# A novel isoform of *ACE2* is expressed in human nasal and bronchial respiratory epithelia and is upregulated in response to RNA respiratory virus infection

Cornelia Blume(1,2)*, Claire L Jackson(1,2,3)*, Cosma Mirella Spalluto(1,4), Jelmer Legebeke(5), Liliya Nazlamova(1), Franco Conforti(1,2), Jeanne-Marie Perotin-Collard(2,6), Martin Frank(7), John Butler(8), Max Crispin(3,8), Janice Coles(1,2), James Thompson(1,2), Robert A Ridley(1,2), Lareb S N Dean(1,2), Matthew Loxham(1,2), Stephanie Reikine (9), Adnan Azim(1,2), Kamran Tariq(1,2), David A Johnston(1,10), Paul J Skipp(3,8), Ratko Djukanovic(1,2), Diana Baralle(2,5), Chris McCormick(1,2), Donna E Davies(1,2,3)#, Jane S Lucas(1,2,3)#, Gabrielle Wheway(3,5)#, Vito Mennella(1,2,3)#

*Joint first authors

#corresponding authors d.e.davies@soton.ac.uk; jlucas1@soton.ac.uk; g.wheway@soton.ac.uk; v.mennella@soton.ac.uk

**Affiliations**

**1.** School of Clinical and Experimental Sciences, University of Southampton Faculty of Medicine, Southampton, UK.

**2.** Southampton NIHR Biomedical Research Centre, University Hospital Southampton NHS Foundation Trust and University of Southampton, Southampton, UK.

**3.** Institute for Life Sciences, University of Southampton, Southampton, UK

**4.** Wessex Investigational Sciences Hub, Faculty of Medicine, University of Southampton Faculty of Medicine, Southampton, UK

**5.** School of Human Development and Health, University of Southampton Faculty of Medicine, Southampton, UK.

**6.** Service de Pneumologie, INSERM U1250, Hôpital Universitaire, Reims, France

**7.** Biognos AB, Generatorsgatan 1, Box 8963, 40274 Göteborg, Sweden

**8.** School of Biological Sciences, University of Southampton Faculty of Environmental and Life Sciences Southampton, UK.

**9.** Nuclera Nucleics Ltd, 137 Cambridge Science Park, Milton Road, Cambridge, UK

**10.** Biomedical Imaging Unit, Faculty of Medicine, University of Southampton, Southampton, UK

**Subject Terms:**

Health sciences/Medical research/Genetics research

Health sciences/Medical research/Experimental models of disease

Biological sciences/Molecular biology/Transcriptomics

**Data availability:**

The RNA sequencing datasets analysed during the current study (Figure 1, Figure 3, Supplementary Figure 1 and Supplementary Figure 6) are available in the Sequence Read Archive repository, Accession: PRJNA650028 ID: 650028 https://www.ncbi.nlm.nih.gov/bioproject/650028 (Figure 1, Figure 3, Supplementary

Figure 1) and SubmissionID: SUB8455806 BioProject ID: PRJNA674784 (Supplementary Figure 6).

## Abstract

Angiotensin-converting enzyme 2 (ACE2) is the main entry point in the airways for SARS-CoV-2. ACE2 binding to SARS-CoV-2 protein Spike triggers viral fusion with the cell membrane, resulting in viral RNA genome delivery into the host. Despite ACE2's critical role in SARS-CoV-2 infection, an understanding of ACE2 expression, including in response to viral infection, remains unclear.

Until now *ACE2* was thought to encode five transcripts and one 805 amino acid protein. Here we identify a novel short isoform of ACE2. Short *ACE2* is expressed in the airway epithelium, the main site of SARS-CoV-2 infection; it is substantially upregulated in response to interferon stimulation and RV infection, but not in response to SARS-CoV-2 infection, and it shows differential regulation in asthma patients. This short isoform lacks SARS-CoV-2 spike glycoprotein high-affinity binding sites and altogether, our data are consistent with a model where short *ACE2* is unlikely to directly contribute to host susceptibility to SARS-CoV-2 infection.


## Introduction

With more than 45 million confirmed cases of COVID-19 and more than a million associated deaths worldwide (WHO; 4$^{th}$ November 2020), there is an urgent need to understand the molecular mechanism of infection and disease to identify patients' susceptibility and targets for therapeutic intervention. A key molecule responsible for SARS-CoV-2 viral entry is the metalloprotease angiotensin-converting enzyme 2 (ACE2), a transmembrane protein encoded by the human *ACE2* gene. *ACE2* consists of 19 exons and encodes five annotated transcripts, two of which encode the same 805 amino acid protein ACE2 (UniProt Q9BYF1) with predicted molecular

mass of 92.4 kDa and observed mass of ~120 kDa due to multiple sites of glycosylation of the N-terminal region[1]. ACE2 consists of an N-terminal extracellular domain and a C-terminal membrane anchor domain[1]. The extracellular domain contains a 17-amino acid signal peptide sequence, an N-terminal catalytic metallopeptidase domain with 41% amino acid sequence identity to ACE[1,2] and a C-terminal domain with 48% amino acid sequence identity to renal amino acid transporter collectrin (TMEM27)[3]. ACE2 extracellular domain can be shed via proteolytic cleavage of specific residues in the ferredoxin-like fold domain (neck dimerization domain) by proteases including ADAM17, TMPRSS11D or TMPRSS2. TMPRSS2 expression and activity has been shown to increase SARS-CoV-2 viral entry[4,5].

ACE2 is the main viral entry point for coronavirus N63, SARS-CoV and SARS-CoV-2, which cause severe-acute respiratory syndromes, the latter being responsible for COVID-19 in humans[6-9]. ACE2 binds to the S1 domain of trimeric SARS-CoV Spike (S) glycoprotein[6] and SARS-CoV-2 Sprotein[10], which is primed by TMPRSS2[11]. Cellular entry of SARS-CoV is dependent upon the extracellular domain of ACE2 being cleaved by TMPRSS2 protease at Arg697 and Lys716, and the transmembrane domain of ACE2 internalised with the virus via the clathrin-mediated[12,13] and clathrin-independent[14] endocytosis pathways.

ACE2 is a carboxypeptidase with several known physiological functions including regulation of blood pressure, salt and water balance in mammals[1,2] amino acid uptake in the small intestine[15,16] and glucose homeostasis and pancreatic beta-cell function[17,18]. Interestingly, ACE2 has been suggested to play an important role in protection from acute lung injury. Ace2 protein levels are reduced in mouse models of acid driven acute lung injury, although *Ace2* gene expression has not been

investigated in this model[19]. It is unclear whether the lower ACE2 protein levels are due to downregulation of *Ace2* gene expression or due to death of *Ace2*-expressing pneumocytes. Interestingly, *Ace2* knockout mice show a more severe acute lung injury phenotype[19]. Moreover, improved outcomes are seen in pig models of lung injury in which *ACE2* is overexpressed[20] and an activator of *ACE2*, XNT, can protect against pulmonary hypertension in rat models[21,22]. Although the molecular mechanism by which ACE2 protects against acute lung injury remains unclear, it is known that its carboxypeptidase function is required to confer this protection and that AT2 (AngII receptor 2) also confers protection[19].

*ACE2* expression in different tissues is controlled by multiple promoter elements[23] including Ikaros homology activating elements around -516/-481 in the heart[24] and under the control of estrogen responsive elements in adipose tissue[25]. In human nasal epithelia and lung tissue ACE2 expression has been reported to be interferon (IFN) regulated with evidence of STAT1, STAT3, IRF8, and IRF1 binding sites between −1500–500 bp within the *ACE2* promoter[26]. Activation of IFN responsive genes is an important antiviral defence pathway in humans, and both interferon and influenza exposure have been reported to increase *ACE2* expression in human airway[26].

Bulk RNA sequencing data[27] detects low-level expression of *ACE2* in testis, small intestine, thyroid, colon, kidney, heart left ventricle and atrial appendage, and visceral adipose. Single cell RNA sequencing (scRNAseq) studies show *ACE2* expression at low levels in airway, cornea, esophagus, ileum, colon, liver, gallbladder, heart, kidney and testis[28]. Using scRNAseq and RNA *in situ* hybridisation, *ACE2* expression in the airways has been observed to be relatively high in nasal epithelium and progressively lower in the bronchial and alveolar

regions[29]. *ACE2* expression correlates with levels of infection of SARS-CoV-2 isolates from patients in different airway compartments[29]. Consistently, SARS-CoV-2 viral loads have been found to be higher in swabs taken from the nose than swabs taken from the throat of COVID-19 patients[30]. Highest *ACE2* expression is seen in goblet and ciliated cells of the nasal epithelium[28], and ACE2 protein localises to the membrane of motile cilia of respiratory tract epithelia[31]. Consistent with this, SARS-CoV-2 has been detected *in situ* in pulmonary pneumocytes, ciliated airway cells and upper airway epithelium in COVID-19 patients examined by autopsy[32]. Airway multiciliated cells appear to be one of the main targets of SARS-CoV-2 infection[29], and it has been demonstrated that SARS-CoV infection occurs through ACE2 on cilia in airway epithelia[31]., possibly because their extension from the cell surface makes ACE2 more accessible to the virus. Altogether, these studies have established the upper airway as the main site of SARS-CoV-2 infection.

Here we detail the identification of a novel isoform of ACE2, which we name short ACE2, that is expressed in human nasal and bronchial respiratory epithelia, the main site of SARS-CoV-2 infection, and is preferentially expressed in asthmatic bronchial epithelium relative to full length ACE2 (long ACE2). In primary airway cells, short *ACE2* is upregulated in response to IFN treatment and infection with rhinovirus, but not SARS-CoV-2.

## Results

### *Identification of novel short ACE2 transcript in nasal and bronchial airway cells*

We analysed the expression of *ACE2* in airway epithelia in our existing RNAseq datasets from nasal brushings and nasal epithelia cultured at air-liquid interface (ALI) by aligning reads to human genome build 38 using STAR 2-pass mapping[33] and

GENCODE v33 gene annotations. We visually analysed mappings to *ACE2* using Integrative Genomics Viewer (IGV)[34], which identified multiple reads mapping to a genomic region between exon 9 and 10 of the constitutive *ACE2* gene build (**Figure 1a**). These mappings showed a discrete 3' junction at GRCh38 chrX:15580281, but variable 5' length suggesting a splice junction with downstream exon 10, but no splicing upstream to exon 8. This suggests that a novel unannotated exon exists between exon 9 and 10, and that this exon is the beginning of a novel transcript distinct from full-length *ACE2* transcripts ACE2-202 (ENST00000427411.1) or ACE2-201 (ENST00000252519.8) in the airway. Evaluation of read mappings to *ACE2* gene in IGV also showed approximately double the number of read support to exons 10-19 compared to exons 1-9 (**Figure 1a**), further suggesting that a novel shorter transcript of *ACE2*, which includes a novel exon plus exons 10-19, is expressed at equal or higher levels than longer *ACE2* transcripts including exons 1-19 (ACE2-202) or 2-19 (ACE2-201). Assembly of transcriptomes from all samples using SCALLOP tool identified novel transcripts including this novel exon to exon 19 (**Figure 1b**). Sashimi plot analysis confirmed splicing between this new exon and downstream exon 10, but complete absence of splicing at the 5' end of the new exon (**Figure 1c**). Analysis of these RNAseq data with code developed by Cummings *et al.*[35] also independently detected a novel splice junction at chrX:15580281. Analysis of splice junctions identified by STAR aligner confirmed multiple uniquely-mapped reads to a novel exon/exon boundary removing a novel intron of coordinates GRCh38 chrX:15578316–15580280. We named this novel exon 9a. Study of the sequence of the exon 9a/intron boundary showed a strong U1-dependent consensus splice site sequence (AG|GTAAGTA) suggesting that it is a strong splice donor site (**Figure 1d**). This splicing event introduces a new in-frame ATG start codon 29

nucleotides upstream of the splice site (**Figure 1d**), and a TATA box 148 nucleotides upstream of the splice site (**Figure 1d**) suggesting that this transcript is protein-coding. Furthermore, a promoter flanking region has been identified at GRCh38 chrX:15581200-15579724 (ENSR00000902026), suggesting active transcription upstream of exon 9a (approx. chrX:15580402 – chrX:15580281). Analysis of this region shows a near consensus ISGF-3 binding site (TgGTTTCAgTTTCCt)[36] 159bp upstream of the splice junction, a near-consensus AP-1 binding site (TGtGTCA)[37] 223bp upstream of the splice site and an NF-kB binding site (GGGTTTTCCC)[38] 787bp upstream of the splice junction. This suggests that the short form of ACE2 is under independent transcriptional control from full-length ACE2 expression, and that this may be controlled by IFN, AP-1 and NF-kB elements.

Expression of this novel transcript was confirmed by RT-PCR using primers specific to exon 1, exon 9a and exon 19 and cDNA from both nasal brushings and differentiated immortalized bronchial epithelial cells BCi-NS1.1 which differentiate robustly into airway multiciliated cells[39] (**Figure 2a, b**). Sanger sequencing confirmed the identity of these PCR amplicons and confirmed sequences spanning exon 9 and 10, and exon 9a and 10 in the amplicons from the constitutive transcript and novel transcript, respectively (**Figure 2c**).

To investigate expression of this novel *ACE2* transcript relative to full-length *ACE2* transcripts (ACE-202 and ACE-201) we quantified reads mapped to exon9a/exon10 junction and reads mapped to exon9/exon10 from the STAR alignment output file and calculated exon 9a inclusion rates relative to inclusion of exon 9. This analysis showed that in nasal epithelia the mean expression level of short and long *ACE2* transcripts was 0.745 (reads mapped to exon9a/10 or exon9/10 per million mapped reads) and relative inclusion of exon 9a was 0.763 (st err 0.0829) identifying that

short *ACE2* expression is significantly higher than long *ACE2* in nasal epithelial cells ($p<0.05$, Student's t-test, n=6). We then designed specific qPCR primers to amplify the short and long transcripts of *ACE2* individually, as well as primers to amplify both transcripts and quantify total levels of *ACE2* expression, and determined the dynamic range of these qPCRs (**Supplementary Figure 2a**) and consistency of our SYBR-green based qPCRs and Taqman probe-based PCRs (**Supplementary Figure 2b**). Expression of long *ACE2* was confirmed in a number of cell lines and primary airway cells, with highest expression being observed in the Vero E6 cell line, differentiated BCi-NS1.1 cells and *in vitro* differentiated nasal epithelial cells grown at ALI, with expression comparable to that observed in *ex vivo* nasal epithelial cells (**Figure 3a,b**). Expression of short *ACE2* was low in Vero E6, HEK293, Caco2, RPE1, H441 and 16HBE cells, and this contrasted with differentiated BCi-NS1.1 cells and *ex vivo* or *in vitro* differentiated nasal cells which exhibited high expression of this novel isoform. Both isoforms were also expressed robustly in *ex vivo* or *in vitro* differentiated primary bronchial cells (**Figure 3a,b**). Assessment of *ACE2* isoform induction during differentiation of nasal epithelial cells grown at ALI *in vitro* showed very low *ACE2* expression on day 0, with expression of both isoforms reaching levels comparable with those observed in primary nasal brushings at day 4 of ALI culture when the start of cilia gene transcription is observed (usually days 4-7), being maintained until day 37, and reducing at day 84 as the cultures became senescent (**Figure 3c and Supplementary Figure 1**). This is consistent with published work showing that *ACE2* expression (and SARS-CoV-2 infection) is dependent on airway epithelial cell differentiation[40]. Given reports that bronchial epithelial cells express lower levels of *ACE2* than nasal cells[26], we also compared expression of the long and short isoforms of *ACE2* in these two cell types from

multiple donors of primary tissue. Consistent with previous reports, total levels of *ACE2* were lower in bronchial epithelial cells, which was due to reduced expression of both long and short forms of *ACE2* (**Figure 3d**).

### *Short ACE2 transcript is expressed in multiple tissues*

Transcript-specific probe-based qPCR on cDNA from a multiple tissue control panel showed robust expression of the long transcript of *ACE2* in all tissues tested except whole brain (**Figure 3e**). The short transcript of *ACE2* was detected in all tissues except whole brain and skeletal muscle although expression level was low in most tissues, with highest expression in lung and kidney (**Figure 3e**). Together these data show that we identified a novel *ACE2* transcript that is expressed in multiple tissues including kidney, and airway suggesting a significant role in these compartments.

### *Novel short* ACE2 *transcript is predicted to encode a protein which lacks most of the SARS-CoV-2 binding interface*

We next sought to investigate whether this novel *ACE2* transcript can be translated into a protein product. We identified a TATA box and in-frame ATG start codon in the new exon suggesting that this transcript would produce a 459 amino acid protein consisting of Arg357 – Phe805 of the full-length long ACE2 protein isoform plus an additional 10 amino acid sequence (M-R-E-A-G-W-D-K-G-G) before Arg357 and lacking the secretion signal sequence present in long ACE2. This is predicted to produce a protein of 52.7 kDa which includes the C-terminal 449 amino acids of long ACE2, but lacks the 356 N-terminal residues. Consistent with our expectations, data mining of proteomics datasets in the public domain also identified sequences corresponding to the novel 10 amino acid peptide at the N-terminus of short ACE2 (M-R-E-A-G-W-D-K-G-G; M-R-E-A-G-W-D-K) in colon, breast and ovarian cancer

proteomes[41] and BLASTP analysis showed that no annotated proteins show 100% homology to this peptide in mammalian species. BLASTN analysis also showed that the 30 nucleotide sequence encoding the N-terminus specific sequence of short ACE2 was not found in any other part of the human genome, consistent with the observed peptide being derived from the short form of ACE2.

We modelled the structure of the predicted translation product of this short transcript of *ACE2*, based on the structure of full-length long ACE2 protein in complex with the receptor binding domain of SARS-CoV-2 resolved by cryo-EM (PDB 6M17)[9]. This analysis highlighted the extent of loss of the SARS-CoV-2 binding region in the predicted protein product of short *ACE2,* with many residues previously shown to be important for viral binding not present in this short ACE2 protein (**Supplementary Figure 3a-c**)[9]. In particular, predicted short ACE2 lacks two entire regions involved in interaction with SARS-CoV-2 spike glycoprotein (aa 30-41 and aa 82-84), including a high-affinity binding site (aa 30-41), and retains only a few residues of a third region involved in this interaction[9]. This latter sequence is replaced by the N-terminal specific sequence of short ACE2, which is predicted to form a disordered/helical secondary structure by PEP-fold, compared to the beta sheet present in long ACE2 further modifying the third binding interface to Spike. Short ACE2 however retains the HEXXH catalytic motif (aa 374-378) and the sequences required for cleavage by ADAM17, TMPRSS11D and TMPRSS2. While the majority of the spike binding domain of ACE2 is missing from short ACE2, the neck/ferredoxin-like fold domain (residues 616 to 726) which is the most important dimerization interface[9] and transmembrane region are present. Assuming that the homologous parts of short ACE2 fold in the same way as full length ACE2, molecular

dynamic simulation of a short ACE2 homodimer suggested that it may form a stable structure (**Supplementary Figure 3d-f; Supplementary Video 1**).

