The aging transcriptome and cellular landscape of the human lung in relation to SARS-CoV-2

Ryan D. Chow [1,2,3,4] and Sidi Chen [1,2,3,4,5,6,7,8,9,10,11,12,#]

Affiliations

1. Department of Genetics, Yale University School of Medicine, New Haven, Connecticut, USA
2. System Biology Institute, Yale University, West Haven, Connecticut, USA
3. Center for Cancer Systems Biology, Yale University, West Haven, Connecticut, USA
4. M.D.-Ph.D. Program, Yale University, West Haven, Connecticut, USA
5. Immunobiology Program, Yale University, New Haven, Connecticut, USA
6. Combined Program in the Biological and Biomedical Sciences, Yale University, New Haven, Connecticut, USA
7. Yale Comprehensive Cancer Center, Yale University School of Medicine, New Haven, Connecticut, USA
8. Department of Neurosurgery, Yale University School of Medicine, New Haven, Connecticut, USA
9. Yale Stem Cell Center, Yale University School of Medicine, New Haven, Connecticut, USA
10. Yale Liver Center, Yale University School of Medicine, New Haven, Connecticut, USA
11. Yale Center for Biomedical Data Science, Yale University School of Medicine, New Haven, Connecticut, USA
12. Yale Center for RNA Science and Medicine, Yale University School of Medicine, New Haven, Connecticut, USA

Correspondence:
SC (sidi.chen@yale.edu)
+1-203-737-3825 (office)
+1-203-737-4952 (lab)

Abstract

Since the emergence of SARS-CoV-2 in December 2019, Coronavirus Disease-2019 (COVID-19) has rapidly spread across the globe. Epidemiologic studies have demonstrated that age is one of the strongest risk factors influencing the morbidity and mortality of COVID-19. Here, we interrogate the transcriptional features and cellular landscapes of the aging human lung through integrative analysis of bulk and single-cell transcriptomics. By intersecting these age-associated changes with prior experimental data on host interactions between SARS-CoV-2 or its relative SARS-CoV, we identify several age-associated factors that may contribute to the heightened severity of COVID-19 in older populations. These analyses illuminate potential avenues for further studies on the relationship between the aging lung and COVID-19 pathogenesis, which may inform strategies to more effectively treat this disease.

Main Text

Age is one of the strongest risk factors for severe outcomes among patients with COVID-19 [1,2]. For instance, the case-fatality rate of patients 50-59 years old was reported to be 1.0% in Italy (as of March 17, 2020) and 1.3% in China (as of February 11, 2020) [3,4]; in contrast, the case-fatality rate of patients ≥ 80 years old was 20.2% in Italy, and 14.8% in China during that same time frame (**Figure 1a**). Similarly, in the United States, the Center for Disease Control (CDC) estimated that from February 12 to March 16, 2020, the case-fatality rate of patients 55-64 years old was 1.4 – 2%, and 10.4 -27.3% for patients ≥ 85 years old [5] (**Figure 1a**). Furthermore, a recent analysis of Chinese CDC data revealed that while patients of all ages shared a similar probability of becoming infected by SARS-CoV-2, the causative agent of COVID-19 [6–8], the clinical manifestations of infection in children (< 18 years old) were less severe than in adults. With the exception of infants and younger children (< 1 years old and 1-5 years old, respectively), most children were asymptomatic or experienced mild illness[9]. Collectively, these observations indicate a strong association between age and COVID-19 morbidity and mortality. However, it must be emphasized that younger patients can still frequently contract the disease, possibly leading to hospitalization, ICU admission, or death (**Supplementary Figure 1**).

While the effects of age on COVID-19 are likely to be multifactorial, involving a complex blend of systemic and local factors, we hypothesized that tissue-intrinsic changes that occur with aging may offer valuable clues. To explore possible mechanisms by which age influences the clinical manifestations of SARS-CoV-2 infection, here we investigate the transcriptomic features and cellular landscape of the aging human lung in relation to SARS-CoV-2. We emphasize that our findings are not intended to direct clinical practice, but rather to generate hypotheses and to further guide research efforts that can improve our understanding of this disease.

Results

We focused our analysis on the Genotype-Tissue Expression (GTEx) project[10,11], a comprehensive public resource of gene expression profiles from non-diseased tissue sites. As the lung is the primary organ affected by COVID-19, we specifically analyzed lung RNA-seq transcriptomes from donors of varying ages (20-79 years old) (**Figure 1b**). A total of 578 lung RNA-seq profiles were compiled, of which 31.66% were from women.

An initial hypothesis for why SARS-CoV-2 differentially affects patients of varying ages is that the expression of host factors essential for SARS-CoV-2 infection may increase with aging. To assess this possibility, we examined the gene expression of *ACE2*, which encodes the protein angiotensin-converting enzyme 2 that is coopted as the host receptor for SARS-CoV-2 [8,12–15]. There was no statistically significant association between age and *ACE2* expression levels (Kruskal-Wallis test, $p = 0.92$) (**Figure 1c**). While ACE2 is the direct cell surface receptor for SARS-CoV-2, transmembrane serine protease 2 (TMPRSS2) and cathepsin L (CTSL) have been demonstrated to facilitate SARS-CoV-2 infection by priming the spike protein for host cell entry[14]. Expression of the corresponding genes *TMPRSS2* and *CTSL* was modestly associated with age ($p = 0.0039$ and $p = 0.026$, respectively) (**Figure 1c**), but *TMPRSS2* and *CTSL* expression levels tended to decrease with age. Of note, biological sex was not significantly associated with expression of *ACE2*, *TMPRSS2*, or *CTSL* (**Supplementary Figure 2a-c**), though it has been observed that males are more likely to be affected by COVID-19 than females[16–18]. Taken together, these data indicate that differential expression of SARS-CoV-2 host entry factors alone is unlikely to explain the relationship between age and severity of COVID-19 illness.

To discern the host cell types involved in COVID-19 entry, we turned to a single cell RNA-seq (scRNA-seq) dataset of 57,020 human lung cells from the Tissue Stability Cell Atlas[19]. In agreement with prior reports, analysis of the single cell lung transcriptomes revealed that alveolar type 2 (AT2) cells were comparatively enriched in *ACE2* and *TMPRSS2*-expressing cells [20,21] (**Supplementary Figure 3a-b**). However, *ACE2*-expressing cells represented only 1.69% of all AT2 cells, while 47.52% of AT2 cells expressed *TMPRSS2*. Alveolar type 1 (AT1) cells also showed detectable expression of *ACE2* and *TMPRSS2*, but at lower frequencies (0.39% and 26.70%). *CTSL* expression could be broadly detected in many different cell types including AT2 cells, but its expression was particularly pronounced in macrophages (**Supplementary Figure 3c**).

Since the expression of host entry factors *ACE2*, *TMPRSS2* and *CTSL* did not increase with age, we next sought to identify all age-associated genes expressed in the human lung (Methods). Using a likelihood-ratio test[22], we pinpointed the genes for which age significantly impacts their expression. With a stringent cutoff of adjusted $p < 0.0001$, we identified two clusters of genes in

which their expression progressively changes with age (**Figure 1d**). Cluster 1 is composed of 643 genes that increase in expression with age, while Cluster 2 contains 642 genes that decrease in expression with age. These 1,285 age-associated genes modestly overlapped with two published aging-related gene sets[23,24], but 92.06% (1,183/1,285) of the age-associated genes identified in this study were not present in either of the published gene sets (**Supplementary Figure 4a-b**).

Gene ontology and pathway analysis of Cluster 1 genes (increasing with age) revealed significant enrichment for cell adhesion, vascular smooth muscle contraction, oxytocin signaling, and platelet activation, in addition to several other pathways (**Figure 1e**). These findings are consistent with known physiologic changes of aging, including decreased pulmonary compliance[25], and heightened risk for thrombotic diseases[26]. Of note, deregulation of the renin-angiotensin system has been implicated in the pathogenesis of acute lung injury induced by SARS-CoV [27,28], the closely related coronavirus responsible for the SARS epidemic of 2002-2003 [29]. In mice, SARS-CoV infection downregulates ACE2, leading to disinhibition of angiotensin II production by ACE [27,28,30–32] and subsequent vasoconstriction. A possible hypothesis is that increased baseline vascular smooth muscle contraction in older patients may predispose the development of acute respiratory distress syndrome (ARDS) in the setting of SARS-CoV and/or SARS-CoV-2 infection. In line with this hypothesis, recent analyses of COVID-19 cohorts in China have found that patients with hypertension were significantly more likely to develop ARDS[17] and require ICU admission[33], though we note that correlative epidemiologic studies do not necessarily demonstrate causality.

Cluster 2 genes (decreasing with age) were significantly enriched for mitochondrion, mitochondrial translation, metabolic pathways, and mitosis, among other pathways (**Figure 1f**), which is consistent with prior observations of progressive mitochondrial dysfunction with aging[34–36]. Of note, Cluster 2 was also enriched for genes involved in lipid metabolism, fatty acid metabolism, peroxisome, and lysosomal membranes. Age-associated alterations in lipid metabolism could impact SARS-CoV-2 infection, as SARS-CoV can enter cells through cholesterol-rich lipid rafts [37–40]. Similarly, age-associated alterations in lysosomes could influence late endocytic viral entry, as the protease cathepsin L cleaves SARS-CoV spike proteins from within lysosomes [41,42].

Having compiled a high-confidence set of age-associated genes, we sought to identify the lung cell types that normally express these genes, using the human lung single cell transcriptomics dataset from the Tissue Stability Cell Atlas[19]. By examining the scaled percentage of expressing cells within each cell subset, we identified age-associated genes predominantly enriched in different cell types. Cell types with highly enriched expression for certain Cluster 1 genes (increasing with age) included fibroblasts, muscle cells, and lymph vessels (**Figure 2a**). In contrast, cell types with highly enriched expression for certain Cluster 2 genes (decreasing with age) included macrophages, dividing dendritic cells (DCs)/monocytes, and AT2 cells (**Figure 2b**). Examining the muscle-enriched genes that increased in expression with age, gene ontology analysis revealed enrichment for vascular smooth muscle contraction, cGMP-PKG signaling, Z-disc, and actin cytoskeleton, among other pathways **(Figure 2c**). As for the AT2-enriched genes that decreased in expression with age, gene ontology analysis revealed enrichment for metabolic pathways, biosynthesis of antibiotics, lipid metabolism, extracellular exosome, and mitochondrial matrix (**Figure 2d**). A subset of these enriched gene ontologies had also been identified by the bulk RNA-seq analysis (**Figure 1e-f**). Thus, integrative analysis of bulk and single-cell transcriptomes revealed that many of the age-associated transcriptional changes in human lung can be mapped to specific cell subpopulations, suggesting that either the transcriptional status, or overall abundance of these cell types, or both, may be altered with aging.

