, $N=12$, $n=3$.
- E) Example of gating strategy of ACE2 and TSPAN8 expression. For all the organoids and the individual staining sets, quadrants were defined on the basis on unstained controls.
- F) Spectral flow cytometry plots showing live dsRNA^{neg} cells (in gray) and live dsRNA^{pos} cells (in red- overlaid) in Mock and SARS-CoV-2 infected organoids at 72 hours p.i. (MOI=0.3).
- G) Linear regression showing the relationship between the fraction of TSPAN8+ ACE2+ cell population in SARS-CoV-2 infected organoids and infection rate. Linear regression, R² value and statistical significance stated on graph. Mean values, $N=12$, $n=3$.
- H-I) Flow Cytometry plots (H) and graph of Spectral Flow Cytometry analysis of dsRNA⁺ cells (I) in 2450UL lung organoids, 24 hours post SARS-CoV-2 infection (MOI=0.3), in absence or presence of TSPAN8 or ACE2 blocking antibodies ($n=3$). Bars represent mean, error bars are SEM. Paired t-test, * $p < 0.05$, **** $p < 0.0001$; ns, non-significant.

Supplementary Table 1:

Organoid Donor Records

HTN= hypertension, HLD= hypersensitive lung disease, Afib= atrial fibrillation, CAD= coronary artery disease.

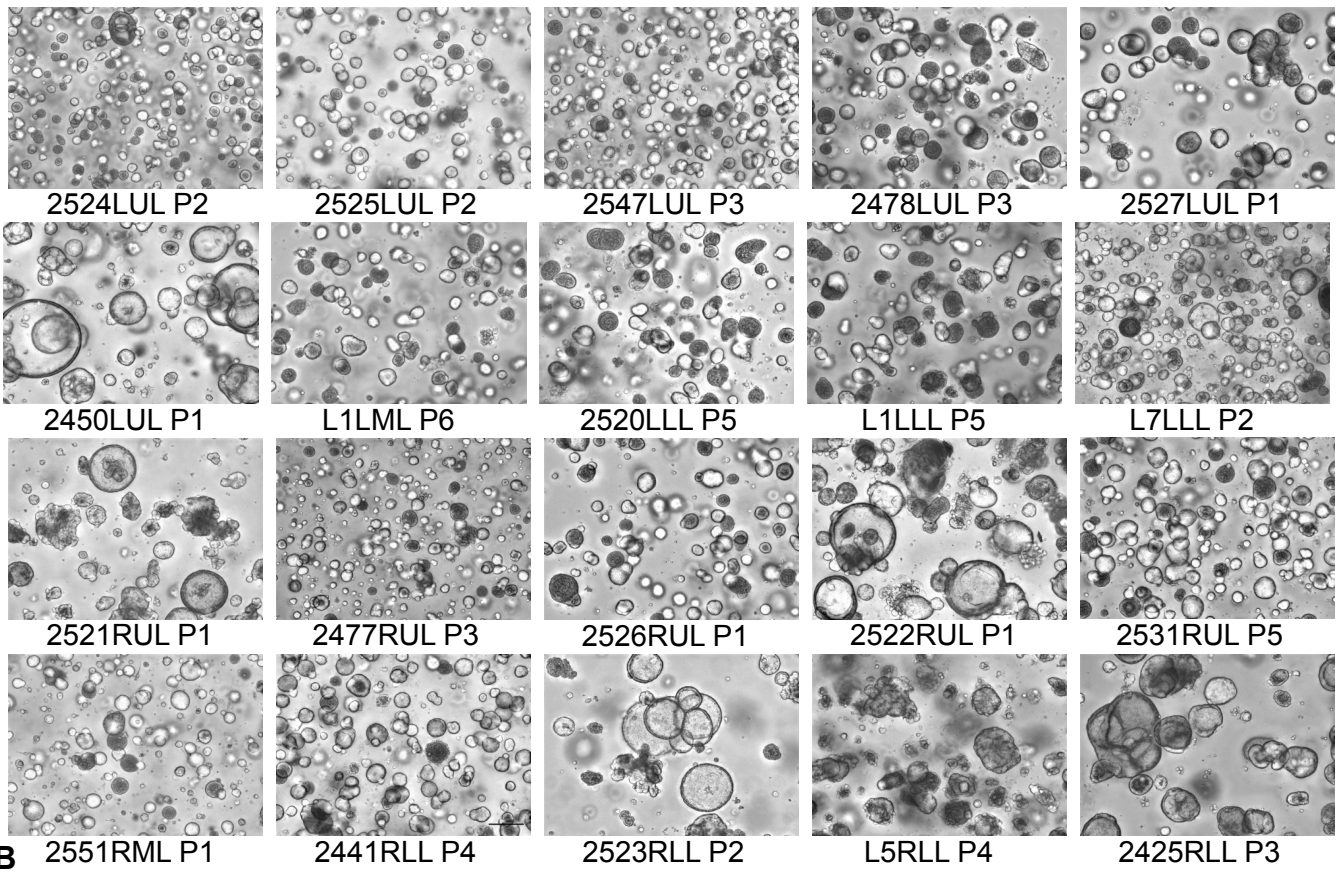
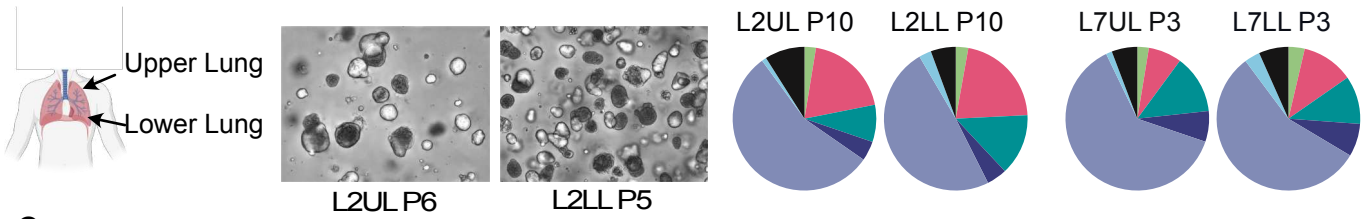
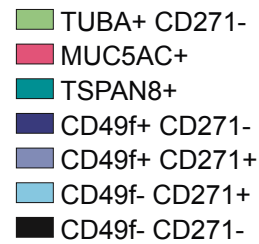
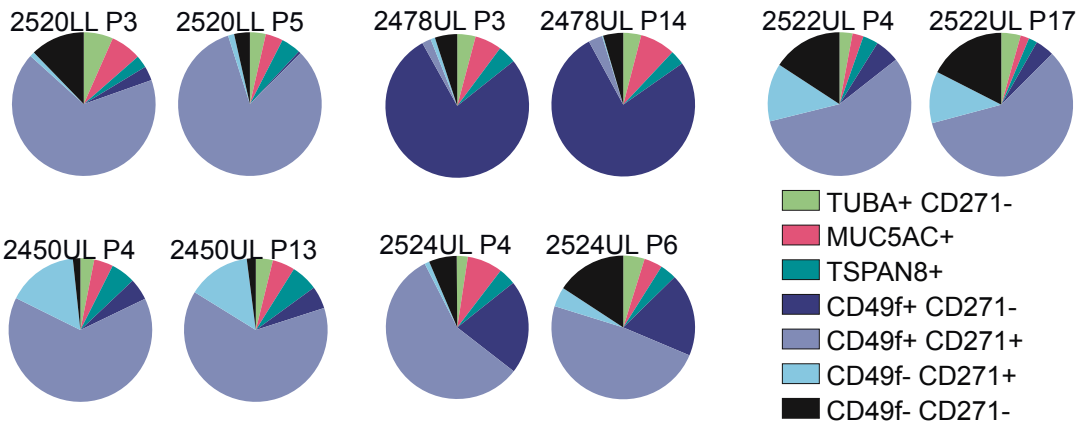
Supplementary Table 2:

