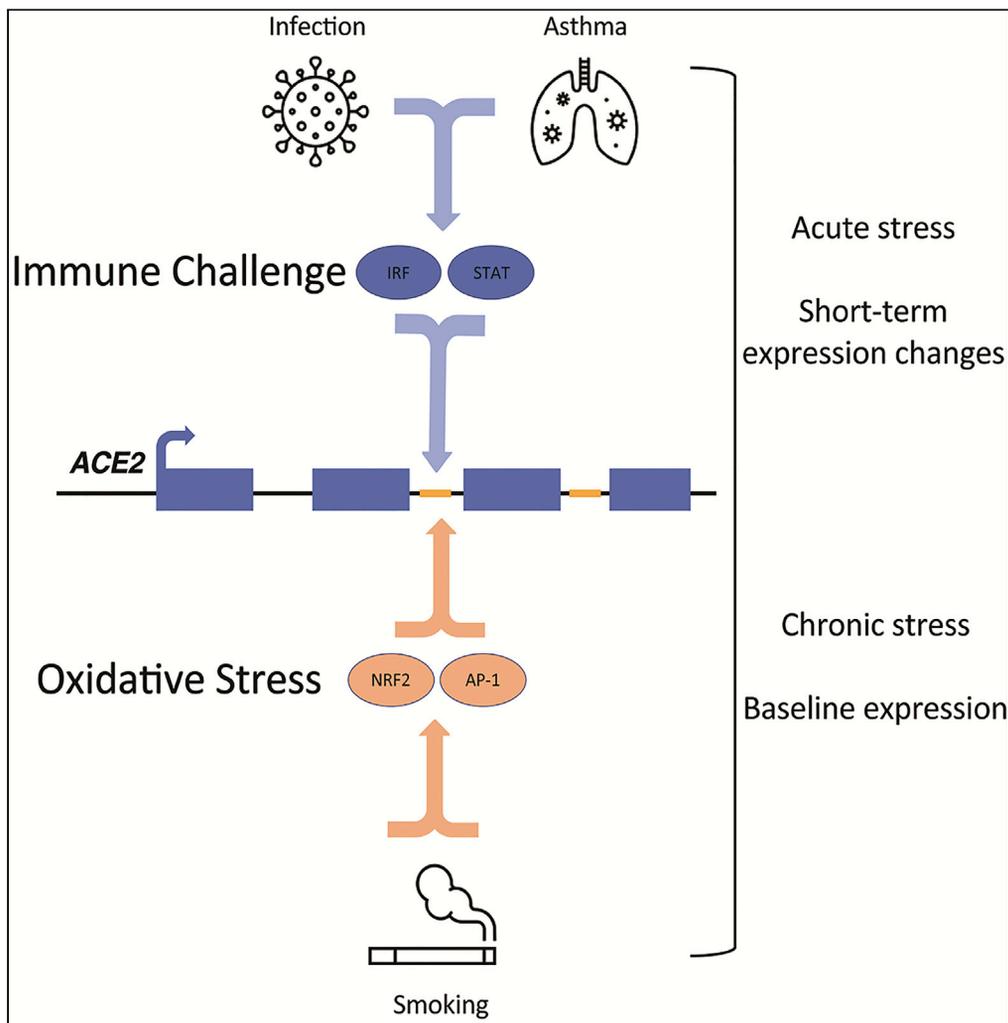**Article**

Intronic regulation of SARS-CoV-2 receptor (ACE2) expression mediated by immune signaling and oxidative stress pathways

Daniel Richard,
Pushpanathan
Muthuirulan,
Jennifer Aguiar, ...,
Karen Mossman,
Jeremy Hirota,
Terence D.
Capellini

tcapellini@fas.harvard.edu

Highlights

Lung expression patterns suggest *ACE2* regulation by immune and oxidative signaling

CRISPR deletion of intronic regulatory elements (REs) alters *ACE2* expression

Effects of RE deletion are modified by immune stimulation and oxidative stress

Propose two mechanisms for regulating *ACE2* at baseline and after immune challenge

Article

Intronic regulation of SARS-CoV-2 receptor (ACE2) expression mediated by immune signaling and oxidative stress pathways

Daniel Richard,[1,8] Pushpanathan Muthuirulan,[1,8] Jennifer Aguiar,[2] Andrew C. Doxey,[2] Arinjay Banerjee,[2,3,4] Karen Mossman,[5] Jeremy Hirota,[2,5,6] and Terence D. Capellini[1,7,9,*]

SUMMARY

The angiotensin-converting enzyme 2 (ACE2) protein is a key catalytic regulator of the renin-angiotensin system (RAS), involved in fluid homeostasis and blood pressure modulation. ACE2 also serves as a cell-surface receptor for some coronaviruses such as *SARS-CoV* and *SARS-CoV-2*. Improved characterization of *ACE2* regulation may help us understand the effects of pre-existing conditions on COVID-19 incidence, as well as pathogenic dysregulation following viral infection. Here, we perform bioinformatic analyses to hypothesize on *ACE2* gene regulation in two different physiological contexts, identifying putative regulatory elements of *ACE2* expression. We perform functional validation of our computational predictions via targeted CRISPR-Cas9 deletions of these elements *in vitro*, finding them responsive to immune signaling and oxidative-stress pathways. This contributes to our understanding of *ACE2* gene regulation at baseline and immune challenge. Our work supports pursuit of these putative mechanisms in our understanding of infection/disease caused by current, and future, SARS-related viruses such as *SARS-CoV-2*.

INTRODUCTION

The angiotensin-converting enzyme 2 (ACE2) protein has been highly studied as a key catalytic regulator of the renin-angiotensin system (RAS), involved in fluid homeostasis and blood pressure modulation (Lavoie and Sigmund, 2003). ACE2 control on this system occurs both directly (i.e., by lowering levels of angiotensin II) and indirectly (i.e., via alternative cleavage products) inhibiting the self-damaging effects of RAS over-activation, including vasoconstriction, fibrosis, and excessive inflammation (Gheblawi et al., 2020). The RAS system functions across different organs (Lavoie and Sigmund, 2003), and similarly, ACE2 is expressed throughout the body (Aguiar et al., 2020; Hikmet et al., 2020) where it mediates its protective effects and impacts tissue function (Gheblawi et al., 2020). This activity has prompted the pursuit of ACE2 as a clinical target for protection and treatment against cardiovascular disease, diabetes mellitus, and acute lung damage (Gheblawi et al., 2020; Wang et al., 2020).

In addition to its important physiological role as a broadly expressed membrane-bound protein (Gheblawi et al., 2020), ACE2 serves as a cell-surface receptor for some viruses—most notably, coronaviruses such as *SARS-CoV* (Li et al., 2003) and *SARS-CoV-2* (Gheblawi et al., 2020; Wang et al., 2020; Zhou et al., 2020). Protein overexpression studies have demonstrated that ACE2 facilitates *SARS-CoV-2* infection (Zhou et al., 2020), while mice with engineered human ACE2 are susceptible to infection (Yinda et al., 2021), and it has been suggested that the distribution of ACE2 receptor expression across different tissues contributes to differential virus susceptibility (e.g., lung tissue and alveolar cells) (Zhang et al., 2020a). The tissue expression of ACE2 may also explain the wide-ranging symptoms of COVID-19 in patients (Clerkin et al., 2020), though alternative means of viral entry has been suggested (Aguiar et al., 2020). During infection, ACE2 proteins bound by *SARS-CoV-2* particles are endocytosed which, along with increased ADAM17 activity and upstream transcriptional changes, lead to a depletion of cell-surface ACE2 localization and reduced angiotensin catalytic activity (Clerkin et al., 2020; Wang et al., 2020). It has been suggested that reduced expression of *ACE2* may lead to an imbalance of the RAS system in patients with COVID-19, which may represent a major pathological outcome of viral infection (Gheblawi et al., 2020; Lanza et al., 2020).

[1]Department of Human Evolutionary Biology, Harvard University, Cambridge, MA 02138, USA

[2]Department of Biology, University of Waterloo, Waterloo, ON N2L3G1, Canada

[3]Vaccine and Infectious Disease Organization, University of Saskatchewan, Saskatoon, SK S7N 5E3, Canada

[4]Department of Veterinary Microbiology, Western College of Veterinary Medicine, University of Saskatchewan, Saskatoon, SK S7N5B4, Canada

[5]Department of Medicine, McMaster University, Hamilton, ON L8N 3Z5, Canada

[6]Division of Respiratory Medicine, Department of Medicine, University of British Columbia, Vancouver, BC V5Z 1M9, Canada

[7]Broad Institute of MIT and Harvard, Cambridge, 02142 MA, USA

[8]These authors contributed equally

[9]Lead contact

*Correspondence: tcapellini@fas.harvard.edu

https://doi.org/10.1016/j.isci. 2022.104614

These findings have prompted intense interest in the use of recombinant ACE2 and other synthetic mimics as potential therapeutics (Hippisley-Cox et al., 2020; Monteil et al., 2021; Zhang et al., 2020b).

Given that levels of ACE2 expression may impact viral susceptibility (Aguiar et al., 2020; Devaux et al., 2020), and that subsequent changes to expression is a likely pathogenic mechanism of disease (Gheblawi et al., 2020; Lanza et al., 2020), an improved understanding of how *ACE2* expression is regulated at the genomic and transcriptional level may help us understand not only how the effects of pre-existing conditions (e.g., chronic obstructive pulmonary disease, (COPD)) may manifest with increased COVID-19 incidence but also the mechanisms that regulate ACE2 levels following viral infection (Bhalla et al., 2020; Ni et al., 2020a; Wang et al., 2022). In this study, we first perform bioinformatic analyses of several datasets to generate hypotheses about *ACE2* gene-regulatory mechanisms in the context of immune signaling and chronic oxidative stress. We next identify putative non-coding regulatory elements within the intronic regions of the *ACE2* gene as potential determinants of *ACE2* expression activity. We then perform functional validation of our computational predictions via targeted deletion of the identified *ACE2* cis-regulatory elements in the context of immunological stimulation and oxidative stress conditions. Our results demonstrate the presence of intronic *ACE2* regulatory elements responsive to both immune signaling and oxidative-stress pathways, contributing to our understanding of how expression of this gene may be modulated at both baseline and during immune challenge. Furthermore, our work supports the further pursuit of these putative mechanisms in our understanding, prevention, and treatment of infection and disease caused by ACE2-utilizing viruses such as *SARS-CoV*, *SARS-CoV-2*, and future emerging SARS-related viruses (Wells et al., 2021).

RESULTS

Upregulation of *ACE2* gene expression in healthy individuals is associated with immune signaling and viral infection

To first examine patterns of baseline *ACE2* gene expression, we analyzed microarray expression datasets from a cohort of healthy, never-smokers (N = 109) (see Table S1 for accessions). In these individuals, *ACE2* was co-expressed with a gene set that is most significantly enriched in immune signaling and virus perturbations (Figure 1, Table S1). The top transcription factors associated with these genes included IRFs and STATs (e.g., *IRF1* and *STAT1*). Consistent with this finding, both *IRF1* and *STAT1* genes were also among the top 200 *ACE2*-correlated genes. Other genes that were associated with these enriched "immune-response", and "viral response" terms, and co-expressed with *ACE2*, include *IFI16, IFI44, IFI35, NLRC5*, and *TLR3.* These findings suggest that in healthy never-smokers, *ACE2* may be a component of an immune signaling pathway, specifically relating to viral sensing and response and potentially mediated by IRF and STAT transcription factors.

Given the above findings on viral sensing, we next looked for the presence of these co-expressed immune response genes in an aggregated resource of COVID-19 RNA-seq studies (Butler et al., 2021; Park et al., 2022). We found that all of the genes highlighted above were also significantly upregulated in infected and diseased patients. More broadly, we also found that interferon signaling genes (e.g., *STAT1*) were upregulated in infected samples (Figure S1). This suggests that our association between *ACE2* co-expression and immune signaling in healthy microarray expression datasets also reflects the importance of immune signaling (possibly via IRF and STAT regulation) during *SARS-CoV-2* infection and disease.

During our analyses, we also detected as co-expressed with *ACE2, DUOX2*, a known response factor to reactive oxygen species (ROS) (Ewald, 2018). This suggests that oxidative stress may be another important mechanism that regulates *ACE2* expression (see below). Also of interest are genes that help identify cell-type-specific regulation of *ACE2*, along with its co-regulated gene network. The third top-correlated gene with *ACE2* in healthy non-smokers was *MUC13*, an epithelial mucin known to be expressed in the large intestine and trachea (Williams et al., 2001) as well as in goblet cells, which are all proposed sites of *SARS-CoV-2* replication (Gheblawi et al., 2020).