As the pathophysiology of viral-induced ARDS involves an intricate interplay of diverse cell types, most notably the immune system[43,44], aging-associated shifts in the lung cellular milieu[25] could contribute an important dimension to the relationship between age and risk of ARDS in patients with COVID-19 [33]. To investigate the cellular landscape of the aging lung, we applied a gene signature-based approach[45] to infer the enrichment of different cell types from the bulk RNA-seq profiles. Since bulk RNA-seq measures the average expression of genes within a cell population, such datasets will reflect the relative proportions of the cell types that comprised the input population, though with the caveat that cell types can have overlapping expression profiles and such profiles may be altered in response to stimuli. Using this approach, we identified age-associated alterations in the enrichment scores of several cell types (**Figure 3a**). Whereas epithelial cells decreased with age, fibroblasts increased with age (**Figure 3b**). This finding is consistent with the progressive loss of lung parenchyma due to reduced regenerative capacity of the aging

lung[46], as well as the increased risk for diseases such as chronic obstructive pulmonary disease and pulmonary fibrosis[47]. In addition, these results are concordant with the findings from analysis of human lung single-cell transcriptomes (**Figure 2a-b**).

Among the innate immune cell populations, the enrichment scores of total macrophages were inversely associated with age (**Figure 3c**). Macrophages are major drivers of innate immune responses in the lung, acting as first-responders against diverse respiratory infections[48]. Thus, the age-associated decrease in macrophage abundance may be a possible factor related to the greater severity of lung pathology in patients with COVID-19. Although macrophage accumulation is often associated with the pathologic inflammation of viral ARDS [49,50], pulmonary macrophages can act to limit the duration and severity of infection by efficiently phagocytosing dead infected cells and released virions [51–53]. Notably, macrophages infected with SARS-CoV have been found to abort the replication cycle of the virus[54,55], further supporting their role in antiviral responses. However, macrophages may suppress antiviral adaptive immune responses [48,56], inhibiting viral clearance in mouse models of SARS-CoV infection [57]. In aggregate, these prior reports suggest that the precise role of lung macrophages in SARS-CoV-2 pathophysiology is likely context-dependent. It is also plausible that the increased numbers of macrophages are not the primary distinction between young and old patients, but rather the functional status of the macrophages. In line with this, we observed that the age-associated changes in macrophages were specifically attributed to the pro-inflammatory M1 macrophage subset but not the immunoregulatory M2 subset [58] (**Figure 3c**), though this binary classification scheme represents an oversimplification of macrophage function. Nevertheless, elucidating the consequences of age-associated changes in lung macrophages may reveal insights into the differential outcomes of older patients with COVID-19. Further studies are needed to investigate whether macrophages or other innate immune cells respond to SARS-CoV-2 infection, and how their numbers or function may change with aging.

Among the adaptive immune cell populations, we observed that Th1 cells and $CD4^+$ Tcm cells trended in opposite directions with aging (**Figure 3d**). While the lungs of younger donors were enriched for Th1 cells, they were comparatively depleted for $CD4^+$ Tcm cells; the inverse was true in the lungs of older donors. Of note, mouse models of SARS-CoV infection have indicated

important roles for CD4+ T cell responses in viral clearance [59,60]. Additionally, Th1 cells are responsive to SARS-CoV vaccines [61] and promote macrophage activation against viruses [62]. It is therefore possible that age-associated shifts in CD4+ T cell subtypes within the lung may influence the subsequent host immune response in response to coronavirus infection. However, future studies will be needed to determine the role of Th1 cells and other adaptive immune cells in the response to SARS-CoV-2, and how these dynamics may change with aging.

We next explored the roles of lung age-associated genes in host responses to viral infection. Since functional screening data with SARS-CoV-2 has not yet been described (as of March 30, 2020), we instead searched for data on SARS-CoV. While these two viruses belong to the same genus (*Betacoronaviridae*) and are conserved to some extent [8], they are nevertheless two distinct viruses with different epidemiological features, indicating unique virology and host biology. Therefore, data from experiments performed with SARS-CoV must be interpreted with caution.

We reassessed the results from a prior *in vitro* siRNA screen of host factors involved in SARS-CoV infection [63]. In this kinase-focused screen, 130 factors were determined to have a significant effect on SARS-CoV replication. Notably, 11 of the 130 factors exhibited age-associated gene expression patterns (**Figure 4a**), with 4 genes in Cluster 1 (increasing with age) and 7 genes in Cluster 2 (decreasing with age). The 4 genes in Cluster 1 were all associated with increased SARS infectivity upon siRNA knockdown; these genes included *CLK1, AKAP6, ALPK2,* and *ITK*. Paradoxically, while knockdown of *CLK1* was associated with increased SARS infectivity, cell viability was also found to be increased (**Figure 4b**). Of the 7 genes in Cluster 2, 6 were associated with increased SARS infectivity and reduced cell viability upon siRNA knockdown (*AURKB, CDKL2, PDIK1L, CDKN3, MST1R,* and *ADK*). Age-related downregulation of these 6 factors could be related to the increased severity of illness in older patients. However, we emphasize that until rigorous follow-up experiments are performed with SARS-CoV-2, the therapeutic potential of targeting these factors in patients with COVID-19 is unknown.

Using the human lung scRNA-seq data, we then determined which cell types predominantly express these host factors. Of the 4 genes in Cluster 1 that had a significant impact on SARS-CoV replication, *CLK1* was universally expressed, while ALPK2 expression was rarely detected

(**Figure 4c**). *ITK* was preferentially expressed in lymphocytic populations, and *AKAP6* was most frequently expressed in ciliated cells and muscle cells. Of the 7 overlapping genes in Cluster 2, *AURKB* and *CDKN3* were predominantly expressed in proliferating immune cell populations, such as macrophages, DCS/monocytes, T cells, and NK cells (**Figure 4d**). *MST1R*, *PDIK1L*, and *PSKH1* were infrequently expressed, through their expression was detected in a portion of AT2 cells (5.54%, 5.30%, and 5.47%, respectively). Finally, *ADK* and *CDKL2* exhibited preferential enrichment in AT2 cells (51.40% and 34.43%). In aggregate, these analyses showed that the age-associated genes with functional roles in SARS-CoV are expressed in specific cell types of the human lung.

We then investigated whether age-associated genes in the human lung interact with proteins encoded by SARS-CoV-2. A recent study interrogated the human host factors that interact with 27 different SARS-CoV-2 proteins [64], revealing the SARS-CoV-2 : Human protein interactome in cell lines expressing recombinant SARS-CoV-2 proteins. By cross-referencing the interacting host factors with the set of age-associated genes, we identified 20 factors at the intersection (**Figure 5a**). 4 of these genes showed an increase in expression with age (i.e. Cluster 1 genes), while 16 decreased in expression with age (Cluster 2 genes). Mapping these factors to their interacting SARS-CoV-2 proteins, we noted that the age-associated host factors which interact with M, Nsp13, Nsp1, Nsp7, Nsp8, Orf3a, Orf8, Orf9c, and Orf10 proteins generally decrease in expression with aging (**Figure 5b**). However, a notable exception was Nsp12, as the age-associated host-factors that interact with Nsp12 both showed increased expression with aging (*CRTC3* and *MYCBP2*) (**Figure 5c**). Nsp12 encodes for the primary RNA-dependent RNA polymerase (RdRp) of SARS-CoV-2, and is a prime target for developing therapies against COVID-19. The observation that *CRTC3* and *MYCBP2* increase in expression with aging is intriguing, as these genes may be related to the activity of Nsp12/RdRp in host cells. Of note, MYCBP2 is a known repressor of cAMP signaling [65,66], and cAMP signaling potently inhibits contraction of airway smooth muscle cells [67]. Thus, age-associated increases in *MYCBP2* could promote smooth muscle contraction, which is concordant with our analyses on age-associated gene signatures in the lung (**Figure 1e**). *MYCBP2* might therefore possibly contribute to COVID-19 pathology by not only interacting with SARS-CoV-2 RdRp, but also through its normal physiologic role in promoting smooth muscle contraction.

To assess the cell type-specific expression patterns of these various factors, we further analyzed the lung scRNA-seq data. Of the SARS-CoV-2 interacting genes that increase in expression with age, *MYCBP2* was frequently expressed across several populations, particularly proliferating immune populations (DC/monocyte, T cells, macrophages), muscle cells, fibroblasts, and lymph vessels (**Figure 5d**). *MYCBP2* was also expressed in 21.86% of AT2 cells. *CEP68* was preferentially expressed in lymph vessels, while *AKAP8L* and *CRTC3* showed relatively uniform expression frequencies across cell types, including a fraction of AT2 cells (8.53% and 9.07% expressing cells, respectively). Of the SARS-CoV-2 interacting genes that decrease in expression age, *NPC2* and *NDUFB9* were broadly expressed in many cell types, including AT2 cells (99.96% and 81.91%, respectively) (**Figure 5e**). AT2 cells also frequently expressed *ATP1B1, ALG5, NEU1,* and *ATP6V1A* (70.21%, 50.64%, 43.24%, and 27.78%). Taken together, these analyses highlight specific age-associated factors that interact with the SARS-CoV-2 proteome, in the context of the lung cell types in which these factors are normally expressed.

Discussion

Here we systematically analyzed the transcriptome of the aging human lung and its potential relationship to SARS-CoV-2. We found that the aging lung is characterized by a wide array of changes that could contribute to the worse outcomes of older patients with COVID-19. On the transcriptional level, we first identified 1,285 genes that exhibit age-associated expression patterns. We subsequently demonstrated that the aging lung is characterized by several gene signatures, including increased vascular smooth muscle contraction, reduced mitochondrial activity, and decreased lipid metabolism. By integrating these data with single cell transcriptomes of human lung tissue, we further pinpointed the specific cell types that normally express the age-associated genes. We showed that epithelial cells, macrophages, and Th1 cells decrease in abundance with age, whereas fibroblasts and pericytes increase in abundance with age. These systematic changes in tissue composition and cell interactions can potentially propagate positive feedback loops that predispose the airways to pathological contraction [68]. Notably, some of the age-associated genes have been previously identified as host factors with a functional role in SARS-CoV replication [63], and a fraction of the age-associated factors have been shown to directly interact with the SARS-

CoV-2 proteome[64]. Whether any of these age-associated changes causally contribute to the heightened susceptibility of COVID-19 in older populations remains to be experimentally tested.

It is important to note that the prior datasets analyzed here were not from patients with COVID-19. Although transcriptomic data from patients will be beneficial for better understanding COVID-19, the health of the patients should continue to be the highest priority. Given the limited data that is currently publicly available, we emphasize that the analyses presented here are exploratory in nature, and as such should not be used to guide clinical practice. Instead, we hope that these analyses will illuminate new directions for subsequent research efforts on SARS-CoV-2 by generating hypotheses for why advanced age is one of the strongest risk factors for COVID-19 morbidity and mortality. Ultimately, we hope this knowledge may help the field to sooner develop rational therapies for COVID-19 that are rooted in concrete biological mechanisms.

Acknowledgments

We thank Akiko Iwasaki, Craig Wilen, Hongyu Zhao, Wei Liu, Wenxuan Deng, Andre Levchenko, Katie Zhu, Ruth Montgomery, Bram Gerriten, Steve Kleinstein, and a number of other colleagues for their critical comments and suggestions, which were incorporated into the analyses and manuscript. We thank Antonio Giraldez, Andre Levchenko, Chris Incarvito, Mike Crair, and Scott Strobel for their support on COVID-19 research. We thank our colleagues in the Chen lab, the Genetics Department, the Systems Biology Institute and various Yale entities. Personally, we also want to thank all of the healthcare workers who are risking their health on the frontlines to treat patients with this disease.

