

Opinion**ACE2 and COVID-19 Susceptibility and Severity**Ming Zheng^{1,2,*}¹Institute of Military Cognition and Brain Sciences, Academy of Military Medical Sciences, Beijing, China²Beijing Institute of Basic Medical Sciences, Beijing, China

[Received May 7, 2021; Revised August 4, 2021; Accepted August 5, 2021]

ABSTRACT: Sick, male, and older populations are more vulnerable to COVID-19. However, it remains unclear whether a common mechanism exists across different demographic characteristics. SARS-CoV-2 infection is initiated by the specific binding of the viral spike protein to angiotensin-converting enzyme 2 (ACE2). This study analyzed the demographics of pulmonary ACE2 expression, Mendelian randomization (MR) of ACE2 and COVID-19, and comparative tropism of SARS-CoV-2. The key features of SARS-CoV-2 tropism, including pulmonary ACE2 expression and ACE2-expressing cell types, showed distinct subphenotypes associated with the demographics of vulnerable COVID-19 populations, suggesting a hypothesis centered on “ACE2” to explain their interplay. Next, by integrating multiple COVID-19 cohorts of genome-wide association studies (GWASs) and cis-expression quantitative trait loci (cis-eQTLs) of ACE2, MR analysis demonstrated that ACE2 played a causal role in COVID-19 susceptibility and severity, suggesting ACE2 as a promising target for early COVID-19 treatment. Next, by analyzing the expression of host cell receptors using single-cell RNA sequencing (scRNA-seq) data of human lung tissues, comparative tropism analysis showed that SARS-CoV-2 and other respiratory viruses, but not non-respiratory viruses, had remarkably overlapping and enriched cellular tropism in alveolar type 2 (AT2) cells. This finding indicates the possibility of coinfection with SARS-CoV-2 and other respiratory viruses, perhaps implying sociovirology at the cellular level. Moreover, the binding of viral entry proteins to the compatible host cell receptors is under strong natural selection pressure. Therefore, comparative tropism might reveal the footprint of natural selection that shapes the virus population, which provides a novel perspective for understanding zoonotic spillover events.

Key words: COVID-19, SARS-CoV-2, ACE2, host cell receptor, viral tropism, GWAS, cis-eQTL, scRNA-seq, Mendelian randomization

Since 2020, the coronavirus disease 2019 (COVID-19) outbreak has become a global pandemic. To date, the number of COVID-19 worldwide has exceeded 200 million (<https://covid19.who.int/>). It has been reported that COVID-19 patients with demographic characteristics of pre-existing disease, male sex, and older age are particularly prone to severe clinical outcomes, such as hospitalization, ICU admission, and death [1]. In the U.S., patients who die from COVID-19 are predominantly individuals with pre-existing diseases (96.4%) and aged

more than 50 years (95.4%) [1]. Despite accumulating evidence demonstrating the substantial variation in the COVID-19 outcomes across demographics, it remains unclear whether a common mechanism exists across different demographic characteristics, including medical conditions, sex, and age.

COVID-19 is caused by a novel coronavirus, designated severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) [2]. SARS-CoV-2 infection is initiated by the specific binding of the viral spike protein to

*Correspondence should be addressed to: Dr. Ming Zheng, Institute of Military Cognition and Brain Sciences, Academy of Military Medical Sciences, Beijing, China. E-mail: mmzheng@fmmu.edu.cn or zhengming_china@163.com.

Copyright: © 2021 Zheng M. This is an open-access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

angiotensin-converting enzyme 2 (ACE2) [3]. It is worth noting that the expression of ACE2 is associated with increased viral load in human cell lines [4, 5] and mice [6]. Moreover, ACE2 expression has been linked to the clinical consequences of COVID-19 [7]. Thus, ACE2

expression plays an essential role in both SARS-CoV-2 infection and COVID-19 manifestations. However, the fundamental question still exists whether ACE2 expression varies across demographics, and the role of ACE2 in COVID-19 remains unclear.

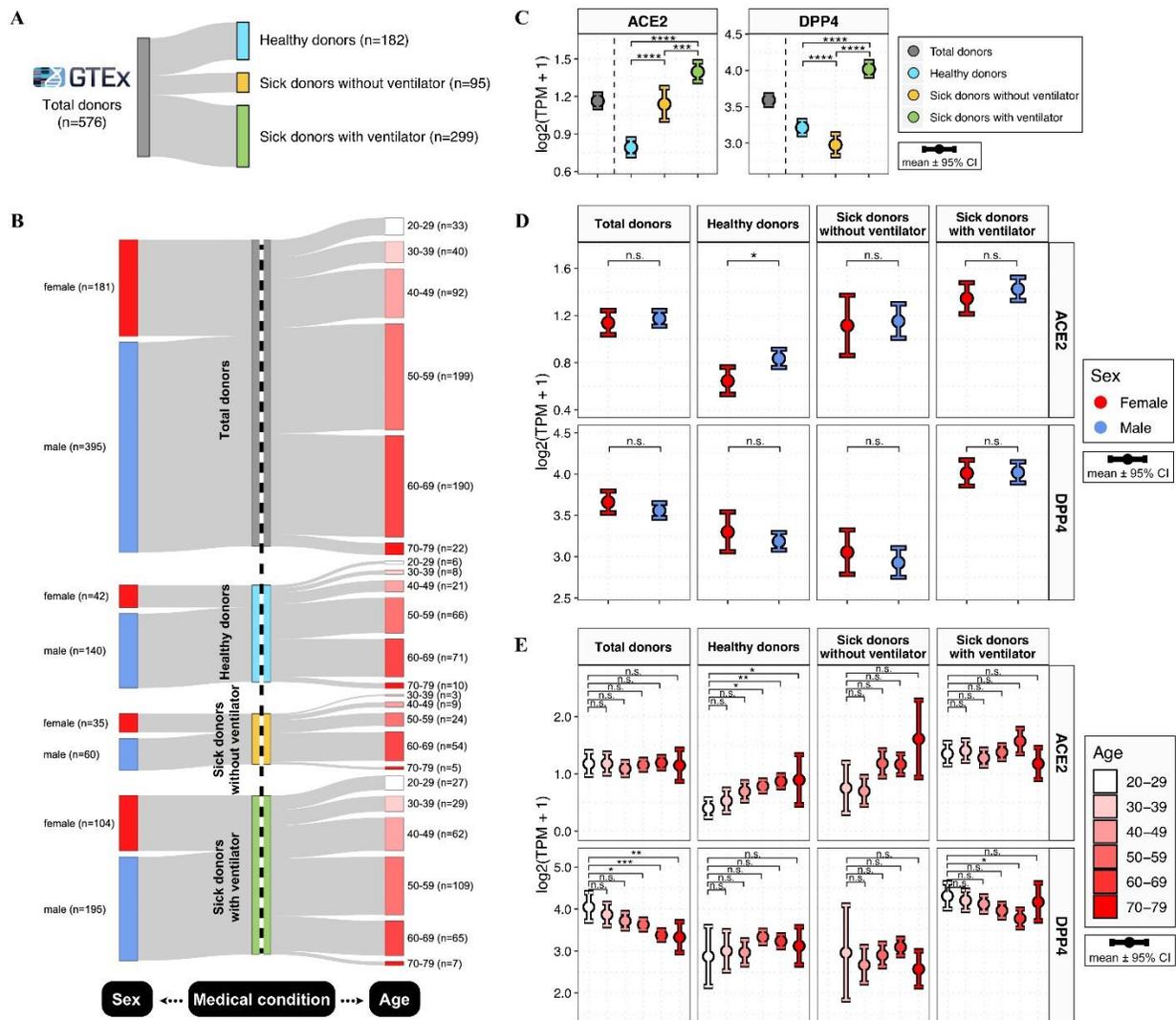


Figure 1. Demographics of pulmonary ACE2 and DPP4 expression. (A-B) Sankey plots show the sample sizes of GTEx donors across demographics, including medical conditions, sex, and age. (C-E) Distributions of ACE2 and DPP4 expression levels in lung tissues among all donors and the subgroups defined by medical condition (C), sex (D), and age (E). Expression levels were measured by log₂-transformed TPM (transcripts per million) values. The error bars represent the 95% confidence interval (CI), with color coding according to medical status (C), sex (D), and age (E). The results were considered statistically significant when $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***), and $P < 0.0001$ (****) using the Kruskal–Wallis test.