We next examined a cohort of N = 136 individuals with asthma to investigate whether the observed associations (e.g., between *ACE2* co-expression and inflammatory signaling) persisted in individuals with chronic inflammatory lung disease (Figure S2). *ACE2* co-expression in asthmatic individuals was also associated with immune signaling, antiviral responses, and IRF and STAT transcription factors. The top *ACE2*-correlated gene in asthmatics was *CD47*, which is involved in the regulation of interferon gamma.
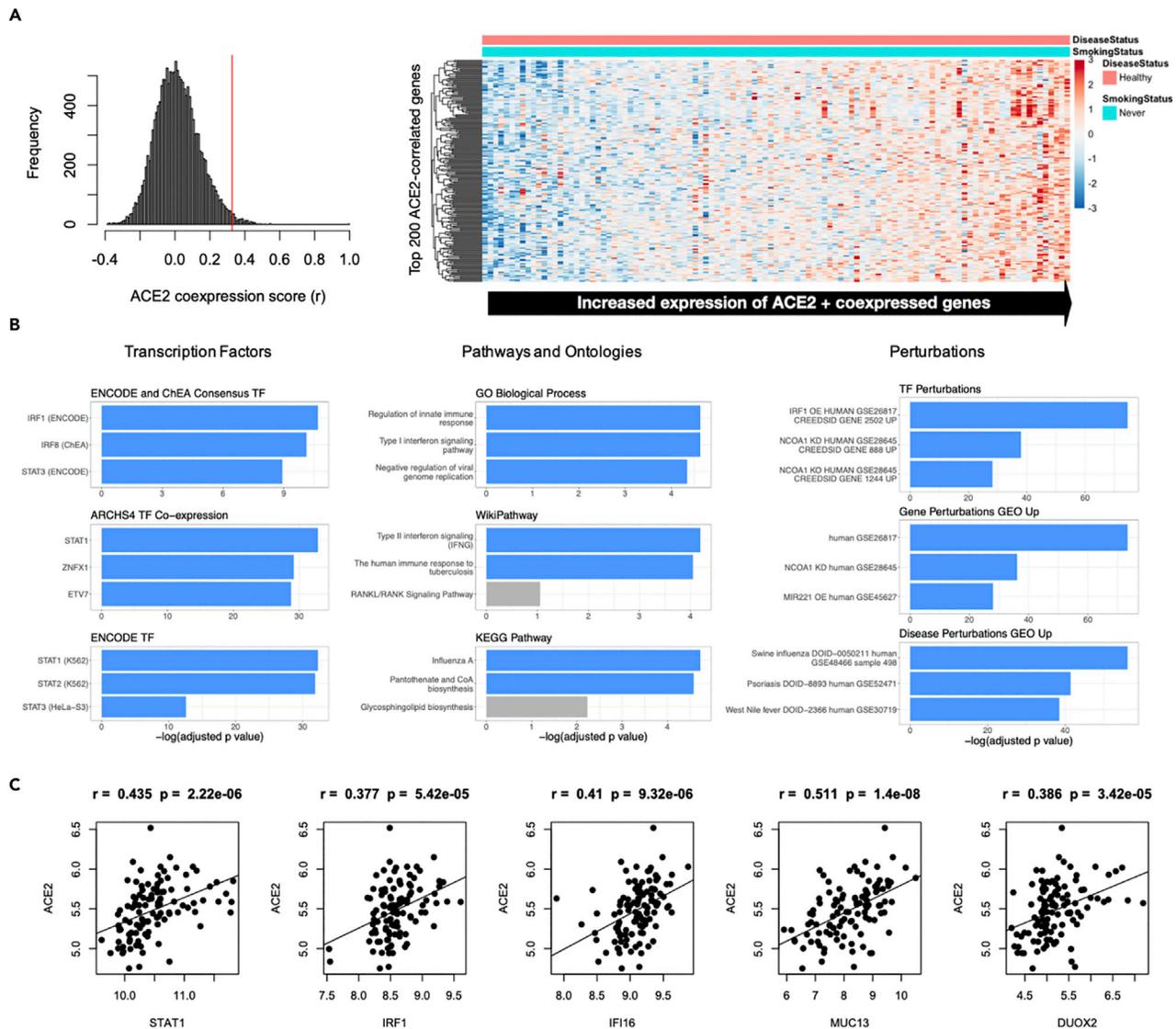
Figure 1. Expression and functional enrichment analysis of *ACE2* and co-expressed genes

(A) Expression of top 200 *ACE2*-correlated genes (including *ACE2*) in healthy non-smokers (N = 109).

(B) Functional enrichment analysis of top 200 *ACE2*-correlated genes (including *ACE2*). Terms are ranked by -log$_2$(FDR-adjusted p value) for nine ontologies/groups of interest.

(C) Pearson correlation of *ACE2* with important interferon-related candidate genes found to be co-expressed with *ACE2*.

Consistent with this finding, *ACE2* and *CD47* are both co-expressed with the interferon-inducible gene *IFI44*, whose expression is regulated by IFN-γ exposure (Zeng et al., 2006). Interestingly, *IFI44* has been suggested as a key target for controlling the cytokine storm-induced immunopathology observed in patients with influenza virus and high pathogenic coronavirus infections (Dediego et al., 2019). Based on our microarray expression analyses, we hypothesized that *ACE2* transcriptional regulation is associated with an immune signaling pathway involving IRF and STAT factors.

An important limitation of microarray data concerning *ACE2* is the inability to discriminate between full-length *ACE2* and the recently discovered short-length isoform *dACE2* (Onabajo et al., 2020). Therefore, the relative contribution of full-length ACE2 versus short-form dACE2 to these expression profiles remains unclear. We therefore sought more explicit, experimental interrogation of *ACE2* gene regulation by considering the *cis*-regulatory landscape of the *ACE2* locus.

Identification of functional intronic *ACE2* regulatory elements with STAT1 and IRF1 binding sites

Gene expression is controlled by regulatory sequences bound by transcription factors. We next examined the regulatory region encompassed by *ACE2*, compiling chromatin-accessibility datasets (i.e., DNase-I Hypersensitivity Sites, (DHS)) from *in vitro* and adult *in vivo* lung samples from the ENCODE project (Davis et al., 2018) (Figure 2). We identified six intronic putative regulatory regions overlapping either cell-line or primary tissue DHS signals, a number of which also possess potential binding motifs for STAT1 and IRF1. We then refined this list to three intronic putative regulatory elements (Regions 1,4, and 5 in Figure 2) that overlap DHS data from both lung cell and tissue data and which contained either predicted STAT1 and IRF1 binding motifs and/or aggregated ChIP-seq datasets for each factor (see STAR Methods). These predicted factor binding motifs may be directly bound by STAT1 (Regions 1 and 5), and one possibly bound by IRF1 (Region 5). Furthermore, these putative regulatory elements were also identified in a previous study of regulatory activity in the *ACE2* locus (Lee et al., 2021).

In order to test the functionality of these three regulatory elements on *ACE2* regulation, we designed CRISPR guide RNAs (sgRNAs) to target and delete each element. We also designed sgRNAs to target predicted STAT1 and IRF1 motifs in Region 4 (Table S2). We tested our targeting strategy *in vitro* on a human lung epithelial cell line (Calu-3). To rule out potential off-target effects, we first confirmed that transfection of sgRNA plasmids did not disrupt the expression of nearby genes. Expression levels of nearby *TMEM-27* and *BMX1* were not significantly altered with deletion of any element or putative binding site (Table S2). Using full-length isoform-specific primers, we next assessed levels of full-length *ACE2* transcripts using qPCR in wild type and CRISPR-deleted cells. We found that full-length *ACE2* expression was significantly decreased with deletions of each individual element, as well as the targeted binding sites within Region 4 (Figure 2C and Table S2). A recent study identified that the dACE2 isoform is regulated upon *SARS-CoV-2* infection (Onabajo et al., 2020). Interestingly, using primers specific to *dACE2*, we found that its transcript levels were also significantly decreased in our deletion experiments, and to a greater degree compared to full-length *ACE2* (Table S2). These results indicate that, in the absence of additional perturbation (i.e., above transfection), each of the three candidate intronic regulatory sequences we tested acts as an enhancer specifically for *ACE2*.

Our bioinformatic analyses of RNA-expression datasets and subsequent motif/ChIP-seq scans suggest that *ACE2* expression is regulated by an immune signaling pathway, possibly through STAT1 and IRF1 binding activity intronic to the *ACE2* locus. We therefore tested the effects of deleting these putative immune-responsive elements and specific binding sites in the context of immune signaling. Type I interferons (IFNs), such as IFN-α are our first line of defense against invading viruses (Stanifer et al., 2020). We used IFN-α treatment to induce intracellular immune signaling pathways that would occur during viral infections (Taniguchi and Takaoka, 2002) (see STAR Methods). We first performed this experiment on wild-type cells and found that this treatment led to a moderate increase of full-length *ACE2* transcripts only after 48H (Figure S3A), with dACE2 levels increasing strongly and significantly at both time points, consistent with previous studies (Lee et al., 2021; Onabajo et al., 2020) (Table S2). We independently confirmed this finding at the protein level (Figure S3B), and further found that additional potent inducers of immune signaling, such as poly(I:C) treatment and direct infection with *SARS-CoV-2*, did not lead to significant upregulation of full-length ACE2 at the protein level.

We next performed IFN-α treatment in the context of enhancer deletion. We observed a significantly decreased effect of CRISPR element deletion on full-length *ACE2* gene expression reduction with IFN-α stimulation compared to expression changes in the absence of stimulation (Figure 2D and Table S2). This was observed across the majority of our element and sub-element (i.e., motif) deletions. This attenuated downregulation was also observed for *dACE2* across stimulation-deletion experiments (Table S2). These findings suggest that these enhancer elements may be in part responsive to immunological stimulation (via IFNs) and play a role in a more complicated, potentially redundant, regulatory mechanism for *ACE2* expression (see Discussion).

ACE2 gene expression in lung epithelial cells is correlated with smoking and COPD disease status and associated with an *NRF2* antioxidant response

While much of the initial medical literature associated with *SARS-CoV-2* patient demographics suggest a link between smoking status and disease severity (Alqahtani et al., 2020; Patanavanich and Glantz, 2020;
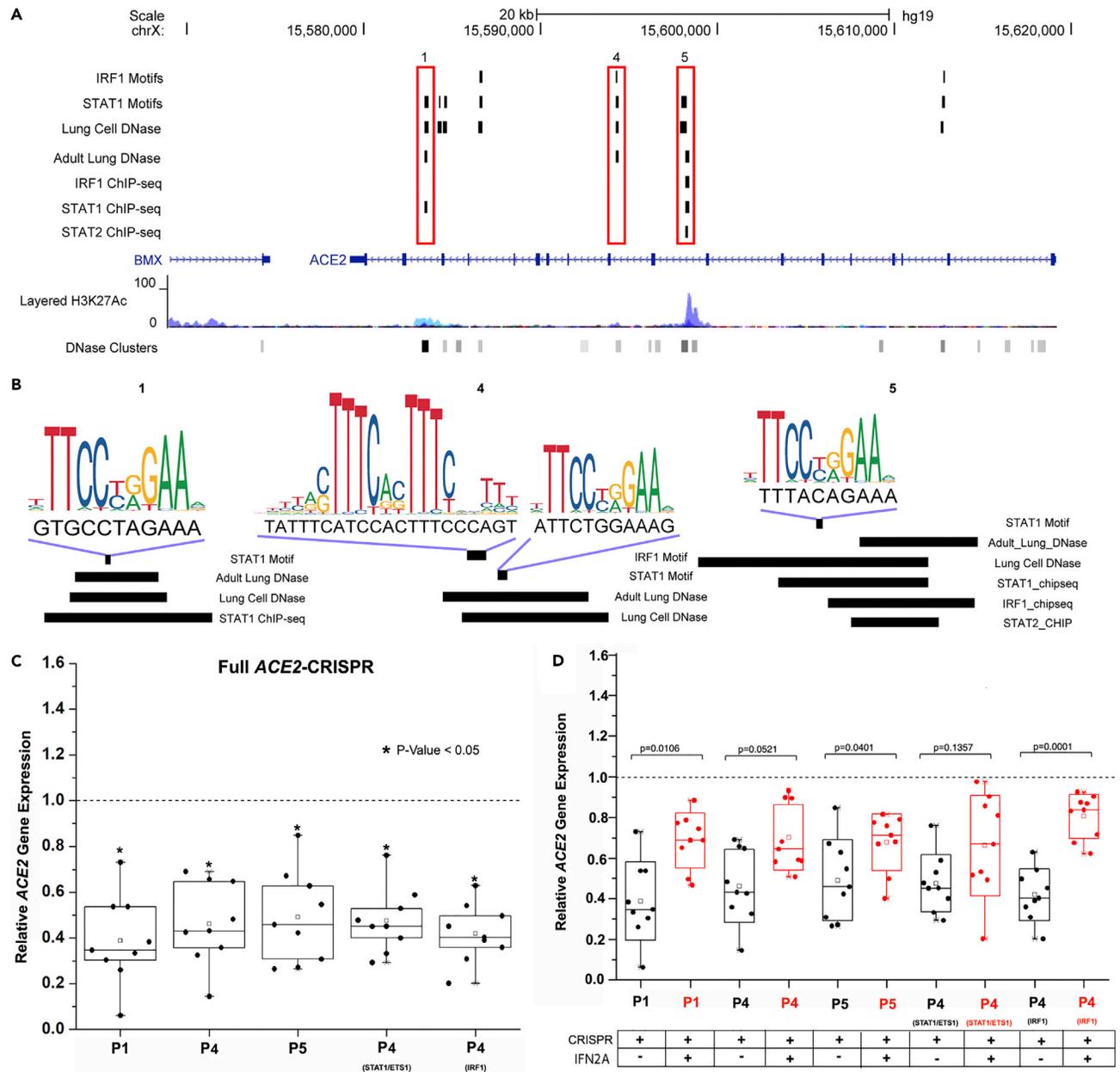
Figure 2. Identification of putative viral-response elements (STAT1 and IRF1 binding sites) in the *ACE2* intronic region

(A) Identified transcription factor binding sites in *ACE2* intronic regions in the human genome (hg19). Three separate regions labeled 1,4, and 5 contain overlapping ChIP-seq peaks including IRF1, STAT1, and STAT2 binding sites, as well as DHS in lung cell lines indicative of open-chromatin and active transcriptional regulation.

