






OPEN

## Comparative transcriptomic analysis of SARS-CoV-2 infected cell model systems reveals differential innate immune responses

Guihua Sun<sup>1</sup>, Qi Cui<sup>2</sup>, Gustavo Garcia Jr.<sup>3,4</sup>, Cheng Wang<sup>2</sup>, Mingzi Zhang<sup>2</sup>, Vaithilingaraja Arumugaswami<sup>3,4</sup>, Arthur D. Riggs<sup>1</sup> & Yanhong Shi<sup>2</sup>

The transcriptome of SARS-CoV-2-infected cells that reflects the interplay between host and virus has provided valuable insights into mechanisms underlying SARS-CoV-2 infection and COVID-19 disease progression. In this study, we show that SARS-CoV-2 can establish a robust infection in HEK293T cells that overexpress human angiotensin-converting enzyme 2 (hACE2) without triggering significant host immune response. Instead, endoplasmic reticulum stress and unfolded protein response-related pathways are predominantly activated. By comparing our data with published transcriptome of SARS-CoV-2 infection in other cell lines, we found that the expression level of hACE2 directly correlates with the viral load in infected cells but not with the scale of immune responses. Only cells that express high level of endogenous hACE2 exhibit an extensive immune attack even with a low viral load. Therefore, the infection route may be critical for the extent of the immune response, thus the severity of COVID-19 disease status.

Coronavirus (CoV) disease 2019 (COVID-19) is a pandemic caused by severe acute respiratory syndrome CoV 2 (SARS-CoV-2)<sup>1,2</sup>. Although enormous information has been gathered about this virus for the past year, there are still many unknowns. Scientists and medical doctors are rigorously developing ways to prevent SARS-CoV-2 infection, manage disease development, and apply relevant therapeutic strategies to patients with different extent of symptoms. Despite of rapid deployment of vaccines that could prevent further spread of the virus and eventually take the pandemic under control, an effective cure is still not available. Unlike the two previous deadly CoV diseases, including SARS that was caused by SARS-CoV and middle east respiratory syndrome (MERS) that was caused by MERS-CoV, COVID-19 caused by SARS-CoV-2 infection exhibits a much wider range of disease symptoms because SARS-CoV-2 is much more contagious and can be incubated longer, but has a lower death rate.

While all three pathogenic CoVs can cause similar symptoms, including fever, cough, difficulty in breathing, SARS-CoV-2 also causes additional symptoms, such as loss of taste or smell<sup>3-6</sup>. The infection by SARS-CoV-2 leads to about 80% non-severe cases (40% asymptomatic and 40% mild symptoms) and 20% severe cases (15% severe illness cases and 5% cases at critical condition that results in around 2 to 3% death globally)<sup>7-9</sup>. In addition, SARS-CoV-2 can cause lingering or recurring post COVID-19 symptoms, the so-called long COVID, including unusual fatigue, brain fog, and insomnia that can persist for a long period of time<sup>10-13</sup>. In some asymptomatic or mildly affected patients and immunocompromised patients, the SARS-CoV-2 virus can persist for a long period of time and accumulate many mutations that may present a hurdle to have the pandemic under control<sup>14-17</sup>.

Although detailed disease mechanisms underlying the wide variety of COVID-19 symptoms are lacking, it is generally accepted that SARS-CoV-2 infection primarily damages the lungs and can also cause severe problems to heart, kidney, gastrointestinal tract, and brain. Viral infection can induce various host responses that are reflected

<sup>1</sup>Department of Diabetes Complications & Metabolism, Beckman Research Institute of City of Hope, Duarte, CA 91010, USA. <sup>2</sup>Division of Stem Cell Biology Research, Department of Developmental and Stem Cell Biology, Beckman Research Institute of City of Hope, Duarte, CA 91010, USA. <sup>3</sup>Department of Molecular and Medical Pharmacology, University of California, Los Angeles, Los Angeles, CA 90095, USA. <sup>4</sup>Eli and Edythe Broad Center of Regenerative Medicine and Stem Cell Research, University of California, Los Angeles, CA 90095, USA. ✉email: VArumugaswami@mednet.ucla.edu; ariggs@coh.org; yshi@coh.org

as changes in host transcriptome, epigenome, and proteome. Identifying these changes can provide critical information at molecular level to understand the cause of symptoms and develop therapeutic strategies accordingly.

Since the beginning of the pandemic, studies that were mostly focused on severe COVID-19 cases have uncovered changes in host transcriptome, epigenome, and proteome upon SARS-CoV-2 infection and have provided strong evidence that the disease state is correlated with inadequate or defective host immune responses<sup>18–26</sup>. The abnormal host immune response includes overreaction of pro-inflammatory genes and lack of adequate anti-inflammatory gene expression, which leads to severe symptoms and systemic hyperinflammation in patients with severe cases. These studies revealed cytokine release syndrome, an acute systemic inflammatory syndrome, as the underlying immunopathology in most severe COVID-19 cases. Results from these omics studies are further supported by animal model studies and clinical data from patients with genetic mutations in immune genes that cannot mount effective Type I or Type III interferon responses<sup>27–31</sup>. However, not much resource has been applied to study symptoms in non-severe cases and the cause of post COVID-19 symptoms remains elusive. The large population of patients with non-severe COVID-19 symptoms represents a significant public health concern, therefore studies to assess these symptoms and to find ways to prevent these symptoms and their potential long-lasting effects, and to develop appropriate treatments are urgently needed.

The published transcriptome data in many cell lines and patient samples have shown that immune response genes are overwhelmingly represented in genes that are differentially expressed (DEGs) between control and SARS-CoV-2-infected cells, which makes it difficult to identify host response genes beyond immunity. In the current study, we chose HEK293T cells that are known to have innate immunity defects to identify non-immune response genes that are induced in response to SARS-CoV-2 infection<sup>32,33</sup>. Despite the very high viral load, SARS-CoV-2 infection in HEK293T cells that overexpress human angiotensin-converting enzyme 2 (hACE2) (HEK293T-hACE2) caused very low level of immune responses. Instead, the unfolded protein response (UPR), a cellular response of the endoplasmic reticulum (ER) stress, was highly activated. A set of gene expression changes observed in this transcriptome matched with gene expression changes in the transcriptome of SARS-CoV-2 infected individuals who do not undergo severe host immune response, such as asymptomatic individuals or COVID-19 patients with mild symptoms or post infection symptoms. Moreover, we have identified a list of host genes that are common to COVID-19 symptoms by comparing transcriptomic datasets in SARS-CoV-2-infected cells of different tissue types with or without host immune response.

## Results

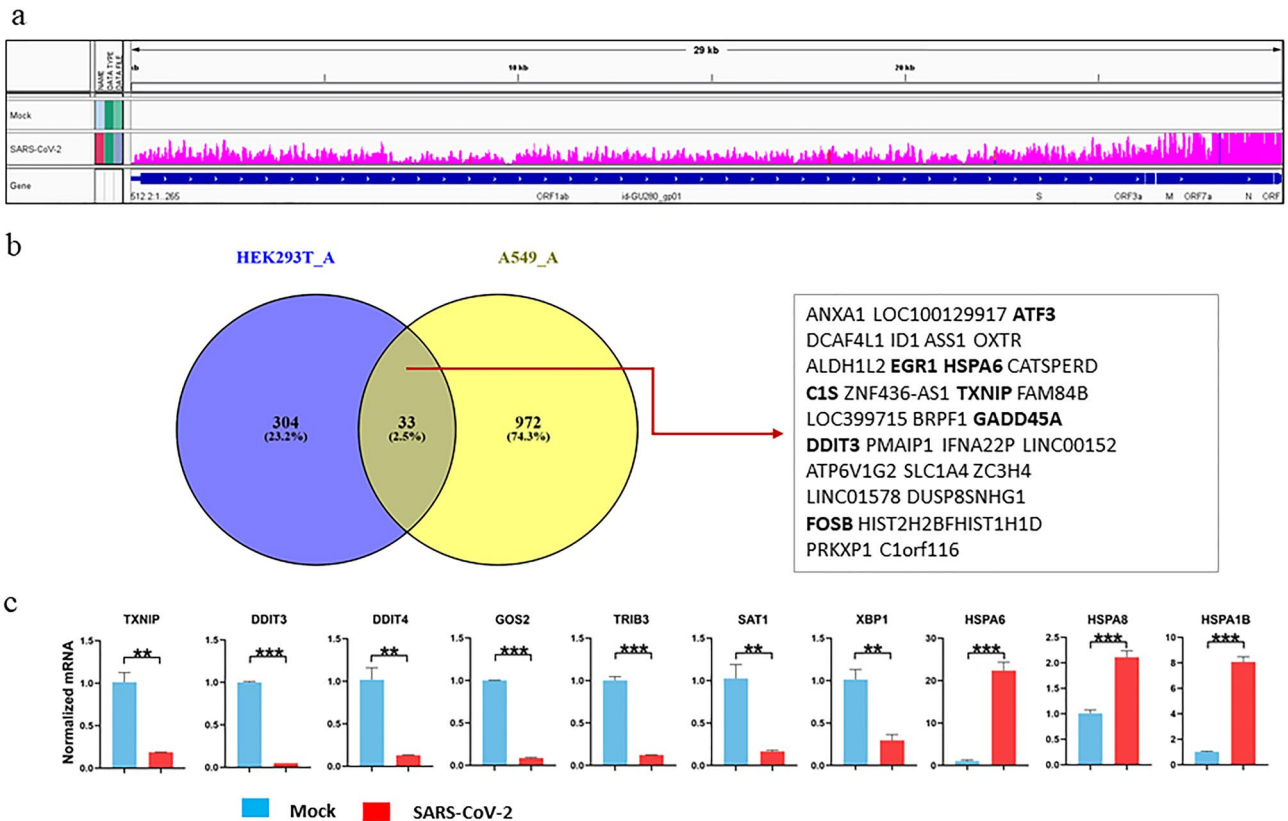
**SARS-CoV-2-infected HEK293T-hACE2 cells generated a transcriptome without strong immune response.** HEK293T cells have been used to screen SARS-CoV-2 entry factors and study the proteome, interactome, and phosphoproteome of SARS-CoV-2 infection<sup>21,34</sup>. However, due to very low level expression of both hACE2 and transmembrane protease serine 2 (TMPRSS2), two major host factors that mediate SARS-CoV-2 cellular entry<sup>35,36</sup>, HEK293T cells exhibit a very low rate of infection by SARS-CoV-2. Overexpression of hACE2 in HEK293T cells led to robust infection by SARS-CoV-2 (Fig. 1a). RNA-seq of mock-infected and SARS-CoV-2-infected HEK293T-hACE2 cells revealed several hundred significantly dysregulated host genes. As expected, we did not find any host immune response genes among the top DEGs, such as cytokines, chemokines, or interferon stimulated genes (Fig. 2a, supplemental file 2 of DEseq2 results).

