The Contrasting Role of Nasopharyngeal Angiotensin Converting Enzyme 2 (*ACE2*) Expression in SARS-CoV-2 Infection: A Cross-Sectional Study of People Tested for COVID-19 in British Columbia

Authors: Aidan M. Nikiforuk[1,4*] (MSc.), Kevin S. Kuchinski[1,2*] (BSc.), David D.W. Twa[2,3*] (PhD), Christine D. Lukac[1] (MPH), Hind Sbihi[1,4] (PhD), C. Andrew Basham[1,4] (MSc.), Christian Steidl[3] (MD), Natalie A. Prystajecky[1,2] (PhD), Agatha N. Jassem[1,2] (PhD), Mel Krajden[1,2] (MD), David M, Patrick[1,4] (MD) and Inna Sekirov[1,2] (PhD).

[*]these authors contributed equally to the work.

[1]British Columbia Centre for Disease Control, Vancouver, V5Z 4R4, British Columbia, Canada
[2]Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, V6T 1Z4, British Columbia, Canada
[3]British Columbia Cancer Research Centre, Vancouver, V5Z 1L3, British Columbia, Canada
[4]School of Population and Public Health, University of British Columbia, Vancouver, V6T 1Z4, British Columbia, Canada

Corresponding Author: Dr. Inna Sekirov MD/ PhD, British Columbia Centre for Disease Control, Vancouver, British Columbia, Canada, V5Z 4R4 and Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, British Columbia, Canada, V6T 1Z4.
email: inna.sekirov@bccdc.ca

Summary

Background

Angiotensin converting enzyme 2 (*ACE2*) serves as the host receptor for SARS-CoV-2, with a critical role in viral infection. We aim to understand population level variation of nasopharyngeal *ACE2* expression in people tested for COVID-19 and the relationship between *ACE2* expression and SARS-CoV-2 viral RNA load, while adjusting for expression of the complementary protease, Transmembrane serine protease 2 (*TMPRSS2)*, soluble *ACE2*, age, and biological sex.

Methods

A cross-sectional study of n=424 participants aged 1-104 years referred for COVID-19 testing was performed in British Columbia, Canada. Participants who tested negative or positive for COVID-19 were matched by age and biological sex. Viral and host gene expression was measured by quantitative reverse-transcriptase polymerase chain reaction. Bivariate analysis and multiple linear regression were performed to understand the role of nasopharyngeal *ACE2* expression in SARS-CoV-2 infection. The *ACE2* gene was targeted to measure expression of transmembrane and soluble transcripts.

Findings

Analysis shows no association between age and nasopharyngeal *ACE2* expression in those who tested negative for COVID-19 (P=0·092). Mean expression of transmembrane (P=1·2e-4), soluble *ACE2* (P<0·0001) and *TMPRSS2* (P<0·0001) differed between COVID-19-negative and -positive groups. In bivariate analysis of COVID-19-positive participants, expression of transmembrane *ACE2* positively correlated with SARS-CoV-2 RNA viral load (P<0·0001), expression of soluble *ACE2* negatively correlated (P<0·0001), and no correlation was found with *TMPRSS2 (P*=0·694*)*. Multivariable analysis showed that the greatest viral RNA loads were observed in participants with high transmembrane *ACE2* expression (*B*=0·886, 95%CI:[0·596 to 1·18]), while expression of soluble *ACE2* may protect against high viral RNA load in the upper respiratory tract (*B*= -0·0990, 95%CI:[-0·176 to -0·0224]).

Interpretation

Nasopharyngeal *ACE2* expression plays a dual, contrasting role in SARS-CoV-2 infection of the upper respiratory tract. Transmembrane *ACE2* positively correlates, while soluble *ACE2* negatively correlates with viral RNA load after adjusting for age, biological sex and expression of TMPRSS2.

Funding

This project (COV-55) was funded by Genome British Columbia as part of their COVID-19 rapid response initiative.

Research in Context

Evidence before this study

We conducted a MEDLINE ® search using the MeSH topic terms: "angiotensin converting enzyme 2 expression" and "SARS" and "age", restricting the search to English-language reports published from January 1st, 2020. The search returned n=98 articles, eighty-eight of which reported primary research; these were further filtered by the MeSH qualifiers "epidemiology" and "virology" to provide n=43 articles. Search results were further restricted to the MeSH headings: age, virus replication, host pathogen interactions, nasal mucosa, RNA viral load and real-time polymerase chain reaction, which returned n=24 results. We read through the abstracts of these twenty-four papers and manually selected n=4 for full review. This review provided evidence that *ACE2* expression is greater in the upper respiratory tract than the lower respiratory tract, when measured by single-cell RNA sequencing, immunohistochemistry and high-sensitivity RNA in-situ mapping. In the upper airway, *ACE2* mRNA abundance closely correlates with protein concentration. A reverse genetics study demonstrated that a variable SARS-CoV-2 infection gradient occurs in the respiratory tract, with highest viral loads expected in the upper airway. In several patient cohorts, upper respiratory expression of *ACE2* was significantly increased in those who smoke, in this analysis multivariable adjustment of age suggested limited confounding.

Added Value of this study

We measured *ACE2* expression in the context of COVID-19 testing to investigate the role of nasopharyngeal *ACE2* in SARS-CoV-2 infection. Our findings support previous work: the strong correlation we observe between nasopharyngeal *ACE2* expression and SARS-CoV-2 load also suggests an infection gradient across the human airway. Greater viral loads are expected in tissue with high transmembrane ACE2 expression. No observed relationship between age and nasopharyngeal *ACE2* in COVID-19-negative participants suggests that upper airway *ACE2* expression is independent of the RAAS pathway. We are the first to measure endogenous expression of soluble *ACE2* in the upper airway of people tested for COVID-19 and include this measure in multivariable analysis. Importantly, nasopharyngeal expression of soluble *ACE2* negatively correlates with viral RNA load, inferring a protective role at the population level.

Implications of all the available evidence

Considering all available evidence, *ACE2* may play a dual, contrasting role in SARS-CoV-2 infection of the upper airway. Expression of transmembrane ACE2 positively correlates with SARS-CoV-2 RNA load, while expression of soluble ACE2 shows a negative association. Expression of nasopharyngeal ACE2 does not seem to correlate with age, as would be expected in the lower respiratory tract. Risk factors such as smoking may affect the baseline risk of high SARS-CoV-2 RNA loads in those infected. Factors associated with endogenous nasopharyngeal expression of soluble *ACE2* require further investigation.

