Cigarette smoke exposure and inflammatory signaling increase the expression of the SARS-CoV-2 receptor ACE2 in the respiratory tract

Joan C. Smith[1,2], Erin L. Sausville[2], Vishruth Girish[2,3], Monet Lou Yuan[2,4], Kristen M. John[2,5], and Jason M. Sheltzer[2]*

1. Google, Inc., New York City, NY, 10011, USA.

2. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 11724, USA.

3. Stony Brook University, Stony Brook, NY, 11794, USA.

4. Johns Hopkins University, Baltimore, MD, 21218, USA.

5. Hofstra University, Hempstead, NY, 11549, USA.

Lead contact: Jason M. Sheltzer, sheltzer@cshl.edu.

* To whom correspondence should be addressed. Email: sheltzer@cshl.edu.

Abstract

The factors mediating fatal SARS-CoV-2 infections are poorly understood. Here, we show that cigarette smoke causes a dose-dependent upregulation of Angiotensin Converting Enzyme 2 (ACE2), the SARS-CoV-2 receptor, in rodent and human lungs. Using single-cell sequencing data, we demonstrate that ACE2 is expressed in a subset of secretory cells in the respiratory tract. Chronic smoke exposure triggers the expansion of this cell population and a concomitant increase in ACE2 expression. In contrast, quitting smoking decreases the abundance of these secretory cells and reduces ACE2 levels. Finally, we demonstrate that ACE2 expression is responsive to inflammatory signaling and can be upregulated by viral infections or interferon treatment. Taken together, these results may partially explain why smokers are particularly susceptible to severe SARS-CoV-2 infections. Furthermore, our work identifies ACE2 as an interferon-stimulated gene in lung cells, suggesting that SARS-CoV-2 infections could create positive-feedback loops that increase ACE2 levels and facilitate viral dissemination.

Introduction

In December 2019, a novel respiratory disease emerged in a seafood market in Wuhan, China[1]. Genomic sequencing demonstrated that the causative agent was a highly-contagious coronavirus, since named SARS-CoV-2[2,3]. The disease, called COVID-19, rapidly spread worldwide, and as of April 2020, more than 3 million people have been infected and more than 200,000 people have died[4]. No clinically-validated treatment or vaccine for COVID-19 is currently available. Thus, understanding the factors that mediate susceptibility to SARS-CoV-2 is crucial for controlling disease transmission.

Molecular analysis has begun to shed light on how SARS-CoV-2 infections occur. Like a related coronavirus that emerged in 2003[5], SARS-CoV-2 enters human cells by binding to the extracellular domain of Angiotensin Converting Enzyme 2 (ACE2)[3,6]. Importantly, ACE2 is both necessary and sufficient for infection by SARS-CoV-2: ACE2-targeting antibodies block viral uptake in permissive cells while transgenic expression of human ACE2 allows viral entry in non-human cells. ACE2 normally functions in the renin-angiotensin system (RAS) by cleaving the vasoconstrictive hormone angiotensin-II into the vasodilator angiotensin 1-7[7]. Sequestration of ACE2 by coronavirus dysregulates the RAS pathway, contributing to morbidity[8]. Additionally, ACE2 levels are capable of influencing disease progression: among a cohort of mice engineered to express human ACE2, mice expressing the highest levels of ACE2 mRNA exhibited the shortest survival time following coronavirus exposure[9]. Thus, the regulation of ACE2 expression likely has a significant effect on SARS-CoV-2 susceptibility.

Epidemiological studies have identified multiple demographic features that correlate with the severity of clinical COVID-19 cases. While fewer than 5% of SARS-CoV-2 infections are fatal[10,11], men and elderly patients are particularly at risk of developing severe disease[12–15]. Additionally, cigarette smokers with COVID-19 are significantly more likely to develop critical illnesses that require aggressive medical intervention[16,17]. For instance, in a study of 1,099 patients with laboratory-confirmed COVID-19, 12.3% of current smokers required mechanical ventilation, were admitted to an ICU, or died, compared to only 4.7% of non-smokers[13]. The causes underlying these differences in outcome are at present unknown.

Several clinical features have also been identified that can differentiate between patients with severe and non-severe coronavirus infections. Notably, a dysregulated immune response has been identified as a crucial mediator of COVID-19 mortality[18]. Patients who present with elevated levels of inflammatory cytokines are more likely to develop critical illnesses[19–21]. These so-called “cytokine storms” cause an increase in vascular permeability that facilitates immune cell efflux into affected tissues, but may also worsen pneumonia[22]. How this immunopathology affects the regulation of the host factors required for coronavirus infections is poorly understood.

Results

ACE2 levels in mammalian lungs are not strongly affected by age or sex

In order to study factors that could potentially influence susceptibility to SARS-CoV-2 infection, we investigated the expression of the coronavirus receptor ACE2. We first assessed the expression of ACE2 in a variety of rodent and human tissues (Figure S1 and Table S1)[23–26]. ACE2 was expressed at high levels in mouse, rat, and human kidneys, consistent with its role as a regulator of the RAS pathway. ACE2 was also expressed in rodent and human lungs, the predominant location of coronavirus infections[27]. Interestingly, significant ACE2 expression was

also evident in the mouse and human small intestine. Viral RNA has been detected in stool samples from patients with COVID-19[28], and gastrointestinal symptoms have been reported in a subset of affected individuals[13], suggesting a potential alternate route for SARS-CoV-2 transmission. However, as SARS-CoV-2 appears to be primarily spread through viral inhalation[27], we focused our study on factors that affect ACE2 expression in the lung and associated respiratory tissue.

Age and male sex are significant risk factors for severe SARS-CoV-2 infections[13,29]. We therefore investigated whether either feature was associated with increased ACE2 expression. In four different cohorts of aging mice and one cohort of aging rats, we did not detect a significant age-dependent change in ACE2 expression in the lung (Figure 1A-B and S1D-F)[25,30–33]. Similarly, ACE2 expression in rodent lungs was not significantly different between sexes (Figure 1C-D)[25,34]. We next assessed the expression of ACE2 in three different human cohorts: 1) lung tissue from the Genotype-Tissue Expression project (GTEx)[23,24], 2) whole-lung tissue samples from organ donors[35], and 3) pathologically-normal lung tissue from a cohort of patients analyzed as part of The Cancer Genome Atlas[36]. These datasets yielded results that were consistent with our rodent analyses: ACE2 expression was equivalent between men and women and between young individuals (<29 years) and elderly individuals (>70 years)(Figure 1E-J). In total, these findings suggest that the increased morbidity of men and older patients with COVID-19 is unlikely to result from inherent differences in the basal level of ACE2 expression in the lung.

Cigarette smoke increases the expression of ACE2 in the mammalian respiratory tract

Cigarette smoking is strongly associated with adverse outcomes from COVID-19[13,16,17,37–40]. To investigate whether smoking could affect ACE2 levels, we first assessed gene expression in mouse lungs. We analyzed a cohort of mice exposed to diluted cigarette smoke for 2, 3, or 4 hours per day for five months[41]. Strikingly, we found a dose-dependent increase in ACE2 expression according to smoke exposure (Figure 2A). Mice exposed to the highest dose of cigarette smoke expressed ~80% more ACE2 in their lungs compared to sham-treated mice. To determine whether this association was present in humans as well, we assessed tissue collected from three cohorts of current smokers and never-smokers[42–44]. For these analyses, lung epithelial cells were sampled by fiberoptic bronchoscopy from either the trachea, the large airways, or the small airways (Figure 2B). In each cohort, we observed that tissue samples from smokers exhibited ~30%-55% more ACE2 compared to tissue from non-smokers (Figure 2C-E). In a combined analysis comprised of data from all three tissue locations, ACE2 was in the top 2% of genes most strongly dysregulated by smoke exposure (Figure 2F).

Next, we sought to determine whether human ACE2 expression showed a dose-dependent relationship with cigarette smoke, as we had observed in mice. To investigate this, we analyzed two human datasets: 1) lung tissue from a cohort of smokers undergoing thoracic surgery for transplantation, lung volume reduction, or nodule resection[45] and 2) pathologically-normal lung tissue from TCGA patients[36]. In both cohorts, lung samples from patients who reported smoking the greatest number of pack-years also expressed the highest levels of ACE2 (Figure 2G-H). For instance, among smokers undergoing thoracic surgery, patients who had smoked more than 80 pack-years exhibited a ~100% increase in ACE2 expression relative to patients who had smoked less than 20 pack-years (Figure 2I).

We then investigated whether other demographic features could explain the upregulation of ACE2 in the lungs of smokers. Multivariate linear regression on the thoracic surgery cohort confirmed that smoking history was a significant predictor of ACE2 expression even when controlling for a patient's age, sex, race, and body-mass index (Table 1). In the TCGA cohort, pack-year history

did not remain significantly associated with ACE2 in a regression with additional clinical variables, which may reflect the very small number of patients who could be included in this analysis (22 patients; data not shown). However, demographic information was available for the tracheal epithelium cohort analyzed above (Figure 2E), and smoking status remained correlated with ACE2 in a regression that included each patient's age, sex, and race (Table 1).

Finally, we examined the effects of quitting smoking on ACE2 expression. In a cohort of patients comprised of either current smokers or former smokers who had refrained from smoking for at least 12 months, quitting smoking was associated with a ~40% decrease in ACE2 expression (Figure 2I). ACE2 was among the top 5% of genes most strongly affected by quitting smoking (Figure 2J). In total, our results demonstrate that exposure to cigarette smoke increases the expression of the coronavirus receptor ACE2 in rodent and human respiratory tissue, and this upregulation is potentially reversible.

Coronavirus infections are facilitated by a set of host proteases that cleave and activate the viral spike (S) protein[46]. SARS-CoV-2 primarily relies on the serine protease TMPRSS2 but can also utilize an alternate pathway involving Cathepsin B/L in TMPRSS2-negative cells[6]. Interestingly, we observed that Cathepsin B expression, but not TMPRSS2 or Cathepsin L expression, was consistently increased in mice and humans exposed to cigarette smoke (Figure S3). In a meta-analysis across the trachea, large airways, and small airways, Cathepsin B was in the top 11% of genes dysregulated in the respiratory tract of cigarette smokers (Figure S3D). Thus, smoking can upregulate both the coronavirus receptor as well as a protease that SARS-CoV-2 is capable of using for viral activation.

ACE2 is expressed at high levels in secretory cells in the lung epithelium

Mammalian lungs harbor more than 30 distinct cell types representing a variety of epithelial, endothelial, stromal, and immune compartments[47]. Of note, the upper respiratory epithelium is comprised of mucociliary cells, including goblet cells, club cells, and ciliated cells, that secrete protective fluids and remove inhaled particles from the airways[48]. The lower respiratory epithelium includes alveolar type 1 cells, which allow gas exchange with the blood, and alveolar type 2 cells, which regulate alveolar fluid balance and can differentiate into type 1 cells following injury[49].

To gain further insight into coronavirus infections, we profiled multiple single-cell RNA-Seq experiments to identify the cell type(s) that express ACE2. We first examined a dataset containing 13,822 cells from normal mouse lungs[50]. We performed unsupervised Leiden clustering to separate the cells into distinct populations and then we assigned cell types to major clusters using established markers[51–54]. ACE2 was expressed solely in the EpCAM+ clusters that comprise the lung epithelium[55], and was not detected in CD45+ immune cells[56], PDGFRA+ mesenchymal cells[57], or TMEM100+ endothelial cells (Figure 3A-B)[58]. We therefore focused on localizing ACE2 within the epithelial lineage. We found that ACE2 was present in a cluster of cells that express secretory markers including MUC5AC, GABRP, and SCGB1A1 that we identified as being comprised of closely-related goblet and club cells (Figure 3C-D)[54,59–61]. ACE2 expression was also observed in a subset of LAMP3+ alveolar type 2 cells[62], but was largely absent from RTKN2+ alveolar type 1 cells[54].

We then extended our findings to an independent single-cell dataset from human lungs[50]. Consistent with our initial observations, we found that ACE2+ cells were found almost exclusively in the EPCAM+ epithelial compartment (Figure 3E-F). Within this lineage, we observed ACE2 expression in the LAMP3+ alveolar type 2 cluster and the MUC5AC+ goblet/club cell cluster (Figure 3G-H). Additionally, in human cells, there was some expression of ACE2 in FOXJ1+

ciliated cells[63]. In total, our combined analysis demonstrates that ACE2 is expressed in the mammalian lung epithelium, and is present at particularly high levels in secretory club and goblet cells as well as in alveolar type 2 cells.

Cigarette smoke triggers an increase in ACE2+ cells by driving secretory cell expansion

Our findings on ACE2's expression pattern suggested a possible explanation for its upregulation in the lungs of cigarette smokers. Chronic exposure to cigarette smoke has been reported to induce the expansion of secretory goblet cells, which produce mucous to protect the respiratory tract from inhaled irritants[64–66]. Thus, the increased expression of ACE2 in smokers' lungs could be a byproduct of smoking-induced secretory cell hyperplasia.

To investigate this hypothesis, we examined a dataset of single-cell transcriptomes collected from the tracheas of current smokers and never-smokers[67]. As these cells were derived from the upper airways, they were highly-enriched for pseudostratified respiratory epithelial cell types and contained few alveolar cells. Consistent with the initial analysis of this dataset, Leiden clustering revealed a large population of TP63+ basal cells (the putative stem cell population for the airway epithelium)[68], an intermediate KRT8^{high} population that showed evidence of ongoing differentiation[67], and populations of mature FOXJ1+ ciliated cells and MUC5AC+ secretory cells (Figure 4A-B). We also identified markers for submucosal gland (SMG) clusters as well as rare epithelial cell types, including CFTR+ ionocytes[69] and CHGA+ neuroendocrine cells[70]. Within this dataset, ACE2 was expressed at the highest levels in the MUC5AC+ secretory cell cluster, followed by SMG-secretory cells, KRT8^{high} intermediate cells, and FOXJ1+ ciliated cells.

We next identified the transcripts whose expression correlated with ACE2. Across all cells, ACE2 levels were strongly correlated with several mucin genes, including MUC1, MUC4, MUC15, and MUC16, as well as other genes associated with barrier epithelia, including ALCAM[71], CLDN7[72], and TJP3[73] (Table S2A). Gene ontology analysis revealed that ACE2-correlated transcripts were enriched for genes involved in secretion, glycosylation, and the response to toxic substances, consistent with the airway epithelium's role as a producer of mucus and a barrier against foreign matter (Figure 4C and Table S2B)[74]. ACE2+ correlates overlapped with but were not identical to goblet/club markers: several genes were widely expressed in both populations (MUC16, CLDN7, S100P), while others were more restricted to single lineages (PIGR, PSCA)(Figure 4D). In general, the ACE2+ signature was expressed in both secretory cells and ciliated cells but not in the basal stem cell compartment.

We next separated the cells harvested from current smokers and never-smokers and then analyzed each population separately. Consistent with previous reports[64–66], we detected a significant expansion of the secretory cell compartment in smokers' airways, as evidenced by both cluster analysis and by the increase in the number of cells expressing the canonical goblet cell marker MUC5AC (Figure 4E-F). FOXJ1+ ciliated cells decreased in abundance, likely reflecting the fact that smoke exposure inhibits ciliogenesis[75]. Smoking caused an increase in both the number of ACE2+ positive cells and in ACE2 expression within ACE2+ cells (Figure 4G-H). This expansion resulted from hyperplasia of the secretory cell compartment: smoke exposure resulted in a two-fold increase in the frequency of MUC5AC+/ACE2+ double-positive cells, while there was no significant change in the frequency of MUC5AC-/ACE2+ cells. In total, these results provide single-cell evidence that the increase in ACE2 expression in smokers' respiratory tracts is caused by the expansion of mucous-secreting goblet cells that co-express ACE2.

To further validate these findings, we derived gene signatures for each of the cell types identified from tracheal single-cell sequencing, and we applied these signatures to analyze the bulk gene

expression datasets from smokers and non-smokers that we had previously investigated (Figure 2). The goblet/club cell gene signature was significantly upregulated in current smokers compared to never smokers, while in former smokers these genes were strongly downregulated (Figure 4I-J). A signature based on ACE2-correlated transcripts showed a similar expression pattern. In contrast, a ciliated cell transcriptional signature was strongly downregulated among current smokers. As some FOXJ1+ ciliated cells co-express ACE2, this smoking-dependent suppression of ciliogenesis may partially blunt the increase in ACE2+ cells caused by smoke-induced goblet cell hyperplasia.