Pulmonary ACE2 expression across demographics

ACE2 is the host cell receptor for both SARS-CoV [8] and SARS-CoV-2 [3]. Both of these viruses cause viral pneumonia, which is the primary reason for fatality [9, 10]. Previously, positive ACE2 expression in human lung tissues was detected using immunohistochemistry

staining [11]. To extend the previous finding, we investigated the expression pattern of ACE2 in human lung tissues across demographic factors of medical condition, sex, and age.

Here, we analyzed the RNA-seq data from the GTEx cohort, which contains individuals of predominantly European ancestry [12]. ACE2 expression was analyzed

in 576 human lung tissue samples from 182 healthy donors and 394 sick donors naive to SARS-CoV-2 infection (Fig. 1A). Sick donors suffered from chronic illnesses, including cerebrovascular, cardiac, hepatic, renal, neurological and respiratory diseases. These common diseases were largely consistent with the pre-existing diseases reported in COVID-19 patients [1]. Additionally, the 394 sick donors included 95 sick donors

without ventilator usage and 299 sick donors with the therapeutic intervention of ventilator usage (Fig. 1A-B). Compared to the healthy donors, the sick donors, with or without ventilator support, had significantly higher ACE2 expression levels (Fig. 1C, left plot). Moreover, sick donors with ventilator support had markedly higher ACE2 expression levels than sick donors without ventilator support.

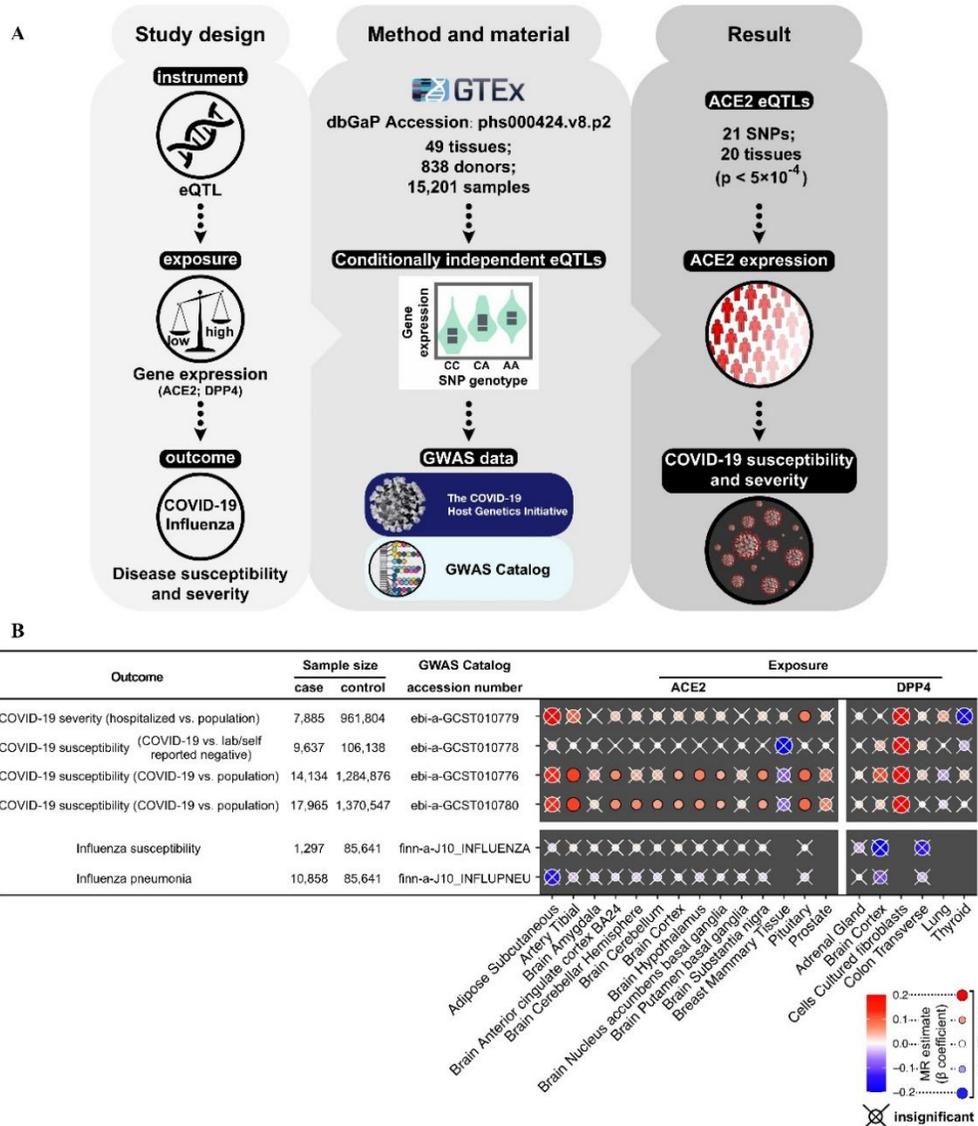


Figure 2. Mendelian randomization study of ACE2 and COVID-19 susceptibility and severity. (A) Graphical representation of the study design, method and material, and result of the Mendelian randomization (MR) study. This MR study used expression quantitative trait loci (eQTLs) to estimate the causal role of ACE2 expression in COVID-19 susceptibility and severity. This MR analysis tested the cis-eQTL proposed instruments of ACE2 and DPP4 against multiple genome-wide association study (GWAS) data of COVID-19 and influenza. (B) MR results show the causal role of ACE2 in the susceptibility and severity of COVID-19 but not influenza. MR estimates were calculated using inverse-variance weighting (IVW) for instruments that contained more than one cis-eQTL and the Wald ratio for instruments with one cis-eQTL. Thus, instruments containing only one cis-eQTL were not tested for IVW. Heatmap shows MR estimates available in at least one tissue and GWAS cohort. The color and size of dots represent the β estimate of the MR. A positive β estimate indicates that high gene expression is associated with an increased risk of disease susceptibility and severity. The results were considered statistically significant when P < 0.05, and the “x” represents insignificant results.