***Short ACE2 encodes a novel protein that is expressed in differentiated airway epithelia***

To investigate whether the predicted short ACE2 protein isoform is expressed in airway epithelial cells and other cell types, we performed western blotting analysis of cell lysates using multiple antibodies to ACE2 recognising epitopes on different regions of the protein (**Figure 4a**). Initially, we tested an antibody raised to the C-terminal domain (CTD) of ACE2 (Abcam 15348), which we anticipated would recognise a common epitope in long and short ACE2. Western blotting of lysates prepared from seven cell lines including Vero cells, which are used for infection assays of SARS-CoV[42] and SARS-CoV-2[43] identified two bands at 100 and 120kDa, consistent with the presence of glycosylated and non-glycosylated forms of full length (long) ACE2 protein[6] (see below). We also detected an additional band at ~52 kDa, the expected molecular weight of short ACE2, in differentiated BCi NIS1.1 cells (**Figure 4b**) and fully differentiated primary nasal and bronchial epithelial cultures (**Figure 4c**). Expression of this protein in other cell lines was low, consistent with the qPCR data (**Figure 4b**). Lack of correlation between long ACE2 protein levels and the intensity of the 52 kDa band suggests that the latter is not a degradation product of long ACE2. Preadsorption of the anti ACE2 CTD antibody with the immunizing peptide, but not a peptide with similar charge, blocked detection of both long and short ACE2 isoforms in the airway cells, confirming specific detection of short ACE2 (**Figure 4c**).

To orthogonally validate antibody specificity, and examine glycosylation of the ACE2 isoforms, we performed additional western blot analyses on cell lysates of both nasal and bronchial epithelial ALI cultures, as well as Vero E6 cells before and after treatment with PNGase F, an enzyme that removes N-linked oligosaccharides from glycoproteins. This confirmed that the ~120 kDa band detected by the anti ACE2 CTD antibody was glycosylated ACE2 whereas the ~100 kDa band was non (or partially) glycosylated ACE2; the mobility of the ~52 kDa band corresponding to short ACE2 did not change suggesting it is not N-glycosylated. Similar results were obtained with an antibody raised against the ectodomain of ACE2 (anti ACE2 ECTO amino acids 18-740), which recognised glycosylated and non-glycosylated isoforms of long ACE2, as well as short ACE2. In contrast, an antibody (anti ACE2 NTD) raised against amino acids 200-300, which are only present in the N-terminal region of long ACE2 recognised glycosylated and non-glycosylated isoforms of long ACE2 but, as expected, it did not recognise short ACE2 (**Figure 4d**). To examine the localisation of the ACE2 isoforms in airway epithelial cells, we performed immunofluorescent staining of differentiated ALI cultures of primary bronchial epithelial cells with anti-ACE2 antibodies visualised by confocal microscopy (**Figure 4e, Supplementary Figure 4a-d**). The antibodies recognising common epitopes in short and long ACE2 showed localisation mainly to the cellular apical regions and motile cilia. Staining with the third antibody that detects only long ACE2 (anti ACE2 NTD antibody) was too weak to interpret.

Thus, to further investigate whether short ACE2 localises to cilia as has been reported for full-length ACE2[44] we purified motile cilia from BCi-NS1.1 cells Consistent with our previous experiments, western blotting of whole cell lysates with anti ACE2 CTD antibody showed a distinct band around 52 kDa in addition to full

length ACE2 (**Figure 4f**). Notably, the band corresponding to short ACE2 was not enriched in the cilia fraction albeit still present in detectable amounts (**Figure 4f**) suggesting that it is predominantly localised to the apical regions of the cells. Densitometric analysis confirmed enrichment of long ACE2 in cilia fraction relative to short ACE2 (**Figure 4g**).

To confirm that the novel short *ACE2* transcript can support protein translation, we first performed *Escherichia coli* cell-free protein synthesis assays comparing expression of the ectodomains of long and short ACE2. These experiments confirmed that both constructs were expressed at comparable levels, with little evidence of degradation (**Supplementary Figure 5a**). In contrast, when GFP-tagged expression constructs were transfected into HEK293, RPE1 or Vero E6 cells, we observed expression of long ACE2-GFP but no expression of short ACE2-GFP, suggesting that within a cellular context, short ACE2 is an unstable protein. However, when we used the H441 airway cell line, short ACE2 expression was consistently observed, albeit at a lower level than long ACE2 (**Supplementary Figure 5b**). Together, these data suggest that short ACE2 is translated into a protein product but that its stability within target cells may be cell-type dependent.

***Novel short ACE2 isoform is upregulated by interferon and rhinovirus (RV) infection***

To begin investigating the functional relevance of short *ACE2* transcript expression, we first assessed whether it is an IFN stimulated gene. As expected, treatment of bronchial epithelial cells with Type I, II or III IFNs caused upregulation of *MX1* and *IP10* (**Supplementary Figure 6a and b**) and upregulation of total *ACE2* (**Figure 5a**) which was largely due to an effect on short *ACE2* rather than long *ACE2* (**Figure**

**5a**). In these experiments, IFN-β>IFN-γ and IFN-λ>IFN-α for induction of short *ACE2* expression (**Figure 5a**); we also noted inhibitory effects of some of the IFNs on expression of long *ACE2* (**Supplementary Figure 6c**), with the result that only IFN-β caused a significant change in the ratio of short *ACE2* relative to long *ACE2* expression (**Figure 5a, Supplementary Figure 6d**). Based on the greater potency of IFN-β, we confirmed induction of short *ACE2* in differentiated bronchial epithelial cells treated with IFN-β (**Supplementary Figure 6e and f**).

These data led us to evaluate the response to viral infection, as it has been reported that *ACE2* expression is upregulated in this condition[26]. We exposed nasal epithelial cells grown at ALI to rhinovirus (RV) and harvested cells 24hr after infection. qPCR analysis showed a significant upregulation of both long and short *ACE2* expression relative to UV-RV treated control (* p=<0.001, non-parametric Wilcoxon Signed Rank test, control vs. RV16 n=11) (**Figure 5b**). While this is consistent with previously published work showing that *ACE2* is upregulated in response to influenza exposure[26], it is notable that we found that it was the short *ACE2* transcript which was upregulated more robustly than long *ACE2* transcript (around 9-fold increase in expression of short *ACE2* compared to around 2.5-fold increase in long *ACE2* expression) (**Figure 5b**). Parallel experiments using bronchial epithelial cells infected with RV, confirmed induction of short *ACE2* but had no significant effect on long *ACE2* (**Figure 5c**).

As long ACE2 has been described as a point of entry for SARS-CoV-2, and that SARS-CoV-2 infection stimulates an increase in *ACE2* expression, we sought to investigate the effect of SARS-CoV-2 infection on the expression of the short transcript of *ACE2*. Differentiated BCi-NS1.1 cells were infected with SARS-CoV-2 and harvested at 1 and 72 hours after infection. qPCR amplification of the CoVN1

nucleocapsid gene expression confirmed significant infection by 72 hours (**Figure 5d**), however we did not observe significant induction of either the long or short isoforms of *ACE2*, or total *ACE2* (**Figure 5e**). As *ACE2* is IFN-regulated, it is likely that lack of *ACE2* induction at this time point is due to the known ability of SARS coronaviruses, including SARS-CoV-2, to inhibit both IFN expression and downstream signalling from the Type I IFN receptor[45-48]. Consistent with these previously published results, we did not observe significant induction of the interferon-stimulated genes, *MX1* or *CXCL10* in the SARS-CoV-2 infected cultures (**Figure 5f**). This was not due to an inability of the BCi-NS1.1 cells to respond to viral infection with IFN production, as RV16 infection led to induction of *MX1* and *IP10*, which was one-two orders of magnitude greater than that seen with SARS CoV-2 (**Supplementary Figure 6g and h**).

As we observed that short *ACE2* was not induced following SARS-CoV-2 infection where IFN expression and signalling are compromised[45-48], we next examined its induction in differentiated bronchial epithelial cultures from severe asthmatic donors which are known to have reduced Type I and Type III IFN responses to RV infection[52,53]. We observed a significant increase in expression of total *ACE2* and short *ACE2* but not long *ACE2* when differentiated bronchial epithelial cultures from either healthy control subjects or people with severe asthma were infected with RV16 (**Supplementary Figure 7a and b**). This resulted in a significant increase in the ratio of short *ACE2* relative to long *ACE2* (**Figure 5g and Supplementary Figure 7c**); however, for the severe asthma-derived cultures this increase was significantly less than observed in cultures from healthy donors (**Figure 5g**). The RV16-induced increase in short *ACE2* expression was positively correlated with

expression of *IFNB1 (***Figure 5h**) (**r=0.58, p=0.003**), whereas there was no similar relationship for long *ACE2* expression (**r=0.05, p=0.80**). Furthermore, as we observed previously[52], RV-induced IFN-λ (IL-29/28) secretion was lower in the severe asthma-derived cultures (**Supplementary Figure 7d**) and levels of IFN-λ were significantly correlated with expression levels of short *ACE2* (**r=0.69, p=0.016**) but not long *ACE2*, (**r=-0.23, p=0.47**) (**Supplementary Figure 7e**).

Finally, as *ACE2* expression has been reported to be increased in some severe asthmatic patients who exhibit type 1 and 2 IFN signatures as a phenotypic trait[54], we also analysed expression of long and short ACE2 in bronchial brushings from healthy controls and patients with severe asthma who were free of respiratory virus infections. This showed statistically significantly lower expression of total *ACE2* and long *ACE2* in patients with severe asthma, but no significant difference in short *ACE2* expression between groups (**Supplementary Figure 7f**). As a consequence, the ratio of short *ACE2* to long *ACE2* was significantly higher in severe asthma patients (**Supplementary Figure 7g**). In these severe asthmatic patients, the IFN-stimulated genes, *OAS1* and *ISG15*, were significantly increased (**Supplementary Figure 7h**), and short *ACE2* expression levels significantly correlated with IFN-γ protein levels in bronchoalveolar lavage fluid (**Supplementary Figure 7j**) suggesting that variation in expression of *ACE2* isoforms *in vivo* may reflect the inflammatory status of the asthmatic subject, independently of a virus-induced exacerbation.

## Discussion

Here we present identification and characterisation of a short 11 exon transcript of human *ACE2*, consisting of a novel previously unannotated first exon, which we name exon 9a, and exons 10-19 of the long *ACE2* transcript. We show that, whilst the long transcript of *ACE2* is expressed in multiple tissues, this short *ACE2* transcript is expressed predominantly in airways. We show highest expression in primary respiratory epithelia, most notably in the nasal epithelium where the level of short transcript expression is higher than long *ACE2* transcript expression. Expression of both transcripts of *ACE2* is dependent on differentiation of epithelial cells, with levels of expression comparable with primary nasal epithelia from days 4 – 63 of differentiation at ALI culture. We show that expression of this short transcript is regulated independently of the long transcript of *ACE2*, with putative promoter elements identified upstream of the transcriptional start site of the short *ACE2* transcript. We confirm that this short transcript is translated into a protein product of around 52 kDa, which is not glycosylated in differentiated airway epithelial cells. Although the transcript lacks a signal peptide, recombinant protein expression studies confirm that the short *ACE2* transcript can be translated into a protein product *in vitro*, however within cells its expression appears to be relatively unstable. As is the case for other unstable proteins, this may be because the standard cell lines used for expression studies are not equipped to accomplish the required post-translational modifications or molecular folding of this molecule. In parallel with our studies, two independent reports identified the novel *ACE2* transcript by extensive mining of public datasets[55,56], yet neither study detected endogenous protein expression of the short ACE2 isoform in a range of cell lines or undifferentiated primary bronchial epithelial cells. Thus, detailed studies of short ACE2 function may

require use of airway cell systems, especially differentiated cell models where short ACE2 appears more abundant, perhaps reflecting a stabilizing environment.

Sequence analysis shows that the novel ACE2 isoform contains a transmembrane domain, collectrin homology domain and portions of the ACE homology domain but lacks the 356 most N-terminal amino acids of full-length ACE2, with 10 novel amino acids in their place. Most of the key residues required for SARS-CoV-2 spike binding are absent from this isoform suggesting that short ACE2 is not a viral entry point for SARS-CoV-2. Consistent with this, one of the two studies referred to above[56] reported successful expression of a GFP-tagged construct of short ACE2 (termed dACE2) which achieved membrane localisation but failed to bind to SARS-CoV-2 spike protein. While the same study also reported that the GFP-tagged recombinant short ACE2 isoform does not cleave angiotensin II, our structural modelling indicates that short ACE2 retains the catalytic motif conferring metallopeptidase function., However, as short ACE2 lacks of the N-terminal residues of long ACE2 which are known to affect substrate specificity (especially R273)[57], short ACE2 may retain enzymatic activity but most likely with different substrate specificity.

Previous studies have reported that *ACE2* is an IFN-regulated gene and have suggested that SARS-CoV-2 could exploit IFN-driven upregulation of *ACE2* to enhance infection[26]. In our study, we show that it is the short transcript of *ACE2* that is more strongly induced by Type I, II and III IFNs and by viral infection than long *ACE2*, findings similar to two recent reports[55,56]. In our study, we performed detailed comparative dose response studies and found that short *ACE2* was not only upregulated by IFN-$\beta$ and IFN-$\lambda$ reflecting innate epithelial responses to infection, but also by IFN-$\gamma$ and, to a much lesser extent, IFN-$\alpha$ suggesting a further contribution from co-resident immune cells such as T-cell and plasmacytoid dendritic

cells *in vivo*. While we observed a small increase in long *ACE2* expression in RV infected nasal epithelial cells, it seems unlikely that this was due to virus-induced IFN production, as we observed either no effect or an inhibitory effect of IFNs on long *ACE2* expression. Given that viral infection also induces expression of cytokines such as IL-1β, which has been reported to induce *ACE2* expression[58] it seems more likely that cytokines other than IFNs induce long *ACE2*, the viral receptor form of ACE2.Given the absence of key residues required for SARS-CoV-2 spike binding and the evidence that short ACE2 cannot bind SARS-CoV-2 spike protein[56] it is unlikely that IFNs have a detrimental effect in the airways by promoting SARS-CoV-2 entry as has been suggested previously[26]. Such a conclusion is supported by a recent clinical trial of inhaled IFN-β in hospitalised patients who showed greater odds of clinical improvement and recovery, as well as reduced breathlessness compared with placebo[59], thus highlighting the antiviral benefits of IFN-β.

While patients with asthma have reduced susceptibility to SARS-CoV-2, and asthma symptoms are not exacerbated by SARS-CoV-2 infection[49-51], it has been reported that *ACE2* expression is higher in bronchial epithelium of a subset of asthma patients with Type-2 low disease and characteristics resembling known risk factors for severe COVID-19 (male sex, history of hypertension, low peripheral blood, and elevated bronchoalveolar lavage lymphocytes)[54]. As this increase in *ACE2* expression was found to be associated with upregulation of viral response genes, it was suggested that therapies targeting the IFN family may be of benefit in this subset of patients[54]. Through analysis of *ACE2* isoforms, our data show that long *ACE2* expression is lower in severe asthmatic subjects compared to healthy controls, even though we were able to identify up-regulation of certain IFN-responsive genes in the asthma group. In contrast, expression of short ACE2 was comparable between the groups

and there was a higher ratio of expression of short *ACE2* : long *ACE2* in severe asthma. These differences in *ACE2* isoform expression might be explained by the epithelial changes observed in asthma and by specific inflammatory subtypes of the disease, although an effect of treatment with corticosteroids cannot be excluded. Thus, in asthma there is goblet cell metaplasia resulting in a reduction in the number of ciliated cells[60] which are the site of expression of both isoforms of *ACE2* while elevated levels of the type 2 cytokine, IL-13, are reported to suppress *ACE2* expression[61], possibly by driving goblet cell metaplasia. Based on the correlation between short *ACE2* and levels of IFNγ in bronchoalveolar lavage fluid, we conclude that those asthmatic subjects with IFN signatures would show preferential up-regulation of short *ACE2* rather than long *ACE2* and that this is unlikely to directly increase susceptibility to SARS CoV-2 infection. Although severe asthmatic subjects receive inhaled or oral corticosteroids as part of their regular care it is not known whether corticosteroids directly modulate *ACE2* expression. However, it has been reported that corticosteroids suppress Type I IFN signalling[62] suggesting that they have the potential to affect short *ACE2* expression in the context of a respiratory viral infection.

Although the function of the short *ACE2* isoform remains unclear, our data clearly show that this isoform is expressed in the airways, particularly in ciliated cells of the nasal and bronchial epithelium. Of note, ACE2 is an homologue of ACE which also utilizes different promoters to produce two distinct isoforms, the full length molecule and a lower molecular mass variant which is found only in testis (tACE)[63] where its expression regulates sperm capacitation[64]. Like the motile cilia of the airways, sperm flagella have a characteristic 9+2 axoneme structure suggesting some important

function of the shorter ACE2 and ACE isoforms, respectively, in relation to these structures. However, we did not find evidence that short ACE2 was enriched in cilia, which contrasts with the long form of ACE2. Instead, we found that short ACE2 was retained within the cell body along with a fraction of long ACE2. Therefore, it is interesting that short *ACE2* is preferentially upregulated in response to viral infection, independently of long *ACE2* expression. We hypothesise that this short isoform of ACE2 plays an important physiological role in the airway and, in addition that it may influence host susceptibility to SARS-CoV-2 infection. While it does not appear to bind SARS CoV-2 spike directly, it retains the sequences required for cleavage by ADAM17, TMPRSS11D and TMPRSS2 and so may compete for other membrane proteases required for viral entry. Furthermore, it retains the neck/ferredoxin-like fold domain which is the most important dimerization interface[9] and so may heterodimerise with long ACE2 (or other accessory proteins) to prevent its trafficking to the exposed tips of the cilia and/or influence the spike binding interaction. In support of this hypothesis, constitutive homodimerization of sACE and tACE has been demonstrated, in addition to heterodimerization of both sACE and tACE with AT2R[65]. Of note, we failed to see induction of short *ACE2* in response to SARS-CoV-2 infection of differentiated airway epithelial cells, a model which closely mimics the route of infection in humans. This finding is similar to that reported for SARS-CoV-2 infection of Calu3 lung adenocarcinoma cells and is distinct from responses in intestinal epithelial cells where short *ACE2* was induced by SARS-CoV-2 infection[56]. These differences may reflect differences in infection levels and/or the ability of SARS-CoV2 to suppress IFN expression and signalling[45-48], however the possibility of cell type differences requires further consideration, especially as the airways are the primary route of SARS CoV-2 infection.

The discovery of short ACE2 may have significant consequences for design of therapeutic approaches targeting ACE2 to tackle COVID-19 [66,67] and has implications for the numerous studies reporting on ACE2 expression levels and differences in levels of expression along airways, across age groups and disease groups[29,54,68,69] including COVID-19 disease severity[69,70]. This finding should be considered when selecting reagents for future studies of *ACE2* expression in tissues relevant to SARS-CoV-2 viral infection, bearing in mind the genomic region targeted by primer sets and the epitopes recognised by antibodies. For example, this may help to resolve the current debate over the relevance of conjunctiva as a site of SARS-CoV-2 viral entry, with IHC and protein studies suggesting expression of ACE2 but some RNA studies disputing this[71-74].

These data emphasise the remaining gaps in our understanding of full complexity of the human transcriptome, particularly cell-specific transcriptomes, and the power of transcript-level analysis of deep bulk RNA sequence data in resolving some of this complexity[75]. Deep, bulk RNA sequence analysis of isolated cell types followed by transcript level analysis, including transcript expression quantification and differential splicing analysis has the power to revolutionise understanding of disease.