Secretory cell differentiation of lung epithelium can be modeled *in vitro* by culturing cells at an air-liquid interface (ALI)[76,77]. Under appropriate conditions, primary respiratory cells growing at an ALI will undergo mucociliary differentiation into a stratified epithelium consisting of ciliated cells, goblet cells, and club cells[78]. As our single-cell analysis suggested that the coronavirus receptor ACE2 is expressed at higher levels in differentiated secretory and ciliated cells compared to basal stem cells, we investigated whether *in vitro* mucociliary differentiation increases ACE2 expression. Indeed, in mouse tracheal extracts[79] and primary human lung cells[80], mucociliary differentiation resulted in a highly-significant upregulation of ACE2 (Figure 4K-L). Finally, to investigate the link between smoking, differentiation, and ACE2 expression, we examined data from human bronchial epithelial cells cultured at an ALI in which cells were either exposed to clean air or to diluted cigarette smoke[81]. Remarkably, treatment with cigarette smoke during *in vitro* differentiation resulted in a significant upregulation of ACE2 relative to cells that were differentiated in clean air (Figure 4M). Smoke exposure increased ACE2 expression by ~42%, comparable to the increases that we observed between the lungs of non-smokers and smokers (Figure 2). Differentiation in the presence of cigarette smoke similarly resulted in an upregulation of the goblet/club cell transcriptional signature and a downregulation of the ciliated cell transcriptional signature (Figure 4N). In full, our results demonstrate that a subset of lung secretory cells express the coronavirus receptor ACE2, and cigarette smoke promotes the expansion of this cell population.

ACE2 is upregulated in smoking-associated diseases and by viral infections

To follow up on these observations, we investigated whether ACE2 expression was affected by other lung diseases and/or carcinogen exposures. Indeed, we observed increased ACE2 expression in multiple cohorts of patients with chronic obstructive pulmonary disease (COPD) and idiopathic pulmonary fibrosis (IPF)(Figure S3A-D)[82–85]. Interestingly, both COPD and IPF are strongly associated with prior cigarette exposure[86,87], and COPD in particular has been identified as a risk factor for severe COVID-19[16,88]. However, ACE2 expression was generally not affected by other lung conditions or toxins. We did not observe a significant difference in ACE2 expression in lung samples from a large cohort of patients with asthma or from patients with the lung disease sarcoidosis (Figure S3E-F)[89,90]. Similarly, ACE2 expression was unaltered in lung tissue from a mouse model of cystic fibrosis and in mice exposed to a variety of carcinogens, including arsenic, ionizing radiation (IR), and 1,3-butadiene (Figure S3G-J)[91–94]. We conclude that ACE2 upregulation in the lung is tightly associated with a history of cigarette smoking and is not a universal response to pulmonary diseases.

So-called “cytokine storms”, characterized by high levels of circulating inflammatory cytokines, have been identified as a cause of COVID-19-related mortality[18,20,21]. Cytokine release can be triggered by viral infections, which serve to induce immune cell activation and expansion[95]. Cigarette smoke is also an inflammatory agent, and smokers tend to exhibit an increase in inflammation-related serological markers[96,97]. To investigate a potential link between inflammation and the expression of the host factors required for coronavirus infections, we first examined the levels of ACE2 in published datasets of respiratory epithelial cells challenged with different

viruses. We observed a highly significant upregulation of ACE2 in airway cells that were infected with two RNA viruses, influenza and Respiratory Syncytial Virus (Figure 5A-B)[98]. Infection with two coronaviruses, SARS and MERS, that are closely related to the 2019 pandemic strain also caused an increase in ACE2 levels (Figure 5C-D)[99,100]. Double-stranded RNA that is produced during viral replication can be detected by host pattern recognition receptors, triggering a strong immune response[101]. Accordingly, transfection of the dsRNA mimic poly(I:C) into airway epithelial cells was sufficient to induce the upregulation of ACE2 (Figure 5E)[102]. These results reveal that inflammatory signals – like those triggered by smoking or by a viral infection – are capable of increasing the expression of ACE2 in the respiratory epithelium.

To further investigate the stimuli that are sufficient to upregulate ACE2, we cultured primary epithelial cells from either the small airways or the trachea in the presence of different compounds. We found that exposure to interferons resulted in a significant upregulation of ACE2 expression (Figure 5F-G). IFN-α, IFN-β, and IFN-γ treatment increased ACE2 expression in tracheal cells, while only IFN-α and IFN-β affected ACE2 in small airway cells. To further verify these results, we re-analyzed a published gene expression dataset of airway epithelial cells exposed to IFN-β, and we found a similar increase in ACE2 levels following interferon treatment (Figure 5H)[103,104]. The results identify ACE2 as an interferon-regulated gene and suggest a potential mechanism by which inflammatory stimuli could facilitate SARS-CoV-2 infections.

ACE2 RNA and protein levels are tightly correlated

This paper describes a series of analyses and experiments to identify factors that drive the expression of the coronavirus receptor ACE2. However, these results could potentially be confounded if changes in ACE2 RNA levels fail to affect the steady-state levels of ACE2 protein. While we lack the data to directly test that link in the lungs of smokers, we conducted three additional analyses to investigate whether ACE2 RNA and protein expression are generally correlated. First, we examined a dataset of human cancer cell lines that had been profiled at the transcriptional level by RNA-Seq and at the protein level by mass spectrometry[105–107]. Across 53 cell lines, ACE2 RNA and protein levels were strongly correlated ($r = 0.82$, $p < .0001$; Figure S4A). Indeed, ACE2's RNA-protein correlation coefficient was higher than 95% of human genes that could be reliably detected (Figure S4B). Next, we compared ACE2 RNA levels with ACE2 immunohistochemistry staining in untransformed human tissues. We similarly observed a strong correlation between ACE2 RNA and protein staining: five of the six tissues with the highest levels of RNA expression exhibited strong staining, while none of the 35 other tissues with lower RNA levels exhibited strong staining (Figure S4C)[108]. Lastly, we directly compared ACE2 expression in seven human cell lines by qRT-PCR and by western blotting. We found that cell lines that displayed the highest levels of ACE2 RNA (LoVo, Caco-2, and Calu3) also exhibited the most ACE2 protein, while cell lines with very low levels of ACE2 RNA (HCT116, CAL148, and A2780) had undetectable levels of ACE2 protein (Figure S4D-E). Thus, while we cannot rule out the possibility that ACE2 RNA and protein levels are differentially regulated in certain circumstances, our analysis reveals that ACE2 RNA and protein levels are generally correlated.

Discussion

While SARS-CoV-2 has infected more than 3 million people worldwide, fewer than 5% of COVID-19 cases are fatal[4,13]. Here, we show that cigarette smokers harbor consistently higher levels of the SARS-CoV-2 receptor ACE2 in their respiratory tracts. This upregulation is likely mediated by the expansion of ACE2+ secretory cells caused by chronic smoke exposure. Certain inflammatory cytokines also trigger ACE2 upregulation, which could further influence ACE2 expression due to

smoking-associated lung inflammation. The overabundance of ACE2 in the lungs of smokers may partially explain why smokers are significantly more likely to develop severe SARS-CoV-2 infections that require aggressive medical interventions[13,16,17,37]. Furthermore, as quitting smoking is associated with a decrease in ACE2 expression, it is possible that giving up cigarettes may reduce susceptibility to deadly COVID-19.

Several contrasting findings exist in the literature on ACE2 and cigarette exposure[109–116]. In particular, it has been reported that nicotine and/or cigarette smoke has the potential to downregulate ACE2 expression in certain tissues or cell types[109–112]. In this manuscript, we focused our analysis on factors affecting ACE2 expression in the mammalian lungs and associated respiratory epithelia. We observed a consistent correlation between smoking history and ACE2 expression that was dose-dependent (Figure 2G-H), that could be recapitulated in mice (Figure 2A) and *in vitro* (Figure 4M), that was detectable in both bulk and single-cell analyses (Figure 2F and 4G-H), and that remained significant when controlling for other demographic variables (Table 1). Thus, we propose that cigarette smoke causes the upregulation of ACE2 expression in the respiratory tract. However, we recognize that individual components found within cigarettes, like nicotine, may have a different effect on ACE2 than whole smoke, and smoking itself could alter ACE2 levels in non-respiratory organs in different ways[110].

Several previous studies have examined the localization of ACE2 within the respiratory tract[116–127]. Consistent with our results, ACE2 expression has been detected in the respiratory epithelium, including prominent staining in alveolar type 2 cells. However, some prior studies have also documented ACE2 expression in lung endothelial and smooth muscle cells[118,120,121], which we did not observe. These findings could represent non-specific staining resulting from the particular antibodies that were used[128]. Alternately, ACE2 may be expressed in mesenchymal or other lineages within the lung, and our inability to detect it may be a limitation of the single-cell samples that we analyzed. Nonetheless, by examining datasets from both mice and humans, and by including cells collected from bronchial brushing, we demonstrated consistently high levels of ACE2 transcripts in the secretory cells that participate in mucociliary clearance[48]. As these cells line the upper respiratory tract, they may represent the initial site of coronavirus infections, followed by an eventual spread and migration into the alveoli.

Cigarette smoke has a profound impact on the lungs of chronic smokers. Notably, prolonged smoke exposure triggers secretory cell hyperplasia, thereby increasing the production of mucous in the respiratory tract[64–66]. We found that this smoking-induced expansion of secretory cells also causes an increase in ACE2 levels, as smoking doubled the number of cells co-expressing ACE2 and MUC5AC. Additionally, consistent with a recent report[122], we found that ACE2 is an interferon-regulated gene, and is over-expressed in lung epithelial cells following viral infection or interferon treatment. Lung damage and inflammation caused by smoking could also contribute to ACE2 upregulation. Additionally, we speculate that the interferon-dependent upregulation of ACE2 could create a positive-feedback loop for SARS-CoV-2 infections. That is, interferon secretion following an initial infection could increase ACE2 expression within neighboring cells, thereby rendering those cells susceptible to SARS-CoV-2 and facilitating viral dissemination. Accordingly, clinical interventions to dampen the immune response could benefit patients in part by breaking this positive-feedback cycle[22,129].

The factors that mediate overall susceptibility to SARS-CoV-2 infections are poorly understood. We speculate that the increased expression of ACE2 that we found in the lungs of smokers could partially contribute to the severe cases of COVID-19 that have been observed in this patient population. Chronic smokers may exhibit a number of co-morbidities, including emphysema, atherosclerosis, and immune dysregulation[130], that are also likely to affect COVID-19 progression.

While the effects of smoking can last for years, smoking cessation causes an improvement in lung function and an overall decrease in disease burden[130]. Quitting smoking leads to a normalization of respiratory epithelial architecture[131], a decrease in hyperplasia[132], and a downregulation of ACE2 levels. Thus, for multiple reasons, smoking cessation could eventually lessen the risks associated with SARS-CoV-2 infections.

Limitations

Our analysis in this manuscript has a number of important limitations. First, the relationship between age, sex, and ACE2 expression remains controversial. In the literature, ACE2 expression in the lung has been reported to both increase[133] and decrease[120,134] during aging. Across five rodent and three human datasets, we did not observe a significant correlation in either direction between ACE2 levels and age. It remains possible that such a difference could be found using tissue from older or younger individuals. Additionally, it is possible that ACE2 expression within single cell types correlates with age, but such differences are not prominent enough to detect using bulk microarray and RNA-Seq analysis.

In order for SARS-CoV-2 to infect a host cell, the virus must bind to ACE2 protein that is present on the cell surface. Here, we have described the expression of ACE2 transcripts in mammalian lungs, but our work does not guarantee that stimuli affecting the levels of ACE2 mRNA will have the same effect on ACE2 protein. In general, we have documented a strong correlation between ACE2 mRNA and protein levels (Figure S4), but it remains possible that ACE2 protein is differentially regulated in mammalian lungs, or that the localization of ACE2 protein to the cell surface is strictly controlled. Immunohistochemistry using tissue samples from smokers and virus-infected patients will be needed to confirm that these factors also increase the levels of ACE2 protein available to interact with SARS-CoV-2 particles.

Finally, the exact role of ACE2 as a mediator of disease severity remains to be determined. Mice that were engineered to express high levels of human ACE2 succumbed to infections with the SARS coronavirus more quickly than mice that expressed low levels of human ACE2, suggesting that increasing ACE2 enhances viral susceptibility[9]. At the same time, ACE2-knockout mice are vulnerable to a variety of lung injuries, and ACE2 expression has been reported to play a protective role in the respiratory tract[8,135,136]. As ACE2 expression is both necessary and sufficient for SARS-CoV-2 infections[6,137], it seems highly likely that an expansion of ACE2+ cells in the lungs will facilitate viral dissemination, but it remains possible that the presence of ACE2 has some beneficial functions as well. Additional work will be required to determine the precise impact of ACE2 expression levels on the clinical course of COVID19.

Author contributions

Conceptualization, J.C.S. and J.M.S.; Methodology, J.C.S. and J.M.S.; Software, J.C.S.; Formal analysis, J.C.S. and J.M.S.; Writing – Original Draft, J.M.S.; Writing – Review & Editing, J.C.S. and J.M.S; Supervision, J.M.S.; Investigation, J.C.S., E.L.S., V.G., M.L.Y., K.M.J., and J.M.S..

Declaration of Interests

J.C.S. is a co-founder of Meliora Therapeutics and is an employee of Google, Inc. This work was performed outside of her affiliation with Google and used no proprietary knowledge or materials from Google. J.M.S. has received consulting fees from Ono Pharmaceuticals, is a member of the Advisory Board of Tyra Biosciences, and is a co-founder of Meliora Therapeutics.

Table 1. Patient characteristics and ACE2 expression.

Multivariate regression between various patient characteristics, smoking status, and ACE2 expression. The thoracic surgery cohort is from GSE76925 and the tracheal epithelium cohort is from GSE13933. Bolding indicates significant variables. *, p < .05; **, p < .005; ***, p < .0005.

		β value	Standard Error	P value
Thoracic surgery cohort	**Age**	0.0014	0.0086	0.8700
	BMI	-0.0199	0.0127	0.1193
	Smoking Pack-years	0.0081	0.0025	**0.0015****
	Race **Caucasian = 0** **Not Caucasian = 1**	0.1910	0.1663	0.2527
	Sex **Female = 0** **Male = 1**	0.3335	0.1338	**0.0138***
Tracheal epithelium cohort	**Age**	-0.0027	0.0080	0.7364
	Smoking status **Never smoker = 0** **Current smoker = 1**	0.5638	0.1269	**7×10^{-5}*** **
	Race **Caucasian = 0** **Not Caucasian = 1**	0.1252	0.1331	0.3527
	Sex **Female = 0** **Male = 1**	0.2538	0.1328	0.0632

Figure Legends

Figure 1. ACE2 expression in the lung is uncorrelated with age or sex.

(A) ACE2 expression in the lungs of young mice (<26 weeks old) and old mice (>78 weeks old).

(B) ACE2 expression in the lungs of young rats (6 weeks old) and old rats (104 weeks old).

(C) ACE2 expression in the lungs of female mice and male mice.

(D) ACE2 expression in the lungs of female rats and male rats.

(E) ACE2 expression in lungs from the GTEx cohort by age.

(F) ACE2 expression in lungs from the GTEx cohort by sex.

(G) ACE2 expression in lungs from a cohort of organ donors by age.

(H) ACE2 expression in lungs from a cohort of organ donors by sex.

(I) ACE2 expression in pathologically normal lung tissue from patients from TCGA by age.

(J) ACE2 expression in pathologically-normal lung tissue from patients from TCGA by sex.

Each panel displays log2-normalized ACE2 expression relative to a control group. Data analyzed in A and C were from GSE132040. Data analyzed in B and D were from GSE53960. Data analyzed in E and F were from www.gtexportal.org. Data analyzed in G and H were from GSE1643. Data analyzed in I and J were from https://gdac.broadinstitute.org/. Additional information on the data sources and sample sizes are included in Table S1.

Figure 2. Cigarette smoke increases the expression of ACE2 in mouse and human lungs.

