EBioMedicine 66 (2021) 103316

Contents lists available at ScienceDirect

EBioMedicine

journal homepage: www.elsevier.com/locate/ebiom

Research paper

The contrasting role of nasopharyngeal angiotensin converting enzyme 2 (*ACE2*) transcription in SARS-CoV-2 infection: A cross-sectional study of people tested for COVID-19 in British Columbia, Canada

Aidan M. Nikiforuk[a,d,#], Kevin S. Kuchinski[a,b,#], David D.W. Twa[b,c,#], Christine D. Lukac[a], Hind Sbihi[a,d], C.Andrew Basham[a,d], Christian Steidl[c], Natalie A. Prystajecky[a,b], Agatha N. Jassem[a,b], Mel Krajden[a,b], David M. Patrick[a,d], Inna Sekirov[a,b,*]

[a] *British Columbia Centre for Disease Control, Vancouver, BC V5Z 4R4, Canada*
[b] *Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC V6T 1Z4, Canada*
[c] *British Columbia Cancer Research Centre, Vancouver, BC V5Z 1L3, Canada*
[d] *School of Population and Public Health, University of British Columbia, Vancouver, BC V6T 1Z4, Canada*

ARTICLE INFO

Article History:
Received 9 December 2020
Revised 6 March 2021
Accepted 15 March 2021
Available online xxx

Keywords:
ACE2 transcription
Cross-sectional study
Viral RNA load
Nasopharynx
Soluble ACE2
COVID-19
SARS-CoV-2

ABSTRACT

Background: Angiotensin converting enzyme 2 (ACE2) protein 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* transcription in people tested for COVID-19 and the relationship between *ACE2* transcription and SARS-CoV-2 viral load, while adjusting for expression of: (i) the complementary protease, Transmembrane serine protease 2 (*TMPRSS2*), (ii) soluble *ACE2*, (iii) age, and (iv) biological sex. The *ACE2* gene was targeted to measure expression of transmembrane and soluble transcripts.

Methods: A cross-sectional study of n = 424 "participants" aged 1−104 years referred for COVID-19 testing was performed in British Columbia, Canada. Patients who tested positive for COVID-19 were matched by age and biological sex to patients who tested negative. Viral load and host gene expression were assessed 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.

Findings: Analysis showed no association between age and nasopharyngeal *ACE2* transcription in those who tested negative for COVID-19 (P = 0•092). Mean relative transcription of transmembrane (P = 0•00012) and soluble (P<0•0001) *ACE2* isoforms, as well as *TMPRSS2* (P<0•0001) was higher in COVID-19-negative participants than COVID–19 positive ones, yielding a negative correlation between targeted host gene expression and positive COVID-19 diagnosis. In bivariate analysis of COVID-19-positive participants, transcription of transmembrane *ACE2* positively correlated with SARS-CoV-2 viral RNA load (B = 0•49, R^2=0•14, P<0•0001), transcription of soluble *ACE2* negatively correlated (B= -0•85, R^2= 0•26, P<0•0001), and no correlation was found with *TMPRSS2* transcription (B= -0•042, R^2=<0•10, P = 0•69). Multivariable analysis showed that the greatest viral RNA loads were observed in participants with high transmembrane *ACE2* transcription (B= 0•89, 95%CI: [0•59 to 1•18]), while transcription of the soluble isoform appears to protect against high viral RNA load in the upper respiratory tract (B= -0•099, 95%CI: [-0•18 to -0•022]).

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

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

1. Introduction

In December 2019, clusters of viral pneumonia were reported in Wuhan, China. A novel highly pathogenic human coronavirus was

* Corresponding author.
E-mail address: inna.sekirov@bccdc.ca (I. Sekirov).
# These authors contributed equally to the work.

https://doi.org/10.1016/j.ebiom.2021.103316

Research in Context

Evidence before this study

We conducted a MEDLINE ® search using the MeSH topic terms: "angiotensin converting enzyme 2 expression", "SARS" and "age", restricting the search to English-language reports published from January 1st, 2020. The search returned 98 articles, 88 of which reported primary research; these were further filtered by the MeSH qualifiers "epidemiology" and "virology" to provide 43 articles. Search results were further restricted to the MeSH terms: age, virus replication, host pathogen interactions, nasal mucosa, viral RNA load and real-time polymerase chain reaction, which returned 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.