List of genes expressed in different clusters of the UMAP in Figure 5A

Supplementary Table 3:

Differentially expressed genes coding for cell receptors in cells negative and positive for SARS-CoV-2 genome reads.

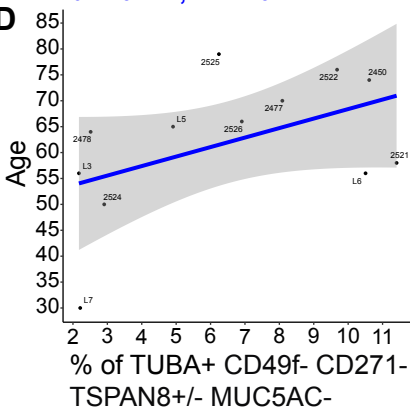
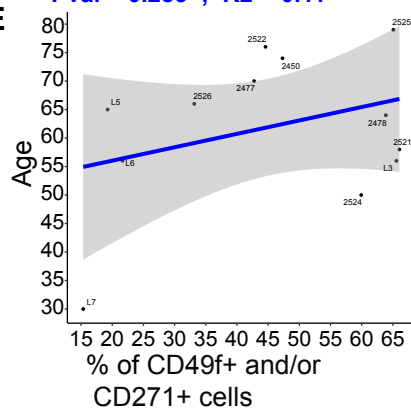
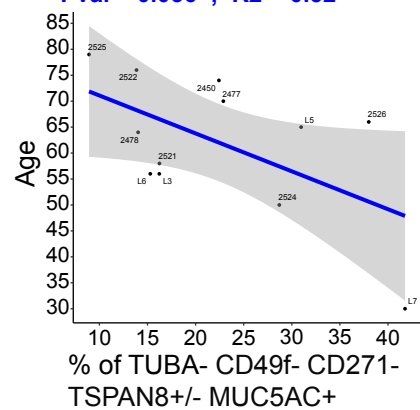
Sample ID	Age	Sex	Smoker (Y/N)	Heart Condition (Y/N)	Other Conditions	Location of Sample	Lung Sample Side	Specific Sample Location
2425LL	81	F	Y	Y	NSCLC, CAD, HTN, HLD	Lower Lobe	Right	
2441LL	49	F	N	N	NSCLC	Lower Lobe	Right	
2450UL	74	F	Y	Y	NSCLC, HTN, HLD, Non-Sustained and Paroxysmal Ventricular Tachycardia (on beta-blocker therapy)	Upper Lobe	Left	
2477UL	70	F	Y	N	NSCLC	Upper Lobe	Right	
2478UL	64	F	Y	Y	NSCLC, HTN, HLD, Afib (on Xarelto)	Upper Lobe	Left	
2520LL	46	F	N	N	NSCLC	Lower Lobe	Left	Anterior Basilar Segment
2521UL	58	M	Y	Y	NSCLC, Afib	Upper Lobe	Right	
2522UL	76	F	Y	N	NSCLC, HTN	Upper Lobe	Right	
2523LL	38	F	N	N	NSCLC, Unconfirmed Tachycardia	Lower Lobe	Right	
2524UL	50	F	N	N	NSCLC	Upper Lobe	Left	
2525UL	79	F	N	N	NSCLC	Upper Lobe	Left	Apical
2526UL	66	F	Y	N	NSCLC	Upper Lobe	Right	
2527UL	70	F	Y	N	NSCLC	Upper Lobe	Left	
2531UL	80	M	Y (former)	N	NSCLC	Upper Lobe	Right	
2547UL	76	F	Y	N	NSCLC, CAD, HTN	Upper Lobe	Left	
2551ML	26	F	Y	N	NSCLC	Middle Lobe	Right	
L1UL, L1ML, L1LL	32	M	Y (meth, marijuana)	N/A	Psychaitric Problems	Upper, Middle, Lower Lobe	Left	
L2UL, L2LL	27	M	Y	N/A	N/A	Upper, Lower Lobe	Left	
L3LL	56	M	N	N/A	Bronchial Asthma	Lower Lobe	Left	
L5LL	65	F	Y	N/A	Post-Op Pnemonia on Emphysema/ Bronchoectasia Background	Lower Lobe	Right	
L6UL	56	F	N	N/A	N/A	Upper Lobe	Left	
L7UL, L7LL	30	M	Y (former)	N/A	Asthma, Diabetes, HTN	Upper, Lower Lobe	Left	