(A) ACE2 expression in the lungs of mice that were sham-treated or that were exposed to diluted cigarette smoke for two, three, or four hours a day (low, medium, and high smoke exposure, respectively).

(B) A diagram showing the approximate locations of tissue samples used in this analysis. Tracheal, large airway epithelial, and small airway epithelial specimens were collected by fiberoptic bronchoscopy as described in their respective publications. Lung resections were collected surgically from various locations.

(C) ACE2 expression in human tracheal epithelia analyzed according to smoking history.

(D) ACE2 expression in human large airway epithelia analyzed according to smoking history.

(E) ACE2 expression in human small airway epithelia analyzed according to smoking history.

(F) A volcano plot comparing gene expression in the respiratory epithelia of current smokers and never-smokers from C, D, and E. The dotted lines indicate various p-value thresholds (e.g., genes located above the 10th percentile have a combined p value greater than 90% of the genes included in the meta-analysis). The location of ACE2 is indicated with a red star.

(G) ACE2 expression in the lungs of a cohort of patients undergoing thoracic surgery analyzed according to the number of pack-years each patient smoked.

(H) ACE2 expression in the lungs of TCGA patients analyzed according to the number of pack-years each patient smoked.

(I) ACE2 expression in respiratory epithelia collected by fiberoptic bronchoscopy among either current smokers or former smokers.

(J) A volcano plot comparing gene expression between current smokers and former smokers. The dotted lines indicate various p-value thresholds (e.g., genes located above the 10th percentile have a combined p value greater than 90% of the genes included in the meta-analysis). The location of ACE2 is indicated with a red star.

Each panel displays log2-normalized ACE2 expression relative to a control group. Data analyzed in A were from GSE18344. Data analyzed in C were from GSE13933. Data analyzed in D were from GSE22047. Data analyzed in E were from GSE64614. Data analyzed in G were from GSE76925. Data analyzed in H were from https://gdac.broadinstitute.org/. Data analyzed in I were from GSE79209. Additional information on the data sources and sample sizes are included in Table S1. *, $p < .05$; **, $p < .005$; ***, $p < .0005$ (Student's t-test).

Figure 3. ACE2 is expressed in secretory club and goblet cells along with alveolar type 2 cells in the mammalian lung.

(A) T-SNE clustering of cells from the mouse lung. Cells expressing ACE2 are highlighted in the right panel.

(B) Cells in the mouse lung that express various lineage markers (TMEM100 for endothelial cells, EPCAM for epithelial cells, PDGFRA for mesenchymal cells, and PTPRC for immune cells) are highlighted.

(C) Cells expressing markers for various epithelial lineages are highlighted: RTKN2 for alveolar type 1 cells, LAMP3 for alveolar type 2 cells, FOXJ1 for ciliated cells, GABRP for both goblet and club cells, MUC5AC for goblet cells, and SCGB1A for club cells.

(D) A track plot displaying the expression of ACE2 and several lineage-related genes in different cell populations obtained from Leiden clustering.

(E) T-SNE clustering of cells from the human lung. Cells expressing ACE2 are highlighted in the right panel.

(F) Cells in the human lung that express various lineage markers (TMEM100 for endothelial cells, EPCAM for epithelial cells, PDGFRA for mesenchymal cells, and PTPRC for immune cells) are highlighted.

(G) Cells expressing markers for various epithelial lineages are highlighted: LAMP3 for alveolar type 2 cells, FOXJ1 for ciliated cells, TP63 for basal cells, GABRP for both goblet and club cells, MUC5AC for goblet cells, and SCGB1A1 for club cells.

(H) A track plot displaying the expression of ACE2 and several lineage-related genes in different cell populations obtained from Leiden clustering.

The gene expression data used in A, B, C, and D are from GSE121611. The gene expression data used in D, E, F, and G are from GSE122960. Additional information on the data sources and sample sizes are included in Table S1.

Figure 4. Cigarette smoke causes the expansion of ACE2+ secretory cells.

(A) T-SNE clustering of the transcriptomes from single cells derived from the airway epithelia of smokers and never-smokers. Cells expressing ACE2 are highlighted in the right panel.

(B) Cells in the human airway that express various lineage markers (FOXJ1 for ciliated cells, TP63 for basal cells, MUC5AC for goblet cells, and SCGB1A1 for club cells) are highlighted.

(C) GO terms enriched among ACE2-correlated transcripts.

(D) Dot plots displaying the expression of the top 10 differentially-expressed marker genes for various airway lineages and for ACE2+ cells.

(E) The fraction of cells expressing the indicated marker genes are displayed. FOXJ1 is a marker for ciliated cells, MUC5AC is a marker for goblet cells, and TP63 is a marker for basal cells.

(F) The fractions of cells found in each cell type cluster are displayed.

(G) The numbers of counts per ACE2+ cell are displayed.

(H) The fraction of ACE2+ cells co-expressing MUC5AC or that are found within the goblet/club cell cluster are displayed.

(I) The 100 top-ranked differentially-expressed genes from each cluster in (A) as well as the 100 genes most strongly correlated with ACE2 were used to re-analyze the bulk gene expression data from smokers and non-smokers in Figure 2F. A volcano plot displays the mean expression change of each cell type signature.

(J) The same transcriptional signatures as in Figure 4I were used to re-analyze the data from current smokers and former smokers in Figure 2J. A volcano plot displays the mean expression change of each cell type signature.

(K) ACE2 expression in mouse tracheal explants undergoing mucociliary differentiation.

(L) ACE2 expression in human airway epithelial cells undergoing mucociliary differentiation.

(M) ACE2 expression in human airway epithelial cells that underwent mucociliary differentiation in the presence of clean air or cigarette smoke.

(N) The same transcriptional signatures as in Figure 4I were used to re-analyze the data from smoke exposure during differentiation from Figure 4M. A volcano plot displaying the mean expression change of each cell type signature is displayed.

Data analyzed in A through H were from GSE131391. Data analyzed in K were from GSE75715. Data analyzed in L were from GSE39059. Data analyzed in M were from GSE135188. Additional information on the data sources and sample sizes are included in Table S1. In E, F, and H, a chi-square test is applied. In G, a Mann-Whitney U test is applied. In K, L, and M, a Student's t-test is applied. *, $p < .05$; **, $p < .005$; ***, $p < .0005$.

Figure 5. **ACE2 is an interferon-stimulated gene that is upregulated by viral infections.**

(A) ACE2 expression in airway epithelial cells that were infected with influenza.

(B) ACE2 expression in airway epithelial cells that were infected with respiratory syncytial virus.

(C) ACE2 expression in airway epithelial cells that were infected with SARS.

(D) ACE2 expression in airway epithelial cells that were infected with MERS.

(E) ACE2 expression in airway epithelial cells that were transfected with the dsRNA mimic poly(I:C).

(F) ACE2 expression in human tracheal cells that were cultured in the presence of the indicated cytokine for 24 hours.

(G) ACE2 expression in human small airway epithelial cells that were cultured in the presence of the indicated cytokine for 24 hours.

(H) ACE2 expression in airway epithelial cells that were cultured in the presence of IFN-β.

Each panel displays log2-normalized ACE2 expression relative to a control group. Data analyzed in A and B were from GSE32138. Data analyzed in C were from GSE47963. Data analyzed in D were from GSE100504. Data analyzed in E were from GSE51392. Data analyzed in H were from GSE19392. Additional information on the data sources and sample sizes are included in Table S1. *, $p < .05$; **, $p < .005$; ***, $p < .0005$ (Student's t-test).

STAR Methods

RESOURCE AVAILABILITY

Further information and requests for resources should be directed to and will be fulfilled by the Lead Contact, Dr. Jason Sheltzer (sheltzer@cshl.edu).

MATERIALS AVAILABILITY

This study did not generate any new unique reagents.

DATA AND CODE AVAILABILITY

All of the data used in this manuscript are publicly accessible and are described in Table S1. All code for performing these analyses is available at github.com/joan-smith/covid19.

EXPERIMENTAL MODELS AND SUBJECT DETAILS

Cell lines

All cell lines used in this study were acquired from ATCC, except for CAL148 cells that were acquired from DSMZ and A2780 cells that were acquired from ECACC. Primary small airway epithelial cells and primary bronchial/tracheal epithelial cells were cultured in Airway Epithelial Cell Basal Medium (ATCC; Cat. No. PCS-300-030) supplemented with the Bronchial Epithelial Cell Growth Kit (ATCC; Cat. No. PCS-300-040™). HCT116, CAL148, and A375 cell lines were cultured in Dulbecco's Modified Eagle Medium (DMEM) (Gibco; Cat. No. 11995073) supplemented with 10% fetal bovine serum (FBS)(Corning; Cat. No. 35-010-CV). LoVo and A2780 were cultured in RPMI 1640 (Lonza; Cat. No. 12-115F) supplemented with 10% FBS. Calu-3 was cultured in Eagle's Minimum Essential Medium (EMEM) (ATCC; Cat. No. 30-2003™) supplemented with 10% FBS. Caco-2 was cultured in EMEM supplemented with 20% FBS. All cell lines were grown in a humidified environment at 37°C and 5% CO2.

METHOD DETAILS

Sample Sizes and Statistical Methodology

For the qPCR assays in Figure 5F-G and Figure S4D, we performed three biological replicates using two ACE2 primer pairs. Gene expression values in Figure 1, Figure S1, Figure 2, Figure S2, Figure 4K-M, Figure S3, and Figure 5 were compared by a two-sided Student's t-test. Gene expression values in Figure 4G were compared by a Mann-Whitney U test. The fractions of cells expressing a particular marker in Figure 4E-F and Figure 4H were compared by a chi-square test. No outliers were excluded from these analyses. Sample sizes of the gene expression datasets that were analyzed are listed in Table S1.

Overall analysis strategy

The analysis described in this paper was performed using Python, Excel, and Graphpad Prism. Gene expression data was acquired from the Gene Expression Omnibus (GEO)[139], the GTEx portal[23], the Broad Institute TCGA Firehose[140], the Human Cell Atlas[141], and the Single-Cell Expression Atlas[142], as described below. For microarray datasets, probeset definitions were downloaded from GEO, and probes mapping to the same gene were collapsed by averaging. For each gene expression comparison, a control population was identified (e.g., young rats, sham-treated mice, non-smokers, etc.), and gene expression values were log2-transformed and normalized by subtraction so that the mean expression of a gene of interest in the control population was 0. Graphs of gene expression values were then generated using Graphpad Prism; all data points are displayed and no outliers were excluded from analysis.

Data sources

The data sources used in this paper are listed in Table S1. In general, pre-processed microarray and RNA-Seq datasets were downloaded. Additional notes on sample selection and processing are included in Table S1.

Multivariate regression and smoking history

Multivariate regression to investigate the relationship between ACE2 expression and smoking history was performed on the GSE76925 lung tissue dataset and the GSE13933 tracheal epithelium dataset. Regressions were performed in Python using ordinary least squares from the statsmodels package[143]. Results reported include the standard errors ('bse'), betas, and p-values.

Single-cell analysis

Single-cell clustering and analysis on the datasets listed in Table S1 was performed in Python using the Scanpy and Multicore-TSNE packages[51,144]. To filter out low-quality cells, only cells in which 500 or more genes were detected were included in this analysis. Before clustering, transcript counts were counts-per-million normalized and log2 transformed. Highly variable genes were selected using the Seurat approach in Scanpy, and these highly variable genes were used to produce the principal component analysis. A t-SNE projection and unsupervised Leiden clustering were then performed on each dataset using nearest neighbors, as described in the associated code. Parameter searches were performed according to the method described in ref[145].

In order to label each cluster, a gene ranking analysis was obtained using Scanpy. The 20 most highly-ranked genes from each cluster (as determined by a t-test with overestimated variance) were identified. These genes were then compared against gold-standard marker lists from multiple sources to produce the cluster labels[51–54,67].

Gene ontology analysis

In order to identify the genes whose expression correlates with ACE2, pairwise Pearson correlations coefficients (PCCs) were calculated between ACE2 and every other expressed gene. Genes whose PCC were more than three standard deviations greater than the average gene's PCC were classified as strongly correlated with ACE2. Gene ontology terms enriched in this group were then identified with GProfiler against the background list of non-strongly correlated genes using a Benjamini-Hochberg FDR of .05[146].

Volcano plots

For the meta-analysis of gene expression changes in the small airways, large airways, and trachea, p values from each independent dataset were combined using Fisher's method. To deconvolve the bulk gene expression datasets based on cell type signatures, the 100 highest ranking genes from each cell type cluster (identified using a t-test with overestimated variance, as described above) were identified and the 100 genes exhibiting the strongest correlation with ACE2 were identified. Volcano plots were generated using the mean expression of the 100 genes comprising each signature and the p-value of a t-test comparing the signature genes to all genes within a dataset.

Cytokine treatments

Early-passage human tracheal epithelial cells and small airways epithelial cells were acquired from ATCC and cultured according to the supplier's recommended method. Cells were grown for 48 hours to approximately 70% confluence, treated with the indicated compound for 24 hours, and then total RNA was harvested.

RNA expression analysis

qRT-PCR was performed as described in ref[147]. In short, RNA was extracted using TRIzol (Life Technologies; Cat. No. 15596018) and then purified with a Qiagen RNeasy Mini Kit (Cat. No. 74106). RNA was converted to cDNA using SuperScript IV VILO Master Mix (ThermoFisher Scientific; Cat. No. 11756500). Quantitative PCR was performed using SYBR Premier Ex Taq (Takara; Cat No. RR420L) and quantified using the QuantStudio 6 Flex Real-Time PCR system (Applied Biosystems). The primers for ACE2 were as follows: (set 1) 5'-CGAAGCCGAAGACCTGTTCTA and 5'-GGGCAAGTGTGGACTGTTCC and (set 2) 5'-CAAGAGCAAACGGTTGAACAC and 5'-CCAGAGCCTCTCATTGTAGTCT. GAPDH was assessed as a control using primers 5'-TGCACCACCAACTGCTTAGC and 5'-GGCATGGACTGTGGTCATGAG.

Western blotting

Western blotting was performed as described in ref[147]. In short, whole cell lysates were harvested and resuspended in RIPA buffer [25 mM Tris, pH 7.4, 150 mM NaCl, 1% Triton X 100, 0.5% sodium deoxycholate, 0.1% sodium dodecyl sulfate, protease inhibitor cocktail (Sigma; Cat. No. 4693159001), and phosphatase inhibitor cocktail (Sigma, Cat. No. 4906845001)]. Protein lysates were loaded onto a 10% SDS-PAGE gel. The Trans-Blot Turbo Transfer System (Bio-Rad) and polyvinylidene difluoride membranes were used for protein transfer. Antibody blocking was done with 5% milk in TBST (19 mM Tris base, NaCl 137 mM, KCl 2.7 mM and 0.1% Tween-20) for 1 hour at room temperature. Primary antibodies used were ACE2 (Abcam; Cat. No. ab108252), used at a 1:1000 dilution in 5% milk, and GAPDH (Santa Cruz Biotechnology; Cat. No. sc-365062), used at a 1:1000 dilution in 5% milk. Blots were incubated overnight at 4°C. Membranes were then washed three times for 10 minutes each before they were incubated in secondary antibodies for an hour at room temperature. For the ACE2 blot, goat anti-rabbit IgG H&L (HRP) (Abcam; Cat. No. ab-97051) was used as a secondary antibody at a 1:20000 dilution in 5% milk. For the GAPDH blot, goat anti-mouse IgG (H + L)-HRP Conjugate (Bio-Rad; Cat. No. 1706516) was used as a secondary antibody at a 1:20000 dilution in 5% milk. Membranes were washed three more times and then developed using ProtoGlow ECl (National Diagnostics; Cat. No. CL-300).

Supplemental Figures and Tables

Figure S1. ACE2 expression varies between different mammalian tissues but is not strongly affected by aging. Related to Figure 1.

(A) The expression of ACE2 in 17 different mouse tissues or cell types, sorted according to mean ACE2 expression.

(B) The expression of ACE2 in 11 different rat tissues, sorted according to mean ACE2 expression.

(C) The expression of ACE2 in 30 different human tissues or cell types, sorted according to mean ACE2 expression.

(D) ACE2 expression in the lungs of a cohort of mice, sorted according to age.

(E) ACE2 expression in the lungs of an independent cohort of young mice (2-month old) and old mice (18-month old and 26-month old).