(B) DNA sequence matches to predicted IRF and STAT transcription factor binding sites in the three regions identified above, with corresponding ChIP-seq peaks indicated as horizontal bars.

(C) Deletion of regions leads to decreased expression of full-length *ACE2* (n = 9; *p < 0.05 to respective empty vector-transfected WT cells, two-tailed Student's T-test). Boxplots indicate upper/lower quartiles and median of experimental data.

(D) Reductions in expression become attenuated when elements are deleted in the presence/absence of IFN-α treatment (n = 9; *p < 0.05 to respective IFN-α untreated CRISPR KO cells, two-tailed Student's T-test). See also Table S2.

Vardavas and Nikitara, 2020), more recent studies have cast doubts as to the strength and significance of this relationship (Cattaruzza et al., 2020; Lippi and Henry, 2020). The relationship between smoking history and respiratory viral infection disease severity has been suggested to be more complicated (Zhao et al., 2020). It is also worth noting that ACE2, in addition to being the primary receptor for *SARS-CoV-2* infection

(Wang et al., 2020; Zhou et al., 2020), serves an important biological role in multiple tissues (Gheblawi et al., 2020), and is present in lung epithelium (Aguiar et al., 2020; Hikmet et al., 2020). Thus, shifts in basal expression levels of this protein, especially over time, may contribute to lung dysfunction in an indirect, more complex manner than can be measured using metrics such as COVID-19 disease severity. Given this possibility, we next assessed expression patterns of bronchial brushing datasets from current and previous smokers, focusing on *ACE2* and other co-expressed genes (e.g., *DUOX2*).

We analyzed a dataset of 159 healthy current smokers versus healthy former smokers, and identified the top 200 *ACE2*-correlated genes (Figure 3, Table S1). Expression patterns for *ACE2* suggest current smoking status is associated with increased *ACE2* levels, consistent with previous observations (Brake et al., 2020; Cai et al., 2020; Leung et al., 2020a, 2020b), and that this also accounts for the increased *ACE2* in patients with COPD (Figure 3A). Expression patterns for *ACE2*-correlated genes alone were able to effectively distinguish smokers from non-smokers (Figure 3B). Functional enrichment analysis showed that in this dataset, genes co-expressed with *ACE2* are significantly associated with the NRF2 pathway, oxidative stress, glutathione metabolism, and TGF-β regulation of the extracellular matrix. *NRF2* is a key transcription factor that regulates the oxidative stress response in the lung (Gebe et al., 2010; Rangasamy et al., 2004). Consistent with this, according to both ChIP-seq data and gene expression perturbation data, *NRF2* (*NFE2L2*) was the top transcription factor identified as a likely regulator of these genes; for example, the NRF2-regulated antioxidant gene *NQO1* was the fourth ranked *ACE2*-correlated gene in this dataset. *ACE2*-correlated genes also overlapped significantly with genes upregulated by the transcription factor ETS1 (GSE21129) (Verschoor et al., 2010); ETS1 is an important regulator of ROS (reactive oxygen species) in response to angiotensin II, linking it to ACE2 function (Ni et al., 2007). Moreover, *ETS1* expression is induced by ROS exposure through an antioxidant response (Wilson et al., 2005). Thus, *ACE2* expression in smokers appears to be associated with oxidative stress gene regulation, likely mediated by NRF2 and ETS1.

To verify these trends, we repeated the same analyses with a second cohort dataset from a different microarray platform associated with 345 healthy smokers versus healthy non-smokers (Figure S4). Notably, this dataset consists predominantly of younger individuals (age <50), whereas the first dataset includes predominantly older individuals (>50). *ACE2* co-expression patterns were highly correlated between the two independent datasets providing support that these are robust signals. As with the first analysis, genes co-expressed with *ACE2* showed the strongest associations with *NRF2* gene targets. Moreover, *NRF2* and *ETS1* formed the top three overlapping datasets according to enrichments for transcription factor perturbation datasets (Figure S4C). This dataset also included a larger number of patients with COPD; that we observed similar patterns of *ACE2* expression and co-expressing genes in this analysis may suggest similar effects on smoking and disease status on *ACE2* regulation (see Discussion).

Identification of functional intronic *ACE2* regulatory elements with possible antioxidant-response element (ARE) activity

Our microarray data analyses led to a predicted association between oxidative stress and *ACE2* levels, which prompted us to consider the existence of antioxidant-response elements (AREs within the *ACE2* locus (Figure 4A). We performed an unbiased analysis of the regulatory regions intronic to *ACE2* (Figure 4B). Our analysis identified AP-1 as the top enriched transcription factor binding site in the locus, suggesting that AP-1 may be an additional regulator of *ACE2*. Klatt and colleagues have shown AP-1-*c*-Jun subunit binding to DNA is dependent on the cellular GSH/GSSG ratio, a marker of cellular ROS levels (Klatt et al., 1999). Looking at the top *ACE2*-correlated genes across all datasets, we identified significant co-expression between *ACE2* and *FOSL2*, as well as *ACE2* and *JUN*. This suggests that AP-1 may be a transcription factor involved in oxidative stress-mediated regulation of *ACE2* levels.

Given the observed motif enrichments for AP-1 and NRF2 transcription factor binding sites in the *ACE2* locus, we next looked at the individual predicted motif hits for AP-1 and NRF2 factors within the six putative regulatory regions defined above (Figure 4B). Three of these enhancers (Regions 1, 4, and 5) contained predicted AP-1 and NRF2 binding sites and were active in both *in vivo* and *in vitro* lung datasets. Two of these regions (Region 1 and 4) also overlapped ChIP-seq datasets for FOS and JUN factors, important co-factors associated with AP-1 complex (Rauscher et al., 1988) and NRF2 (Jeyapaul and Jaiswal, 2000) enhancer binding, respectively. We next sought to delete these elements in the context of oxidative stress, with the expectation that, if these elements act as AREs, that the effects of deletion should be magnified during an ROS response.
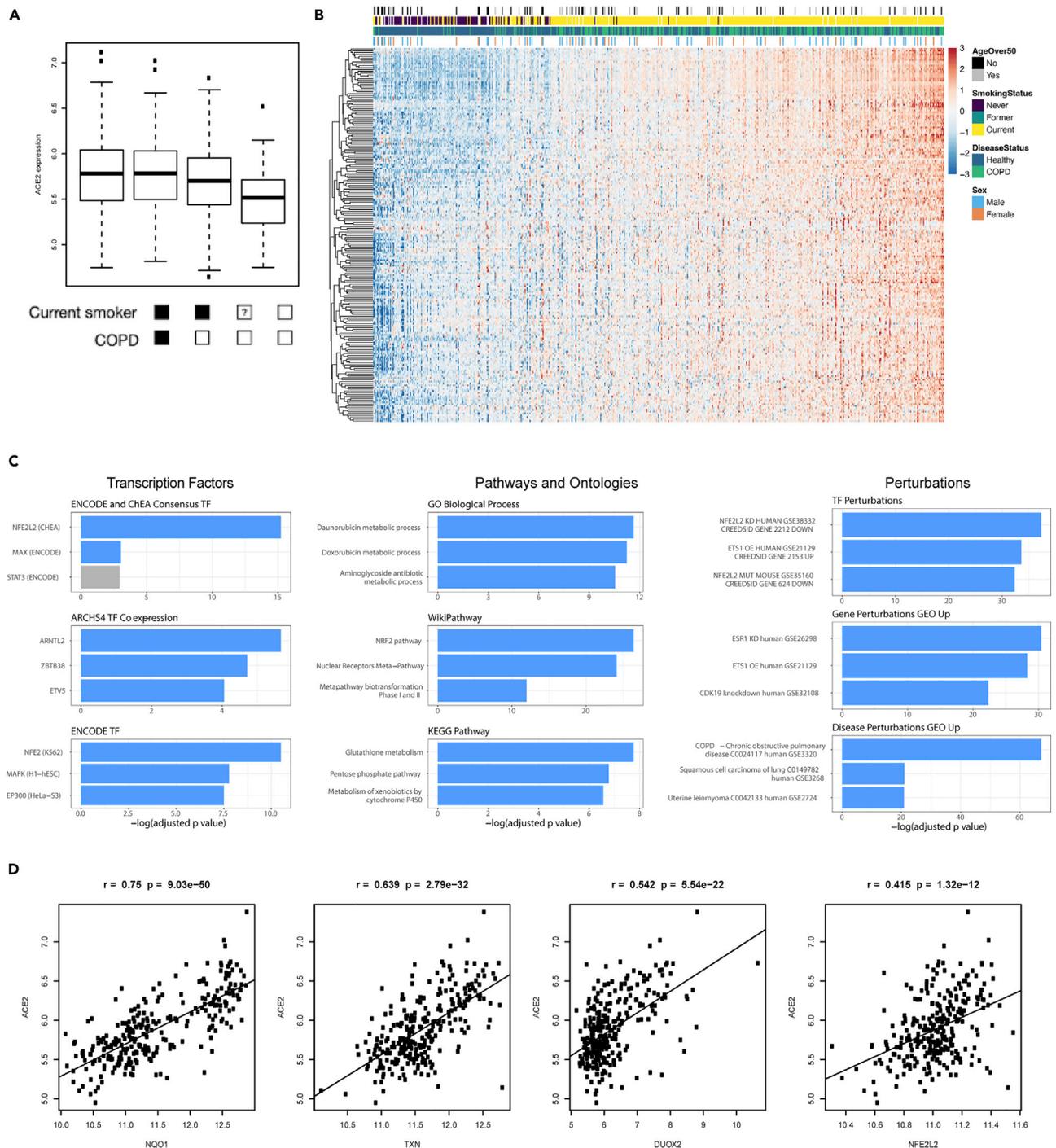
Figure 3. Expression and functional enrichment analysis of *ACE2* and co-expressed genes in smokers and individuals with COPD

(A) Analysis of relative *ACE2* expression with respect to smoking status and COPD diagnosis. Boxplots indicate upper/lower quartiles and median of expression data.

(B) Expression of top 200 *ACE2*-correlated genes (including *ACE2*) in individuals with various smoking status and COPD diagnosis (N = 159).

(C) Functional enrichment analysis of top 200 *ACE2*-correlated genes (including *ACE2*). Terms are ranked by -log$_2$(FDR-adjusted p value) for nine ontologies/groups of interest.