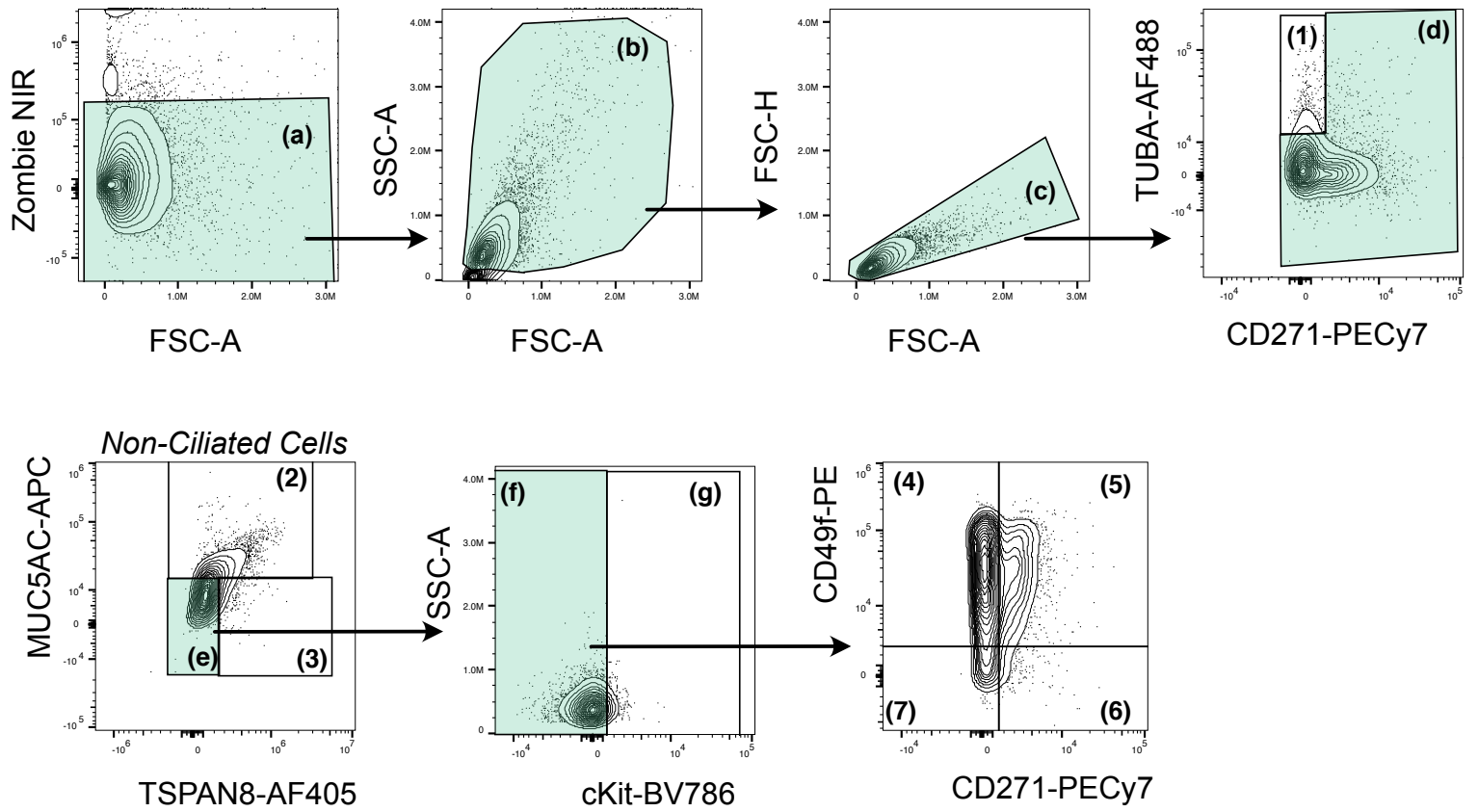
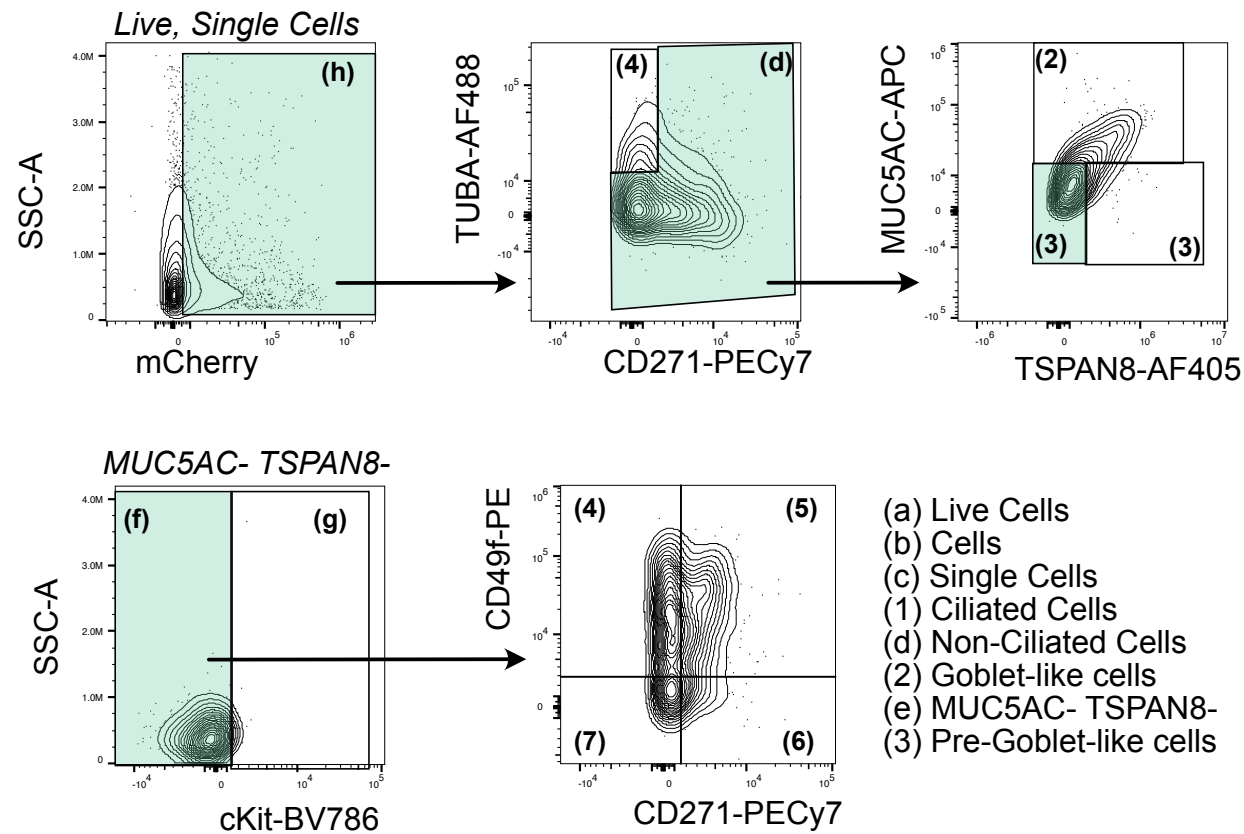
A**B****C**