Like HEK293T cells, A549 cells also have very low level of hACE2 expression and overexpression of hACE2 in A549 (A549-hACE2) cells led to robust infection<sup>27</sup>. In both SARS-CoV-2-infected HEK293T-hACE2 and A549-hACE2 cells, there are over 50% of mapped sequence reads aligned to SARS-CoV-2 genome, but host DEGs in these two cell lines are very different, having only 2.5% in common, with highly upregulated immune response genes only observed in A549-hACE2 cells (Figs. 1b, 2a–c). A list of DEGs in SARS-CoV-2-infected HEK293T-hACE2 cells were validated by quantitative RT-PCR (qRT-PCR) (Fig. 1c).

KEGG pathway analysis revealed that SARS-CoV-2 infection in HEK293T-hACE2 and A549-hACE2 cells triggered different pathways. The most significantly regulated pathway in SARS-CoV-2-infected HEK293T-hACE2 cells is the metabolic pathway, whereas the TNF, IL17, cytokine, chemokine, and the nucleotide-binding oligomerization domain-like (NOD-like) receptor signaling pathways are among the top upregulated pathways in the infected A549-hACE2 cells (Fig. 2d,e). Gene ontology (GO) analysis revealed that many of the DEGs in SARS-CoV-2-infected HEK293T-hACE2 cells are involved in cellular response to unfolded proteins, response to ER stress, and chaperone cofactor-dependent protein refolding, whereas many of the DEGs in SARS-CoV-2-infected A549-hACE2 cells are involved in the regulation of apoptotic signaling pathway, rhythmic process, and cellular response to biotic stimulus (Fig. S1–3). These data indicate that HEK293T-hACE2 and A549-hACE2 have a different host transcriptome upon SARS-CoV-2 infection although both have hACE2 overexpression and can achieve robust SARS-CoV-2 infection.

DEGs in HEK293T-hACE2-infected cells were clustered into three groups based on the expression level measured by transcripts per million reads (TPM). The SARS-CoV-2 host entry factors, ACE2, neuropilin-1 (NRP1), and TMPRSS2, were added to the DEG list for comparison purpose, and they were clustered with genes in the high, medium, and low level of expression groups, respectively (Fig. 3a)<sup>36–38</sup>. Many DEGs are involved in the top 30 GO biological processes (GO-BP) (Fig. 3b).

Among the significantly upregulated DEGs, many are heat shock protein-encoding genes, such as heat shock 70 kDa protein 6 (HSPA6) and heat shock 70 kDa protein 1B (HSPA1B). These heat shock proteins are likely used by SARS-CoV-2 to mediate the folding of newly translated viral proteins. In response to SARS-CoV-2 infection, host molecular chaperones of DnaJ homolog subfamily B (DNAJB) members, DNAJB1 and DNAJB9, were also upregulated. Among genes that are significantly downregulated in SARS-CoV-2-infected HEK293T-hACE2 cells, the tribbles homolog 3 (TRIB3) gene can be induced by NF- $\kappa$ B and negatively regulates NF- $\kappa$ B expression by a feedback mechanism. TRIB3 makes cells sensitive to TNF- and TRAIL-induced apoptosis and has been



**Figure 1.** SARS-CoV-2 infection of HEK293T-hACE2 (HEK293T\_A) cells and qPCR validation of DEGs. **(a)** IGV view of SARS-CoV-2 infection in HEK293T\_A cells: reads mapped to SARS-CoV-2 genome. **(b)** DEGs overlapped in HEK293T\_A versus A549-hACE2 (A549\_A) infection. **(c)** qPCR validation of DEGs in HEK293T\_A infection. Statistical significance was analyzed using GraphPad Prism 8. Unpaired Student's *t* test was used. For all test, *p* values were presented as \*  $P < 0.05$ , \*\*  $P < 0.01$ , and \*\*\*  $P < 0.001$ . Error bar stands for SD.

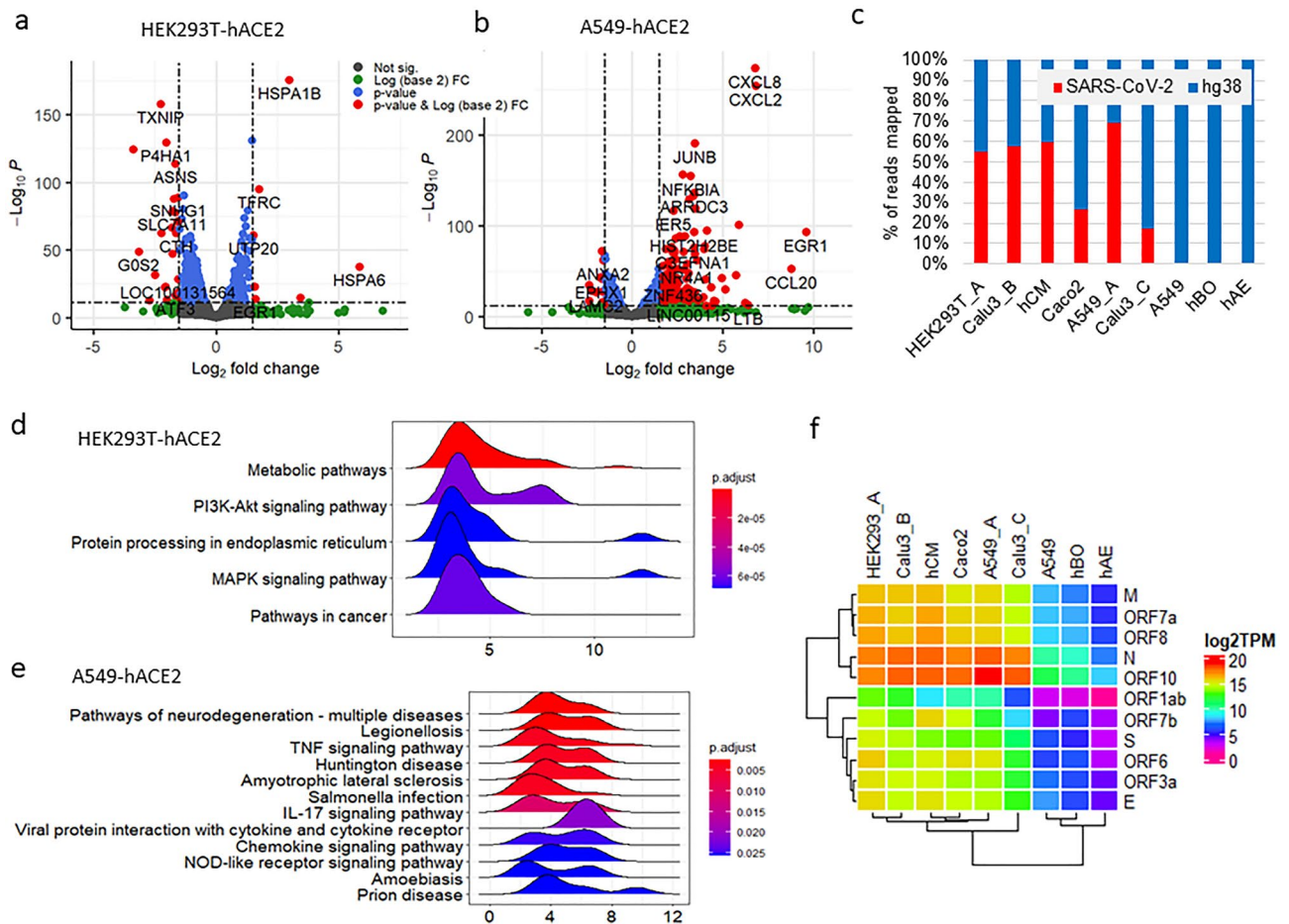
reported to decrease expression during aging<sup>39</sup>. The thioredoxin-interacting protein (TXNIP) gene plays an important role in redox regulation and is a marker gene for diabetes that responds to glucose metabolism and inflammation<sup>40</sup>. The asparagine synthetase gene (ASNA) and the G0/G1 Switch Gene 2 (GOS2) are involved in cell cycle control. Activating transcription factor 3 (ATF3), usually induced during physiological stress, is a member of the cAMP-responsive element-binding protein family<sup>41</sup>. DNA damage-inducible transcript 3 (DDIT3) is a pro-apoptotic transcription factor that has an important role in cellular stress response by blocking the binding of C/EBP members to DNA<sup>42</sup>. DDIT4 is a negative regulator of mTOR that regulates cell growth, proliferation, and autophagy<sup>43</sup>.