In conclusion, we have identified short ACE2, a novel isoform of ACE2 that has lost the majority of the SARS-CoV-2 spike binding sites but retains the main dimerization domain. It is predominantly expressed in differentiated airway epithelial cells, especially in cells of the upper airways which are the main site of SARS-CoV-2 infection. Our data suggests that the transcript encodes a 52 kDa protein which can be detected in airway epithelial cells, although because it lacks a signal peptide, it may be a relatively unstable protein. We demonstrate that it is this isoform, rather than full length ACE2 that is IFN-regulated and inducible upon rhinovirus infection.

However, in conditions of IFN suppression, as observed during SARS-CoV-2 infection, or IFN deficiency, as in asthma, short *ACE2* is not induced to the same degree as normal. While the function of short ACE2 is unknown, its regulation by IFN suggests it may play an essential role in innate anti-viral defence mechanisms in the airways.

## Materials and Methods

*Collection of airway samples*

Nasal epithelial cells were isolated by brushing the inferior turbinate with a sterile 3.0 mm cytology brush (Conmed). Cells were processed into RNAlater for subsequent RNAseq analysis or were stored in liquid nitrogen prior to cell culture. Bronchial epithelial cells were harvested by bronchoscopic brushings for primary bronchial epithelial cell culture. To assess inflammatory status some individuals underwent bronchial-alveolar lavage (BAL) at time of bronchoscopy. Cytokines in BAL fluid supernatants were analysed using a V-PLEX Proinflammatory Panel 1 Human Kit (Meso Scale Diagnostics). Airway samples for the study were collected following approval by South-Central Hampshire A, Research Ethics Committee, UK (reference numbers: 07/Q1702/109, 13/SC/0182 and 14/WM/1226) and all participants gave their informed consent.

*Cell culture*

Nasal cells were expanded and differentiated at air-liquid interface (ALI) culture as previously described[76]. All cultureware were pre-coated in 1:10 diluted PureCol collagen (5005-B CellSystems) throughout each step. Cells were cultured using Pneumacult Ex-Plus (Stemcell Technologies) and at ALI on 6.5mm 0.4µm polyester membrane transwell permeable supports (Corning Life Sciences) for up to 84 days

590 using Pneumacult ALI media (Stemcell Technologies), with apical surface washed
591 (HBSS) and medium changes 3 times weekly. Cell were cultured in 100% relative
592 humidity, 5% $CO_2$ at 37$^{o}$C. First ciliation was observed by microscopy from day 7 at
593 ALI and maintained until harvested. Primary nasal epithelial cells were fully
594 differentiated and ciliated by 28 days.

595 Primary bronchial epithelial cells were expanded in Airway Epithelial Cell Growth
596 medium (Promocell) up to passage 1 as previously described[77]. At passage 2 cells
597 were either cultured submerged as monolayers or differentiation was induced by
598 plating cells on 6.5mm 0.4µm polyester membrane transwell permeable supports
599 (Corning Life Sciences) and differentiated at ALI for 21 days. Transepithelial
600 electrical resistance was monitored weekly using an EVOM Voltohmmeter (World
601 Precision Instruments) and cells with a TER ≥330WΩm2 on day 21 were used for
602 experiments.
603 hTERT transformed bronchial epithelial cell line BCi-NS1.1, provided by Walters et
604 al[39,78] were expanded in PneumaCult-Ex Plus Basal Medium (Stem Cell
605 Technologies) supplemented with Pneumacult Ex Plus supplements (Stem Cell
606 Technologies), hydrocortisone, nystatin and penicillin/streptomycin. BCi-NS1.1 cells
607 were grown at air-liquid interface in PneumaCult-ALI Basal Medium (Stem Cell
608 Technologies) supplemented with Pneumacult ALI supplement (Stem Cell
609 Technologies), hydrocortisone, PneumaCult ALI maintenance supplement, heparin,
610 nystatin and penicillin/streptomycin. BCi-NS1.1 cell ciliation was observed by
611 microscopy and cells were fully differentiated and ciliated by 42 days at ALI.
612 Vero E6 (ECACC Vero C1008) cells were cultured in DMEM medium supplemented
613 with 10% FBS and penicillin/streptomycin (Gibco) at 37°C with 5% $CO_2$. When 70-

80% confluent (every 5-7 days) cells were passaged by washing with HBSS before detaching with 0.2% trypsin EDTA.

NCI-H441 [H441] (ATCC HTB-174) cells were cultured in RPMI 1640 medium (Gibco) supplemented with 10% FCS, sodium pyruvate, L-glutamine and penicillin/streptomycin at 37°C with 5% $CO_2$ and passaged every 3-4 days. Cells were washed with modified HBSS with calcium and magnesium (HyClone) before detaching with 0.2% trypsin EDTA.

hTERT RPE-1 (ATCC CRL-4000) were cultured in DMEM/F12 medium (Gibco) supplemented with 10% FCS at 37°C with 5% $CO_2$ and passaged every 4-6 days.

293 [HEK-293] (ATCC® CRL-1573™) were cultured in DMEM high glucose supplemented with 10% FCS at 5% $CO_2$ and passaged every 4-6 days at a ratio of 1:8.

*RNA extraction and quality control for RNAseq*

RNA was extracted from nasal brushings and primary nasal epithelial cells grown at ALI using RNeasy Plus Mini kit (Qiagen). RNA was extracted from epithelial brushings using miRNeasy Mini Kit and RNase-Free DNase Set (Qiagen). RNA quality and concentration were measured using an RNA Nano chip on the Bioanalyzer 2100 (Agilent). Samples with total RNA concentration ≥20ng/µl, RIN ≥9.6 and OD 260/280 were taken forward for cDNA library preparation and sequencing.

*cDNA library preparation, sequencing and data quality control*

cDNA libraries from primary nasal epithelial brushings and primary nasal epithelial cells grown at ALI were prepared using Ribo-Zero Magnetic Kit for rRNA depletion and NEBNext Ultra Directional RNA Library Prep Kit library prep kit by Novogene Inc. cDNA libraries from primary bronchial brushings were prepared using NEBNext

Ultra (non-stranded) mRNA library prep kit with polyA pulldown for mRNA enrichment. Library quality was assessed using a broad range DNA chip on the Agilent Bioanalyser 2100. Library concentration was assessed using Qubit and q-PCR. Libraries were pooled, and paired-end 150bp sequencing to a depth of 20M-100M reads per sample was performed on an Illumina HiSeq2500 by Novogene Inc. Raw FASTQ reads were subjected to adapter trimming and quality filtering (reads containing N > 10%, reads where >50% of read has Qscore<= 5) by Novogene Inc. Quality of sequence was assessed using FastQC v0.11.5 (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). No further data filtering or trimming was applied. Raw FASTQ reads after adapter trimming and quality filtering (reads containing N > 10%, reads where >50% of read has Qscore<= 5) were deposited on the Sequence Read Archive, SRA accession to be provided upon publication.

*Alignment to reference genome and quality control*

Paired FASTQ files were aligned to GRCh38 human genome reference using GENCODE v33 gene annotations and STAR v2.6.0a splice aware aligner[33], using ENCODE recommended options (3.2.2 in the STAR manual (https://github.com/alexdobin/STAR/blob/master/doc/STARmanual.pdf). The two-pass alignment method was used.

Alignment files were assessed for saturation of known splice junctions using RSeqQC v3.0.1[79].

*Transcriptome assembly*

Unique transcripts were assembled from merged alignment files, and a merged transcriptome reference formed from the unique transcripts and GENCODE v33 reference transcriptome using SCALLOP tool v0.10.5[80].

*Alignment to reference transcriptome and transcript level abundance estimates*

SALMON tool v1.3.0[81] was used to perform transcript abundance estimates from raw FASTQ files using selective alignment with a decoy-aware transcriptome assembled using Scallop tool. Integrative genome viewer (IGV) v.2.3.93[34] was used to visualise alignment files.

*Differential splicing analysis*

A Mendelian RNA-seq method for identifying and filtering splice junctions developed by Cummings *et al.* [35] was used to detect aberrant and novel splice events. No changes were made to this code. The individual sample splice junction discovery output files were combined into an overall splice junction discovery file used for splice junction normalisation.

*RNA extraction and cDNA production*

From nasal and bronchial brushings, cDNA was synthesised from excess RNA purified for RNAseq using High Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific) following manufacturer's instructions.

From cell lines, RNA was isolated from cell lysates using standard phenol-chloroform extraction, and reverse transcribed to cDNA using a Precision Reverse Transcription kit (PrimerDesign, Southampton, UK) according to the manufacturer's instructions. From BCi-NS1.1, RNA was isolated from cell lysates using standard QIAzol extraction, and reverse transcribed to cDNA using High Capacity cDNA Reverse Transcription kit (Thermo Fisher Scientific) following manufacturer's instructions.

*Long-range RT-PCR and Sanger sequencing*

Phusion High-Fidelity PCR Master Mix with HF Buffer (NEB) was used to amplify the novel and annotated transcript using custom primers (IDT) from cDNA produced from nasal brushings and BCi-NS1.1 cells. Manufacturer's instructions and

recommended thermocycling conditions were followed, with annealing temperature (64°C) calculated using NEB Tm calculator. RT-PCR Primer sequences: For amplifying short ACE transcript (exon 9a - exon 19): forward 5'-3' ATTGAGGAGAGCTCTGAGGC, reverse 5'-3' TCTCTCCTTGGCCATGTTGT. For amplifying long ACE2 transcript (exon 1 - exon 19): forward 5'-3' TGCTAACGGACCCAGGAAAT, reverse 5'-3' TCTCTCCTTGGCCATGTTGT. Samples were size separated against Hyperladder 1kb (BioLine). Gel extracted PCR products were sequenced using forward and reverse PCR primers at 3.2μM by Source Biosciences. Electropherograms were visualised using 4Peaks.

*RT-qPCR*

cDNA from cell cultures and from human multiple tissues control cDNA panel I (TakaraBio) was amplified by qPCR (cycling conditions 95 ºC 10 min, then 50 cycles of 95 ºC 15 s, 60 ºC 1 min) using *ACE2* primer pairs (see below). Data were normalised to the geometric mean of the housekeeping genes (ubiquitin C and glyceraldehyde 3-phosphate dehydrogenase, probe-based duplex primer mix, PrimerDesign) and fold change in gene expression relative to controls was determined using the ΔΔCt method. Probe-based primers were used to determine expression levels of human MX1 (DD-hu-600-MX1, PrimerDesign) IP10 (PP-hu-900-IP10, PrimerDesign) and IFNB1 (DD-hu-600-IFNB1, PrimerDesign) (see below for sequences). To detect SARS-CoV-2 in BCi-NS1.1, Taqman gene expression assays were used against 2019-nCoV_N1 (primers sequences from Public Health Service Centers for Disease Control and Prevention (CDC), 20 January 2020 copy) normalised to the genes HPRT, 18S and RNase P expressions using the dCt method.

qPCR primer sequences: For amplifying long *ACE2* transcript only: Forward 5'-3' CAAGAGCAAACGGTTGAACAC, Reverse 5'-3' CCAGAGCCTCTCATTGTAGTCT (from Harvard PCR primer bank). For amplifying short *ACE2* transcript only: forward 5'-3' GTGAGAGCCTTAGGTTGGATTC, reverse 5'-3' TAAGGATCCTCCCTCCTTTGT. For amplifying both transcripts: forward 5'-3' TGGGACTCTGCCATTTACTTAC, reverse 5'-3' CCCAACTATCTCTCGCTTCATC

Probe-based qPCR: Total *ACE2* primers to exon 17/18 boundary plus FAM-MGB probe (Thermo Fisher Scientific, cat Hs01085333_m1, sequences proprietary); long *ACE2* primers to exon 2/3 boundary plus FAM-MGB probe (Thermo Fisher Scientific, cat Hs01085335_m1, sequences proprietary); short *ACE2* primers to exon9a/10 boundary forward 5'-3' GTGAGAGCCTTAGGTTGGATTC, reverse 5'-3' TAAGGATCCTCCCTCCTTTGT (IDT) plus FAM-MGB probe 5'-3' TCATTGAGGAGAGCTCTGAGGCAGA (Thermo Fisher Scientific). IP10 primers forward 5'-3' CAGAGGAACCTCCAGTCTCAG, reverse 5'-3' GGTACTCCTTGAATGCCACTTA, probe 5'-3' ACTGCGATTCTGATTTGCCTTATCTTTCTGtcgcagt (PrimerDesign). IFNB1 primers forward 5'-3' TTACTTCATTAACAGACTTACAGGT, reverse 5'-3' TACATTAGCCATCAGTCACTTAAAC, probe 5'-3' CCTCCGAAACTGAAGATCTCCTAGCCTGTGCCaagtttcg. MX1 primers forward 5'-3' CCCCAGTAATGTGGACATCG, reverse 5'-3' ACCTTGTCTTCAGTTCCTTTGT, probe 5'-3' CGTCAACATTCCGATGGTCCTGTCTCCCTCttgacg.

*Electrophoresis gels*

All gels were 1.5% agarose in 1x TAE buffer and staining with ethidium bromide (Sigma, 5 µL per 50 ml gel). Gels were run for 75 minutes at 90 V. PCR products were loaded with 6X purple gel loading dye (B7025, NEB) and electrophoresed

738 alongside a low molecular weight ladder (range 25 to 766 base pairs, N3233, NEB)
739 or HyperLadder 1 kb (range 200 bp to 10 kb, BIO-33053, Bioline) to determine
740 product sizes.

741 *Cilia extraction*

742 Cilia were extracted from differentiated ALI cultures on 24-Transwell inserts (Costar)
743 following a protocol modified from[82]. Cells on ice were washed with ice cold PBS,
744 then incubated on the apical surface for 15 minutes in 100 µL washes with deciliation
745 buffer (20 mM Tris hydrochloride (pH 7.5), 0.05 M sodium chloride, 10 mM calcium
746 chloride, 1 mM EDTA, 7 mM 2-mercaptoethanol and 0.1% triton X-100)[83] containing
747 additional protease inhibitor cocktail (Sigma) 10uL/ml buffer. Washes were pooled
748 and centrifuged for 2 min at 1000xg to pellet. Supernatant fractions were centrifuged
749 at 16,000xg for 8 mins. Cilia pellets were frozen before Western blot procedure.
750 Immunofluorescence labelling confirmed cilia enrichment, and detachment of cilia on
751 ALI membranes. Briefly, ice cold methanol fixed and 4% dried milk blocked cilia
752 pellets and deciliated cell membranes (excised from Transwell inserts) were labelled
753 for 1 hour at room temperature with a mouse anti-alpha-tubulin antibody (T9026,
754 Sigma) diluted 1:500 in PBST. Following three PBST washes a secondary goat anti-
755 mouse Alexa488 antibody (Molecular Probes) was incubated for 30 minutes at room
756 temperature before PBST washes. Deciliated cells on membranes were additionally
757 DAPI stained before mounting. Cilia pellets and membranes were mounted in
758 Mowiol between two glass coverslips and imaged using a Leica SP8 laser scanning
759 confocal microscope and LAS X software.

760 *Western blot*

761 Cells were washed in phosphate-buffered saline (PBS) or HBSS and lysed in 20 mM
762 Tris-HCl pH 8.0, 137 mM NaCl, 1% (w/v) NP-40, 2mM EDTA supplemented with
763 cOmplete™ protease inhibitor (Sigma). Samples were diluted with 10x denaturing
764 buffer (5% SDS, 400mM DTT) and 6x non-reducing SDS sample buffer (Boston
765 BioProducts), incubated at 60C for 10min and separated on an SDS-PAGE gel and
766 transferred to polyvinylidene fluoride membranes (BioRad). After blocking in 5%
767 milk/TBST membranes were probed with primary anti-ACE2 antibody (Abcam 15348
768 for total ACE2 or Novus #NBP2-67692 for long ACE2) followed by the appropriate
769 secondary HRP-conjugated antibody (Dako). Bound antibody was detected using
770 Clarity ECL Western Blotting Substrate (Bio-Rad) with the image digitally captured
771 using an Amersham Imager 600 (GE Healthcare Life Sciences). HRP-conjugated
772 anti-β-actin antibody (Sigma) was used as a loading control. ImageJ was used for
773 densitometry.

774 *Peptide blocking assay*

775 A 1:500 dilution of ab15348 was incubated with 10ug/ml immunizing peptide
776 (CKGENNPGFQNTDDVQTSF) or a control peptide (GMEHLREVRAVTSANIQEF)
777 for 2h at RT with agitation before incubating the blots at 4C overnight under rotation.

778 *PNGase F treatment*

779 PNGase F (NEB) was used to remove N-linked oligosaccharides from glycoproteins
780 following manufacturer's instructions.

781 *Immunofluorescence cell staining*

782 After apical wash with HBSS, cells were fixed with 4% PFA, permeabilized with 0.1%
783 Triton X-100 and blocked with 1% BSA in PBS. Membranes were cut from the
784 inserts and epithelial cells were stained with anti-ACE2 antibodies (ab15348
785 (Abcam), AF933 (R&D systems) and #NBP2-67692 (Novus)), anti-alpha tubulin

(T9026 Sigma), and appropriate fluorescently labelled secondary antibodies (Alexa-488 labelled anti-mouse (Invitrogen), Alexa649 labelled anti-goat (Abcam), DyLight 647 labelled anti-rabbit (Biolegend)). Actin filaments were stained using Alexafluor-555 phalloidin (Cytoskeleton Inc) and nuclei with DAPI. Confocal images were taken using a Leica SP8 laser scanning confocal microscope with LAS X software.

*Antibodies*

| Antigen | Species | Immunogen | Manufacturer and product code |
|---|---|---|---|
| ACE-2 | goat polyclonal | Gln18-Ser740 | R&D AF933 (lot HOK0320041) |
| ACE-2 | rabbit polyclonal | C-terminal domain aa 788-805 | Abcam ab15348 (lot GR3333640-/) |
| ACE-2 | rabbit polyclonal | N-terminal domain aa 200-300 | Abcam ab108252 (GR3338009-3) |
| Beta actin | mouse monoclonal | anti-β-actin-peroxidase | clone AC-15, Sigma, A3854-200UL |
| Alpha tubulin | mouse monoclonal | Alpha tubulin | Sigma T9026 |

*Human rhinovirus (RV) 16 propagation and titration*

Human rhinovirus (HRV16; ATCC VR-283™, Teddington, UK) was amplified using H1 HeLa cells as previously described[84,85]. Infectivity of stocks and release of infective virions in cell culture supernatants was determined using a HeLa titration assay and 50% tissue culture infective dose assay ($TCID_{50}$/ml). Ultraviolet-irradiated

798 virus controls (UV-RV16) were prepared by exposure of virus stocks to UV light at
799 1200 mJ/cm$^2$ on ice for 50 min.

800 *Rhinovirus infection of differentiated human nasal and bronchial epithelial cells*

801 Fully differentiated nasal and bronchial epithelial cells (28 or 21 days after ALI) were
802 apically infected with human rhinovirus 16 (RV16) at a multiplicity of Infection (MOI)
803 of 1 for 6h, washed apically 3x using HBSS and incubated for additional 18h at the
804 air-liquid interface (24h in total). Cells were washed 3x with HBSS and lysed using
805 TriZol (Invitrogen) for RNA and protein extraction.