Author contributions

RC and SC conceived and designed the study. RC developed the analysis approach, performed all data analyses, and created the figures. RC and SC prepared the manuscript. SC supervised the work.

Declaration of interests

No competing interests related to this study.

The authors have no competing interests relate to this work, or other interests that might be perceived to influence the interpretation of the article. The authors are committed to freely share all COVID-19 related data, knowledge and resources to the community to facilitate the development of new treatment or prevention approaches against SARS-CoV-2 / COVID-19 as soon as possible.

As a note for full disclosure, SC is a co-founder, funding recipient and scientific advisor of EvolveImmune Therapeutics, which is not related to this study.

Figure Legends

Figure 1: Identification of age-associated genes in the human lung

a. Case-fatality statistics in China, Italy, and the United States. Data are from the Chinese CDC (as of February 11, 2020), the Italian ISS (as of March 17, 2020), and the US CDC (as of March 16, 2020), respectively. Error bars for case-fatality in the United States represent lower and upper bounds of the estimated rates, given the preliminary nature of the data.

b. Demographics of the human lung RNA-seq profiles in the GTEx dataset, detailed by sex and age group (total n = 578).

c. Tukey boxplots detailing the expression of SARS-CoV-2 entry factors *ACE2*, *TMPRSS2*, and *CTSL* in the lung. Data are shown as $\log_2$ transcripts per million (TPM). Statistical significance of the expression variation across all age groups is shown on top of each plot (Kruskal-Wallis test). Pairwise comparisons of each age group to the 20-29 year-old group are also shown, color-coded by the age group (two-tailed Mann-Whitney test).

d. Dot plot of age-associated genes in human lung. Cluster 1 contains genes that increase in expression with aging (left, n = 643), while cluster 2 contains genes that decrease in expression with aging (right, n = 642). Statistical significance was assessed by DESeq2 likelihood ratio test (adj. $p < 0.0001$). Data are shown in terms of median z-score, with lines connecting the same gene across different age groups.

e. DAVID gene ontology and pathway analysis of Cluster 1 genes (increasing with age).

f. DAVID gene ontology and pathway analysis of Cluster 2 genes (decreasing with age).

Figure 2: Lung single-cell transcriptomics pinpoints cell type-specific expression of age-associated genes

a. Heatmap showing the percentage of cells expressing each of the Cluster 1 genes (increasing with age), scaled by gene across the different cell types.

b. Heatmap showing the percentage of cells expressing each of the Cluster 2 genes (decreasing with age), scaled by gene across the different cell types.

c. DAVID gene ontology and pathway analysis of Cluster 1 age-associated genes that exhibit enriched expression in muscle cells.

d. DAVID gene ontology and pathway analysis of Cluster 2 age-associated genes that exhibit enriched expression in alveolar type 2 (AT2) cells.

Figure 3: The evolving cellular landscape of the aging human lung.

a. Heatmap of cell type enrichment scores in the human lung across different age groups. Data are expressed as median z-scores within each subpopulation.

b. Tukey boxplots detailing the enrichment scores of epithelial cells (top) and fibroblasts (bottom) in the lung across age groups. Statistical significance of age-associated variation was assessed by Kruskal-Wallis test.

c. Tukey boxplots detailing the enrichment scores of macrophages, M1 macrophages, and M2 macrophages in the lung across age groups. Statistical significance of age-associated variation was assessed by Kruskal-Wallis test.

d. Tukey boxplots detailing the enrichment scores of Th1 cells, naive $CD4^+$ T cells, and $CD4^+$ Tcm cells in the lung across age groups. Statistical significance of age-associated variation was assessed by Kruskal-Wallis test.

Figure 4: Age-associated genes in human lung that influence SARS-CoV replication

a. Scatter plot of data from a published siRNA screen of host factors that influence SARS-CoV pathogenesis (from de Wilde et al., 2015). Data are shown in terms of relative SARS-CoV infectivity and relative cell viability upon perturbation of various host factors. Values >1.0 for relative SARS-CoV infectivity suggest that knockdown of the target gene promotes infection. Values >1.0 for relative cell viability suggest that knockdown of the target gene promotes cell survival. The size of each point is scaled by the statistical strength of the association between expression of the gene and aging. Genes identified in Cluster 1 (increasing with age) and Cluster 2 (decreasing with age) are additionally color-coded (blue and orange, respectively).

b. Tukey boxplots detailing the expression of select age-associated genes with potential roles in SARS-CoV pathogenesis, highlighted in **(d)**. Data are shown as $\log_2$ transcripts per million (TPM) for *CLK1, CDKL2, PDIK1L, MST1R, ADK,* and *AURKB.* Statistical significance of the expression variation across all age groups was assessed by Kruskal-Wallis test**.**

c. Heatmap showing the percentage of cells expressing each of the Cluster 1 genes (increasing with age) with an effect on SARS-CoV, as highlighted in **(a)**.

d. Heatmap showing the percentage of cells expressing each of the Cluster 2 genes (decreasing with age) with an effect on SARS-CoV, as highlighted in **(a)**.

Figure 5: Age-associated genes in human lung and their relation to SARS-CoV-2 proteins.
a. Venn diagram of the intersection between age-associated genes in human lung and the SARS-CoV-2 : Human protein interactome (Gordon et al., 2020). Of the 20 age-associated genes that were found to also interact with SARS-CoV-2, 4 of them increased in expression with age, while 16 decreased with age.
b. Age-associated genes in human lung and their interaction with SARS-CoV-2 proteins, where each block contains a SARS-CoV-2 protein (underlined) and its interacting age-associated factors. Blocks are colored by the dominant directionality of the age association (orange, decreasing with age; blue, increasing with age). Gene targets with already approved drugs, investigational new drugs, or preclinical molecules are additionally denoted with an asterisk.
c. Tukey boxplots detailing the expression of select age-associated genes that interact with SARS-CoV-2 proteins, highlighted in **(b)**. Data are shown as $\log_2$ transcripts per million (TPM) for *ATP1B1, HOOK1, PRIM1, AGPS, NARS2, CRTC3, ALG5, NPC2, NDUFB9,* and *PPT1*. The SARS-CoV-2 interacting protein is also annotated in parentheses. Statistical significance of the expression variation across all age groups was assessed by Kruskal-Wallis test.
d. Heatmap showing the percentage of cells expressing each of the Cluster 1 genes (increasing with age) that interact with SARS-CoV-2 proteins, as highlighted in **(b)**.
e. Heatmap showing the percentage of cells expressing each of the Cluster 2 genes (decreasing with age) that interact with SARS-CoV-2 proteins, as highlighted in **(b)**.

Supplementary Figure 1: Severe outcomes of COVID-19 in the United States
Rates of hospitalization, ICU admission, and case-fatality from COVID-19 in the United States, stratified by age group. Data from the US CDC (as of March 16, 2020). Error bars represent lower and upper bounds of the estimated rates, given the preliminary nature of the data.

Supplementary Figure 2: Sex-specific analysis of SARS-CoV-2 host entry factor expression
a. Expression of *ACE2* in lungs from female or male donors, stratified by age. Statistical significance was assessed by two-tailed Mann-Whitney test.
b. Expression of *TMPRSS2* in lungs from female or male donors, stratified by age. Statistical significance was assessed by two-tailed Mann-Whitney test.

c. Expression of *CTSL* in lungs from female or male donors, stratified by age. Statistical significance was assessed by two-tailed Mann-Whitney test.

Supplementary Figure 3: Expression of SARS-CoV-2 host factors across lung cell types

a. Left, normalized expression of *ACE2* across different lung cell types. Right, bar plot showing the percentage of *ACE2*-expressing cells, grouped by cell type.

b. Left, normalized expression of *TMPRSS2* across different lung cell types. Right, bar plot showing the percentage of *TMPRSS2*-expressing cells, grouped by cell type.

c. Left, normalized expression of *CTSL* across different lung cell types. Right, bar plot showing the percentage of *CTSL*-expressing cells, grouped by cell type.

Supplementary Figure 4: Comparison of the lung age-associated gene set with other published gene sets

a. Venn diagram of the intersection between age-associated genes in human lung (this study) and the GenAge database [23] of age-associated genes. Statistical significance of the overlap was assessed by two-tailed hypergeometric test.

b. Venn diagram of the intersection between age-associated genes in human lung (this study) and a previously published age-associated gene set [24]. Statistical significance of the overlap was assessed by two-tailed hypergeometric test.

Methods

Data accession

The Genotype-Tissue Expression (GTEx) Project was supported by the Common Fund of the Office of the Director of the National Institutes of Health, and by NCI, NHGRI, NHLBI, NIDA, NIMH, and NINDS [10,11]. RNA-seq raw counts and normalized TPM matrices were downloaded from the GTEx Portal (https://gtexportal.org/home/index.html) on March 18, 2020, release v8. All accessed data used in this study are publicly available on the web portal and have been de-identified, except for patient age range and gender.

Case-fatality rates in China and Italy were from the Chinese CDC and Italian ISS, respectively, as described in a recent publication[3,4]. Statistics on COVID-19 cases in the United States from February 12 – March 16, 2020 were compiled from the CDC Morbidity and Mortality Weekly Report dated March 18, 2020 [5]. Given the preliminary nature of the data, only estimated lower and upper bounds were reported.

Single cell transcriptomes of human lungs were obtained from the Tissue Stability Cell Atlas (https://www.tissuestabilitycellatlas.org/) [19].

RNA-seq gene expression visualization and statistical analysis

For visualization of RNA-seq expression data, the TPM values were $\log_2$ transformed and plotted in R (v3.6.1). All boxplots are Tukey boxplots, with interquartile range (IQR) boxes and 1.5 × IQR whiskers. Pairwise statistical comparisons in the plots were assessed by two-tailed Mann-Whitney test, while statistical comparisons across all age groups were performed by Kruskal-Wallis test.

Identification of age-associated genes in human lung

To identify age-associated genes, the raw counts values were analyzed by DESeq2 (v1.24.0) [22], using the likelihood ratio test (LRT). Age-associated genes were determined at a significance threshold of adjusted $p < 0.0001$. Genes passing the significance threshold were then scaled to z-scores and clustered using the degPatterns function from the R package DEGreport (v1.20.0). Gene clusters with progressive and consistent trends with age were retained for downstream analysis.

Comparison of age-associated gene sets

The lung age-associated gene set identified in this study was compared with other published age-associated gene sets. Specifically, we accessed the GenAge dataset [23] (build 20) on March 24, 2020 (https://genomics.senescence.info/genes/human.html), as well as another previously published gene list in human lung [24], which had analyzed an earlier release of the GTEx dataset (v3) using a different statistical methodology. Statistical significance of overlaps between the gene sets was assessed by hypergeometric test, assuming 20,449 total genes in the human genome as annotated in the GRCh38.p13 Ensembl assembly.