These DEGs are mainly enriched in genes involved in response to UPRs and ER stress, including chaperons and cofactors, apoptotic signaling pathways as shown in gene network, GO-BP, GO cellular components (GO-CC), and GO molecular function (GO-MF) analysis. KEGG disease-related pathway analysis revealed that SARS-CoV-2 infection in HEK293T-hACE2 cells mainly affects metabolic pathways and can lead to several disease conditions, such as severe sepsis, anoxia, autoimmune disease, hyperlipidemia, Alzheimer disease, and immunodeficiency (Fig. S2).

The difference in the transcriptional response to SARS-CoV-2 infection between HEK293T-hACE2 and A549-hACE2 cells prompted us to further compare our dataset of HEK293T-hACE2 with several additional RNAseq datasets that have been deposited into GEO, including primary cells such as human pluripotent stem cell-derived cardiomyocytes (hCM)<sup>23</sup>, primary human airway epithelial cells (hAE)<sup>44</sup>, and human lung bronchial organoid cells (hBO)<sup>45</sup> (Table 1, Fig. 2c). The activation of many host immune response genes upon SARS-CoV-2 infection has been reported in cells, including cell lines such as Caco2<sup>46</sup>, A549-hACE2<sup>27</sup> and Calu3 (Calu3\_B<sup>46</sup> refer to dataset from the bioRxiv preprint and published on iScience by Wyler et al.<sup>26</sup>, Calu3\_C refer to dataset from the Cell paper by Blanco-Melo et al.<sup>27</sup>) cells, and primary cells such as hCM<sup>23</sup>, hAE<sup>44</sup>, and hBO<sup>45</sup> (Fig. 2f).

To compare gene expression profile of SARS-CoV-2-infected HEK293T-hACE2 cells with other infected cells, the log<sub>2</sub> fold changes (log<sub>2</sub>FC) value of all genes from these infected cells were sorted in descending order according to the value in HEK293T-hACE2 cells before generating a heatmap. As shown in the heatmap, many of the genes that are altered upon SARS-CoV-2 infection in other cell types did not exhibit substantial expression change in infected HEK293T-hACE2 cells (Fig. 4a).

Next, we selected a list of DEGs in HEK293T-hACE2 cells with log<sub>2</sub>FC ≥ 2 and log<sub>2</sub>FC ≤ -1.5 (we used 1.5-fold as the cut-off for down-regulated genes to increase the number of downregulated genes in order to balance the DEG list because there are more genes in the upregulated list) and false discovery rate (FDR, *P*-adj value in



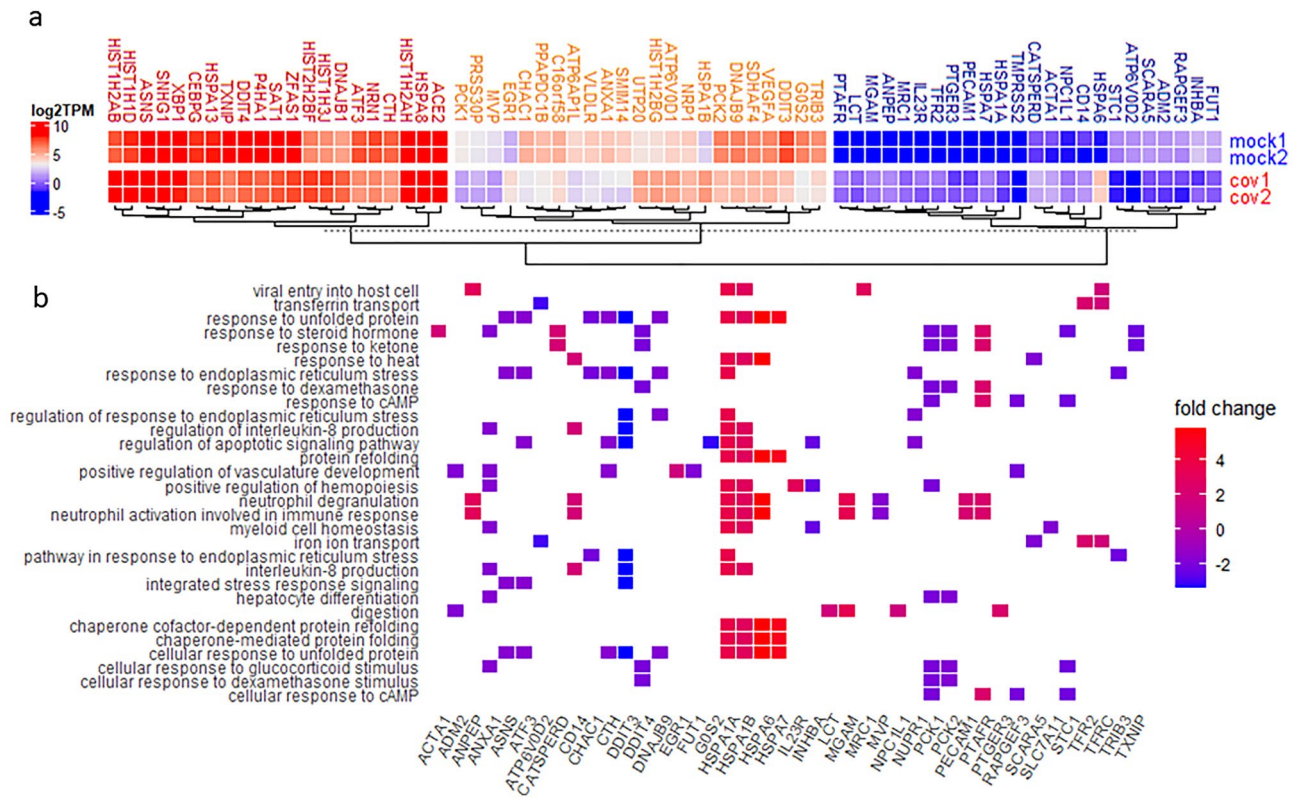
**Figure 2.** The transcriptome of SARS-CoV-2 infection in HEK293T-hACE2 cells. **(a)** Volcano Plot of DEGs of SARS-CoV-2 infection in HEK293T-hACE2 cells. **(b)** Volcano Plot of DEGs of SARS-CoV-2 infection in A549-hACE2 cells. **(c)** Percentage of reads mapped to SARS-CoV-2 genome and human hg38 genome from the nine RNAseq datasets of SARS-CoV-2 infection in cell lines used in this study. **(d)** Major KEGG pathways enriched with DEGs of SARS-CoV-2 infection in HEK293T-hACE2. **(e)** Major KEGG pathways enriched with DEGs of SARS-CoV-2 infection in A549-hACE2. **(f)** Heatmap of SARS-CoV-2 gene expression SARS-CoV-2 infected cells by transcripts per million (TPM). HEK293T\_A and A549\_A represent HEK293T or A549 with hACE2 overexpression, respectively. Calu3\_B and Calu3\_C represent data from the RNAseq data by Wyler E. (bioRxiv) and Blanco-Melo D. (Cell).

Deseq2 results)  $< 0.05$ , and then compared to their  $\log_2FC$  values in other data sets (Fig. 4b). As shown in the heatmap, many of the significantly dysregulated genes in SARS-CoV-2-infected HEK293T-hACE2 cells were absent in other infected cell lines.

By comparing DEGs with absolute value of  $\log_2FC$  [ $abs(\log_2FC)] \geq 1$  and  $FDR < 0.05$  in A549, A549\_hACE2, Calu3\_C, and HEK293T-hACE2 cells datasets, we identified an eight-gene signature that is common to the datasets from SARS-CoV-2-infected cells of different tissue types and with different host immune response status (Fig. 4c). These eight genes may represent non-immune genes that are commonly altered upon SARS-CoV-2 infection. The unique host transcriptome of SARS-CoV-2 infection in HEK293T-hACE2 cells allowed us to identify genes and pathways regulated by viral infection beyond host immune response.

**Identifying host response factors beyond immune response genes.** It is conceivable that there are a number of host transcriptomic, epigenomic, and proteomic changes associated with the acute and long-term symptoms of COVID-19 that have been masked by the extensive host immune response in severe cases. Therefore, we set to identify major factors beyond those mediating the host immune response by comparing data from SARS-CoV-2-infected HEK293T-hACE2 cells that are defective of host immune response with the datasets from cells with robust host immune response.