806 *SARS-CoV-2 infection*

807 BCi-NS1.1 cells (42-63 days after ALI) were apically infected for 1 hour with 100,000
808 pfu SARS-CoV-2 strain BetaCoV/Australia/VIC01/2020 (obtained from Public Health
809 England (PHE) UK and propagated in Vero E6 cells for no more than 2 passages
810 before use) and washed 2X using HBSS. Cells were harvested into QIAzol (Qiagen)
811 at 1h following HBSS wash and at 72h post infection for RNA extraction.

812 *Interferon-treatment*

813 PBECs monolayer cultures were stimulated with 100IU or 1000IU/ml with
814 Recombinant Human IFN-α A (α 2a) Protein (R&D, cat # 11100-1), IFN-β(a gift from
815 Synairgen Research Ltd), IFN-γ (Peprotech, cat# 300-02) or IFN-λ (National institute
816 of Biological standards and Control (NIBSC), cat# 10-176)
817 at a confluency of 70% and differentiated PBEC cultures at ALI were stimulated
818 basolateral with 1000IU/ml IFN-β. After 24h RNA was isolated using Monarch Total
819 RNA miniprep Kit (NEB) and reverse transcribed to cDNA using a Precision Reverse
820 Transcription kit (PrimerDesign, Southampton, UK) according to the manufacturer's
821 instructions.

822 *Exogenous protein expression cloning*

Full length sequence-verified human ACE2 cDNA expression clone with C-terminal GFP tag was purchased from Origene (rg208442). Nucleotides encoding amino acids 1-356 were removed from the vector by BamHI (NEB) digestion and cut ends dephosphorylated using rSAP (NEB). Cut vector was isolated using 0.5% agarose gel electrophoresis and band extraction using QIAquick Gel Extraction Kit (Qiagen). BamHI digestion also removes the ribosome binding site and Kozak consensus sequence. The ribosome binding site, Kozak consensus sequence, and nucleotides encoding 10 novel amino acids were inserted into the vector using a phosphorylated Ultramer Duplex with BamHI overhangs from IDT of the form:

Top 5'-3'

GATCCGAGGAGATCTGCCGCCGCGATCGCCATGAGGGAAGCAGGCTGGGACA

AAGGAGGGAG

Bot 5'-

3'GATCCTCCCTCCTTTGTCCCAGCCTGCTTCCCTCATGGCGATCGCGGCGGCA

GATCTCCTCG

Insert and digested vector were ligated using T4 DNA ligase (NEB), transformed into competent *E. coli* (Agilent) and multiple colonies picked from ampicillin LB agar plates for sequence verification using T7 promoter forward sequencing primer. Sequencing was performed by Source Biosciences. Clones with insert in correct orientation were selected for amplification in *E. coli* in liquid LB media and purification using plasmid midi kit (Qiagen).

*Transfection*

H441 and Vero E6 cells were transfected with short-ACE2-turboGFP, long-ACE2-turboGFP or control GFP constructs using Lipofectamine 3000 (Thermo Fischer Scientific) following manufacturer's protocol. HEK293 cells were transfected with

PEI, and RPE1 cells were nucleofected using a Lonza nucleofector 4D, using programme EA104. All cells were imaged live after 24 hours.

*Cell Free Expression*

The sequences for the ectodomains (10-740) of WT and short ACE2 isoforms were cloned into pTXTL_p70a vectors containing C terminal His6 and GFP11 sequences. To generate the linear DNA used for cell free expression, the plasmids were amplified with primers generating an extra 500 bp of flaking regions (5'-3' AAAGGGAATAAGGGCGACACG and 5'-3' TTGAGAAAGCGCCACGCTTC). The cell free protein expression was performed in a 24 µL scale using the MyTXTL Linear DNA Expression kit (Arbor Bioscience) with 5 nM DNA for 4 hours at 29oC. Following expression, a sample was taken for gel analysis, the remaining cell lysate was spun down, and the pellet was resuspended in reducing loading dye. The total and pellet samples were run on a reducing 4-20% SDS-PAGE gel (BioRad) and stained with InstantBlue Coomassie Protein Stain (Abcam).

*Structural analysis*

PyMOL Molecular Graphics System, Version 2.0 (Schrödinger, LLC) or UCSF Chimera[86] was used to model the novel protein isoform based on full-length isoform 6M17. Molecular dynamic simulation was performed by preparing an initial 3D model of the novel protein isoform based on PDB entry 6M17 using the homology modelling function of YASARA[87]. Molecular dynamics simulations in explicit solvent were performed using YASARA with GPU acceleration[88] on an Intel i9-9940X CPU (using 28 Threads) and GeForce RTX 2080 Ti. The molecular trajectory was sampled for 320 ns under NPT conditions at 310 K in 0.1% NaCl solution at pH 7.4 using periodic boundary conditions. Pymol, Chimera[86] and VMD[89] were used for molecular display and animation.

*Statistics*

Statistical analyses were performed in GraphPad Prism v7.02 (GraphPad Software Inc., San Diego, CA, USA) unless otherwise indicated. For each experiment, sample size reflects the number of independent biological replicates and is provided in the figure legend. Normality of data was assessed using Shapiro test to inform whether to apply parametric or non-parametric statistical tests. Statistical analyses of single comparisons of two groups utilized Student's t-test or Wilcoxon Signed Rank test for parametric and non-parametric data respectively. Correlations between groups of data were analysed using Pearson or Spearman rank correlation coefficients for parametric and non-parametric data respectively. Results were considered significant if $P<0.05$, where * $P<0.05$, ** $P<0.01$, *** $P<0.001$, ****$P<0.0001$.

## Contributions

VM and JSL conceptualised and supervised the initial study to interrogate RNAseq data for *ACE2* isoforms. GW identified the novel transcript, from which point CB, CLJ, DED, JSL, GW and VM conceptualised and supervised the remainder of the study. CLJ, CB, GW, DED and VM designed the experiments and analysed data. CLJ and CB performed most of the experimental work with additional contributions from GW, CMS, LN, JB, JL, FC, JC, JT, DJ, CMcC, RAR, LSND, PJS, SR and ML. GW, CB, JSL, VM, DB, DED, RD J-MP-C, AA, PS and KT provided samples and/or resources. MF and MC developed predictive structural models. GW, VM and DED wrote the manuscript text with CB, CLJ and JSL. GW and VM prepared figures with CB, CLJ, DED and DAJ. All authors approved the final submission.

## Acknowledgments

Research was supported by NIHR Southampton Biomedical Research Centre (BRC), NIHR Wellcome Trust Clinical Research Facility and AAIR Charity. GW is supported by a Wellcome Trust Seed Award in Science (204378/Z/16/Z). DB is supported by NIHR Research Professorship RP-2016-07-011. JSL, CLJ, VM, JT and JC are supported by the NHS England PCD National Service.

CB is a University of Southampton Career Track Fellow and FC is a Medical Research Foundation Fellow. ML is a BBSRC Future Leader Fellow and an NIHR Southampton BRC Senior Research Fellow. FC is supported by Medical Research Foundation grant MRF-091-0003-RG-CONFO.

We are grateful to healthy volunteers and respiratory patients who donated airway cells and to Synairgen Research Ltd who provided cells and reagents to support these studies.

## Data and materials availability

The RNA sequencing datasets analysed during the current study in Figure 1, Figure 3 and Supplementary Figure 1) are available in the Sequence Read Archive repository, Accession: PRJNA650028 ID: 650028 https://www.ncbi.nlm.nih.gov/bioproject/650028. The RNA sequencing dataset in Supplementary Figure 6 is available in the Sequence Read Archive repository SubmissionID: SUB8455806

BioProject ID: PRJNA674784

## Figure legends

**Fig 1. A novel short transcript of *ACE2* is expressed in airway epithelia**

1a. IGV plot showing RNA sequencing reads from one nasal brushing sample mapped to a 72kb region on chromosome Xp22 (chr X:15,543,782-15,617,034). Sequencing reads can be seen to be mapped to exons 1 - 19 of *ACE2*, and also to a region between exons 9 and 10, which we call exon 9a (red arrows). The red horizontal line across the mapped reads serves to illustrate that approximately twice the number of reads map to exons 9a – 19 compared to the number of reads mapped to exons 1 – 9, suggesting that two transcripts are expressed; one encompassing exons 1-19, and a second encompassing exons 9a-19.

1b. GENCODE v33 gene build exons and novel SCALLOP transcriptome build exons, showing novel exon 9a in a novel transcript encompassing exons 9a-19, assembled by SCALLOP tool

1c. Sashimi plot showing splice junction between exons 9 and 10, and between exons 9a and 10 counted from RNA sequencing reads from one nasal brushing sample. This shows that 29 reads map to the junction between exon 9 and exon 10, and that there is upstream splicing of exon 9 to exon 8, whereas there are 76 reads mapping to the junction between exon 9a and exon 10, with no further splicing upstream of exon 9a, suggesting that exon 9a is the first exon in this transcript.

1d. Nucleotide sequence of novel exon 9a, plus 5' UTR, start codon and splice junction

**Figure 2. RT-PCR and Sanger sequencing confirm short *ACE2* transcript expression in nasal and bronchial epithelial cells**

2a. Graphic of *ACE2* transcripts and RT-PCR primer locations

945 2b. Agarose gel electrophoresis image of long-range transcript-specific PCR
946 products amplifying full short *ACE2* transcript and exons 9-19 of long *ACE2*
947 transcript from nasal epithelial brushings and BCi-NS1.1 cells
948 2c. Sanger sequencing electropherogram traces showing sequence at exon/exon
949 boundaries of long *ACE2* transcript exon 9-10 and short *ACE2* transcript exon 9a-10.
950 Amino acid translation is shown below
951
952 **Figure 3. Short *ACE2* is expressed in different cell types, and predominantly in**
953 **the upper airway epithelium**
954 3a. Agarose gel electrophoresis image of transcript-specific *ACE2* RT-PCRs from
955 different cell types studied. Size standard = NEB Low Molecular Weight ladder.
956 3b. RT-qPCR analysis of transcript-specific PCRs for long and short *ACE2*
957 expression in cell lines and airway cells. Analysis was done at least in duplicate and
958 from different passages or donors in all lines but RPE1 and HEK293 cells where
959 analysis was done in duplicate from one passage.
960 3c. Graphs showing relative expression of short *ACE2* transcript and long *ACE2*
961 transcript in nasal epithelial cells at different stages of differentiation at air-liquid
962 interface (Day 1, 4, 8, 14, week 4, week 9, week 12) (n=3 for each time point) and
963 primary nasal brushings (n=6). Scale = reads mapped to exon/exon boundary per
964 million mapped reads. Error bars = standard error of the mean.
965 3d. Graphs showing relative expression of short *ACE2* transcript and long *ACE2*
966 transcript in nasal (n=11) and bronchial (n=11) ALI cultures from healthy donors, as
967 determined using transcript specific qPCR. Data were analysed using Mann
968 Whitney U test.

3e. Graphs showing relative expression of short *ACE2*, long *ACE2* and total *ACE2* transcript in a Multiple Tissue cDNA panel 1 (636742, Takara), as determined using transcript-specific probe-based *ACE2* RT-qPCR. n.d.= not detected.

**Figure 4. Short ACE2 protein is expressed and is not enriched on motile cilia relative to long ACE2 protein**

4a. Schematic illustration of predicted long and short protein isoforms of ACE2 and position of antigen sequences used to generate antibodies used.

4b. Representative western blot (n=3) of lysates prepared from Vero E6, HEK293, Caco2, RPE1, H441, 16HBE and BCi-NS1.1 cell lines immunoblotted with ACE2 C-terminal domain antibody (anti-ACE2 CTD). Grey arrow points to what is presumed to be glycosylated long ACE2, black arrow points to what is presumed to be unglycosylated long ACE2, red arrow points to what is presumed to be short ACE2

4c. Representative western blot of lysates prepared from Vero E6, *in vitro* differentiated nasal and bronchial epithelial cells, immunoblotted with ACE2 C-terminal domain antibody (anti-ACE2 CTD) preadsorbed with the immunizing (blocking) peptide (right) (n=3) or control peptide with similar charge (left) (n=1). Grey arrow points to what is presumed to be glycosylated long ACE2, black arrow points to what is presumed to be unglycosylated long ACE2, red arrow points to what is presumed to be short ACE2

4d. Left panel: Representative western blot (n=3) of Vero cells and *in vitro* differentiated nasal and bronchial cells, lysed and incubated with or without PNGase F, blotted with ACE2 antibody raised to C-terminal domain (amino acids 788-805). Right panel: Representative western blot (n=3) of Vero E6 cells, *in vitro* differentiated nasal and bronchial cells western blotted with ACE2 antibody raised to N-terminal

994 domain (amino acids 200-300). Grey arrow points to glycosylated long ACE2, black
995 arrow points to unglycosylated long ACE2, red arrow points to short ACE2
996 4e. Representative IF confocal images (n=4) of *in vitro* differentiated primary
997 bronchial epithelial cells stained with anti-alpha tubulin (red), Alexafluor-555
998 phalloidin (blue), DAPI (grey) and ACE2 (green) detected with antibody detecting
999 epitopes across the protein (anti-ACE2-ECTO) (top panel) or C-terminal domain
1000 antibody (anti-ACE2 CTD) (bottom panel)
1001 4f. Schematic illustration of deciliation protocol using calcium shock (left) and
1002 western blot of whole and deciliated BCi-NS1.1 cells, deciliation wash, and cilia
1003 pellet (right). Grey arrow points to glycosylated long ACE2, red arrow points to short
1004 ACE2 which not enriched on cilia relative to long ACE2 enrichment.
1005 4g. Box and whisker plot of semi-quantitative analysis of the Western blots by
1006 densitometric analysis (n=4).
1007
1008
1009 **Figure 5. Short *ACE2* is upregulated in response to IFN and rhinovirus (RV16)**
1010 **infection but not SARS-CoV-2 infection**
1011 5a. Undifferentiated primary bronchial epithelial cell (PBEC) monolayer cultures were
1012 treated with IFN-α (n=3 donors), IFN-β (n=7 donors), IFN-γ (n=4 donors) or IFN-λ
1013 (n=4 donors) at the doses indicated for 24h and total *ACE2* transcripts (left panel)p,
1014 long *ACE2* transcripts (middle ) or short *ACE2* transcripts (right panel) were
1015 measured by RT-qPCR with transcript-specific primers. Data were analysed using
1016 Mann-Whitney test.
1017 5b. *In vitro* differentiated (ALI) nasal epithelia cells (NEC) (n=11 donors) were
1018 infected with rhinovirus (RV16) (MOI of 1) or mock-infected using a UV-irradiated

control (UV-RV16). Nasal cells were collected from 3 female, 8 male patients with a mean age of 45.31+/-3.23 (SEM). After 24h, induction of *ACE2* isoform expression was assessed by RT-qPCR with transcript-specific primers. Data were analysed using non-parametric Wilcoxon test.

5c. *In vitro* differentiated primary bronchial epithelia cells (PBEC) (n=13 donors) were infected with rhinovirus (RV16) (MOI of 1) or mock-infected using a UV-irradiated control (UV-RV16). Cells were collected from 6 female and 7 male patients with a mean age of 36.69+/-4.02 (mean+/-SEM). After 24h, induction of *ACE2* isoform expression was assessed by RT-qPCR with transcript-specific primers. Data were analysed using non-parametric Wilcoxon test.

5d. BCi-NS1.1 cells were grown at ALI and then infected for 1 hour with 100,000 pfu of SARS-CoV-2 strain nCoV/Victoria/1/2020. After 1h or 72h SARS-CoV-2 infection was confirmed by CoV-N1 RT-qPCR. Data were analysed using non-parametric Wilcoxon test (n=4).

5e. 1 hour and 72 hours after SARS-CoV-2 infection of BCi-NS1.1 cells at ALI, induction of *ACE2* transcript expression was assessed by RT-qPCR with transcript-specific primers. . Data were analysed using non-parametric Wilcoxon test (n=4).

5f. 1 hour and 72 hours after SARS-CoV-2 infection of BCi-NS1.1 cells at ALI, induction of *MX1* and *IP10* transcript expression was assessed by RT-qPCR. Data were analysed using non-parametric Wilcoxon test (n=4).

5g. *In vitro* differentiated primary bronchial epithelia cells (PBEC) from healthy (n=13) or severe asthmatic (n=12) donors were infected with rhinovirus (RV16) (MOI of 1) or mock-infected using a UV-irradiated control (UV-RV16). After 24h, induction of *ACE2* isoform expression was assessed by RT-qPCR with transcript-specific primers and

the ratio of short *ACE2* to long *ACE2* was calculated as delta -dCT values of short and long ACE2 expression. Data were analysed using Mann-Whitney test.

5h. Correlation between expression of short *ACE2* (left) or long *ACE2* (right) and *IFNB1* gene expression in response to RV infection. Upward pointing open triangles represent healthy controls, downward pointing grey-filled triangles represent severe asthmatic individuals.

**Supplementary Figure 1**

Boxplot and whisker showing median, quartiles and range of gene expression data for 6 selected cilia genes in primary nasal brushings (blue) and primary nasal epithelial cells cultured at ALI for 1, 4, 8, 14, 21, 28 and 63 days (orange), to demonstrate early activation of genes associated with ciliogenesis and cilium function from day 4.

**Supplementary Figure 2**

a) Plots of CT value against cDNA input into RT-qPCR reactions to determine dynamic range of RT-qPCRs for 3 SyBr green assays using primers targeting total *ACE2* (top left), long *ACE2* (top right) and short *ACE2* (bottom left) and one probe-based Taqman assay using primers targeting short *ACE2*. CT values below the indicated dotted line were regarded as positive results.

b) SyBr green qPCR results compared to probe-based Taqman assay results to show consistency of short ACE2 qPCR assays.

**Supplementary Figure 3**

a. Graphical representation of long and short ACE2 protein with position of relevant spike binding regions and protease cleavage domains. TM = transmembrane domain, CYT = cytoplasmic domain.

b. Long ACE2 homodimer (teal) in complex with SARS-Cov-2 spike protein (orange) (right) from cryo-EM resolved structure PDB 6M17; predicted short ACE2 homodimer (right) based on this structure. In both, residues essential for cleavage by ADAM17 are shown in yellow and residues essential for cleavage by TMPRSS11D and TMPRSS2 are shown in fuchsia. The residues present in long ACE2 but absent in short ACE2 are shown in royal blue on the long ACE2 structure.

c. High magnification of the putative interaction region of long ACE2 with SARS-CoV-2 spike protein (top) and the same region in short ACE2 (bottom).

d. Snapshot of MD simulation of short ACE2 obtained with YASARA (**Supplementary Video 1**)

e. DSSP analysis of short ACE2 chain A and B over the course of the 300 ns MD simulation.

f. Analysis of helical content (left) and secondary structure variation (right) of short ACE2 over the course of the 300 ns MD simulation.

**Supplementary Figure 4**

Representative IF confocal images of ALI-differentiated primary bronchial epithelial cells stained with 2 commercial anti ACE 2antibodies (green), anti alpha-tubulin (red), Actin-stain 555 phalloidin (blue) and DAPI (grey). Overlay and separate channels.

a. XY imaging plane and XZ and YZ orthogonal confocal slices along white lines using Abcam ab15348 anti-ACE2 C terminal domain (raised against amino acids 788-805). Scale bar = 20 µm.

b. XY imaging plane and XZ and YZ orthogonal confocal slices along white lines using R&D AF933 anti-ACE2 ectodomain (raised against amino acids 18-740). Scale bar = 20 μm.

c. Enlarged orthogonal slices using Abcam ab15348 anti-ACE2 C-terminal domain showing ACE2 staining extending beyond the ciliary axoneme. Overlay and separate channels. Scale bar = 10 μm.

d. Enlarged orthogonal slices using R&D AF933 anti-ACE2 ectodomain showing ACE2 staining extending beyond the ciliary axoneme. Overlay and separate channels. Scale bar = 10 μm.