(F) ACE2 expression in the lungs of an independent cohort of young mice (7- to 9-week old) and old mice (14-month old).

Each panel displays log2-normalized ACE2 expression relative to a control group. Data analyzed in A and D were from GSE132040. Data analyzed in B were from GSE53960. Data analyzed in C were from www.gtexportal.org. Data analyzed in E were from GSE6591. Data analyzed in F were from GSE80680. Additional information on the data sources and sample sizes are included in Table S1.

Figure S2. Cigarette smoke increases the expression of the coronavirus-associated protease Cathepsin B. Related to Figure 2.

(A) The expression of TMPRSS2, Cathepsin B, and Cathepsin L in the lungs of mice that were sham-treated or that were exposed to diluted cigarette smoke for two, three, or four hours a day (low, medium, and high smoke exposure, respectively).

(B) The expression of TMPRSS2, Cathepsin B, and Cathepsin L in human trachea collected by fiberoptic brushing and analyzed according to smoking history.

(C) The expression of TMPRSS2, Cathepsin B, and Cathepsin L in human lung tissue in a cohort of smokers and analyzed according to the number of pack-years smoked.

(D) A volcano plot comparing gene expression in the respiratory epithelia of current smokers and never-smokers from Figure 2C, 2D, and 2E. The dotted lines indicate various p-value thresholds (e.g., genes located above the 10th percentile have a combined p value greater than 90% of the genes included in the meta-analysis).

Each panel displays log2-normalized expression of the indicated gene relative to a control group. Data analyzed in A were from GSE18344. Data analyzed in B were from GSE13933. Data analyzed in C were from GSE76925. Additional information on the data sources and sample sizes are included in Table S1. *, $p < .05$; **, $p < .005$; ***, $p < .0005$ (Student's t-test).

Figure S3. ACE2 is not commonly upregulated by lung diseases or carcinogen exposure. Related to Figure 4.

(A) ACE2 expression in normal lung tissue and in lung tissue from patients with COPD.

(B) ACE2 expression in normal lung tissue and in lung tissue from patients with COPD.

(C) ACE2 expression in lung tissue from patients with idiopathic pulmonary fibrosis or from histologically-normal lung tissue from patients with lung cancer.

(D) ACE2 expression in tissue from the lungs of transplant recipients with idiopathic pulmonary fibrosis or from control donor lungs.

(E) The expression of ACE2 in bronchial brushing from individuals with normal lung function, mild asthma, or severe asthma.

(F) The expression of ACE2 in disease-free lung tissue resected from patients with sarcoidosis or control tissue from healthy lung donors.

(G) The expression of ACE2 in the lungs of a cystic fibrosis mouse model (*Cftr*$^{-/-}$) or wild-type littermate controls.

(H) The expression of ACE2 in the lungs of mice provided with normal water, water with 10 parts per billion sodium arsenite, or water with 100 parts per billion sodium arsenite (low and high arsenic, respectively).

(I) The expression of ACE2 in the lungs of mice exposed to normal air or exposed to the carcinogen 1,3-butadiene.

(J) The expression of ACE2 in the lungs of mice that were sham-treated, exposed to 5 Gy of ionizing radiation to the thorax, or exposed to 17.5 Gy of ionizing radiation to the thorax (low and high IR, respectively).

Each panel displays log2-normalized ACE2 expression relative to a control group. Data analyzed in A were from GSE57148. Data analyzed in B were from GSE103174. Data analyzed in C were from GSE2052. Data analyzed in D were from GSE47460. Data analyzed in E were from GSE43696. Data analyzed in F were from GSE16538. Data analyzed in G were from GSE3100. Data analyzed in H were from GSE11056. Data analyzed in I were from GSE86623. Data analyzed in J were from GSE41789. Additional information on the data sources and sample sizes are included in Table S1. *, $p < .05$; **, $p < .005$; ***, $p < .0005$ (Student's t-test).

Figure S4. ACE2 RNA levels and protein levels are tightly correlated. Related to Figure 5.

(A) A scatter plot comparing the expression of ACE2 RNA, as measured by RNA-Seq, and ACE2 protein, as measured by mass spectrometry, across human cancer cell lines. The black line indicates a linear regression plotted against the data.

(B) A bar graph comparing the Pearson correlation coefficients between RNA and protein expression across 11,362 human genes. The grey bars denote the fraction of genes with

an indicated correlation coefficient while the orange line indicates the cumulative distribution of those coefficients. ACE2's position is indicated with a red star.

(C) A graph comparing ACE2 immunohistochemistry staining in various tissues vs ACE2 RNA expression in those same tissues.

(D) qRT-PCR of ACE2 expression in the indicated cell lines, relative to GAPDH expression.

(E) Western blotting of ACE2 expression in the indicated cell lines. GAPDH was analyzed as a loading control.

The data in A and B are from ref[106]. The data in C is from the Human Protein Atlas[138]. Additional information on these sources is included in Table S1.

Table S1. Data sources used for this analysis.

Table S2. ACE2-correlated genes and GO terms.

References

1. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *New England Journal of Medicine* **382**, 727–733 (2020).
2. Wu, F. *et al.* A new coronavirus associated with human respiratory disease in China. *Nature* **579**, 265–269 (2020).
3. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
4. Total confirmed deaths due to COVID-19. *Our World in Data* https://ourworldindata.org/grapher/total-deaths-covid-19.
5. Li, W. *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450–454 (2003).
6. 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) doi:10.1016/j.cell.2020.02.052.
7. Jiang, F. *et al.* Angiotensin-converting enzyme 2 and angiotensin 1–7: novel therapeutic targets. *Nature Reviews Cardiology* **11**, 413–426 (2014).
8. Kuba, K. *et al.* A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus–induced lung injury. *Nature Medicine* **11**, 875–879 (2005).
9. McCray, P. B. *et al.* Lethal Infection of K18-hACE2 Mice Infected with Severe Acute Respiratory Syndrome Coronavirus. *Journal of Virology* **81**, 813–821 (2007).
10. Wu, Z. & McGoogan, J. M. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72 314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA* (2020) doi:10.1001/jama.2020.2648.

11. Ruan, S. Likelihood of survival of coronavirus disease 2019. *The Lancet Infectious Diseases* **0**, (2020).
12. Luo, X. *et al.* Characteristics of patients with COVID-19 during epidemic ongoing outbreak in Wuhan, China. *medRxiv* 2020.03.19.20033175 (2020) doi:10.1101/2020.03.19.20033175.
13. Guan, W.-J. *et al.* Clinical Characteristics of Coronavirus Disease 2019 in China. *N. Engl. J. Med.* (2020) doi:10.1056/NEJMoa2002032.
14. Remuzzi, A. & Remuzzi, G. COVID-19 and Italy: what next? *The Lancet* **0**, (2020).
15. Rabin, R. C. In Italy, Coronavirus Takes a Higher Toll on Men. *The New York Times* (2020).
16. Zhao, Q. *et al.* The impact of COPD and smoking history on the severity of Covid-19: A systemic review and meta-analysis. *J. Med. Virol.* (2020) doi:10.1002/jmv.25889.
17. Patanavanich, R. & Glantz, S. A. Smoking is Associated with COVID-19 Progression: A Meta-Analysis. *medRxiv* 2020.04.13.20063669 (2020) doi:10.1101/2020.04.13.20063669.
18. Pedersen, S. F. & Ho, Y.-C. SARS-CoV-2: a storm is raging. *J Clin Invest* **130**, (2020).
19. Yang, Y. *et al.* Exuberant elevation of IP-10, MCP-3 and IL-1ra during SARS-CoV-2 infection is associated with disease severity and fatal outcome. *medRxiv* 2020.03.02.20029975 (2020) doi:10.1101/2020.03.02.20029975.
20. Shi, Y. *et al.* Immunopathological characteristics of coronavirus disease 2019 cases in Guangzhou, China. *medRxiv* 2020.03.12.20034736 (2020) doi:10.1101/2020.03.12.20034736.
21. Chen, G. *et al.* Clinical and immunological features of severe and moderate coronavirus disease 2019. *J. Clin. Invest.* (2020) doi:10.1172/JCI137244.
22. Zhang, C., Wu, Z., Li, J.-W., Zhao, H. & Wang, G.-Q. The cytokine release syndrome (CRS) of severe COVID-19 and Interleukin-6 receptor (IL-6R) antagonist Tocilizumab may be the key to reduce the mortality. *International Journal of Antimicrobial Agents* 105954 (2020) doi:10.1016/j.ijantimicag.2020.105954.

23. Lonsdale, J. *et al.* The Genotype-Tissue Expression (GTEx) project. *Nature Genetics* **45**, 580–585 (2013).
24. Carithers, L. J. *et al.* A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreservation and Biobanking* **13**, 311–319 (2015).
25. Yu, Y. *et al.* A rat RNA-Seq transcriptomic BodyMap across 11 organs and 4 developmental stages. *Nature Communications* **5**, 1–11 (2014).
26. Schaum, N. *et al.* The murine transcriptome reveals global aging nodes with organ-specific phase and amplitude. *bioRxiv* 662254 (2019) doi:10.1101/662254.
27. Wölfel, R. *et al.* Virological assessment of hospitalized patients with COVID-2019. *Nature* 1–10 (2020) doi:10.1038/s41586-020-2196-x.
28. Xu, Y. *et al.* Characteristics of pediatric SARS-CoV-2 infection and potential evidence for persistent fecal viral shedding. *Nature Medicine* 1–4 (2020) doi:10.1038/s41591-020-0817-4.
29. Incidence, clinical characteristics and prognostic factor of patients with COVID-19: a systematic review and meta-analysis | medRxiv. https://www.medrxiv.org/content/10.1101/2020.03.17.20037572v1.
30. Jonker, M. J. *et al.* Life spanning murine gene expression profiles in relation to chronological and pathological aging in multiple organs. *Aging Cell* **12**, 901–909 (2013).
31. Misra, V. *et al.* Global expression profiles from C57BL/6J and DBA/2J mouse lungs to determine aging-related genes. *Physiol. Genomics* **31**, 429–440 (2007).
32. Soundararajan, R. *et al.* Global gene profiling of aging lungs in Atp8b1 mutant mice. *Aging (Albany NY)* **8**, 2232–2252 (2016).
33. Consortium, T. T. M. *et al.* A Single Cell Transcriptomic Atlas Characterizes Aging Tissues in the Mouse. *bioRxiv* 661728 (2020) doi:10.1101/661728.
34. Crowley, J. J. *et al.* Analyses of allele-specific gene expression in highly divergent mouse crosses identifies pervasive allelic imbalance. *Nat. Genet.* **47**, 353–360 (2015).

35. Gruber, M. P. *et al.* Human lung project: evaluating variance of gene expression in the human lung. *Am. J. Respir. Cell Mol. Biol.* **35**, 65–71 (2006).
36. The Cancer Genome Atlas Research Network *et al.* The Cancer Genome Atlas Pan-Cancer analysis project. *Nat Genet* **45**, 1113–1120 (2013).
37. Qi, D. *et al.* Epidemiological and clinical features of 2019-nCoV acute respiratory disease cases in Chongqing municipality, China: a retrospective, descriptive, multiple-center study. *medRxiv* 2020.03.01.20029397 (2020) doi:10.1101/2020.03.01.20029397.
38. Zhang, R. *et al.* CT features of SARS-CoV-2 pneumonia according to clinical presentation: a retrospective analysis of 120 consecutive patients from Wuhan city. *Eur Radiol* (2020) doi:10.1007/s00330-020-06854-1.
39. Wang, R. *et al.* Epidemiological and clinical features of 125 Hospitalized Patients with COVID-19 in Fuyang, Anhui, China. *Int J Infect Dis* (2020) doi:10.1016/j.ijid.2020.03.070.
40. Hu, L. Risk Factors Associated with Clinical Outcomes in 323 COVID-19 Patients in Wuhan, China | medRxiv. *medRxiv* (2020).
41. Gebel, S. *et al.* The transcriptome of Nrf2-/- mice provides evidence for impaired cell cycle progression in the development of cigarette smoke-induced emphysematous changes. *Toxicol. Sci.* **115**, 238–252 (2010).
42. Turetz, M. L. *et al.* Trachea epithelium as a ‘canary’ for cigarette smoking-induced biologic phenotype of the small airway epithelium. *Clin Transl Sci* **2**, 260–272 (2009).
43. Butler, M. W. *et al.* Modulation of cystatin A expression in human airway epithelium related to genotype, smoking, COPD, and lung cancer. *Cancer Res.* **71**, 2572–2581 (2011).
44. Yang, J. *et al.* Smoking-Dependent Distal-to-Proximal Repatterning of the Adult Human Small Airway Epithelium. *Am. J. Respir. Crit. Care Med.* **196**, 340–352 (2017).
45. Morrow, J. D. *et al.* Functional interactors of three genome-wide association study genes are differentially expressed in severe chronic obstructive pulmonary disease lung tissue. *Scientific Reports* **7**, 1–11 (2017).

46. Millet, J. K. & Whittaker, G. R. Host cell proteases: Critical determinants of coronavirus tropism and pathogenesis. *Virus Research* **202**, 120–134 (2015).

47. Franks, T. J. *et al.* Resident cellular components of the human lung: current knowledge and goals for research on cell phenotyping and function. *Proc Am Thorac Soc* **5**, 763–766 (2008).

48. Whitsett, J. A. Airway Epithelial Differentiation and Mucociliary Clearance. *Annals ATS* **15**, S143–S148 (2018).

49. Whitsett, J. A. & Weaver, T. E. Alveolar Development and Disease. *Am J Respir Cell Mol Biol* **53**, 1–7 (2015).

50. Reyfman, P. A. *et al.* Single-Cell Transcriptomic Analysis of Human Lung Provides Insights into the Pathobiology of Pulmonary Fibrosis. *Am J Respir Crit Care Med* **199**, 1517–1536 (2018).

51. Wolf, F. A., Angerer, P. & Theis, F. J. SCANPY: large-scale single-cell gene expression data analysis. *Genome Biology* **19**, 15 (2018).

52. Ardini-Poleske, M. E. *et al.* LungMAP: The Molecular Atlas of Lung Development Program. *American Journal of Physiology-Lung Cellular and Molecular Physiology* **313**, L733–L740 (2017).

53. Plasschaert, L. W. *et al.* A single-cell atlas of the airway epithelium reveals the CFTR-rich pulmonary ionocyte. *Nature* **560**, 377–381 (2018).

54. Angelidis, I. *et al.* An atlas of the aging lung mapped by single cell transcriptomics and deep tissue proteomics. *Nature Communications* **10**, 1–17 (2019).

55. Kasper, M., Behrens, J., Schuh, D. & Müller, M. Distribution of E-cadherin and Ep-CAM in the human lung during development and after injury. *Histochem Cell Biol* **103**, 281–286 (1995).

56. Thomas, M. L. The leukocyte common antigen family. *Annu. Rev. Immunol.* **7**, 339–369 (1989).

57. Li, R. *et al.* Pdgfra marks a cellular lineage with distinct contributions to myofibroblasts in lung maturation and injury response. *eLife* **7**, e36865 (2018).

58. Moon, E.-H. *et al.* Generation of mice with a conditional and reporter allele for Tmem100. *genesis* **48**, 673–678 (2010).

59. Hovenberg, H. W., Davies, J. R. & Carlstedt, I. Different mucins are produced by the surface epithelium and the submucosa in human trachea: identification of MUC5AC as a major mucin from the goblet cells. *Biochem J* **318**, 319–324 (1996).

60. Rawlins, E. L. *et al.* The role of Scgb1a1+ Clara cells in the long-term maintenance and repair of lung airway, but not alveolar, epithelium. *Cell Stem Cell* **4**, 525–534 (2009).

61. Guha, A. *et al.* Analysis of Notch Signaling-Dependent Gene Expression in Developing Airways Reveals Diversity of Clara Cells. *PLoS One* **9**, (2014).

62. Salaun, B. *et al.* CD208/Dendritic Cell-Lysosomal Associated Membrane Protein Is a Marker of Normal and Transformed Type II Pneumocytes. *Am J Pathol* **164**, 861–871 (2004).

63. Yu, X., Ng, C. P., Habacher, H. & Roy, S. Foxj1 transcription factors are master regulators of the motile ciliogenic program. *Nat. Genet.* **40**, 1445–1453 (2008).

