

## **NF- $\kappa$ B inhibitor alpha has a cross-variant role during SARS-CoV-2 infection in ACE2-overexpressing human airway organoids**

Camille R. Simoneau<sup>1,2,14</sup>, Pei-Yi Chen<sup>1,14</sup>, Galen K. Xing<sup>3,4,14</sup>, Mir M. Khalid<sup>1</sup>, Nathan L. Meyers<sup>1</sup>, Jennifer M. Hayashi<sup>1</sup>, Taha Y. Taha<sup>1</sup>, Kristoffer E. Leon<sup>1,5</sup>, Tal Ashuach<sup>4</sup>, Krystal A. Fontaine<sup>1</sup>, Lauren Rodriguez<sup>6</sup>, Bastian Joehnk<sup>7</sup>, Keith Walcott<sup>7</sup>, Sreelakshmi Vasudevan<sup>8</sup>, Xiaohui Fang<sup>9</sup>, Mazharul Maishan<sup>9</sup>, Shawn Schultz<sup>4</sup>, Jeroen Roose<sup>10</sup>, Michael A. Matthay<sup>9</sup>, Anita Sil<sup>7</sup>, Mehrdad Arjomandi<sup>8,11</sup>, Nir Yosef<sup>4,12\*</sup>, and Melanie Ott<sup>1,3,13\*</sup>

<sup>1</sup>Gladstone Institute of Virology, San Francisco, CA, USA

<sup>2</sup>Biomedical Sciences Graduate Program, University of California San Francisco, San Francisco, CA, USA

<sup>3</sup>Chan-Zuckerberg Biohub, San Francisco, CA, USA

<sup>4</sup>Center for Computational Biology, University of California, Berkeley, Berkeley CA, USA

<sup>5</sup>Medical Scientist Training Program, University of California San Francisco, San Francisco, CA, USA

<sup>6</sup>ImmunoX CoLabs, University of California San Francisco, San Francisco, CA, USA

<sup>7</sup>Department of Microbiology and Immunology, University of California San Francisco, San Francisco, CA, USA

<sup>8</sup>Medical Service, San Francisco VA Healthcare System, San Francisco, CA, USA

<sup>9</sup>Department of Medicine and Department of Anesthesia, Cardiovascular Research Institute, University of California San Francisco, San Francisco, CA, USA

<sup>10</sup>Department of Anatomy, University of California San Francisco, San Francisco, CA, USA

<sup>11</sup>Division of Pulmonary and Critical Care, Department of Medicine, University of California San Francisco, San Francisco, CA, USA

<sup>12</sup>Department of Systems Immunology, Weizmann Institute of Science, Rehovot, Israel

<sup>13</sup>Department of Medicine, University of California, San Francisco, San Francisco, CA, USA.

<sup>14</sup>Equal contribution

\*Correspondence to [melanie.ott@gladstone.ucsf.edu](mailto:melanie.ott@gladstone.ucsf.edu) and [niryosef@berkeley.edu](mailto:niryosef@berkeley.edu)



## Abstract

As SARS-CoV-2 continues to spread worldwide, tractable primary airway cell models that accurately recapitulate the cell-intrinsic response to arising viral variants are needed. Here we describe an adult stem cell-derived human airway organoid model overexpressing the ACE2 receptor that supports robust viral replication while maintaining 3D architecture and cellular diversity of the airway epithelium. ACE2-OE organoids were infected with SARS-CoV-2 variants and subjected to single-cell RNA-sequencing. NF- $\kappa$ B inhibitor alpha was consistently upregulated in infected epithelial cells, and its mRNA expression positively correlated with infection levels. Confocal microscopy showed more I $\kappa$ B $\alpha$  expression in infected than bystander cells, but found concurrent nuclear translocation of NF- $\kappa$ B that I $\kappa$ B $\alpha$  usually prevents. Overexpressing a nondegradable I $\kappa$ B $\alpha$  mutant reduced NF- $\kappa$ B translocation and increased viral infection. These data demonstrate the functionality of ACE2-OE organoids in SARS-CoV-2 research and identify an incomplete NF- $\kappa$ B feedback loop as a rheostat of viral infection that may promote inflammation and severe disease.

## Introduction

Understanding how the SARS-CoV-2 virus replicates and causes damage in the upper and lower airways has been the focus of intense scientific research for the past 2 years. The virus replicates in airway epithelial cells, and its receptor ACE2 is expressed throughout the airways and lung, but less so from the airway to the alveolar space.<sup>1</sup> Notably, replication of the Omicron variant appears to be confined to the upper airway tract, a finding that correlates with less severe disease.<sup>2,3</sup>

Removal and neutralization of potentially harmful substances from inhaled air are main functions of the airway epithelium, which consists of multiple cell types. 1) Basal cells are the progenitor cells that differentiate into the epithelial cell types. 2) Club cells secrete the immunomodulatory club cell secretory protein and also fulfill regenerative functions. 3) Goblet cells secrete mucins onto the internal surface of the respiratory tract, thereby forming a liquid layer (termed mucus) to protect the underlying epithelium. 4) Ciliated cells are located across the apical surface and facilitate the movement of mucus across the airway tract.<sup>4</sup> In the lungs, alveolar epithelial cells (ATI and II cells) line the alveoli where gas exchange takes place.

The viral spike protein has been found in the airway epithelium and lung tissue of deceased COVID-19 patients.<sup>5</sup> In *ex vivo* infections, ciliated cells were identified as natural targets of SARS-CoV-2.<sup>6</sup> However, evidence of infection was also noted in secretory and basal cells, in both *in vivo* and *ex vivo* infections.<sup>7</sup> This includes viral budding in cells with secretory vesicles in infected human bronchial epithelial cells;<sup>8</sup> others found goblet cells infected in *in vitro* human airway epithelial cell cultures and, as a response to infection, showed increased mucus production.<sup>9</sup>

To recapitulate the complex cellularity of the airway epithelium and provide a physiological, yet workable model of the infection in humans, multiple stem-cell derived 3D organoid models have been applied to virology<sup>7,10-12</sup>. Organoids consist of multiple cell types grown in a 3D structure to mimic organ structure. They can be derived from embryonic stem cells (ESCs), pluripotent stem cells (iPSCs) or progenitor cells in adult tissues. SARS-CoV-2 replicates in iPSC- or ESC-derived human airway organoids (HAOs)<sup>13,14</sup> and adult stem-cell HAOs that encompass both airway and alveolar cells<sup>15,16</sup>. In the latter model, donor lungs are digested



and cultured in Matrigel and a medium that allows for the outgrowth of basal, club and goblet cells, which can then include ciliated cells with additional *ex vivo* differentiation steps.<sup>4,17</sup>

Differentiation, often performed at the air-liquid interface, produces a pseudostratified configuration that most closely resembles the *in vivo* situation, but is lengthy, requires large volumes of cells, and loses 3D structure. The undifferentiated organoids provide an intermediate model system that is more structured and heterogeneous than a cell line, yet less rigid than the fully differentiated model. These undifferentiated HAOs are tractable and easy to work with, while not being cancerously transformed and maintaining cellular diversity that resembles the airway epithelium. The undifferentiated organoid model has been used to study respiratory syncytial virus<sup>17</sup> and SARS-CoV-2<sup>6,18,19</sup>. However, work with the latter is plagued with variable, donor-dependent infection rates where most organoids carry less than 5% infected cells<sup>18,19</sup>, a major obstacle for further analysis (i.e., single-cell genomics).

Single-cell RNA sequencing (scRNA-seq) has provided important insight into the biology of SARS-CoV-2 and is particularly valuable in heterogeneous cell systems where it effectively separates infected and bystander responses across multiple cell types. ScRNA-seq data were generated throughout the pandemic in both clinical and laboratory settings to determine cell-intrinsic responses to the virus, map COVID-19 changes in the lung, and understand the tissue milieu of responses.<sup>20-22</sup> These experiments, for example, distinguished interferon-stimulated gene (ISG) signatures in COVID-19 patients with mild to moderate phenotypes from patients with severe COVID-19 illness.<sup>20</sup> At the same time, advances in computational modeling for single-cell genomics permit an integration of samples from multiple conditions or cohorts, in a manner that allows direct comparisons of cellular compositions and gene expression programs (viral and host)<sup>23</sup>.

SARS-CoV-2 is no longer a uniform strain as multiple variants have arisen, both globally and regionally. The D614G mutation arose early in 2020, and the beginning of 2021 was characterized by the rise and fall of additional common variants, such as Alpha and Beta, which spread globally and eventually became regionally dominant. Others, such as Epsilon, spread locally before being outcompeted by a new variant<sup>24</sup>. Delta dominated infections by late 2021, until it was overtaken by an ever-evolving cast of Omicron variants<sup>25,26</sup>. This

series of variants offers unique opportunities to delineate common and specific traits of SARS-CoV-2 infection and to better treat or predict clinical disease severity.

To address these challenges and opportunities for studying SARS-CoV-2 infection, we generated and explored an adult stem cell-derived undifferentiated HAO model. Our model was genetically modified with ACE2 overexpression to overcome obstacles with low infection rates and long and variable differentiation times. We studied the effects of infection in this system using a genetically diverse set of variants, and applied scRNA-seq to understand the cell-intrinsic response to SARS-CoV-2 in this primary cell model. We identified *NFKBIA*, a critical regulator of the NF- $\kappa$ B pathway, as a universal gene that marks highly infected cells across all viral variants, underscoring the recognized link between inflammation and SARS-CoV-2 infection that dominates acute respiratory distress syndrome and severe disease in the lung.

## Results

HAOs were grown from digested uninfected donor lungs, according to a published protocol (Fig. 1A).<sup>17</sup> As expected and confirmed by light sheet microscopy, the majority of cells (~80%) in one organoid expressed P63, the basal cell marker, but a minority (~20%) represented secretory cells, including MUC5AC<sup>+</sup> goblet and CC10<sup>+</sup> club cells, which aligns with published reports (Fig. 1B).<sup>17</sup> In the absence of cell differentiation, very few FOXJ1<sup>+</sup> ciliated cells were observed (not shown). Compared to primary lung epithelium, markers for basal cells (P63, KRT5) were statistically overrepresented in quantitative RT-PCR analysis of organoids, and those for ciliated cells (FOXJ1), goblet cells (MUC5AC), and club cells (CC10) were underrepresented (Fig. 1C). Expression of ACE2 mRNA was also higher in primary lung tissue than undifferentiated organoids, and no significant difference was noted in TMPRSS2 mRNA levels (Fig. 1D,E). No ACE2 protein was detected in the organoids by western blotting, which explains their poor infection efficiency in previous studies (Fig. 1F,G).<sup>18</sup>

To overcome these issues, organoids were transduced with a lentiviral vector expressing the *ACE2* open reading frame. ACE2-overexpressing organoids (ACE2-OE) were selected and maintained robust ACE2 levels over at least nine culture passages as shown by western blotting (Fig 1F, G). This was also confirmed by confocal immunofluorescence microscopy where ACE2 protein expression was visible on the surface of most

cells in the ACE2-OE organoids (Fig. 1H). ACE2 overexpression did not affect the ability of organoids to differentiate at the air-liquid interface into all four cell types similarly to wildtype (WT) organoids as assessed by RT-qPCR for epithelial cell markers (Fig. 1I).