Gene ontology and pathway analysis of lung age-associated genes

Gene ontology and pathway enrichment analysis was performed using DAVID (v6.8) [69] (https://david.ncifcrf.gov/), separating the age-associated genes into the two clusters (increasing or decreasing with age), as described above.

scRNA-seq data analysis

scRNA-seq data were analyzed in R (v3.6.1) using Seurat[70,71] and custom scripts. Of the 1,285 age-associated genes identified from GTEx bulk transcriptomes, 1,049 genes could be matched in the scRNA-seq dataset. To determine the percentage of cells expressing a given gene, the expression matrices were converted to binary matrices by setting a threshold of expression > 0. Cell type-specific expression frequencies for each gene were then calculated using the provided cell type annotations. To identify genes preferentially expressed in a specific cell type, we further scaled the expression frequencies in R to obtain z-scores. Data were visualized in R using the NMF package [72]. Where applicable, gene ontology analysis was performed with DAVID (v6.8) [69], using genes with z-score > 2 in the cell type of interest for analysis.

Inferring the cellular compositions of the human lung

To infer the cellular composition of each lung sample, we analyzed the TPM expression matrices using the xCell algorithm[45]. The resultant cell type enrichment tables were analyzed in R. For data visualization, cell type enrichment scores were scaled to z-scores, and the median z-score for each age group was expressed as a heatmap, using the superheat package [73]. Age-association was assessed across all age groups by Kruskal-Wallis test.

Assessing functional roles of lung age-associated genes in SARS-CoV

To assess whether any age-associated genes affect host responses to SARS-CoV (a coronavirus related to SARS-CoV-2), we analyzed the data from a published siRNA screen of host factors influencing SARS-CoV [63] (Data Set S1 in the publication; accessed on March 20, 2020). For data visualization, each point corresponding to a target gene was size-scaled and color-coded according to the age-association statistical analyses described above.

Age-associated genes that interact with the SARS-CoV-2 proteome

To assess whether any lung age-associated genes encode proteins that interact with the SARS-CoV-2 proteome, we compiled the data from a preprint manuscript detailing the human host factors that interact with 27 different proteins in the SARS-CoV-2 proteome [64] (accessed on March 23, 2020).

Statistical information summary

Comprehensive information on the statistical analyses used are included in various places, including the figures, figure legends and results, where the methods, significance, p-values and/or tails are described. All error bars have been defined in the figure legends or methods.

Code availability

Codes used for data analysis or generation of the figures related to this study are available upon request to the corresponding author and will be deposited to GitHub upon publication for free public access.

Data and resource availability

All relevant processed data generated during this study are included in this article and its supplementary information files. Raw data are from various sources as described above. All data and resources related to this study are freely available upon request to the corresponding author.

Supplementary Tables

Table S1: Demographics of donors for GTEx lung samples.

Table S2: Normalized expression matrix for SARS-CoV-2 host entry factors in human lung.

Table S3: Cell type-specific expression frequencies for SARS-CoV-2 host entry factors in human lung.

Table S4: Likelihood-ratio test table for evaluating age-associated genes.

Table S5: Final set of age-associated genes.

Table S6: DAVID analysis table for genes that increase in expression with age.

Table S7: DAVID analysis table for genes that decrease in expression with age.

Table S8: Cell type-specific expression frequencies for genes that increase in expression with age.

Table S9: Cell type-specific expression frequencies for genes that increase in expression with age, transformed to z-scores.

Table S10: Cell type-specific expression frequencies for genes that decrease in expression with age.

Table S11: Cell type-specific expression frequencies for genes that decrease in expression with age, transformed to z-scores.

Table S12: DAVID analysis table for muscle-enriched genes that increase in expression with age.

Table S13: DAVID analysis table for AT2-enriched genes that decrease in expression with age.

Table S14: Table of scaled cell enrichment scores, median value of each age group.

Table S15: Age-association statistics for cell enrichment scores.

Table S16: Annotation of a prior SARS-CoV siRNA screen (de Wilde et al., 2015) with age-association statistics from this study.

Table S17: Annotation of the SARS-CoV-2 : Human protein interactome (Gordon et al., 2020) with age-association statistics from this study.

References

1. Wu, J. T. *et al.* Estimating clinical severity of COVID-19 from the transmission dynamics in Wuhan, China. *Nat Med* 1–5 (2020) doi:10.1038/s41591-020-0822-7.
2. Riou, J., Hauser, A., Counotte, M. J. & Althaus, C. L. Adjusted age-specific case fatality ratio during the COVID-19 epidemic in Hubei, China, January and February 2020. *medRxiv* 2020.03.04.20031104 (2020) doi:10.1101/2020.03.04.20031104.
3. Onder, G., Rezza, G. & Brusaferro, S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA* (2020) doi:10.1001/jama.2020.4683.
4. Livingston, E. & Bucher, K. Coronavirus Disease 2019 (COVID-19) in Italy. *JAMA* (2020) doi:10.1001/jama.2020.4344.
5. CDCMMWR. Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) — United States, February 12–March 16, 2020. *MMWR Morb Mortal Wkly Rep* **69**, (2020).
6. Zhu, N. *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N. Engl. J. Med.* **382**, 727–733 (2020).
7. Lu, R. *et al.* Genomic characterisation and epidemiology of 2019 novel coronavirus: implications for virus origins and receptor binding. *Lancet* **395**, 565–574 (2020).
8. Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273 (2020).
9. Dong, Y. *et al.* Epidemiological Characteristics of 2143 Pediatric Patients With 2019 Coronavirus Disease in China. *Pediatrics* e20200702 (2020) doi:10.1542/peds.2020-0702.
10. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580–585 (2013).
11. Carithers, L. J. *et al.* A Novel Approach to High-Quality Postmortem Tissue Procurement: The GTEx Project. *Biopreserv Biobank* **13**, 311–319 (2015).

12. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260–1263 (2020).
13. Wan, Y., Shang, J., Graham, R., Baric, R. S. & Li, F. Receptor Recognition by the Novel Coronavirus from Wuhan: an Analysis Based on Decade-Long Structural Studies of SARS Coronavirus. *Journal of Virology* **94**, (2020).
14. 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.
15. Walls, A. C. *et al.* Structure, Function, and Antigenicity of the SARS-CoV-2 Spike Glycoprotein. *Cell* (2020) doi:10.1016/j.cell.2020.02.058.
16. Remuzzi, A. & Remuzzi, G. COVID-19 and Italy: what next? *The Lancet* **0**, (2020).
17. Wu, C. *et al.* Risk Factors Associated With Acute Respiratory Distress Syndrome and Death in Patients With Coronavirus Disease 2019 Pneumonia in Wuhan, China. *JAMA Intern Med* (2020) doi:10.1001/jamainternmed.2020.0994.
18. Team, T. N. C. P. E. R. E. The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) — China, 2020. *CCDCW* **2**, 113–122 (2020).
19. Madissoon, E. *et al.* scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biology* **21**, 1 (2019).
20. 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.
21. Qi, F., Qian, S., Zhang, S. & Zhang, Z. Single cell RNA sequencing of 13 human tissues identify cell types and receptors of human coronaviruses. *Biochemical and Biophysical Research Communications* (2020) doi:10.1016/j.bbrc.2020.03.044.

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

23. Magalhães, J. P. de & Toussaint, O. GenAge: a genomic and proteomic network map of human ageing. *FEBS Letters* **571**, 243–247 (2004).

24. Yang, J. *et al.* Identification of human age-associated gene co-expressions in functional modules using liquid association. *Oncotarget* **9**, 1063–1074 (2017).

25. Sharma, G. & Goodwin, J. Effect of aging on respiratory system physiology and immunology. *Clin Interv Aging* **1**, 253–260 (2006).

26. Mohebali, D., Kaplan, D., Carlisle, M., Supiano, M. A. & Rondina, M. T. Alterations in Platelet Functions During Aging: Clinical Correlations with Thrombo-Inflammatory Disease in Older Adults. *J Am Geriatr Soc* **62**, 529–535 (2014).

27. Kuba, K. *et al.* A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus–induced lung injury. *Nat Med* **11**, 875–879 (2005).

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

29. Drosten, C. *et al.* Identification of a Novel Coronavirus in Patients with Severe Acute Respiratory Syndrome. *New England Journal of Medicine* **348**, 1967–1976 (2003).

30. Donoghue Mary *et al.* A Novel Angiotensin-Converting Enzyme–Related Carboxypeptidase (ACE2) Converts Angiotensin I to Angiotensin 1-9. *Circulation Research* **87**, e1–e9 (2000).

31. Tipnis, S. R. *et al.* A Human Homolog of Angiotensin-converting Enzyme CLONING AND FUNCTIONAL EXPRESSION AS A CAPTOPRIL-INSENSITIVE CARBOXYPEPTIDASE. *J. Biol. Chem.* **275**, 33238–33243 (2000).

32. Crackower, M. A. *et al.* Angiotensin-converting enzyme 2 is an essential regulator of heart function. *Nature* **417**, 822–828 (2002).

33. Wang, D. *et al.* Clinical Characteristics of 138 Hospitalized Patients With 2019 Novel Coronavirus–Infected Pneumonia in Wuhan, China. *JAMA* **323**, 1061–1069 (2020).

34. Srivastava, S. The Mitochondrial Basis of Aging and Age-Related Disorders. *Genes (Basel)* **8**, (2017).

35. Sun, N., Youle, R. J. & Finkel, T. The Mitochondrial Basis of Aging. *Mol Cell* **61**, 654–666 (2016).

36. Petersen, K. F. *et al.* Mitochondrial dysfunction in the elderly: possible role in insulin resistance. *Science* **300**, 1140–1142 (2003).

37. Lu, Y., Liu, D. X. & Tam, J. P. Lipid rafts are involved in SARS-CoV entry into Vero E6 cells. *Biochem. Biophys. Res. Commun.* **369**, 344–349 (2008).

38. Wang, H. *et al.* SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. *Cell Res.* **18**, 290–301 (2008).

39. Li, G.-M., Li, Y.-G., Yamate, M., Li, S.-M. & Ikuta, K. Lipid rafts play an important role in the early stage of severe acute respiratory syndrome-coronavirus life cycle. *Microbes Infect.* **9**, 96–102 (2007).

40. Pike, L. J. Lipid rafts: heterogeneity on the high seas. *Biochem. J.* **378**, 281–292 (2004).

41. Huang, I.-C. *et al.* SARS coronavirus, but not human coronavirus NL63, utilizes cathepsin L to infect ACE2-expressing cells. *J. Biol. Chem.* **281**, 3198–3203 (2006).

42. Simmons, G. *et al.* Inhibitors of cathepsin L prevent severe acute respiratory syndrome coronavirus entry. *Proc. Natl. Acad. Sci. U.S.A.* **102**, 11876–11881 (2005).

43. Gralinski, L. E. & Baric, R. S. Molecular pathology of emerging coronavirus infections. *J Pathol* **235**, 185–195 (2015).

44. Peiris, J. S. M. *et al.* Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. *Lancet* **361**, 1767–1772 (2003).

45. Aran, D., Hu, Z. & Butte, A. J. xCell: digitally portraying the tissue cellular heterogeneity landscape. *Genome Biol* **18**, (2017).

46. Navarro, S. & Driscoll, B. Regeneration of the Aging Lung: A Mini-Review. *Gerontology* **63**, 270–280 (2017).

47. Lowery, E. M., Brubaker, A. L., Kuhlmann, E. & Kovacs, E. J. The aging lung. *Clin Interv Aging* **8**, 1489–1496 (2013).