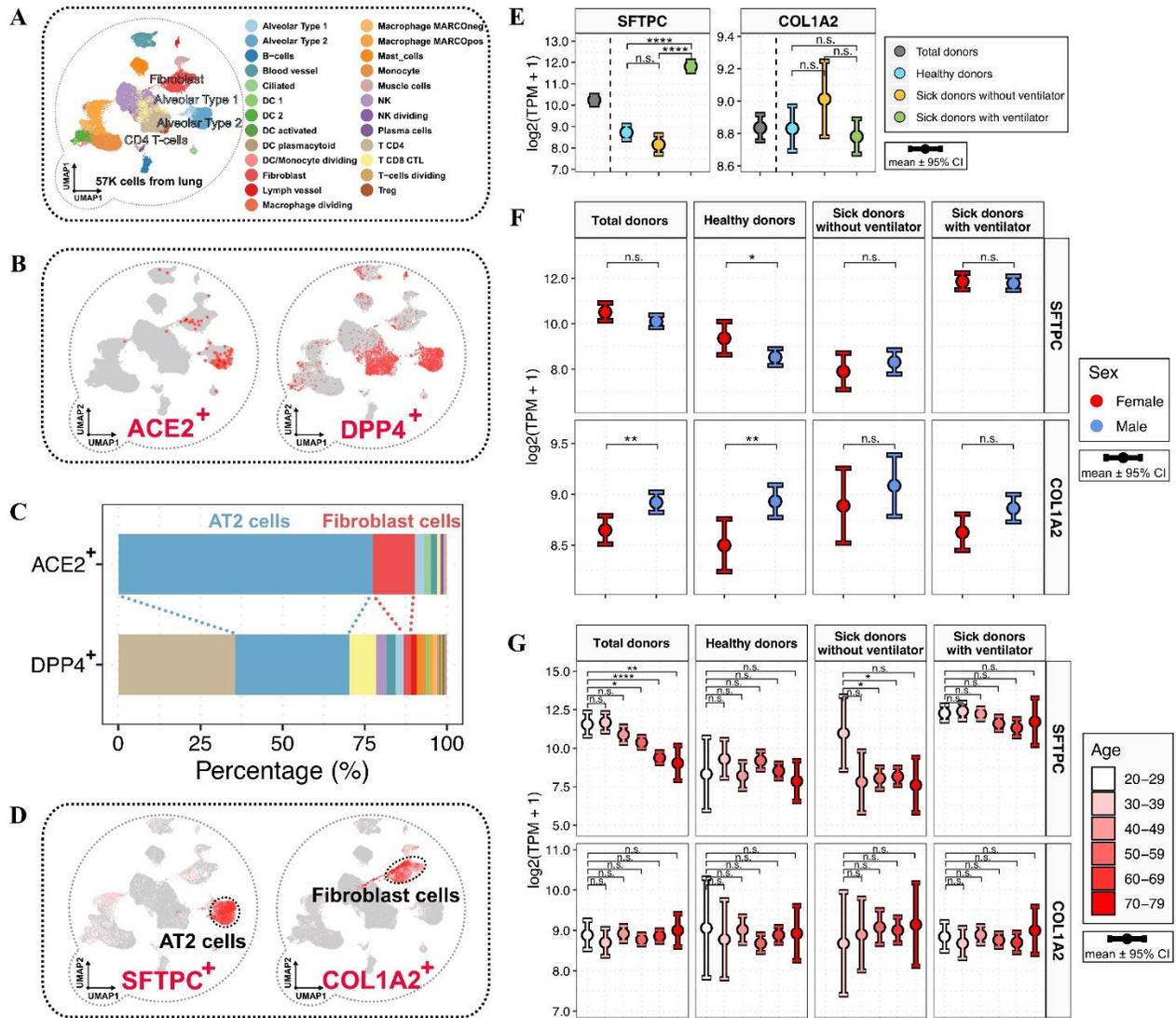


Figure 3. Pulmonary ACE2⁺ and DPP4⁺ cell types and the relative abundances of alveolar type 2 and fibroblast cells across demographics. (A) The uniform manifold approximation and projection (UMAP) plot shows the landscape of lung cells at the single-cell level, with color coding according to cell type. (B) UMAP plots show the expression levels of ACE2 and DPP4, with color coding according to the gene expression level. (C) The stack bar plot shows the percentages of pulmonary ACE2⁺ and DPP4⁺ cells, with color coding according to cell type. (D) UMAP plots show the expression levels of SFTPC and COL1A2, which were exclusively expressed in alveolar type 2 (AT2) and fibroblast cells, respectively. (E-G) Distributions of pulmonary SFTPC and COL1A2 expression levels in total donors and the subgroups defined by medical condition (E), sex (F), and age (G). The error bars represent the 95% confidence interval (CI), with color coding according to medical status (E), sex (F), and age (G). The results were considered statistically significant when $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***), and $P < 0.0001$ (****) using the Kruskal–Wallis test.

Next, we explored the sex- and age-related differences in ACE2 expression. Among all donors, males had slightly higher ACE2 expression levels than females, but the difference was insignificant (Fig. 1D, upper left plot). Next, among all subgroups defined by medical condition, ACE2 expression levels were upregulated in males compared with females, but a statistically significant difference was observed only in the healthy subgroup (Fig. 1D, upper plot). Additionally, the ACE2

expression levels were significantly increased with age only in the healthy subgroup (Fig. 1E, upper plot).

In parallel, we analyzed the expression of dipeptidyl peptidase 4 (DPP4), the host cell receptor for Middle East respiratory syndrome coronavirus (MERS-CoV) [13]. Compared to the healthy donors, the DPP4 expression levels were significantly lower in the sick donors without ventilator support but remarkably higher in the sick donors with ventilator support (Fig. 1C, right plot). Moreover, DPP4 expression showed no significant sex-

related difference (Fig. 1D, lower plot). The DPP4 expression levels were significantly decreased with age among all donors and the subgroup of sick donors with ventilator support (Fig. 1E, lower plot). Thus, ACE2 and DPP4 had different expression patterns across demographics, indicating different tissue tropisms between SARS-CoV-2 and MERS-CoV.

The above results revealed upregulated ACE2 expression levels in sick, male, and older populations before SARS-CoV-2 infection. It has been reported that the demographic characteristics of pre-existing diseases, male sex, and older age are risk factors for severe COVID-19 [1]. Thus, the expression of ACE2 coincides with the clinical manifestations of COVID-19 across demographics, suggesting a hypothesis centered on “ACE2” for COVID-19 risk.

Mendelian randomization study of ACE2 and COVID-19 susceptibility and severity

To better understand the relationship between ACE2 and COVID-19 risk, we conducted a Mendelian randomization (MR) study. MR, built on Mendel’s second law, is a commonly used genetics technique that implicates the causality of exposure for a disease outcome [14]. MR analysis uses genetic variants as instrumental variables. Since genetic variants are randomly assorted during meiosis and typically unassociated with potential confounders, MR results are less prone to the bias of confounding and reverse causality [14].

Here, we conducted MR to infer the causal effect of ACE2 on COVID-19 manifestations, including disease susceptibility and severity. Fig. 2A describes the study design, method and material, and the results of our MR study. Two-sample MR analysis was conducted using TwoSampleMR software (<https://mrcieu.github.io/TwoSampleMR/>), as previously reported [15]. We selected conditionally independent cis-expression quantitative trait loci (cis-eQTLs; p -value $< 5 \times 10^{-4}$) of ACE2 and DPP4 from the GTEx v8 dataset [16]. Linkage disequilibrium (LD) clumping was performed for pruning correlated cis-eQTLs ($r^2 < 0.001$) using the PLINK clumping method [17, 18]. Next, we tested the cis-eQTL proposed instruments of ACE2 and DPP4 against multiple genome-wide association study (GWAS) data of COVID-19 and influenza from the COVID-19 Host Genetics Initiative [19] and GWAS Catalog [20]. The GWAS cohorts were categorized into disease susceptibility and severity (Fig. 2B). COVID-19 severity was indicated by hospitalization, while influenza severity was indicated by influenza pneumonia.

MR estimates were calculated using inverse-variance weighting for instruments that contained more than one cis-eQTL and Wald ratio for instruments with

one cis-eQTL. All the significant MR estimates (β -coefficients) of ACE2 were positive for both the susceptibility and severity of COVID-19 (Fig. 2B), indicating that upregulated ACE2 expression is associated with an increased risk of COVID-19. In contrast, all the MR β -coefficients for DPP4 exposure and influenza outcome were insignificant. Thus, our MR study discovered that ACE2 played a causal role in COVID-19 susceptibility and severity, indicating ACE2 as a potential therapeutic target of COVID-19.

Overlapping virus-permissive cells between MERS-CoV and SARS-CoV-2

The cellular expression of host cell receptors is a prerequisite of virus infection, and thus forms the basis of the understanding of on which cells a virus can exert its pathogenic effects. Based on this theory, previous studies have demonstrated the feasibility of mapping virus-permissive cells using single-cell RNA-seq data (scRNA-seq) [7, 21, 22]. Here, by analyzing the scRNA-seq data of 57,020 cells from healthy human lung tissues from a previous study [23], we investigated the identities of ACE2⁺ and DPP4⁺ cells (Fig. 3A-B). According to the original study, we retained the information of cell clustering and cell type annotations (Fig. 3A) [23]. Compared to ACE2, DPP4 was expressed by a broader distribution of cell types (Fig. 3B-C). The most dominant cell types for ACE2⁺ cells were alveolar type 2 (AT2) and fibroblast cells (Fig. 3C). It is worth noting that 77.5% of ACE2⁺ cells and 34.8% of DPP4⁺ cells were AT2 cells, albeit at much lower proportions, 12.7% of ACE2⁺ cells and 2.2% of DPP4⁺ cells were fibroblast cells (Fig. 3C). These results indicate the substantial overlap of virus-permissive cells between MERS-CoV and SARS-CoV-2.