Introduction

In December 2019, clusters of viral pneumonia were reported in Wuhan, China. A novel highly pathogenic human coronavirus was isolated and named severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the etiological agent of coronavirus disease (COVID-19).[1] SARS-CoV-2 utilizes the same host receptor and protease for cell entry as the human coronaviruses SARS-CoV and HCoV-NL63. The receptor angiotensin converting enzyme 2 (*ACE2*) mediates cellular entry, while transmembrane serine protease 2 (TMPRSS2) serves as a complementary host factor.[2,3] In human physiology, *ACE2* has a cardiovascular protective and anti-inflammatory role, as a constituent of the renin-angiotensin-aldosterone-system (RAAS).[4] Interestingly,

expression of *ACE2* in the nasopharyngeal tract exceeds that in alveolar tissue, explaining initiation of SARS-CoV-2 infection in the upper respiratory tract.[5] Transcription of *ACE2* produces at least two dominant mRNA transcripts responsible for translation into soluble and membrane-bound protein isoforms.[6] Although the transmembrane isoform has been shown to be crucial for viral entry into host cells, the role of soluble *ACE2* remains uncharacterized, though some evidence exists that it protects against SARS-CoV-2 infection.[7,8] Transmembrane serine protease 2 (*TMPRSS2*) contributes to SARS-CoV-2 cell entry by cleaving the viral spike protein into a conformational form necessary for membrane fusion.[2] Unlike *ACE2*, *TMPRSS2* expression occurs more stably across upper airway tissue and alternative enzymes such as cathepsin B/L may perform its role in viral infection interchangeably.[5] To understand the importance of *ACE2* expression in SARS-CoV-2 infection, we performed a cross-sectional study of people tested for COVID-19 in British Columbia, Canada. The study aims to investigate the relationship between i) nasopharyngeal expression of *ACE2* and age in COVID-19-negative participants, ii) nasopharyngeal expression of host genes by COVID-19 test result and iii) nasopharyngeal expression of transmembrane *ACE2* and viral RNA load in those who tested COVID-19-positive adjusting for age, biological sex, expression of soluble *ACE2* and *TMPRSS2*. Characterizing expression of nasopharyngeal *ACE2* in a large group of participants tested for COVID-19 will increase our knowledge of host genes involved in viral infection and may allow for assessment of baseline risk.

Methods

Study Design and Participants

We performed cross-sectional sampling of people tested for COVID-19 at the British Columbia Center for Disease Control Public Health Laboratory (BCCDC-PHL) from 24/3/2020-9/5/2020. At the time of sampling, provincial health guidelines required a clinical indication for referral of a COVID-19 test. Inclusion criteria were applied to select study participants whose diagnostic specimens were: tested centrally at the BCCDC-PHL, collected by nasopharyngeal swab, the first test administered by provincial health number, suspended in Hologic Aptima™ media, negative for concurrent Influenza A, B or Respiratory syncytia virus infection, stored at -80°C following RNA extraction and for whom host gene expression was successfully measured by quantitative reverse-transcriptase polymerase chain reaction (qRT-PCR). People meeting the

inclusion criteria (n=444), were excluded (n=16) if samples had a SARS-CoV-2 E gene cycle threshold (Ct) value of >=38 by qRT-PCR. COVID-19-positive cases (n=212) were matched in a 1:1 ratio with those that tested negative for COVID-19 by age and biological sex (Figure S1).

Ethical approval for the study was obtained from the University of British Columbia human ethics board (H20-01110). Demographic variables of age and biological sex were drawn from public health laboratory data. Laboratory methods were performed in a College of American Pathologists accredited laboratory with externally validated qRT-PCR assays. [9–12]

Procedures

Nasopharyngeal samples collected in Hologic Aptima™ media were stored at 4°C before RNA extraction using the Viral RNA isolation kit on the MagMAX-96™ platform (ThermoFisher).[13]

Host and viral gene targets were assayed by qRT-PCR on the Applied Biosystems 7500 Real-Time PCR platform using TaqMan FastVirus 1-step polymerase (ThermoFisher). Total reaction volumes equaled 20*u*l, with 5*u*l of RNA template, 1*u*l of 20x primer/probe, 5*u*l Fast Virus and 9*u*l of nuclease free water per reaction. Cycling conditions were set to: 50°C for 5 minutes, 95°C for 20 seconds followed by 40 cycles of 95°C for 15 seconds and 60°C for 1 minute. A multiplex qRT-PCR reaction targeting SARS-CoV-2 envelope (E) and host ribonuclease P (RNaseP) was used to diagnosis acute viral infection by presence of viral RNA.[14] Expression of RNaseP was used to ensure sample quality and control for sampling variation.[14] Participants were diagnosed as COVID-19 positive with an E gene Ct value of <38. The E gene Ct values were transformed to genome equivalents per millilitre by making a 5-fold, 1:10 standard curve of SARS-CoV-2 synthetic RNA- MN908947·3 (Twist Bioscience) (Table S1). Commercially available primer probe sets were used to amplify the host gene targets TMPRSS2[12] (Hs00237175_m1) (ThermoFisher) and *ACE2* [10] (Hs01085333_m1, HS01085340_m1) by multiplex qRT-PCR with a glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) control[11] (Hs02758991_g1) (ThermoFisher). Relative gene expression was calculated between *TMPRSS2*, *ACE2* and *GAPDH* using the $2^{-\Delta\Delta Ct}$ method.[15] Transmembrane *ACE2* was defined from the Hs01085333_m1 (ThermoFisher) gene target, which exon spans the transmembrane domain. Soluble *ACE2* was defined as absolute relative expression between the Hs01085333_m1 (ThermoFisher) and HS01085340_m1 (ThermoFisher) gene targets (Figure

S3). The HS01085340_m1 target, does not include the transmembrane domain and is positioned between two motifs essential for SARS-CoV receptor binding (Figure S3).

Statistical Analysis

De-identified data reporting participant COVID-19 test result, site of sample collection, type of collection media, accessioning laboratory, age and biological sex was accessed from public health laboratory records. Matching of COVID-19-positive people, who met the described inclusion, exclusion criteria, to COVID-19-negative participants by age and biological sex in a 1:1 ratio was performed by a nearest neighbor algorithm (n=424).[16] Bivariate analysis was performed between: age, biological sex, viral RNA load, *TMPRSS2*, soluble or transmembrane *ACE2*, and COVID-19 test result. The balance of covariates between test groups was examined post matching.

The relationship between age and nasopharyngeal expression of transmembrane *ACE2* was examined in COVID-19-negative participants over the age of 18 by both linear regression and categorization of age into 9-year intervals. Differences in *ACE2* expression by age category were tested by analysis of variance.

Differences in mean expression of host genes by COVID-19 diagnosis was further examined by Levene's test for equal variance and a two-tailed, t-test assuming non-equal variance.[17] Correlation between host gene expression and viral RNA load was analyzed by simple linear regression. Multivariable analysis was performed by multiple linear regression, variable importance was assessed by the partial f-test.[18] Collinearity was examined by the variable inflation factor with a cut-off of 10.[19] The common cause criterion was applied to control for measured confounding, effect modification was incorporated when found statistically significant.[20] All analysis was performed in RStudio version 1.2.5042 using the packages: car, ggsci, tidyverse, dataexplorer, ggpubr, lmtest, publish, forcats, matchit, tableone and effects. [21]

Role of the Funding Source and Data Stewards

The funder and data stewards played no role in the study design, analysis or interpretation of the results. As such, interpretation of the results does not reflect the views of the funding organization or data stewards. The study was performed as research at the University of British Columbia and British Columbia Centre for Disease Control.

Results

The analytic dataset contains an age- and biological sex-matched sample of n=424 participants tested for COVID-19 in British Columbia from 24/3/2020-9/5/2020 (Figure S1). Participant characteristics are shown in Table 1. The mean age of n=212 COVID-19-positive participants was 61·6, 47·6% were biologically male. In n=212 COVID-19-negative participants, the mean age was 62·1, 48·6% were biologically male. Bivariate analysis between the examined covariates and diagnosis (COVID-19-negative or -positive) shows balance of age (P=0·83) and biological sex (P=0·92) between groups. Viral RNA (E gene Ct) was only detected in COVID-19 positive participants, with an average Ct value of 28·2. Relative expression of transmembrane *ACE2*, soluble *ACE2*, and *TMPRSS2* statistically differed by COVID-19 test result (Table 1).

Table 1: Structure and Characteristics of Analytic Data Stratified by COVID-19 Test Result (n=424).