48. Byrne, A. J., Mathie, S. A., Gregory, L. G. & Lloyd, C. M. Pulmonary macrophages: key players in the innate defence of the airways. *Thorax* **70**, 1189–1196 (2015).

49. Ding, Y. *et al.* The clinical pathology of severe acute respiratory syndrome (SARS): a report from China. *J. Pathol.* **200**, 282–289 (2003).

50. Hsiao, C. H. *et al.* Evolution of pulmonary pathology in severe acute respiratory syndrome. *J. Formos. Med. Assoc.* **104**, 75–81 (2005).

51. Braciale, T. J., Sun, J. & Kim, T. S. Regulating the adaptive immune response to respiratory virus infection. *Nat. Rev. Immunol.* **12**, 295–305 (2012).

52. Holt, P. G., Strickland, D. H., Wikström, M. E. & Jahnsen, F. L. Regulation of immunological homeostasis in the respiratory tract. *Nat Rev Immunol* **8**, 142–152 (2008).

53. Allard, B., Panariti, A. & Martin, J. G. Alveolar Macrophages in the Resolution of Inflammation, Tissue Repair, and Tolerance to Infection. *Front Immunol* **9**, (2018).

54. Yip, M. S. *et al.* Antibody-dependent infection of human macrophages by severe acute respiratory syndrome coronavirus. *Virol. J.* **11**, 82 (2014).

55. Cheung, C. Y. *et al.* Cytokine Responses in Severe Acute Respiratory Syndrome Coronavirus-Infected Macrophages In Vitro: Possible Relevance to Pathogenesis. *J Virol* **79**, 7819–7826 (2005).

56. Thepen, T., Van Rooijen, N. & Kraal, G. Alveolar macrophage elimination in vivo is associated with an increase in pulmonary immune response in mice. *J Exp Med* **170**, 499–509 (1989).

57. Zhao, J., Zhao, J., Van Rooijen, N. & Perlman, S. Evasion by stealth: inefficient immune activation underlies poor T cell response and severe disease in SARS-CoV-infected mice. *PLoS Pathog.* **5**, e1000636 (2009).

58. Shapouri-Moghaddam, A. *et al.* Macrophage plasticity, polarization, and function in health and disease. *J. Cell. Physiol.* **233**, 6425–6440 (2018).

59. Zhao, J., Zhao, J. & Perlman, S. T cell responses are required for protection from clinical disease and for virus clearance in severe acute respiratory syndrome coronavirus-infected mice. *J. Virol.* **84**, 9318–9325 (2010).

60. Chen, J. *et al.* Cellular immune responses to severe acute respiratory syndrome coronavirus (SARS-CoV) infection in senescent BALB/c mice: CD4+ T cells are important in control of SARS-CoV infection. *J. Virol.* **84**, 1289–1301 (2010).

61. Jin, H. *et al.* Induction of Th1 type response by DNA vaccinations with N, M, and E genes against SARS-CoV in mice. *Biochem. Biophys. Res. Commun.* **328**, 979–986 (2005).

62. Swain, S. L., McKinstry, K. K. & Strutt, T. M. Expanding roles for CD4+ T cells in immunity to viruses. *Nat Rev Immunol* **12**, 136–148 (2012).

63. Wilde, A. H. de *et al.* A Kinome-Wide Small Interfering RNA Screen Identifies Proviral and Antiviral Host Factors in Severe Acute Respiratory Syndrome Coronavirus Replication, Including Double-Stranded RNA-Activated Protein Kinase and Early Secretory Pathway Proteins. *Journal of Virology* **89**, 8318–8333 (2015).

64. Gordon, D. E. *et al.* A SARS-CoV-2-Human Protein-Protein Interaction Map Reveals Drug Targets and Potential Drug-Repurposing. *bioRxiv* 2020.03.22.002386 (2020) doi:10.1101/2020.03.22.002386.

65. Scholich, K., Pierre, S. & Patel, T. B. Protein associated with Myc (PAM) is a potent inhibitor of adenylyl cyclases. *J. Biol. Chem.* **276**, 47583–47589 (2001).

66. Pierre, S. C., Häusler, J., Birod, K., Geisslinger, G. & Scholich, K. PAM mediates sustained inhibition of cAMP signaling by sphingosine-1-phosphate. *EMBO J* **23**, 3031–3040 (2004).

67. Billington, C. K., Ojo, O. O., Penn, R. B. & Ito, S. cAMP Regulation of Airway Smooth Muscle Function. *Pulm Pharmacol Ther* **26**, 112–120 (2013).

68. Kilic, O. *et al.* A microphysiological model of the bronchial airways reveals the interplay of mechanical and biochemical signals in bronchospasm. *Nat Biomed Eng* **3**, 532–544 (2019).

69. Huang, D. W., Sherman, B. T. & Lempicki, R. A. Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat Protoc* **4**, 44–57 (2009).

70. Butler, A., Hoffman, P., Smibert, P., Papalexi, E. & Satija, R. Integrating single-cell transcriptomic data across different conditions, technologies, and species. *Nat Biotechnol* **36**, 411–420 (2018).

71. Stuart, T. *et al.* Comprehensive Integration of Single-Cell Data. *Cell* **177**, 1888-1902.e21 (2019).

72. Gaujoux, R. & Seoighe, C. A flexible R package for nonnegative matrix factorization. *BMC Bioinformatics* **11**, 367 (2010).

73. Barter, R. L. & Yu, B. Superheat: An R Package for Creating Beautiful and Extendable Heatmaps for Visualizing Complex Data. *Journal of Computational and Graphical Statistics* **27**, 910–922 (2018).

Figure 1

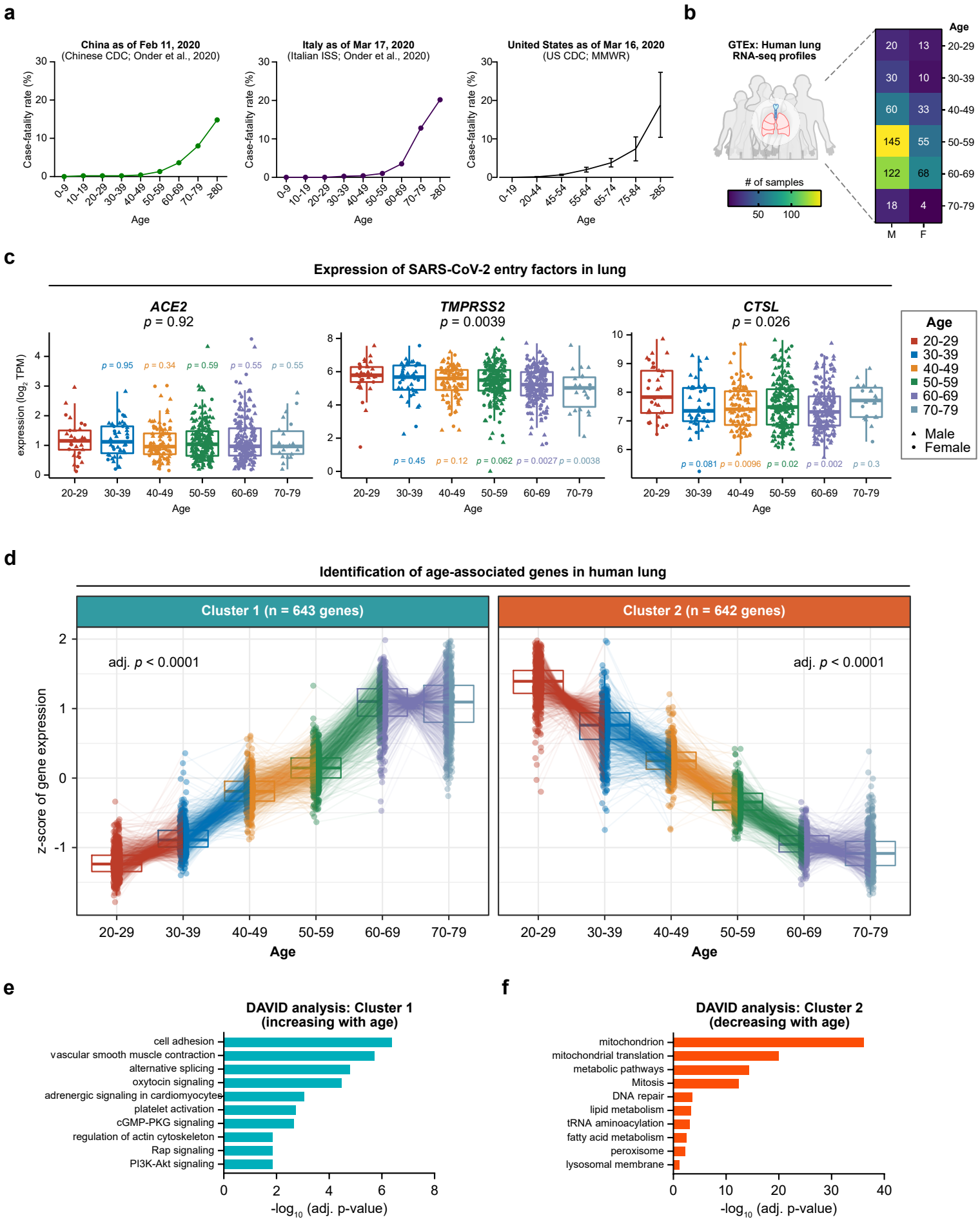

a
China as of Feb 11, 2020
(Chinese CDC; Onder et al., 2020)
Italy as of Mar 17, 2020
(Italian ISS; Onder et al., 2020)
United States as of Mar 16, 2020
(US CDC; MMWR)
Case-fatality rate (%)
Age
b
GTEx: Human lung
RNA-seq profiles
# of samples
Age
M
F
c
Expression of SARS-CoV-2 entry factors in lung
ACE2
p = 0.92
TMPRSS2
p = 0.0039
CTSL
p = 0.026
expression (log2 TPM)
Age
Male
Female
d
Identification of age-associated genes in human lung
Cluster 1 (n = 643 genes)
Cluster 2 (n = 642 genes)
adj. p < 0.0001
z-score of gene expression
Age
e
DAVID analysis: Cluster 1
(increasing with age)
cell adhesion
vascular smooth muscle contraction
alternative splicing
oxytocin signaling
adrenergic signaling in cardiomyocytes
platelet activation
cGMP-PKG signaling
regulation of actin cytoskeleton
Rap signaling
PI3K-Akt signaling
-log10 (adj. p-value)
f
DAVID analysis: Cluster 2
(decreasing with age)
mitochondrion
mitochondrial translation
metabolic pathways
Mitosis
DNA repair
lipid metabolism
tRNA aminoacylation
fatty acid metabolism
peroxisome
lysosomal membrane
-log10 (adj. p-value)