64. Cosio, M. G., Hale, K. A., Niewoehner, D. E. & Markert, M. Morphologic and Morphometric Effects of Prolonged Cigarette Smoking on the Small Airways. *Am Rev Respir Dis* **122**, 265–271 (1980).

65. Saetta, M. *et al.* Goblet Cell Hyperplasia and Epithelial Inflammation in Peripheral Airways of Smokers with Both Symptoms of Chronic Bronchitis and Chronic Airflow Limitation. *Am J Respir Crit Care Med* **161**, 1016–1021 (2000).

66. Liu, Y. & Di, Y. P. Effects of second hand smoke on airway secretion and mucociliary clearance. *Front. Physiol.* **3**, (2012).

67. Goldfarbmuren, K. C. *et al.* Dissecting the cellular specificity of smoking effects and reconstructing lineages in the human airway epithelium. *bioRxiv* 612747 (2019) doi:10.1101/612747.

68. Rock, J. R., Randell, S. H. & Hogan, B. L. M. Airway basal stem cells: a perspective on their roles in epithelial homeostasis and remodeling. *Disease Models & Mechanisms* **3**, 545–556 (2010).

69. Montoro, D. T. *et al.* A revised airway epithelial hierarchy includes CFTR-expressing ionocytes. *Nature* **560**, 319–324 (2018).

70. Gu, X. *et al.* Chemosensory Functions for Pulmonary Neuroendocrine Cells. *Am J Respir Cell Mol Biol* **50**, 637–646 (2013).

71. Tomita, K., van Bokhoven, A., Jansen, C. F. J., Bussemakers, M. J. G. & Schalken, J. A. Coordinate Recruitment of E-Cadherin and ALCAM to Cell–Cell Contacts by α-Catenin. *Biochemical and Biophysical Research Communications* **267**, 870–874 (2000).

72. Farkas, A. E. *et al.* HNF4α Regulates Claudin-7 Protein Expression during Intestinal Epithelial Differentiation. *The American Journal of Pathology* **185**, 2206–2218 (2015).

73. González-Mariscal, L., Betanzos, A., Nava, P. & Jaramillo, B. E. Tight junction proteins. *Progress in Biophysics and Molecular Biology* **81**, 1–44 (2003).

74. Fahy, J. V. & Dickey, B. F. Airway Mucus Function and Dysfunction. *New England Journal of Medicine* **363**, 2233–2247 (2010).

75. Tamashiro, E. *et al.* Cigarette Smoke Exposure Impairs Respiratory Epithelial Ciliogenesis. *Am J Rhinol□Allergy* **23**, 117–122 (2009).

76. Upadhyay, S. & Palmberg, L. Air-Liquid Interface: Relevant In Vitro Models for Investigating Air Pollutant-Induced Pulmonary Toxicity. *Toxicol Sci* **164**, 21–30 (2018).

77. Jiang, D., Schaefer, N. & Chu, H. W. Air–Liquid Interface Culture of Human and Mouse Airway Epithelial Cells. in *Lung Innate Immunity and Inflammation: Methods and Protocols* (eds. Alper, S. & Janssen, W. J.) 91–109 (Springer, 2018). doi:10.1007/978-1-4939-8570-8_8.

78. Ross, A. J., Dailey, L. A., Brighton, L. E. & Devlin, R. B. Transcriptional Profiling of Mucociliary Differentiation in Human Airway Epithelial Cells. *Am J Respir Cell Mol Biol* **37**, 169–185 (2007).
79. Nemajerova, A. *et al.* TAp73 is a central transcriptional regulator of airway multiciliogenesis. *Genes Dev.* **30**, 1300–1312 (2016).
80. Martinez-Anton, A. *et al.* Changes in microRNA and mRNA expression with differentiation of human bronchial epithelial cells. *Am. J. Respir. Cell Mol. Biol.* **49**, 384–395 (2013).
81. GEO Accession viewer. https://www.ncbi.nlm.nih.gov/geo/query/acc.cgi?acc=GSE135188.
82. McDonough, J. E. *et al.* Transcriptional regulatory model of fibrosis progression in the human lung. *JCI Insight* **4**, (2019).
83. Pardo, A. *et al.* Up-Regulation and Profibrotic Role of Osteopontin in Human Idiopathic Pulmonary Fibrosis. *PLOS Medicine* **2**, e251 (2005).
84. Cruz, T. *et al.* Multi-level immune response network in mild-moderate Chronic Obstructive Pulmonary Disease (COPD). *Respiratory Research* **20**, 152 (2019).
85. Kim, W. J. *et al.* Comprehensive Analysis of Transcriptome Sequencing Data in the Lung Tissues of COPD Subjects. *Int J Genomics* **2015**, 206937 (2015).
86. Baumgartner, K. B., Samet, J. M., Stidley, C. A., Colby, T. V. & Waldron, J. A. Cigarette smoking: a risk factor for idiopathic pulmonary fibrosis. *Am J Respir Crit Care Med* **155**, 242–248 (1997).
87. Laniado-Laborín, R. Smoking and Chronic Obstructive Pulmonary Disease (COPD). Parallel Epidemics of the 21st Century. *Int J Environ Res Public Health* **6**, 209–224 (2009).
88. Lippi, G. & Henry, B. M. Chronic obstructive pulmonary disease is associated with severe coronavirus disease 2019 (COVID-19). *Respir Med* (2020) doi:10.1016/j.rmed.2020.105941.
89. Voraphani, N. *et al.* An airway epithelial iNOS–DUOX2–thyroid peroxidase metabolome drives Th1/Th2 nitrative stress in human severe asthma. *Mucosal Immunology* **7**, 1175–1185 (2014).

90. Crouser, E. D. *et al.* Gene Expression Profiling Identifies MMP-12 and ADAMDEC1 as Potential Pathogenic Mediators of Pulmonary Sarcoidosis. *Am J Respir Crit Care Med* **179**, 929–938 (2009).
91. Citrin, D. E. *et al.* Role of type II pneumocyte senescence in radiation-induced lung fibrosis. *J. Natl. Cancer Inst.* **105**, 1474–1484 (2013).
92. Kozul Courtney D. *et al.* Chronic Exposure to Arsenic in the Drinking Water Alters the Expression of Immune Response Genes in Mouse Lung. *Environmental Health Perspectives* **117**, 1108–1115 (2009).
93. Chappell Grace A. *et al.* Variation in DNA-Damage Responses to an Inhalational Carcinogen (1,3-Butadiene) in Relation to Strain-Specific Differences in Chromatin Accessibility and Gene Transcription Profiles in C57BL/6J and CAST/EiJ Mice. *Environmental Health Perspectives* **125**, 107006.
94. Haston, C. K., Cory, S., Lafontaine, L., Dorion, G. & Hallett, M. T. Strain-dependent pulmonary gene expression profiles of a cystic fibrosis mouse model. *Physiol. Genomics* **25**, 336–345 (2006).
95. Mogensen, T. H. & Paludan, S. R. Molecular Pathways in Virus-Induced Cytokine Production. *Microbiol Mol Biol Rev* **65**, 131–150 (2001).
96. Gan, W. Q., Man, S. F. P. & Sin, D. D. The Interactions Between Cigarette Smoking and Reduced Lung Function on Systemic Inflammation. *Chest* **127**, 558–564 (2005).
97. Arnson, Y., Shoenfeld, Y. & Amital, H. Effects of tobacco smoke on immunity, inflammation and autoimmunity. *Journal of Autoimmunity* **34**, J258–J265 (2010).
98. Ioannidis, I. *et al.* Plasticity and virus specificity of the airway epithelial cell immune response during respiratory virus infection. *J. Virol.* **86**, 5422–5436 (2012).
99. Menachery, V. D. *et al.* MERS-CoV Accessory ORFs Play Key Role for Infection and Pathogenesis. *mBio* **8**, (2017).

100. Mitchell, H. D. *et al.* A network integration approach to predict conserved regulators related to pathogenicity of influenza and SARS-CoV respiratory viruses. *PLoS ONE* **8**, e69374 (2013).

101. Gantier, M. P. & Williams, B. R. G. The response of mammalian cells to double-stranded RNA. *Cytokine Growth Factor Rev* **18**, 363–371 (2007).

102. Wagener, A. H. *et al.* dsRNA-induced changes in gene expression profiles of primary nasal and bronchial epithelial cells from patients with asthma, rhinitis and controls. *Respir Res* **15**, 9 (2014).

103. Rusinova, I. *et al.* INTERFEROME v2.0: an updated database of annotated interferon-regulated genes. *Nucleic Acids Res* **41**, D1040–D1046 (2013).

104. Shapira, S. D. *et al.* A physical and regulatory map of host-influenza interactions reveals pathways in H1N1 infection. *Cell* **139**, 1255–1267 (2009).

105. Nusinow, D. P. & Gygi, S. P. A Guide to the Quantitative Proteomic Profiles of the Cancer Cell Line Encyclopedia. *bioRxiv* 2020.02.03.932384 (2020) doi:10.1101/2020.02.03.932384.

106. Nusinow, D. P. *et al.* Quantitative Proteomics of the Cancer Cell Line Encyclopedia. *Cell* **180**, 387-402.e16 (2020).

107. Barretina, J. *et al.* The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**, 603–607 (2012).

108. Uhlén, M. *et al.* Tissue-based map of the human proteome. *Science* **347**, 1260419 (2015).

109. Ferrari, M. F. R., Raizada, M. K. & Fior-Chadi, D. R. Nicotine modulates the renin-angiotensin system of cultured neurons and glial cells from cardiovascular brain areas of Wistar Kyoto and spontaneously hypertensive rats. *J. Mol. Neurosci.* **33**, 284–293 (2007).

110. Ferrari, M. F. R., Raizada, M. K. & Fior-Chadi, D. R. Differential regulation of the renin-angiotensin system by nicotine in WKY and SHR glia. *J. Mol. Neurosci.* **35**, 151–160 (2008).

111. Yue, X. *et al.* Nicotine Downregulates the Compensatory Angiotensin-Converting Enzyme 2/Angiotensin Type 2 Receptor of the Renin–Angiotensin System. *Annals ATS* **15**, S126–S127 (2018).

112. Oakes, J. M., Fuchs, R. M., Gardner, J. D., Lazartigues, E. & Yue, X. Nicotine and the renin-angiotensin system. *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology* **315**, R895–R906 (2018).

113. Cai, G. Tobacco-Use Disparity in Gene Expression of ACE2, the Receptor of 2019-nCov. (2020) doi:10.20944/preprints202002.0051.v1.

114. Brake, S. J. *et al.* Smoking Upregulates Angiotensin-Converting Enzyme-2 Receptor: A Potential Adhesion Site for Novel Coronavirus SARS-CoV-2 (Covid-19). *Journal of Clinical Medicine* **9**, 841 (2020).

115. Leung, J. M. *et al.* ACE-2 Expression in the Small Airway Epithelia of Smokers and COPD Patients: Implications for COVID-19. *Eur Respir J* (2020) doi:10.1183/13993003.00688-2020.

116. Yilin, Z., Yandong, N. & Faguang, J. Role of angiotensin-converting enzyme (ACE) and ACE2 in a rat model of smoke inhalation induced acute respiratory distress syndrome. *Burns* **41**, 1468–1477 (2015).

117. Ding, Y. *et al.* Organ distribution of severe acute respiratory syndrome (SARS) associated coronavirus (SARS-CoV) in SARS patients: implications for pathogenesis and virus transmission pathways. *The Journal of Pathology* **203**, 622–630 (2004).

118. Hamming, I. *et al.* Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J. Pathol.* **203**, 631–637 (2004).

119. Jia, H. P. *et al.* ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia. *J. Virol.* **79**, 14614–14621 (2005).

120. Xudong, X., Junzhu, C., Xingxiang, W., Furong, Z. & Yanrong, L. Age- and gender-related difference of ACE2 expression in rat lung. *Life Sciences* **78**, 2166–2171 (2006).

121. Wiener, R. S., Cao, Y. X., Hinds, A., Ramirez, M. I. & Williams, M. C. Angiotensin converting enzyme 2 is primarily epithelial and is developmentally regulated in the mouse lung. *Journal of Cellular Biochemistry* **101**, 1278–1291 (2007).

122. 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* **0**, (2020).

123. Sungnak, W. *et al.* SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nature Medicine* 1–7 (2020) doi:10.1038/s41591-020-0868-6.

124. Lukassen, S. *et al.* SARS-CoV-2 receptor ACE2 and TMPRSS2 are primarily expressed in bronchial transient secretory cells. *The EMBO Journal* **n/a**, e105114 (2020).

125. Travaglini, K. J. *et al.* A molecular cell atlas of the human lung from single cell RNA sequencing. *bioRxiv* 742320 (2020) doi:10.1101/742320.

126. Wu, C. & Zheng, M. Single-Cell RNA Expression Profiling Shows that ACE2, the Putative Receptor of Wuhan 2019-nCoV, Has Significant Expression in the Nasal, Mouth, Lung and Colon Tissues, and Tends to be Co-Expressed with HLA-DRB1 in the Four Tissues. *Preprints* (2020).

127. Zhao, Y. *et al.* Single-cell RNA expression profiling of ACE2, the putative receptor of Wuhan 2019-nCov. *bioRxiv* 2020.01.26.919985 (2020) doi:10.1101/2020.01.26.919985.

128. Bradbury, A. & Plückthun, A. Reproducibility: Standardize antibodies used in research. *Nature News* **518**, 27 (2015).

129. Roumier, M., Paule, R., Groh, M., Vallee, A. & Ackermann, F. Interleukin-6 blockade for severe COVID-19. *medRxiv* 2020.04.20.20061861 (2020) doi:10.1101/2020.04.20.20061861.

130. Office of the Surgeon General (US) & Office on Smoking and Health (US). *The Health Consequences of Smoking: A Report of the Surgeon General*. (Centers for Disease Control and Prevention (US), 2004).

131. Bertram, J. F. & Rogers, A. W. Recovery of bronchial epithelium on stopping smoking. *Br Med J (Clin Res Ed)* **283**, 1567–1569 (1981).

132. Lee, J. J. *et al.* Long-Term Impact of Smoking on Lung Epithelial Proliferation in Current and Former Smokers. *J Natl Cancer Inst* **93**, 1081–1088 (2001).

133. Muus, C. *et al.* Integrated analyses of single-cell atlases reveal age, gender, and smoking status associations with cell type-specific expression of mediators of SARS-CoV-2 viral entry and highlights inflammatory programs in putative target cells. *bioRxiv* 2020.04.19.049254 (2020) doi:10.1101/2020.04.19.049254.

134. Decrease in ACE2 mRNA expression in aged mouse lung | bioRxiv. https://www.biorxiv.org/content/10.1101/2020.04.02.021451v1.

135. Imai, Y. *et al.* Angiotensin-converting enzyme 2 protects from severe acute lung failure. *Nature* **436**, 112–116 (2005).

136. Rey-Parra, G. J. *et al.* Angiotensin converting enzyme 2 abrogates bleomycin-induced lung injury. *J Mol Med* **90**, 637–647 (2012).

137. Zang, R. *et al.* TMPRSS2 and TMPRSS4 mediate SARS-CoV-2 infection of human small intestinal enterocytes. *bioRxiv* 2020.04.21.054015 (2020) doi:10.1101/2020.04.21.054015.

138. Uhlen, M. *et al.* Towards a knowledge-based Human Protein Atlas. *Nat. Biotechnol.* **28**, 1248–1250 (2010).

139. Edgar, R., Domrachev, M. & Lash, A. E. Gene Expression Omnibus: NCBI gene expression and hybridization array data repository. *Nucl. Acids Res.* **30**, 207–210 (2002).

140. Broad GDAC Firehose. https://gdac.broadinstitute.org/.

141. Regev, A. *et al.* The Human Cell Atlas. *Elife* **6**, (2017).

142. Papatheodorou, I. *et al.* Expression Atlas update: from tissues to single cells. *Nucleic Acids Res* **48**, D77–D83 (2020).

143. Seabold, S. & Perktold, J. Statsmodels: Econometric and Statistical Modeling with Python. in 92–96 (2010). doi:10.25080/Majora-92bf1922-011.