			COVID-19 Test Result			
Variable Name	**Level**	**Total (n)**	**Negative**	**Positive**	**Chi X^2**	**T-Test**
	..	424	212	212	..	..
	..		..	..	..	..
Age (mean [SD])	..	424	62.2 [24.3]	61.7 [23.5]	..	0.83
Sex (n [%])	..	..	..	..	0.92	..
	Male	204	103 [48.6]	101 [47.6]	..	..
	Female	220	109 [51.4]	111 [52.4]	..	..
E Gene Ct (mean [SD])	..	424	..	28.2 [7.18]	..	..
	..	..	..	..	..	..
Transmembrane ACE2 (mean [SD])	..	424	(0.00) [1.08]	(-0.609) [1.90]	..	1.2e-4
Soluble ACE2 (mean [SD])	..	424	(0.00) [0.625]	(-0.898) [1.47]	..	<0.0001
TMPRSS2 Total (mean [SD])	..	424	(0.00) [0.797]	(-1.35) [1.60]	..	<0.0001

Participants who tested negative or positive by COVID-19 were matched by age and biological sex, balance was checked between groups post matching by chi X^2 or t-test (α=0·05). Relative gene expression of ACE2 and TMPRSS2 was normalized to the negative group by the $2^{-\Delta\Delta Ct}$ method. [15]

Relationship between ACE2 expression and age in COVID-19-negative participants

No relationship was found between nasopharyngeal *ACE2* expression and age in COVID-19 negative participants between 19 and 98 years of age with age as a continuous variable by

linear regression (P=0·076). This finding was reproduced when age was categorized into nine-year intervals (P=0·092) (Figure 1).

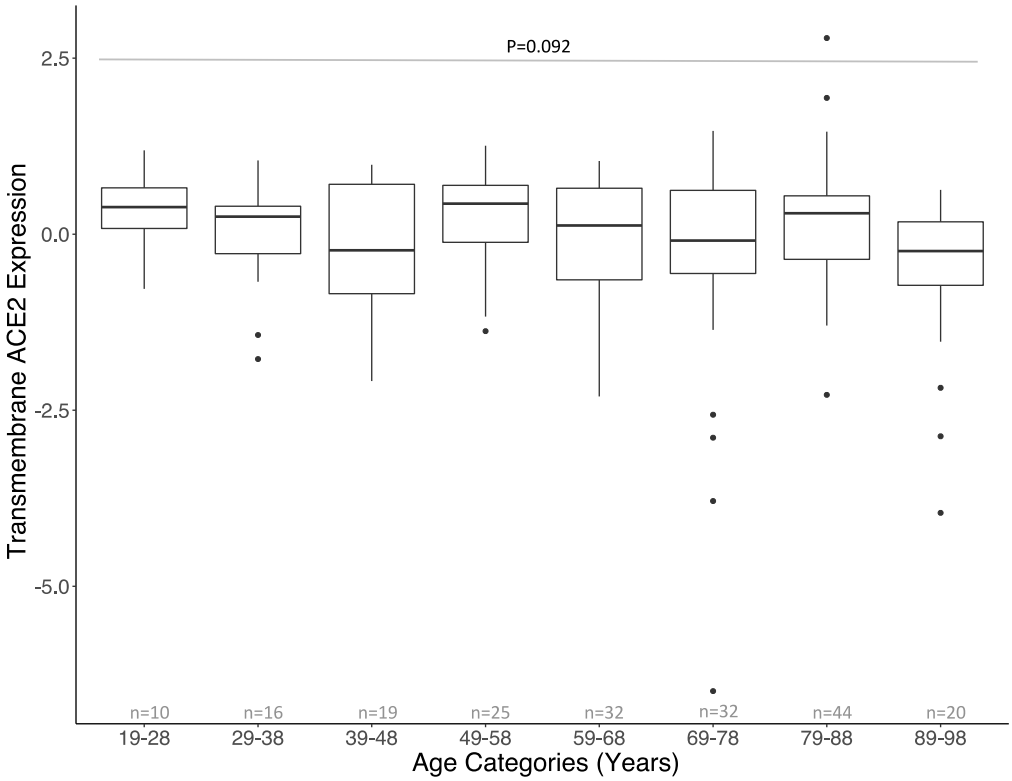

Figure 1: Relationship between age and nasopharyngeal transmembrane *ACE2* expression in COVID-19-negative participants. Boxplots of transmembrane *ACE2* expression by nine-year age categories in COVID-19-negative participants between the ages of 19 and 98 (n=198); boxes represent the Q1-Q3 interquartile range, whiskers represent 1.5x the Q1 or Q3 and horizontal lines the median transmembrane ACE2 expression by age category. Nine-year categories were selected to optimize the distribution of observations between groups. Participants who tested negative younger than 19 or older than 98 were excluded based on n<10 observations per age group. No difference was detected in mean transmembrane *ACE2* expression among age categories (ANOVA, P=0·092).

Nasopharyngeal expression of ACE2 and TMPRSS2 by COVID-19 test result

Bivariate analysis showed a significant mean difference in transmembrane, soluble *ACE2* and *TMPRSS2* expression between COVID-19-negative and -positive participants (Figure 2). These differences were further examined to detect non-equal variance for all host gene targets by COVID-19 test result. Assuming non-equal variance, mean expression differed for transmembrane *ACE2* (P=1·2e-4), soluble *ACE2* (P<0·0001) and *TMPRSS2* (P<0·0001) between

those testing COVID-19-negative or -positive (Figure 2). Expression of all three host genes was lower in COVID-19-positive participants.