Relative abundances of pulmonary AT2 and fibroblast cells across demographics

Due to the dominant proportion of AT2 and fibroblast cells in ACE2⁺ cells, we investigated whether the relative abundances of AT2 and fibroblast cells were related to medical condition, sex, and age. In the scRNA-seq data of lung tissues, surfactant protein C (SFTPC) and collagen type I alpha 2 chain (COL1A2), the well-acknowledged cell markers of AT2 and fibroblast cells, were expressed dominantly by AT2 and fibroblast cells (Fig. 3D). Next, we measured the relative abundances of AT2 and fibroblast cells through the expression of SFTPC and COL1A2 in lung tissues. We observed quantitative fluctuations in AT2 and fibroblast cell abundances in sick, male, and older populations (Fig. 3E-G), indicating the importance of investigating both ACE2 expression and ACE2-expressing cell types.

To capture the above-stated changes in the key features of viral tropism for SARS-CoV-2, we used radar charts to graphically represent the feature variables of SARS-CoV-2 tropism, including ACE2 expression and the most dominant SARS-CoV-2-permissive cells, AT2 and fibroblast cells (Fig. 4A). The radar charts successfully capture the distinct patterns of SARS-CoV-2 tropism across the sick, male, and older populations (Fig. 4B-D). Compared to the healthy subgroup, the ACE2 expression levels were increased in the sick subgroups and increased even more in the sick subgroup with ventilator usage (Fig. 4B). Of note, ventilator usage was

associated with an inverse relationship in the abundances of AT2 and fibroblast cells. Next, we found that males had increased fibroblast cell abundances and higher ACE2 expression levels than females. These findings were concordant among all donors as well as all three subgroups defined by medical condition (Fig. 4C). In contrast, we observed no concordant age-related differences in the ACE2-expressing cells, whereas consistently higher ACE2 expression levels were observed in the older subgroup (≥ 50 years) than in the younger subgroup (< 50 years) (Fig. 4D).

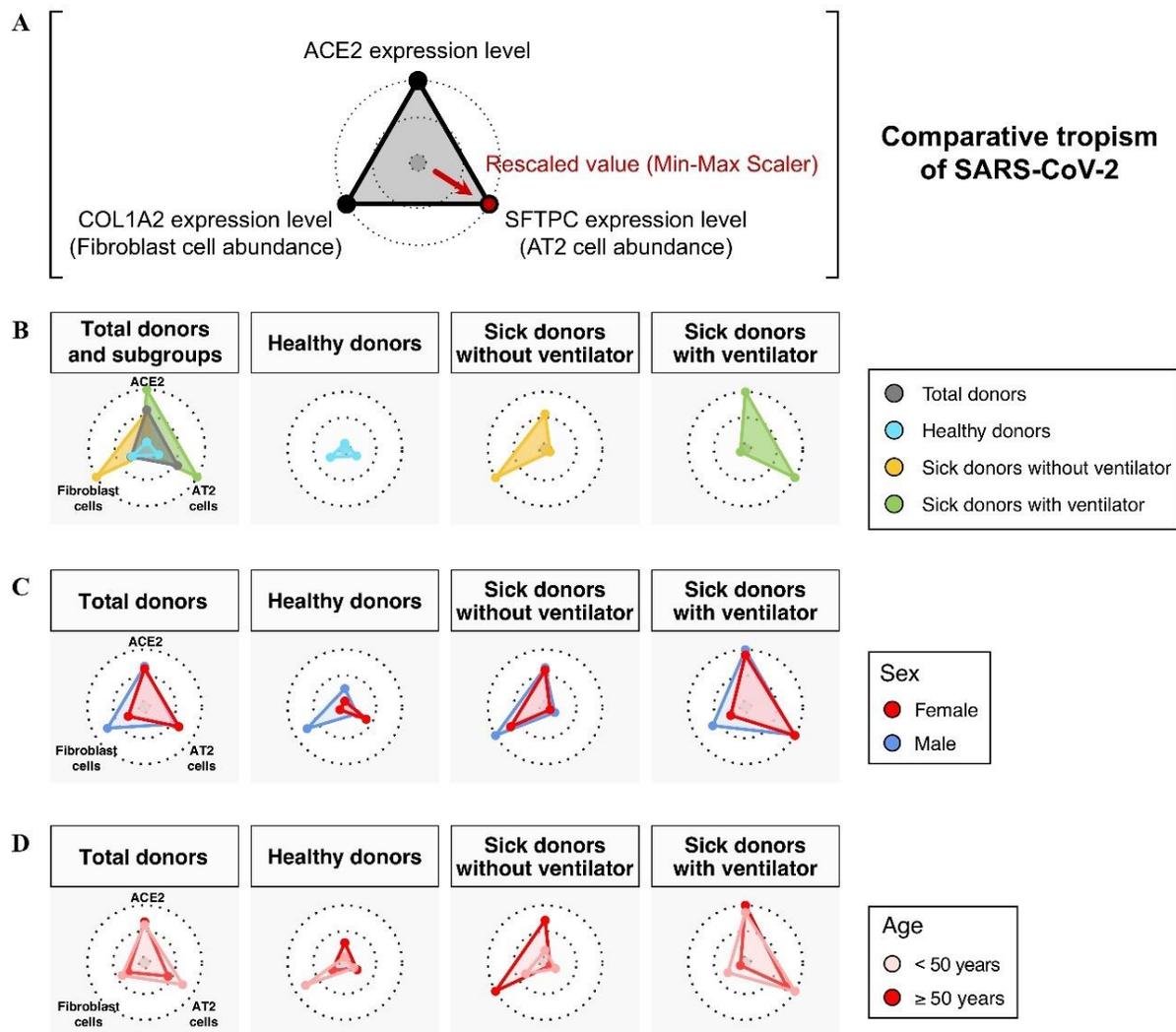


Figure 4. Key features of viral tropism for SARS-CoV-2 across demographics. (A) Comparative analysis of viral tropism for SARS-CoV-2 through phenotyping the ACE2 expression level and relative abundances of AT2 and fibroblast cells. The relative abundances of AT2 and fibroblast cells were measured by the expression of their cell markers, SFTPC and COL1A2. The average expression levels of ACE2, SFTPC, and COL1A2 were scaled using the min-max scaler (MMS) method for radar charts. (B-D) Radar charts show the feature variables of SARS-CoV-2 tropism (ACE2 expression level, relative abundances of AT2 and fibroblast cells) in total donors and the subgroups defined by medical condition (B), sex (C), and age (D).

Overlapping cellular tropism in human respiratory viruses

The global view of a comparative tropism across human viruses is still lacking. This hole in our knowledge is due in part to the technical difficulty of investigating comparative tropism systematically. Recently, we developed a bioinformatic framework for mapping viral tropism through the expression of host cell receptors [7]. Here, by applying our bioinformatic framework to the scRNA-seq data of human lung tissues from the previous study [23], we conducted a comparative tropism analysis across different human respiratory viruses. We characterized the virus-permissive cells in 14 human respiratory viruses, including SARS-CoV, SARS-CoV-2, MERS-CoV, rhinovirus A, rhinovirus B, rhinovirus C, reovirus, adenovirus, coxsackievirus, poliovirus 1, Old World arenavirus, New World arenavirus, Hendra and Nipah viruses, and measles virus (Fig. 5A). These respiratory viruses employ various host cell receptors, including ACE2 [3, 8], DPP4 [13], ICAM1 [24, 25], LDLR [26], CDHR3 [27], F11R [28], RTN4R [29], CXADR [30], CD55 [31, 32], PVR [33], DAG1 [34], TFRC [35], EFNB2 [36], SLAMF1 [37], and CD46 [38]. Using scRNA-seq data from healthy human lung tissues, we analyzed the percentage of virus-permissive cells with the expression of different host cell receptors (Fig. 5B-C). Notably, all these respiratory viruses showed overlapping cellular tropisms in AT2 and fibroblast cells. Moreover, AT2 and fibroblast cells were the first and third most dominant cell types, respectively (Fig. 5D).

reovirus, adenovirus, coxsackievirus, poliovirus 1, Old World arenavirus, New World arenavirus, Hendra and Nipah viruses, and measles virus (Fig. 5A). These respiratory viruses employ various host cell receptors, including ACE2 [3, 8], DPP4 [13], ICAM1 [24, 25], LDLR [26], CDHR3 [27], F11R [28], RTN4R [29], CXADR [30], CD55 [31, 32], PVR [33], DAG1 [34], TFRC [35], EFNB2 [36], SLAMF1 [37], and CD46 [38]. Using scRNA-seq data from healthy human lung tissues, we analyzed the percentage of virus-permissive cells with the expression of different host cell receptors (Fig. 5B-C). Notably, all these respiratory viruses showed overlapping cellular tropisms in AT2 and fibroblast cells. Moreover, AT2 and fibroblast cells were the first and third most dominant cell types, respectively (Fig. 5D).