(D) Pearson correlation of *ACE2* with important interferon-related candidate genes found to be co-expressed with *ACE2*.
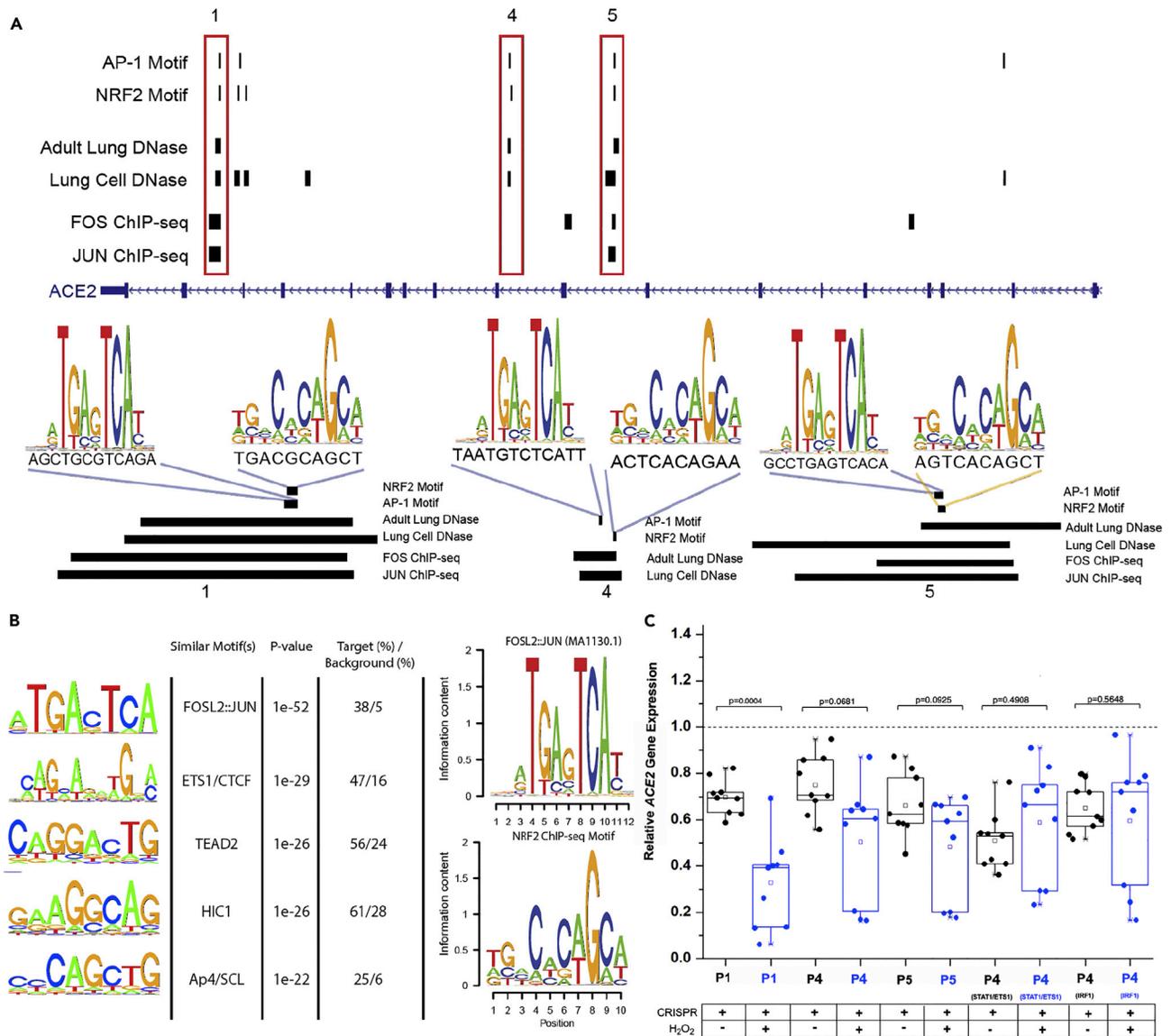
Figure 4. Putative antioxidant response elements (AREs) in the *ACE2* regulatory region

(A) Predicted AP-1 and NRF2 binding sites in *ACE2* intronic regions in the human genome. Three separate regions labeled 1, 4, and 5 (red boxes) contain overlapping ChIP-seq peaks including FOS/JUN binding sites, as well as DNase hypersensitivity peaks in adult lung tissue and cell line samples indicative of open-chromatin and active transcriptional regulation. Shown below are DNA sequence matches of predicted AP-1 and NRF2 transcription factor binding sites in the three regions identified above, with corresponding ChIP-seq peaks indicated as horizontal bars.

(B) (Left) Top statistically over-represented motifs in *ACE2* non-coding regulatory regions and their top matches to known transcription factor binding preferences. The top two binding motifs identified bear strong resemblance to FOSL2:JUN (AP1) and (ETS1/CTCF). (Right) The JASPAR-database FOSL2:JUN binding motif (MA1130.1) was enriched in intragenic *ACE2* elements. Also shown is an NRF2 binding motif defined using ChIP-seq data. Both motifs in (b) were used to scan intragenic *ACE2* element sequences in (a).

(C) Deletion of regions leads to decreased expression of full-length *ACE2* in the presence/absence of oxidative stress (blue boxes). P-values indicate results of two-tailed Student's T-test. Boxplots indicate median, upper, and lower quartiles of expression data. See also Table S2.

We first examined *ACE2* transcript levels after exposing wild-type Calu-3 cells to hydrogen peroxide (0.5mM), a potent ROS (Boardman et al., 2004) (see STAR Methods). We found that exposure to hydrogen peroxide led to significant decreases in expression of both full-length *ACE2* and *dACE2* (Figure S5, Table S2). Previous mouse studies have also demonstrated a downregulation of *ACE2* levels and activity following acute ROS exposure (Fang et al., 2019). However, *ACE2* levels are upregulated with chronic oxidative stresses (Gebe et al., 2010; Smith et al., 2020); this may suggest a more complicated regulation of *ACE2*, possibly as a function of time (see Discussion).

We next deleted each regulatory element and assessed *ACE2* expression in the presence/absence of hydrogen peroxide treatment (see STAR Methods). We observed a significant decrease in basal levels of full-length *ACE2* transcripts for the majority of regions/sites deleted in the absence of external stimulus (Table S2). In the context of exogenous oxidative stress, we observed a further significant decrease only for Region 1, while all others trended downward (Table S2). For this first region, the magnitude of down-regulation was significantly greater under oxidative stress when compared to the unstimulated deletion change (Figure 4C), while for other regions magnitudes were similarly larger despite the lack of significance (potentially due to the increased variability in expression observed in ROS-stressed cells). This first region contains predicted NRF2 and AP-1 motifs, and also overlaps with both FOS and JUN ChIP-seq signals, possibly explaining the increased effect of deletion under ROS conditions. Finally, we again saw these differences to be accentuated when considering levels of the *dACE2* transcript (Table S2).

DISCUSSION

Regulation of *ACE2* expression at the transcriptional level may impact susceptibility to viral infection. Subsequently, changes to *ACE2* expression during viral infection can lead to an imbalance in renin-angiotensin system (RAS) signaling contributing to the manifestation of clinical symptoms such as excessive inflammation (Mahmudpour et al., 2020), myocardial injury (Clerkin et al., 2020), and lung injury (Ni et al., 2020b). It has recently been suggested that dysregulation of the RAS system during initial infection may be partially responsible for the activation of the cytokine storm observed in some severe cases of COVID-19 (Mahmudpour et al., 2020), wherein the loss of surface ACE2 promotes the release of inflammatory cytokines via enhanced Ang II signaling. Elevated Ang II has also been suggested to mediate consequences of RAS imbalance tied to COVID-19 (Lanza et al., 2020), including severe hypoxia and lymphopenia. These findings indicate that targeting transcriptional inhibition of *ACE2* expression may be a therapeutic avenue for prevention of severe COVID-19 infection (Qiao et al., 2020), while counteracting infection-induced *ACE2* downregulation may act as a therapeutic treatment to reduce disease severity (Chatterjee and Thakur, 2020; Zhang et al., 2020b).

In the first part of our study, we utilized microarray expression datasets from healthy non-smokers and identified genes whose expression patterns significantly correlated with that of *ACE2*. Groups of correlated genes may suggest shared upstream regulators; gene set enrichment analyses indicated that *ACE2* and correlated genes may be under the control of immune signaling pathways integrating on the STAT and IRF families of transcription factors—namely, STAT1 and IRF1. These results were also observed when performing a separate analysis on asthmatics individuals (Figure S2). Moreover, expression data from COVID-19-infected samples further substantiated this association with immune signaling pathways (Figure S1). Interestingly, a recent study of microarray datasets from *SARS-CoV-2* infected cells suggested the importance of JAK-STAT signaling (Luo et al., 2021). In that study, the authors found that JAK-STAT components (e.g., *STAT1*) were co-expressed with *ACE2* in *SARS-CoV-2*-infected cell lines, and suggested that ACE2 could act to regulate JAK-STAT activity. Given our computational and experimental findings, we suggest that immune signaling pathways such as JAK-STAT act upstream to regulate *ACE2*. This hypothesis has been further corroborated by an independent study, which found that the transcriptional response of *ACE2* (as well as *dACE2*) to interferon stimulation was mitigated by pharmacological inhibition of JAK (Lee et al., 2021). Interestingly, this study also found that the expression of genes identified in our correlated expression analysis, including *STAT1*, *IFI44*, and *IRF1*, were also under control of immune signaling pathways involved in interferon response (Lee et al., 2021).

Given our computational findings, we considered *cis*-regulatory elements within the *ACE2* locus which may be proximate mediators of this immune-response regulatory mechanism. After identifying six such intronic regions, we prioritized and then experimentally deleted three putative intronic enhancers, containing predicted STAT1 and IRF1 binding motifs. Deletion of these three elements (Regions 1, 4, and 5) separately led to a consistent downregulation of *ACE2* transcripts. The downregulation of *ACE2* upon deletion, relative to mock-transfected controls, suggests that these enhancers may contribute to basal expression both individually but possibly collectively. This latter possibility could not be examined because experimental deletion in tandem of enhancers spanning separate introns would likely generate *ACE2* loss-of-function contexts as well as complex gene/regulatory element interactions at the locus. Interestingly, performing individual deletion of each enhancer in the context of immune stimulation, we did observe a significant attenuation of the downregulation caused by our deletions, while immune stimulation in wild-type cells caused moderate changes to *ACE2* expression (Figure S3). These results corroborate our previous findings that

SARS-CoV-2 infection does not significantly upregulate transcript levels of *ACE2* in Calu-3 cells in spite of significant increases in type I and type III IFNs, as well as upregulation of known interferon-stimulated genes (e.g., *IFIT1*, *IRF7*, and *OAS2*) (Banerjee et al., 2021). The observed attenuation effect may suggest an upper-threshold or "saturation" of *ACE2* expression from baseline, such that immune signaling does not lead to a substantial increase. However, following deletion of these putative enhancers, proximate regulators of immune signaling (e.g., STAT1 and IRF1) acting elsewhere in the *ACE2* locus (e.g., at the promoter level (Ziegler et al., 2020)) may be able to compensate for the loss/reduction in enhancer activity. Further experimentation (e.g., using a viral-infection model system) into this complex regulatory system may elucidate the role that these enhancers play in upregulating *ACE2* during an immune response.

An understanding of the mechanisms regulating *ACE2* expression during viral infection is important from a disease-pathology point of view, given that this may inhibit the protective effects of ACE2 activity. In addition, understanding the regulatory mechanisms controlling *ACE2* expression prior to viral exposure may be of equal importance from the perspective of disease prevention, given that baseline levels of *ACE2* in high-exposure tissues (e.g., lung) may modify viral susceptibility (Devaux et al., 2020).

It has been suggested, though not conclusively shown, that chronic smokers are at elevated risk to both *SARS-CoV-2* infection as well as severe disease (Alqahtani et al., 2020; Cattaruzza et al., 2020; Patanavanich and Glantz, 2020). This follows with previous studies of other coronaviruses, e.g., *MERS-CoV*, for which epidemiological evidence does suggest smoking status as a key risk factor (Arcavi and Benowitz, 2004; Park et al., 2018). In terms of increased susceptibility, it may be that smokers have elevated baseline *ACE2* expression in lung tissues, increasing the likelihood that *SARS-CoV-2* may bind their target receptors (Qiao et al., 2020). In the second part of our study, we analyzed microarray expression data from two independent datasets consisting of current smokers, non-smokers, and patients with COPD. Considering the expression of *ACE2*, we also observed previously reported increases in baseline expression within smokers (Qiao et al., 2020; Smith et al., 2020). With this, as well as genes showing similar transcriptional behaviors, we identified enrichments for oxidative stress-response pathways, including transcriptional regulators such as NRF2 and the AP-1 complex. These signals are indicative of another potential regulatory mechanism acting on the *ACE2* locus, and are expected given the chronic oxidative stresses experienced by habitual smokers and patients with COPD (Pierrou et al., 2007).

Looking again at the *ACE2* locus, we found enrichment in open-chromatin regions (putative enhancers) for DNA sequences bearing similarity to known FOSL2:JUN binding motifs, further suggesting the regulatory effects of oxidative-response signaling at this locus. We therefore performed another set of targeted deletion experiments of intronic enhancers most likely to behave as antioxidant-response elements (AREs) (e.g., contain NRF2 motifs, AP-1 ChIP-seq signals, etc.). Performing these deletions in the context of exogenous oxidative stress yielded a substantial decrease in *ACE2* expression for the first element tested (Region 1), with a fold-change decrease below that observed in wild-type cells following treatment.