Next, WT and ACE2-OE HAOs were infected with the SARS-CoV-2 WA1 ancestral strain at a multiplicity of infection (MOI) of 1. Culture supernatants were collected at 24 and 72 hours and subjected to plaque assays. ACE2 overexpression resulted in approximately two log higher infectious particle production at 24 hours and a log difference at 72 hours (Fig. 2A-C). As a second measure of active viral infection, infected organoids were subjected to confocal immunofluorescence microscopy after staining with an antibody against double-stranded (ds) RNA. This antibody specifically stains the RNA replication centers in infected cells containing positive-negative strand RNA hybrids.<sup>27</sup> dsRNA<sup>+</sup> cells in WT and ACE2-OE organoids had two- to threefold greater ACE2 expression than WT organoids (Fig. 2D,E). These results show a robust increase in infected cell numbers with ACE2 overexpression that may now allow downstream analysis of single cells.

To test this, WT and ACE2-OE, organoids were subjected to 10X scRNA-seq after infection with the SARS-CoV-2 WA1 strain. Raw sequencing data were aligned to the human genome with the SARS-CoV-2 genome appended, run through scVI<sup>28,23</sup>, and represented in a joint (batch-corrected) low dimensional space (see Methods). In that space, the WT infected (WT\_I), ACE2 uninfected (ACE2\_U) and WT uninfected (WT\_U) conditions clustered together, and a large subset of the ACE2 infected (ACE2\_I) cells clustered separately (Fig. 3A, B).

To further validate the presence of cell types in the organoids identified by microscopy, the Lung Cell Atlas<sup>29</sup> was used as a reference to assign cell types to our scRNA-seq data. In agreement with our previous results (Fig. 1), most cells were basal cells, and the remainder were respiratory goblet cells- a secretory cell type present in the lungs (Fig. 3E). Only a small number were identified as ciliated cells (Fig. 3E). Furthermore, this annotation revealed no significant difference in proportion of each cell type when ACE2 and WT organoids were compared, indicating that ACE2 overexpression does not change the cellular composition of organoids (Supp. Fig. 1A).

To further confirm that overexpression of *ACE2* did not change the biology of the organoids, a differential gene expression analysis was conducted between uninfected WT and *ACE2*-OE samples (Supp. Fig. 1B). *ACE2* was identified as the highest differentially expressed gene, and only four other genes (*C20orf85*, *BPIFA1*, *TMEM190*, *UTP14C*) were differentially expressed with a q-value less than 0.01. Considering SARS-CoV-2 RNA, we found expression primarily in cells from the *ACE2\_I* condition (Fig. 3C, Supp. Fig1C). These results are consistent with previous data (Fig. 2), showing that SARS-CoV-2 infection levels are higher in the *ACE2* condition than the WT. They further suggest that the difference in infection levels between the organoids is most likely due to *ACE2* overexpression and not altered abundance of other receptors or cofactors.

Considering only cells from the infected conditions (*ACE2\_I*, *WT\_I*), we observed a broad variation in the abundance of viral products, presumably reflecting different levels of infection. To distinguish cells that had productive SARS-CoV-2 infections from those that had little to no infection, we fitted a Gaussian Mixture Model to classify cells from the infected samples into high, low and no infection groups (accounting for possible ambient viral product; see Methods and Fig. 3D). As expected, most of successfully infected cells (low and high infection groups in our model) came from the *ACE2\_I* condition. Conversely, a subset of *ACE2\_I* cells from the “no infection” group of cells clustered together with cells from the uninfected samples (*ACE2\_U*, *WT\_U*), suggesting that some cells in the infected organoid were not only uninfected but also not influenced by the infection of other cells.

To determine whether the successfully infected cells were detecting the viral replication, we assessed their levels of interferon gene expression. We found no expression of IFN- $\alpha$ , - $\beta$ , or - $\gamma$  in our scRNA-seq data. This result is similar to data published from COVID-19 patient samples in which IFN- $\lambda$  is the primary response.<sup>30</sup> Interestingly, while we found IFN- $\lambda$  expression only in our infected samples, the highest amounts were observed in cells with low levels of infection, with substantially lower expression in cells from the no-infection or high-infection groups (Fig. 3F-G). While these data help to demonstrate that these organoids recapitulate the response of lungs during infection, they also suggest that the anti-viral IFN- $\lambda$  response is only active in cells with low-level infections and shrinks as the amounts of intracellular viral products increase.

Unlike the ACE2\_I sample, we found comparably little expression of interferon genes in the WT\_I sample (Fig. 3G). To further interrogate this, we looked at the levels of interferon-stimulated genes (ISGs) in the different groups of cells. Consistently, we found that ISGs were primarily expressed in cells with low levels of viral RNA and not in the other infection groups or in the WT cells (Fig. 3H-I). Finally, for an unbiased analysis, we examined the differential expression in infected and uninfected samples in each of our genetic backgrounds (ACE2 OE and WT). Using gene set enrichment analysis, we found strong enrichment for interferon and anti-viral response program in ACE2 case, and much less so in the WT sample (Supp Fig. 2A-B).

Our results suggest that the lack of viral replication in WT cells is not due to an immune response, but more likely is a result of limited viral entry and replication. It therefore supports the notion that the WT system is not a good model for studying cellular response to SARS-CoV-2 infection, as the infection level is too low to stimulate an immune response. Conversely, due to the higher levels of replication in the ACE2 organoids, a measurable cellular response in those cells, and no difference in baseline biology between the ACE2 and WT organoids, we chose the ACE2 OE condition for follow-up scRNA-seq experiments.

Next, we compared transcriptome responses at the single-cell level in a second set of ACE2-OE organoids infected with various viral variants. Besides WA1 (using similar conditions as above), we tested Alpha (B.1.1.7), two separate isolates of the Beta variant (B.1.351), and the California-resident Epsilon variant (B.1.429) to capture the emerging diversity in viral sequences at the time of our study. Irrespective of cell type, the cells from the infected organoids were well mixed, and most uninfected cells clustered separately (Fig. 4A), indicating no clear difference or outlier among the infections (Fig. 4B). Basal cells remained the overwhelming cell type in this experiment, with a smaller representation of goblet cells (Fig. 4C). We again use our GMM approach for classifying the level of infection within each cell (Fig. 4D). As in Figure 3, we saw activation of INF- $\lambda$  and ISGs across all samples infected by the variants but not in the uninfected sample (Fig. 4E-F).

In addition, all variants had similar infection levels with one of the Beta strains having a slightly higher level (Fig. 4J). However, a comparison of the replication dynamics over time of Beta B, Beta A and WA-1 showed no significant differences, suggesting that the higher levels of virus are due to variability and not to an

intrinsic difference in these viruses (Supp. Fig. 3). Our results suggest that the different strains have comparable effects on our organoid system.

To determine which genes were most consistently upregulated with infection across the variants, we computed the correlation between gene expression and viral product (over cells) for each variant. The genes were then ranked by their level of correlation, separately for each variant (Fig. 4G; Table S1). Interestingly, we found a high level of consistency among the resulting lists of top ranked genes, with *CXCL3* being the top ranked in four out of the five variants, *TNFAIP3* in the top three in four of the variants, and *NFKBIA* in the top five in all variants. *NFKBIA* had the highest expression of the top-correlated genes, and had high expression in the high infection cluster (Fig. 4H). This cluster also had the highest levels of SARS-COV-2 viral products (Fig. 4I). *NFKBIA* is a quintessential NF- $\kappa$ B response gene and encodes the I $\kappa$ B $\alpha$  protein, which in turn dampens NF- $\kappa$ B activity by preventing its translocation into the nucleus.<sup>31</sup> This usually generates a negative feedback loop where NF- $\kappa$ B activity is terminated or temporarily restrained by the upregulation of I $\kappa$ B $\alpha$ . The fact that *NFKBIA* is the top-upregulated gene in the analysis indicates strong activation of the NF- $\kappa$ B signaling pathway in infected cells across variants.

To investigate this finding in greater depth, we confirmed upregulation of the I $\kappa$ B $\alpha$  protein by confocal microscopy. For better resolution, we used A549 epithelial cells overexpressing ACE2 as they form a homogenous monolayer on the cover slips. Cells were infected with SARS-CoV-2 WA-1 and co-stained for dsRNA and I $\kappa$ B $\alpha$ . Cells were imaged, and the mean fluorescence intensities (MFI) of I $\kappa$ B $\alpha$  were compared in dsRNA<sup>+</sup> and dsRNA<sup>-</sup> cells. I $\kappa$ B $\alpha$  protein expression was enhanced in infected (dsRNA<sup>+</sup>) versus noninfected bystander (dsRNA<sup>-</sup>) cells, mirroring the induction of *NFKBIA* transcript levels (Fig. 5A). In a parallel experiment, we co-stained infected A549-ACE2 cells with antibodies against dsRNA and the p65/RELA subunit of NF- $\kappa$ B and quantified p65 nuclear localization. 75% of dsRNA<sup>+</sup> cells showed p65 nuclear localization, but only a few dsRNA<sup>-</sup> cells did (Fig. 5B, C). This confirms a strong activation of the NF- $\kappa$ B signaling pathway only in infected, and not in uninfected, cells, which explains the consequent upregulation of *NFKBIA* as a response gene. However, the resultant increase in I $\kappa$ B $\alpha$  activity appears insufficient to

downregulate or terminate NF- $\kappa$ B signaling in cells infected with SARS-CoV-2, as the majority of infected cells localize p65 in the cell nucleus.

The I $\kappa$ B $\alpha$  protein can be phosphorylated by upstream kinases in a process that destabilizes the inhibitor and facilitates p65 nuclear translocation and transcriptional activity.<sup>32</sup> To determine the role of I $\kappa$ B $\alpha$  in the infectious process, we stably overexpressed an I $\kappa$ B $\alpha$  construct in A549-ACE2 cells with the serine residues at positions 32 and 36 mutated to alanine to prevent phosphorylation.<sup>33</sup> This generates a “super-inhibitor” that is unresponsive to cell signals and is highly active in retaining p65 in the cytoplasm, preventing its transcriptional activity in the nucleus.<sup>33</sup> We confirmed the proper function of the mutated inhibitor by stimulating uninfected A549-ACE2 cells with TNF $\alpha$ , a strong activator of the NF- $\kappa$ B pathway through activation of relevant upstream kinases, leading to phosphorylation and degradation of I $\kappa$ B $\alpha$  and consequent nuclear NF- $\kappa$ B translocation.<sup>34</sup> In the parent cell line with endogenous I $\kappa$ B $\alpha$ , 100% of cells had nuclear translocation of p65 in response to TNF $\alpha$  (Fig. 5E). However, in cells expressing nondegradable I $\kappa$ B $\alpha$ , only 25% of cells showed p65 nuclear translocation, confirming that the super-inhibitor remained stable, keeping p65 in the cytoplasm. Co-staining of p65 and the HA-tagged I $\kappa$ B $\alpha$  mutant revealed that the remaining p65 translocation occurred in cells with low expression of the stable I $\kappa$ B $\alpha$  mutant (Fig 5 E).