Figure 2

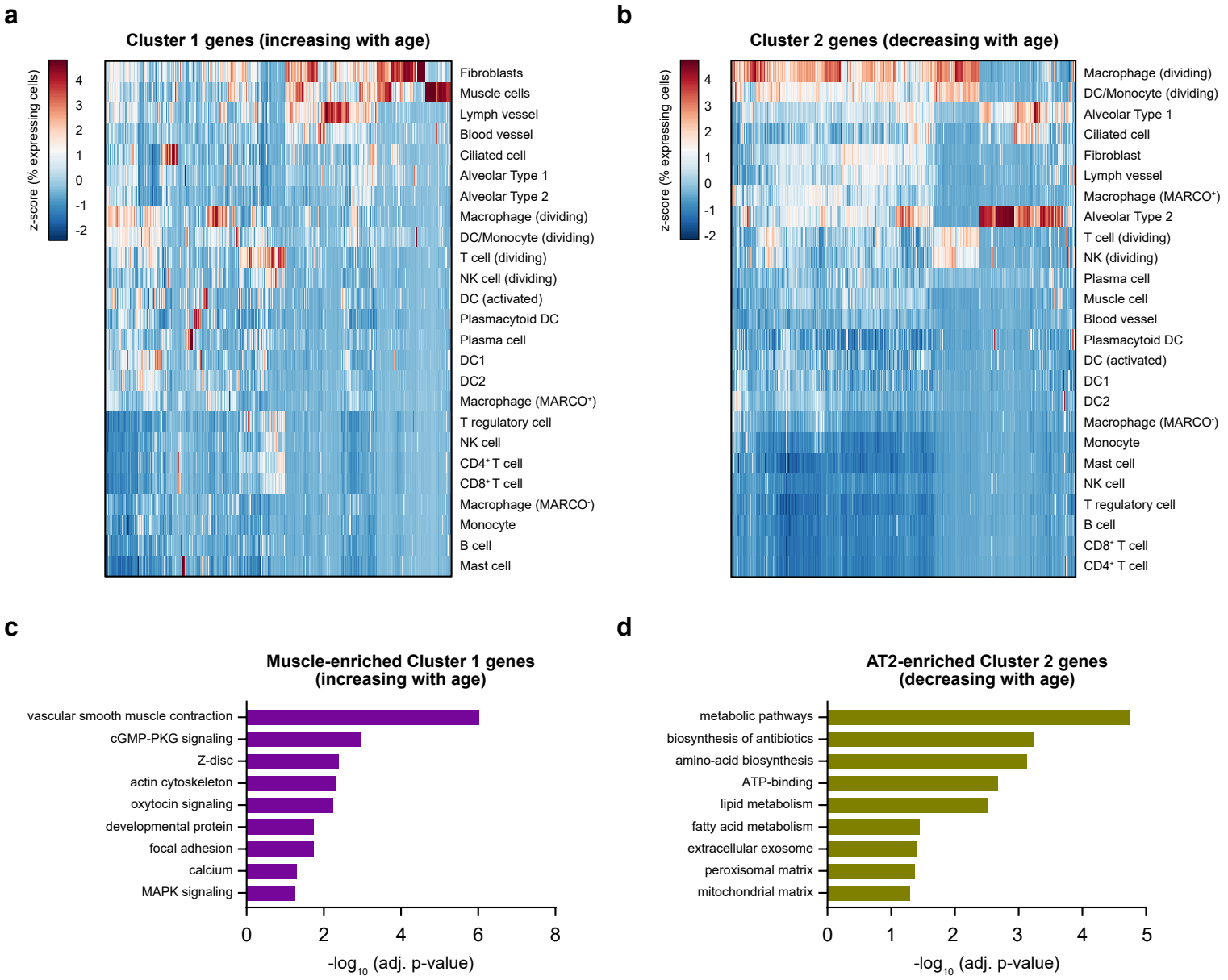

a
Cluster 1 genes (increasing with age)
z-score (% expressing cells)
4
3
2
1
0
-1
-2
Fibroblasts
Muscle cells
Lymph vessel
Blood vessel
Ciliated cell
Alveolar Type 1
Alveolar Type 2
Macrophage (dividing)
DC/Monocyte (dividing)
T cell (dividing)
NK cell (dividing)
DC (activated)
Plasmacytoid DC
Plasma cell
DC1
DC2
Macrophage (MARCO+)
T regulatory cell
NK cell
CD4+ T cell
CD8+ T cell
Macrophage (MARCO-)
Monocyte
B cell
Mast cell
b
Cluster 2 genes (decreasing with age)
z-score (% expressing cells)
4
3
2
1
0
-1
-2
Macrophage (dividing)
DC/Monocyte (dividing)
Alveolar Type 1
Ciliated cell
Fibroblast
Lymph vessel
Macrophage (MARCO+)
Alveolar Type 2
T cell (dividing)
NK (dividing)
Plasma cell
Muscle cell
Blood vessel
Plasmacytoid DC
DC (activated)
DC1
DC2
Macrophage (MARCO-)
Monocyte
Mast cell
NK cell
T regulatory cell
B cell
CD8+ T cell
CD4+ T cell
c
Muscle-enriched Cluster 1 genes
(increasing with age)
vascular smooth muscle contraction
cGMP-PKG signaling
Z-disc
actin cytoskeleton
oxytocin signaling
developmental protein
focal adhesion
calcium
MAPK signaling
0
2
4
6
8
-log10 (adj. p-value)
d
AT2-enriched Cluster 2 genes
(decreasing with age)
metabolic pathways
biosynthesis of antibiotics
amino-acid biosynthesis
ATP-binding
lipid metabolism
fatty acid metabolism
extracellular exosome
peroxisomal matrix
mitochondrial matrix
0
1
2
3
4
5
-log10 (adj. p-value)

Figure 3

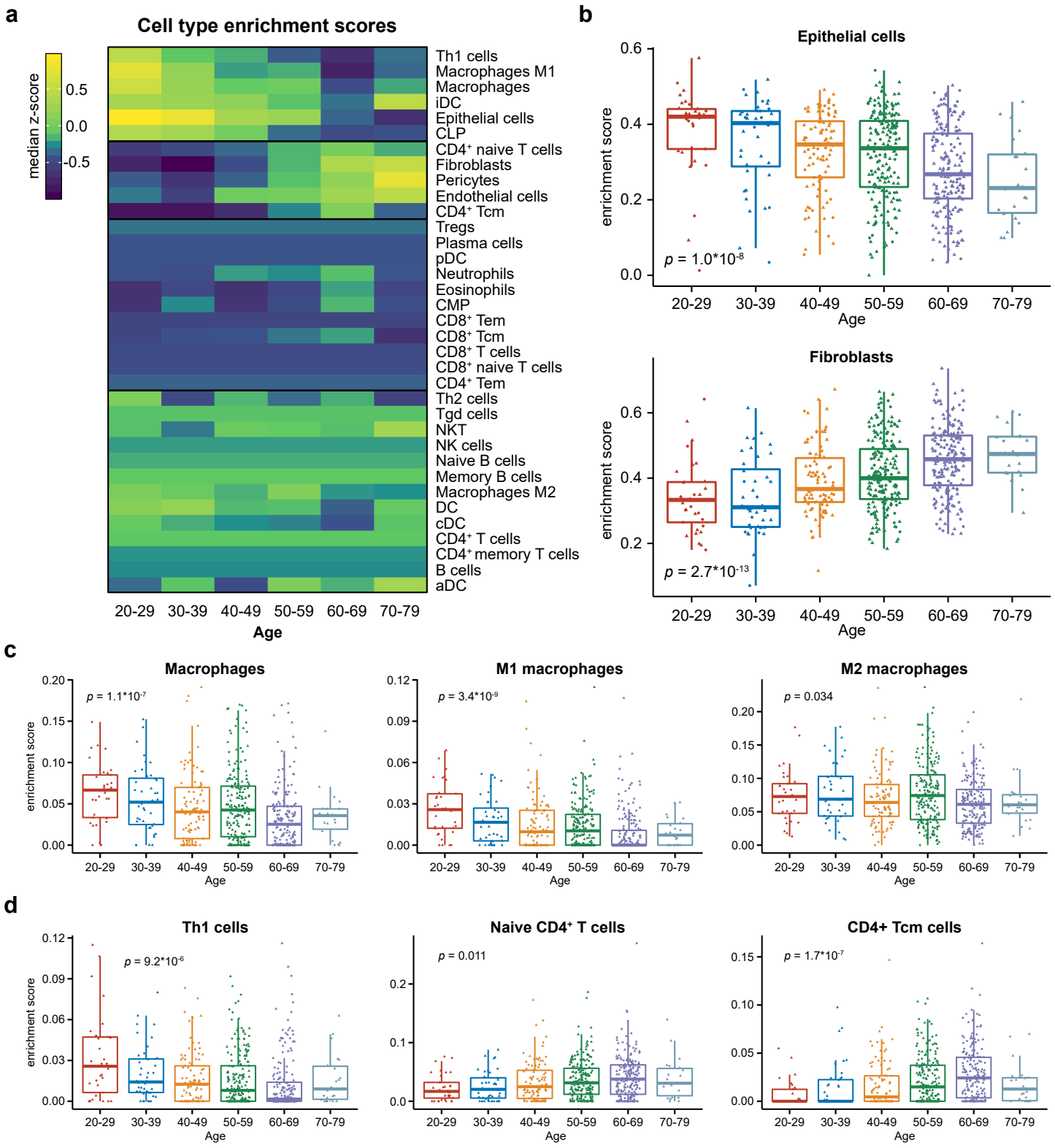

a
Cell type enrichment scores
median z-score
0.5
0.0
-0.5
Th1 cells
Macrophages M1
Macrophages
iDC
Epithelial cells
CLP
CD4+ naive T cells
Fibroblasts
Pericytes
Endothelial cells
CD4+ Tcm
Tregs
Plasma cells
pDC
Neutrophils
Eosinophils
CMP
CD8+ Tem
CD8+ Tcm
CD8+ T cells
CD8+ naive T cells
CD4+ Tem
Th2 cells
Tgd cells
NKT
NK cells
Naive B cells
Memory B cells
Macrophages M2
DC
cDC
CD4+ T cells
CD4+ memory T cells
B cells
aDC
20-29 30-39 40-49 50-59 60-69 70-79
Age
b
Epithelial cells
enrichment score
p = 1.0*10-8
Age
Fibroblasts
enrichment score
p = 2.7*10-13
Age
c
Macrophages
p = 1.1*10-7
M1 macrophages
p = 3.4*10-9
M2 macrophages
p = 0.034
enrichment score
Age
d
Th1 cells
p = 9.2*10-6
Naive CD4+ T cells
p = 0.011
CD4+ Tcm cells
p = 1.7*10-7
enrichment score
Age