We suggest that this putative ARE, and potentially others which trend in the same direction, act to counter the inhibitory effects of oxidative stress on *ACE2* expression, which has been previously observed in a mouse model of hyperoxia (Fang et al., 2019), preventing a more deleterious loss of ACE2 protein following acute exposure. ACE2 plays an important role in mitigating acute oxidative stress (Xia et al., 2011; Zheng et al., 2014), particularly in the context of cardiovascular and lung disease (Rabelo et al., 2011; Shenoy et al., 2011). We further propose that the repression of *ACE2* upon acute oxidative stress, when repeated on the order of decades in chronic smokers, may lead to an "over-compensation" of baseline *ACE2* expression—establishing higher levels of ACE2 protein to protect lung tissues from further damage. This process could be mediated by a number of oxidative stress-response mechanisms; in particular, our observed enrichments for NRF2-regulated genes co-expressed with *ACE2*, along with the presence of NRF2 motifs within intronic *ACE2* enhancers, follow with the protective role of NRF2 signaling induced in response to cigarette smoke (Ma, 2013). Furthermore, a mouse-model study of cigarette smoke found significant increases in ACE2 activity only after three weeks of exposure (Hung et al., 2016), while additional studies have found dose-response effects with increased treatment time (Gebe et al., 2010; Liu et al., 2021). Human smokers also exhibit a dose-response effect of *ACE2* expression with increasing pack-years (Smith et al., 2020). However, we acknowledge the speculative nature of this proposed over-compensating effect and note the importance of additional experimental testing. While the links between smoking status and COVID-19 severity are controversial, the link

between COPD status and COVID-19 severity may be clearer (Leung et al., 2020a, 2020b; Olloquequi, 2020; Zhao et al., 2020). More generally, it has been suggested that the detrimental effects of smoking, most notably attenuation of antiviral innate immune responses (Sopori, 2002), can increase susceptibility to pathogen infection (Arcavi and Benowitz, 2004).

In summary, here, we explored the regulatory mechanisms which may act on the *ACE2* locus in the context of both immune stimulation as well as oxidative stress, leading us to identify two putative pathways which may mediate this transcriptional regulation. It is important to note that these pathways are not mutually exclusive; the links between immune signaling and oxidative stress are well established (Rahman et al., 1996), and this is particularly true for ACE2 given its biological role in RAS regulation (Gheblawi et al., 2020; Wang et al., 2020). We suggest that further experimental testing is warranted to confirm these predicted mechanisms, and furthermore to develop potential strategies taking advantage of this knowledge to modify susceptibility and disease severity of coronavirus infections, particularly *SARS-CoV-2*.

Limitations of the study

The microarray analysis performed is limited to the detection of transcripts for which corresponding probes exist on the given chip, meaning that splice isoforms and rare variants are unlikely to be detected. This limitation impacts our study by preventing the detection of dACE2 and differentiation between this isoform and the canonical ACE2 within the sample population. Given that our identified intronic elements span multiple different exons in *ACE2*, it is extremely difficult to generate serial deletions in *cis* of these enhancers without generating any *ACE2* loss-of-function scenarios. Thus, we cannot comment on the potential combinatorial effects of different intronic *ACE2* regulatory elements to regulating expression.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead contact
  - Materials availability
  - Data and code availability
- METHOD DETAILS
  - *ACE2* co-expression and functional enrichment analysis with public microarray data
  - *ACE2* regulatory region analyses
  - *In-vitro* experimental validation studies
  - SARS-CoV-2 infection
  - Immunoblots
  - Cell line and culture condition
  - CRISPR targeting of *ACE2* regulatory elements *in vitro*
  - CRISPR deletion of *ACE2* regulatory elements under interferon treatment
  - CRISPR deletion of *ACE2* regulatory elements under H_2O_2 treatment
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2022.104614.

ACKNOWLEDGMENTS

AUTHOR CONTRIBUTIONS

T.D.C. supervised project. T.D.C., D.R., A.D., and A.B. conceived and designed the project. J.A. and A.D. performed microarray expression analyses. P.M. designed CRISPR constructs, performed CRISPR-deletion experiments, and performed expression quantification. A.B. performed Western blot experiments. K.M. and J.H. provided expert feedback. D.R. and T.D.C. wrote/edited the manuscript with input from all authors.

DECLARATION OF INTERESTS

The authors declare no competing interests.

REFERENCES

Aguiar, J.A., Tremblay, B.J.M., Mansfield, M.J., Woody, O., Lobb, B., Banerjee, A., Chandiramohan, A., Tiessen, N., Cao, Q., Dvorkin-Gheva, A., et al. (2020). Gene expression and in situ protein profiling of candidate SARS-CoV-2 receptors in human airway epithelial cells and lung tissue. Eur. Respir. J. *56*, 2001123. https://doi.org/10.1183/13993003.01123-2020.

Alqahtani, J.S., Oyelade, T., Aldhahir, A.M., Alghamdi, S.M., Almehmadi, M., Alqahtani, A.S., Quaderi, S., Mandal, S., and Hurst, J.R. (2020). Prevalence, severity and mortality associated with COPD and smoking in patients with COVID-19: a rapid systematic review and meta-analysis. PLoS One *15*, e0233147. https://doi.org/10.1371/journal.pone.0233147.

Arcavi, L., and Benowitz, N.L. (2004). Cigarette smoking and infection. Arch. Intern. Med. *164*, 2206–2216. https://doi.org/10.1001/archinte.164.20.2206.

Bailey, T.L., Boden, M., Buske, F.A., Frith, M., Grant, C.E., Clementi, L., Ren, J., Li, W.W., and Noble, W.S. (2009). MEME Suite: tools for motif discovery and searching. Nucleic Acids Res. *37*, W202–W208. https://doi.org/10.1093/nar/gkp335.

Banerjee, A., El-Sayes, N., Budylowski, P., Mcarthur, A.G., Doxey, A.C., and Mossman, K. (2021). Experimental and natural evidence of SARS-CoV-2-infection-induced activation of type I interferon responses. iScience *24*, 102477. https://doi.org/10.1016/j.isci.2021.102477.

Banerjee, A., Nasir, J.A., Budylowski, P., Yip, L., Aftanas, P., Christie, N., Ghalami, A., Baid, K., Raphenya, A.R., Hirota, J.A., et al. (2020a). Isolation, sequence, infectivity, and replication kinetics of severe acute respiratory syndrome coronavirus 2. Emerg. Infect. Dis. *26*, 2054–2063. https://doi.org/10.3201/eid2609.201495.

Banerjee, A., Zhang, X., Yip, A., Schulz, K.S., Irving, A.T., Bowdish, D., Golding, B., Wang, L.F., and Mossman, K. (2020b). Positive selection of a serine residue in bat IRF3 confers enhanced antiviral protection. iScience *23*, 100958. https://doi.org/10.1016/j.isci.2020.100958.

Bhalla, V., Blish, C.A., and South, A.M. (2020). A historical perspective on ACE2 in the COVID-19 era. J. Hum. Hypertens. *35*, 935–939. https://doi.org/10.1038/S41371-020-00459-3.

Boardman, K.C., Aryal, A.M., Miller, W.M., and Waters, C.M. (2004). Actin Re-distribution in response to hydrogen peroxide in airway epithelial cells. J. Cell. Physiol. *199*, 57–66. https://doi.org/10.1002/jcp.10451.

Brake, S.J., Barnsley, K., Lu, W., McAlinden, K.D., Eapen, M.S., and Sohal, S.S. (2020). Smoking upregulates angiotensin-converting enzyme-2 receptor: a potential adhesion site for novel coronavirus SARS-CoV-2 (Covid-19). J. Clin. Med. *9*, 841. https://doi.org/10.3390/jcm9030841.

Butler, D., Mozsary, C., Meydan, C., Foox, J., Rosiene, J., Shaiber, A., Danko, D., Afshinnekoo, E., MacKay, M., Sedlazeck, F.J., et al. (2021). Shotgun transcriptome, spatial omics, and isothermal profiling of SARS-CoV-2 infection reveals unique host responses, viral diversification, and drug interactions. Nat. Commun. *12*, 1–17. https://doi.org/10.1038/S41467-021-21361-7.

Cai, G., Bossé, Y., Xiao, F., Kheradmand, F., and Amos, C.I. (2020). Tobacco smoking increases the lung gene expression of ACE2, the Receptor of SARS-CoV-2. Am. J. Respir. Crit. Care Med. *201*, 1557–1559. https://doi.org/10.1164/rccm.202003-0693LE.

Cattaruzza, M.S., Zagà, V., Gallus, S., D'Argenio, P., and Gorini, G. (2020). Tobacco smoking and COVID-19 pandemic: old and new issues. A summary of the evidence from the scientific literature. Acta Biomed. *91*, 106–112. https://doi.org/10.23750/abm.v91i2.9698.

Chatterjee, B., and Thakur, S.S. (2020). ACE2 as a potential therapeutic target for pandemic COVID-19. RSC Adv. *10*, 39808–39813. https://doi.org/10.1039/D0RA08228G.

Chorley, B.N., Campbell, M.R., Wang, X., Karaca, M., Sambandan, D., Bangura, F., Xue, P., Pi, J., Kleeberger, S.R., and Bell, D.A. (2012). Identification of novel NRF2-regulated genes by ChiP-Seq: influence on retinoid X receptor alpha. Nucleic Acids Res. *40*, 7416–7429. https://doi.org/10.1093/nar/gks409.

Clerkin, K.J., Fried, J.A., Raikhelkar, J., Sayer, G., Griffin, J.M., Masoumi, A., Jain, S.S., Burkhoff, D., Kumaraiah, D., Rabbani, L.R., et al. (2020). COVID-19 and cardiovascular disease. Circulation *141*, 1648–1655. https://doi.org/10.1161/CIRCULATIONAHA.120.046941.

Davis, C.A., Hitz, B.C., Sloan, C.A., Chan, E.T., Davidson, J.M., Gabdank, I., Hilton, J.A., Jain, K., Baymuradov, U.K., Narayanan, A.K., et al. (2018). The Encyclopedia of DNA elements (ENCODE): data portal update. Nucleic Acids Res. *46*, D794–D801. https://doi.org/10.1093/nar/gkx1081.

Dediego, M.L., Nogales, A., Martinez-Sobrido, L., and Topham, D.J. (2019). Interferon-induced protein 44 interacts with cellular fk506-binding protein 5, negatively regulates host antiviral responses, and supports virus replication. mBio *10*, e01839-19. https://doi.org/10.1128/mBio.01839-19.

Devaux, C.A., Rolain, J.M., and Raoult, D. (2020). ACE2 receptor polymorphism: susceptibility to SARS-CoV-2, hypertension, multi-organ failure, and COVID-19 disease outcome. J. Microbiol. Immunol. Infect. *53*, 425–435. https://doi.org/10.1016/j.jmii.2020.04.015.

Ewald, C.Y. (2018). Redox signaling of nadph oxidases regulates oxidative stress responses, immunity and aging. Antioxidants *7*, 130. https://doi.org/10.3390/antiox7100130.

Fang, Y., Gao, F., and Liu, Z. (2019). Angiotensin-converting enzyme 2 attenuates inflammatory response and oxidative stress in hyperoxic lung injury by regulating NF-jB and Nrf2 pathways. QJM *112*, 914–924. https://doi.org/10.1093/qjmed/hcz206.

Gebe, S., Dieh, S., Pype, J., Friedrichs, B., Weiler, H., Schüller, J., Xu, H., Taguchi, K., Yamamoto, M., and Müller, T. (2010). The transcriptome of Nrf2-/- mice provides evidence for impaired cell cycle progression in the development of cigarette smoke-induced emphysematous changes. Toxicol. Sci. 115, 238–252. https://doi.org/10.1093/toxsci/kfq039.

Gheblawi, M., Wang, K., Viveiros, A., Nguyen, Q., Zhong, J.C., Turner, A.J., Raizada, M.K., Grant, M.B., and Oudit, G.Y. (2020). Angiotensin-converting enzyme 2: SARS-CoV-2 receptor and regulator of the renin-angiotensin system: celebrating the 20th anniversary of the discovery of ACE2. Circ. Res. 126, 1456–1474. https://doi.org/10.1161/CIRCRESAHA.120.317015.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y.C., Laslo, P., Cheng, J.X., Murre, C., Singh, H., and Glass, C.K. (2010). Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol. Cell 38, 576–589.