We focused on a list of 37 genes based on DEGs that have  $abs(\log_2FC) \geq 2$ , baseMean (mean gene expression reported in Deseq2 results) value over 20, and  $FDR$  below  $10^{-5}$  (Fig. 4d,e). We found that three heat shock genes, *HSPA6*, *HSPA1A*, and *HSPA1B*, were highly upregulated in infected HEK293T-hACE2 cells, but downregulated in the rest of infected cells analyzed except in hCM cells. It is interesting that the downregulated genes in SARS-CoV-2-infected HEK293T-hACE2 cells were either upregulated or unchanged in other infected cells. There is



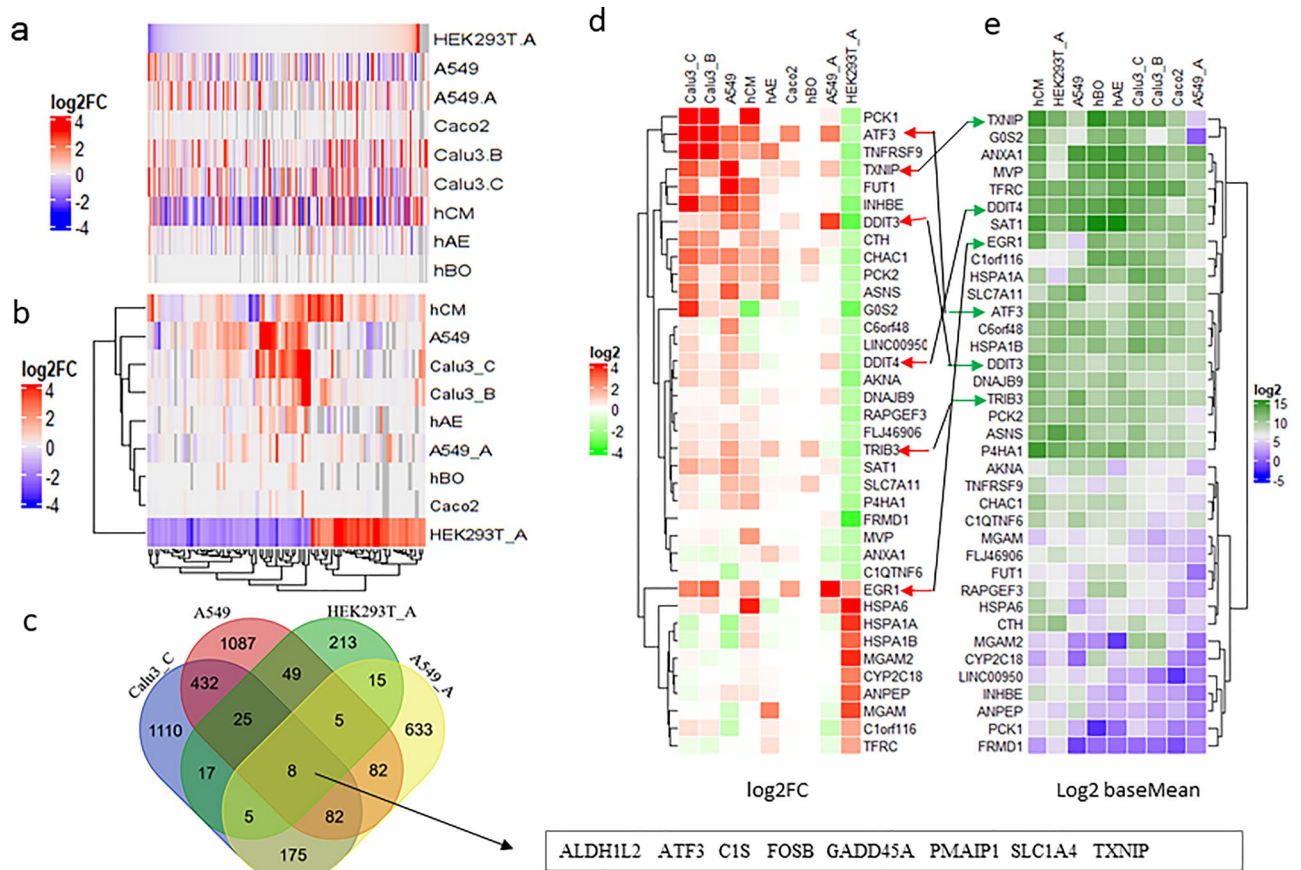
**Figure 3.** DEGs of SARS-CoV-2 infection of HEK293T-hACE2 cells. **(a)** Heatmap of DEGs in mock or SARS-CoV-2 infected HEK293T-hACE2 cells. Mock1 and mock2 indicate two replicates of mock infection and cov1 and cov2 indicate two replicates of SARS-CoV-2 infections. **(b)** Heatmap view of GO-BP enriched by DEGs from mock or SARS-CoV-2 infected HEK293T-hACE2 cells.

Datasets	Infection duration (hours)	MOI	% of viral reads	GSE ID
HEK293T-hACE2	72	1	54.77	169,158
A549-hACE2 <sup>27</sup>	24	0.2	69.21	147,507
hCM <sup>23</sup>	72	0.1	59.58	150,392
Calu3_B <sup>46</sup>	12	0.3	59.59	148,729
Caco2 <sup>46</sup>	12 and 24	0.3	26.43	148,729
Calu3_C <sup>27</sup>	24	2	17.39	147,507
A549 <sup>27</sup>	24	2	0.10	147,507
hBO <sup>45</sup>	120	5.0 × 10 <sup>4</sup> PFU/well	0.06	150,819
hAE <sup>44</sup>	48	0.25	0.01	153,970

**Table 1.** RNAseq datasets from GEO database.

only one exception that the early growth response proteins (EGR) encoding gene 1, *EGR1*, which was recently identified as an essential factor for SARS-CoV-2 infection by a pooled genome-wide CRISPR-Cas9 knockout screen<sup>47</sup>, stands out in that it is upregulated in all infected cells with high viral load. Among genes that were upregulated in infected cells with high viral load (Calu3\_B, Calu3\_C, A549-hACE2, and hCM), we found three members of EGRs, *EGR1*, *EGR2*, and *EGR3*, that were present in the common gene list with genes of cytokines, genes involved in NF-κB, MAPK, TNF pathways, further supporting a critical role of EGRs in SARS-CoV-2 infection (Fig. S4).

Next, we looked at the 10 DEGs that were commonly induced in infected Calu3\_C, A549, hAE, and hBO, regardless of the viral load. We looked at their presence in other RNAseq data that have been curated in coronavirus RNAseq database by Metascape<sup>48</sup>. Metascape analysis showed that they are present in the upregulated gene list of several RNAseq datasets and these genes are involved in inflammation and inflammatory disorder (Fig. S5a-b). Next, we combined these 10 genes with the 8 common genes from HEK293T-hACE2, A549, A549-hACE2 and Calu3\_C. Most of these genes (14 of 18) are present in RNAseq samples as upregulated genes (Fig. S5c). There is a total of 362 curated datasets (including 125 patient samples) of host transcriptome, phosphoproteome, interactome, ubiquitinome, proteome, and translome in coronavirus database by Metascape. The *TXNIP* gene



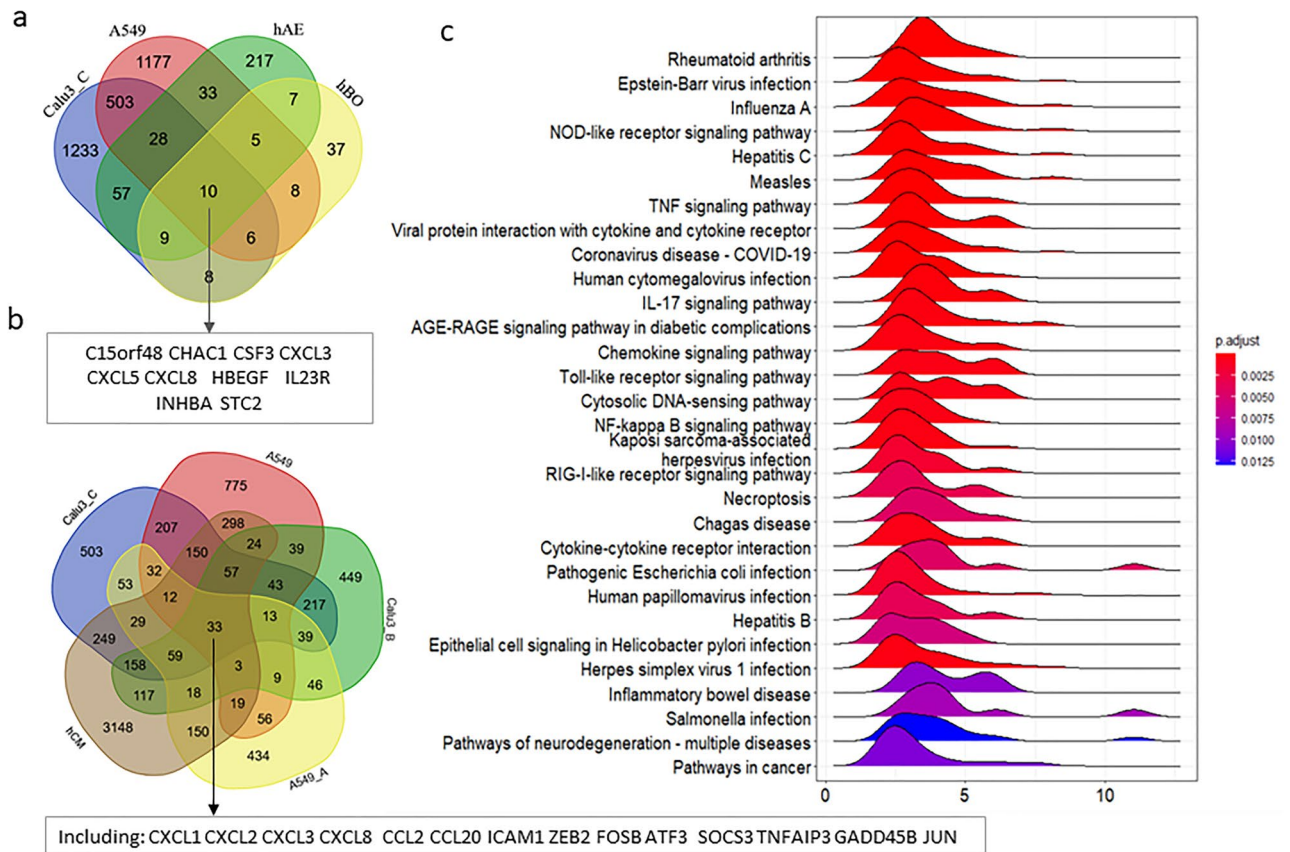
**Figure 4.** Common DEGs and genes of interest. **(a)** Heatmap view of the  $\log_2FC$  value of all genes in SARS-CoV-2 infected HEK293T-hACE2 cells versus those in other infected cell lines. Grey color indicates that the  $\log_2FC$  value are not available in the cell lines. **(b)** Heatmap view of genes with the  $\log_2FC$  value  $\geq 2$  or  $\log_2FC \leq -1.5$  in SARS-CoV-2 infected HEK293T-hACE2 cells versus those in other infected cell lines. Grey color indicates the  $\log_2FC$  value is not available in the cell lines. **(c)** Venn Diagram showing 8 common DEGs of infection in HEK293T-hACE2, A549, A549-hACE2, and Calu3\_C cells. **(d)** Heatmap view of  $\log_2FC$  value of genes of interest. **(e)** Heatmap view of baseMean value (mean gene expression level calculated by DESeq2) of genes of interest. HEK293T\_A/HEK293T.A and A549\_A/A549.A represent HEK293T or A549 with hACE2 overexpression, respectively. Calu3\_B/Calu3.B and Calu3\_C/Calu3.C represent data by Wyler E. (bioRxiv) and Blanco-Melo D. (Cell), respectively.