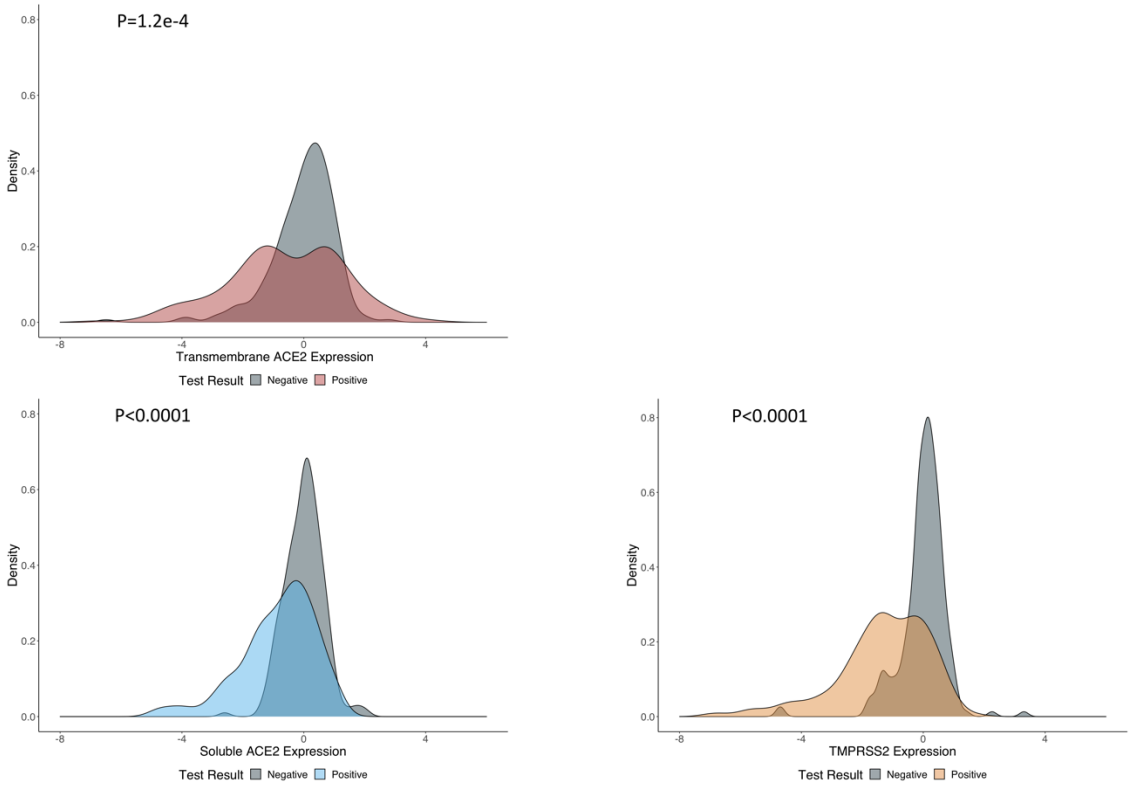

Figure 2: Relative nasopharyngeal gene expression of targeted host genes by COVID-19 test result. Gene expression is portrayed in kernel density plots stratified by COVID-19 test result. Probability densities of relative host gene expression are shown by positive (red, transmembrane *ACE2*; blue, soluble *ACE2*; yellow, *TMPRSS2*) and negative COVID-19 test results (grey). Levene's test was used to detect non-equal variance in gene expression for all host targets between COVID-19-negative and -positive participants. A two-tailed, t-test was used to examine mean difference in host gene expression by COVID-19 test result assuming unequal variance: transmembrane *ACE2* (P=1·2e-4), soluble *ACE2* (P<0·0001) and *TMPRSS2* (P<0·0001).

Association between host gene expression and viral RNA load in COVID-19-positive participants

Correlations between each host gene and viral RNA load were examined by simple linear regression. Nasopharyngeal expression of transmembrane *ACE2* positively correlated with viral RNA (P<0·0001). Expression of soluble *ACE2* negatively correlated with viral RNA (P<0·0001). No correlation was found between *TMPRSS2* expression and viral RNA (P=0·69)

(Figure S2). Multiple linear regression estimated that a one-unit change in transmembrane *ACE2* expression increases viral RNA load by 0·886 Log_{10} GE/mL (95%CI: 0·596 to 1·18) adjusting for age, biological sex, expression of soluble *ACE2* and *TMPRSS2* (Table 2). No collinearity was detected between covariates. Biological sex could have been dropped from the model as evident from comparison between nested models (P=0·76), but was kept for validity. A partial f-test indicated effect modification between transmembrane and soluble *ACE2* expression (P=0·010) (Table 2). No effect modification was observed between transmembrane expression of *ACE2* and *TMPRSS2* (P=0·23). High expression of soluble *ACE2* decreases viral RNA load (*B*= -0·0990, 95%CI: [-0·176 to -0·0224]) (Table 2). The association between transmembrane *ACE2* expression and viral RNA in nasopharyngeal tissue differs by the concomitant level of soluble *ACE2* expressed (Figure 3). Effect modification was visualized by categorizing soluble *ACE2* by the relative mean expression +/- 2 standard deviations of all study participants (Low= -2·86, Mean= -0·444, High= 1·98, SD=2·43) (Figure 3).

Table 2: Unadjusted, Adjusted and Effect Modification Linear Regression Models of SARS-CoV-2 Viral RNA Load

Variables	Unadjusted, *B*(95%CI)	Adjusted, *B*(95%CI)	Effect Modification, *B*(95%CI)
Intercept [a]	5.44 (5.12 to 5.77)	2.55 (1.71 to 3.40)	2.48 (1.65 to 3.31)
Transmembrane ACE2 Expression [b]	0.493 (0.329 to 0.656)	0.886 (0.596 to 1.18)	0.624 (0.274 to 0.975)
Soluble ACE2 Expression [c]	--	-0.243 (-0.494 to 0.00834)	-0.484 (-0.794 to - 0.174)
TMPRSS2 Expression [d]	--	-0.873 (-1.21 to -0.533)	-0.713 (-1.07 to -0.336)
Age [e]	--	0.0276 (0.0162 to 0.0389)	0.0259 (0.0146 to 0.0372)
Biological Sex [f]	--	0.0778 (-0.424 to 0.579)	0.0735 (-0.422 to 0.569)
Transmembrane ACE2* Soluble ACE2 Expression [g]	--	--	-0.0990 (-0.176 to -0.0224)

Regression of SARS-CoV-2 viral RNA load by Transmembrane *ACE2*, Soluble *ACE2* and *TMPRSS2* Expression, Age and Biological Sex.

[a] Viral RNA load in females, measured by Log_{10} viral genome equivalents per millilitre (Log_{10} GE/mL) targeting the SARS-CoV-2 *E* gene.

[b] Transmembrane *ACE2* Expression ($2^{-\Delta\Delta Ct}$) in the nasopharyngeal epithelium of SARS-CoV-2 positive participants.

[c] Soluble *ACE2* expression ($2^{-\Delta\Delta Ct}$) in the nasopharyngeal epithelium of SARS-CoV-2 positive participants.

[d] *TMPRSS2* expression ($2^{-\Delta\Delta Ct}$) in the nasopharyngeal epithelium of SARS-CoV-2 positive participants.

[e] Age, measured in years.

[f] Biological sex, effect of being male.

[g] Effect modification term, the effect of transmembrane *ACE2* expression on viral RNA load differs by expression of soluble *ACE2*.