When cells expressing the I $\kappa$ B $\alpha$  mutant were infected with SARS-CoV-2 WA1, infected cultures showed two-fold more dsRNA<sup>+</sup> cells than the parent cell line (Fig. 5F). In support of this model, we found significantly less p65 translocation in cultures overexpressing the I $\kappa$ B $\alpha$  mutant than in the parent cell line (Fig. 5G). However, I $\kappa$ B $\alpha$  expression (measured as MFI) was increased in dsRNA<sup>+</sup> cells in both mutant and WT I $\kappa$ B $\alpha$ -expressing cell lines, suggesting residual activation of the NF- $\kappa$ B signaling pathway despite the super-inhibitor and resulting upregulation of endogenous *NFKBIA* in infected cells in both conditions (Fig 5H).

## Discussion

Here we generated a genetically modified airway organoid system that overcomes natural variations in ACE2 levels and allows high infection efficiency without lengthy differentiation. We anticipate that the tripling of



infection rates for SARS-CoV-2 and its variants will also be observed for other ACE2-utilizing coronaviruses, such as SARS-CoV and hCoV-NL63. The overall increase in infected cells allowed meaningful single-cell transcriptomics analysis with sufficient infected versus uninfected bystander cell counts recorded for computational analysis. We found infection in three cell types: stem-like basal cells and secretory goblet and club cells, with high viral replication associated with infection level. Transcriptional profiling in all three cell types yielded common transcriptional changes as indicated by clustering according to infection status by different viral variants. *NFKBIA* stood out as a gene of interest in controlling SARS-CoV-2 infection across variants and a highly induced gene whose mRNA levels positively correlated with viral RNA levels. We confirmed that the encoded I $\kappa$ B $\alpha$  protein is expressed at increased levels in infected cells and that, despite this upregulation, NF- $\kappa$ B was found in the cell nucleus in infected, but not uninfected, cells as a sign of continuous activation of NF- $\kappa$ B signaling (Fig. 6).

A similar incomplete feedback loop in NF- $\kappa$ B control was previously seen for respiratory syncytial virus infection<sup>35,37,38</sup>. However, the cause and benefit of this phenotype for viral infections remain unknown. It could represent an ongoing ‘arms race’ between the host and virus, where the host upregulates antiviral NF- $\kappa$ B signaling, but continuous viral presence overwhelms the feedback loop despite high rates of *NFKBIA* transcript and I $\kappa$ B $\alpha$  protein production through continuous degradation of the I $\kappa$ B $\alpha$  inhibitor. These data are consistent with published reports showing high levels of expression of NF- $\kappa$ B signaling-related genes in SARS-CoV-2-infected cells<sup>35</sup>. The overexpression of a mutant I $\kappa$ B $\alpha$  protein that cannot be phosphorylated and consequently degraded and the resulting two log increase in virus production support this model and identify I $\kappa$ B $\alpha$  as a proviral factor (Fig. 6). This is corroborated by studies by Sunshine et al., who identified *NFKBIA* in a Perturb-Seq study as a host protein supporting viral replication.<sup>36</sup>

It remains to be determined what the relevant upstream and downstream signals of this proviral phenotype are and whether the continuous presence of NF- $\kappa$ B in the cell nucleus provides benefits for the SARS-CoV-2 lifecycle. Our findings indicate that after overexpression of the mutant I $\kappa$ B $\alpha$  protein, nuclear p65 levels were lower in infected cells, supporting a model in which I $\kappa$ B $\alpha$  promotes viral infection by constraining,



albeit not completely, the well-known antiviral activities of NF- $\kappa$ B. However, knockdown studies of p65 have pointed to a positive role of NF- $\kappa$ B signaling in SARS-CoV-2 infection that needs further mechanistic exploration.<sup>37</sup> The fact that nuclear p65 is a striking differentiator of infected from bystander cells in our confocal microscopy studies underscores a possible role of nuclear NF- $\kappa$ B in SARS-CoV-2 infection.

While the causes of the observed phenotype remain a matter of speculation, the consequences are more predictable: a constitutively active NF- $\kappa$ B signaling pathway is likely to turn infected cells in the airway into inflammatory “factories” with overproduction of pro-inflammatory cytokines and chemokines that may activate the immune system and contribute to the cytokine storm that is characteristic of severe COVID-19.<sup>38</sup> Indeed, we identify various TNF $\alpha$ -induced proteins, such as CXCL8 and other cytokines, as gene products correlated with increased infection (Table S1).

The organoid model presented here adds another tool to the studies of SARS-CoV-2 infection in primary airway cells. It overcomes the hurdle of donor-dependent variable infection rates<sup>18</sup> and presents a tractable intermediate between cell line studies and fully differentiated primary cell models. It also offers insight into the transcriptional reprogramming occurring in progenitor cells of the airways in response to SARS-CoV-2 infection that may contribute to post-acute sequelae of viral infection<sup>7,38,39</sup>. We confirm that, in non-ACE2-OE organoids, infection of the three predominant cell types occurs, albeit at a much lower level than in ACE2-OE organoids. Our data showing that ACE2-OE organoids recapitulate transcriptional changes seen *in vivo*, such as the induction of IFN- $\lambda$ , support the validity of the model.

As I $\kappa$ B $\alpha$  is a common factor in all variants tested, we envision universal application for our findings across variants. We speculate that expression of *NFKB1A* mRNA, I $\kappa$ B $\alpha$  protein or the degree of nuclear translocation of NF- $\kappa$ B may be markers of infection rates and possibly disease severity in infected individuals; these parameters may also distinguish more severe variants, such as Delta, from less pathogenic variants, such as Omicron. Caution is warranted when therapeutically interfering with NF- $\kappa$ B activity: inhibiting NF- $\kappa$ B signaling with drugs, such as bortezomib, a proteasome inhibitor that stabilizes the I $\kappa$ B $\alpha$  protein may enhance SARS-CoV-2 infection. In contrast, therapeutic interventions that destabilize the I $\kappa$ B $\alpha$  protein should block

viral infection but may also further exacerbate pro-inflammatory cytokine production in infected cells. Future studies will test these and other potential clinical applications of our findings.

## Materials and Methods

**Cell lines.** A549 cells were cultured in complete DMEM (DMEM with 10% FBS, 1% penicillin-streptomycin, 1% glutamine). A549-ACE2 cells were made as described<sup>40</sup> and cultured in complete DMEM supplemented with 10 µg/mL blasticidin.

A549-ACE2- IκBα (SS32,36AA) were transduced with a lentiviral vector expressing IκBα (SS32,36AA) and selected in complete DMEM with 2 µg/mL of puromycin and 10 µg/mL of blasticidin. pLVX-EF1α- IκBα (SS32,36AA)-IRES-Puro was cloned by PCR amplification of HA-tagged IκBα (SS32,36AA) from pCMV4-3 HA/IκBα (SS32,36AA) and Gibson assembly, pCMV4-3 HA/IκBα (SS32,36AA) was a gift from Warner Greene (Addgene plasmid #24143) and pLVX-EF1α-IRES-Puro was purchased from Clontech Laboratories (plasmid #631988).

293T-HA-R-Spondin1-Fc cells were purchased from Trevigen (Catalogue Number 3710-001-K) and cultured according to the manufacturer's protocol to generate conditioned medium of R-spondin-1. Briefly, cells were grown in selection growth media (DMEM with 10% FBS, 1% penicillin-streptomycin, 1% glutamine, and 100 mg/mL Zeocin) for >5 days until they were >90% confluent. Medium was replaced with organoid basal media (Advanced DMEM/F12 from Invitrogen supplemented with 1% penicillin-streptomycin, 1% Glutamax, and 10 mM HEPES). After 3 days, cell supernatant (i.e., R-spondin-1 conditioned medium) was collected, centrifuged at 3,000 x g for 15 min, filtered through a 0.22-µm filter, and frozen at -20°C in 10-mL aliquots. This process was repeated by adding fresh organoid basal medium to the cells and collecting supernatant after 4 days.

**Culture of human airway organoids.** To generate HAOs, input cells were used from whole-lung lavages or from digestion of whole lungs. HAOs were generated from these cells as described<sup>17,41</sup>. Briefly, single cells were suspended in 65% reduced growth factor BME2 (Basement Membrane Extract, Type 2, Trevigen,

catalogue number 3533-001-02). From this mixture, 50- $\mu$ L drops containing 1,000–40,000 cells were seeded in 24-well suspension culture plates (GreinerBio-one, catalogue number 662-102). Drops were incubated at 37°C for >20 min and solidified. After this, 500  $\mu$ L of HAO medium was added to each well. HAO medium is organoid basal media (Advanced DMEM/F12 supplemented with 1% penicillin-streptomycin, 1% Glutamax, and 10 mM HEPES) supplemented with 10% (vol/vol) R-spondin1 conditioned medium, 1% B27 (Gibco), 25 ng/mL Noggin (Peprotech), 1.25 mM N-acetylcysteine (Sigma-Aldrich), 10 mM nicotinamide (Sigma-Aldrich), 5 nM Heregulin beta-1 (Peprotech), and 100  $\mu$ g/mL of Primocin (InvivoGen). HAO medium was further supplemented with 5  $\mu$ M Y-27632, 500 nM A83-01, 500 nM SB202190, 25 ng/mL of FGF-7, and 100 ng/mL of FGF-10 (all from Stem Cell Technologies), and HAO medium was replaced every 3–4 days.

After 14–21 days, organoids were passaged. For this, cold basal medium was used to collect organoids in 15-mL Falcon tubes and dissolved in BME2. Tubes were centrifuged at 250 x g for 5 min at 4 °C. Medium was aspirated, 10x TrypLE Select (Gibco) was added to the organoids, and the mixture was incubated at 37°C for 5–10 min. Organoids were further dissociated by pipette mixing and then diluted in cold basal medium. After another spin and medium aspiration, cells were mixed with BME2 and seeded into new drops. After this initial passage, organoids were passaged every 10–21 days. Stocks of early-passage (P1-P3) organoid lines were prepared by dissociating organoids, mixing them with recovery cell culture-freezing medium (Gibco), and freezing them by standard procedures. These samples could be thawed and immediately cultured in HAO medium.

**Creation of HAO-ACE2 lines.** HAO-ACE2 lines stably expressing hACE2 were generated by lentiviral transduction with plasmid LV-ACE2<sup>42</sup>, and selected with 2  $\mu$ g/mL of blasticidin, as described in a study from our group.<sup>43</sup>

**Differentiation of HAOs at the air-liquid interface.** HAOs were cultured at the air-liquid interface (ALI) for differentiation. Organoids were collected in cold basal medium and dissociated as described above. After spin and medium aspiration, cells were resuspended in warm basal medium with 2% (v/v) BME2 and seeded onto

pre-coated 2% BME2 in 6.5-mm insert of a 24-well plate Transwell Permeable Support (Costar). Cells were incubated for 1 hour at 37°C and then 500 µL of warm HAO medium was added to the bottom wells of the 24-well plate. Cells were grown to confluency for 1 week before the ALI was established by removal of apical side medium and substitution of basolateral medium for 1:1 (v/v) HAO medium and with Pneumacult ALI basal medium (Stem Cell Technologies, 05002). Basolateral medium was changed every 3–4 days, and mucus was washed from the apical side. Cells were differentiated at the ALI for at least 1 month.