144. Ulyanov, D. *DmitryUlyanov/Multicore-TSNE*. (2020).

145. Kobak, D. & Berens, P. The art of using t-SNE for single-cell transcriptomics. *Nature Communications* **10**, 1–14 (2019).

146. Raudvere, U. *et al.* g:Profiler: a web server for functional enrichment analysis and conversions of gene lists (2019 update). *Nucleic Acids Res* **47**, W191–W198 (2019).

147. Vasudevan, A. *et al.* Single-Chromosomal Gains Can Function as Metastasis Suppressors and Promoters in Colon Cancer. *Developmental Cell* **52**, 413-428.e6 (2020).

Figure 1

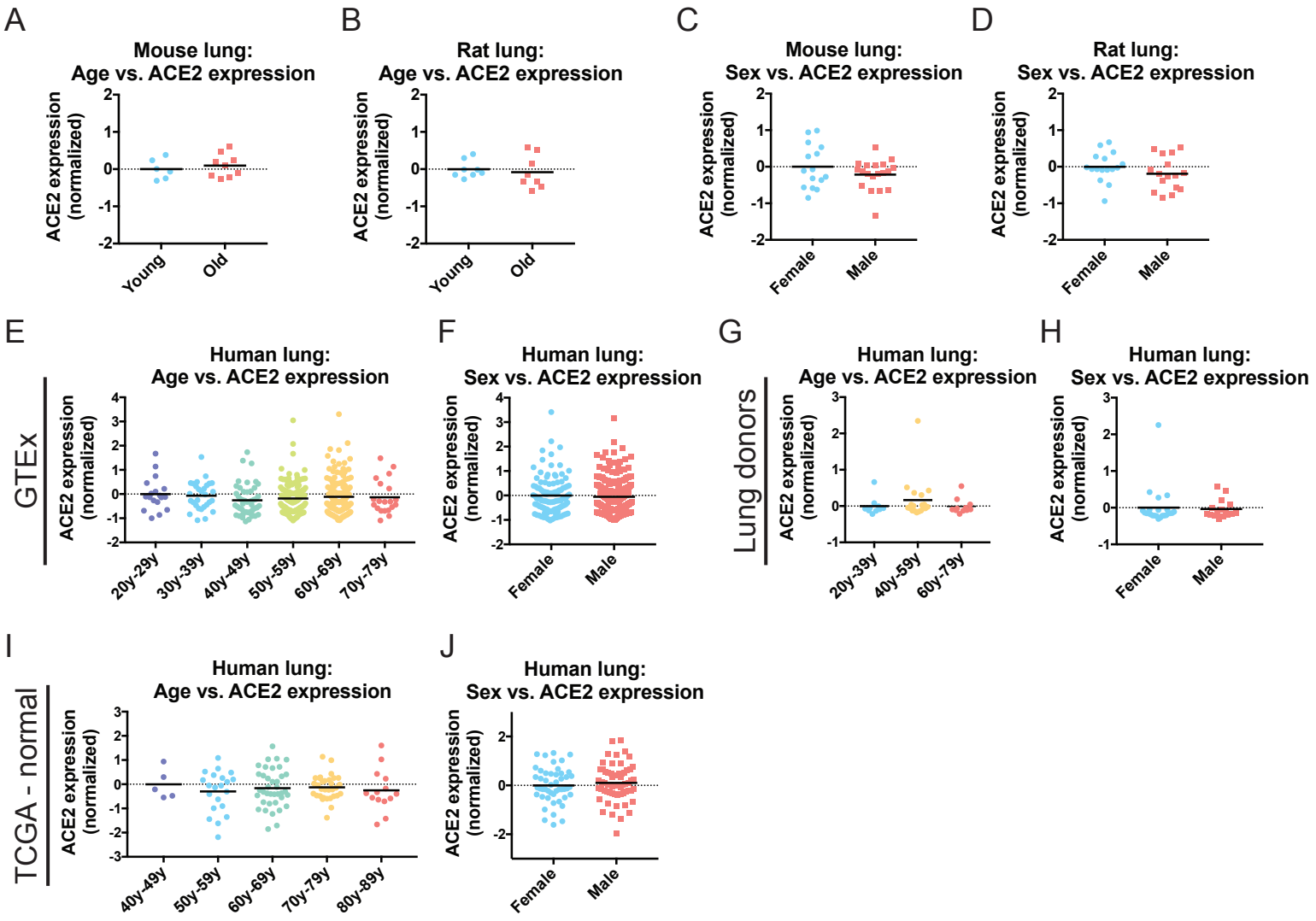
A
Mouse lung:
Age vs. ACE2 expression
ACE2 expression (normalized)
Young
Old
B
Rat lung:
Age vs. ACE2 expression
ACE2 expression (normalized)
Young
Old
C
Mouse lung:
Sex vs. ACE2 expression
ACE2 expression (normalized)
Female
Male
D
Rat lung:
Sex vs. ACE2 expression
ACE2 expression (normalized)
Female
Male
E
GTEx
Human lung:
Age vs. ACE2 expression
ACE2 expression (normalized)
20y-29y
30y-39y
40y-49y
50y-59y
60y-69y
70y-79y
F
Human lung:
Sex vs. ACE2 expression
ACE2 expression (normalized)
Female
Male
G
Lung donors
Human lung:
Age vs. ACE2 expression
ACE2 expression (normalized)
20y-39y
40y-59y
60y-79y
H
Human lung:
Sex vs. ACE2 expression
ACE2 expression (normalized)
Female
Male
I
TCGA - normal
Human lung:
Age vs. ACE2 expression
ACE2 expression (normalized)
40y-49y
50y-59y
60y-69y
70y-79y
80y-89y
J
Human lung:
Sex vs. ACE2 expression
ACE2 expression (normalized)
Female
Male

Figure 2

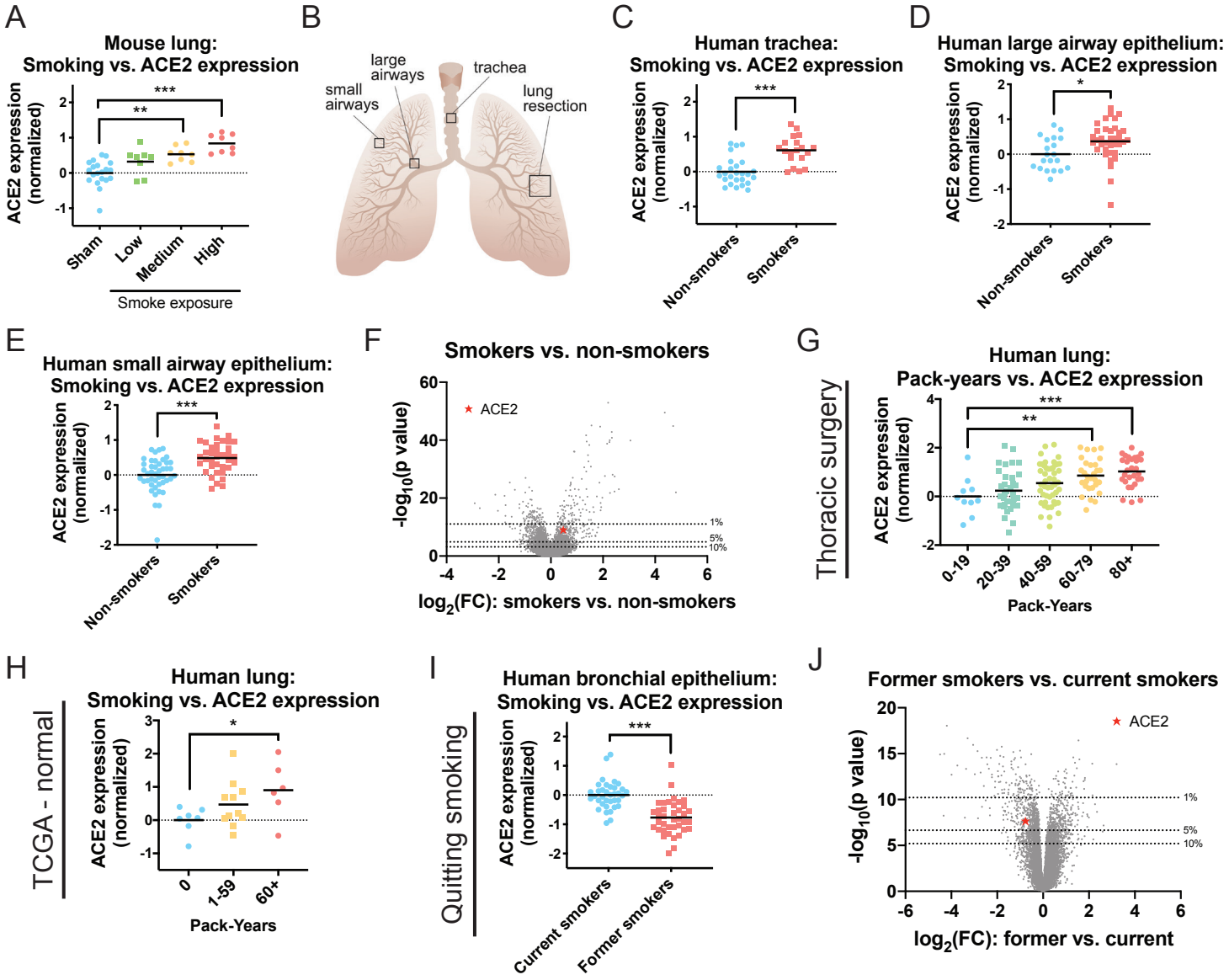
A
Mouse lung:
Smoking vs. ACE2 expression
ACE2 expression (normalized)
Sham
Low
Medium
High
Smoke exposure
B
large airways
small airways
trachea
lung resection
C
Human trachea:
Smoking vs. ACE2 expression
Non-smokers
Smokers
D
Human large airway epithelium:
Smoking vs. ACE2 expression
E
Human small airway epithelium:
Smoking vs. ACE2 expression
F
Smokers vs. non-smokers
ACE2
-log10(p value)
log2(FC): smokers vs. non-smokers
1%
5%
10%
G
Thoracic surgery
Human lung:
Pack-years vs. ACE2 expression
0-19
20-39
40-59
60-79
80+
Pack-Years
H
TCGA - normal
Human lung:
Smoking vs. ACE2 expression
0
1-59
60+
Pack-Years
I
Quitting smoking
Human bronchial epithelium:
Smoking vs. ACE2 expression
Current smokers
Former smokers
J
Former smokers vs. current smokers
ACE2
-log10(p value)
log2(FC): former vs. current

Figure 3

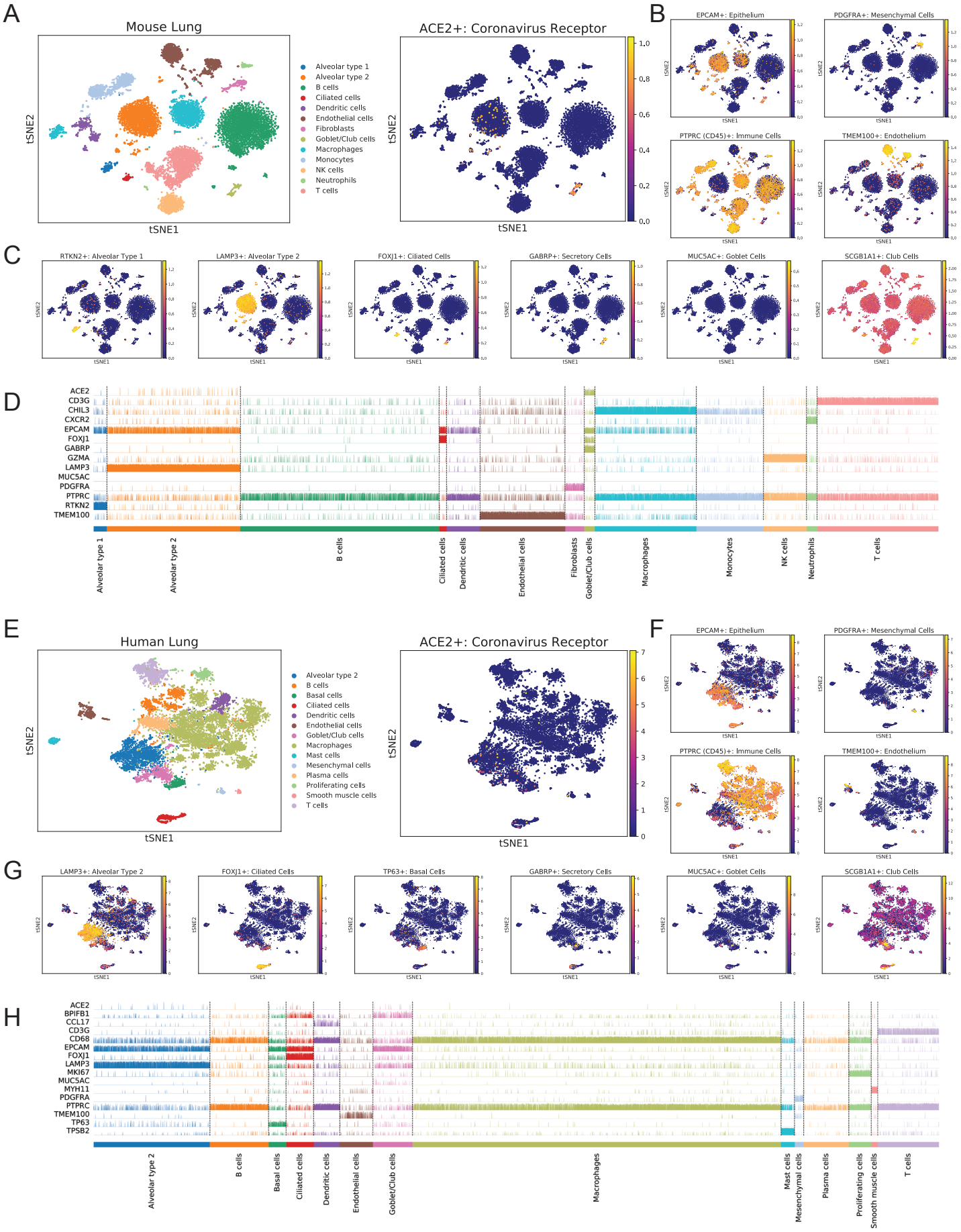
A
Mouse Lung
tSNE2
tSNE1
Alveolar type 1
Alveolar type 2
B cells
Ciliated cells
Dendritic cells
Endothelial cells
Fibroblasts
Goblet/Club cells
Macrophages
Monocytes
NK cells
Neutrophils
T cells
ACE2+: Coronavirus Receptor
B
EPCAM+: Epithelium
PDGFRA+: Mesenchymal Cells
PTPRC (CD45)+: Immune Cells
TMEM100+: Endothelium
C
RTKN2+: Alveolar Type 1
LAMP3+: Alveolar Type 2
FOXJ1+: Ciliated Cells
GABRP+: Secretory Cells
MUC5AC+: Goblet Cells
SCGB1A1+: Club Cells
D
ACE2
CD3G
CHIL3
CXCR2
EPCAM
FOXJ1
GABRP
GZMA
LAMP3
MUC5AC
PDGFRA
PTPRC
RTKN2
TMEM100
Alveolar type 1
Alveolar type 2
B cells
Ciliated cells
Dendritic cells
Endothelial cells
Fibroblasts
Goblet/Club cells
Macrophages
Monocytes
NK cells
Neutrophils
T cells
E
Human Lung
Alveolar type 2
B cells
Basal cells
Ciliated cells
Dendritic cells
Endothelial cells
Goblet/Club cells
Macrophages
Mast cells
Mesenchymal cells
Plasma cells
Proliferating cells
Smooth muscle cells
T cells
ACE2+: Coronavirus Receptor
F
EPCAM+: Epithelium
PDGFRA+: Mesenchymal Cells
PTPRC (CD45)+: Immune Cells
TMEM100+: Endothelium
G
LAMP3+: Alveolar Type 2
FOXJ1+: Ciliated Cells
TP63+: Basal Cells
GABRP+: Secretory Cells
MUC5AC+: Goblet Cells
SCGB1A1+: Club Cells
H
ACE2
BPIFB1
CCL17
CD3G
CD68
EPCAM
FOXJ1
LAMP3
MKI67
MUC5AC
MYH11
PDGFRA
PTPRC
TMEM100
TP63
TPSB2
Alveolar type 2
B cells
Basal cells
Ciliated cells
Dendritic cells
Endothelial cells
Goblet/Club cells
Macrophages
Mast cells
Mesenchymal cells
Plasma cells
Proliferating cells
Smooth muscle cells
T cells