Figure 4

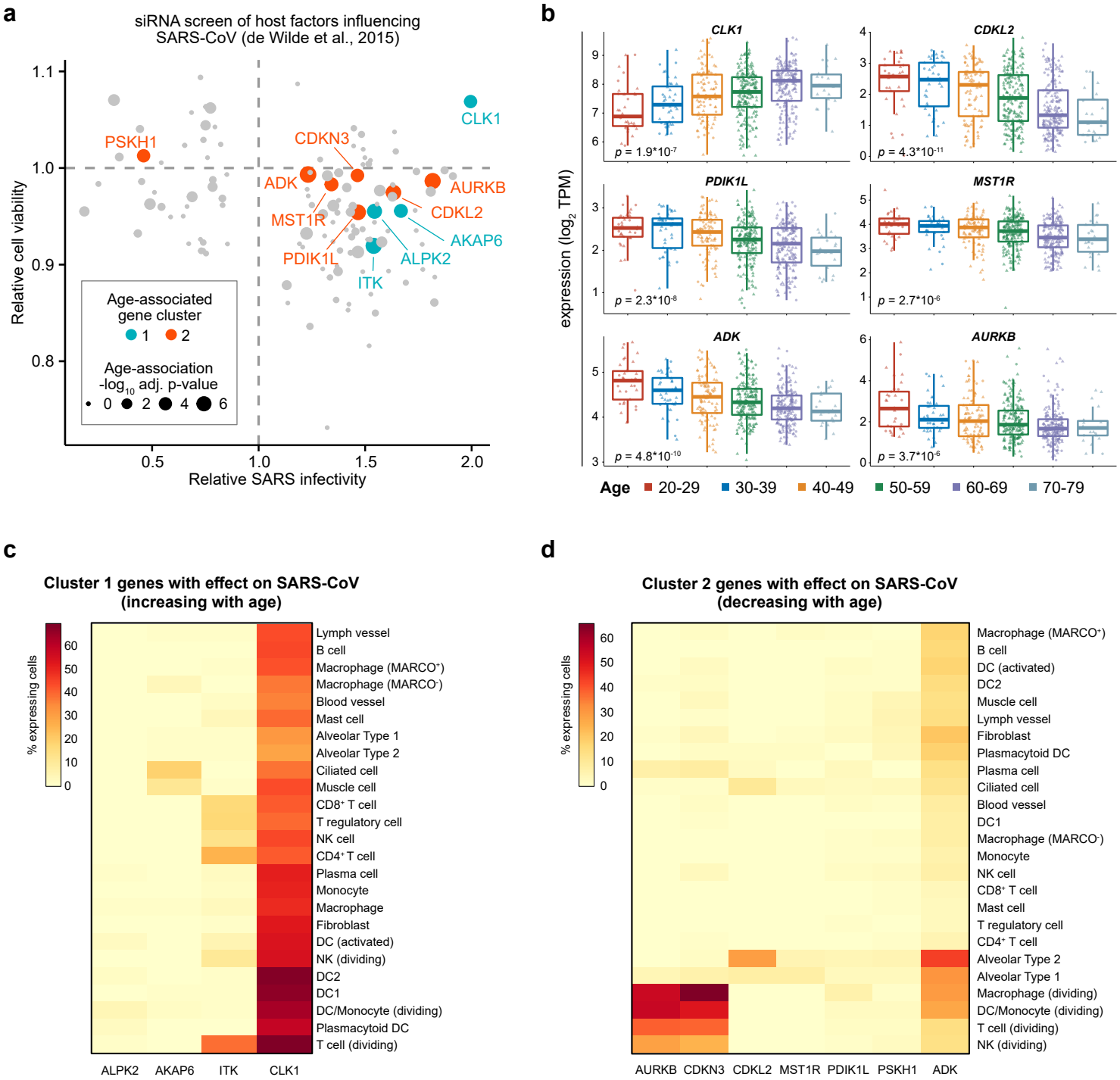
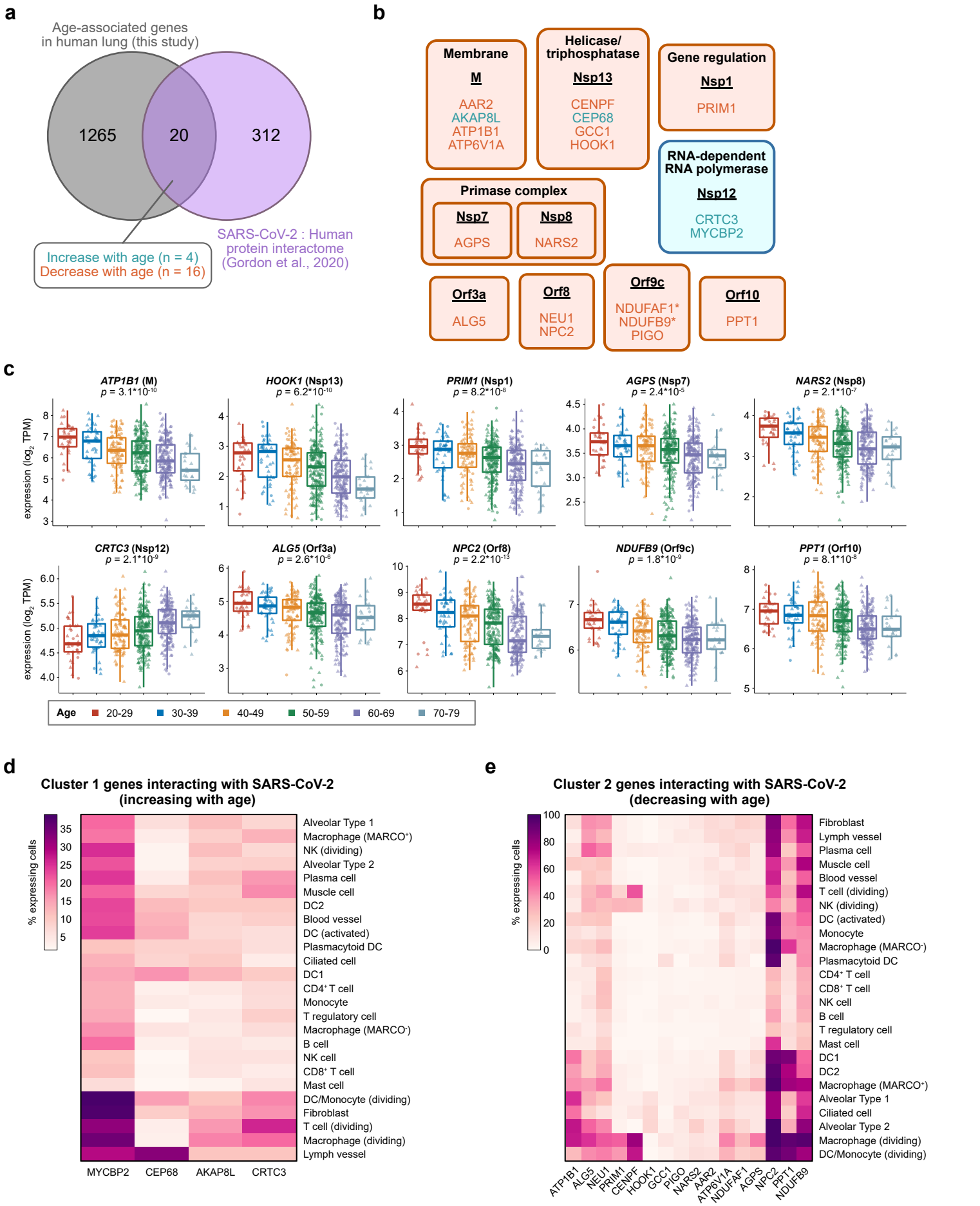
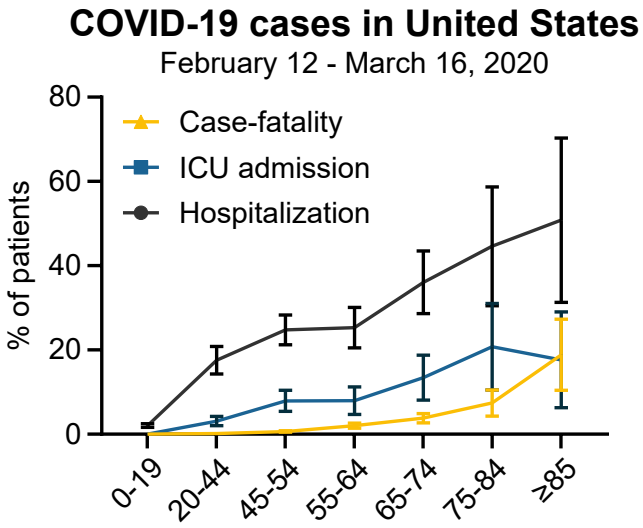

a
siRNA screen of host factors influencing SARS-CoV (de Wilde et al., 2015)
Relative cell viability
Relative SARS infectivity
PSKH1
CDKN3
CLK1
ADK
AURKB
MST1R
CDKL2
AKAP6
PDIK1L
ALPK2
ITK
Age-associated gene cluster
1
2
Age-association -log10 adj. p-value
0
2
4
6
b
CLK1
CDKL2
PDIK1L
MST1R
ADK
AURKB
expression (log2 TPM)
p = 1.9*10-7
p = 4.3*10-11
p = 2.3*10-8
p = 2.7*10-6
p = 4.8*10-10
p = 3.7*10-6
Age
20-29
30-39
40-49
50-59
60-69
70-79
c
Cluster 1 genes with effect on SARS-CoV (increasing with age)
% expressing cells
Lymph vessel
B cell
Macrophage (MARCO+)
Macrophage (MARCO-)
Blood vessel
Mast cell
Alveolar Type 1
Alveolar Type 2
Ciliated cell
Muscle cell
CD8+ T cell
T regulatory cell
NK cell
CD4+ T cell
Plasma cell
Monocyte
Macrophage
Fibroblast
DC (activated)
NK (dividing)
DC2
DC1
DC/Monocyte (dividing)
Plasmacytoid DC
T cell (dividing)
ALPK2
AKAP6
ITK
CLK1
d
Cluster 2 genes with effect on SARS-CoV (decreasing with age)
% expressing cells
Macrophage (MARCO+)
B cell
DC (activated)
DC2
Muscle cell
Lymph vessel
Fibroblast
Plasmacytoid DC
Plasma cell
Ciliated cell
Blood vessel
DC1
Macrophage (MARCO-)
Monocyte
NK cell
CD8+ T cell
Mast cell
T regulatory cell
CD4+ T cell
Alveolar Type 2
Alveolar Type 1
Macrophage (dividing)
DC/Monocyte (dividing)
T cell (dividing)
NK (dividing)
AURKB
CDKN3
CDKL2
MST1R
PDIK1L
PSKH1
ADK