**Real-time quantitative PCR.** The Qiagen RNEasy and Zymo DirectZol RNA MiniPrep were used for RNA extraction and isolation according to manufacturer's instructions. Final RNA concentrations were measured with a NanoDrop ND-1000. Total RNA was reverse-transcribed using oligo(dT)<sub>18</sub> primers (Thermo Scientific), random hexamers primers (Thermo Scientific), and AMV reverse-transcriptase (Promega). cDNA was diluted to 5 ng/µL. Gene expression was assayed by real-time quantitative PCR using Maxima SYBR Green qPCR Master Mix (Thermo Scientific) on a Biorad C1000 real-time PCR system. The SYBR Green qPCR reactions contained 10 µL of 2x SYBR Green Master Mix, 2 µL of diluted cDNA, and 8 pmol of both forward and reverse primers. The reactions were run using the following conditions: 50°C for 2 min, 95°C for 10 mins, followed by 40 cycles of 95°C for 5 secs and 60°C for 30 secs. Relative values for each transcript were normalized to 18S rRNA. Gene primers used are listed in Table S2. For every qPCR run, three technical replicates per sample were used for each gene.

**Western blots.** Organoids were collected from culture plate with cold basal medium, washed, and centrifuged. The resulting cell pellet was re-suspended in 100 µL of RIPA buffer with HALT protease. Sample was incubated on ice for 30 min, before sonicating at 40% power for 10 sec (Sonic Dismembrator Model 500). Sonicated cells were spun down, and supernatants were collected as cell lysates.

Cell-lysate protein concentrations were determined by Bio-Rad DC Assay. For each sample, 16-20 µg of protein was loaded. Appropriate volume of cell lysate was lysed with Tris-glycine SDS sample buffer, heated at 95°C for 5 min. Lysate samples were immediately loaded onto 4–20% gradient gels (Bio-Rad, catalogue #

4561095) for SDS-PAGE for 90 min at 120V. Gel was wet-transferred onto a 0.45  $\mu$ m nitrocellulose membrane for 1 hour at 100V. Blots were blocked in 5% milk in 1x TBST solution for 1 hour in room temperature with gentle rocking, then incubated with primary antibodies in 5% milk in 1x TBST overnight at 4°C with gentle rocking. Blot was washed three times with 1xTBST and incubated with goat anti-rabbit and goat anti-mouse IgG HRP-conjugated secondary antibodies diluted at 1:5000 with 5% milk in 1xTBST for 1 hour at room temperature with gentle shaking. After washing the blots three times with 1xTBST, blots were developed with Lumi-Light Western Blotting Substrate (Roche) for 5 min in the dark and visualized on a chemiluminescence imager (Bio-Rad ChemiDoc MP).

**Whole-mount organoid staining.** Organoids were processed for imaging as described <sup>44</sup>. Briefly, organoids were removed from BME2 with 3x cold PBS washes then fixed in 2–3% paraformaldehyde for 30 min on ice and washed 3x in PBS. Fixed organoid samples were stored at 4°C for up to 2 months.

For staining, organoids were blocked in PBS supplemented with 0.5% Triton X-100, 1% DMSO, 1% BSA, and 1% donkey or goat serum. Organoids were blocked for several hours at room temperature. Blocking solution was removed and replaced with blocking solution containing primary antibodies diluted 1:500 or 1:250. Organoids were incubated with primary antibodies for 24 hours at 4°C. After this, organoids were washed 3x in PBS and incubated with secondary antibodies diluted 1:250 in PBS at room temperature for several hours. Organoids were washed 3x in PBS and stained with Hoescht before visualization. Organoids were imaged on both Zeiss Axio Observer Z.1 and Zeiss Lightsheet Z.1 microscopes. Images were processed using a combination of the Zeiss software, ImageJ 1.51f, and Imaris 9.3. Primary antibodies we used are listed in Table S2.

**Confocal imaging of HAOs.** HAOs were dissociated as described above. Organoids were resuspended in warm HAO medium with 2% BME2 and seeded onto glass microscopy chamber slides (Thermo Fisher, 154453) pre-coated with 1% (v/v) Geltrex (Gibco, A1413301). Organoids were allowed to embed in the Geltrex layer before aspirating the media and fixed in 4% paraformaldehyde for 30 min on ice and washed 3x in PBS.

For staining, organoids were permeabilized in 1% Triton X-100 for 10 min at room temperature. Organoids were then blocked in 100  $\mu$ L of blocking solution (5% donkey or goat serum, 1% BSA, 0.1% cold fish skin gelatin, 0.1% Triton X-100, 0.05% Tween 20) for 2 hours with gentle rocking at room temperature. Organoids were then incubated in primary antibodies diluted in antibody solution (3% serum, 0.1% cold fish skin gelatin, 0.3% Triton X-100, 0.05% Tween 20) overnight at 4°C with gentle rocking. After, organoids were washed three times in 1x PBS with 0.1% BSA, and incubated with secondary antibodies conjugated with fluorophores, diluted 1:400 in antibody solution, for 2 hours at room temperature with gentle rocking while protected from light. Organoids were washed three times in 1x PBS with 0.1% BSA and stained with Hoescht33342 for 10 min at room temperature. Polyester gasket was removed from the glass slide, and rectangular glass slide was mounted onto the HAOs with Gold antifade reagent (Invitrogen, P36934).

Images were taken on the Olympus FV3000RS confocal microscope and processed with Imaris 9.3.

**Infection of HAOs with SARS-CoV-2.** HAOs were dissociated and re-suspended to final density of 100,000 cells per 300  $\mu$ L of HAO medium and seeded onto culture plates pre-coated with BME2 at 37°C for 1 hour. The cells were allowed to settle and embed onto BME2 coating for 20-60 min at 37°C before adding HAO medium. 24 hours later, the organoids were infected at an MOI of 1 for 2 hours, and then then the virus-containing medium was removed, the cells were washed, fresh medium was added, and the cells were incubated until post-infection processing.

**Propagation of SARS-CoV-2 variants.** All SARS-CoV-2 variants were propagated on Vero-E6 cells expressing human TMPRSS2 in an aerosol biosafety level-3 lab (aBSL3) lab. Stocks were titered via plaque assay on Vero-E6 TMPRSS2 cells as described<sup>45</sup> and sequenced to confirm no novel cell-culture mutations.

**Single-cell RNA sequencing.** HAOs were infected with SARS-CoV-2 as described above for 24 hours. Organoids were removed from the plate by dissolving the geltrax with cold basal medium and spun down. A single-cell suspension was made by digesting the organoids with 10X TrypLE for 15 min and passing them

through a 40- $\mu$ M filter. Library prep was done according to the standard 10X protocol with 6,000 cells loaded into the 10X Chromium Controller.

Libraries were sequenced on a NovaSeq6000, and the RNA was aligned with Cell Ranger v.6.0.0 to the human GRCh38 reference (10X Genomics, STAR aligner with the complete SARS-CoV-2 genome (MN985325.1) appended).<sup>46</sup>

**Analysis of SARS-CoV-2 sequencing.** After alignment, cells that had >20% Unique Molecular Identifiers (UMIs) coming from mitochondrial genes were removed. Cells were also removed if they had fewer than 200 UMIs or greater than 300,000 UMIs. After filtering, the single-cell RNA-seq dataset contained 38462 cells and 36611 genes.

**scRNA-seq analysis with scVI.** We used the Seurat v3 method<sup>47</sup> as implemented in Scanpy v.1.8.1<sup>48</sup> to select the top 4000 highly variable genes from the dataset, excluding the SARS-CoV-2 genes and ACE2. Then we ran scVI<sup>28</sup> as implemented in scvi-tools v.0.13.0<sup>23</sup> with default parameters and with each sequencing round treated a batch covariate. This resulted in a single latent space for all the data. We visualized the data by running the Scanpy function `scanpy.pp.neighbor`, followed by `scanpy.tl.umap` for each round of sequencing independently on the scVI latent space.

**Cell-type annotation.** We first used seven different automated cell-type annotation methods to get a consensus prediction using the Lung Cell Atlas as a reference dataset.<sup>29</sup> The seven methods were: 1. OnClass.<sup>49</sup> 2. SVM.<sup>50</sup> 3. Random Forest (as implemented in sklearn).<sup>51</sup> 4. KNN after batch correction with BBKNN.<sup>52</sup> 5. KNN after batch correction with Scanorama.<sup>53</sup> 6. KNN trained on the Lung Cell Atlas<sup>29</sup> scVI latent space<sup>28</sup>. 7. scANVI.<sup>54</sup> The resulting prediction was the majority prediction over the seven methods. We also derived a consensus score, based on the number of methods that agreed with the consensus prediction. After manual inspection, we determined that the only cell types present in the dataset were either: “Basal Cells”, “Goblet Cells”, and “Lung Ciliated Cells”. For all cells with a consensus score less than 5 or were not predicted as a basal/goblet/ciliated



cell, we reclassified with a KNN trained on cells with consensus score greater than 5 and predicted as a basal/goblet/ciliated cell.

**Correlation analysis.** To identify the top genes correlated with increased SCV2 levels, we computed the Pearson correlation as implemented in SciPy v1.4.1 (`scipy.stats.pearsonr`). We calculated the correlation for each gene in each sample individually between the log-normalized gene counts as implemented in Scanpy<sup>48</sup> v.1.8.1 (`scanpy.pp.normalize_total(adata, target_sum=1e4)` followed by `sc.pp.log1p`) with the sum of all SCV2 RNA transcripts per sample. Then for each sample, we filtered for genes expressed in at least 20% of cells in the sample.

**DE testing and GO analysis.** We used DESeq2 (v1.34.0) to identify genes differentially expressed between WT HAOs and ACE2-overexpressed HAOs by pseudobulking the transcript counts of each sample.<sup>55</sup> We used Scanpy v.1.8.1 to identify differentially expressed genes between cells in the infected vs uninfected conditions.<sup>48</sup> Metascape (v 3.5) with default Express Analysis settings were used to identify enriched Gene Ontology Terms.<sup>56</sup>

**Classifying infection with Gaussian Mixture Models.** To classify whether a cell was actually infected with SCV2 (not just exposed to the virus), we used a Gaussian Mixture Model (GMM) to classify whether the SCV2 mRNA UMIs were due to background viral mRNA or from replicating virus within a cell. We used the scikit-learn v.1.0.1 GMM implementation (`sklearn.mixture.GaussianMixture`) with default parameters. Bayesian information criterion (BIC)<sup>57</sup> as implemented in scikit-learn was used for model selection; defined as  $BIC = -2\log(L) + \log(N)d$ , where L is the maximum likelihood of the GMM, N is the number of samples, and d is the degrees of freedom. For each experimental condition, we tested the optimal GMM for two vs three components as well as whether a single model should be trained on the log (sum of viral mRNA per cell + 1) or a separate model for each individual log(viral mRNA counts + 1) per viral gene. When training a separate model individually, we trained a GMM for each viral gene, summed the posterior probability of each component for



each model, and for each cell assigned the component with the greatest total posterior probability. The BIC was then calculated as the mean BIC for each individual model. For each experimental condition, we selected the model with the lowest BIC. In all the uninfected experimental conditions (cells not exposed to virus), the optimal GMM model was a two-component GMM trained on the log (sum of viral mRNA per cell + 1), while the optimal model for all the infected experimental conditions (cells exposed to virus), was a three-component GMM trained on the log(viral mRNA counts + 1) for each viral gene. For the two-component GMM, we interpreted the component with a lower mean as cells with “No Infection” while the higher component was “Yes Infection”. It is worth noting that all cells in the uninfected condition were classified as “No Infection,” which tracks with it being a negative control. For the three-component GMM, we interpreted the component with the lowest mean as “No Infection”, the middle mean as “Low Infection”, and the highest mean as “High Infection”.