Figure 4

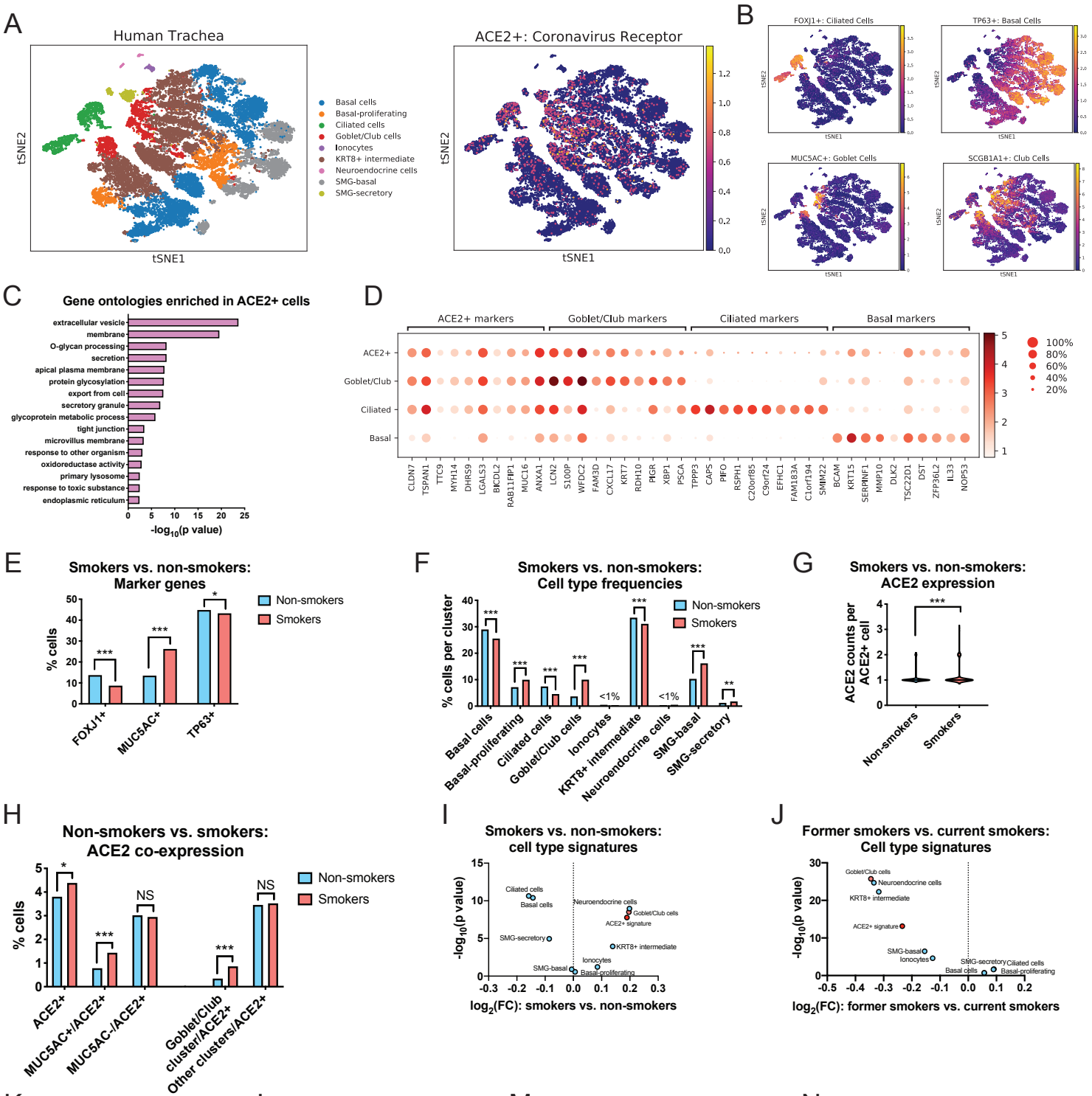
A
Human Trachea
tSNE2
tSNE1
Basal cells
Basal-proliferating
Ciliated cells
Goblet/Club cells
Ionocytes
KRT8+ intermediate
Neuroendocrine cells
SMG-basal
SMG-secretory
ACE2+: Coronavirus Receptor
B
FOXJ1+: Ciliated Cells
TP63+: Basal Cells
MUC5AC+: Goblet Cells
SCGB1A1+: Club Cells
C
Gene ontologies enriched in ACE2+ cells
extracellular vesicle
membrane
O-glycan processing
secretion
apical plasma membrane
protein glycosylation
export from cell
secretory granule
glycoprotein metabolic process
tight junction
microvillus membrane
response to other organism
oxidoreductase activity
primary lysosome
response to toxic substance
endoplasmic reticulum
-log10(p value)
D
ACE2+ markers
Goblet/Club markers
Ciliated markers
Basal markers
ACE2+
Goblet/Club
Ciliated
Basal
100%
80%
60%
40%
20%
CLDN7 TSPAN1 TTC9 MYH14 DHRS9 LGALS3 BICDL2 RAB11FIP1 MUC16 ANXA1 LCN2 S100P WFDC2 FAM3D CXCL17 KRT7 RDH10 PIGR XBP1 PSCA TPPP3 CAPS PIFO RSPH1 C20orf85 C9orf24 EFHC1 FAM183A C1orf194 SMIM22 BCAM KRT15 SERPINF1 MMP10 DLK2 TSC22D1 DST ZFP36L2 IL33 NOP53
E
Smokers vs. non-smokers:
Marker genes
% cells
Non-smokers
Smokers
FOXJ1+
MUC5AC+
TP63+
F
Smokers vs. non-smokers:
Cell type frequencies
% cells per cluster
Basal cells
Basal-proliferating
Ciliated cells
Goblet/Club cells
Ionocytes
KRT8+ intermediate
Neuroendocrine cells
SMG-basal
SMG-secretory
<1%
G
Smokers vs. non-smokers:
ACE2 expression
ACE2 counts per ACE2+ cell
Non-smokers
Smokers
H
Non-smokers vs. smokers:
ACE2 co-expression
% cells
NS
ACE2+
MUC5AC+/ACE2+
MUC5AC-/ACE2+
Goblet/Club cluster/ACE2+
Other clusters/ACE2+
I
Smokers vs. non-smokers:
cell type signatures
-log10(p value)
log2(FC): smokers vs. non-smokers
J
Former smokers vs. current smokers:
Cell type signatures
log2(FC): former smokers vs. current smokers
K
Mouse tracheal explants:
Differentiation vs. ACE2 expression
ACE2 expression (normalized)
Primary cells
Mucociliary differentiation
L
Human airway epithelial cells:
Differentiation vs. ACE2 expression
M
Human airway epithelial cells:
Smoke exposure during differentiation
Sham
+Smoke
N
Human airway epithelial cells:
Smoke exposure during differentiation
log2(FC): smoke vs. air

Figure 5

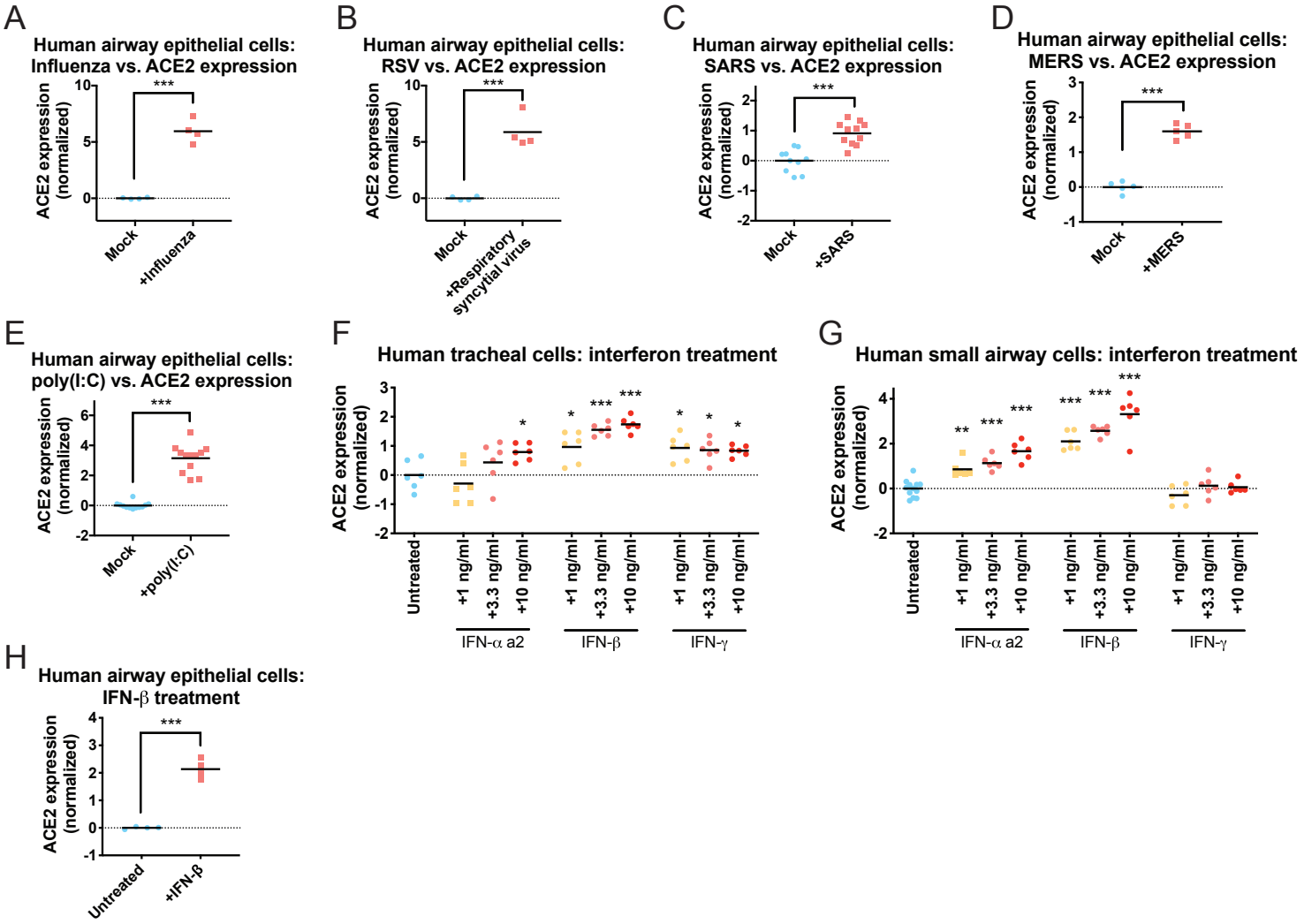
A
Human airway epithelial cells: Influenza vs. ACE2 expression
ACE2 expression (normalized)
Mock
+Influenza
B
Human airway epithelial cells: RSV vs. ACE2 expression
+Respiratory syncytial virus
C
Human airway epithelial cells: SARS vs. ACE2 expression
+SARS
D
Human airway epithelial cells: MERS vs. ACE2 expression
+MERS
E
Human airway epithelial cells: poly(I:C) vs. ACE2 expression
+poly(I:C)
F
Human tracheal cells: interferon treatment
G
Human small airway cells: interferon treatment
Untreated
+1 ng/ml
+3.3 ng/ml
+10 ng/ml
IFN-α a2
IFN-β
IFN-γ
H
Human airway epithelial cells: IFN-β treatment
+IFN-β

Figure S1

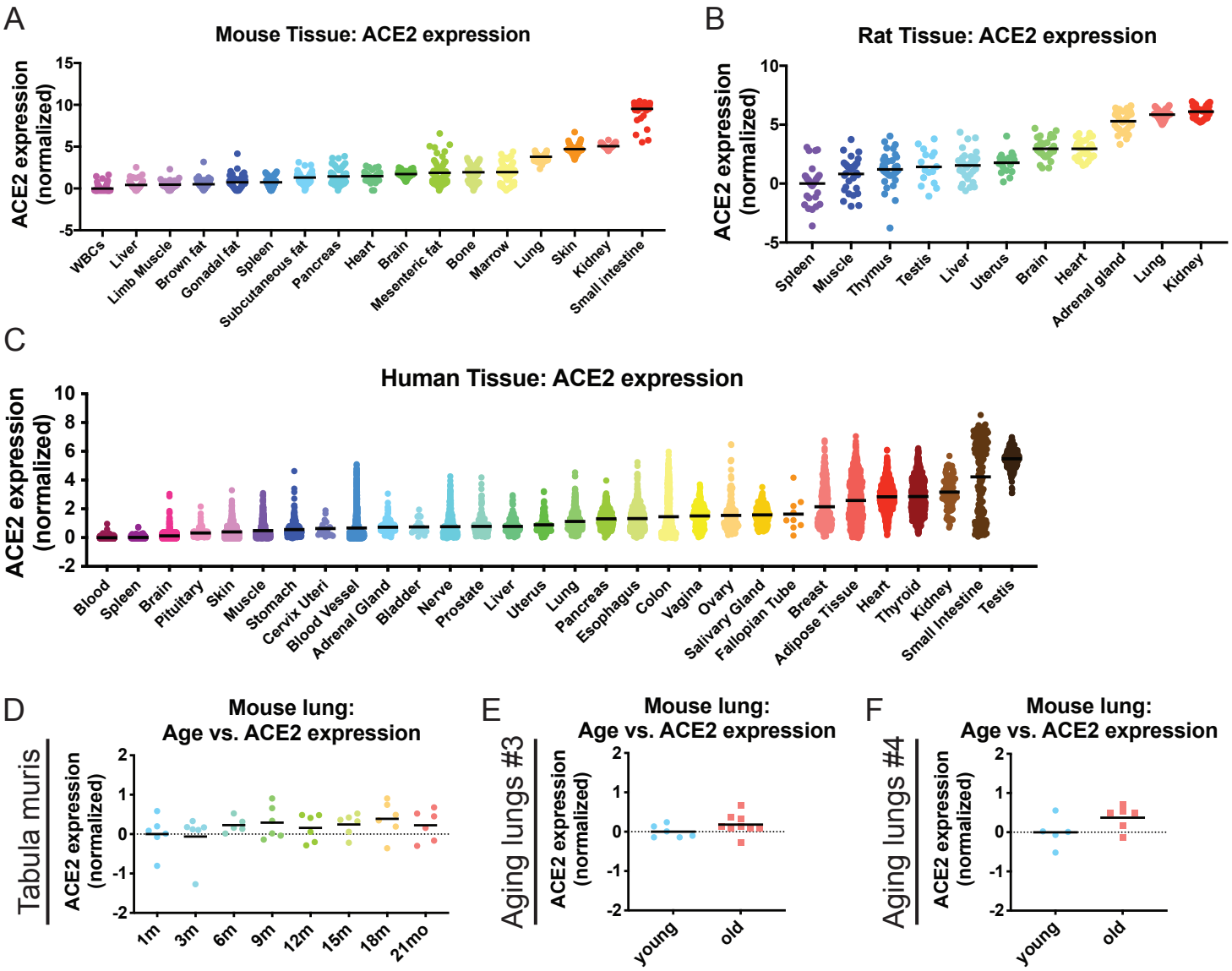

Figure S1. ACE2 expression varies between different mammalian tissues but is not strongly affected by aging. Related to Figure 1.

(A) The expression of ACE2 in 17 different mouse tissues or cell types, sorted according to mean ACE2 expression.

(B) The expression of ACE2 in 11 different rat tissues, sorted according to mean ACE2 expression.

(C) The expression of ACE2 in 30 different human tissues or cell types, sorted according to mean ACE2 expression.

(D) ACE2 expression in the lungs of a cohort of mice, sorted according to age.

(E) ACE2 expression in the lungs of an independent cohort of young mice (2-month old) and old mice (18-month old and 26-month old).

(F) ACE2 expression in the lungs of an independent cohort of young mice (7- to 9-week old) and old mice (14-month old).

Each panel displays log2-normalized ACE2 expression relative to a control group. Data analyzed in A and D were from GSE132040. Data analyzed in B were from GSE53960. Data analyzed in C were from www.gtexportal.org. Data analyzed in E were from GSE6591. Data analyzed in F were from GSE80680. Additional information on the data sources and sample sizes are included in Table S1.