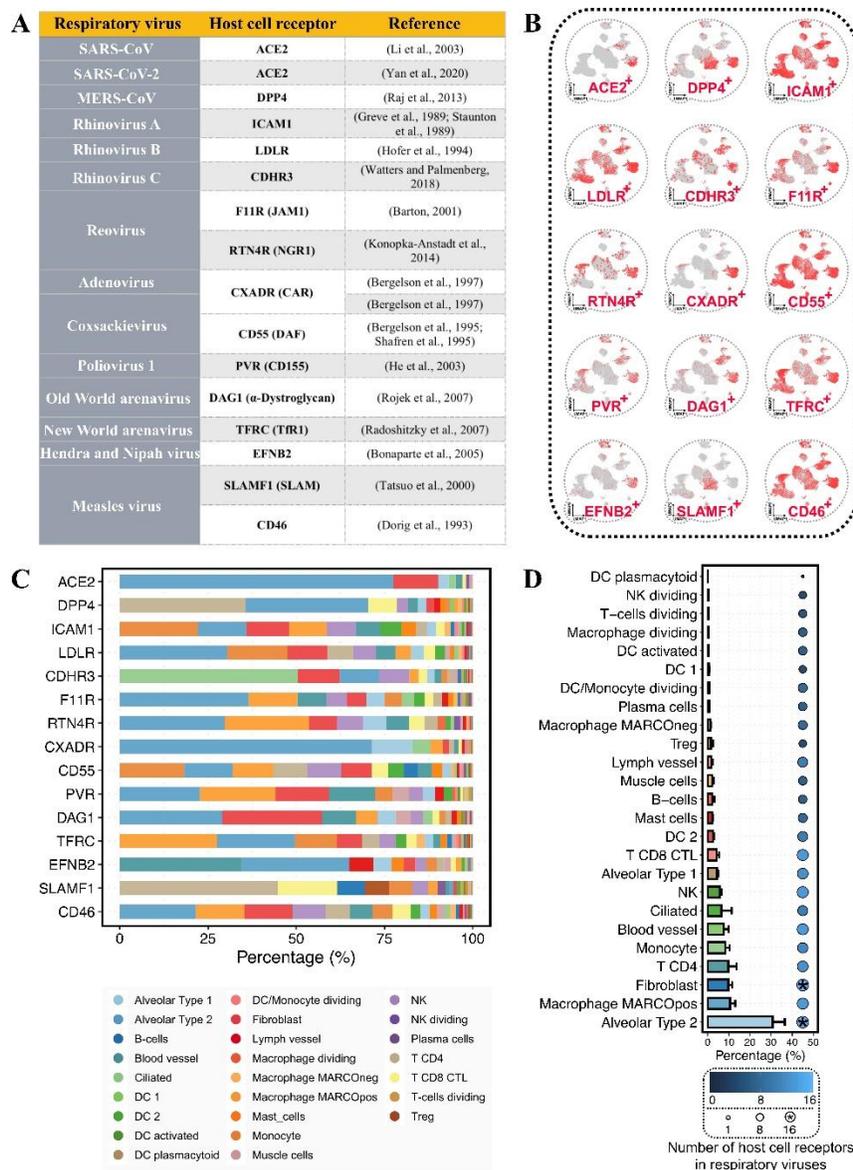


Figure 5. Overlapping cellular tropism in human respiratory viruses. (A) Respiratory viruses and the host cell receptors. **(B)** UMAP plots show the pulmonary expression of host cell receptors from human respiratory viruses. **(C)** The percentages of virus-permissive cell types with the expression of host cell receptors. **(D)** Bar plot shows the average percentages of virus-permissive cell types for respiratory viruses. The error bar indicates the standard error. The dot color and size represent the number of viral-specific host cell receptors expressed by the cell type. The "*" shape represents the cell type expressing all the host cell receptors.

Comparative cellular tropism between human respiratory and non-respiratory viruses

Next, we investigated the comparative cellular tropism between human respiratory and non-respiratory viruses. We analyzed six non-respiratory viruses for comparison, including human immunodeficiency virus (HIV), Zika virus, Ebola virus, human T cell leukemia virus (HTLV), Epstein–Barr virus (EBV), and rabies virus (RABV) (Fig. 6A). The corresponding host cell receptors included CD4 [39], AXL [40], HAVCR1 [41], SLC2A1 [42], CR2 [43],

GRM2 [44], CHRNA1 [45], NCAM1 [46], and NGFR [47]. By analyzing the expression of host cell receptors, we calculated the percentage of virus-permissive cells for human non-respiratory viruses (Fig. 6B–C). The average proportion of AT2 cells was 22.4-fold higher in respiratory viruses than in non-respiratory viruses (p -value < 0.001 , Fig. 6D), but no significant difference was observed in fibroblast cells. This finding implies the important role of virus-permissive AT2 cells in respiratory rather than non-respiratory viruses.