that was downregulated in SARS-CoV-2-infected HEK293T-hACE2 cells was present in 30 of the 125 patient samples, all of which are RNASeq datasets (Figs. 1c, 2a, S5d). Therefore, TXINP may play a role in COVID-19 development and its therapeutic potential deserve further study.

**Human ACE2 expression and host immune reaction to SARS-CoV-2 infection.** Interestingly, in the above reanalyzed data, we noticed that infection with very low viral load can activate some host immune response pathways similar to infection with very high viral load (Figure S3: A549-hACE2, S3: A549, S7: Calu3\_C, S8: hCM, and S9: hAE). We performed similar  $\log_2FC$  value plot of significantly regulated genes in several samples as we did in HEK293T-hACE2 cells. Due to the high number of genes in the list, we used genes with  $abs(\log_2FC) \geq 2$  and  $FDR < 0.05$  for A549, A549-hACE2, Calu3\_B, Calu3\_C, Caco2, and hCM for plotting.

First, we noticed that there are almost equal number of upregulated or downregulated genes in SARS-CoV-2-infected A549 and hCM cells, whereas most of the significantly dysregulated gene in SARS-CoV-2-infected Calu3 and Caco2 cells were upregulated. Of particular interest, data from infected A549 cells showed very low viral load (total viral reads by percentage is less than 0.01% in total reads mapped to hg38 genome or SARS-CoV-2 genome), but still triggered expression level change in a large number of genes.

Some of the highly upregulated genes in infected A549 cells were also present in the dataset of SARS-CoV-2-infected A549-hACE2 cells, Calu3, and hCM cells (Fig. S10a-f). This prompted us to compare DEGs of Calu3\_C, A549, hAE, and hBO, which had low level of viral infection (Fig. 2c,f). Because SARS-CoV-2 infection caused very low level gene expression change in hBO, to have enough number of DEGs for comparison, we adjusted the  $abs(\log_2FC)$  cutoff for hBO to 0.7 (genes with over 1.5-fold change) and  $FDR < 0.05$  to get its DEG list (Fig. S10g).



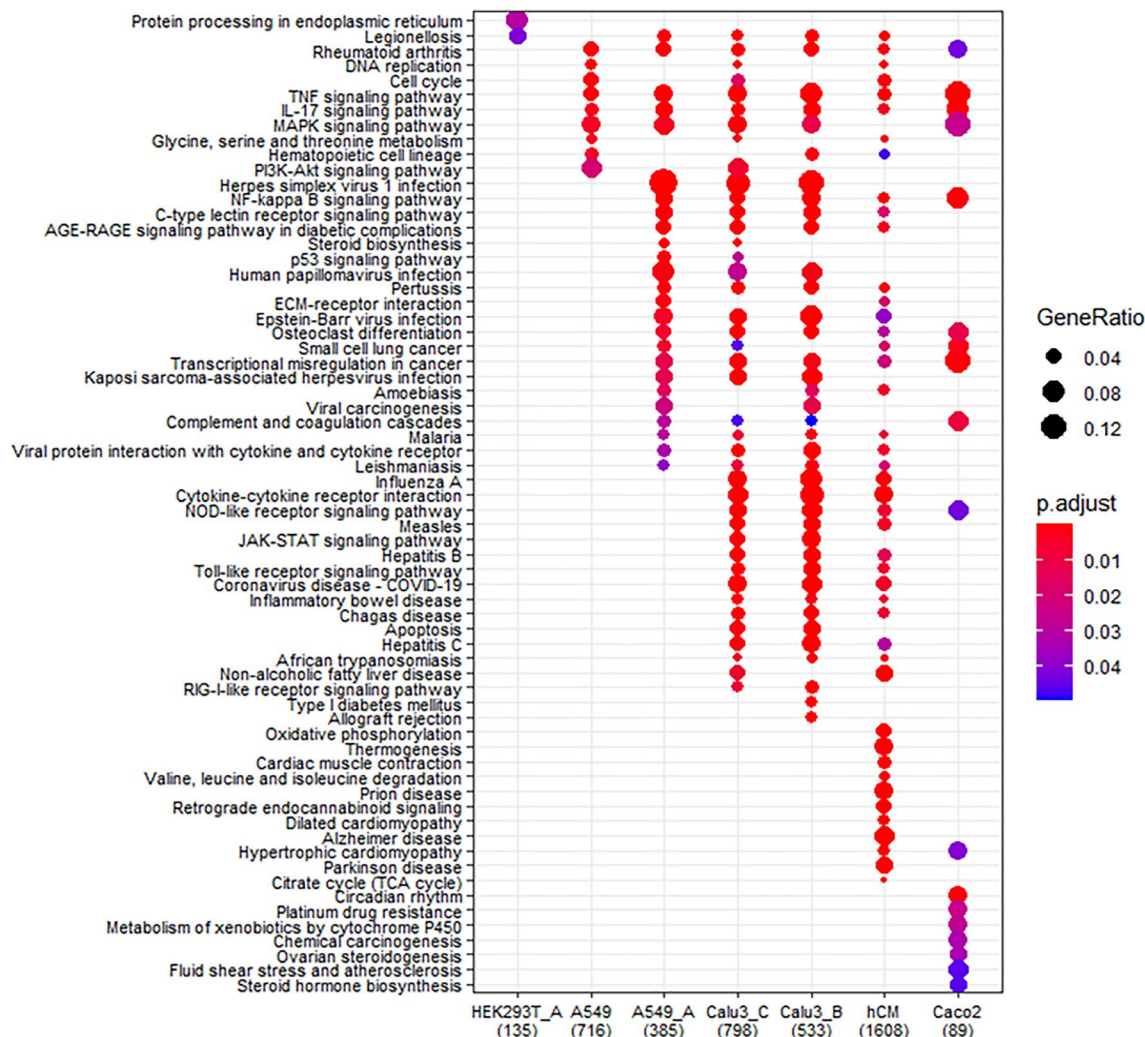
**Figure 5.** hACE2 expression and SARS-CoV-2 infection. (a) A Venn Diagram showing the 10 common DEGs in the 4 samples with low level of infection, including A549, hAE, hBO, and Calu3\_C cells. (b) A Venn Diagram showing the partial list of common DEGs in the 2 samples with low level of infection (A549 and Calu3\_C cells) and 3 samples with high level of infection (A549-hACE2, hCM, and Calu3\_B cells). (c) Top 30 KEGG pathways enriched with DEGs of SARS-CoV-2 infection in Calu3\_C cells. A549\_A represents A549 with hACE2 overexpression. Calu3\_B and Calu3\_C represent data by Wyler E. (bioRxiv) and Blanco-Melo D. (Cell), respectively.

We found that the immune response genes *CXCL3*, *CXCL5*, and *CXCL8* are among the 10 common genes in these four cell types (Fig. 5a). Comparing DEGs of infected A549 cells (low viral load) with DEGs of infected Calu3, hCM, and A549-hACE2 cells (high viral load), we found the immune response genes *CXCL1*, *CXCL2*, *CXCL3*, *CXCL8*, *CCL2*, and *CCL20* are among the 33 DEGs common to these cell lines (Fig. 5b).