**Supplementary Figure 5**

a) Coomassie stained polyacrylamide gel showing purification of short (left) and long (right) ACE2 ectodomains from an *E. coli* cell-free system. Tot = total lysate loading, P = purified protein. Black arrows point to purified proteins at expected sizes.

b) Wide field epifluorescence and phase contrast image overlay of H441 cells transfected with exogenous long ACE2-turboGFP (left) or exogenous short ACE2-turboGFP (right). Scale bar = 50 μM. White arrows point to membrane localization of long ACE2-turboGFP

**Supplementary Figure 6**

a and b) Box and whisker plots showing relative expression of *MX1* (a) and *IP10* (b) in undifferentiated primary bronchial epithelial cell (PBEC) monolayer cultures treated with Type I, II or III IFNs. N=3 for IFN-α, n=4 for γ and λ experiments, n=7 for IFN-β experiments. Data were analysed using Mann-Whitney test.

c and d) Box and whisker plots showing relative expression of long *ACE2* (c) and ratio of short ACE2:long ACE2 expression (d) in undifferentiated primary bronchial

epithelial cell (PBEC) monolayer cultures treated with Type I, II or III IFNs. N=3 for IFN-α, n=4 for γ and λ experiments, n=7 for IFN-β experiments. Data were analysed using Mann-Whitney test.

e) Histograms showing relative expression of total *ACE2*, long *ACE2* and short *ACE2* and f) *MX1* and *IP10* (right) in *in vitro* differentiated (ALI) PBEC cultures (N=3) treated with IFN-β (1000 IU/ml). Data were analysed using Students t-test.

g) Histograms showing relative expression of total *ACE2*, long *ACE2* and short *ACE2* and *MX1* and h) *IP10* in BCi-NS1.1 cells infected with RV16 (n=3).

**Supplementary Figure 7**

a and b) Box and whisker plots of total *ACE2*, long *ACE2* and short *ACE2* expression (measured by transcript-specific RT-qPCR using $2^{-ddCT}$ method) in *in vitro* differentiated primary bronchial epithelia cells (PBEC) from healthy (a) (n=13) or severe asthmatic (b) (n=11) donors 24 hours after infection with rhinovirus (RV16) (MOI of 1) or mock-infection using a UV-irradiated control (UV-RV16).

c) Box and whisker plots showing data from a) and b) as the ratio of short *ACE2* : long *ACE2* expression. After testing for normality, data were analysed using Mann-Whitney test.

d). Box and whisker plots showing level of basolateral secretion of IL29/IL28 (pg/ml) in response to RV16-infection or mock-infection of *in vitro* differentiated PBECs from healthy (n=14) or severe asthmatic (n=8) donors. IL29/IL28 was determined by ELISA. After testing for normality, data were analysed using Student's t-test.

e) Correlation analysis of short *ACE2* (left) and long *ACE2* (right) transcript expression (measured by transcript-specific RT-qPCR using $2^{-ddCT}$ method) with level

of secreted IL29/IL28 (pg/ml) (measured by ELISA) in *in vitro* differentiated PBECs from healthy (n=7) or severe asthmatic (n=5) donors (left) in response to RV16 infection. After testing for normality, data were analysed to calculate Spearman's rank correlation coefficient.

f) Expression of total *ACE2*, long *ACE2* and short *ACE2* (measured by transcript-specific RT-qPCR using -dCT method) in bronchial epithelial brushes from healthy controls (n=13) or severe asthmatic (n=12) donors. After testing for normality, data were analysed using non-parametric Mann-Whitney test.

g) Box and whisker plot showing ration of expression of total *ACE2*, long *ACE2* and short *ACE2* (measured by transcript-specific RT-qPCR using -dCT method) in bronchial epithelial brushes from healthy controls (n=13) or severe asthmatic (n=11) donors. Data were analysed using non-parametric Mann-Whitney test.

h). Expression levels of IFN-response genes *OAS1* and *ISG1* (measured in transcripts per million (TPM) in RNAseq data) in bronchial brushings from healthy volunteers (n=13) compared to severe asthmatic subjects (n=10). After testing for normality, data were analysed using unpaired Student's t-test.

i) Correlation analysis of short *ACE2* transcript expression (measured in transcripts per million (TPM) in RNAseq data) in bronchial brushings with IFN-γ levels (pg/ml) in bronchoalveolar lavage (BAL) fluid harvested from the same donor at the time of bronchial brushing. After testing for normality, data were analysed to calculate Spearman's rank correlation coefficient.

## References

1. Tipnis, S.R. *et al.* A human homolog of angiotensin-converting enzyme. Cloning and functional expression as a captopril-insensitive carboxypeptidase. *J Biol Chem* **275**, 33238-43 (2000).

2. Donoghue, M. *et al.* A novel angiotensin-converting enzyme-related carboxypeptidase (ACE2) converts angiotensin I to angiotensin 1-9. *Circ Res* **87**, E1-9 (2000).
3. Zhang, H. *et al.* Collectrin, a collecting duct-specific transmembrane glycoprotein, is a novel homolog of ACE2 and is developmentally regulated in embryonic kidneys. *J Biol Chem* **276**, 17132-9 (2001).
4. Shulla, A. *et al.* A transmembrane serine protease is linked to the severe acute respiratory syndrome coronavirus receptor and activates virus entry. *J Virol* **85**, 873-82 (2011).
5. Heurich, A. *et al.* TMPRSS2 and ADAM17 cleave ACE2 differentially and only proteolysis by TMPRSS2 augments entry driven by the severe acute respiratory syndrome coronavirus spike protein. *J Virol* **88**, 1293-307 (2014).
6. Li, W. *et al.* Angiotensin-converting enzyme 2 is a functional receptor for the SARS coronavirus. *Nature* **426**, 450-454 (2003).
7. Matsuyama, S. *et al.* Efficient Activation of the Severe Acute Respiratory Syndrome Coronavirus Spike Protein by the Transmembrane Protease TMPRSS2. *Journal of Virology* **84**, 12658 (2010).
8. Kuba, K. *et al.* A crucial role of angiotensin converting enzyme 2 (ACE2) in SARS coronavirus–induced lung injury. *Nature Medicine* **11**, 875-879 (2005).
9. Yan, R. *et al.* Structural basis for the recognition of SARS-CoV-2 by full-length human ACE2. *Science (New York, N.Y.)* **367**, 1444-1448 (2020).
10. Wrapp, D. *et al.* Cryo-EM structure of the 2019-nCoV spike in the prefusion conformation. *Science* **367**, 1260-1263 (2020).
11. 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.e8 (2020).
12. Blau, D.M. & Holmes, K.V. Human coronavirus HCoV-229E enters susceptible cells via the endocytic pathway. *Adv Exp Med Biol* **494**, 193-8 (2001).
13. Inoue, Y. *et al.* Clathrin-dependent entry of severe acute respiratory syndrome coronavirus into target cells expressing ACE2 with the cytoplasmic tail deleted. *Journal of virology* **81**, 8722-8729 (2007).
14. Wang, H. *et al.* SARS coronavirus entry into host cells through a novel clathrin- and caveolae-independent endocytic pathway. *Cell Research* **18**, 290-301 (2008).
15. Camargo, S.M. *et al.* Tissue-specific amino acid transporter partners ACE2 and collectrin differentially interact with hartnup mutations. *Gastroenterology* **136**, 872-82 (2009).
16. Kowalczuk, S. *et al.* A protein complex in the brush-border membrane explains a Hartnup disorder allele. *Faseb j* **22**, 2880-7 (2008).
17. Niu, M.-J., Yang, J.-K., Lin, S.-S., Ji, X.-J. & Guo, L.-M. Loss of angiotensin-converting enzyme 2 leads to impaired glucose homeostasis in mice. *Endocrine* **34**, 56-61 (2008).
18. Bindom, S.M., Hans, C.P., Xia, H., Boulares, A.H. & Lazartigues, E. Angiotensin I-converting enzyme type 2 (ACE2) gene therapy improves glycemic control in diabetic mice. *Diabetes* **59**, 2540-8 (2010).
19. Imai, Y. *et al.* Angiotensin-converting enzyme 2 protects from severe acute lung failure. *Nature* **436**, 112-116 (2005).

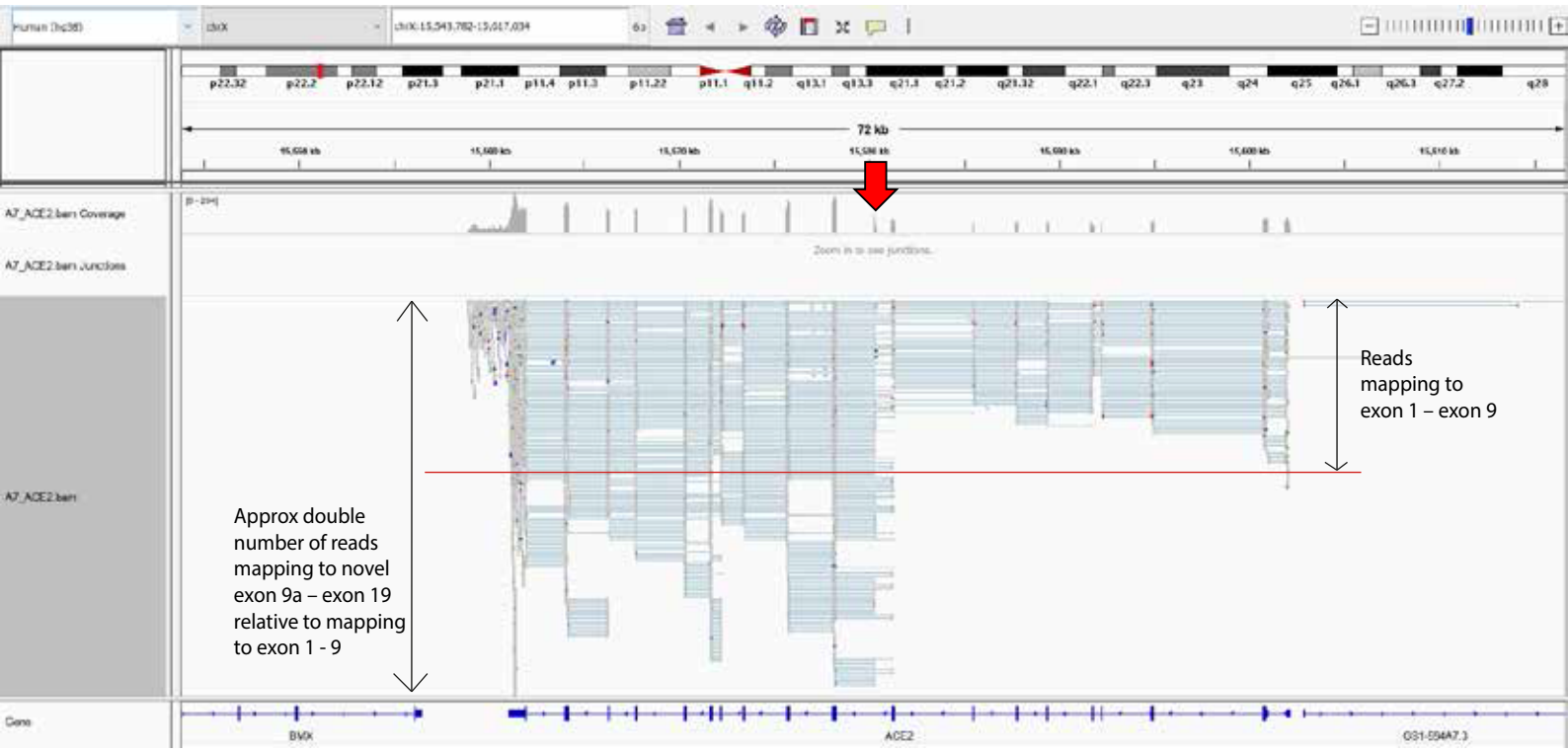
20. Treml, B. *et al.* Recombinant angiotensin-converting enzyme 2 improves pulmonary blood flow and oxygenation in lipopolysaccharide-induced lung injury in piglets. *Crit Care Med* **38**, 596-601 (2010).
21. Ferreira, A.J. *et al.* Evidence for angiotensin-converting enzyme 2 as a therapeutic target for the prevention of pulmonary hypertension. *American journal of respiratory and critical care medicine* **179**, 1048-1054 (2009).
22. Yamazato, Y. *et al.* Prevention of pulmonary hypertension by Angiotensin-converting enzyme 2 gene transfer. *Hypertension* **54**, 365-71 (2009).
23. Pedersen, K.B., Chhabra, K.H., Nguyen, V.K., Xia, H. & Lazartigues, E. The transcription factor HNF1α induces expression of angiotensin-converting enzyme 2 (ACE2) in pancreatic islets from evolutionarily conserved promoter motifs. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms* **1829**, 1225-1235 (2013).
24. Kuan, T.C. *et al.* Identifying the regulatory element for human angiotensin-converting enzyme 2 (ACE2) expression in human cardiofibroblasts. *Peptides* **32**, 1832-9 (2011).
25. Wang, Y. *et al.* Administration of 17β-estradiol to ovariectomized obese female mice reverses obesity-hypertension through an ACE2-dependent mechanism. *Am J Physiol Endocrinol Metab* **308**, E1066-75 (2015).
26. 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.e19 (2020).
27. The Genotype-Tissue Expression (GTEx) project. *Nat Genet* **45**, 580-5 (2013).
28. Sungnak, W. *et al.* SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat Med* **26**, 681-687 (2020).
29. Hou, Y.J. *et al.* SARS-CoV-2 Reverse Genetics Reveals a Variable Infection Gradient in the Respiratory Tract. *Cell* (2020).
30. Zou, L. *et al.* SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. in *N Engl J Med*, Vol. 382 1177-1179 (2020).
31. Sims, A.C. *et al.* Severe acute respiratory syndrome coronavirus infection of human ciliated airway epithelia: role of ciliated cells in viral spread in the conducting airways of the lungs. *Journal of virology* **79**, 15511-15524 (2005).
32. Schaefer, I.M. *et al.* In situ detection of SARS-CoV-2 in lungs and airways of patients with COVID-19. *Mod Pathol*, 1-11 (2020).
33. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15-21 (2013).
34. Robinson, J.T. *et al.* Integrative genomics viewer. *Nature Biotechnology* **29**, 24 (2011).
35. Cummings, B.B. *et al.* Improving genetic diagnosis in Mendelian disease with transcriptome sequencing. *Sci Transl Med* **9**(2017).
36. Fu, X.Y., Kessler, D.S., Veals, S.A., Levy, D.E. & Darnell, J.E., Jr. ISGF3, the transcriptional activator induced by interferon alpha, consists of multiple interacting polypeptide chains. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 8555-8559 (1990).
37. Isern, E. *et al.* The activator protein 1 binding motifs within the human cytomegalovirus major immediate-early enhancer are functionally redundant and act in a cooperative manner with the NF-{kappa}B sites during acute infection. *J Virol* **85**, 1732-46 (2011).

38. Wan, F. & Lenardo, M.J. Specification of DNA binding activity of NF-kappaB proteins. *Cold Spring Harbor perspectives in biology* **1**, a000067-a000067 (2009).
39. Walters, M.S. *et al.* Generation of a human airway epithelium derived basal cell line with multipotent differentiation capacity. *Respiratory research* **14**, 135-135 (2013).
40. Jia, H.P. *et al.* ACE2 receptor expression and severe acute respiratory syndrome coronavirus infection depend on differentiation of human airway epithelia. *Journal of virology* **79**, 14614-14621 (2005).
41. Wen, B., Wang, X. & Zhang, B. PepQuery enables fast, accurate, and convenient proteomic validation of novel genomic alterations. *Genome Res* **29**, 485-493 (2019).
42. Drosten, C. *et al.* Identification of a novel coronavirus in patients with severe acute respiratory syndrome. *N Engl J Med* **348**, 1967-76 (2003).
43. Matsuyama, S. *et al.* Enhanced isolation of SARS-CoV-2 by TMPRSS2-expressing cells. *Proceedings of the National Academy of Sciences* **117**, 7001-7003 (2020).
44. Lee, I.T. *et al.* Robust ACE2 protein expression localizes to the motile cilia of the respiratory tract epithelia and is not increased by ACE inhibitors or angiotensin receptor blockers. *medRxiv* (2020).
45. Kindler, E., Thiel, V. & Weber, F. Interaction of SARS and MERS Coronaviruses with the Antiviral Interferon Response. *Adv Virus Res* **96**, 219-243 (2016).
46. Yuen, C.K. *et al.* SARS-CoV-2 nsp13, nsp14, nsp15 and orf6 function as potent interferon antagonists. *Emerg Microbes Infect* **9**, 1418-1428 (2020).
47. Busnadiego, I. *et al.* Antiviral Activity of Type I, II, and III Interferons Counterbalances ACE2 Inducibility and Restricts SARS-CoV-2. *mBio* **11**, e01928-20 (2020).
48. Acharya, D., Liu, G. & Gack, M.U. Dysregulation of type I interferon responses in COVID-19. *Nature Reviews Immunology* **20**, 397-398 (2020).
49. Garcia-Pachon, E. *et al.* Asthma prevalence in patients with SARS-CoV-2 infection detected by RT-PCR not requiring hospitalization. *Respir Med* **171**, 106084 (2020).
50. Grandbastien, M. *et al.* SARS-CoV-2 pneumonia in hospitalized asthmatic patients did not induce severe exacerbation. *J Allergy Clin Immunol Pract* (2020).
51. Chhiba, K.D. *et al.* Prevalence and characterization of asthma in hospitalized and non-hospitalized patients with COVID-19. *J Allergy Clin Immunol* (2020).
52. Contoli, M. *et al.* Role of deficient type III interferon-lambda production in asthma exacerbations. *Nat Med* **12**, 1023-6 (2006).
53. Wark, P.A. *et al.* Asthmatic bronchial epithelial cells have a deficient innate immune response to infection with rhinovirus. *J Exp Med* **201**, 937-47 (2005).
54. Camiolo, M.J., Gauthier, M., Kaminski, N., Ray, A. & Wenzel, S.E. Expression of SARS-CoV-2 Receptor ACE2 and Coincident Host Response Signature Varies by Asthma Inflammatory Phenotype. *J Allergy Clin Immunol* (2020).
55. Ng, K.W. *et al.* Tissue-specific and interferon-inducible expression of nonfunctional ACE2 through endogenous retroelement co-option. *Nature Genetics* (2020).
56. Onabajo, O.O. *et al.* Interferons and viruses induce a novel truncated ACE2 isoform and not the full-length SARS-CoV-2 receptor. *Nature Genetics* (2020).