Pearson : Pval = 0.11 ;
 Corr = 0.485
 Linear Model : $y = 1.83x + 50.05$;
 Pval = 0.11 ; R2 = 0.24

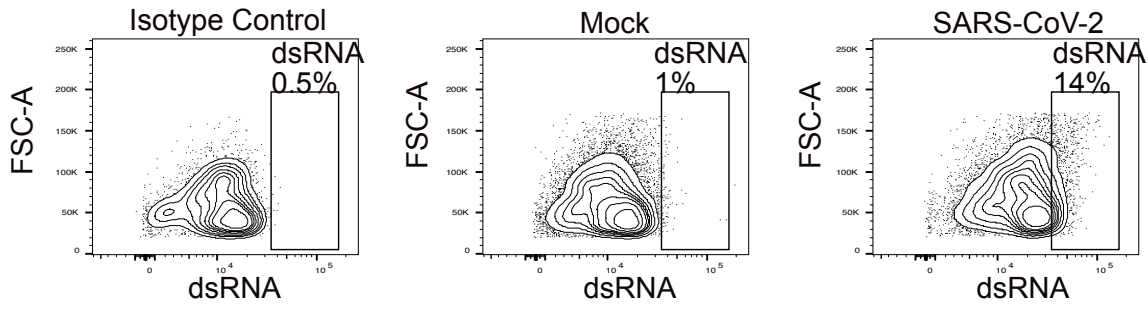
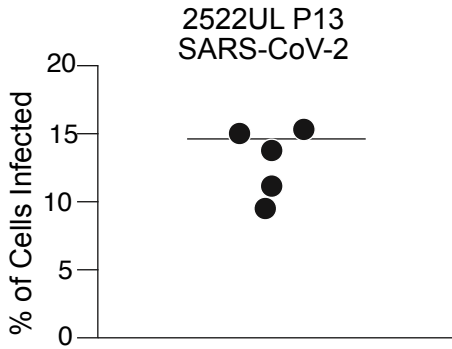
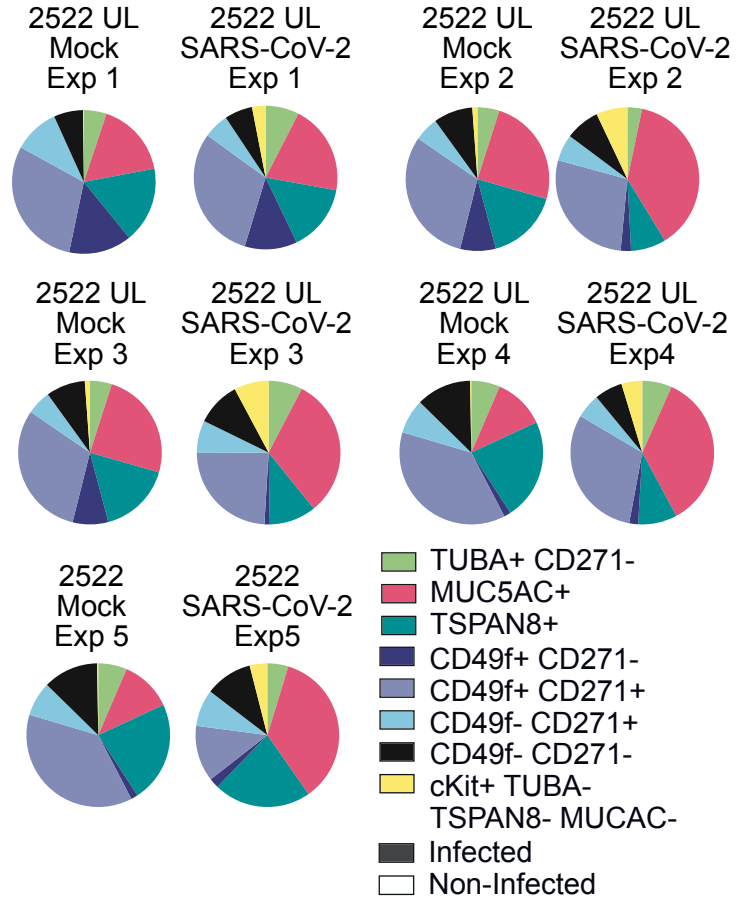
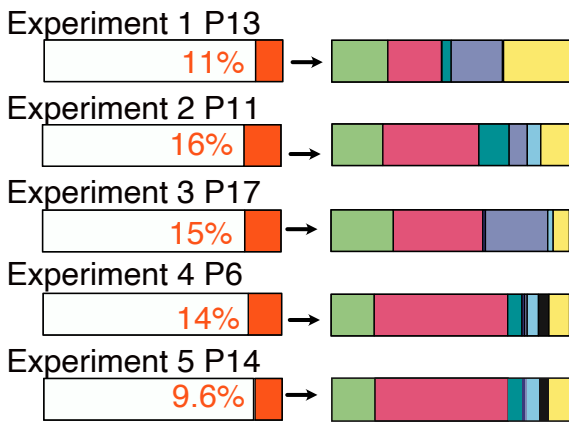
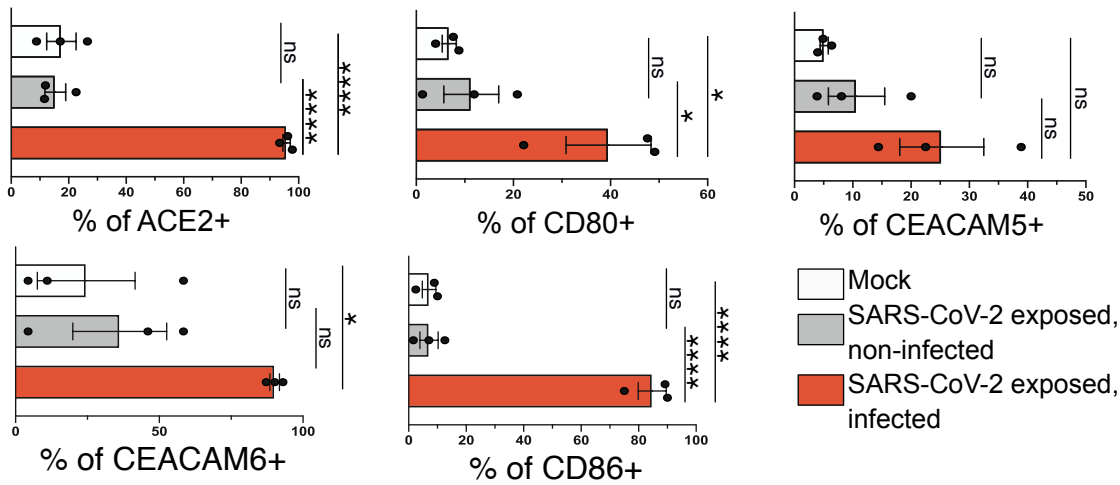
Pearson : Pval = 0.2826 ;
 Corr = 0.338
 Linear Model : $y = 0.24x + 51.33$;
 Pval = 0.283 ; R2 = 0.11