These data indicate that even very low viral load can trigger host immune response by activating the host cytokine and chemokine-related pathways. Although there are more immune response genes that are activated in SARS-CoV-2-infected A549-hACE2 cells than that in infected A549 cells, there is not a direct correlation between the extent of host immune response and the level of viral load when comparing infected A549-hACE2 cells with infected Calu3\_C cells (Fig. S10). There was a much broader spectrum of immune response in infected Calu3\_C cells than that in infected A549-hACE2 cells even though infected Calu3\_C cells only had about 1/5 viral load of that in infected A549-hACE2 cells (Fig. 2c). Like infected hCM, Calu3\_B cells, and Caco2 cells, infected Calu3\_C cells activated many of the cellular response, of which defense response to type I interferon, type I interferon signaling pathways, response to viral infection, and activation of interferon response-related pathways are among the top 5 GO categories (Fig. 5c, S7).

To compare the difference at pathway level, we compared GO categories in cells with low viral load (A549) with that in cells with high viral load and hACE2 overexpression (A549-hACE2 and HEK293T-hACE2), cells with high viral load and high level of endogenous ACE2 expression (Calu3\_B and hCM), and cells with medium viral load and high level of endogenous ACE2 expression (Calu3\_C and Caco2) (Figs. 2c, 6; Table 1).

Regardless of viral load, SARS-CoV-2 infection triggered the activation of TNF, IL-17, and MAPK signaling pathways in all cells except HEK293T-hACE2. A549 and Calu3\_C also triggered the PI3K-AKT signaling pathway that controls cell survival and proliferation. This pathway was shut down when viral load reached higher level in Caco2, A549-hACE2, Calu3\_B, and hCM cells. These data indicate that TNF, IL-17, and MAPK signaling pathways that are proinflammatory and can lead to cell apoptosis and symptoms such as asthma and fever, are very sensitive to SARS-CoV-2 infection and can produce the earliest immune response factors when the infection is still at early stage and the viral load is still low.



**Figure 6.** Comparison of GO enrichment analyses of DEGs with  $abs(\log_2FC) \geq 1$  and  $FDR < 0.05$ . DEGs in HEK293T-hACE2 (HEK293T\_A), A549, A549-hACE2 (A549\_A), Calu3\_B, Calu3\_C, hCM, and Caco2 cells. Calu3\_B and Calu3\_C represent data by Wyler E. (bioRxiv) and Blanco-Melo D. (Cell), respectively.

When the viral load reached higher level, SARS-CoV-2 infection triggered several virus-induced immune reactions and NF- $\kappa$ B pathway in the host. Of interest, it also triggered the AGE-RAGE (advanced glycation end products-receptor of advanced glycation end products) signaling pathway, a pathway that has been shown to be involved in diabetic complications, pulmonary inflammatory responses, and cognitive impairment<sup>49–51</sup>. However, the Influenza A, Hepatitis B and C, and Measles-like immune reaction was only activated in naturally permissive cells with high viral load when compared to A549-hACE2 infection. In cells that express high level of endogenous hACE2, cytokine-cytokine receptor interaction, NOD-like receptor JAK-STAT, Toll-like receptor, and RIG-I-like signaling pathways were activated upon SARS-CoV-2 infection. The interferon-induced genes, such as *OAS* genes, *IFIT* genes, *INFB* genes, were only up-regulated in the three highest viral load datasets with high endogenous hACE2 expression, Calu3\_B, Calu3\_C, and hCM, but not in A549-hACE2 cells in which hACE2 was overexpressed ectopically. These data indicate that once viral load peaked in cells that express high level of endogenous hACE2, it will trigger the activation of cytokine response, response to exogenous dsRNA, activation or reinforcing JNK, NF- $\kappa$ B, MAPK, JAK-STAT signaling pathways, then the infected cells will mount a full spectrum of immune response, such as cytokine release syndrome that can lead to severe COVID-19 symptoms (Figs. 6, S11).

## Discussion

Despite it has now been over a year since the start of the COVID-19 pandemic, many key questions about SARS-CoV-2 infection and different COVID-19 symptoms it causes remain unanswered. An intriguing question about COVID-19 is how SARS-CoV-2 interplays with host during infection and how SARS-CoV-2 infection can cause so many disease symptoms. The COVID-19 symptom study app has documented over 60 symptoms and around 20% recovered patients have experienced long COVID symptoms. However, there are around 80% SARS-CoV-2



infection that does not lead to symptoms or only cause mild flu-like symptoms. Mechanisms underlying the development of chronic post-COVID-19 syndrome also remain largely unknown.

The transcriptome of SARS-CoV-2-infected cells that reflects the interplay between host and virus has provided valuable insights into mechanisms underlying SARS-CoV-2 infection and COVID-19 disease progression. However, these published works showing change in host transcriptome or proteome have been mainly focused on samples with acute and severe symptoms caused by SARS-CoV-2 infection and pointed to the dysregulated host immune reactions as the major cause for the severe outcome of SARS-CoV-2 infection. In the current research, we reported an “asymptomatic” like transcriptome of SARS-CoV-2 infection in HEK293T-hACE2 cells, in which SARS-CoV-2 infection activated many ER stress-related UPR pathways instead of mounting a substantial immune response to the SARS-CoV-2 infection due to the defect of innate immunity in HEK293T cells. The dysregulated UPR could be one of the major causes of COVID-19 symptoms besides host immune response because ER stress could lead to aberrant metabolism, organelle dysfunction, and inflammation<sup>52</sup>. Such pathways are masked by the overwhelming host immune response in cell types that exhibit extensive host immune response. Single cell RNAseq in limited number of COVID-19 patients with asymptomatic, moderate, or severe conditions identified activation of immune response genes as a prominent feature of severe COVID-19, a response that distinguishes the disease outcome from severe to mild or asymptomatic in infected individuals<sup>51</sup>. Therefore, the induction of genes and pathways other than the immune response we detected in HEK293T-hACE2 cell matches the gene expression profile in COVID patients who are asymptomatic or with mild symptoms, both of which lack severe immune response.

Our data showed that very low level of viral infection can also trigger a large-scale host immune response and lead to dysregulation of hundreds and thousands of genes, suggesting that the COVID-19 disease status may not correlate well with the viral load in an infected individual. This hypothesis can be supported by the result showing that the viral load within asymptomatic individuals was indistinguishable from that in symptomatic individuals<sup>53</sup>.

Our analysis of different datasets supports the notion that hACE2 expression level is directly correlated with the viral load but not with the scale of the immune response. We found that there is a clear difference in response to SARS-CoV-2 infection in naturally permissive cells with high level of endogenous ACE2 expression and not very permissive cells with low level of endogenous ACE2 expression. Only naturally permissive cells with high endogenous ACE2 expression mount a full-scale immune attack in response to SARS-CoV-2 infection. In this sense, it is conceivable that the type of tissues that are initially infected by SARS-CoV-2 in COVID-19 patients may impact the magnitude of the immune response, thus the extent of disease progression. Therefore, the infection route may be critical for the severity of COVID-19 disease. It is expected that infection by SARS-CoV-2 through routes with high number of virally permissive cells, such as the alveolar epithelial cells in the upper airway of respiratory system that have high level of ACE2 expression, could lead to COVID-19 progression to severe symptoms faster. On the other hand, infection through cells in the eyes, gastrointestinal tract, or reproductive system may need longer incubation time for the progression of disease and the onset of symptoms, which could explain why the development of COVID-19 have a wide range of incubation time and symptoms<sup>54</sup>. The difference in the extent of the immune response between highly permissive and slightly permissive cells suggests that once the full-scale immune response is triggered by highly permissive cells, irreversible and severe damage to cells and tissues can occur, consequently patients can progress into the severe state. Hence, antiviral drugs may not help severely ill patients, but early antiviral intervention in patients before reaching the severe state may have protective effects by blocking or slowing down viral spreading into tissues with highly permissive cells.

HEK293T cells have been widely used in SARS-CoV-2 studies, including screening for SARS-CoV-2 entry factors, replication factors, generating the infection proteome, interactome, phosphoproteome, and testing antiviral drugs. The current study indicates that those studies address situations without host immune response, thus interpretation of those results need to take the immune response into consideration.

## Materials and method

**Cell lines.** Vero-E6 cell line (ATCC CRL1586) and HEK293T (ATCC CRL-11268) cell lines were obtained from the American Type Culture Collection (ATCC).

The hACE2-HEK293T cells were made by transient transfecting HEK293T cells with hACE2 expressing vector (Addgene #1786) using calcium phosphate precipitation method, and the cells were subjected to SARS-CoV-2 infection, and RNA preparation for RNA-seq.

All cells were cultured in DMEM growth media (10% fetal bovine serum, 2 mM L-glutamine, penicillin (100 units/ml), streptomycin (100 units/ml), and 10 mM HEPES), and cultured at 37 °C and 5% CO<sub>2</sub>.

**SARS-CoV-2 infection, expansion, and quantification of viral RNA.** SARS-CoV-2 related works, such as virus expansion and quantification of viral RNA were performed as previously described<sup>23,55</sup>. SARS-CoV-2 isolate USA-WA1/2020, was obtained from the Biodefense and Emerging Infections (BEI) Resources of the National Institute of Allergy and Infectious Diseases. All procedures involving SARS-CoV-2 infection were conducted within a Biosafety Level 3 facility at UCLA. SARS-CoV-2 was passaged once in Vero-E6 cells and viral stocks were aliquoted and stored at -80 °C. Virus titer was measured in Vero-E6 cells by TCID<sub>50</sub> assay. For SARS-CoV-2 infection of HEK293T-hACE2 cells, viral inoculum (MOI of 1, or 1 plaque forming unit per cell) was prepared using serum-free medium. Culture medium was removed and replaced with 250 μL of prepared inoculum in each well. For mock infection, serum-free medium (250 μL/well) alone was added. The inoculated plates were incubated at 37 °C with 5% CO<sub>2</sub> for 1 h. The inoculum was spread by gently tilting the plate sideways at every 15 min. At the end of incubation, the inoculum was replaced with fresh medium. Cells remained at 37 °C with 5% CO<sub>2</sub> for 72 h. At the end of incubation, Trizol reagent (Thermo Fisher) was added to the well for total RNA isolation.