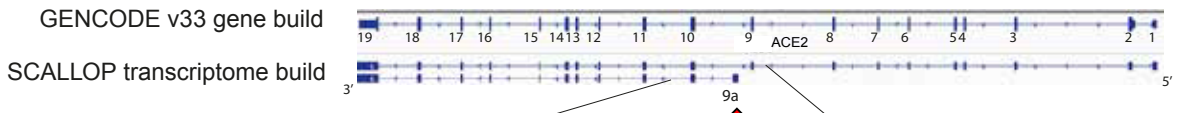
57. Towler, P. *et al.* ACE2 X-ray structures reveal a large hinge-bending motion important for inhibitor binding and catalysis. *J Biol Chem* **279**, 17996-8007 (2004).
58. Clarke, N.E., Belyaev, N.D., Lambert, D.W. & Turner, A.J. Epigenetic regulation of angiotensin-converting enzyme 2 (ACE2) by SIRT1 under conditions of cell energy stress. *Clin Sci (Lond)* **126**, 507-16 (2014).
59. Monk, P.D. *et al.* Safety and efficacy of inhaled nebulised interferon beta-1a (SNG001) for treatment of SARS-CoV-2: a randomised, double-blind, placebo-controlled phase 2 trial. **Submitted**(2020).
60. Takeyama, K., Fahy, J.V. & Nadel, J.A. Relationship of epidermal growth factor receptors to goblet cell production in human bronchi. *Am J Respir Crit Care Med* **163**, 511-6 (2001).
61. Kimura, H. *et al.* Type 2 inflammation modulates ACE2 and TMPRSS2 in airway epithelial cells. *J Allergy Clin Immunol* **146**, 80-88.e8 (2020).
62. Flammer, J.R. *et al.* The type I interferon signaling pathway is a target for glucocorticoid inhibition. *Mol Cell Biol* **30**, 4564-74 (2010).
63. Howard, T.E., Shai, S.Y., Langford, K.G., Martin, B.M. & Bernstein, K.E. Transcription of testicular angiotensin-converting enzyme (ACE) is initiated within the 12th intron of the somatic ACE gene. *Mol Cell Biol* **10**, 4294-302 (1990).
64. Ojaghi, M., Kastelic, J. & Thundathil, J. Testis-specific isoform of angiotensin-converting enzyme (tACE) is involved in the regulation of bovine sperm capacitation. *Mol Reprod Dev* **84**, 376-388 (2017).
65. Abrie, J.A. *et al.* Investigation into the Mechanism of Homo- and Heterodimerization of Angiotensin-Converting Enzyme. *Mol Pharmacol* **93**, 344-354 (2018).
66. Ho, M. Perspectives on the development of neutralizing antibodies against SARS-CoV-2. *Antib Ther* **3**, 109-114 (2020).
67. Inal, J.M. Decoy ACE2-expressing extracellular vesicles that competitively bind SARS-CoV-2 as a possible COVID-19 therapy. *Clin Sci (Lond)* **134**, 1301-1304 (2020).
68. Saheb Sharif-Askari, N. *et al.* Airways Expression of SARS-CoV-2 Receptor, ACE2, and TMPRSS2 Is Lower in Children Than Adults and Increases with Smoking and COPD. *Mol Ther Methods Clin Dev* **18**, 1-6 (2020).
69. Pinto, B.G.G. *et al.* ACE2 Expression is Increased in the Lungs of Patients with Comorbidities Associated with Severe COVID-19. *J Infect Dis* (2020).
70. Emilsson, V. *et al.* ACE2 levels are altered in comorbidities linked to severe outcome in COVID-19. *medRxiv* (2020).
71. Zhou, L. *et al.* ACE2 and TMPRSS2 are expressed on the human ocular surface, suggesting susceptibility to SARS-CoV-2 infection. *Ocul Surf* **18**, 537-44 (2020).
72. Collin, J. *et al.* Co-expression of SARS-CoV-2 entry genes in the superficial adult human conjunctival, limbal and corneal epithelium suggests an additional route of entry via the ocular surface. *Ocul Surf* (2020).
73. Lange, C. *et al.* Expression of the COVID-19 receptor ACE2 in the human conjunctiva. *J Med Virol* (2020).
74. Ma, D. *et al.* Expression of SARS-CoV-2 receptor ACE2 and TMPRSS2 in human primary conjunctival and pterygium cell lines and in mouse cornea. *Eye (Lond)*, 1-8 (2020).

75. Zhang, D. *et al.* Incomplete annotation has a disproportionate impact on our understanding of Mendelian and complex neurogenetic disorders. *Science Advances* **6**, eaay8299 (2020).
76. Hirst, R.A. *et al.* Culture of primary ciliary dyskinesia epithelial cells at air-liquid interface can alter ciliary phenotype but remains a robust and informative diagnostic aid. *PLoS One* **9**, e89675 (2014).
77. Xiao, C. *et al.* Defective epithelial barrier function in asthma. *J Allergy Clin Immunol* **128**, 549-56.e1-12 (2011).
78. Kuek, L.E. *et al.* Identification of an Immortalized Human Airway Epithelial Cell Line with Dyskinetic Cilia. *Am J Respir Cell Mol Biol* **59**, 375-382 (2018).
79. Wang, L., Wang, S. & Li, W. RSeQC: quality control of RNA-seq experiments. *Bioinformatics* **28**, 2184-5 (2012).
80. Shao, M. & Kingsford, C. Accurate assembly of transcripts through phase-preserving graph decomposition. *Nature Biotechnology* **35**, 1167-1169 (2017).
81. Patro, R., Duggal, G., Love, M.I., Irizarry, R.A. & Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat Methods* **14**, 417-419 (2017).
82. Ostrowski, L.E. *et al.* A Proteomic Analysis of Human Cilia. *Molecular & Cellular Proteomics* **1**, 451 (2002).
83. Hastie, A.T. *et al.* Isolation of cilia from porcine tracheal epithelium and extraction of dynein arms. *Cell Motil Cytoskeleton* **6**, 25-34 (1986).
84. Calvén, J. *et al.* Viral stimuli trigger exaggerated thymic stromal lymphopoietin expression by chronic obstructive pulmonary disease epithelium: role of endosomal TLR3 and cytosolic RIG-I-like helicases. *J Innate Immun* **4**, 86-99 (2012).
85. Zhao, W. *et al.* Peroxisome proliferator-activated receptor gamma negatively regulates IFN-beta production in Toll-like receptor (TLR) 3- and TLR4-stimulated macrophages by preventing interferon regulatory factor 3 binding to the IFN-beta promoter. *J Biol Chem* **286**, 5519-28 (2011).
86. Pettersen, E.F. *et al.* UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* **25**, 1605-12 (2004).
87. Krieger, E. *et al.* Improving physical realism, stereochemistry, and side-chain accuracy in homology modeling: Four approaches that performed well in CASP8. *Proteins* **77 Suppl 9**, 114-22 (2009).
88. Krieger, E. & Vriend, G. New ways to boost molecular dynamics simulations. *J Comput Chem* **36**, 996-1007 (2015).
89. Humphrey, W., Dalke, A. & Schulten, K. VMD: visual molecular dynamics. *J Mol Graph* **14**, 33-8, 27-8 (1996).

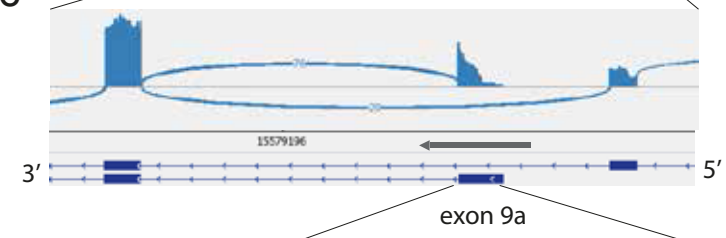
a


72 kb
A7_ACE2.bam Coverage
A7_ACE2.bam Junctions
A7_ACE2.bam
Gene
BVDX
ACE2
Reads mapping to exon 1 – exon 9
Approx double number of reads mapping to novel exon 9a – exon 19 relative to mapping to exon 1 - 9