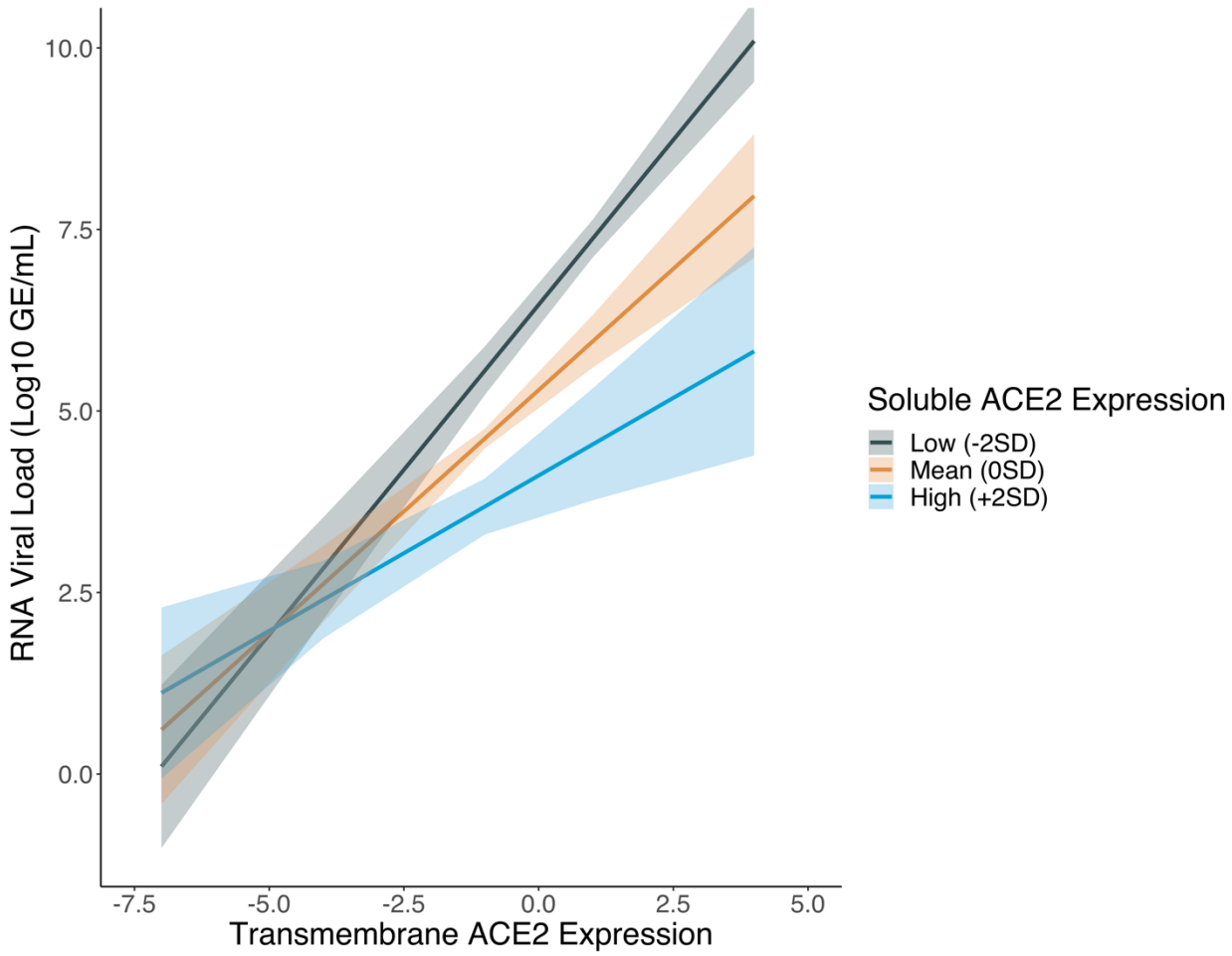

Figure 3: The association between transmembrane *ACE2* expression and SARS-CoV-2 RNA load in nasopharyngeal tissue differs by the amount of soluble *ACE2* expression. Soluble *ACE2* expression was categorized into low, mean and high levels to demonstrate the relationship. The low category (grey) represents soluble *ACE2* expression two standard deviations below the mean expression. The mean category (orange) codes for the mean soluble *ACE2* expression, zero standard deviations. The high category (blue) indicates soluble *ACE2* expression two standard deviations above mean expression. Shaded areas represent 95% confidence intervals, solid lines represent *B* coefficients from multiple linear regression.

Discussion

In the described study, we measured nasopharyngeal *ACE2* and *TMPRRS2* relative gene expression by qRT-PCR in a cross-sectional sample of n=424 participants tested for COVID-19 in British Columbia. Analysis was performed to understand: the relationship between *ACE2* expression and age in COVID-19-negative participants, differences in host gene expression between COVID-19-negative and -positive participants and the role of *ACE2* in SARS-CoV-2

infection in those who tested positive for COVID-19.

Among participants tested negative for COVID-19, we found no relationship between nasopharyngeal *ACE2* expression and age, in a group of adults between 19-98 years of age. This result, agrees with findings from other efforts to characterize nasopharyngeal *ACE2* expression at the population level. Bunyavanich et al. reported a difference in *ACE2* expression levels between children younger than ten years old and young adults from eighteen to twenty-four years old. [22] No difference was evident between young adults, and adults twenty-five or more years in age. [22] Lack of a relationship between nasopharyngeal *ACE2* expression and progressive aging suggests that expression of *ACE2* in the upper airway does not correlate with expression of *ACE2* in the RAAS. The physiological importance of *ACE2* in the RAAS suggests that its expression would increase in response to age as older people have higher prevalence of cardiovascular disease or similar comorbidities (i.e. diabetes). [23] If nasopharyngeal expression of *ACE2* occurs independently of the RAAS pathway, future studies should ascertain what factors if any regulate expression of soluble *ACE2* in nasopharyngeal tissue and variation of *ACE2* expression in COVID-19-negative people over time. Some populations, such as young children, may have differential levels of nasopharyngeal *ACE2* expression which protect against SARS-CoV-2 viral infection.[22,24]

Participants who tested positive for COVID-19 were matched to those who tested negative by age and biological sex to estimate the direct relationship between host gene expression and SARS-CoV-2 test result. This analysis demonstrated that mean expression of transmembrane, soluble *ACE2* and *TMPRSS2* decreased in COVID-19-positive participants, while variation increased. Unfortunately, the limitation of reverse-causality in cross-sectional study design prevents us from understanding whether the presence of SARS-CoV-2 in positive samples affects host gene expression or host gene expression puts people at risk of viral infection. However, if we make the plausible assumption that at least a single round of viral replication must occur in order for qRT-PCR to generate a positive test result, it suggests that the observed variation in host gene expression of participants who test positive for COVID-19 results from viral replication.[25] Numerous studies have demonstrated that coronavirus replication disrupts cellular transcription, as resources required by the cell to produce mRNA are instead sequestered by virus to replicate its genome. [26] Reduced transcription of *ACE2* by SARS-CoV, but not HCoV-NL63, in an *in vitro* model of Vero E6 cell infection was previously

suggested as a pathological mechanism.[3] Putative SARS-CoV-2 disruption of *ACE2* expression could partially explain the apparent association between hypertension, diabetes, and severe COVID-19.[27] Answering these questions; however, requires a longitudinal study design to test the temporal effect of SARS-CoV-2 replication on the expression of *ACE2* and other implicated host genes.

The sample was then restricted to participants who tested positive for COVID-19 to investigate the association between nasopharyngeal *ACE2* expression and SARS-CoV-2 RNA load. Bivariate and multivariable analysis both suggest that *ACE2* plays a dual, contrasting role in SARS-CoV-2 infection. Transmembrane *ACE2* positively correlates with viral RNA load, while soluble *ACE2* may limit viral infection by reducing viral RNA load. Interestingly, effect modification was observed between transmembrane and soluble *ACE2* implicating that the proportion between these molecules may have more importance than absolute expression of either individually. For example, between two people with similar above-average nasopharyngeal transmembrane *ACE2* expression, we would expect a higher viral RNA load in the one with lower soluble *ACE2* expression. Though we are not equipped to characterize the mechanism by which soluble *ACE2* plays a protective role in SARS-CoV-2 infection, identification of this effect at the population level warrants further investigation of the underlying mechanism. Previous work has postulated that soluble *ACE2* restricts SARS-CoV/SARS-CoV-2 infection by acting as a decoy substrate.[28] In a study of virus-*ACE2* dynamics in engineered human tissue, instead of binding to transmembrane *ACE2*, virus particles bound to soluble *ACE2* and were unable to infect susceptible cells.[28] Moreover, recombinant soluble *ACE2* was recently administered to a COVID-19 patient requiring ventilation.[8] In the case report, administration of soluble *ACE2* decreased viral RNA load in the patient's plasma, tracheal aspirate and nasopharyngeal specimens. Additionally, soluble *ACE2* did not interfere with production of IgG and assumed its cardiovascular protective role in the RAAS.[8]

Soluble *ACE2* may also influence the usage of Neuropilin-1 as a co-receptor or alternative receptor for SARS-CoV-2. Function of Neuropilin-1 for cell entry and infectivity may depend upon high SARS-CoV-2 viral load; therefore, expression of soluble *ACE2* may limit SARS-CoV-2 tropism. [29]