Hikmet, F., Méar, L., Edvinsson, Å., Micke, P., Uhlén, M., and Lindskog, C. (2020). The protein expression profile of ACE2 in human tissues. Mol. Syst. Biol. 16, e9610. https://doi.org/10.15252/msb.20209610.

Hippisley-Cox, J., Young, D., Coupland, C., Channon, K.M., Tan, P.S., Harrison, D.A., Rowan, K., Aveyard, P., Pavord, I.D., and Watkinson, P.J. (2020). Risk of severe COVID-19 disease with ACE inhibitors and angiotensin receptor blockers: cohort study including 8.3 million people. Heart 106, 1503–1511. https://doi.org/10.1136/heartjnl-2020-317393.

Hung, Y.H., Hsieh, W.Y., Hsieh, J.S., Liu, F.C., Tsai, C.H., Lu, L.C., Huang, C.Y., Wu, C.L., and Lin, C.S. (2016). Alternative roles of STAT3 and MAPK signaling pathways in the MMPs activation and progression of lung injury induced by cigarette smoke exposure in ACE2 knockout mice. Int. J. Biol. Sci. 12, 454–465. https://doi.org/10.7150/ijbs.13379.

Jeyapaul, J., and Jaiswal, A.K. (2000). Nrf2 and c-Jun regulation of antioxidant response element (ARE)-mediated expression and induction of γ-glutamylcysteine synthetase heavy subunit gene. Biochem. Pharmacol. 59, 1433–1439. https://doi.org/10.1016/S0006-2952(00)00256-2.

Karolchik, D., Barber, G.P., Casper, J., Clawson, H., Cline, M.S., Diekhans, M., Dreszer, T.R., Fujita, P.A., Guruvadoo, L., Haeussler, M., et al. (2013). The UCSC genome browser database: 2014 update. Nucleic Acids Res. 42, D764–D770.

Kent, W.J., Sugnet, C.W., Furey, T.S., Roskin, K.M., Pringle, T.H., Zahler, A.M., and Haussler, D. (2002). The human genome browser at UCSC. Genome Res. 12, 996–1006.

Klatt, P., Molina, E.P., Lacoba, M.G., Padilla, C.A., Martínez-Galisteo, E., Barcena, J., and Lamas, S. (1999). Redox regulation of c-Jun DNA binding by reversible S-glutathiolation. FASEB J. 13, 1481–1490. https://doi.org/10.1096/fasebj.13.12.1481.

Lanza, K., Perez, L.G., Costa, L.B., Cordeiro, T.M., Palmeira, V.A., Ribeiro, V.T., and Simões e Silva, A.C. (2020). Covid-19: the renin–angiotensin system imbalance hypothesis. Clin. Sci. 134, 1259. https://doi.org/10.1042/CS20200492.

Lavoie, J.L., and Sigmund, C.D. (2003). Minireview: overview of the renin-angiotensin system - an endocrine and paracrine system. In Endocrinology (Oxford Academic), pp. 2179–2183. https://doi.org/10.1210/en.2003-0150.

Lee, H.K., Jung, O., and Hennighausen, L. (2021). JAK inhibitors dampen activation of interferon-stimulated transcription of ACE2 isoforms in human airway epithelial cells. Commun. Biol. 4, 654. https://doi.org/10.1038/S42003-021-02167-1.

Leek, J.T., Johnson, W.E., Parker, H.S., Jaffe, A.E., and Storey, J.D. (2012). The sva package for removing batch effects and other unwanted variation in high-throughput experiments. Bioinformatics 28, 882. https://doi.org/10.1093/BIOINFORMATICS/BTS034.

Leung, J.M., Niikura, M., Yang, C.W.T., and Sin, D.D. (2020a). COVID-19 and COPD. Eur. Respir. J. 56, 2002180. https://doi.org/10.1183/13993003.02108-2020.

Leung, J.M., Yang, C.X., Tam, A., Shaipanich, T., Hackett, T.L., Singhera, G.K., Dorscheid, D.R., and Sin, D.D. (2020b). ACE-2 expression in the small airway epithelia of smokers and COPD patients: implications for COVID-19. Eur. Respir. J. 55, 2000688. https://doi.org/10.1183/13993003.00688-2020.

Li, W., Moore, M.J., Vasllieva, N., Sui, J., Wong, S.K., Berne, M.A., Somasundaran, M., Sullivan, J.L., Luzuriaga, K., Greeneugh, T.C., et al. (2003). Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. Nature 426, 450–454. https://doi.org/10.1038/nature02145.

Lippi, G., and Henry, B.M. (2020). Active smoking is not associated with severity of coronavirus disease 2019 (COVID-19). Eur. J. Intern. Med. 75, 107–108. https://doi.org/10.1016/j.ejim.2020.03.014.

Liu, A., Zhang, X., Li, R., Zheng, M., Yang, S., Dai, L., Wu, A., Hu, C., Huang, Y., Xie, M., and Chen, Q. (2021). Overexpression of the SARS-CoV-2 receptor ACE2 is induced by cigarette smoke in bronchial and alveolar epithelia. J. Pathol. 253, 17–30. https://doi.org/10.1002/path.5555.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods 25, 402–408. https://doi.org/10.1006/METH.2001.1262.

Luo, J., Lu, S., Yu, M., Zhu, L., Zhu, C., Li, C., Fang, J., Zhu, X., and Wang, X. (2021). The potential involvement of JAK-STAT signaling pathway in the COVID-19 infection assisted by ACE2. Gene 768, 145325. https://doi.org/10.1016/J.GENE.2020.145325.

Ma, Q. (2013). Role of Nrf2 in oxidative stress and toxicity. Annu. Rev. Pharmacol. Toxicol. 53, 401–426. https://doi.org/10.1146/annurev-pharmtox-011112-140320.

Mahmudpour, M., Roozbeh, J., Keshavarz, M., Farrokhi, S., and Nabipour, I. (2020). COVID-19 cytokine storm: the anger of inflammation. Cytokine 133, 155151. https://doi.org/10.1016/J.CYTO.2020.155151.

Mathelier, A., Fornes, O., Arenillas, D.J., Chen, C.Y., Denay, G., Lee, J., Shi, W., Shyr, C., Tan, G.,

Worsley-Hunt, R., et al. (2016). JASPAR 2016: a major expansion and update of the open-access database of transcription factor binding profiles. Nucleic Acids Res. 44, D110–D115. https://doi.org/10.1093/nar/gkv1176.

McLeay, R.C., and Bailey, T.L. (2010). Motif Enrichment Analysis: a unified framework and an evaluation on ChIP data. BMC Bioinf. 11, 165. https://doi.org/10.1186/1471-2105-11-165.

Monteil, V., Dyczynski, M., Lauschke, V.M., Kwon, H., Wirnsberger, G., Youhanna, S., Zhang, H., Slutsky, A.S., Hurtado del Pozo, C., Horn, M., et al. (2021). Human soluble ACE2 improves the effect of remdesivir in SARS-CoV-2 infection. EMBO Mol. Med. 13, e13426. https://doi.org/10.15252/emmm.202013426.

Ni, W., Yang, X., Yang, D., Bao, J., Li, R., Xiao, Y., Hou, C., Wang, H., Liu, J., Yang, D., et al. (2020a). Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. Crit. Care 24, 1–10. https://doi.org/10.1186/S13054-020-03120-0/FIGURES/3.

Ni, W., Yang, X., Yang, D., Bao, J., Li, R., Xiao, Y., Hou, C., Wang, H., Liu, J., Yang, D., et al. (2020b). Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. Crit. Care 24, 422. https://doi.org/10.1186/s13054-020-03120-0.

Ni, W., Zhan, Y., He, H., Maynard, E., Balschi, J.A., and Oettgen, P. (2007). Ets-1 is a critical transcriptional regulator of reactive oxygen species and p47phox gene expression in response to angiotensin II. Circ. Res. 101, 985–994. https://doi.org/10.1161/CIRCRESAHA.107.152439.

Oki, S., Ohta, T., Shioi, G., Hatanaka, H., Ogasawara, O., Okuda, Y., Kawaji, H., Nakaki, R., Sese, J., and Meno, C. (2018). ChIP-Atlas: a data-mining suite powered by full integration of public ChIP-seq data. EMBO Rep. 19, e46255. https://doi.org/10.15252/embr.201846255.

Olloquequi, J. (2020). COVID-19 Susceptibility in chronic obstructive pulmonary disease. Eur. J. Clin. Invest. 50, e13382. https://doi.org/10.1111/eci.13382.

Onabajo, O.O., Banday, A.R., Stanifer, M.L., Yan, W., Obajemu, A., Santer, D.M., Florez-Vargas, O., Piontkivska, H., Vargas, J.M., Ring, T.J., et al. (2020). Interferons and viruses induce a novel truncated ACE2 isoform and not the full-length SARS-CoV-2 receptor. Nat. Genet. 52, 1283–1293. https://doi.org/10.1038/s41588-020-00731-9.

Park, J., Foox, J., Hether, T., Danko, D.C., Warren, S., Kim, Y., Reeves, J., Butler, D.J., Mozsary, C., Rosiene, J., et al. (2022). System-wide transcriptome damage and tissue identity loss in COVID-19 patients. Cell Rep. Med. 3, 100522. https://doi.org/10.1016/J.XCRM.2022.100522/ATTACHMENT/6D64EBBB-8BE3-45B8-B074-AFBD8F2615F3/MMC4.XLSX.

Park, J.E., Jung, S., and Kim, A. (2018). MERS transmission and risk factors: a systematic review. BMC Publ. Health 18, e2122240. https://doi.org/10.1186/s12889-018-5484-8.

Patanavanich, R., and Glantz, S.A. (2020). Smoking is associated with COVID-19 progression: a meta-analysis. Nicotine Tob. Res. 22, 1653–1656. https://doi.org/10.1093/ntr/ntaa082.

Pierrou, S., Broberg, P., O'Donnell, R.A., Pawłowski, K., Virtala, R., Lindqvist, E., Richter, A., Wilson, S.J., Angco, G., Möller, S., et al. (2007). Expression of genes involved in oxidative stress responses in airway epithelial cells of smokers with chronic obstructive pulmonary disease. Am. J. Respir. Crit. Care Med. *175*, 577–586. https://doi.org/10.1164/rccm.200607-931OC.

Qiao, Y., Wang, X.M., Mannan, R., Pitchiaya, S., Zhang, Y., Wotring, J.W., Xiao, L., Robinson, D.R., Wu, Y.M., Tien, J.C.Y., et al. (2020). Targeting transcriptional regulation of SARS-CoV-2 entry factors ACE2 and TMPRSS2. Proc. Natl. Acad. Sci. U S A *118*. e2021450118. https://doi.org/10.1073/pnas.2021450118.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841–842. https://doi.org/10.1093/bioinformatics/btq033.

R Development Core Team (2008). R: A Language and Environment for Statistical Computing.

Rabelo, L.A., Alenina, N., and Bader, M. (2011). ACE2-angiotensin-(1-7)-Mas axis and oxidative stress in cardiovascular disease. Hypertens. Res. *34*, 154–160. https://doi.org/10.1038/hr.2010.235.

Rahman, I., Morrison, D., Donaldson, K., and Macnee, W. (1996). Systemic oxidative stress in asthma, COPD, and smokers. Am. J. Respir. Crit. Care Med. *154*, 1055–1060. https://doi.org/10.1164/ajrccm.154.4.8887607.

Ran, F.A., Hsu, P.D., Wright, J., Agarwala, V., Scott, D.A., and Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. Nat. Protoc. *8*, 2281.

Rangasamy, T., Cho, C.Y., Thimmulappa, R.K., Zhen, L., Srisuma, S.S., Kensler, T.W., Yamamoto, M., Petrache, I., Tuder, R.M., and Biswal, S. (2004). Genetic ablation of Nrf2 enhances susceptibility to cigarette smoke–induced emphysema in mice. J. Clin. Invest. *114*, 1248–1259. https://doi.org/10.1172/jci21146.

Rauscher, F.J., Voulalas, P.J., Franza, B.R., and Curran, T. (1988). Fos and Jun bind cooperatively to the AP-1 site: reconstitution in vitro. Genes Dev. *2*, 1687–1699. https://doi.org/10.1101/gad.2.12b.1687.

Sean, D., and Meltzer, P.S. (2007). GEOquery: a bridge between the gene expression Omnibus (GEO) and BioConductor. Bioinformatics *23*, 1846–1847. https://doi.org/10.1093/bioinformatics/btm254.

Shenoy, V., Qi, Y., Katovich, M.J., and Raizada, M.K. (2011). ACE2, a promising therapeutic target for pulmonary hypertension. Curr. Opin. Pharmacol. *11*, 150–155. https://doi.org/10.1016/j.coph.2010.12.002.