b

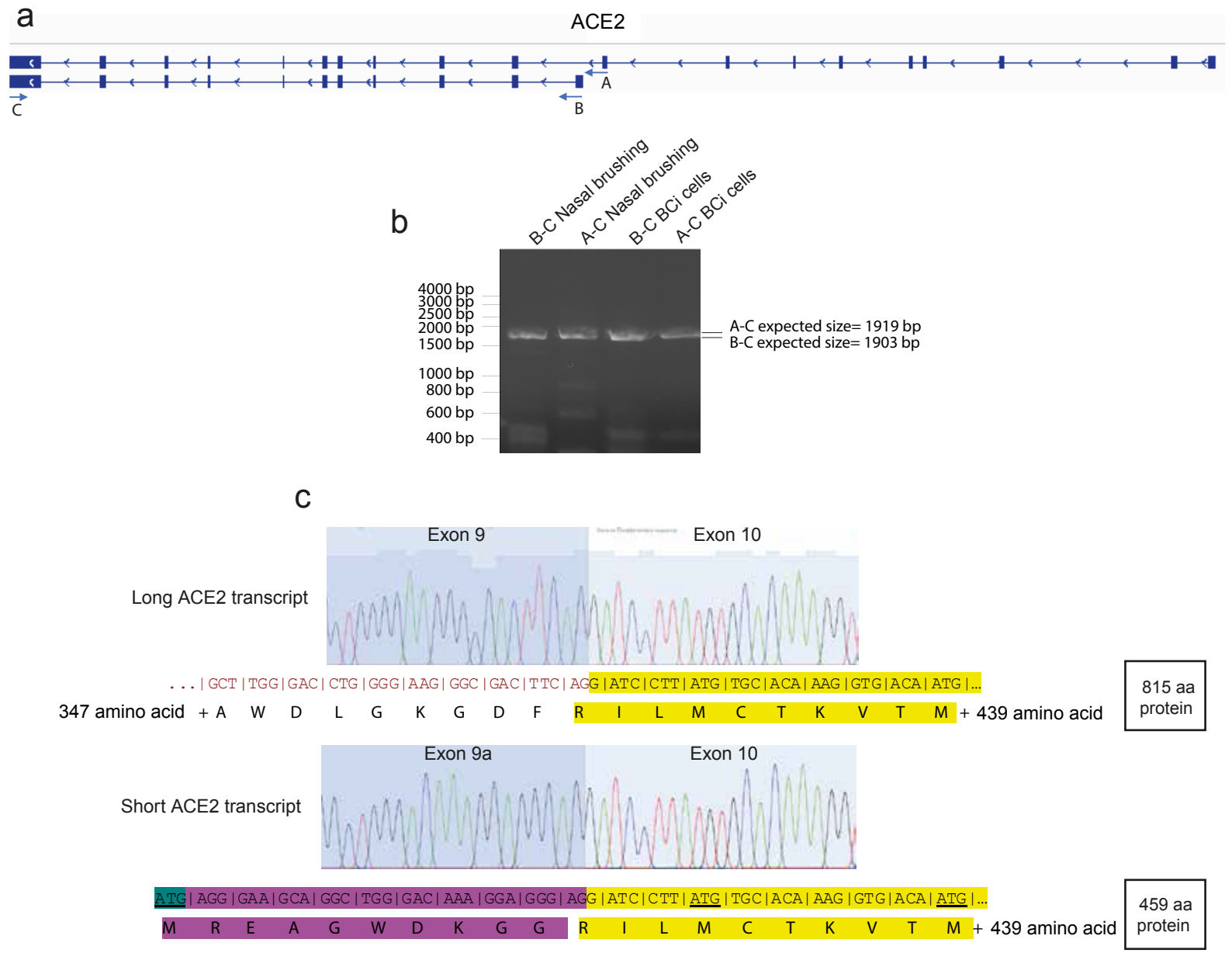

GENCODE v33 gene build
SCALLOP transcriptome build
ACE2
9a

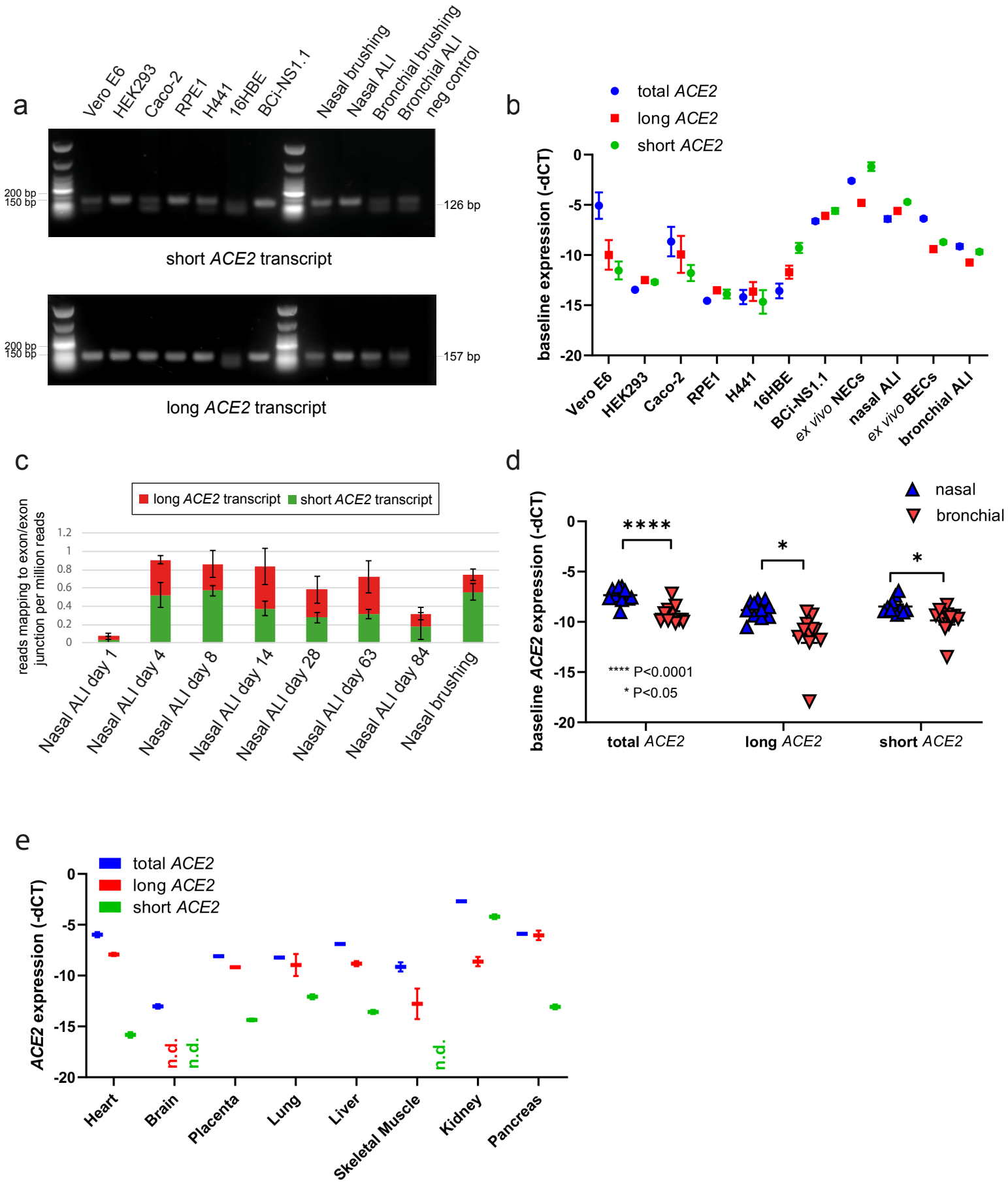
c

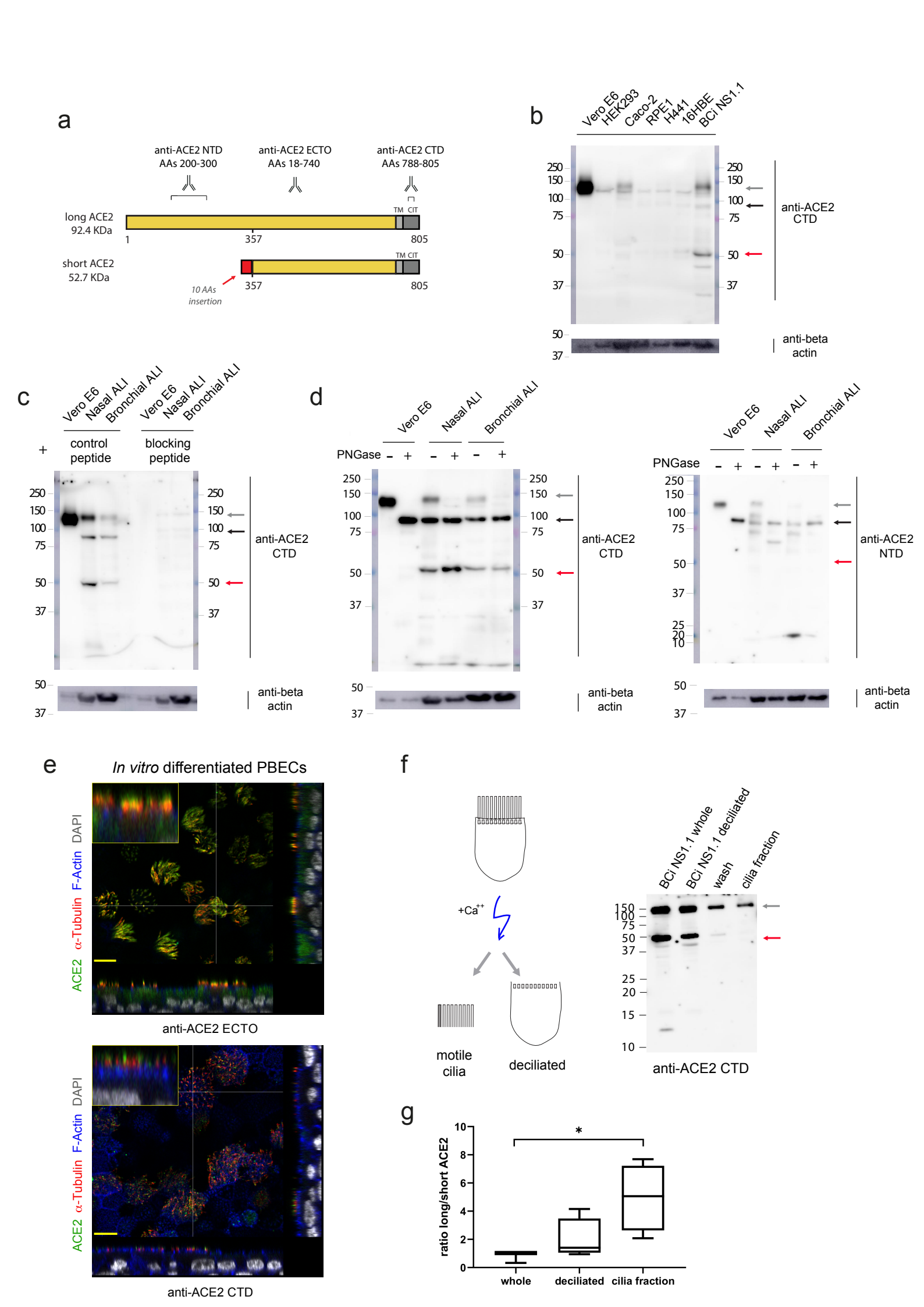

15579196
3′
5′
exon 9a

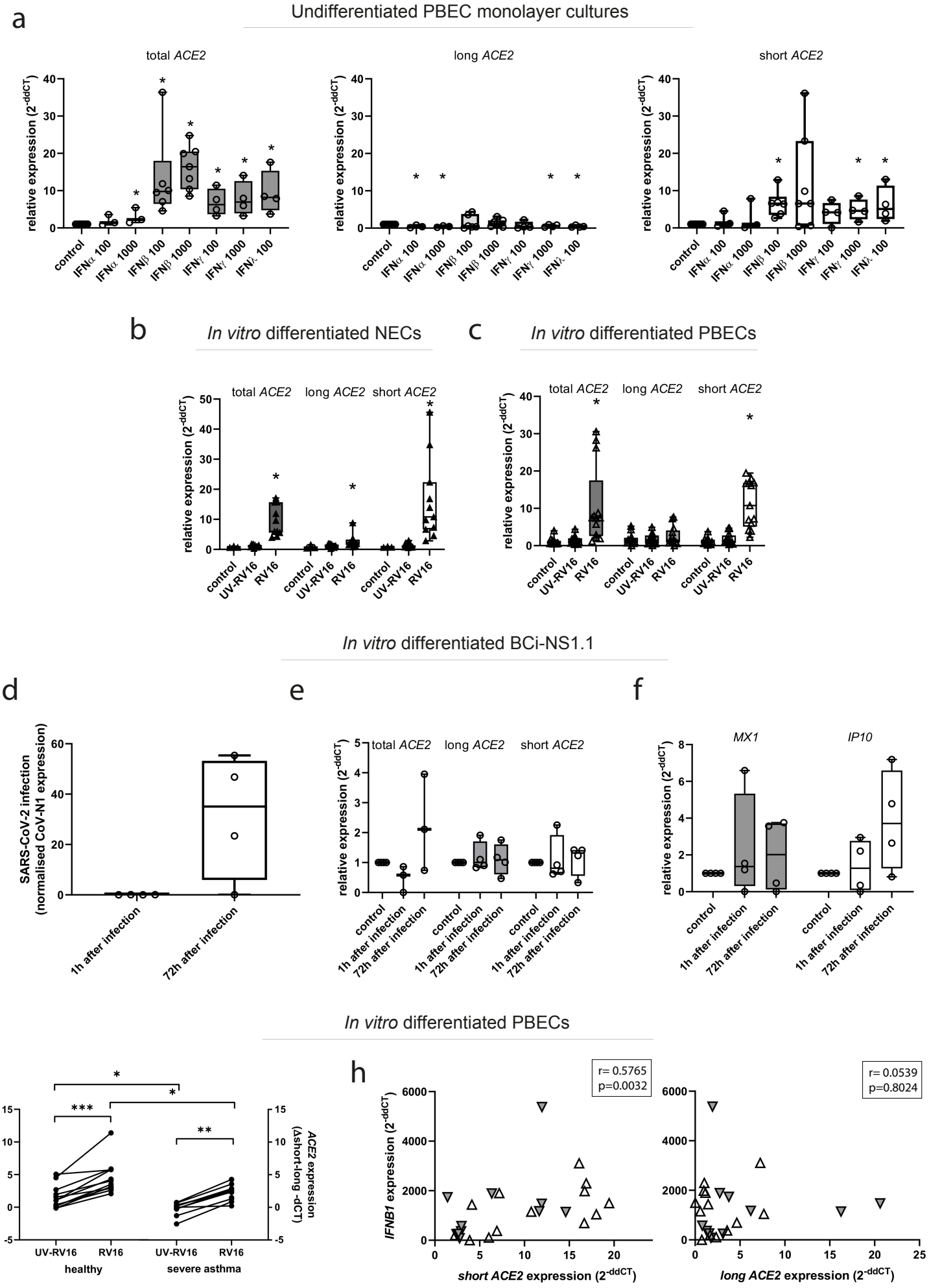
d

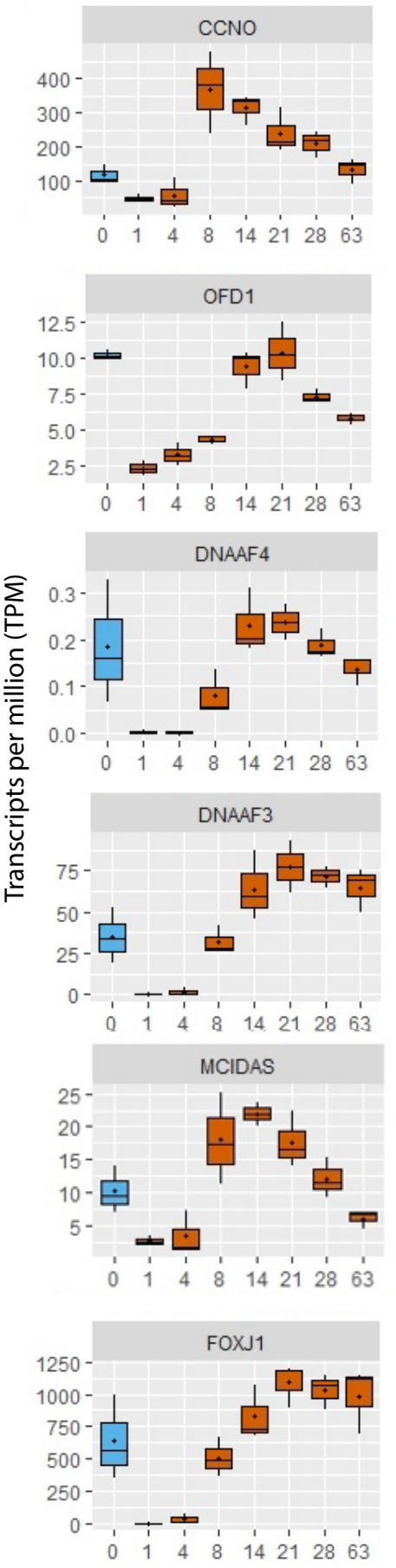

5′
TATAAAGTAGAATAGTAATTCCCAGGTTGCAGGCTTGTGAGAGCCTTAGGTTGGATTCCCTAGCTTGAAAAGGAGATCGTTTTACAAGTGCTTCATTGAGGAGAGCTCTGAGGCAGAGGGGAATGAGGGAAGCAGGCTGGGACAAAGGAGGGAGGTAAGTAAGAA
3′
TATA box
ChrX:15580402 Transcription start site
Start codon
AG|GT
Consensus 5′ splice site (splice donor site)


a
ACE2
C
B
A
b
B-C Nasal brushing
A-C Nasal brushing
B-C BCi cells
A-C BCi cells
4000 bp
3000 bp
2500 bp
2000 bp
1500 bp
1000 bp
800 bp
600 bp
400 bp
A-C expected size= 1919 bp
B-C expected size= 1903 bp
c
Exon 9
Exon 10
Long ACE2 transcript
...|GCT|TGG|GAC|CTG|GGG|AAG|GGC|GAC|TTC|AGG|ATC|CTT|ATG|TGC|ACA|AAG|GTG|ACA|ATG|...
347 amino acid + A W D L G K G D F R I L M C T K V T M + 439 amino acid
815 aa protein
Exon 9a
Exon 10
Short ACE2 transcript
ATG|AGG|GAA|GCA|GGC|TGG|GAC|AAA|GGA|GGG|AGG|ATC|CTT|ATG|TGC|ACA|AAG|GTG|ACA|ATG|...
M R E A G W D K G G R I L M C T K V T M + 439 amino acid
459 aa protein

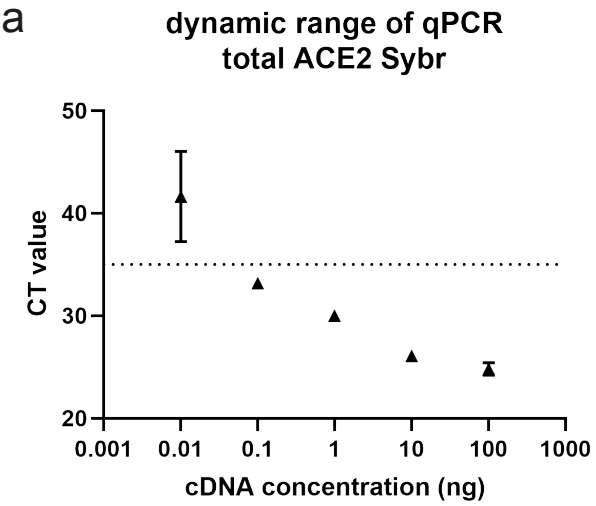

a
Vero E6
HEK293
Caco-2
RPE1
H441
16HBE
BCi-NS1.1
Nasal brushing
Nasal ALI
Bronchial brushing
Bronchial ALI
neg control
200 bp
150 bp
126 bp
short ACE2 transcript
200 bp
150 bp
157 bp
long ACE2 transcript
b
total ACE2
long ACE2
short ACE2
baseline expression (-dCT)
0
-5
-10
-15
-20
Vero E6
HEK293
Caco-2
RPE1
H441
16HBE
BCi-NS1.1
ex vivo NECs
nasal ALI
ex vivo BECs
bronchial ALI
c
long ACE2 transcript
short ACE2 transcript
reads mapping to exon/exon junction per million reads
1.2
1
0.8
0.6
0.4
0.2
0
Nasal ALI day 1
Nasal ALI day 4
Nasal ALI day 8
Nasal ALI day 14
Nasal ALI day 28
Nasal ALI day 63
Nasal ALI day 84
Nasal brushing
d
nasal
bronchial
baseline ACE2 expression (-dCT)
0
-5
-10
-15
-20
****
*
*
**** P<0.0001
* P<0.05
total ACE2
long ACE2
short ACE2
e
total ACE2
long ACE2
short ACE2
ACE2 expression (-dCT)
0
-5
-10
-15
-20
n.d.
n.d.
n.d.
Heart
Brain
Placenta
Lung
Liver
Skeletal Muscle
Kidney
Pancreas

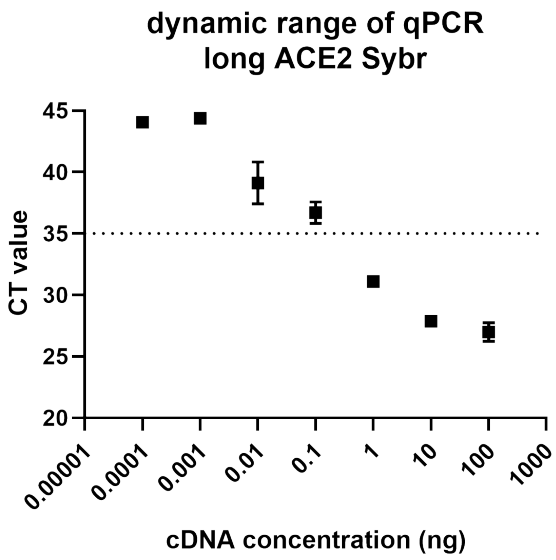

a
anti-ACE2 NTD
AAs 200-300
anti-ACE2 ECTO
AAs 18-740
anti-ACE2 CTD
AAs 788-805
long ACE2
92.4 KDa
TM CIT
1
357
805
short ACE2
52.7 KDa
10 AAs
insertion
b
Vero E6
HEK293
Caco-2
RPE1
H441
16HBE
BCi NS1.1
anti-ACE2
CTD
anti-beta
actin
c
Vero E6
Nasal ALI
Bronchial ALI
control
peptide
blocking
peptide
anti-ACE2
CTD
anti-beta
actin
d
PNGase
anti-ACE2
CTD
anti-ACE2
NTD
anti-beta
actin
e
In vitro differentiated PBECs
DAPI
F-Actin
α-Tubulin
ACE2
anti-ACE2 ECTO
anti-ACE2 CTD
f
+Ca++
motile
cilia
deciliated
BCi NS1.1 whole
BCi NS1.1 deciliated
wash
cilia fraction
anti-ACE2 CTD
g
ratio long/short ACE2
whole
deciliated
cilia fraction
*

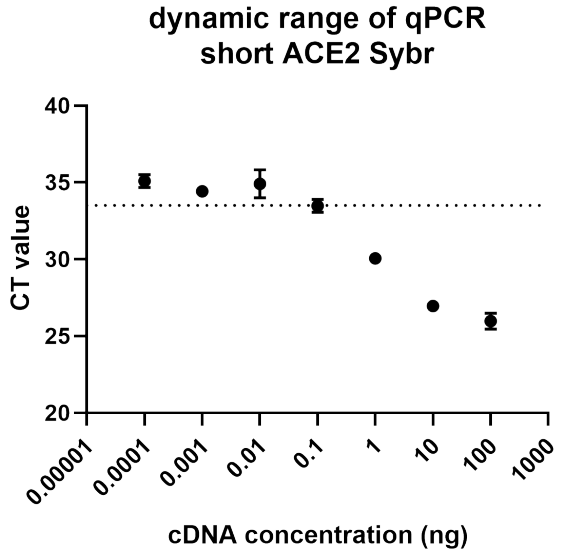

a
Undifferentiated PBEC monolayer cultures
total ACE2
long ACE2
short ACE2
relative expression (2-ddCT)
control
IFNα 100
IFNα 1000
IFNβ 100
IFNβ 1000
IFNγ 100
IFNγ 1000
IFNλ 100
b
In vitro differentiated NECs
c
In vitro differentiated PBECs
total ACE2
long ACE2
short ACE2
control
UV-RV16
RV16
In vitro differentiated BCi-NS1.1
d
SARS-CoV-2 infection
(normalised CoV-N1 expression)
1h after infection
72h after infection
e
f
MX1
IP10
control
1h after infection
72h after infection
In vitro differentiated PBECs
g
ACE2 expression
(Δshort-long -dCT)
UV-RV16
RV16
healthy
severe asthma
h
IFNB1 expression (2-ddCT)
short ACE2 expression (2-ddCT)
long ACE2 expression (2-ddCT)
r= 0.5765
p=0.0032
r= 0.0539
p=0.8024

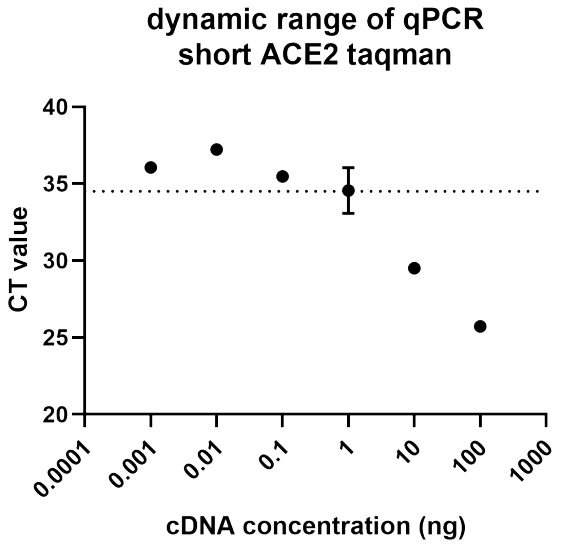

CCNO
400
300
200
100
0 1 4 8 14 21 28 63
OFD1
12.5
10.0
7.5
5.0
2.5
0 1 4 8 14 21 28 63
DNAAF4
0.3
0.2
0.1
0.0
0 1 4 8 14 21 28 63
DNAAF3
75
50
25
0
0 1 4 8 14 21 28 63
MCIDAS
25
20
15
10
5
0 1 4 8 14 21 28 63
FOXJ1
1250
1000
750
500
250
0
0 1 4 8 14 21 28 63
Transcripts per million (TPM)
Timepoint of RNA extraction
from ALI culture

a

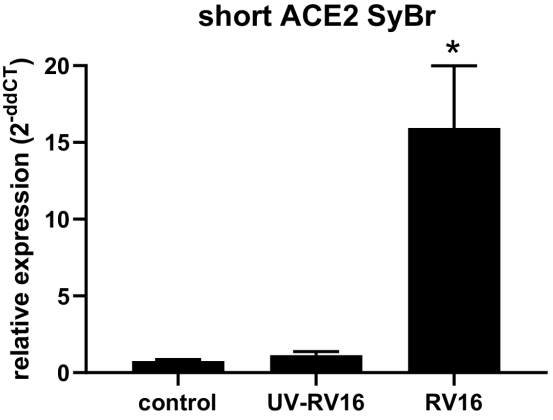

dynamic range of qPCR
total ACE2 Sybr
CT value
50
40
30
20
0.001
0.01
0.1
1
10
100
1000
cDNA concentration (ng)

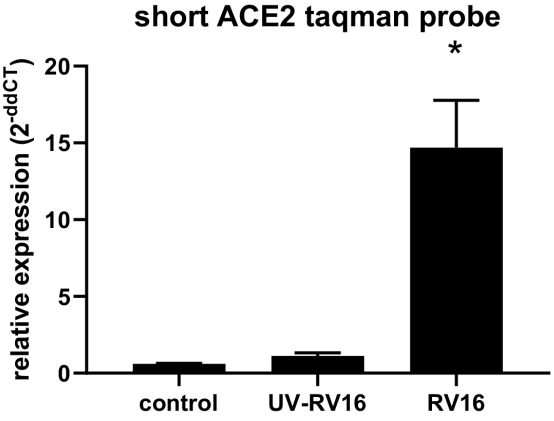

dynamic range of qPCR
long ACE2 Sybr
CT value
45
40
35
30
25
20
0.00001
0.0001
0.001
0.01
0.1
1
10
100
1000
cDNA concentration (ng)