We acknowledge that the described study has several limitations. In our sample, the age distribution possesses a left skew which prevents us from examining the relationship between

age and nasopharyngeal expression of *ACE2* in participants younger than 18 years of age. The difference in mean host gene expression observed between participants who tested negative or positive for COVID-19 cannot be resolved in a cross-sectional study design due to reverse causality. We attempted to control for immunological gene regulation by excluding people with endemic respiratory co-infections from Influenza A/B or Respiratory syncytia virus. We describe associations between transmembrane, soluble *ACE2* and SARS-CoV-2 RNA load which suggest the dual, contrasting role of *ACE2* in viral infection. However, we cannot provide evidence of the mechanism responsible for this association. Additionally, the measure of viral RNA load over-approximates infectious viral titre and viral RNA may be isolated in the absence of infectious virus.[30]

In conclusion, we have characterized nasopharyngeal *ACE2* expression in a sample of people tested for COVID-19. Analysis shows no relationship between age and nasopharyngeal *ACE2* expression in COVID-19-negative participants 19-98 years old. Expression of nasopharyngeal *ACE2* and *TMPRSS2* differs between COVID-19-positive and –negative groups. The role of nasopharyngeal *ACE2* in SARS-CoV-2 infection of the upper-airway may be differentiated by gene isoform expression; transmembrane *ACE2* positively correlates with viral RNA load, while soluble *ACE2* shows a negative association. Nasopharyngeal expression of *ACE2* possesses a dual, contrasting role in SARS-CoV-2 infection at the population level.

Contributors

AMN, DDWT, KSK, NAP, ANJ, MK, DMP, and IS conceived, designed, and obtained funding for the project. AMN wrote the manuscript, editing was provided by Karen Chu. AMN, DDWT and KSK performed the experiments. CDL and HS obtained and cleaned laboratory data. AMN, KSK, DDWT, CAB, CDL, and HS analyzed the data under the direction of CS, NAP, ANJ, MK, DMP and IS. All authors interpreted the data, edited the manuscript and provided their approval to publish.

Declaration of Interests

The authors have no conflicts of interest to declare.

Acknowledgments

We would like to acknowledge the work of all our clinical colleagues at the British Columbia Centre for Disease Control and across the globe in responding to the COVID-19 pandemic. This work would not have been possible without funding from Genome British Columbia. We would like to thank Michael Donoghue for managing this multi-stakeholder project and Karen Chu for assisting as a medical copy editor.

References

1 Zhu N, Zhang D, Wang W, *et al.* A Novel Coronavirus from Patients with Pneumonia in China, 2019. *N Engl J Med* 2020; **382**: 727–33.

2 Hoffmann M, Kleine-Weber H, Schroeder S, *et al.* SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* 2020; **181**: 271-280.e8.

3 Glowacka I, Bertram S, Herzog P, *et al.* Differential Downregulation of ACE2 by the Spike Proteins of Severe Acute Respiratory Syndrome Coronavirus and Human Coronavirus NL63. *J Virol* 2010; **84**: 1198–205.

4 South AM, Diz DI, Chappell MC. COVID-19, ACE2, and the cardiovascular consequences. *Am J Physiol Circ Physiol* 2020; **318**: H1084–90.

5 Sungnak W, Huang N, Bécavin C, *et al.* SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nat Med* 2020; **26**: 681–7.

6 Harmer D, Gilbert M, Borman R, Clark KL. Quantitative mRNA expression profiling of ACE 2, a novel homologue of angiotensin converting enzyme. *FEBS Lett* 2002; **532**: 107–10.

7 Yan R, Zhang Y, Li Y, Xia L, Guo Y, Zhou Q. Structural basis for the recognition of

SARS-CoV-2 by full-length human ACE2. *Science (80-)* 2020; **367**: 1444–8.

8 Zoufaly A, Poglitsch M, Aberle JH, *et al.* Human recombinant soluble ACE2 in severe COVID-19. *Lancet Respir Med* 2020; published online Sept. DOI:10.1016/S2213-2600(20)30418-5.

9 Corman VM, Landt O, Kaiser M, *et al.* Detection of 2019 novel coronavirus (2019-nCoV) by real-time RT-PCR. *Eurosurveillance* 2020; **25**. DOI:10.2807/1560-7917.ES.2020.25.3.2000045.

10 Messmann R, Dietl A, Wagner S, *et al.* Alterations of the renin angiotensin system in human end-stage heart failure before and after mechanical cardiac unloading by LVAD support. *Mol Cell Biochem* 2020; **472**: 79–94.

11 Kitajima T, Okita Y, Kawamura M, *et al.* The relationship between preoperative T helper cytokines in the ileal mucosa and the pathogenesis of pouchitis. *BMC Gastroenterol* 2020; **20**: 277.

12 Shirato K, Kawase M, Matsuyama S. Wild-type human coronaviruses prefer cell-surface TMPRSS2 to endosomal cathepsins for cell entry. *Virology* 2018; **517**: 9–15.

13 Hologic. Instructions for Using the Aptima Swab Specimen Collection Kit for Patient-Collected Specimens. 2020.

14 LeBlanc JJ, Gubbay JB, Li Y, *et al.* Real-time PCR-based SARS-CoV-2 detection in Canadian laboratories. *J Clin Virol* 2020; **128**: 104433.

15 Livak KJ, Schmittgen TD. Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the 2−ΔΔCT Method. *Methods* 2001; **25**: 402–8.

16 Ho DE, Imai K, King G, Stuart EA. MatchIt : Nonparametric Preprocessing for Parametric Causal Inference. *J Stat Softw* 2011; **42**. DOI:10.18637/jss.v042.i08.

17 Gastwirth JL, Gel YR, Miao W. The Impact of Levene's Test of Equality of Variances on Statistical Theory and Practice. *Stat Sci* 2009; **24**: 343–60.

18 Jamshidian M, Jennrich RI, Liu W. A study of partial F tests for multiple linear regression models. *Comput Stat Data Anal* 2007; **51**: 6269–84.

19 Fox J, Monette G. Generalized Collinearity Diagnostics. *J Am Stat Assoc* 1992; **87**: 178.

20 VanderWeele TJ. Principles of confounder selection. *Eur J Epidemiol* 2019; **34**: 211–9.

21 Rstudio Team. RStudio. 2019.

22 Bunyavanich S, Do A, Vicencio A. Nasal Gene Expression of Angiotensin-Converting

Enzyme 2 in Children and Adults. *JAMA* 2020; published online May 20. DOI:10.1001/jama.2020.8707.

23 Ni W, Yang X, Yang D, *et al.* Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. *Crit Care* 2020; **24**: 422.

24 Dong Y, Mo X, Hu Y, *et al.* Epidemiology of COVID-19 Among Children in China. *Pediatrics* 2020; **145**: e20200702.

25 Bar-On YM, Flamholz A, Phillips R, Milo R. SARS-CoV-2 (COVID-19) by the numbers. *Elife* 2020; **9**. DOI:10.7554/eLife.57309.

26 Nakagawa K, Lokugamage KG, Makino S. Viral and Cellular mRNA Translation in Coronavirus-Infected Cells. 2016: 165–92.

27 Obukhov AG, Stevens BR, Prasad R, *et al.* SARS-CoV-2 Infections and ACE2: Clinical Outcomes Linked With Increased Morbidity and Mortality in Individuals With Diabetes. *Diabetes* 2020; **69**: 1875–86.

28 Monteil V, Kwon H, Prado P, *et al.* Inhibition of SARS-CoV-2 Infections in Engineered Human Tissues Using Clinical-Grade Soluble Human ACE2. *Cell* 2020; **181**: 905-913.e7.

29 Cantuti-Castelvetri L, Ojha R, Pedro LD, *et al.* Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. *Science (80-)* 2020; : eabd2985.