Added value of this study

We measured *ACE2* transcription in the context of COVID-19 testing to investigate the role of nasopharyngeal ACE2 protein in SARS-CoV-2 infection. Our findings support previous work: the strong correlation we observe between nasopharyngeal *ACE2* transcription 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 protein expression. No observed relationship between age and nasopharyngeal *ACE2* in COVID-19-negative participants suggests that upper airway *ACE2* expression may be independent of the renin-angiotensin-aldosterone-system. We are the first to measure *ACE2* transcription in nasopharyngeal specimens from a large sample of COVID-19 patients representing a broad range of ages and clinical manifestations, adjusting for age, sex, and transcription of *TMPRSS2* in a multivariable analysis. We demonstrate that nasopharyngeal transcription of soluble *ACE2* negatively correlates with viral load, suggesting a potential protective role at the population level.

Implications of all the available evidence

Considering all available evidence, ACE2 protein may play a dual, contrasting role in SARS-CoV-2 infection of the upper airway. Transmembrane *ACE2* transcript positively correlates with SARS-CoV-2 viral load, while soluble *ACE2* transcript shows a negative association. Total nasopharyngeal *ACE2* transcript does not correlate with age, as would be expected in the lower respiratory tract. Genetic and environmental factors such as *ACE2* single nucleotide polymorphisms or smoking, and whether they regulate the expression of transmembrane and soluble ACE2 protein in healthy nasopharyngeal tissue, require further investigation.

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]. TMPRSS2 protein 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 or furin, may perform its role in viral infection interchangeably [4,5]. In human physiology, ACE2 has a cardiovascular protective and anti-inflammatory role, as a constituent of the renin-angiotensin-aldosterone-system (RAAS) [6]. 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 [4]. Transcription of *ACE2* produces at least two dominant mRNA transcript variants responsible for translation into soluble and membrane-bound protein isoforms [7–10]. The soluble variant (cAug10) shares the same start codon as the transmembrane variant (aAug10), but terminates at nucleotide position 1665 or amino acid position 555 [9]. Although the transmembrane isoform has been shown to be crucial for viral entry into host cells, the role of soluble ACE2 remains uncharacterized, though evidence exists that it may protect against SARS-CoV-2 infection [11,12]. Production of soluble ACE2 also occurs post-transcription via shedding of ACE2 from the cellular surface. The metallopeptidase domain 17 (ADAM17) cleaves membrane anchored ACE2 downstream of the cAug10 transcript termination site between amino acid positions 716 and 741 [13]. Therefore, there are at least two independent mechanisms of soluble ACE2 protein production, alternative spicing and ectodomain shedding. The relationship between alternative splicing and ectodomain shedding to produce soluble ACE2 protein is not well understood in human physiology or in the context of COVID-19 infection; the two mechanisms could work mutually or competitively to regulate expression of soluble ACE2.

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 transcription of *ACE2* and age in COVID-19-negative participants, (ii) nasopharyngeal transcription of host genes by COVID-19 test result, and (iii) nasopharyngeal transcription of transmembrane *ACE2* and viral RNA load in those who tested COVID-19-positive adjusting for age, biological sex, transcription of soluble *ACE2* and *TMPRSS2*. Our study builds on previous work characterizing *ACE2* transcription in older, hospitalized COVID-19 patients, drawing from a larger sample across broader ages and clinical manifestations with adjustment for important covariates [14]. These findings increase our knowledge of host genes involved in SARS-CoV-2 infection and may allow for assessment of baseline COVID-19 risk.

2. Methods

2.1. 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, the individual's first test administered by personal health number, collected by nasopharyngeal swab, suspended in Hologic Aptima™ media, negative for concurrent Influenza A, B or Respiratory syncytial virus infection, stored at −80 °C following RNA extraction, and for whom host gene transcription 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). 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 [15–18].

2.2. Ethics statement

Ethical approval for the study was obtained from the University of British Columbia human ethics board (H20-01110). Written informed consent was not required as per ethics board approval. Participant data was de-identified prior to analysis; the results of this non-interventional observational study were not linked back to any identifying patient records. The study was deemed to be of minimal risk.

2.3. 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) [19].