**RT-qPCR and statistical analysis for qPCR data.** RT-qPCR analysis was performed as described<sup>55</sup>. Specifically, Complementary DNA (cDNA) was synthesized using Tetro cDNA Synthesis Kit (Thomas Scientific). qRT-PCR was performed using SYBR Green Master Mix (Thermo Scientific) on the Step One Plus Real-Time PCR Instrument (Applied Biosystems). *ACTIN* was used as the reference gene. qPCR primers were designed using online primer design tool (<https://www.ncbi.nlm.nih.gov/tools/primer-blast/>). Statistical significance was analyzed using GraphPad Prism 8. Unpaired Student's t test was used. For all test, p values were presented as \*  $P < 0.05$ , \*\*  $P < 0.01$ , and \*\*\*  $P < 0.001$ . Error bar stands for SD.

**RNA Sequencing (RNAseq).** Total RNA and the KAPA RNA HyperPrep Kit with RiboErase (Roche) was used for RNAseq library preparation. Sequencing was done on an Illumina HiSeq 2500 machine with a single read at 51 bp read length and coverage of ~30 M reads/sample.

**RNAseq data processing, analysis, and statistics.** To have the data comparable, we reanalyzed the RNAseq datasets using the same procedure, the same data processing software, and the same genome for mapping and reads counting (Supplemental file 1. Subread genes count.csv). We then fed the raw mapped reads count into DESeq2 to generate DEG results that include the log<sub>2</sub> fold changes (log<sub>2</sub>FC) for all genes of infected samples versus corresponding mock infected samples (Supplemental file 2. DESeq2 results.csv).

Briefly, raw reads obtained from RNAseq were mapped to the human hg38 genome (iGenomes from Illumina) and SARS-CoV-2 reference genome (GenBank: NC\_045512) using subread aligner and read counting to genes were obtained with featureCounts. Subread aligner and featureCounts in Subread package release 2.0.1 were used with default parameters<sup>56</sup>. Integrative Genomics Viewer (IGV) 2.8 was used to view reads alignments in genome. Unnormalized read count from featureCounts were input into Bioconductor package DESeq2 (version 1.30.0) in R version 4.0.2 to generate the differentially expressed gene results (infected versus mock) at default setting [the Wald-test was applied to assess the p value for differential gene expressions and the adjusted p value (p-adj) was done by Benjamini and Hochberg method]<sup>57</sup>. Unless specified in the figure legends,  $P\text{-adj} < 0.05$  was used as cutoff for gene expression with significant changes.

Many Linux freeware, R or Bioconductor packages were used for data analysis and visualization. Such as ComplexHeatmap for heatmaps<sup>58</sup>, EnhancedVolcano for volcano plots, ClusterProfiler and Enrichment plot for GO and KEGG analysis<sup>59</sup>. Patient related COVID-19 gene list in Coronascope was downloaded from Metascope website<sup>48</sup>.

## Data availability

RNAseq data was deposited into GEO database under access number GSE169158. Raw gene counts from sub-read2 for HEK293T-hACE2 are listed in Supplemental file 1: HEK293T-hACE2\_genes\_count.xlsx. DESeq2 produced differential gene expression (viral infection versus mock infection) results for all datasets are listed in Supplemental file 2: DESeq2 results\_condition\_cov\_vs\_mock.xlsx.

Received: 4 May 2021; Accepted: 4 August 2021

Published online: 25 August 2021

## References

- Zhou, P. *et al.* A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270–273. <https://doi.org/10.1038/s41586-020-2012-7> (2020).
- Fehr, A. R. & Perlman, S. Coronaviruses: an overview of their replication and pathogenesis. *Methods Mol. Biol.* **1282**, 1–23. [https://doi.org/10.1007/978-1-4939-2438-7\\_1](https://doi.org/10.1007/978-1-4939-2438-7_1) (2015).
- Long, Q. X. *et al.* Clinical and immunological assessment of asymptomatic SARS-CoV-2 infections. *Nat. Med.* **26**, 1200–1204. <https://doi.org/10.1038/s41591-020-0965-6> (2020).
- Sanyaolu, A. *et al.* Comorbidity and its Impact on Patients with COVID-19. *SN Compr. Clin. Med.* <https://doi.org/10.1007/s42399-020-00363-4> (2020).
- Guan, W. J. *et al.* Comorbidity and its impact on 1590 patients with COVID-19 in China: a nationwide analysis. *Eur Respir J* <https://doi.org/10.1183/13993003.00547-2020> (2020).
- Kaur, N. *et al.* Epidemiological and clinical characteristics of 6635 COVID-19 patients: a pooled analysis. *SN Compr. Clin. Med.* <https://doi.org/10.1007/s42399-020-00393-y> (2020).
- Huang, C. *et al.* Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *Lancet* **395**, 497–506. [https://doi.org/10.1016/S0140-6736\(20\)30183-5](https://doi.org/10.1016/S0140-6736(20)30183-5) (2020).
- Song, Z. *et al.* From SARS to MERS, thrusting coronaviruses into the spotlight. *Viruses* **11**, 1–9. <https://doi.org/10.3390/v11010059> (2019).
- Kissler, S. M., Tedijanto, C., Goldstein, E., Grad, Y. H. & Lipsitch, M. Projecting the transmission dynamics of SARS-CoV-2 through the postpandemic period. *Science* **368**, 860–868. <https://doi.org/10.1126/science.abb5793> (2020).
- Carfi, A., Bernabei, R., Landi, F. & Gemelli, A. C.-P.-A.C.S.G. Persistent symptoms in patients after acute COVID-19. *JAMA* <https://doi.org/10.1001/jama.2020.12603> (2020).
- Bellan, M. *et al.* Respiratory and psychophysical sequelae among patients with COVID-19 four months after hospital discharge. *JAMA Netw. Open* **4**, e2036142. <https://doi.org/10.1001/jamanetworkopen.2020.36142> (2021).
- Huang, C. *et al.* 6-month consequences of COVID-19 in patients discharged from hospital: a cohort study. *Lancet* **397**, 220–232. [https://doi.org/10.1016/S0140-6736\(20\)32656-8](https://doi.org/10.1016/S0140-6736(20)32656-8) (2021).
- Soriano, V., Ganado-Pinilla, P., Sanchez-Santos, M. & Barreiro, P. Unveiling long COVID-19 disease. *AIDS Rev.* **22**, 227–228. <https://doi.org/10.24875/AIDSRev.M20000039> (2020).
- Baang, J. H. *et al.* Prolonged severe acute respiratory syndrome coronavirus 2 replication in an immunocompromised patient. *J. Infect. Dis.* **223**, 23–27. <https://doi.org/10.1093/infdis/jiaa666> (2021).
- Choi, B. *et al.* Persistence and evolution of SARS-CoV-2 in an Immunocompromised Host. *N. Engl. J. Med.* **383**, 2291–2293. <https://doi.org/10.1056/NEJMc2031364> (2020).
- Doykov, I. *et al.* The long tail of Covid-19-The detection of a prolonged inflammatory response after a SARS-CoV-2 infection in asymptomatic and mildly affected patients. *F1000Res* **9**, 1349. <https://doi.org/10.12688/f1000research.27287.2> (2020).