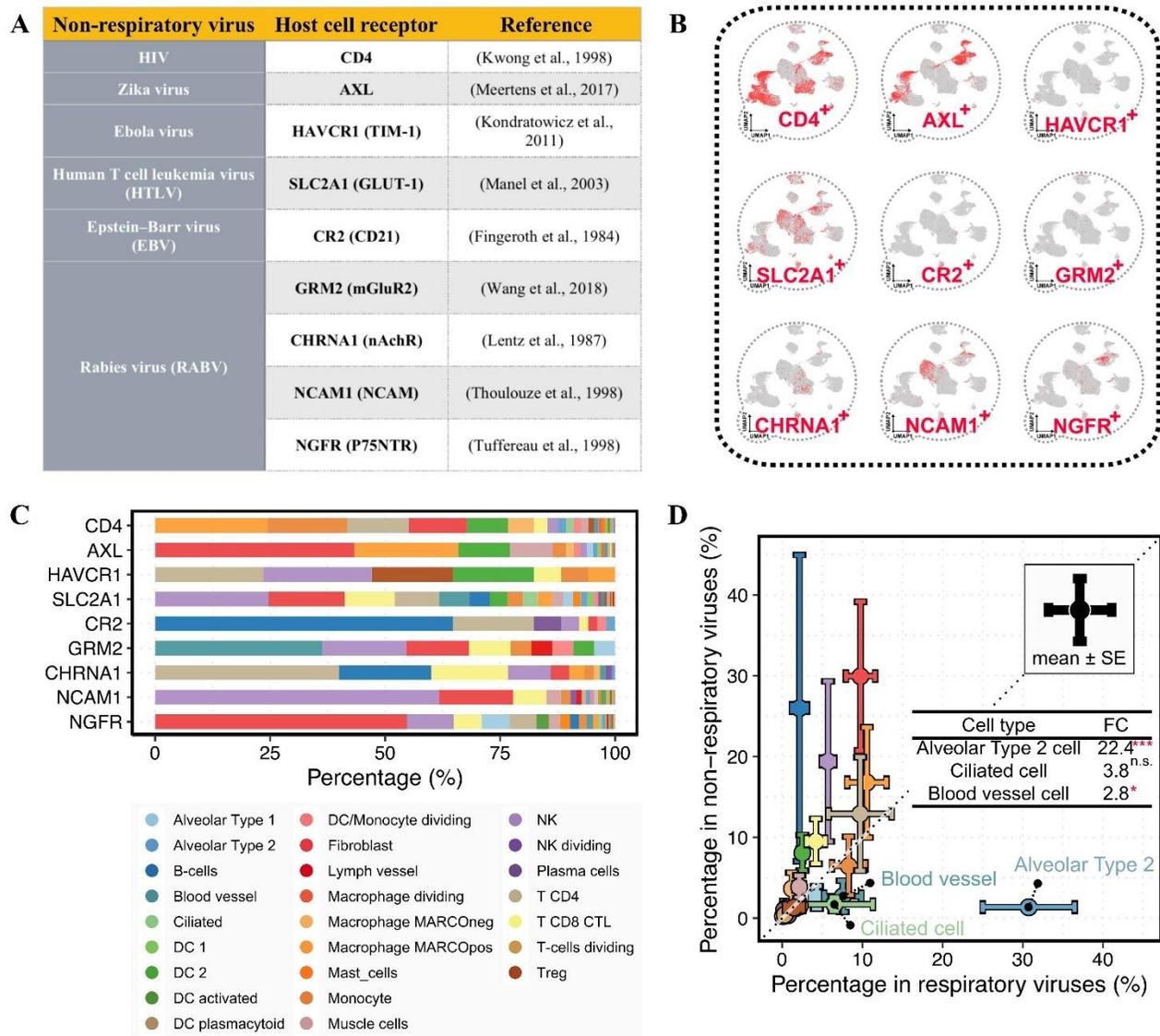


Figure 6. Comparative cellular tropism between respiratory and non-respiratory viruses. (A) Non-respiratory viruses and the host cell receptors. (B) UMAP plots show the pulmonary expression of host cell receptors from human non-respiratory viruses. (C) The percentages of virus-permissive cell types with the expression of host cell receptors. (D) Comparison between the average percentages of virus-permissive cell types for respiratory and non-respiratory viruses. The error bar indicates the standard error. The dashed line represents the 1.0 slope of complete concordance. The result was considered statistically significant when $P < 0.05$ (*), $P < 0.01$ (**), $P < 0.001$ (***), and $P < 0.0001$ (****) using the Kruskal–Wallis test. FC represents fold change.

Relationships between heterogeneous ACE2 expression, COVID-19 manifestations, and comparative tropism of SARS-CoV-2 across demographics

Due to the causal role of ACE2 in COVID-19, sick, male, and older populations with higher ACE2 expression levels were more vulnerable to SARS-CoV-2 infection, leading to higher COVID-19 susceptibility and severity among these subgroups (Fig. 7). Therefore, ACE2 is centrally associated with the demographic risk factors for COVID-19. Next, comparative tropism analysis revealed the demographic features of susceptible cell types for SARS-CoV-2 infection, indicating fundamental changes in the virus-host interaction across demographics.

A previous study reported the causal role of ACE2 in COVID-19 severity through MR analysis [15]. In this study, we discovered that ACE2 had a significant causal role in not only the severity but also the susceptibility to COVID-19. It is worth noting that, although both previous and our MR studies show strong evidence that higher ACE2 expression increases the risk of COVID-19, both studies have a limitation in that no pulmonary cis-eQTL instrument was available for analyzing ACE2 in lung tissues. Nevertheless, both previous and our MR results strongly suggest ACE2 as one of the most promising actionable and druggable targets for the early treatment of COVID-19.

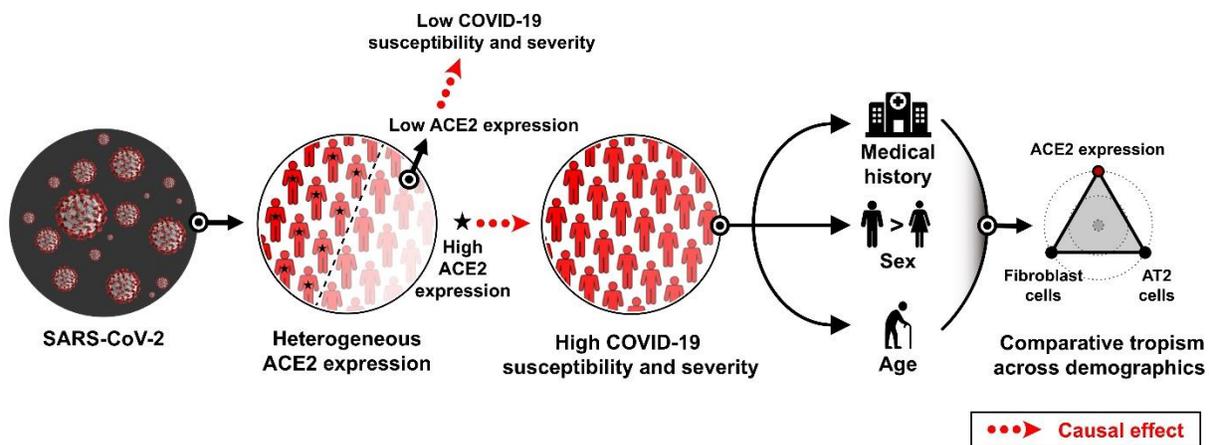


Figure 7. Causal role of ACE2 in COVID-19 susceptibility and severity, and comparative tropism of SARS-CoV-2 across demographics. Sick, male, and older populations have higher ACE2 expression levels, leading to increased COVID-19 susceptibility and severity. The current definitions of severe COVID-19 populations, including sick, male, and older populations, might reflect the subphenotypes of SARS-CoV-2 tropism across demographics. The dashed arrow shows the causal effect of ACE2 on COVID-19 susceptibility and severity.

ACE2 is the prerequisite of viral infection for both SARS-CoV [8] and SARS-CoV-2 [3]. This study found that a substantial proportion (77.5%) of pulmonary ACE2⁺ cells were AT2 cells. AT2 cells, rather than AT1 cells, are preferentially infected by both SARS-CoV and SARS-CoV-2 [5, 48]. Moreover, in a postmortem pathological study of a ready-for-discharge COVID-19 patient, SARS-CoV-2 virus nucleic acid was negative in nasopharyngeal swabs but positive in the lungs, and electron microscopic testing clearly showed coronavirus particles in AT2 cells, implying that AT2 cells retain SARS-CoV-2 infection for prolonged periods [49]. Thus, ACE2-expressing cells successfully indicate the infectable cell types, and the proportional distribution of those virus-permissive cells has profound implications for the preference of viral infection.

In this study, we discovered that donors on mechanical ventilation had the most remarkably increased

ACE2 expression levels and AT2 cell abundances. It has been reported that mechanical ventilation causes the proliferation of AT2 cells [50], which might partly explain the increased abundance of AT2 cells. Recently, there has been a strong argument regarding whether ventilator support should be employed early in COVID-19 patients to prevent disease progression [51]. The major concern for this argument is that vigorous mechanical ventilation can rapidly lead to ventilator-induced lung injury. It is worth noting that AT2 cells play a critical role in decreasing lung injury through controlling the immune response and repairing the injured lung [52]. However, AT2 cells are the primary target of SARS-CoV-2 infection. COVID-19 patients with SARS-CoV-2-infected and impaired AT2 cells might be particularly vulnerable to ventilator-induced lung injury. Therefore, ventilator support in patients with early COVID-19 should be used cautiously, and further clinical study on this issue is required.

It has been reported that the influenza virus preferentially infects AT2 cells [53], and AT2 cells maintain the replication of influenza virus for more prolonged periods [54]. This evidence indicates the possibility of coinfection between SARS-CoV-2 and influenza virus at the cellular level. A previous study found that the coinfection rate of SARS-CoV-2 and influenza virus was as high as 57.3% during the COVID-19 outbreak period in Wuhan [55]. Moreover, coinfection with influenza virus significantly enhances SARS-CoV-2 infection, leading to more severe lung damage in mice [56]. Thus, when the current COVID-19 pandemic merges with the flu season, the coinfection of influenza virus and SARS-CoV-2 might cause a more severe threat to public health.

Furthermore, AT2 cells are also the preferential host cells for other respiratory viruses. Through the comparative tropism analysis of 14 respiratory viruses, we found that, first, albeit using different host cell receptors, all the respiratory viruses showed overlapping cellular tropism to the AT2 cells; second, among the heterogeneous host cells targeted by respiratory viruses, the AT2 cells were the most proportionally dominant cell type; and third, the non-respiratory viruses had extremely limited capability to target the AT2 cells. These findings indicate the essential role of AT2 cells in viruses transmitted through the respiratory route.

Here, we discussed the probable reasons why respiratory viruses preferentially targeted AT2 cells. First, AT2 cells are located on the gas exchange surface of the alveolar epithelium, which is the critical interface exposed to airborne viruses [57]. Second, as highly metabolic secretory cells, AT2 cells have abundant cytoplasm and an extensive endoplasmic reticulum and Golgi apparatus, which provide important machinery for viral replication; additionally, SARS-CoV-2 leverages the secretory pathway to transport viral particles to the plasma membrane and then is secreted from the cell [58]. Thus, we assume that the physiological location and cell traits might jointly contribute to the preferential role of AT2 cells for respiratory viruses.