## **Acknowledgements**

The following reagents were obtained through BEI Resources, NIAID, NIH: SARS-Related Coronavirus 2, Isolate USA-WA1/2020, NR-52281, SARS-Related Coronavirus 2, Isolate hCoV-19/South Africa/KRISP-EC-K005321/2020, NR-54008, contributed by Alex Sigal and Tulio de Oliveira (Beta\_A), SARS-Related Coronavirus 2, Isolate hCoV-19/South Africa/KRISP-K005325/2020, NR-54009, contributed by Alex Sigal and Tulio de Oliveira. (Beta\_B).

We thank Veronica Fonseca for administrative support, Mauricio Montano for BSL-3 laboratory support, Sara Gardner and John Carroll for graphics design, Dr. Norma Neff and Dr. Amy Kistler for sequencing support, the California Department of Public Health, especially Dr. Mary Kate Morris and Dr. Carl Hansen, as well as Dr. Charles Chiu, Dr. Raul Andino and Dr. Miguel Garcia Knight for the Epsilon variant and Alpha variant. We thank Adam Gayoso for feedback about single cell data analysis. We thank members from the laboratories of Melanie Ott, Nir Yosef, and Aaron Streets for general feedback. We gratefully acknowledge support from the CZ Biohub, the Innovative Genomics Institute and the Pendleton Foundation (M.O.), and Administrative

COVID-19 Supplement 3P01AI091580-09S1 (J.P.R) on the parent NIH/NIAID P01-AI091580 (Weiss). SV and MA efforts were funded by the California Tobacco-related Disease Research Program (28FT-0020 to SV and T29IR0715 to MA) and the Departments of Defense (W81XWH-20-1-0158 to MA) and Veterans Affairs (CXV-00125 to MA).

## Author Contributions

CRS, P-YC, and GKK contributed equally. CRS: Conceptualization, Validation, Formal Analysis, Investigation, Writing-Original Draft, Writing-Review and Edit; P-YC: Investigation, Validation, Writing-Original Draft, GKK: Software, formal analysis, Writing-Original Draft, Writing-Review and Edit; MMK, JMH, TYT, KEL, SV, XF, MM: Investigation; TA, SS: Software; NLM, KAF: Conceptualization, Investigation; LR, BJ, KW: Resources; JR, MAM, AS, SV, MA: Resources, Supervision; NY, MO: Conceptualization, Funding acquisition, Supervision, Writing-Review and Edit.

## Figure legends

**Figure 1: Characterization of human airway organoids (HAOs) and overexpression of ACE2** (A) Schematic detailing the establishment of organoid culture and ACE2 overexpression. (B) Light sheet microscopy of HAOs with staining for the cell markers CC10 (club cells) MUC5AC (goblet cells) and p63 (basal cells) with a brightfield image of a single organoid in the lower right-hand corner. (C) RT- qPCR comparison of cell-marker gene expression between digested primary lung tissue and organoids, (N = 3, organoids, N =4, input, +/- SD shown). (D) RT-qPCR comparing ACE2 expression level between organoids and digested lung tissue. (E) RT-qPCR comparing TMPRSS2 expression level between organoids and digested lung tissue (N = 3, organoids, N =4, input, +/- SD shown). (F) Representative western blot showing ACE2 protein levels in organoids at baseline and with ACE2 overexpression. (G) Quantification of ACE2 protein levels in wildtype organoids and organoids overexpressing ACE2. (H) Confocal imaging comparing ACE2 expression levels in wildtype and ACE2-overexpressing organoids. (I) RT-qPCR comparison of cell-marker

genes after differentiation at the air-liquid interface of wildtype and ACE2 overexpressing organoids (N = 2, +/- SD shown).

**Figure 2: Infection of HAOs with SARS-CoV-2** Organoids were infected at an MOI of 1 for 2 hours, and then washed and fresh medium added. Organoids were incubated for the length of time indicated (A) Plaque assay comparing wildtype and ACE2 overexpressing organoids at 24 hours post-infection. (B) Plaque assay comparing wildtype and ACE2 overexpressing organoids at 72 hours post-infection. (C) Representative images of plaque assays performed on supernatant from SARS-CoV-2 infection of wildtype and ACE2 overexpressing organoids. (D) Representative images of infected HAOs at 24 and 72 hours stained for dsRNA (green). (E) Quantification of percent of total DAPI+ cells expression dsRNA across three experiments at 24 and 72 hours.

**Figure 3: Single-cell RNA-sequencing of WT and ACE2 OE organoids infected with SARS-CoV-2** (A) UMAP colored by overexpression condition of organoids. (B) UMAP colored by overexpression condition and infection condition. (C) UMAP colored by log expression of SARS-CoV-2 RNA. (D) UMAP colored by GMM infection classification. (E) UMAP colored by cell type identification. (F) Dot plot showing levels of interferon genes and IL-6 RNA grouped by infection level. (G) Dot plot showing levels of interferon genes and IL-6 grouped by overexpression condition. (H) Dot plot showing levels of top differentially expressed ISGs grouped by infection level. ISGs were selected from the list of top 50 differentially expressed genes between the high-infected and uninfected cells. (I) Dot plot showing levels of top differentially expressed ISGs grouped by overexpression background.

**Figure 4: Single-cell RNA-sequencing of airway organoids infected with SARS-CoV-2 variants** (A) UMAP showing distribution of cells from each variant infection. (B) UMAP colored by infected and uninfected conditions. (C) UMAP colored by cell type (D) UMAP colored by SARS-CoV-2 infection level (via GMM classification). (E) Dot plot showing levels of interferon genes and IL-6 grouped by variant. (F) Dot plot showing levels of ISG expression grouped by variant. (G) Heatmap showing top five genes correlated to SARS-

CoV-2 RNA level grouped by variant. The number in the box shows the ranking of the gene based on correlation for the variant and coloring shows the mean expression of the gene listed on the left. (H) UMAP colored by log expression of *NFKBIA* RNA. (I) UMAP colored by log expression of SARS-CoV-2 RNA. (J) Infection level proportions per variant.

**Figure 5: Role of I $\kappa$ B $\alpha$  in SARS-CoV-2 infection** (A) Quantification of p65 translocation in SARS-CoV-2 infected dsRNA<sup>+</sup> and dsRNA<sup>-</sup> A549-ACE2 cells (n = 2) (B) Representative image of p65 translocation (C) Quantification of MFI of I $\kappa$ B $\alpha$  in dsRNA<sup>+</sup> and dsRNA<sup>-</sup> cells (D) Representative image of I $\kappa$ B $\alpha$  and dsRNA staining in SARS-CoV-2 infected A549-ACE2 cells (E) Quantification of p65 translocation in cells expressing phosphorylation resistant I $\kappa$ B $\alpha$  after stimulation with TNF $\alpha$  and cells with functional I $\kappa$ B $\alpha$ . Three replicates were done, and a minimum of 5,000 cells were analyzed per condition; standard deviation is indicated (F) Representative images of DAPI, p65 and HA staining after TNF $\alpha$  stimulation (G) Quantification of dsRNA in mutant and non-mutant I $\kappa$ B $\alpha$  A549-ACE2 lines (H) Quantification of dsRNA in mutant and non-mutant I $\kappa$ B $\alpha$  A549-ACE2 lines. Three replicates were done, and a minimum of 5,000 cells were analyzed per condition; standard deviation is indicated (I) Quantification of p65 translocation in mutant and non-mutant I $\kappa$ B $\alpha$  A549-ACE2 lines (J) Quantification of I $\kappa$ B $\alpha$  MFI in mutant and non-mutant I $\kappa$ B $\alpha$  A549-ACE2 lines.

**Figure 6: Model of incomplete NF- $\kappa$ B feedback loop in SARS-CoV-2-infected airways cells.** While cells with endogenous I $\kappa$ B $\alpha$  maintain active NF- $\kappa$ B transcription factors in the cell nucleus and viral replication is attenuated (left), the negative feedback loop is strengthened with overexpression of an I $\kappa$ B $\alpha$  mutant, a process that enhances production of infectious SARS-CoV-2.

**Supplemental Figure 1: Related to Figure 3. Similarities and differences in ACE2 OE and WT organoids**

(A) Proportion of goblet and basal cells in the ACE2 OE and WT backgrounds. Small numbers of ciliated cells (<25) were removed. (B) DEseq2 differential expression results of ACE2 vs WT cells. Highlighted cells have a

q>0.01 (C) GMM infection classification of cells in the infected condition, broken down by cell type and genetic background. (D) Values of GMM infection classification of cells in the infected condition. (E) Proportion of goblet and basal cells in the infected vs uninfected conditions. (F) Proportion of goblet and basal cells in each of the classified infection levels (from GMM) and the uninfected condition.

**Supplemental Figure 2: Related to Figure 3. Gene Ontology enrichment analysis.** Metascape analysis showing top functional processes of genes differentially upregulated in cells in the infected vs uninfected conditions in an (A) ACE2 overexpression and (B) WT background.

**Supplemental Figure 3: Replication dynamics of Beta variants** (A) Quantification of dsRNA<sup>+</sup> cells in A549-ACE2 at the indicated time in Beta variants and WA-1. (N =3, MOI = 1).

**Supplemental Figure 4: Related to Figure 3: Expression of viral entry factors.** Violin plots of SCV2 entry factors between ACE2 OE and WT HAOs in the infected and uninfected conditions. Plots show log normalized expression of (A) ACE2, (B) BSG, (C) NRP1, (D) TMPRSS2, (E) TMPRSS4, and (F) TSPAN8.