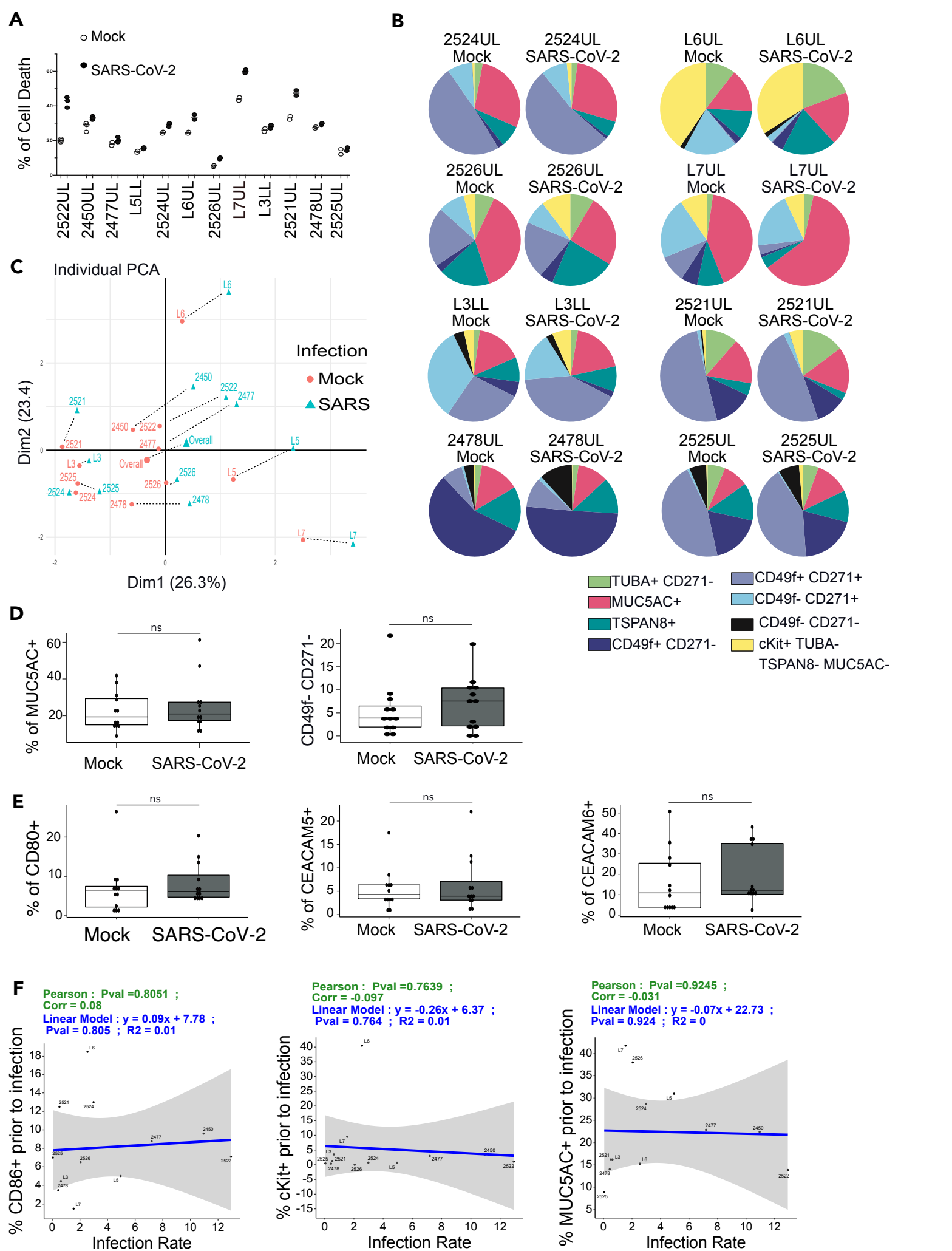
Pearson : Pval = 0.0556 ;
 Corr = -0.565
 Linear Model : $y = -0.73x + 78.39$;
 Pval = 0.056 ; R2 = 0.32