Figure S2

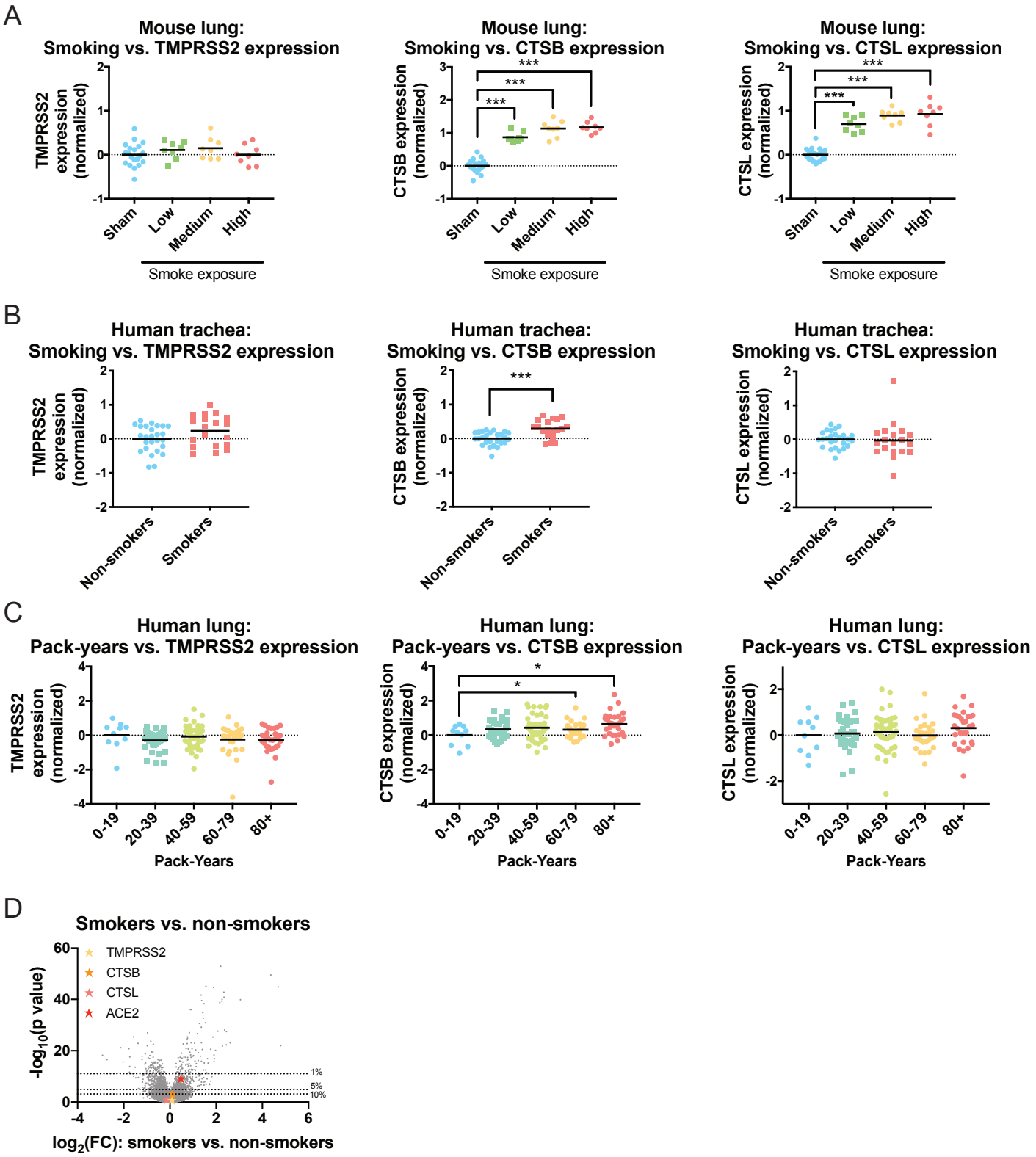
A
Mouse lung:
Smoking vs. TMPRSS2 expression
TMPRSS2 expression (normalized)
Sham
Low
Medium
High
Smoke exposure
Mouse lung:
Smoking vs. CTSB expression
CTSB expression (normalized)
***
Mouse lung:
Smoking vs. CTSL expression
CTSL expression (normalized)
B
Human trachea:
Smoking vs. TMPRSS2 expression
Non-smokers
Smokers
Human trachea:
Smoking vs. CTSB expression
Human trachea:
Smoking vs. CTSL expression
C
Human lung:
Pack-years vs. TMPRSS2 expression
0-19
20-39
40-59
60-79
80+
Pack-Years
Human lung:
Pack-years vs. CTSB expression
*
Human lung:
Pack-years vs. CTSL expression
D
Smokers vs. non-smokers
TMPRSS2
CTSB
CTSL
ACE2
1%
5%
10%
-log10(p value)
log2(FC): smokers vs. non-smokers

Figure S2. Cigarette smoke increases the expression of the coronavirus-associated protease Cathepsin B. Related to Figure 2.

(A) The expression of TMPRSS2, Cathepsin B, and Cathepsin L in the lungs of mice that were sham-treated or that were exposed to diluted cigarette smoke for two, three, or four hours a day (low, medium, and high smoke exposure, respectively).

(B) The expression of TMPRSS2, Cathepsin B, and Cathepsin L in human trachea collected by fiberoptic brushing and analyzed according to smoking history.

(C) The expression of TMPRSS2, Cathepsin B, and Cathepsin L in human lung tissue in a cohort of smokers and analyzed according to the number of pack-years smoked.

(D) A volcano plot comparing gene expression in the respiratory epithelia of current smokers and never-smokers from Figure 2C, 2D, and 2E. The dotted lines indicate various p-value thresholds (e.g., genes located above the 10th percentile have a combined p value greater than 90% of the genes included in the meta-analysis).

Each panel displays log2-normalized expression of the indicated gene relative to a control group. Data analyzed in A were from GSE18344. Data analyzed in B were from GSE13933. Data analyzed in C were from GSE76925. Additional information on the data sources and sample sizes are included in Table S1. *, $p < .05$; **, $p < .005$; ***, $p < .0005$ (Student's t-test).

Figure S3

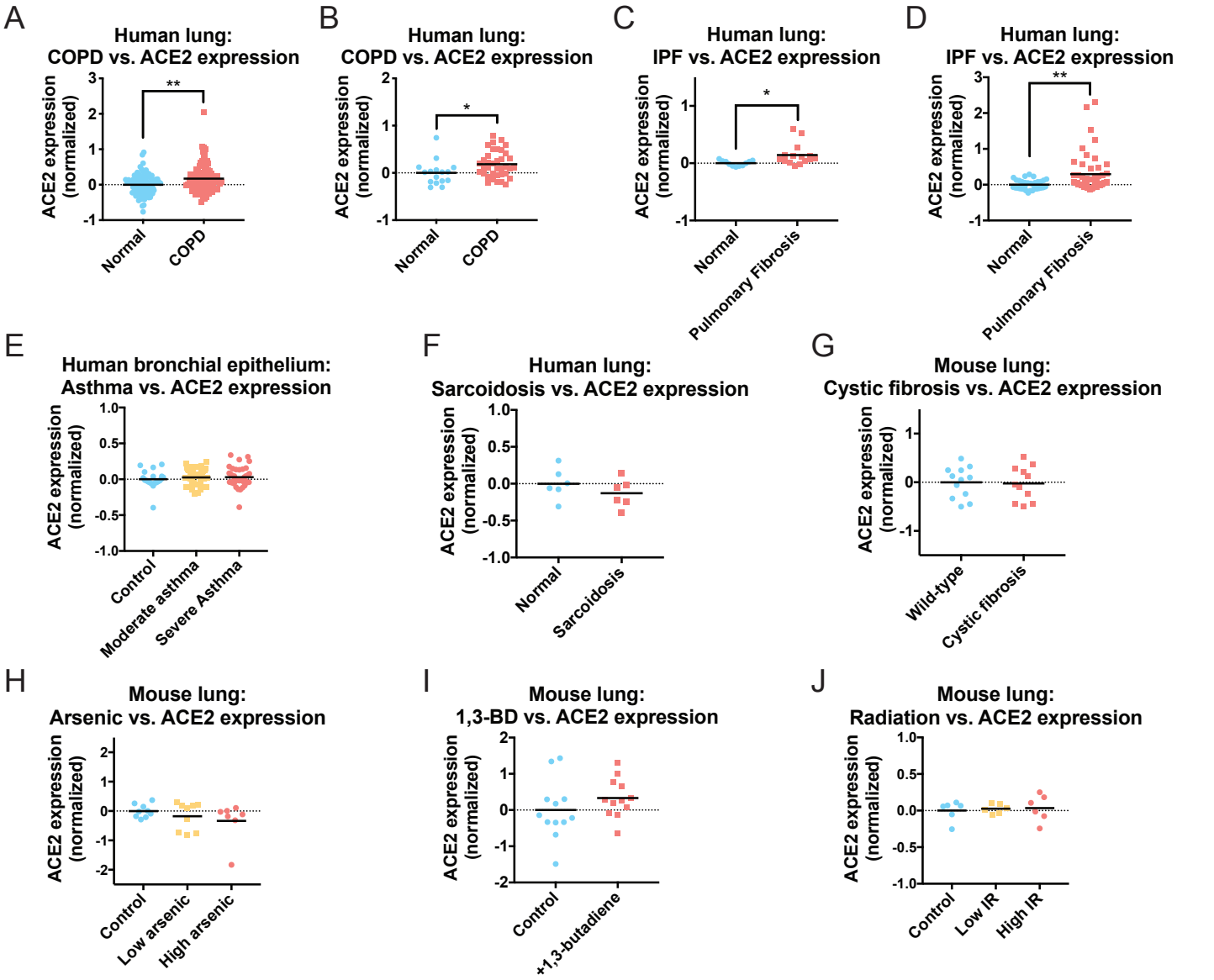

Figure S3. ACE2 is not commonly upregulated by lung diseases or carcinogen exposure. Related to Figure 4.

(A) ACE2 expression in normal lung tissue and in lung tissue from patients with COPD.

(B) ACE2 expression in normal lung tissue and in lung tissue from patients with COPD.

(C) ACE2 expression in lung tissue from patients with idiopathic pulmonary fibrosis or from histologically-normal lung tissue from patients with lung cancer.

(D) ACE2 expression in tissue from the lungs of transplant recipients with idiopathic pulmonary fibrosis or from control donor lungs.

(E) The expression of ACE2 in bronchial brushing from individuals with normal lung function, mild asthma, or severe asthma.

(F) The expression of ACE2 in disease-free lung tissue resected from patients with sarcoidosis or control tissue from healthy lung donors.

(G) The expression of ACE2 in the lungs of a cystic fibrosis mouse model (Cftr-/-) or wild-type littermate controls.

(H) The expression of ACE2 in the lungs of mice provided with normal water, water with 10 parts per billion sodium arsenite, or water with 100 parts per billion sodium arsenite (low and high arsenic, respectively).

(I) The expression of ACE2 in the lungs of mice exposed to normal air or exposed to the carcinogen 1,3-butadiene.

(J) The expression of ACE2 in the lungs of mice that were sham-treated, exposed to 5 Gy of ionizing radiation to the thorax, or exposed to 17.5 Gy of ionizing radiation to the thorax (low and high IR, respectively).

Each panel displays log2-normalized ACE2 expression relative to a control group. Data analyzed in A were from GSE57148. Data analyzed in B were from GSE103174. Data analyzed in C were from GSE2052. Data analyzed in D were from GSE47460. Data analyzed in E were from GSE43696. Data analyzed in F were from GSE16538. Data analyzed in G were from GSE3100. Data analyzed in H were from GSE11056. Data analyzed in I were from GSE86623. Data analyzed in J were from GSE41789. Additional information on the data sources and sample sizes are included in Table S1. *, $p < .05$; **, $p < .005$; ***, $p < .0005$ (Student's t-test).

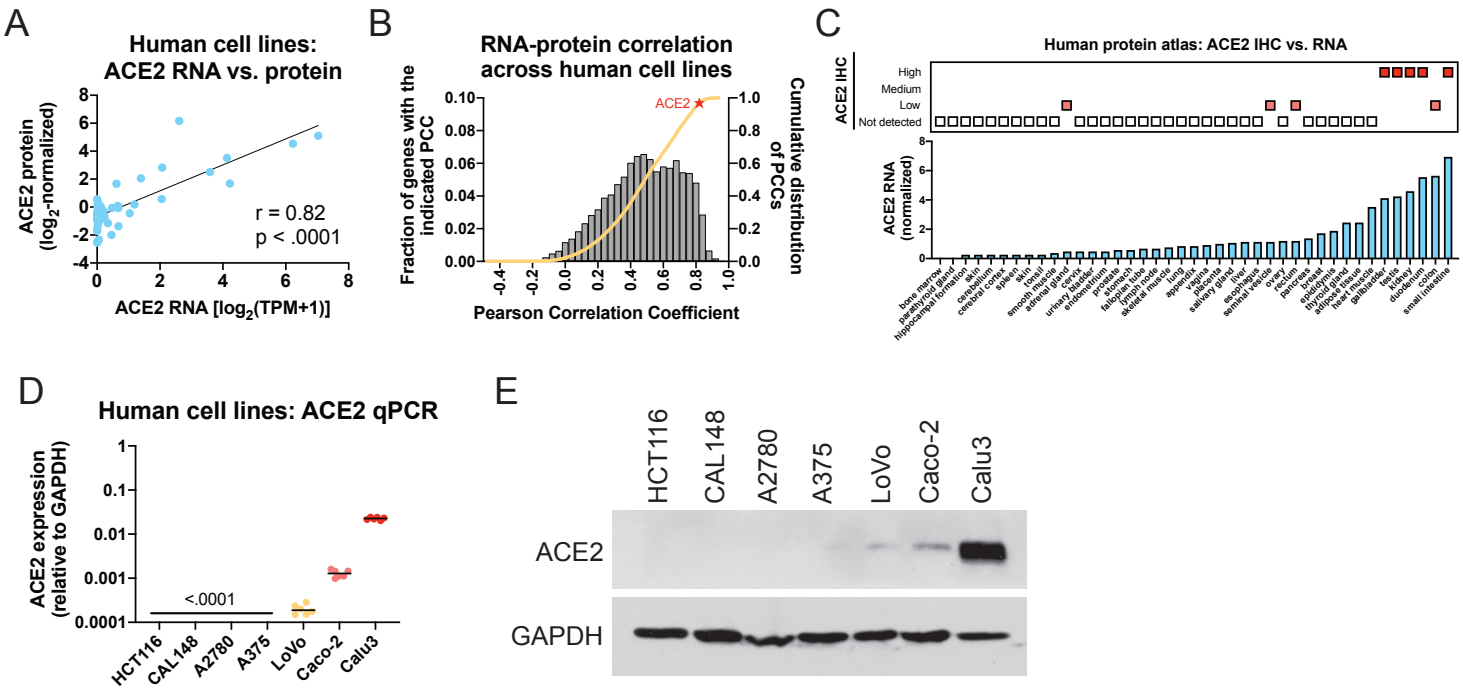

Figure S4. ACE2 RNA levels and protein levels are tightly correlated. Related to Figure 5.

(A) A scatter plot comparing the expression of ACE2 RNA, as measured by RNA-Seq, and ACE2 protein, as measured by mass spectrometry, across human cancer cell lines. The black line indicates a linear regression plotted against the data.

(B) A bar graph comparing the Pearson correlation coefficients between RNA and protein expression across 11,362 human genes. The grey bars denote the fraction of genes with an indicated correlation coefficient while the orange line indicates the cumulative distribution of those coefficients. ACE2's position is indicated with a red star.

(C) A graph comparing ACE2 immunohistochemistry staining in various tissues vs ACE2 RNA expression in those same tissues.

(D) qRT-PCR of ACE2 expression in the indicated cell lines, relative to GAPDH expression.

(E) Western blotting of ACE2 expression in the indicated cell lines. GAPDH was analyzed as a loading control.

The data in A and B are from reference 106. The data in C is from the Human Protein Atlas (reference 138). Additional information on these sources is included in Table S1.