SARS-CoV-2 is considered to originate from bats, and its zoonotic spillover to the human population caused the COVID-19 outbreak [59]. In fact, the majority of human viruses are from zoonotic. In the process of zoonotic spillover, the virus population encounters strong natural selection [60], and only limited viruses accomplish interspecies transmission. For overcoming the species barrier and accomplishing interspecies transmission, there are two essential factors: (i) the binding of viral entry proteins to the compatible host cell receptors mediates successful viral entry into host cells; (ii) the infected host cells support viral replication and the release of viral particles. Thus, it is well-acknowledged that the

binding of viral entry proteins to the compatible host cell receptors is under strong natural selection pressure. Moreover, certain cell types, such as AT2 cells, appear to be more favorable for sustaining viral replication [5, 48, 54]. Consequently, by analyzing the cell-type-specific expression of host cell receptors, certain virus-permissive cell types appeared more often than others, which might reveal the footprint of natural selection that shapes the virus population.

Taken together, the results of this study revealed demographic heterogeneity in the key features of SARS-CoV-2 tropism. This finding indicates that the current definition of vulnerable COVID-19 populations, including sick, male, and older populations, might reflect distinct subphenotypes of SARS-CoV-2 tropism. We believe this finding is essential for improving our understanding of the pathogenic mechanisms of COVID-19. Moreover, we demonstrated that ACE2 expression, the central variable of SARS-CoV-2 tropism, has a causal effect on COVID-19 susceptibility and severity, suggesting the clinical implications of ACE2 in COVID-19 treatment.

To the best of our knowledge, this is the first study revealing the power of comparative tropism analysis for mapping the host cell preferences across multiple viruses. Based on this analysis, we discovered remarkably overlapping and enriched tropism in AT2 cells across different respiratory viruses. These findings have important implications for sociovirology [61], implying conflict, cooperation, and communication among respiratory viruses, suggesting the possibility of coinfection with SARS-CoV-2 and other respiratory viruses. Moreover, with increasing insight into emerging viruses, comparative tropism analysis will help in the understanding of zoonotic spillover events from a novel perspective.

Acknowledgments

M.Z. conceived the project, developed the method, conducted data analysis, and wrote the manuscript. We thank the Genotype Tissue Expression (GTEx) Consortium for the RNA-seq data of human lung tissues.

Ethical Approval and Consent to participate

There is no direct involvement of human subjects in this study. All the data are from prior studies. Therefore, ethical oversight and patient consent were not handled in this study.

Conflicts of interest

The authors disclose no potential conflicts of interest.

References

- [1] Zheng M, Song L (2021). Shift in the Distributions of Pre-existing Medical Condition, Gender and Age across Different COVID-19 Outcomes. *Aging and disease*, 12:327-329.
- [2] Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, *et al.* (2020). A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med*, 382:727-733.
- [3] Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q (2020). Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science*, 367:1444-1448.
- [4] Jia HP, Look DC, Shi L, Hickey M, Pewe L, Netland J, *et al.* (2005). ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia. *J Virol*, 79:14614-14621.
- [5] Hou YJ, Okuda K, Edwards CE, Martinez DR, Asakura T, Dinnon KH, 3rd, *et al.* (2020). SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. *Cell*, 182:429-446 e414.
- [6] McCray PB, Jr., Pewe L, Wohlford-Lenane C, Hickey M, Manzel L, Shi L, *et al.* (2007). Lethal infection of K18-hACE2 mice infected with severe acute respiratory syndrome coronavirus. *J Virol*, 81:813-821.
- [7] Zheng M (2021). Cellular Tropism of SARS-CoV-2 across Human Tissues and Age-related Expression of ACE2 and TMPRSS2 in Immune-inflammatory Stromal Cells. *Aging and disease*, 12:718-725.
- [8] Li W, Moore MJ, Vasilieva N, Sui J, Wong SK, Berne MA, *et al.* (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature*, 426:450-454.
- [9] Peiris JS, Guan Y, Yuen KY (2004). Severe acute respiratory syndrome. *Nat Med*, 10:S88-97.
- [10] Huang C, Wang Y, Li X, Ren L, Zhao J, Hu Y, *et al.* (2020). Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet*, 395:497-506.
- [11] Hamming I, Timens W, Bulthuis ML, Lely AT, Navis G, van Goor H (2004). Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. *J Pathol*, 203:631-637.
- [12] Consortium GT (2015). Human genomics. The Genotype-Tissue Expression (GTEx) pilot analysis: multitissue gene regulation in humans. *Science*, 348:648-660.
- [13] Raj VS, Mou H, Smits SL, Dekkers DH, Muller MA, Dijkman R, *et al.* (2013). Dipeptidyl peptidase 4 is a functional receptor for the emerging human coronavirus-EMC. *Nature*, 495:251-254.
- [14] Emdin CA, Khera AV, Kathiresan S (2017). Mendelian Randomization. *JAMA*, 318:1925-1926.
- [15] Gaziano L, Giambartolomei C, Pereira AC, Gaulton A, Posner DC, Swanson SA, *et al.* (2021). Actionable druggable genome-wide Mendelian randomization identifies repurposing opportunities for COVID-19. *Nat Med*, 27:668-676.
- [16] Consortium GT (2013). The Genotype-Tissue Expression (GTEx) project. *Nat Genet*, 45:580-585.
- [17] Chang CC, Chow CC, Tellier LC, Vattikuti S, Purcell SM, Lee JJ (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience*, 4:7.
- [18] Gaunt TR, Rodriguez S, Day IN (2007). Cubic exact solutions for the estimation of pairwise haplotype frequencies: implications for linkage disequilibrium analyses and a web tool 'CubeX'. *BMC Bioinformatics*, 8:428.
- [19] Initiative C-HG (2020). The COVID-19 Host Genetics Initiative, a global initiative to elucidate the role of host genetic factors in susceptibility and severity of the SARS-CoV-2 virus pandemic. *Eur J Hum Genet*, 28:715-718.
- [20] Buniello A, MacArthur JAL, Cerezo M, Harris LW, Hayhurst J, Malangone C, *et al.* (2019). The NHGRI-EBI GWAS Catalog of published genome-wide association studies, targeted arrays and summary statistics 2019. *Nucleic Acids Res*, 47:D1005-D1012.
- [21] Sungnak W, Huang N, Becavin C, Berg M, Queen R, Litvinukova M, *et al.* (2020). SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat Med*, 26:681-687.
- [22] Ziegler CGK, Allon SJ, Nyquist SK, Mbanjo IM, Miao VN, Tzouanas CN, *et al.* (2020). SARS-CoV-2 Receptor ACE2 Is an Interferon-Stimulated Gene in Human Airway Epithelial Cells and Is Detected in Specific Cell Subsets across Tissues. *Cell*, 181:1016-1035 e1019.
- [23] Madisson E, Wilbrey-Clark A, Miragaia RJ, Saeb-Parsy K, Mahbubani KT, Georgakopoulos N, *et al.* (2019). scRNA-seq assessment of the human lung, spleen, and esophagus tissue stability after cold preservation. *Genome Biol*, 21:1.
- [24] Greve JM, Davis G, Meyer AM, Forte CP, Yost SC, Marlor CW, *et al.* (1989). The major human rhinovirus receptor is ICAM-1. *Cell*, 56:839-847.
- [25] Staunton DE, Merluzzi VJ, Rothlein R, Barton R, Marlin SD, Springer TA (1989). A cell adhesion molecule, ICAM-1, is the major surface receptor for rhinoviruses. *Cell*, 56:849-853.
- [26] Hofer F, Gruenberger M, Kowalski H, Machat H, Huettinger M, Kuechler E, *et al.* (1994). Members of the low density lipoprotein receptor family mediate cell entry of a minor-group common cold virus. *Proc Natl Acad Sci U S A*, 91:1839-1842.
- [27] Watters K, Palmenberg AC (2018). CDHR3 extracellular domains EC1-3 mediate rhinovirus C interaction with cells and as recombinant derivatives, are inhibitory to virus infection. *PLoS Pathog*, 14:e1007477.
- [28] Barton E (2001). Forrest JC, Connolly JL, Chappell JD, Liu Y, Schnell FJ, Nusrat A, Parkos CA, and Dermody TS. Junction adhesion molecule is a receptor for reovirus. *Cell*, 104:441-451.