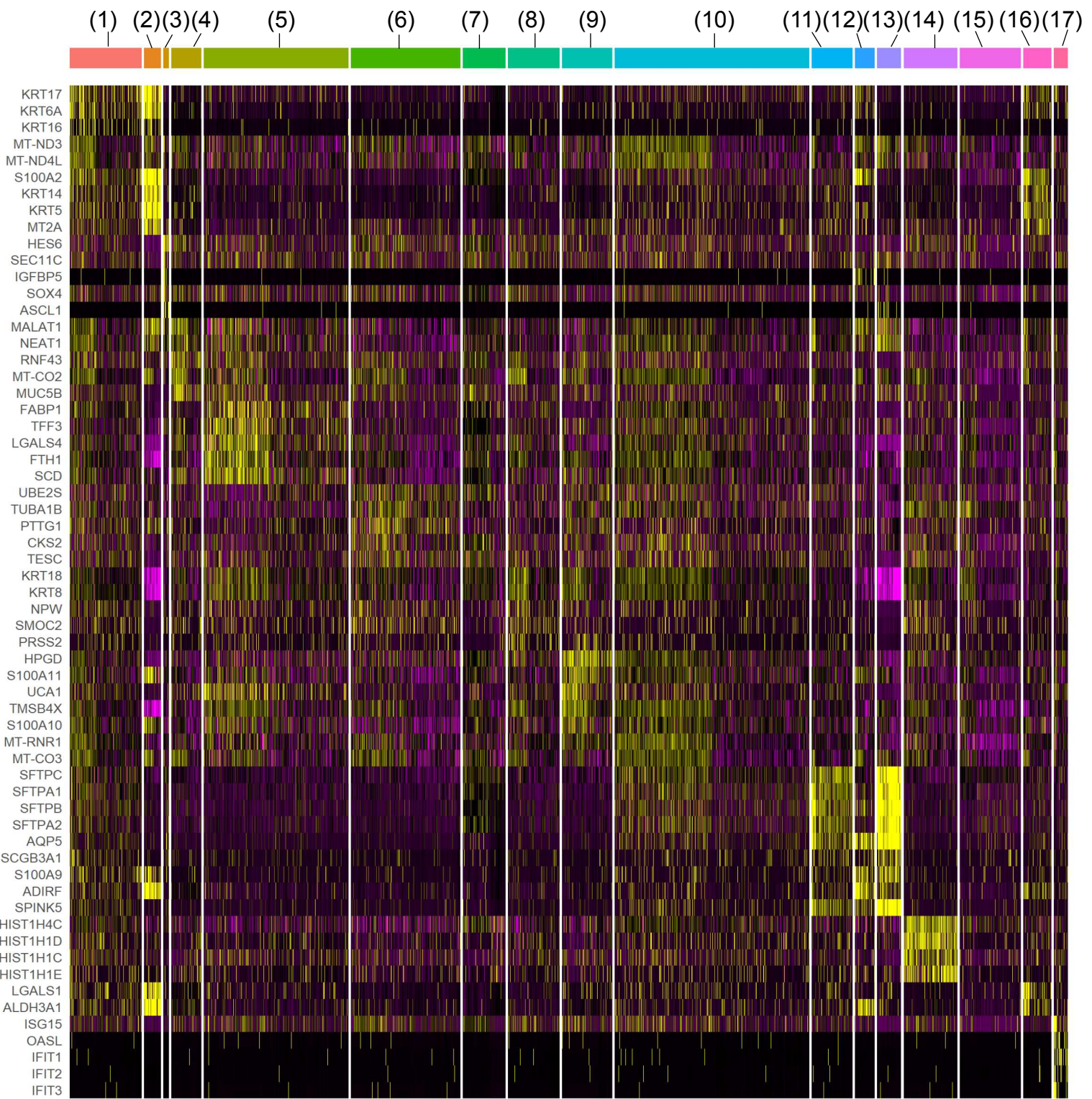
D**E****F**

A**B**

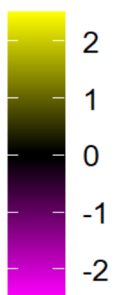
- (a) Live Cells
- (b) Cells
- (c) Single Cells
- (1) Ciliated Cells
- (d) Non-Ciliated Cells
- (2) Goblet-like cells
- (e) MUC5AC- TSPAN8-
- (3) Pre-Goblet-like cells
- (f) cKit-
- (g) cKit+
- (4) CD49f+ CD271-
- (5) CD49f+ CD271+
- (6) CD49f- CD271-
- (7) CD49f- CD271+
- (h) mCherry+ Cells