## References

- 1 Hou, Y. J. *et al.* SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. *Cell* **182**, 429-446 e414, doi:10.1016/j.cell.2020.05.042 (2020).
- 2 Zhao, H. *et al.* Fusion-inhibition peptide broadly inhibits influenza virus and SARS-CoV-2, including Delta and Omicron variants. *Emerg Microbes Infect* **11**, 926-937, doi:10.1080/22221751.2022.2051753 (2022).
- 3 Meng, B. *et al.* Altered TMPRSS2 usage by SARS-CoV-2 Omicron impacts infectivity and fusogenicity. *Nature* **603**, 706-714, doi:10.1038/s41586-022-04474-x (2022).
- 4 Davis, J. D. & Wypych, T. P. Cellular and functional heterogeneity of the airway epithelium. *Mucosal Immunol* **14**, 978-990, doi:10.1038/s41385-020-00370-7 (2021).
- 5 Borczuk, A. C. *et al.* COVID-19 pulmonary pathology: a multi-institutional autopsy cohort from Italy and New York City. *Mod Pathol* **33**, 2156-2168, doi:10.1038/s41379-020-00661-1 (2020).
- 6 Lamers, M. M. *et al.* SARS-CoV-2 productively infects human gut enterocytes. *Science* **369**, 50-54, doi:10.1126/science.abc1669 (2020).
- 7 Simoneau, C. R. & Ott, M. Modeling Multi-organ Infection by SARS-CoV-2 Using Stem Cell Technology. *Cell Stem Cell* **27**, 859-868, doi:10.1016/j.stem.2020.11.012 (2020).
- 8 Robinot, R. *et al.* SARS-CoV-2 infection induces the dedifferentiation of multiciliated cells and impairs mucociliary clearance. *Nat Commun* **12**, 4354, doi:10.1038/s41467-021-24521-x (2021).
- 9 Morrison, C. B. *et al.* SARS-CoV-2 infection of airway cells causes intense viral and cell shedding, two spreading mechanisms affected by IL-13. *Proc Natl Acad Sci U S A* **119**, e2119680119, doi:10.1073/pnas.2119680119 (2022).
- 10 Ramani, S., Crawford, S. E., Blutt, S. E. & Estes, M. K. Human organoid cultures: Transformative New Tools for Human Virus Studies. *Curr Opin Virol* **29**, 79-86, doi:10.1016/j.coviro.2018.04.001 (2018).
- 11 Bar-Ephraim, Y. E., Kretzschmar, K. & Clevers, H. Organoids in immunological research. *Nat Rev Immunol* **20**, 279–293, doi:10.1038/s41577-019-0248-y (2019).

- 12 van der Vaart, J., Lamers, M. M., Haagmans, B. L. & Clevers, H. Advancing lung organoids for COVID-19 research. *Dis Model Mech* **14**, doi:10.1242/dmm.049060 (2021).
- 13 Han, Y. *et al.* Identification of SARS-CoV-2 Inhibitors using Lung and Colonic Organoids. *Nature*, doi:10.1038/s41586-020-2901-9 (2020).
- 14 Samuel, R. *et al.* Androgen Regulates SARS-CoV-2 Receptor Levels and Is Associated with Severe COVID-19 Symptoms in Men. *Cell Stem Cell* (2020).
- 15 Salahudeen, A. A. *et al.* Progenitor identification and SARS-CoV-2 infection in human distal lung organoids. *Nature* **588**, 670-675, doi:10.1038/s41586-020-3014-1 (2020).
- 16 Tindle, C. *et al.* Adult stem cell-derived complete lung organoid models emulate lung disease in COVID-19. *Elife* **10**, doi:10.7554/eLife.66417 (2021).
- 17 Sachs, N. *et al.* Long-term expanding human airway organoids for disease modeling. *EMBO J* **38**, doi:10.15252/emj.2018100300 (2019).
- 18 Hysenaj, L. *et al.* SARS-CoV-2 infection studies in lung organoids identify TSPAN8 as novel mediator. *bioRxiv*, 2021.2006.2001.446640, doi:10.1101/2021.06.01.446640 (2021).
- 19 Ekanger, C. T. *et al.* Human Organotypic Airway and Lung Organoid Cells of Bronchiolar and Alveolar Differentiation Are Permissive to Infection by Influenza and SARS-CoV-2 Respiratory Virus. *Front Cell Infect Microbiol* **12**, 841447, doi:10.3389/fcimb.2022.841447 (2022).
- 20 Combes, A. J. *et al.* Global absence and targeting of protective immune states in severe COVID-19. *Nature* **591**, 124-130, doi:10.1038/s41586-021-03234-7 (2021).
- 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* **27**, 890-904 e898, doi:10.1016/j.stem.2020.10.005 (2020).
- 22 Stephenson, E. *et al.* Single-cell multi-omics analysis of the immune response in COVID-19. *Nat Med* **27**, 904-916, doi:10.1038/s41591-021-01329-2 (2021).
- 23 Gayoso, A. *et al.* A Python library for probabilistic analysis of single-cell omics data. *Nat Biotechnol* **40**, 163-166, doi:10.1038/s41587-021-01206-w (2022).



- 24 Siddiqui, S., Alhamdi, H. W. S. & Alghamdi, H. A. Recent Chronology of COVID-19 Pandemic. *Front Public Health* **10**, 778037, doi:10.3389/fpubh.2022.778037 (2022).
- 25 Tian, D., Sun, Y., Zhou, J. & Ye, Q. The Global Epidemic of the SARS-CoV-2 Delta Variant, Key Spike Mutations and Immune Escape. *Front Immunol* **12**, 751778, doi:10.3389/fimmu.2021.751778 (2021).
- 26 Callaway, E. Why does the Omicron sub-variant spread faster than the original? *Nature* **602**, 556-557, doi:10.1038/d41586-022-00471-2 (2022).
- 27 Weber, F., Wagner, V., Rasmussen, S. B., Hartmann, R. & Paludan, S. R. Double-stranded RNA is produced by positive-strand RNA viruses and DNA viruses but not in detectable amounts by negative-strand RNA viruses. *J Virol* **80**, 5059-5064, doi:10.1128/JVI.80.10.5059-5064.2006 (2006).
- 28 Lopez, R., Regier, J., Cole, M. B., Jordan, M. I. & Yosef, N. Deep generative modeling for single-cell transcriptomics. *Nat Methods* **15**, 1053-1058, doi:10.1038/s41592-018-0229-2 (2018).
- 29 Travaglini, K. J. *et al.* A molecular cell atlas of the human lung from single-cell RNA sequencing. *Nature* **587**, 619-625, doi:10.1038/s41586-020-2922-4 (2020).
- 30 Park, A. & Iwasaki, A. Type I and Type III Interferons - Induction, Signaling, Evasion, and Application to Combat COVID-19. *Cell Host Microbe* **27**, 870-878, doi:10.1016/j.chom.2020.05.008 (2020).
- 31 Yu, H., Lin, L., Zhang, Z., Zhang, H. & Hu, H. Targeting NF-kappaB pathway for the therapy of diseases: mechanism and clinical study. *Signal Transduct Target Ther* **5**, 209, doi:10.1038/s41392-020-00312-6 (2020).
- 32 Wang, X. *et al.* Post-translational Modifications of IkappaBalpha: The State of the Art. *Front Cell Dev Biol* **8**, 574706, doi:10.3389/fcell.2020.574706 (2020).
- 33 Sun, S., Elwood, J. & Greene, W. C. Both amino- and carboxyl-terminal sequences within I kappa B alpha regulate its inducible degradation. *Mol Cell Biol* **16**, 1058-1065, doi:10.1128/MCB.16.3.1058 (1996).
- 34 Hayden, M. S. & Ghosh, S. Regulation of NF-kappaB by TNF family cytokines. *Semin Immunol* **26**, 253-266, doi:10.1016/j.smim.2014.05.004 (2014).



- 35 Fiedler, M. A. & Wernke-Dollries, K. Incomplete regulation of NF-kappaB by IkappaBalpha during respiratory syncytial virus infection in A549 cells. *J Virol* **73**, 4502-4507, doi:10.1128/JVI.73.5.4502-4507.1999 (1999).
- 36 Sunshine, S. *et al.* Systematic functional interrogation of SARS-CoV-2 host factors using Perturb-seq. *bioRxiv*, 2022.2007.2015.500120, doi:10.1101/2022.07.15.500120 (2022).
- 37 Nilsson-Payant, B. E. *et al.* The NF-kappaB Transcriptional Footprint Is Essential for SARS-CoV-2 Replication. *J Virol* **95**, e0125721, doi:10.1128/JVI.01257-21 (2021).
- 38 Ragab, D., Salah Eldin, H., Taeimah, M., Khattab, R. & Salem, R. The COVID-19 Cytokine Storm; What We Know So Far. *Front Immunol* **11**, 1446, doi:10.3389/fimmu.2020.01446 (2020).
- 39 Li, S. *et al.* When stem cells meet COVID-19: recent advances, challenges and future perspectives. *Stem Cell Res Ther* **13**, 9, doi:10.1186/s13287-021-02683-1 (2022).
- 40 Wang, R. *et al.* Genetic Screens Identify Host Factors for SARS-CoV-2 and Common Cold Coronaviruses. *Cell* **184**, 106-119 e114, doi:10.1016/j.cell.2020.12.004 (2021).
- 41 Zhou, J. *et al.* Differentiated human airway organoids to assess infectivity of emerging influenza virus. *Proc Natl Acad Sci U S A* **115**, 6822-6827, doi:10.1073/pnas.1806308115 (2018).
- 42 Khanna, K. *et al.* Thiol drugs decrease SARS-CoV-2 lung injury in vivo and disrupt SARS-CoV-2 spike complex binding to ACE2 in vitro. *bioRxiv*, doi:10.1101/2020.12.08.415505 (2021).
- 43 Deng, X. *et al.* Transmission, infectivity, and neutralization of a spike L452R SARS-CoV-2 variant. *Cell* **184**, 3426-3437 e3428, doi:10.1016/j.cell.2021.04.025 (2021).
- 44 Broutier, L. *et al.* Culture and establishment of self-renewing human and mouse adult liver and pancreas 3D organoids and their genetic manipulation. *Nat Protoc* **11**, 1724-1743, doi:10.1038/nprot.2016.097 (2016).
- 45 Perez-Bermejo, J. A. *et al.* SARS-CoV-2 infection of human iPSC-derived cardiac cells reflects cytopathic features in hearts of patients with COVID-19. *Sci Transl Med* **13**, doi:10.1126/scitranslmed.abf7872 (2021).

- 46 Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21, doi:10.1093/bioinformatics/bts635 (2013).
- 47 Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902 e1821, doi:10.1016/j.cell.2019.05.031 (2019).
- 48 Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biol* **19**, 15, doi:10.1186/s13059-017-1382-0 (2018).
- 49 Wang, S. *et al.* Leveraging the Cell Ontology to classify unseen cell types. *Nat Commun* **12**, 5556, doi:10.1038/s41467-021-25725-x (2021).
- 50 Cortes, C. & Vapnik, V. Support-vector networks. *Machine Learning* **20**, 273-297, doi:10.1007/BF00994018 (1995).
- 51 Breiman, L. Arcing Classifiers. *The Annals of Statistics* **26**, 801-824 (1998).
- 52 Polanski, K. *et al.* BBKNN: fast batch alignment of single cell transcriptomes. *Bioinformatics* **36**, 964-965, doi:10.1093/bioinformatics/btz625 (2020).
- 53 Hie, B., Bryson, B. & Berger, B. Efficient integration of heterogeneous single-cell transcriptomes using Scanorama. *Nat Biotechnol* **37**, 685-691, doi:10.1038/s41587-019-0113-3 (2019).
- 54 Xu, C. *et al.* Probabilistic harmonization and annotation of single-cell transcriptomics data with deep generative models. *Mol Syst Biol* **17**, e9620, doi:10.15252/msb.20209620 (2021).
- 55 Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 550, doi:10.1186/s13059-014-0550-8 (2014).
- 56 Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat Commun* **10**, 1523, doi:10.1038/s41467-019-09234-6 (2019).
- 57 Schwarz, G. Estimating the Dimension of a Model. *The Annals of Statistics* **6**, 461-464, 464 (1978).