- [29] Konopka-Anstadt JL, Mainou BA, Sutherland DM, Sekine Y, Strittmatter SM, Dermody TS (2014). The Nogo receptor NgR1 mediates infection by mammalian reovirus. *Cell Host Microbe*, 15:681-691.
- [30] Bergelson JM, Cunningham JA, Droguett G, Kurt-Jones EA, Krithivas A, Hong JS, *et al.* (1997). Isolation of a common receptor for Coxsackie B viruses and adenoviruses 2 and 5. *Science*, 275:1320-1323.
- [31] Bergelson JM, Mohanty JG, Crowell RL, St John NF, Lublin DM, Finberg RW (1995). Coxsackievirus B3 adapted to growth in RD cells binds to decay-accelerating factor (CD55). *J Virol*, 69:1903-1906.
- [32] Shafren DR, Bates RC, Agrez MV, Herd RL, Burns GF, Barry RD (1995). Coxsackieviruses B1, B3, and B5 use decay accelerating factor as a receptor for cell attachment. *Journal of virology*, 69:3873-3877.
- [33] He Y, Mueller S, Chipman PR, Bator CM, Peng X, Bowman VD, *et al.* (2003). Complexes of poliovirus serotypes with their common cellular receptor, CD155. *J Virol*, 77:4827-4835.
- [34] Rojek JM, Campbell KP, Oldstone MB, Kunz S (2007). Old World arenavirus infection interferes with the expression of functional alpha-dystroglycan in the host cell. *Mol Biol Cell*, 18:4493-4507.
- [35] Radoshitzky SR, Abraham J, Spiropoulou CF, Kuhn JH, Nguyen D, Li W, *et al.* (2007). Transferrin receptor 1 is a cellular receptor for New World haemorrhagic fever arenaviruses. *Nature*, 446:92-96.
- [36] Bonaparte MI, Dimitrov AS, Bossart KN, Cramer G, Mungall BA, Bishop KA, *et al.* (2005). Ephrin-B2 ligand is a functional receptor for Hendra virus and Nipah virus. *Proc Natl Acad Sci U S A*, 102:10652-10657.
- [37] Tatsuo H, Ono N, Tanaka K, Yanagi Y (2000). SLAM (CDw150) is a cellular receptor for measles virus. *Nature*, 406:893-897.
- [38] Dorig RE, Marcil A, Chopra A, Richardson CD (1993). The human CD46 molecule is a receptor for measles virus (Edmonston strain). *Cell*, 75:295-305.
- [39] Kwong PD, Wyatt R, Robinson J, Sweet RW, Sodroski J, Hendrickson WA (1998). Structure of an HIV gp120 envelope glycoprotein in complex with the CD4 receptor and a neutralizing human antibody. *Nature*, 393:648-659.
- [40] Meertens L, Labeau A, Dejarnac O, Cipriani S, Sinigaglia L, Bonnet-Madin L, *et al.* (2017). Axl Mediates ZIKA Virus Entry in Human Glial Cells and Modulates Innate Immune Responses. *Cell Rep*, 18:324-333.
- [41] Kondratowicz AS, Lennemann NJ, Sinn PL, Davey RA, Hunt CL, Moller-Tank S, *et al.* (2011). T-cell immunoglobulin and mucin domain 1 (TIM-1) is a receptor for Zaire Ebolavirus and Lake Victoria Marburgvirus. *Proc Natl Acad Sci U S A*, 108:8426-8431.
- [42] Manel N, Kim FJ, Kinet S, Taylor N, Sitbon M, Battini J-L (2003). The Ubiquitous Glucose Transporter GLUT-1 Is a Receptor for HTLV. *Cell*, 115:449-459.
- [43] Fingerroth JD, Weis JJ, Tedder TF, Strominger JL, Biro PA, Fearon DT (1984). Epstein-Barr virus receptor of human B lymphocytes is the C3d receptor CR2. *Proc Natl Acad Sci U S A*, 81:4510-4514.
- [44] Wang J, Wang Z, Liu R, Shuai L, Wang X, Luo J, *et al.* (2018). Metabotropic glutamate receptor subtype 2 is a cellular receptor for rabies virus. *PLoS Pathog*, 14:e1007189.
- [45] Lentz TL, Hawrot E, Wilson PT (1987). Synthetic peptides corresponding to sequences of snake venom neurotoxins and rabies virus glycoprotein bind to the nicotinic acetylcholine receptor. *Proteins*, 2:298-307.
- [46] Thoulouze MI, Lafage M, Schachner M, Hartmann U, Cremer H, Lafon M (1998). The neural cell adhesion molecule is a receptor for rabies virus. *J Virol*, 72:7181-7190.
- [47] Tuffereau C, Benejean J, Blondel D, Kieffer B, Flamand A (1998). Low-affinity nerve-growth factor receptor (P75NTR) can serve as a receptor for rabies virus. *EMBO J*, 17:7250-7259.
- [48] Mossel EC, Wang J, Jeffers S, Edeen KE, Wang S, Cosgrove GP, *et al.* (2008). SARS-CoV replicates in primary human alveolar type II cell cultures but not in type I-like cells. *Virology*, 372:127-135.
- [49] Yao XH, He ZC, Li TY, Zhang HR, Wang Y, Mou H, *et al.* (2020). Pathological evidence for residual SARS-CoV-2 in pulmonary tissues of a ready-for-discharge patient. *Cell Res*, 30:541-543.
- [50] Tsuno K, Miura K, Takeya M, Kolobow T, Morioka T (1991). Histopathologic pulmonary changes from mechanical ventilation at high peak airway pressures. *Am Rev Respir Dis*, 143:1115-1120.
- [51] Marini JJ, Gattinoni L (2020). Management of COVID-19 Respiratory Distress. *JAMA*, 323:2329-2330.
- [52] Olajuyin AM, Zhang X, Ji HL (2019). Alveolar type 2 progenitor cells for lung injury repair. *Cell Death Discov*, 5:63.
- [53] Weinheimer VK, Becher A, Tonnies M, Holland G, Knepper J, Bauer TT, *et al.* (2012). Influenza A viruses target type II pneumocytes in the human lung. *J Infect Dis*, 206:1685-1694.
- [54] Yu WC, Chan RW, Wang J, Travanty EA, Nicholls JM, Peiris JS, *et al.* (2011). Viral replication and innate host responses in primary human alveolar epithelial cells and alveolar macrophages infected with influenza H5N1 and H1N1 viruses. *J Virol*, 85:6844-6855.
- [55] Yue H, Zhang M, Xing L, Wang K, Rao X, Liu H, *et al.* (2020). The epidemiology and clinical characteristics of co-infection of SARS-CoV-2 and influenza viruses in patients during COVID-19 outbreak. *J Med Virol*, 92:2870-2873.
- [56] Bai L, Zhao Y, Dong J, Liang S, Guo M, Liu X, *et al.* (2021). Coinfection with influenza A virus enhances SARS-CoV-2 infectivity. *Cell Res*, 31:395-403.
- [57] Mason RJ (2006). Biology of alveolar type II cells. *Respirology*, 11 Suppl:S12-15.
- [58] Sicari D, Chatziioannou A, Koutsandreas T, Sitia R, Chevet E (2020). Role of the early secretory pathway in SARS-CoV-2 infection. *J Cell Biol*, 219.

- [59] Zhou P, Shi ZL (2021). SARS-CoV-2 spillover events. *Science*, 371:120-122.
- [60] Borderia AV, Isakov O, Moratorio G, Henningsson R, Aguera-Gonzalez S, Organtini L, *et al.* (2015). Group Selection and Contribution of Minority Variants during Virus Adaptation Determines Virus Fitness and Phenotype. *PLoS Pathog*, 11:e1004838.
- [61] Diaz-Munoz SL, Sanjuan R, West S (2017). Sociovirology: Conflict, Cooperation, and Communication among Viruses. *Cell Host Microbe*, 22:437-441.