A**B****C****D****E**

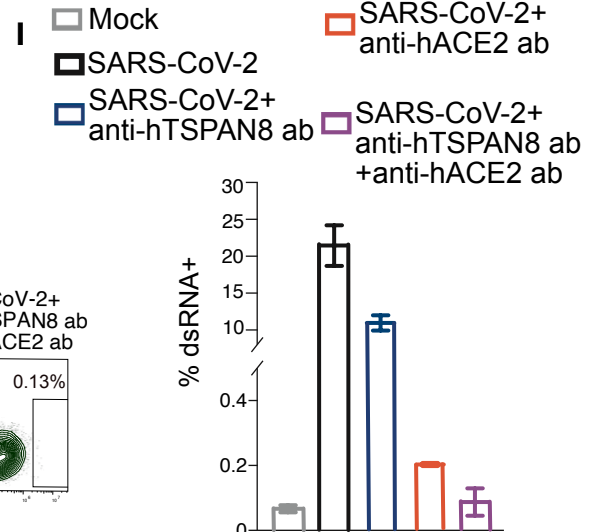
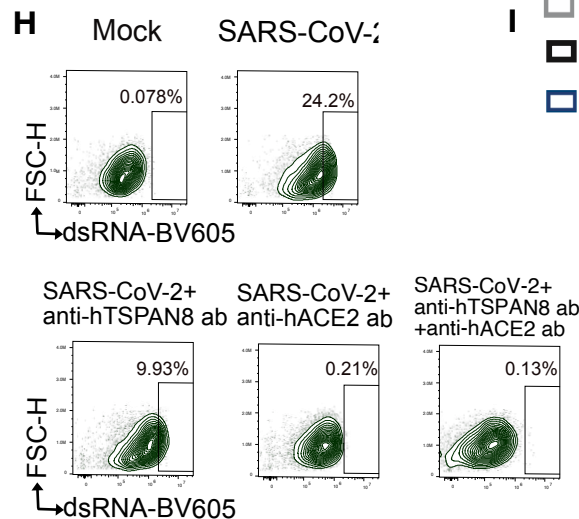
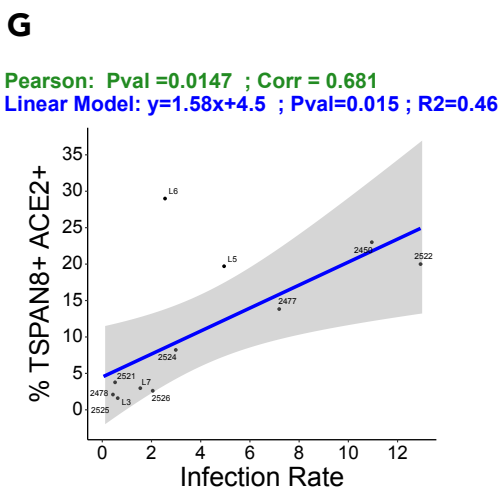
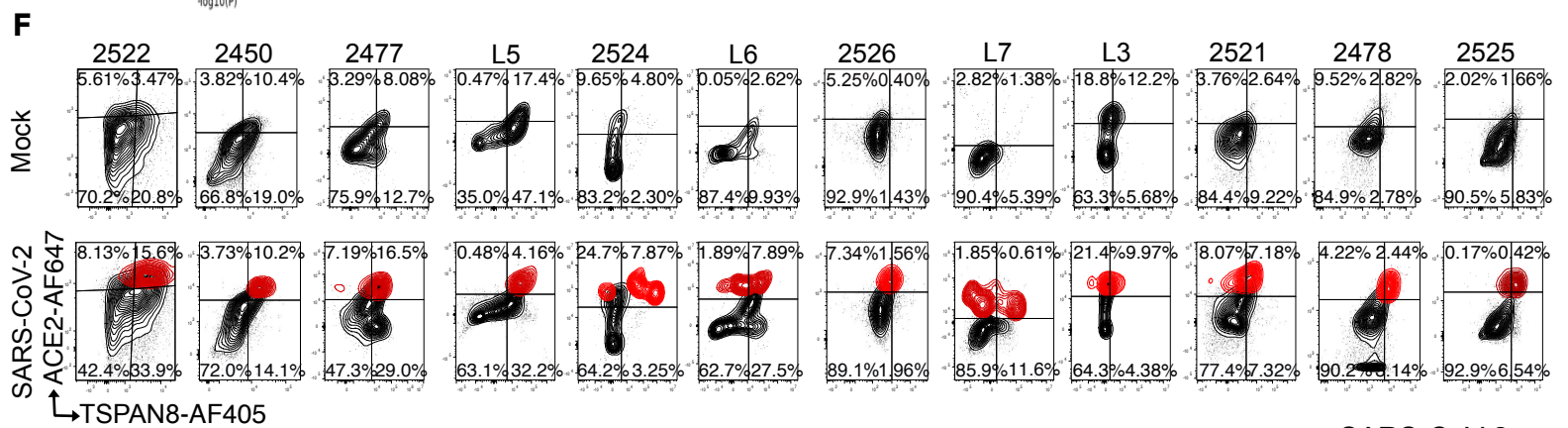
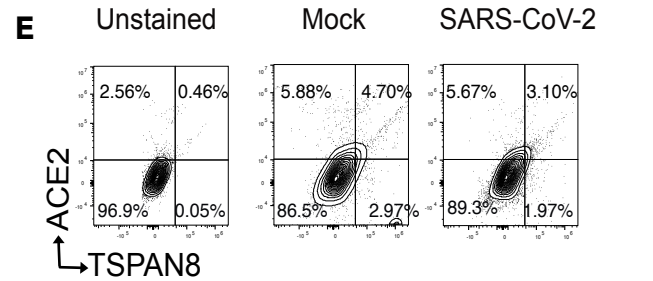
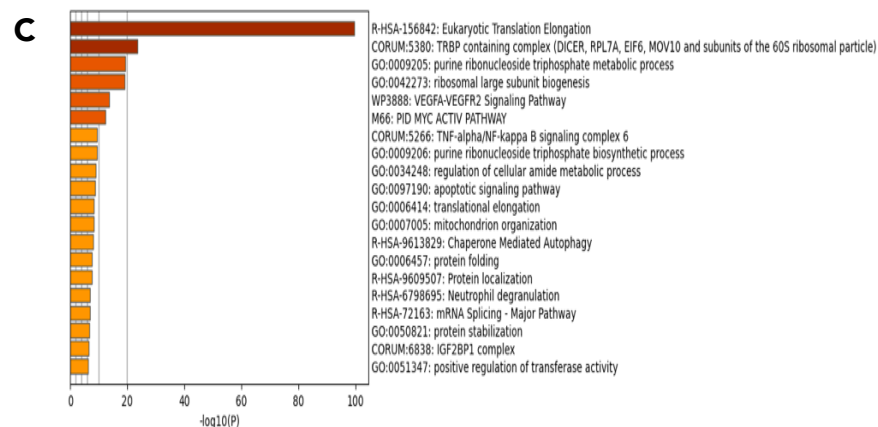
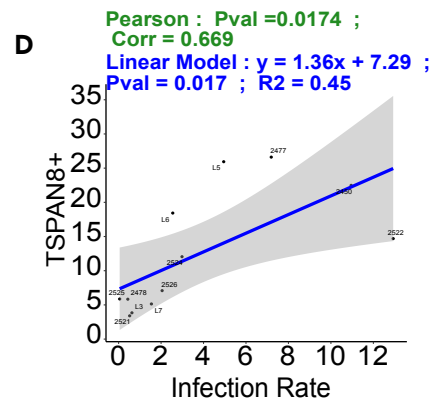
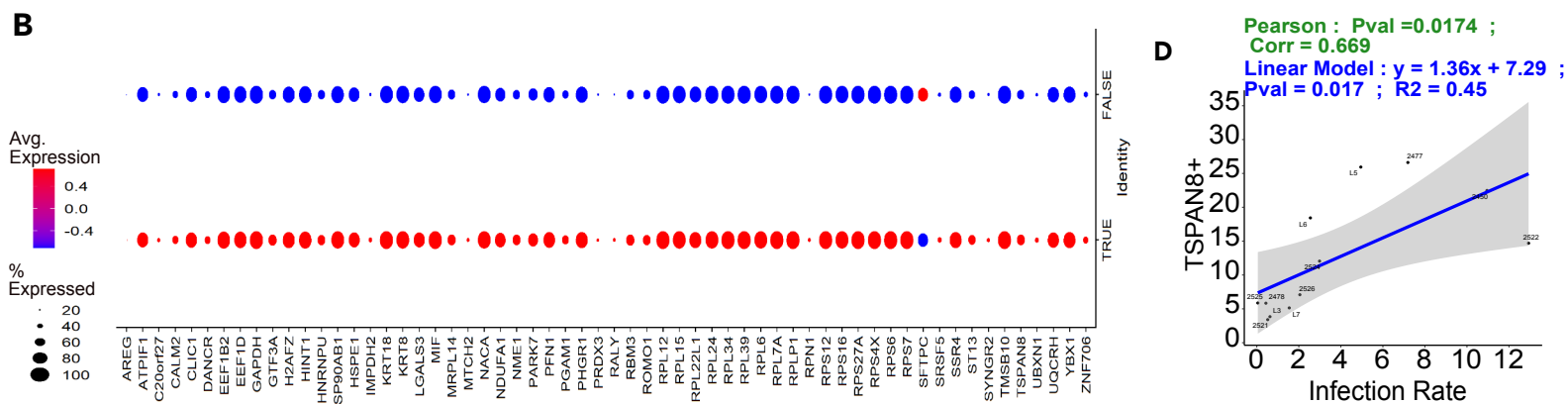
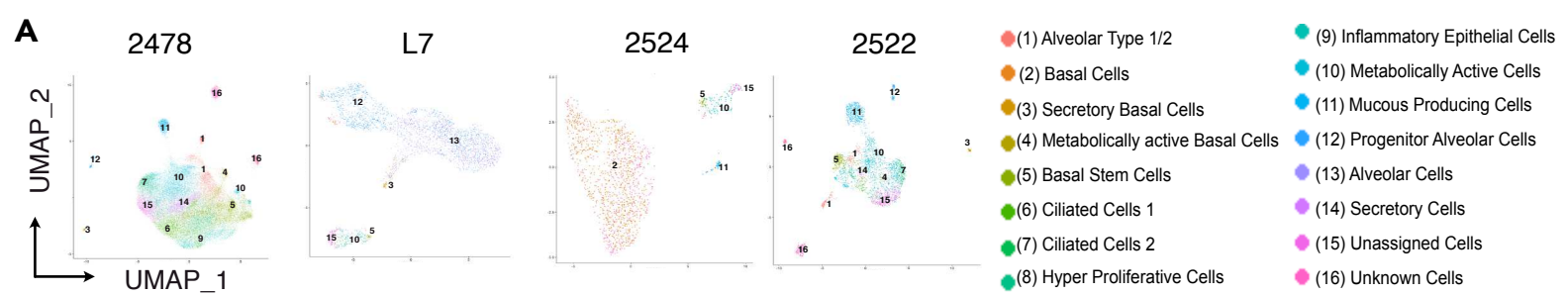




Expression



- (1) Alveolar Type 1/2
- (2) Basal Cells
- (3) Secretory Basal Cells
- (4) Metabolically active Basal Cells
- (5) Basal Stem Cells
- (6) Ciliated Cells 1
- (7) Ciliated Cells 2
- (8) Hyper Proliferative Cells
- (9) Inflammatory Epithelial Cells
- (10) Metabolically Active Cells
- (11) Mucous Producing Cells
- (12) Progenitor Alveolar Cells
- (13) Alveolar Cells
- (14) Secretory Cells
- (15) Unassigned Cells
- (16) Unknown Cells



Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Table2.xls](#)
- [Table3.xls](#)