30 Bullard J, Dust K, Funk D, *et al.* Predicting Infectious Severe Acute Respiratory Syndrome Coronavirus 2 From Diagnostic Samples. *Clin Infect Dis* 2020; published online May 22. DOI:10.1093/cid/ciaa638.

Supplementary Material

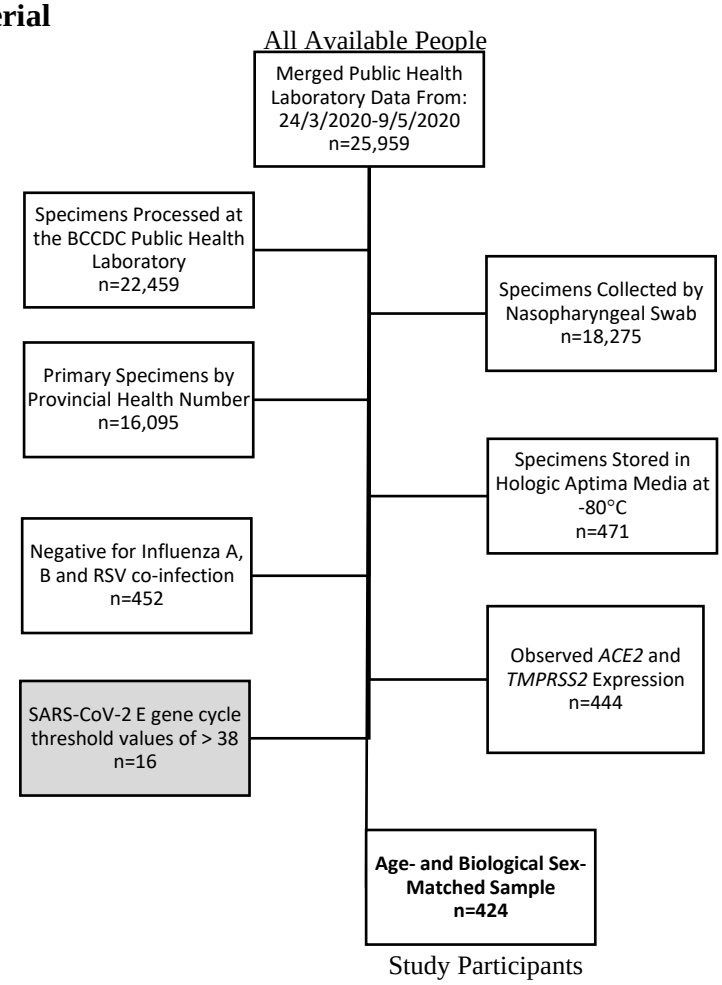

Figure S1: Inclusion, exclusion criteria applied to all people tested for COVID-19 in British Columbia from 24/3/2020-9/5/2020 to select study participants (analytic data). Specimens from people tested for COVID-19 with the following criteria were included (white box): tested at the BCCDC-PHL, specimen was collected by nasopharyngeal swab, test was the primary test by provincial health number, specimen was suspended in Hologic Aptima™ media and stored at -80°C following testing, negative for Influenza A, B or Respiratory syncytia virus infection and possessing host gene expression data. Those with a COVID-19-positive diagnosis, were further restricted by excluding (grey box) SARS-CoV-2 E gene Ct values of > 38 by qRT-PCR (n=212), and matched in a 1:1 ratio with COVID-19-negative people (n=212) by age and biological sex (bold text).

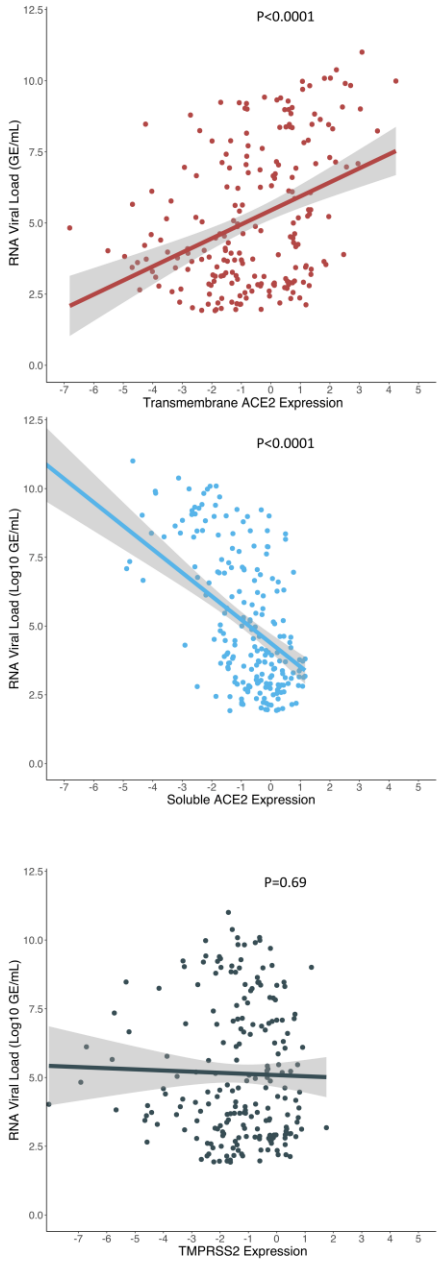

Figure S2: Correlation between host gene expression in COVID-19-positive participants and SARS-CoV-2 RNA load. The level of transmembrane *ACE2* (red), soluble *ACE2* (blue) and *TMPRSS2* (black) expression in positive cases of COVID-19 and its relationship with SARS-CoV-2 RNA is shown in Log_{10} genome equivalents per millilitre (GE/mL). Host gene expression was measured by quantitative real-time PCR and normalized to expression of the housekeeping gene GAPDH by the ($2^{-\Delta\Delta Ct}$) method. Soluble *ACE2* was defined as absolute relative expression between the Hs01085333_m1 (ThermoFisher) and HS01085340_m1 (ThermoFisher) gene targets (Figure S3).

Simple linear regression was used estimate correlations, the reported P values are for the *B*eta-coefficients of host gene expression.

Table S1: Serial Dilution of SARS-CoV-2 RNA for Generation of a Standard Curve by Quantitative Real-Time Polymerase Chain Reaction.

Log_{10} Dilution	**10^0**	**10^1**	**10^2**	**10^3**	**10^4**
~GE/mL[τ]	**1**	**10**	**100**	**1000**	**10,000**
R1	36.1	32.6	28.9	25.9	22.8
R2	35.8	31.9	28.8	25.8	22.8
R3	34.8	32.3	28.8	25.8	22.8
R4	35.6	31.7	28.8	25.7	22.8
R5	34.5	31.9	28.6	25.7	22.3
R6	34.0	32.1	28.7	25.5	22.7
R7	34.4	31.9	28.6	25.6	22.8
R8	34.1	31.4	28.8	25.6	22.8
R9	34.7	31.4	28.7	25.6	22.9

a 5-fold, 1:10 serial dilution of SARS-CoV-2 synthetic RNA was performed over nine independent replicates (n=9) to measure a standard curve by quantitative real-time PCR targeting the viral E gene.

[τ] a simple linear regression model was fit to n=9 replicates of the SARS-CoV-2 RNA standard curve: GE/mL=b_0 + b_1x_1. Where, Genome Equivalents per millilitre (GE/mL) = 34.5 -1.27ln(x) and R^2= 0.999.