**Supplementary Table 1:** Top 25 genes correlated with SCV2 levels per variant.

	Alpha	Beta_A	Beta_B	Epsilon	WA-1
0	CXCL3	CXCL3	NFKBIA	CXCL3	DRD5
1	TNFAIP3	ICAM1	PLAU	TNFAIP3	CD33
2	NFKBIA	NFKBIA	TNFAIP3	PLAU	AC103808.1
3	CXCL2	AC091544.4	CXCL3	NFKBIA	AL357873.1
4	DUSP8	TNFAIP3	EGR1	CXCL2	LRRIQ4
5	PLAU	UBASH3A	CXCL8	CXCL8	LINC02284
6	CXCL8	NRAP	AC127526.5	KLHL5	BEST3
7	KLF6	AC010624.3	NCOA7	DUSP8	U91328.3
8	ICAM1	AC008115.4	EDN1	CXCL1	AC084816.1
9	CXCL1	DUSP8	HES1	NFKBIZ	CXCL3
10	KLHL5	AL450270.1	KLF6	LCN2	TNFAIP3
11	EDN1	CXCL1	CXCL2	EGR1	ICAM1
12	BIRC3	CXCL8	NFKBIZ	TM4SF1	NFKBIA
13	NFKBIZ	EDN1	JUN	KLF6	EGR1
14	STX19	KLHL5	CXCL1	PIM3	LINC01354
15	CCN1	NR1D1	AP003498.2	ICAM1	AL139118.1
16	EGR1	NFKBIZ	DUSP8	AC127526.5	AP001001.1
17	TMEM156	CXCL2	ARRDC3	AP003498.2	RAB40AL
18	PIM3	BHLHE41	ICAM1	AC055720.1	AP006261.1
19	AC127526.5	EGR1	TM4SF1	AC010255.1	AC032044.1
20	IL6	BIRC3	EFNA1	LINC01814	NXP3
21	IL32	PLAU	ZC3H12A	AC127537.1	MYRFL
22	CD83	KLF6	AL627171.2	EDN1	AC022509.2
23	NCOA7	HIST1H2BD	KLHL5	AC005062.1	TNF
24	DUSP10	IPCEF1	MT-ND3	SFRP4	AC021483.2

## Supplementary Table 2: qPCR primers

Gene	Forward	Reverse	Probe
FOXJ1	TCGTATGCCACGCTCATCTG	CGGATTGAATTCTGCCAGG T	
MUC5AC	CCTACAAAGCTGAGGCCTGT	GACCCTCCTCTCAATGGTGC	
CC10	TCCTCCACCATGAAACTCGC	AGGAGGGTTTCGATGACAC G	
p63	ACCTGGAAAACAATGCCCAGA	ACGAGGAGCCGTTCTGAAT C	
KRT5	GCATCACCGTTCCTGGGTAA	GACACACTTGACTGGCGAG A	
ACE2	CGAGTGGCTAATTTGAAACCAAG AA	ATTGATACGGCTCCGGGAC A	
TMPRSS2	ATGGCTTTGAACTCAGGGTC	TTAGCCCGTCTGCCCTCATT T	
SARS-CoV-2 N	AAATTTTGGGGACCAGGAAC	TGGCACCTGTGTAGGTCAA C	ATGTCGCGCATTGGCATGGA
SARS-CoV-2 E	ACAGGTACGTTAATAGTTAATAG CGT	ATATTGCAGCAGTACGCAC ACA	ACACTAGCCATCCTTACTGCGCT TCG

## Figure 1

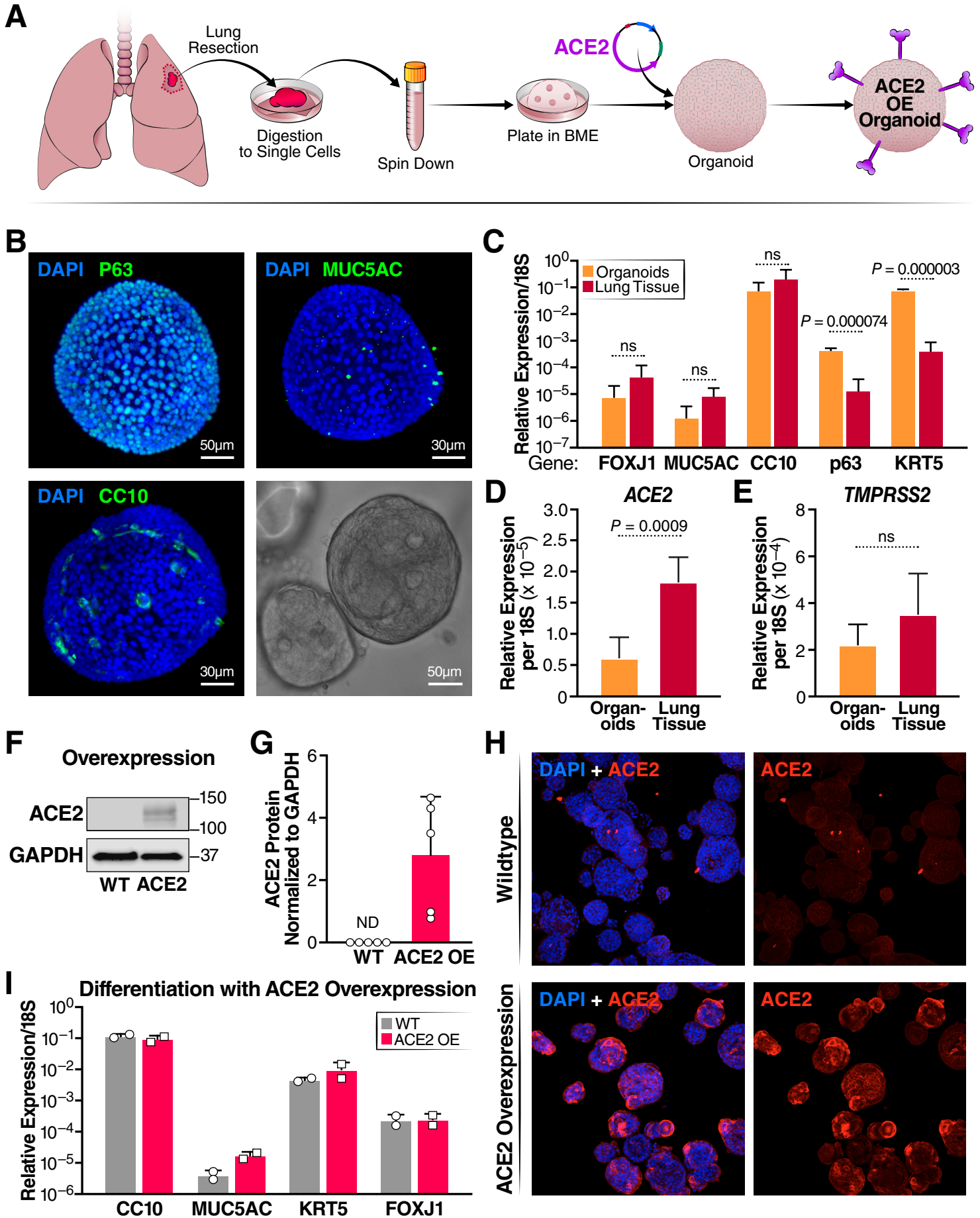
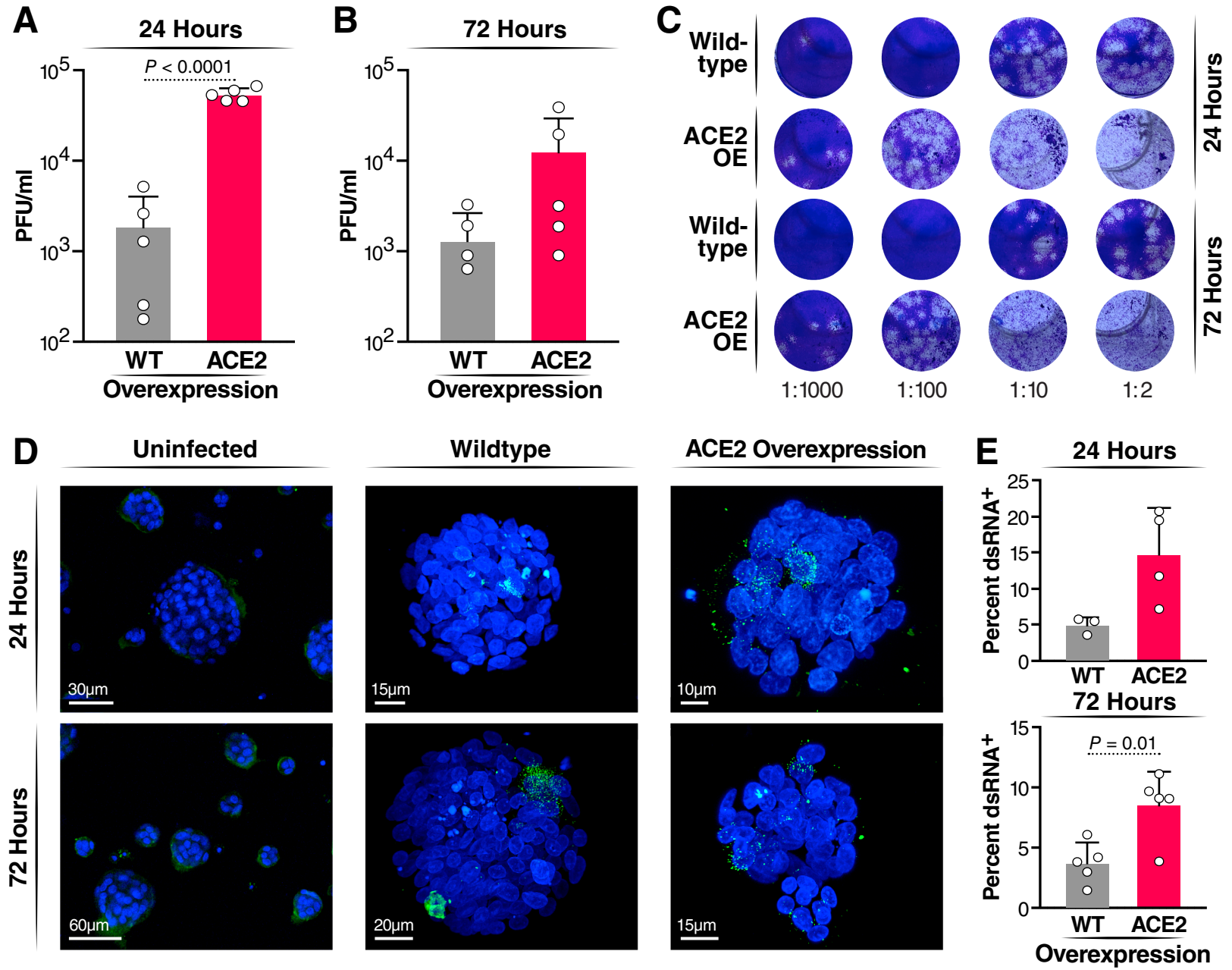
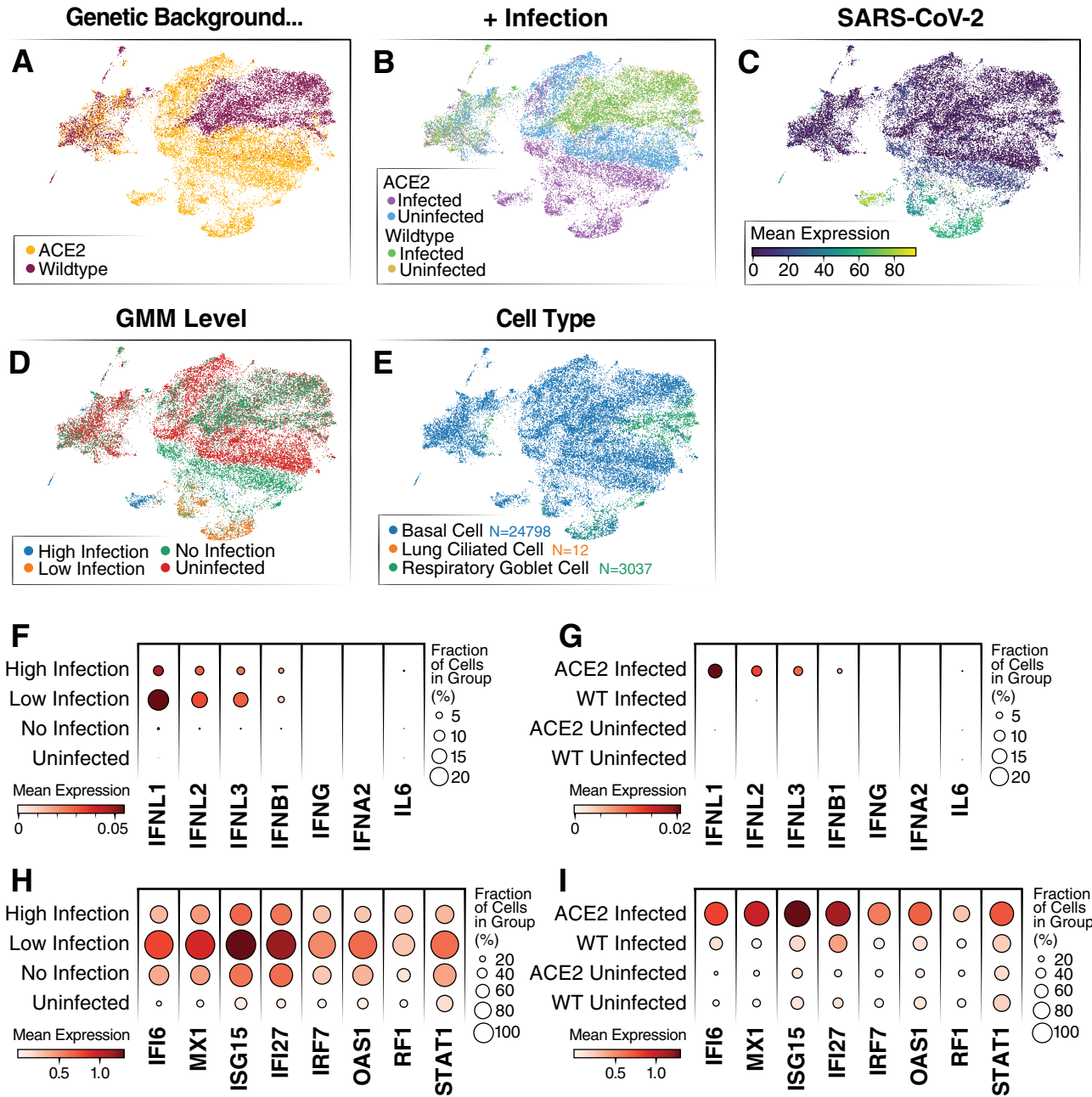