17. McCarthy, K. R. *et al.* Recurrent deletions in the SARS-CoV-2 spike glycoprotein drive antibody escape. *Science* <https://doi.org/10.1126/science.abf6950> (2021).
18. Zhou, Z. *et al.* Heightened innate immune responses in the respiratory tract of COVID-19 patients. *Cell Host Microbe* **27**, 883–890 e882. <https://doi.org/10.1016/j.chom.2020.04.017> (2020).
19. Winkler, E. S. *et al.* SARS-CoV-2 infection in the lungs of human ACE2 transgenic mice causes severe inflammation, immune cell infiltration, and compromised respiratory function. *bioRxiv* <https://doi.org/10.1101/2020.07.09.196188> (2020).
20. Bouhaddou, M. *et al.* The global phosphorylation landscape of SARS-CoV-2 infection. *Cell* **182**, 685–712 e619. <https://doi.org/10.1016/j.cell.2020.06.034> (2020).
21. Gordon, D. E. *et al.* A SARS-CoV-2 protein interaction map reveals targets for drug repurposing. *Nature* **583**, 459–468. <https://doi.org/10.1038/s41586-020-2286-9> (2020).
22. Lamers, M. M. *et al.* SARS-CoV-2 productively infects human gut enterocytes. *Science* **369**, 50–54. <https://doi.org/10.1126/science.abc1669> (2020).
23. Sharma, A. *et al.* Human iPSC-derived cardiomyocytes are susceptible to SARS-CoV-2 infection. *Cell Rep. Med.* **1**, 100052. <https://doi.org/10.1016/j.xcrm.2020.100052> (2020).
24. Hoffmann, H. H. *et al.* Functional interrogation of a SARS-CoV-2 host protein interactome identifies unique and shared coronavirus host factors. *bioRxiv* <https://doi.org/10.1101/2020.09.11.291716> (2020).
25. Corley, M. J. *et al.* Genome-wide DNA methylation profiling of peripheral blood reveals an epigenetic signature associated with severe COVID-19. *J. Leukoc Biol.* <https://doi.org/10.1002/JLB.5HI0720-466R> (2021).
26. Wyler, E. *et al.* Transcriptomic profiling of SARS-CoV-2 infected human cell lines identifies HSP90 as target for COVID-19 therapy. *iScience* **24**, 102151. <https://doi.org/10.1016/j.isci.2021.102151> (2021).
27. Blanco-Melo, D. *et al.* Imbalanced host response to SARS-CoV-2 drives development of COVID-19. *Cell* **181**, 1036–1045 e1039. <https://doi.org/10.1016/j.cell.2020.04.026> (2020).
28. Zhang, Q. *et al.* Inborn errors of type I IFN immunity in patients with life-threatening COVID-19. *Science* <https://doi.org/10.1126/science.abd4570> (2020).
29. Bastard, P. *et al.* Autoantibodies against type I IFNs in patients with life-threatening COVID-19. *Science* <https://doi.org/10.1126/science.abd4585> (2020).
30. Israelow, B. *et al.* Mouse model of SARS-CoV-2 reveals inflammatory role of type I interferon signaling. *J Exp Med* <https://doi.org/10.1084/jem.20201241> (2020).
31. Liao, M. *et al.* Single-cell landscape of bronchoalveolar immune cells in patients with COVID-19. *Nat. Med.* **26**, 842–844. <https://doi.org/10.1038/s41591-020-0901-9> (2020).
32. Burdette, D. L. *et al.* STING is a direct innate immune sensor of cyclic di-GMP. *Nature* **478**, 515–518. <https://doi.org/10.1038/nature10429> (2011).
33. Lau, L., Gray, E. E., Brunette, R. L. & Stetson, D. B. DNA tumor virus oncogenes antagonize the cGAS-STING DNA-sensing pathway. *Science* **350**, 568–571. <https://doi.org/10.1126/science.aab3291> (2015).
34. Gordon, D. E. *et al.* Comparative host-coronavirus protein interaction networks reveal pan-viral disease mechanisms. *Science* **370**, 11. <https://doi.org/10.1126/science.abe9403> (2020).
35. Harcourt, J. *et al.* Isolation and characterization of SARS-CoV-2 from the first US COVID-19 patient. *bioRxiv* <https://doi.org/10.1101/2020.03.02.972935> (2020).
36. Hoffmann, M. *et al.* SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. *Cell* **181**, 271–280 e278. <https://doi.org/10.1016/j.cell.2020.02.052> (2020).
37. Ziegler, C. G. K. *et al.* SARS-CoV-2 Receptor ACE2 is an interferon-stimulated gene in human airway epithelial cells and is detected in specific cell subsets across tissues. *Cell* **181**, 1016–1035 e1019. <https://doi.org/10.1016/j.cell.2020.04.035> (2020).
38. Daly, J. L. *et al.* Neuropilin-1 is a host factor for SARS-CoV-2 infection. *Science* **370**, 861–865. <https://doi.org/10.1126/science.abd3072> (2020).
39. Richmond, L. & Keeshan, K. Pseudokinases: a tribble-edged sword. *FEBS J.* **287**, 4170–4182. <https://doi.org/10.1111/febs.15096> (2020).
40. Cao, X., He, W., Pang, Y., Cao, Y. & Qin, A. Redox-dependent and independent effects of thioredoxin interacting protein. *Biol. Chem.* **401**, 1215–1231. <https://doi.org/10.1515/hsz-2020-0181> (2020).
41. Rohini, M., Haritha, M. A. & Selvamurugan, N. Role of activating transcription factor 3 and its interacting proteins under physiological and pathological conditions. *Int. J. Biol. Macromol.* **120**, 310–317. <https://doi.org/10.1016/j.ijbiomac.2018.08.107> (2018).
42. Ron, D. & Habener, J. F. CHOP, a novel developmentally regulated nuclear protein that dimerizes with transcription factors C/EBP and LAP and functions as a dominant-negative inhibitor of gene transcription. *Genes Dev.* **6**, 439–453. <https://doi.org/10.1101/gad.6.3.439> (1992).
43. Sofer, A., Lei, K., Johannessen, C. M. & Ellisen, L. W. Regulation of mTOR and cell growth in response to energy stress by REDD1. *Mol. Cell Biol.* **25**, 5834–5845. <https://doi.org/10.1128/MCB.25.14.5834-5845.2005> (2005).
44. Vanderheiden, A. *et al.* Type I and Type III interferons restrict SARS-CoV-2 infection of human airway epithelial cultures. *J. Virol.* <https://doi.org/10.1128/JVI.00985-20> (2020).
45. Suzuki, T. *et al.* Generation of human bronchial organoids for SARS-CoV-2 research. *bioRxiv* <https://doi.org/10.1101/2020.05.25.115600> (2020).
46. Emanuel, W. *et al.* (2020) Bulk and single-cell gene expression profiling of SARS-CoV-2 infected human cell lines identifies molecular targets for therapeutic intervention. *bioRxiv* <https://doi.org/10.1101/2020.05.05.079194> (2020).
47. Wang, R. *et al.* Genetic screens identify host factors for SARS-CoV-2 and common cold coronaviruses. *Cell* **184**, 106–119 e114. <https://doi.org/10.1016/j.cell.2020.12.004> (2021).
48. Zhou, Y. *et al.* Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523. <https://doi.org/10.1038/s41467-019-09234-6> (2019).
49. Kay, A. M., Simpson, C. L. & Stewart, J. A. Jr. The role of AGE/RAGE signaling in diabetes-mediated vascular calcification. *J. Diabetes Res.* **2016**, 6809703. <https://doi.org/10.1155/2016/6809703> (2016).
50. Gasparotto, J. *et al.* Receptor for advanced glycation end products mediates sepsis-triggered amyloid-beta accumulation, Tau phosphorylation, and cognitive impairment. *J. Biol. Chem.* **293**, 226–244. <https://doi.org/10.1074/jbc.M117.786756> (2018).
51. Oczypok, E. A., Perkins, T. N. & Oury, T. D. All the “RAGE” in lung disease: The receptor for advanced glycation endproducts (RAGE) is a major mediator of pulmonary inflammatory responses. *Paediatr. Respir. Rev.* **23**, 40–49. <https://doi.org/10.1016/j.prrv.2017.03.012> (2017).
52. Hetz, C. & Papa, F. R. The unfolded protein response and cell fate control. *Mol. Cell* **69**, 169–181. <https://doi.org/10.1016/j.molcel.2017.06.017> (2018).
53. Yang, Q. *et al.* Just 2% of SARS-CoV-2-positive individuals carry 90% of the virus circulating in communities. *medRxiv* <https://doi.org/10.1101/2021.03.01.21252250> (2021).
54. Burlacu, A., Genovesi, S., Popa, I. V. & Crisan-Dabija, R. Unpuzzling COVID-19 prothrombotic state: are preexisting thrombophilic risk profiles responsible for heterogenous thrombotic events?. *Clin. Appl. Thromb. Hemost.* **26**, 1076029620952884. <https://doi.org/10.1177/1076029620952884> (2020).
55. Wang, C. *et al.* ApoE-isoform-dependent SARS-CoV-2 neurotropism and cellular response. *Cell Stem Cell* **28**, 331–342 e335. <https://doi.org/10.1016/j.stem.2020.12.018> (2021).

56. Liao, Y., Smyth, G. K. & Shi, W. The R package Rsubread is easier, faster, cheaper and better for alignment and quantification of RNA sequencing reads. *Nucleic Acids Res.* **47**, e47. <https://doi.org/10.1093/nar/gkz114> (2019).
57. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550. <https://doi.org/10.1186/s13059-014-0550-8> (2014).
58. Gu, Z., Eils, R. & Schlesner, M. Complex heatmaps reveal patterns and correlations in multidimensional genomic data. *Bioinformatics* **32**, 2847–2849. <https://doi.org/10.1093/bioinformatics/btw313> (2016).
59. Yu, G., Wang, L. G., Han, Y. & He, Q. Y. clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* **16**, 284–287. <https://doi.org/10.1089/omi.2011.0118> (2012).

## Acknowledgements

This work was supported by the Louise and Herbert Horvitz Charitable Foundation, and California Institute for Regenerative Medicine DISC2-12172 to Y.S. Research reported in this publication included RNAseq performed in the Integrate Genomics Core of City of Hope supported by the National Cancer Institute of the National Institutes of Health under award number P30CA33572.

## Author contributions

G.S., V.A., A.R., and Y.S. designed the study. Q.C. prepared HEK293T-hACE2 cells for viral infection and extracted RNA from control and infected cells for RNA-seq. G.G. performed SARS-CoV-2 infection and cell collection. C.W. and M.Z. performed qRT-PCR. G.S. and Y.S. wrote the manuscript with inputs from other authors. All authors read and approved the final manuscript.

## Competing interests

The authors declare no competing interests.

## Additional information

**Supplementary Information** The online version contains supplementary material available at <https://doi.org/10.1038/s41598-021-96462-w>.

**Correspondence** and requests for materials should be addressed to V.A., A.D.R. or Y.S.

**Reprints and permissions information** is available at [www.nature.com/reprints](http://www.nature.com/reprints).

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021