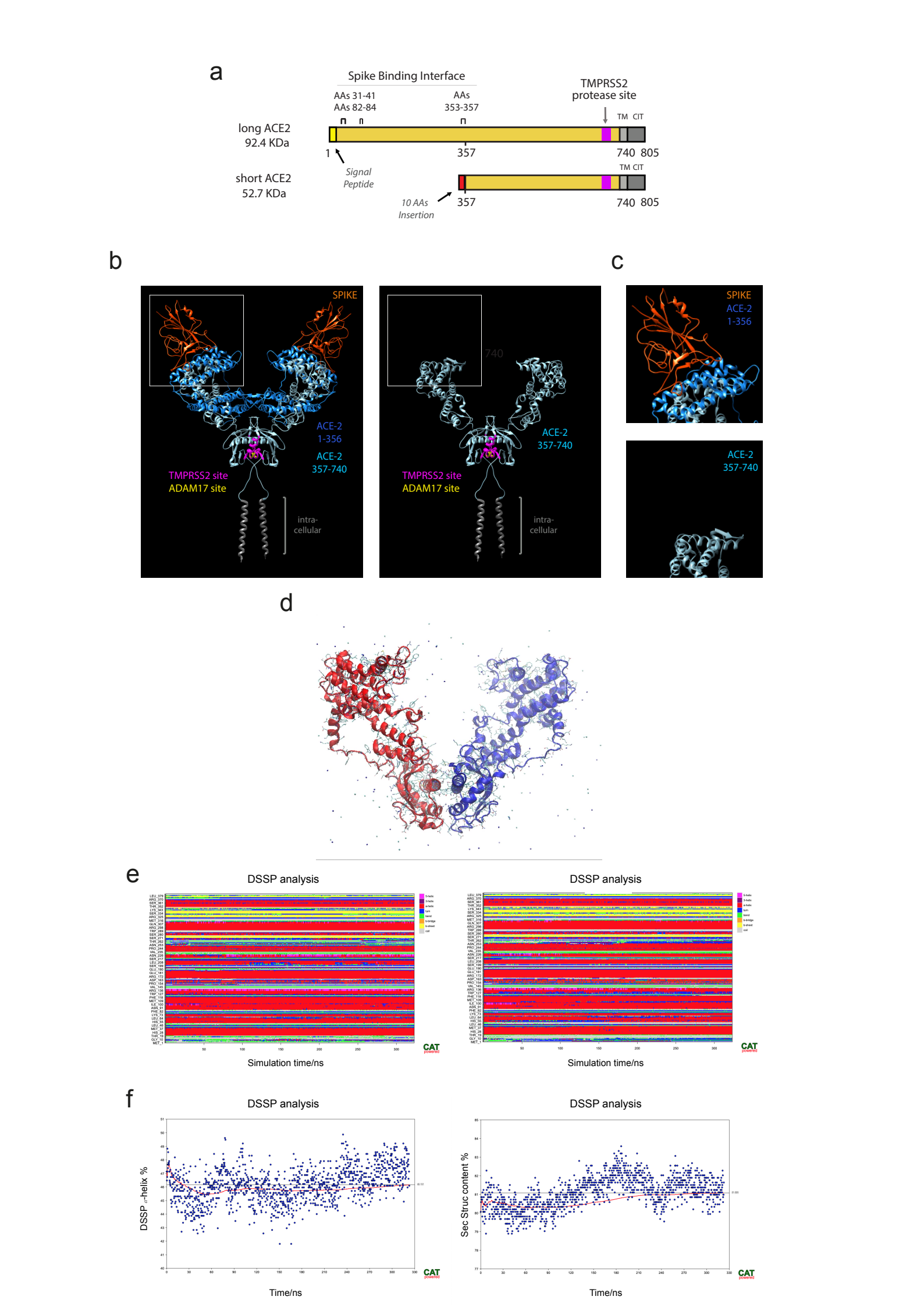

dynamic range of qPCR
short ACE2 Sybr
CT value
40
35
30
25
20
0.00001
0.0001
0.001
0.01
0.1
1
10
100
1000
cDNA concentration (ng)


dynamic range of qPCR
short ACE2 taqman
CT value
40
35
30
25
20
0.0001
0.001
0.01
0.1
1
10
100
1000
cDNA concentration (ng)

b


short ACE2 SyBr
relative expression (2-ddCT)
20
15
10
5
0
*
control
UV-RV16
RV16

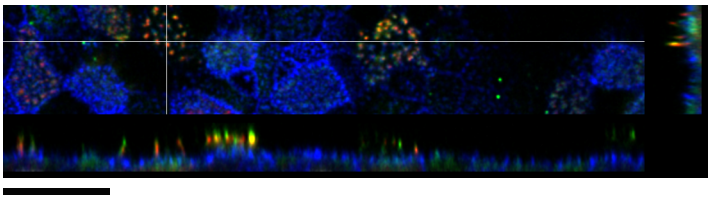

short ACE2 taqman probe
relative expression (2-ddCT)
20
15
10
5
0
*
control
UV-RV16
RV16

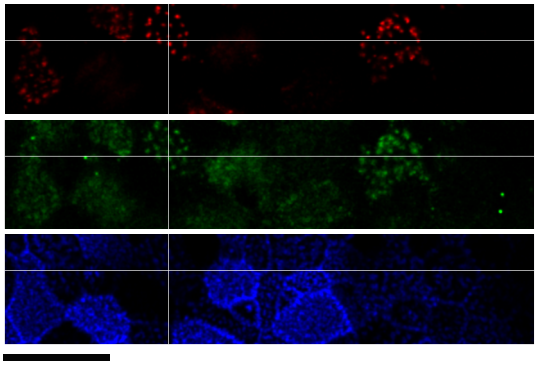

a
Spike Binding Interface
AAs 31-41
AAs 82-84
AAs
353-357
TMPRSS2
protease site
TM CIT
long ACE2
92.4 KDa
1
357
740 805
Signal
Peptide
short ACE2
52.7 KDa
TM CIT
10 AAs
Insertion
357
740 805
b
SPIKE
ACE-2
1-356
ACE-2
357-740
TMPRSS2 site
ADAM17 site
intra-
cellular
740
ACE-2
357-740
TMPRSS2 site
ADAM17 site
intra-
cellular
c
SPIKE
ACE-2
1-356
ACE-2
357-740
d
e
DSSP analysis
Simulation time/ns
DSSP analysis
Simulation time/ns
f
DSSP analysis
DSSP α-helix %
Time/ns
DSSP analysis
Sec Struc content %
Time/ns

a

ACE2 ALPHA-TUBULIN F-ACTIN DAPI

anti-ACE2 CTD

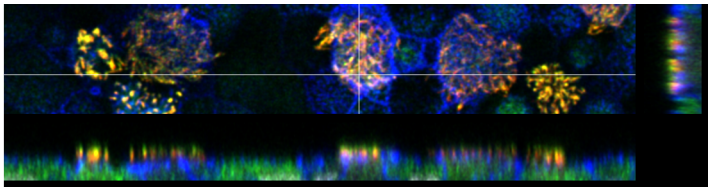


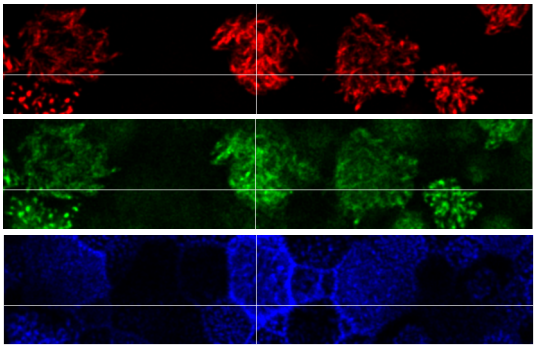


b

ACE2 ALPHA-TUBULIN F-ACTIN DAPI

anti-ACE2 ECTO

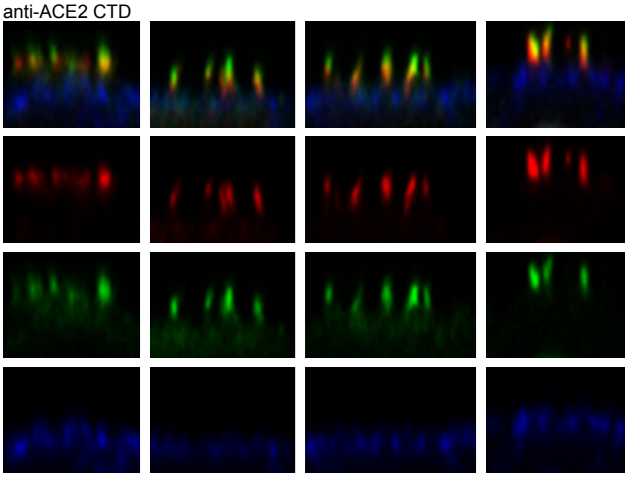




c

ACE2 ALPHA-TUBULIN F-ACTIN DAPI

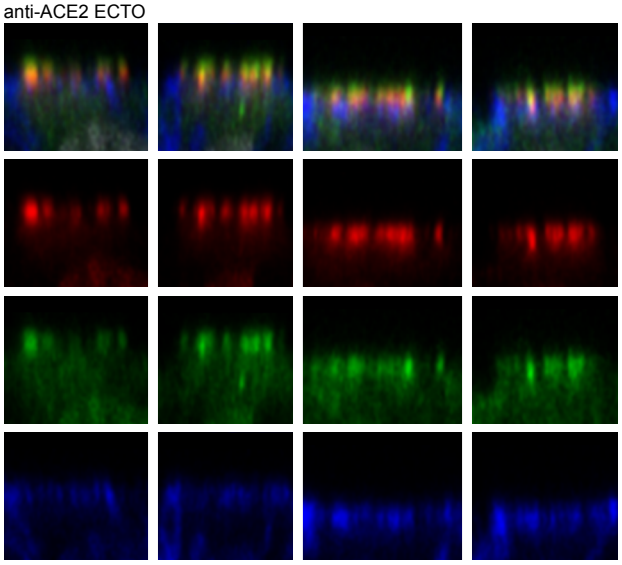
anti-ACE2 CTD



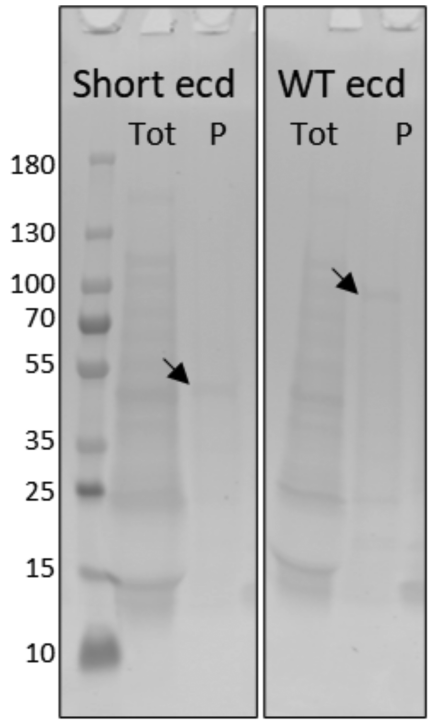
d

ACE2 ALPHA-TUBULIN F-ACTIN DAPI

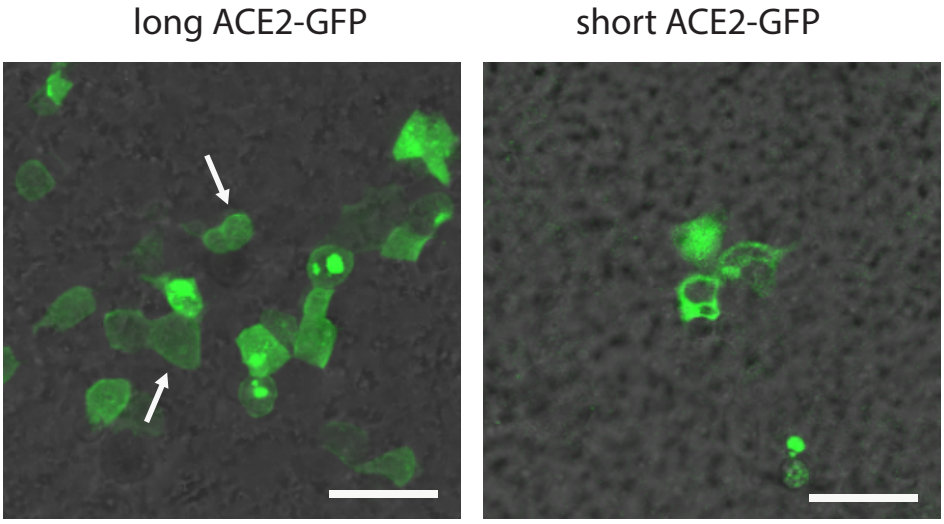
anti-ACE2 ECTO



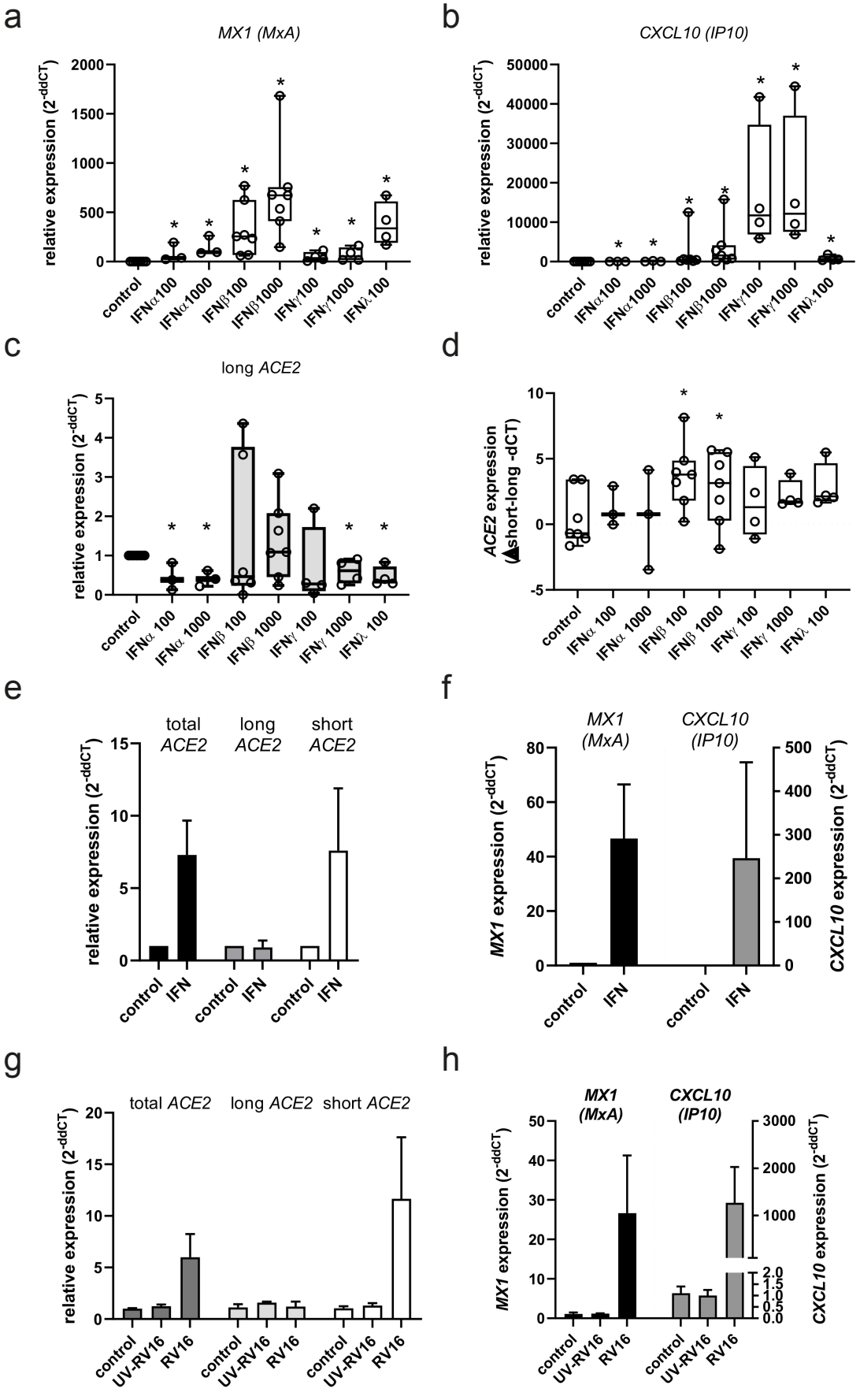
a

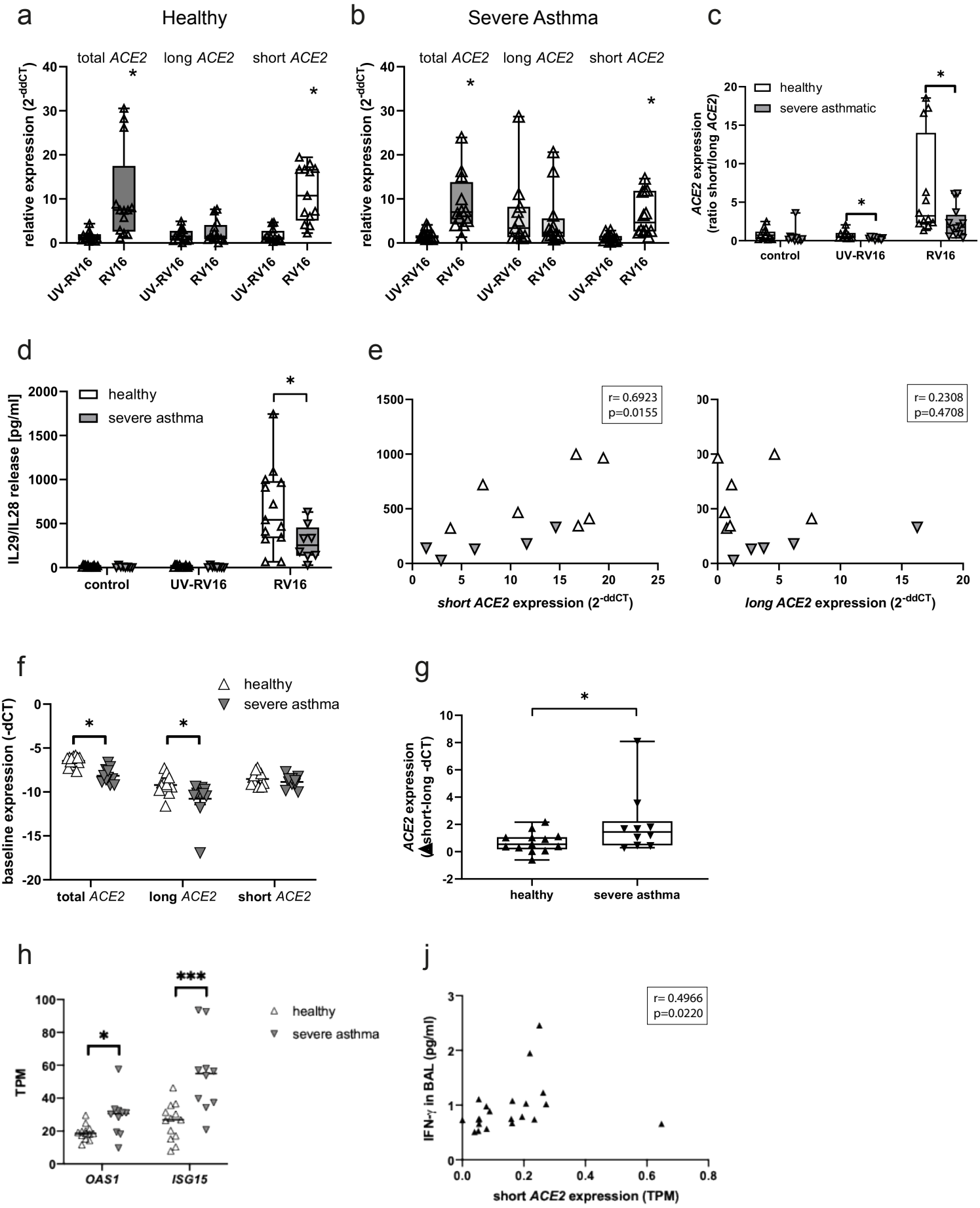

Short ecd
Tot
P
WT ecd
Tot
P
180
130
100
70
55
35
25
15
10

b


long ACE2-GFP
short ACE2-GFP

H441


a
MX1 (MxA)
relative expression (2-ddCT)
control
IFNα100
IFNα1000
IFNβ100
IFNβ1000
IFNγ100
IFNγ1000
IFNλ100
b
CXCL10 (IP10)
c
long ACE2
d
ACE2 expression (Δshort-long -dCT)
e
total ACE2
long ACE2
short ACE2
IFN
f
MX1 (MxA)
CXCL10 (IP10)
MX1 expression (2-ddCT)
CXCL10 expression (2-ddCT)
g
UV-RV16
RV16
h


a
Healthy
total ACE2
long ACE2
short ACE2
relative expression (2-ddCT)
UV-RV16
RV16
b
Severe Asthma
c
healthy
severe asthmatic
ACE2 expression (ratio short/long ACE2)
control
UV-RV16
RV16
d
healthy
severe asthma
IL29/IL28 release [pg/ml]
e
r= 0.6923
p=0.0155
short ACE2 expression (2-ddCT)
r= 0.2308
p=0.4708
long ACE2 expression (2-ddCT)
f
baseline expression (-dCT)
g
ACE2 expression (short-long -dCT)
h
TPM
OAS1
ISG15
j
r= 0.4966
p=0.0220
IFN-γ in BAL (pg/ml)
short ACE2 expression (TPM)