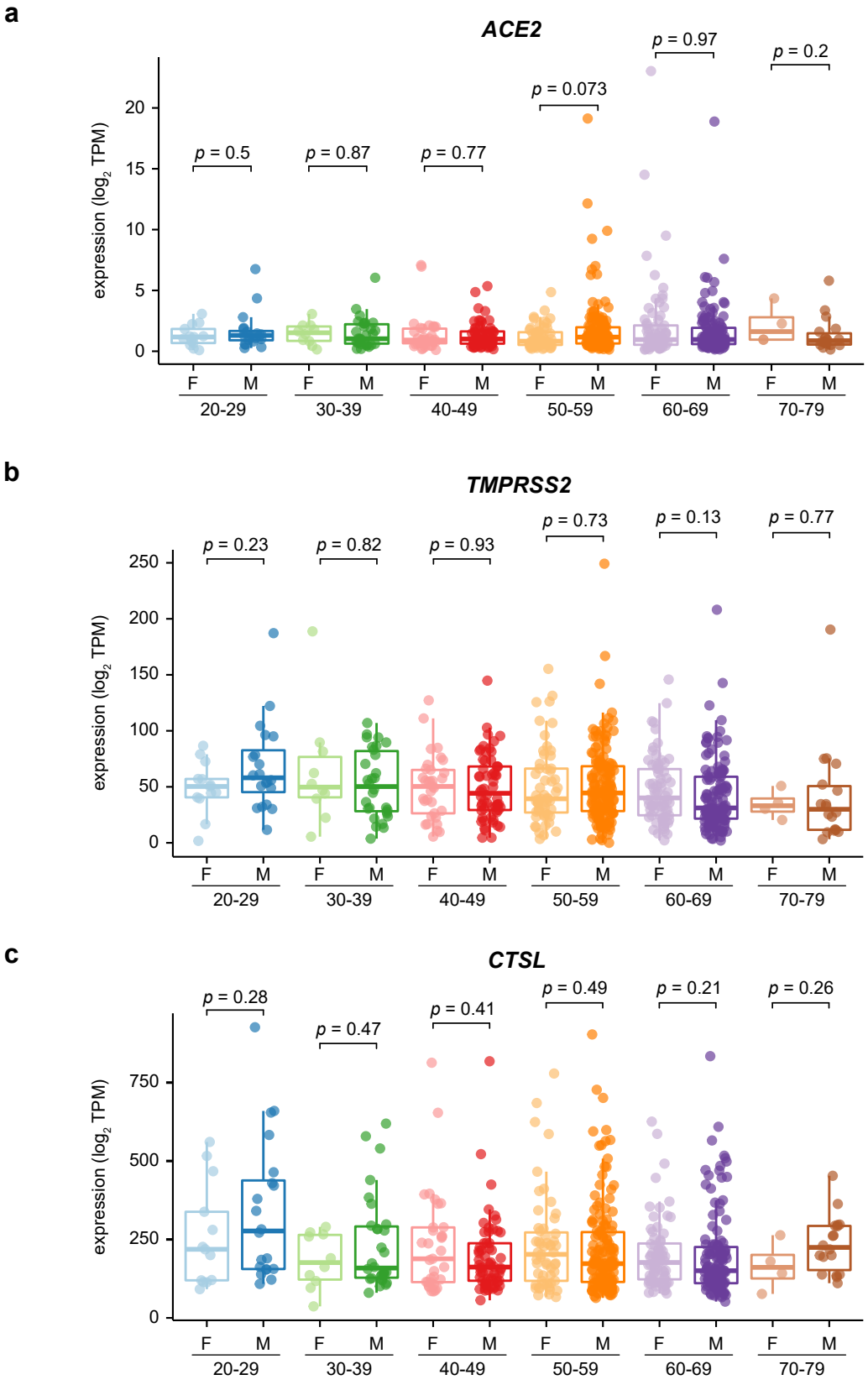
Figure 5


a
Age-associated genes in human lung (this study)
1265
20
312
SARS-CoV-2 : Human protein interactome (Gordon et al., 2020)
Increase with age (n = 4)
Decrease with age (n = 16)
b
Membrane
M
AAR2
AKAP8L
ATP1B1
ATP6V1A
Helicase/ triphosphatase
Nsp13
CENPF
CEP68
GCC1
HOOK1
Gene regulation
Nsp1
PRIM1
RNA-dependent RNA polymerase
Nsp12
CRTC3
MYCBP2
Primase complex
Nsp7
AGPS
Nsp8
NARS2
Orf3a
ALG5
Orf8
NEU1
NPC2
Orf9c
NDUFAF1*
NDUFB9*
PIGO
Orf10
PPT1
c
ATP1B1 (M) p = 3.1*10-10
HOOK1 (Nsp13) p = 6.2*10-10
PRIM1 (Nsp1) p = 8.2*10-8
AGPS (Nsp7) p = 2.4*10-5
NARS2 (Nsp8) p = 2.1*10-7
CRTC3 (Nsp12) p = 2.1*10-9
ALG5 (Orf3a) p = 2.6*10-6
NPC2 (Orf8) p = 2.2*10-13
NDUFB9 (Orf9c) p = 1.8*10-9
PPT1 (Orf10) p = 8.1*10-8
expression (log2 TPM)
Age 20-29 30-39 40-49 50-59 60-69 70-79
d
Cluster 1 genes interacting with SARS-CoV-2 (increasing with age)
% expressing cells
Alveolar Type 1
Macrophage (MARCO+)
NK (dividing)
Alveolar Type 2
Plasma cell
Muscle cell
DC2
Blood vessel
DC (activated)
Plasmacytoid DC
Ciliated cell
DC1
CD4+ T cell
Monocyte
T regulatory cell
Macrophage (MARCO-)
B cell
NK cell
CD8+ T cell
Mast cell
DC/Monocyte (dividing)
Fibroblast
T cell (dividing)
Macrophage (dividing)
Lymph vessel
MYCBP2 CEP68 AKAP8L CRTC3
e
Cluster 2 genes interacting with SARS-CoV-2 (decreasing with age)
% expressing cells
Fibroblast
Lymph vessel
Plasma cell
Muscle cell
Blood vessel
T cell (dividing)
NK (dividing)
DC (activated)
Monocyte
Macrophage (MARCO-)
Plasmacytoid DC
CD4+ T cell
CD8+ T cell
NK cell
B cell
T regulatory cell
Mast cell
DC1
DC2
Macrophage (MARCO+)
Alveolar Type 1
Ciliated cell
Alveolar Type 2
Macrophage (dividing)
DC/Monocyte (dividing)
ATP1B1 ALG5 NEU1 PRIM1 CENPF HOOK1 GCC1 PIGO NARS2 AAR2 ATP6V1A NDUFAF1 AGPS NPC2 PPT1 NDUFB9

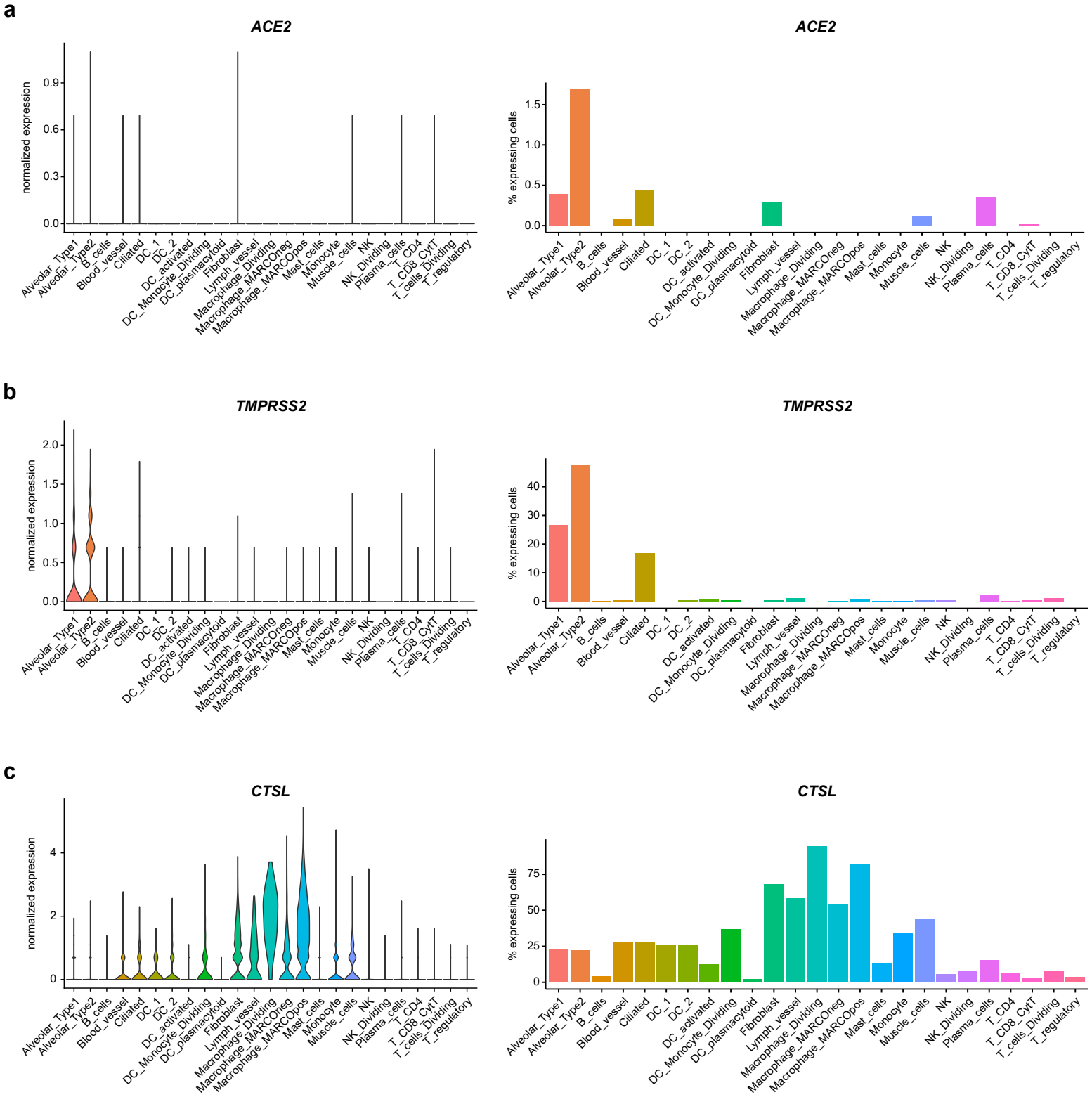
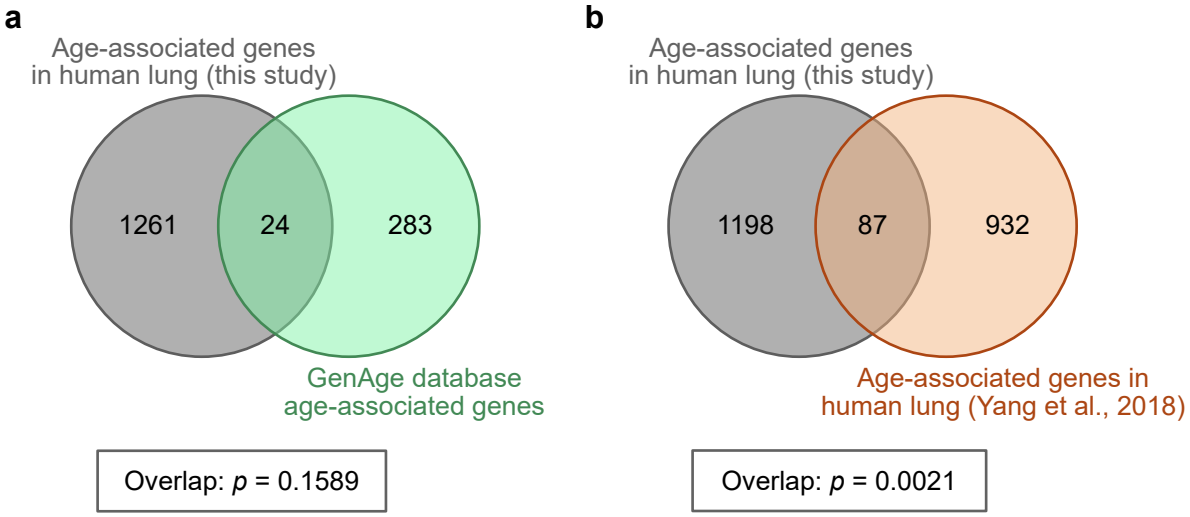
Supplementary Figure 1



Supplementary Figure 2



Supplementary Figure 3



Supplementary Figure 4