Figure 2





**Figure 3**



**Figure 4**

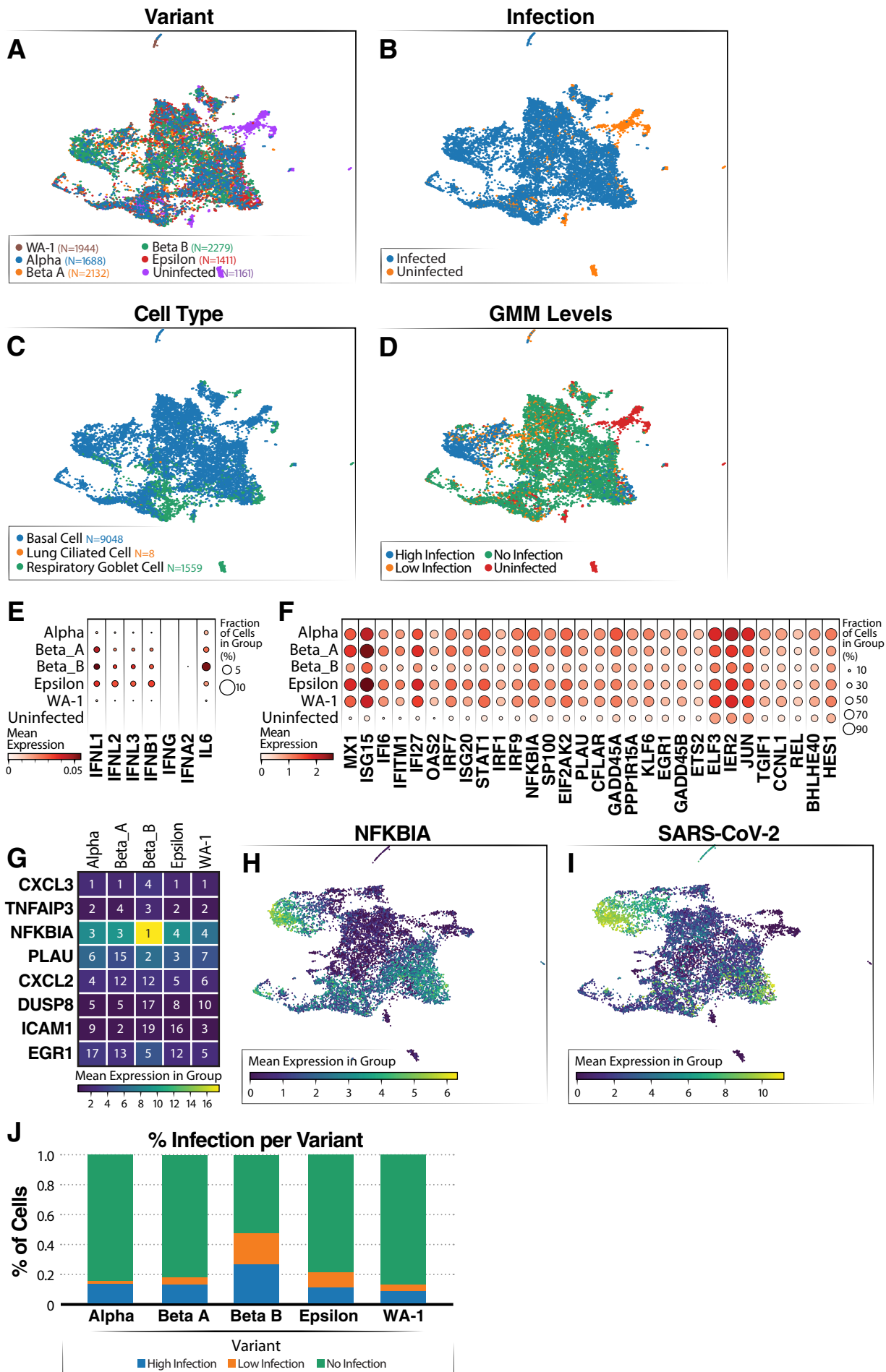
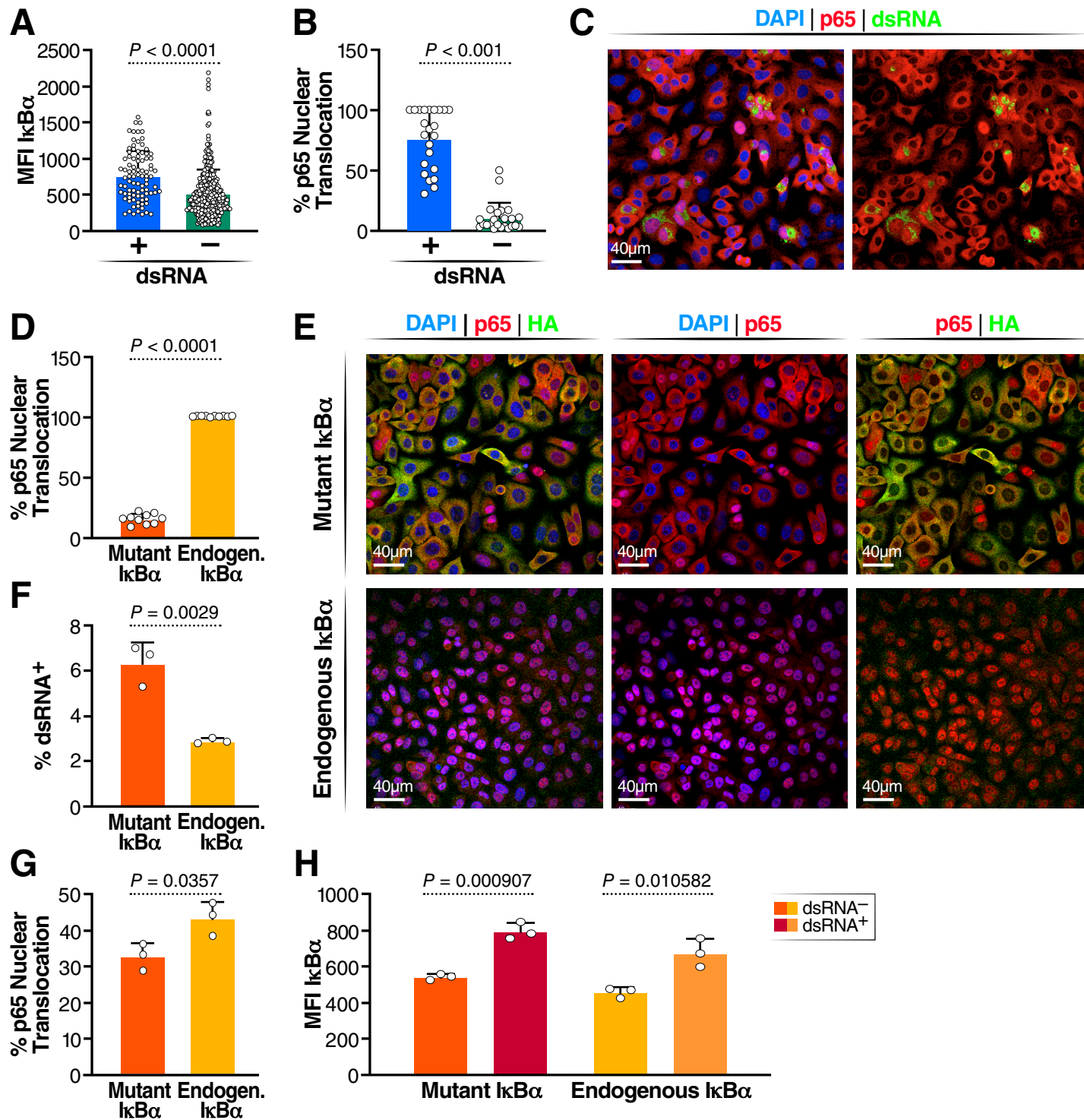




Figure 5



**Figure 6**