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 20ul, with 5ul of RNA template, 1ul of 20x primer/probe, 5ul Fast Virus and 9ul of nuclease free water per reaction. Cycling conditions were set to: 50 °C for 5 min, 95 °C for 20 s followed by 40 cycles of 95 °C for 15 s and 60 °C for 1 min. A multiplex qRT-PCR reaction targeting SARS-CoV-2 envelope (*E*) and host ribonuclease P (*RNaseP*) was used to diagnose acute viral infection by presence of viral RNA [20]. Expression of *RNaseP* was used to measure sampling variation, ensuring sample quality [20]. 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 milliliter by making a 5-fold, 1:10 standard curve of SARS-CoV-2 synthetic RNA- MN908947•3 (Twist Bioscience) (Table S1) [21]. Commercially available primer probe sets were used to amplify the host gene targets *TMPRSS2* [18] (Hs00237175_m1) (ThermoFisher) and *ACE2* [16] (Hs01085333_m1, HS01085340_m1) by multiplex qRT-PCR with a glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*) control [17] (Hs02758991_g1) (ThermoFisher). The full length *ACE2* reference sequence (NC_000023.11) was downloaded from the National Centre for Biotechnology Information GenBank and annotated to show the (Hs01085333_m1, HS01085340_m1) qRT-PCR targets using Geneious Prime 2021.3 (Figure S3) [47]. Relative gene transcription was calculated between *TMPRSS2*, *ACE2* and *GAPDH* using the $2^{-\Delta\Delta Ct}$ method [22]. Transmembrane *ACE2* was defined from the Hs01085333_m1 (ThermoFisher) gene target, which exon spans the transmembrane domain. Soluble *ACE2* was defined as the absolute difference in transcript expression between the HS01085340_m1 (ThermoFisher) and Hs01085333_m1 (ThermoFisher) gene targets (HS01085340_m1-Hs01085333_m1) (Figure S3). Subtraction of the Hs01085333_m1 target from HS01085340_m1 is mathematically equivalent to the use of Hs01085333_m1 instead of *GAPDH* in the $2^{-\Delta\Delta Ct}$ equation to calculate the relative expression of HS01085340_m1. The HS01085340_m1 target does not include the transmembrane domain (Figure S3).

2.4. Statistical analysis

Sample size estimation was performed in the design stage of the study. Recruitment of n = 424 participants, or n = 212 participants per COVID-19 test result, provides evaluation of a Cohen minimum effect size (h = 0•20), assuming: alpha=0•05 and power=0•80 for comparison between groups. For comparison within groups, n = 212 provided 0•85 power assuming a minimum effect size (h = 0•20), standard deviation (σ =1), and alpha of 0•05 [23].

De-identified data reporting participant COVID-19 test result, site of sample collection, type of collection media, accessioning laboratory, age and biological sex were accessed from public health laboratory records. Matching of COVID-19-positive people, who met the described inclusion and exclusion criteria, to COVID-19-negative participants by age and biological sex in a 1:1 ratio was performed by a nearest neighbor algorithm, caliper = 0 SD (n = 424) [24]. Bivariate analyses were performed between: age, biological sex, viral RNA load, *TMPRSS2* transcription, soluble or transmembrane *ACE2* transcription, and COVID-19 test result. The balance of covariates between test groups was examined post-matching by the standardized mean difference (SMD) [25]. Parametric statistical tests were used given the large sample size of the study [26].

The relationship between age and nasopharyngeal transcription of transmembrane *ACE2*, soluble *ACE2*, and *TMPRSS2* was examined in unmatched COVID-19-negative participants over the age of 18 and less than 98 years by both linear regression and categorization of age into 10-year intervals. Differences in *ACE2* and *TMPRSS2* transcription by age category were tested by analysis of variance. The bivariate relationship between transmembrane *ACE2* expression and biological sex was analyzed by using a t-test in unmatched COVID-19-negative participants.

Differences in mean transcription of host genes by COVID-19 diagnosis was further examined in the matched sample by a two-tailed, paired t-test assuming non-equal variance [27]. Correlation between host gene transcription 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 [28]. Collinearity was examined by the variable inflation factor with a cut-off of 10 [29]. The common cause criterion was applied to control for measured confounding [30]. Effect modification was reported when found statistically significant and supported conceptually [8]. All analysis was performed in R version 3.6.3 using the packages: car, ggsci, tidyverse, dataexplorer, ggpubr, lmtest, publish, forcats, matchit, tableone, emmeans and effects [31].