Smith, J.C., Sausville, E.L., Girish, V., Yuan, M.L., Vasudevan, A., John, K.M., and Sheltzer, J.M. (2020). Cigarette smoke exposure and inflammatory signaling increase the expression of the SARS-CoV-2 receptor ACE2 in the respiratory tract. Dev. Cell *53*, 514–529.e3. https://doi.org/10.1016/j.devcel.2020.05.012.

Sopori, M. (2002). Effects of cigarette smoke on the immune system. Nat. Rev. Immunol. *2*, 372–377. https://doi.org/10.1038/nri803.

Stanifer, M.L., Guo, C., Doldan, P., and Boulant, S. (2020). Importance of type I and III interferons at respiratory and intestinal barrier surfaces. Front. Immunol. *11*, 3240. https://doi.org/10.3389/FIMMU.2020.608645/BIBTEX.

Tan, G., and Lenhard, B. (2016). TFBSTools: an R/bioconductor package for transcription factor binding site analysis. Bioinformatics *32*, 1555–1556. https://doi.org/10.1093/bioinformatics/btw024.

Taniguchi, T., and Takaoka, A. (2002). The interferon-α/β system in antiviral responses: a multimodal machinery of gene regulation by the IRF family of transcription factors. Curr. Opin. Immunol. *14*, 111–116. https://doi.org/10.1016/S0952-7915(01)00305-3.

Vardavas, C.I., and Nikitara, K. (2020). COVID-19 and smoking: a systematic review of the evidence. Tob. Induc. Dis. *18*, 20. https://doi.org/10.18332/tid/119324.

Verschoor, M.L., Wilson, L.A., Verschoor, C.P., and Singh, G. (2010). Ets-1 regulates energy metabolism in cancer cells. PLoS One *5*, e13565. https://doi.org/10.1371/journal.pone.0013565.

Wang, J., Zhao, H., and An, Y. (2022). ACE2 shedding and the role in COVID-19. Front. Cell. Infect. Microbiol. *11*, 1422. https://doi.org/10.3389/FCIMB.2021.789180/BIBTEX.

Wang, K., Gheblawi, M., and Oudit, G.Y. (2020). Angiotensin converting enzyme 2: a double-edged sword. Circulation *142*, 426–428. https://doi.org/10.1161/CIRCULATIONAHA.120.047049.

Wells, H.L., Letko, M., Lasso, G., Ssebide, B., Nziza, J., Byarugaba, D.K., Navarrete-Macias, I., Liang, E., Cranfield, M., Han, B.A., et al. (2021). The evolutionary history of ACE2 usage within the coronavirus subgenus Sarbecovirus. Virus Evol. *7*, 7. https://doi.org/10.1093/ve/veab007.

Williams, S.J., Wreschner, D.H., Tran, M., Eyre, H.J., Sutherland, G.R., and McGuckin, M.A. (2001). MUC13, a novel human cell surface mucin expressed by epithelial and hemopoietic cells. J. Biol. Chem. *276*, 18327–18336. https://doi.org/10.1074/jbc.M008850200.

Wilson, L.A., Gemin, A., Espiritu, R., and Singh, G. (2005). Ets-1 is transcriptionally up-regulated by H2O2 via an antioxidant response element. FASEB J. *19*, 2085–2087. https://doi.org/10.1096/fj.05-4401fj.

Xia, H., Suda, S., Bindom, S., Feng, Y., Gurley, S.B., Seth, D., Navar, L.G., and Lazartigues, E. (2011). ACE2-Mediated reduction of oxidative stress in the central nervous system is associated with improvement of autonomic function. PLoS One *6*, e22682. https://doi.org/10.1371/journal.pone.0022682.

Yinda, C.K., Port, J.R., Bushmaker, T., Owusu, I.O., Purushotham, J.N., Avanzato, V.A., Fischer, R.J., Schulz, J.E., Holbrook, M.G., Hebner, M.J., et al. (2021). K18-hACE2 mice develop respiratory disease resembling severe COVID-19. PLoS Pathog. *17*, e1009195. https://doi.org/10.1371/journal.ppat.1009195.

Yue, F., Cheng, Y., Breschi, A., Vierstra, J., Wu, W., Ryba, T., Sandstrom, R., Ma, Z., Davis, C., Pope, B.D., et al. (2014). A comparative encyclopedia of DNA elements in the mouse genome. Nature *515*, 355–364. https://doi.org/10.1038/nature13992.

Zeng, W., Miyazato, A., Chen, G., Kajigaya, S., Young, N.S., and Maciejewski, J.P. (2006). Interferon-γ-induced gene expression in CD34 cells: identification of pathologic cytokine-specific signature profiles. Blood *107*, 167–175. https://doi.org/10.1182/blood-2005-05-1884.

Zhang, H., Penninger, J.M., Li, Y., Zhong, N., and Slutsky, A.S. (2020a). Angiotensin-converting enzyme 2 (ACE2) as a SARS-CoV-2 receptor: molecular mechanisms and potential therapeutic target. Intensive Care Med. *46*, 586–590. https://doi.org/10.1007/s00134-020-05985-9.

Zhang, J., Xie, B., and Hashimoto, K. (2020b). Current status of potential therapeutic candidates for the COVID-19 crisis. Brain Behav. Immun. *87*, 59–73. https://doi.org/10.1016/j.bbi.2020.04.046.

Zhao, Q., Meng, M., Kumar, R., Wu, Y., Huang, J., Lian, N., Deng, Y., and Lin, S. (2020). The impact of COPD and smoking history on the severity of COVID-19: a systemic review and meta-analysis. J. Med. Virol. *92*, 1915–1921. https://doi.org/10.1002/jmv.25889.

Zheng, J., Li, G., Chen, S., Bihl, J., Buck, J., Zhu, Y., Xia, H., Lazartigues, E., Chen, Y., and Olson, J.E. (2014). Activation of the ACE2/Ang-(1-7)/Mas pathway reduces oxygen-glucose deprivation-induced tissue swelling, ROS production, and cell death in mouse brain with angiotensin II over-production. Neuroscience *273*, 39–51. https://doi.org/10.1016/j.neuroscience.2014.04.060.

Zhou, P., Yang, X.L., Wang, X.G., Hu, B., Zhang, L., Zhang, W., Si, H.R., Zhu, Y., Li, B., Huang, C.L., et al. (2020). 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.

Ziegler, C.G.K., Allon, S.J., Nyquist, S.K., Mbano, I.M., Miao, V.N., Tzouanas, C.N., Cao, Y., Yousif, A.S., Bals, J., Hauser, B.M., 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.e19. https://doi.org/10.1016/J.CELL.2020.04.035.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Antibodies		
1:1000 mouse anti-SARS/SARS-CoV-2 N	ThermoFisher Scientific	MA5-29981; RRID: AB_2785780
1:1000 rabbit anti-beta-actin	Abcam	ab8227; RRID: AB_2305186
mouse anti-ACE2	R&D Systems	MAB933; RRID: AB_2223153
1:5000 donkey anti-rabbit 800	LI-COR Biosciences	926-32213; RRID: 621848
1:5000 goat anti-mouse 680	LI-COR Biosciences	Catalogue number: 925-68070; RRID: AB_2651128
Bacterial and virus strains		
SARS-CoV-2	Clinical isolate	SARS-CoV-2/SB3
Biological samples		
Chemicals, peptides, and recombinant proteins		
Human recombinant interferon alpha 2	OriGene Technologies	TP723881
Recombinant human IFNβ1	ThermoFisher Scientific	R69007
Critical commercial assays		
Direct-zol RNA Miniprep kit	Zymo Research Corporation	R2072
Applied Biosystems Power SYBR master mix	ThermoFisher Scientific	4368577
Lipofectamine 2000	Invitrogen	D5210
SuperScript III First Strand cDNA Synthesis Reaction kit	Life Technologies	18090010
Deposited data		
Human reference genome NCBI build 37, GRCh37	Genome Reference Consortium	https://www.ncbi.nlm.nih.gov/projects/genome/assembly/grc/human/
See Table S1 for list of accessions for publicly-available microarray expression datasets used in this study.		
NRF2 ChIP-seq dataset	Chorley et al., 2012	GEO GSE37589
ENCODE DNase-seq datasets used	Yue et al., 2014	ENCFF271JAF, ENCFF331SYD, ENCFF681UOZ, ENCFF165ZIA, ENCFF334RSR, ENCFF338GII, ENCFF446FTN, ENCFF460ZFL, ENCFF546QUZ, ENCFF644XOI, ENCFF957JQC
JASPAR Database	Mathelier et al., 2016	https://jaspar.genereg.net/
ChIP-ATLAS	Oki et al., 2018	https://chip-atlas.org/
Supplemental information	This study; Mendeley Data	https://doi.org/10.17632/wmv6f24xm2.1
Experimental models: Cell lines		
Calu-3	ATCC	HTB-55
Oligonucleotides		
See Table S2 for all PCR primers.		
See Table S2 for all oligonucleotides used in CRISPR experiments.		
Recombinant DNA		
PX458 plasmid	Addgene	101731
Software and algorithms		
Bedtools version 2.26.0	Quinlan and Hall, 2010	https://bedtools.readthedocs.io/en/latest/
HOMER version 4.10	Heinz et al., 2010	http://homer.ucsd.edu/homer/

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
R base version 3.6.3	R Development Core Team, 2008	https://www.r-project.org/
COVID-19 Genes	Butler et al., 2021, Park et al., 2022	https://covidgenes.weill.cornell.edu/
GEOquery version 2.52.0	Sean and Meltzer, 2007	https://www.bioconductor.org/packages/release/bioc/html/GEOquery.html
sva version 3.32.1	Leek et al., 2012	https://www.bioconductor.org/packages/release/bioc/html/sva.html

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources should be directed to and will be fulfilled by the lead contact, Terence D. Capellini (tcapellini@fas.harvard.edu).

Materials availability

This study did not generate new unique reagents.

Data and code availability

- Accessions for publicly-available datasets used in this study are described in previous publications and in the key resources table.

- This paper does not report original code. Code used to generate figures is available upon reasonable request from the lead contact.

- Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

- Additional Supplemental Items are available from Mendeley Data: https://data.mendeley.com/datasets/wmv6f24xm2/1.

Richard, Daniel (2022) "Intronic regulation of SARS-CoV-2 receptor (ACE2) expression mediated by immune signaling and oxidative stress pathways - Richard et al., 2022", Mendeley Data, V1: https://doi.org/10.17632/wmv6f24xm2.1.

METHOD DETAILS

ACE2 co-expression and functional enrichment analysis with public microarray data

Public microarray experiments using Affymetrix chips (HuGene-1.0-st-v1 and HG-U133 Plus 2) on airway epithelial cell samples were obtained from the NCBI Gene Expression Omnibus (GEO) database. This resulted in a total of 1859 individual samples from 33 different experiments (See Table S1). Within this dataset, disease status (Healthy: 504, COPD: 338, Asthma: 136) and/or smoking status (Never: 409, Former: 139, Current: 956) information was included for 1716 samples. For all datasets, raw intensity values and annotation data were downloaded using the *GEOquery* R package (version 2.52.0) (Sean and Meltzer, 2007) from the Bioconductor project. Probe definition files were downloaded from Bioconductor and probes were annotated using Bioconductor's *annotate* package. All gene expression data were unified into a single dataset that was then RMA-normalized, and only genes present in both of the Affymetrix platforms (N = 16,105) were kept for subsequent analyses. Correction of experiment-specific batch effects was performed using the ComBat method implemented using the *sva* R package (version 3.32.1) (Leek et al., 2012).

Top *ACE2* co-expressed genes were identified based on the 200 highest Pearson correlation (*r*) values to *ACE2*. Heat maps for top 200 *ACE2*-correlated genes across samples were generated with the *pheatmap* R package (version 1.0.12). For display only, expression values were row-normalized (across gene) using the 'scale' function in base R, and converted to Z-scores. For heatmap coloring, a "ceiling" and "floor" of +3 and −3 was applied to the Z-scores. Histograms and *ACE2* correlation scatterplots were generated with the *base* R package (version 3.6.3) (R Development Core Team, 2008). Gene lists are available in Table S1.

The top 200 *ACE2*-correlated genes were analyzed using Enrichr(71) to identify enriched pathways and functional ontologies. Terms were ranked within ontologies by FDR-adjusted p value (calculated by Enrichr by running the Fisher exact test for random gene sets in order to compute a mean rank and standard deviation from the expected rank for each term in the gene set library) and the top 3 terms for ontologies of interest were selected. Functional enrichment bar plots were generated with the *ggplot2* R package (version 3.2.1).