Homo sapiens angiotensin I converting enzyme 2 (ACE2), transcript variant 2, mRNA NCBI Reference Sequence: NM_021804.3

```
5' 1 ggcactcata catacactct ggcaatgagg acactgagct cgcttctgaa atttgacaag
  61 ataaccacta aaatctcttt gaattctatg ttgttgtgat cccatggcta cagaggatca
 121 ggagttgaca tagatactct ttggatttca taccatgtgg aggctttctt acttccacgt
 181 gaccttgact gagttttgaa tagcgcccaa cccaagttca aaggctgata agagagaaaa
 241 tctcatgagg aggttttagt ctagggaaag tcattcagtg gatgtgatct tggctcacag
 301 gggacgatgt caagctcttc ctggctcctt ctcagccttg ttgctgtaac tgctgctcag
 361 tccaccattg aggaacaggc caagacattt ttggacaagt ttaaccacga agccgaagac
 421 ctgttctatc aaagttcact tgcttcttgg aattataaca ccaatattac tgaagagaat
 481 gtccaaaaca tgaataatgc tggggacaaa tggtctgcct ttttaaagga acagtccaca
 541 cttgcccaaa tgtatccact acaagaaatt cagaatctca cagtcaagct tcagctgcag
 601 gctcttcagc aaaatgggtc ttcagtgctc tcagaagaca agagcaaacg gttgaacaca
 661 attctaaata caatgagcac catctacagt actggaaaag tttgtaaccc agataatcca
 721 caagaatgct tattacttga accaggtttg aatgaaataa tggcaaacag tttagactac
 781 aatgagaggc tctgggcttg ggaaagctgg agatctgagg tcggcaagca gctgaggcca
 841 ttatatgaag agtatgtggt cttgaaaaat gagatggcaa gagcaaatca ttatgaggac
 901 tatggggatt attggagagg agactatgaa gtaaatgggg tagatggcta tgactacagc
 961 cgcggccagt tgattgaaga tgtggaacat acctttgaag agattaaacc attatatgaa
1021 catcttcatg cctatgtgag ggcaaagttg atgaatgcct atccttccta tatcagtcca
1081 attggatgcc tccctgctca tttgcttggt gatatgtggg gtagattttg gacaaatctg
1141 tactctttga cagttccctt tggacagaaa ccaaacatag atgttactga tgcaatggtg
1201 gaccaggcct gggatgcaca gagaatattc aaggaggccg agaagttctt tgtatctgtt
1261 ggtcttccta atatgactca aggattctgg gaaaattcca tgctaacgga cccaggaaat
1321 gttcagaaag cagtctgcca tcccacagct tgggacctgg ggaagggcga cttcaggatc
1381 cttatgtgca caaaggtgac aatggacgac ttcctgacag ctcatcatga gatggggcat
1441 atccagtatg atatggcata tgctgcacaa ccttttctgc taagaaatgg agctaatgaa
1501 ggattccatg aagctgttgg ggaaatcatg tcactttctg cagccacacc taagcattta
1561 aaatccattg gtcttctgtc acccgatttt caagaagaca atgaaacaga aataaacttc
1621 ctgctcaaac aagcactcac gattgttggg actctgccat ttacttacat gttagagaag
1681 tggaggtgga tggtctttaa aggggaaatt cccaaagacc agtggatgaa aaagtggtgg
1741 gagatgaagc gagagatagt tggggtggtg gaacctgtgc cccatgatga aacatactgt
1801 gaccccgcat ctctgttcca tgtttctaat gattactcat tcattcgata ttacacaagg
1861 acccttttacc aattccagtt tcaagaagca ctttgtcaag cagctaaaca tgaaggccct
1921 ctgcacaaat gtgacatctc aaactctaca gaagctggac agaaactgtt caatatgctg
1981 aggcttggaa aatcagaacc ctggacccta gcattggaaa atgttgtagg agcaaagaac
2041 atgaatgtaa ggccactgct caactacttt gagcccttat ttacctggct gaaagaccag
2101 aacaagaatt cttttgtggg atggagtacc gactggagtc catatgcaga ccaaagcatc
2161 aaagtgagga taagcctaaa atcagctctt ggagataaag catatgaatg gaacgacaat
2221 gaaatgtacc tgttccgatc atctgttgca tatgctatga ggcagtactt tttaaaagta
2281 aaaaatcaga tgattctttt tggggaggag gatgtgcgag tggctaattt gaaaccaaga
2341 atctccttta atttctttgt cactgcacct aaaaatgtgt ctgatatcat tcctagaact
2401 gaagttgaaa aggccatcag gatgtcccgg agccgtatca atgatgcttt ccgtctgaat
2461 gacaacagcc tagagtttct ggggatacag ccaacacttg gacctcctaa ccagcccccct
2521 gtttccatat ggctgattgt ttttggagtt gtgatgggag tgatagtggt tggcattgtc
2581 atcctgatct tcactgggat cagagatcgg aagaagaaaa ataaagcaag aagtggagaa
2641 aatccttatg cctccatcga tattagcaaa ggagaaaata atccaggatt ccaaaacact
2701 gatgatgttc agacctcctt ttagaaaaat ctatgttttt cctcttgagg tgattttgtt
2761 gtatgtaaat gttaatttca tggtatagaa aatataagat gataaagata tcattaaatg
2821 tcaaaactat gactctgttc agaaaaaaaa ttgtccaaag acaacatggc caaggagaga
2881 gcatcttcat tgacattgct ttcagtattt atttctgtct ctggatttga cttctgttct
2941 gtttcttaat aaggattttg tattagagta tattagggaa agtgtgtatt tggtctcaca
3001 ggctgttcag ggataatcta aatgtaaatg tctgttgaat ttctgaagtt gaaaacaagg
3061 atatatcatt ggagcaagtg ttggatcttg tatggaatat ggatggatca cttgtaagga
3121 cagtgcctgg gaactggtgt agctgcaagg attgagaatg gcatgcatta gctcactttc
3181 atttaatcca ttgtcaagga tgacatgctt tcttcacagt aactcagttc aagtactatg
3241 gtgatttgcc tacagtgatg tttggaatcg atcatgcttt cttcaaggtg acaggtctaa
3301 agagagaaga atccagggaa caggtagagg acattgcttt ttcacttcca aggtgcttga
3361 tcaacatctc cctgacaaca caaaactaga gccaggggcc tccgtgaact cccagagcat
3421 gcctgataga aactcatttc tactgttctc taactgtgga gtgaatggaa attccaactg
3481 tatgttcacc ctctgaagtg ggtacccagt ctcttaaatc ttttgtattt gctcacagtg
3541 tttgagcagt gctgagcaca aagcagacac tcaataaatg ctagatttac acactc 3'
```

Legend:
Transmembrane Domain
SARS-CoV Binding Domain
ACE-2 HS01085333_m1 3'
https://www.thermofisher.com/order/genome-database/?pearUXVerSuffix=pearUX2&elcanoForm=true#!/ge/assays/ge_all/?keyword=HS01085333_m1
ACE-2 HS01085340_m1 5'
https://www.thermofisher.com/order/genome-database/?pearUXVerSuffix=pearUX2&elcanoForm=true#!/ge/assays/ge_all/?keyword=HS01085340_m1

Figure S3: Nasopharyngeal expression of *ACE2* was determined by two quantitative real-time polymerase chain reaction targets, one target (yellow) exon spans the *ACE2* transmembrane domain. The other does not include the transmembrane domain (pink) and lays nested between two motifs known to interact with the SARS-CoV receptor binding domain (blue). The measure of soluble *ACE2* represents normalized expression of the pink to yellow target, or the proportion of non-transmembrane *ACE2* to transmembrane *ACE2*.