To parse transcriptional datasets from SARS-CoV-2 infected samples, we made use of the aggregated meta-analysis dataset and analyses performed by Butler et al. (2021) and Park et al. (2022) (Butler et al., 2021; Park et al., 2022). We queried the 'COVID19 Genes' online interface hosted by the Mason lab at Cornell University (https://covidgenes.weill.cornell.edu/) for immune-response genes detailed in our results section (those with strong correlated expression with ACE2), finding that all of these genes were significantly elevated in CoV-2 infected samples relative to their respective dataset controls. Given our finding of interferon signaling as an enriched pathway within our set of top 200 ACE2-correlated genes, we used the 'Pathway Enrichment' tool from the web interface, using the term 'REACTOME_INTEFERON_SIGNALING'. The heatmaps shown in Figure S1 were generated from the 'Pathway Heatmap, WCM NP' and 'Pathway Heatmap, WCM Autopsy Lung' output tabs.

ACE2 regulatory region analyses

ENCODE (Yue et al., 2014) DNase-seq datasets were obtained for adult lung (File accessions: ENCFF271JAF, ENCFF331SYD, ENCFF681UOZ) and primary cell (ENCFF165ZIA, ENCFF334RSR, ENCFF338GII, ENCFF446FTN, ENCFF460ZFL, ENCFF546QUZ, ENCFF644XOI, ENCFF957JQC) samples as processed hg19 bed-formatted files. Called peaks (putative regulatory elements) were subsequently intersected within lung/cell sets using *bedtools* (Quinlan and Hall, 2010) (version 2.26.0), requiring that a peak be replicated in at least two samples for retention. To capture a broader regulatory region around the *ACE2* gene, peaks falling within 1MB upstream/downstream of the *ACE2* promoter were collected and pooled across lung/cell sets. Human hg19 reference sequences were obtained for these elements using UCSC (Kent et al., 2002).

Sequences were subsequently used for *de novo* motif analysis using HOMER (version 4.10) (Heinz et al., 2010) using a 10x random shuffling as a background set. *De novo* motifs were compared to a vertebrate motif library included with HOMER which incorporates the JASPAR (Mathelier et al., 2016) database. Matches are scored using Pearson's correlation coefficient of vectorized motif matrices (PWMs), with neutral frequencies (0.25) substituted for non-overlapping (e.g., gapped) positions. Best-matching motif PWMs obtained from this analysis are shown in Figure 4B. The highest-rank motif bore close similarity to the *FOSL2::JUN* PWM from the JASPAR database (MA1130.1). This reference PWM was subsequently used for motif scanning. The AME program (McLeay and Bailey, 2010), part of the MEME-Suite (Bailey et al., 2009), was also used with these sequences to look for enrichments of known transcription factor (TF) motifs. Focusing on elements intragenic to *ACE2* as more proximate candidate regulatory regions, these reference sequences were also tested for enriched motifs using AME. The *FOSL2::JUN* (MA1130.1) PWM was used to scan these sequences using TFBSTools (Tan and Lenhard, 2016) (version 1.24.0). Hits of minimum sequence scores of 80% were retained, and subsequently filtered for Benjamini-Hochberg adjusted p value <0.05. For illustrative purposes, the best-scoring hit for each hit-containing element was selected. JASPAR-database motif matrices were also obtained for *STAT1* (MA0137.3), and *IRF1* (MA0050.2) and similarly used to scan intragenic *ACE2* elements using TFBSTools as described. Given the observed expression data for *NRF2*, a ChIP-seq dataset for this factor (GSE37589) (Chorley et al., 2012) was downloaded as called hg18 peaks. These were lifted-over to hg19 using the UCSC liftOver utility (Kent et al., 2002) with sulforaphane and vehicle-treatment datasets pooled and merged for a final set of 919 peaks. Reference hg19 sequences were obtained for these peaks and used with HOMER to define an *NRF2 de novo* motif; the resulting PWM was subsequently used with TFBSTools to scan the putative *ACE2* intragenic regulatory sequences as described above. ChIP-seq data for indicated factors (*IRF1*, *STAT1*, *STAT2*, *FOS* and *JUN*) were obtained from ChIP-ATLAS (Oki et al., 2018) as an aggregate across all cell types using a significance threshold of 50 (*q* value < 1 × 10^{-5}). Peak files (hg19) were sorted and merged for overlapping peaks using bedtools. DNase, ChIP-seq, and motif hit tracks were loaded into the UCSC Genome Browser (Karolchik et al., 2013) for visualization.

In-vitro experimental validation studies

Poly(I:C) transfection and IFN treatment

Calu-3 cells were mock transfected with 4 μL of lipofectamine 3000 (ThermoFisher Scientific) in Opti-MEM (ThermoFisher Scientific) only or transfected with 100 ng of poly(I:C) (InvivoGen) for 6 h. Recombinant human IFNβ1 was generated using *Drosophila* Schneider 2 (S2) cells following manufacturer's recommendation and by using ThermoFisher Scientific's *Drosophila* Expression system (ThermoFisher Scientific). Recombinant IFNβ1 was collected as part of the cell culture supernatant from S2 cells and total protein was measured using Bradford assay. Total protein concentration was used for subsequent experiments. To demonstrate that S2 cell culture media did not contain non-specific stimulators of mammalian antiviral responses, we also generated recombinant green fluorescent protein (GFP) using the same protocol and used the highest total protein concentration (2 mg/mL) for mock treated cells. SARS-CoV-2 infected cells were treated with supernatant containing IFNβ1 or GFP for 6 h.

SARS-CoV-2 infection

Clinical isolate of SARS-CoV-2 (SARS-CoV-2/SB3) was propagated on Vero E6 cells and validated by next generation sequencing (Banerjee et al., 2020a). Virus stocks were thawed once and used for an experiment. A fresh vial was used for each experiment to avoid repeated freeze-thaws.

Immunoblots

Calu-3 cells were seeded at a density of 3×10^5 cells/well in 12-well plates. Cells were infected with SARS-CoV-2 at an MOI of 1. Control cells were sham infected. Twelve hours post incubation, cells were transfected or treated with poly(I:C) or IFNβ, respectively for 6 h. Cell lysates were harvested for immunoblots and analyzed on reducing gels as mentioned previously (Banerjee et al., 2020b, 2021). Briefly, samples were denatured in a reducing sample buffer and analyzed on a reducing gel. Proteins were blotted from the gel onto polyvinylidene difluoride (PVDF) membranes (Immobilon, EMD Millipore) and detected using primary and secondary antibodies. Primary antibodies used were: 1:1000 mouse anti-SARS/SARS-CoV-2 N (ThermoFisher Scientific; Catalog number: MA5-29981; RRID: AB_2785780), 1:1000 rabbit anti-beta-actin (Abcam; Catalog number: ab8227; RRID: AB_2305186), and 2 μg/mL of mouse anti-ACE2 (R&D Systems; Catalog: MAB933; RRID: AB_2223153). Secondary antibodies used were: 1:5000 donkey anti-rabbit 800 (LI-COR Biosciences; Catalogue number: 926-32213; RRID: 621848) and 1:5000 goat anti-mouse 680 (LI-COR Biosciences; Catalogue number: 925-68070; RRID: AB_2651128). Blots were observed and imaged using Image Studio (LI-COR Biosciences) on the Odyssey CLx imaging system (LI-COR Biosciences).

Cell line and culture condition

The human lung adenocarcinoma cell line, Calu-3 (ATCC HTB-55) were grown in Minimum Essential Medium (MEM)-α (Gibco, Gaithersburg, Maryland) supplemented with 10% fetal bovine serum (FBS), and 1% penicillin-streptomycin (P/S) in 5% CO_2 at 37°C. Media was replaced every 2 days and the cells were subcultured every 5 days. The cells were passaged and used in experimental assays without additional STR authentication or mycoplasma testing.

CRISPR targeting of *ACE2* regulatory elements *in vitro*

All sgRNAs flanking human *ACE2* regulatory elements and sub-elements containing TF binding sites were designed using the Genetic Perturbation Platform (GPP) sgRNA design tool from Broad Institute (https://portals.broadinstitute.org/gpp/public/), synthesized by Integrated DNA Technologies, Inc (Coralville, Iowa), and cloned into the PX458 vector following published protocols (Ran et al., 2013). The sequence of all sgRNAs along with their chromosomal locations (hg19) are listed in Table S2.

Guide RNAs (see Table S2), flanking the A*CE2* regulatory elements and sub elements containing TF binding sites, were first tested for the ability to induce efficient deletions of the enhancer elements/sub elements in cultured Calu-3 cells (N = 3 biological replicates per assay). Calu-3 cells were maintained in MEM α media supplied with 10% FBS (Gibco) and 1% Pen/Strep (0.025%) and seeded in a six-well plate for 1- day prior to transfection. After culturing at 37°C with 5% CO_2, the cells were scanned for GFP fluorescence under GFP-microscope at 24 h to verify successful transfection efficiency (i.e., >70% of the cells were GFP positive). After 48 h of CRISPR experiment, DNA was extracted from the CRISPR-cas9 targeted Calu-3 cells using E.Z.N.A Tissue DNA Kit (Omega Bio-Tek, Norcross, GA), and the *ACE2* regulatory element/sub-element region was amplified using PCR primers flanking each sgRNA location (listed in

Table S2). PCR amplified products were purified from 1% agarose gel using E.Z.N.A Gel Extraction Kit (Omega Bio-Tek, Norcross, GA). Sanger sequencing was used to verify successful deletion of the target region. To examine effects on *ACE2* and nearby gene expression (*TMEM-27* and *BMX1*), RNA was extracted from control and CRISPR-Cas9 targeted Calu-3 cells (N = 3 biological replicates, with 3 technical replicates per experiment per condition) and prepared using Trizol Reagent (Thermo Fisher Scientific, Springfield Township, New Jersey) and Direct-zol™ RNA Miniprep kit (ZYMO). Two micrograms of total RNA were used to synthesize first-strand cDNA using Super-Script III First-Strand Synthesis System (Thermo Fisher Scientific). qRT-PCR analysis was then performed with gene specific primers and Applied Biosystems Power SYBR master mix (Thermo Fisher Scientific) with *ACTB* house-keeping gene as an internal control. sgRNAs and primers used for qRT-PCR are listed in Table S2.

CRISPR deletion of *ACE2* regulatory elements under interferon treatment

Briefly, the Calu-3 cells were cultured in MEM α media supplemented with 10% FBS (Gibco) and 1% Pen/Strep (0.025%), plated in 6-well plates and utilized at ~70% confluence. The cells were then subjected to CRISPR-Cas9 mediated deletion of *ACE2* regulatory elements/sub-elements in the presence/absence of recombinant protein of human interferon, alpha 2 (IFNA2) (OriGene Technologies Inc, Atlanta, GA) (100 ng/mL) for 48 h. Following CRISPR-deletion under interferon treatment, DNA and RNA were extracted from the Calu-3 cells and used for genotyping and gene expression analysis, respectively.

CRISPR deletion of *ACE2* regulatory elements under H_2O_2 treatment

The Calu-3 cells were cultured in MEM α media as described above and subjected to CRISPR-cas9 mediated deletion of *ACE2* regulatory elements/sub elements for 48 h. CRISPR-cas9 targeted Calu-3 cells were treated with hydrogen peroxide (H_2O_2) using the protocol described previously (Boardman et al., 2004). Following CRISPR deletion of *ACE2* regulatory elements/sub elements, the Calu-3 cells were challenged with or without oxidative stress by exposure to 0.5 mM (initial dose) of H_2O_2. As Calu-3 cells rapidly metabolize H_2O_2 in 1 h, H_2O_2 treatments were performed for 2 h, with additional bolus of H_2O_2 every 60 min for times longer than 1 h. DNA and RNA were extracted from the CRISPR-cas9 targeted Calu-3 cells subjected to H_2O_2 treatment, and used for genotyping and gene expression analysis, respectively.

QUANTIFICATION AND STATISTICAL ANALYSIS

For CRISPR targeting experiments, all genes were quantified by real-time PCR using gene specific primers. *ACE2* gene expression data was normalized relative to *ACTB* house-keeping gene expression and compared between control and experimental condition (e.g., putative enhancer deletion or putative enhancer deletion in the presence/absence of IFN-α and H_2O_2 treatment). Normalization of *ACE2* gene expression and relative fold change calculation were performed by $2^{-\Delta\Delta CT}$ method of Livak and Schmittgen (Livak and Schmittgen, 2001). All data are presented as the mean $\pm$ SEM, or in boxplot form indicating median, upper and lower quartiles. Individual pairwise comparisons between control and experimental condition were analyzed by two-sample, two-tailed Student's T-test unless otherwise noted, with $p < 0.05$ regarded as significant. N numbers listed in figure legends (N = 9 biological replicates per comparison).