2.5. 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 research was performed at the University of British Columbia and British Columbia center for Disease Control.

3. Results

The analytic dataset contains age- and biological sex-matched participants tested for COVID-19 in British Columbia from 24/3/2020–9/5/2020 (n = 424; Figure S1). Participant characteristics are shown in Table 1. The mean age of COVID-19-positive participants was 61•69 years, 47•60% were biologically male. In COVID-19-negative participants, the mean age was 62•18 years, 48•60% were biologically male. Age (SMD=0•020) and biological sex (SMD=0•019) were balanced between COVID-19-negative and -positive groups. Viral RNA (*E* gene Ct) was only detected in COVID-19-positive participants, with an average Ct value of 28•24. Mean relative transcription of transmembrane *ACE2*, soluble *ACE2*, and *TMPRSS2* statistically differed by COVID-19 test result; on average, transcription of each gene was lower in samples from COVID-19-positive participants (Table 1).

3.1. Relationship between ACE2 expression, age, and sex in unmatched COVID-19-negative participants

To assess ACE2 expression in uninfected nasopharyngeal tissue, we compared *ACE2* transcription and age in unmatched COVID-19-

Table 1
Characteristics of analytic data stratified by COVID-19 test result (n = 424).

Variable Name	Level	Total (n)	COVID-19 Test Result Negative	Positive	SMD [a]	P-Value [b]
		424	212	212		
E Gene Ct (mean [SD])		424	.. [c]	28•24 [7•18]		
Viral RNA Load (mean [SD]) [d]		424	..	5•14 [2•46]		
Age (mean [SD])		424	62.18 [24.29]	61.69 [23.49]	0.020	0.83
Biological Sex (n [%])					0.019	0.92
	Male	204	103 [48.60]	101 [47.60]		
	Female	220	109 [51.40]	111 [52.40]		
Transmembrane *ACE2* (mean [SD])		424	(0.00) [1.08]	(−0.61) [1.90]		0.00012
Soluble *ACE2* (mean [SD])		424	(0.00) [0.62]	(−0.89) [1.47]		<0.0001
TMPRSS2 Total (mean [SD])		424	(0.00) [0.79]	(−1.35) [1.60]		<0.0001

Participants who tested positive for COVID-19 by qRT-PCR were matched with COVID-19-negative participants by age and biological sex, balance was checked between groups post matching by calculating the standardized mean difference (SMD) (cutoff=0.20). Relative transcription of *ACE2* and *TMPRSS2* was normalized to the negative group by the $2^{-\Delta\Delta Ct}$ method.

[a] the standardized mean difference (SMD) was calculated to determine balance of age and biological sex between participants who tested negative or positive for COVID-19 post-matching (cutoff=0.20).

[b] p-values are reported for parametric tests used for continuous (*t*-test) and categorical variables (x^2 test).

[c] no SARS-CoV-2 E Gene Ct values were reported from qRT-PCR for COVID-19-negative participants.

[d] measured as Log_{10} genome equivalents/ml (Log10 GE/ml), viral RNA load data is not available from patients who tested negative.

negative participants between 19 and 98 years of age. No relationship was found between nasopharyngeal transmembrane *ACE2* transcription and age with age as a continuous variable in linear regression (P = 0•076). This finding was reproduced when age was categorized into ten-year intervals (ANOVA, P = 0•092) (Fig. 1). Similarly, no association was observed between age and soluble *ACE2* (ANOVA, P = 0•081) or *TMPRSS2* (ANOVA, P = 0•173) transcription. We also assessed the relationship between sex and transmembrane *ACE2* transcription but found no relationship in unmatched COVID-19-negative participants (*t*-test, P = 0•832).

3.2. *Nasopharyngeal transcription of* ACE2 *and* TMPRSS2 *by COVID-19 test result*

Bivariate analysis showed a significant mean difference in transmembrane *ACE2*, soluble *ACE2* and *TMPRSS2* transcription between matched COVID-19-negative and -positive participants (Fig. 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, paired *t*-tests showed lower mean transcription of transmembrane *ACE2* (P = 0•00,012, Fig. 2a), soluble *ACE2* (P<0•0001,

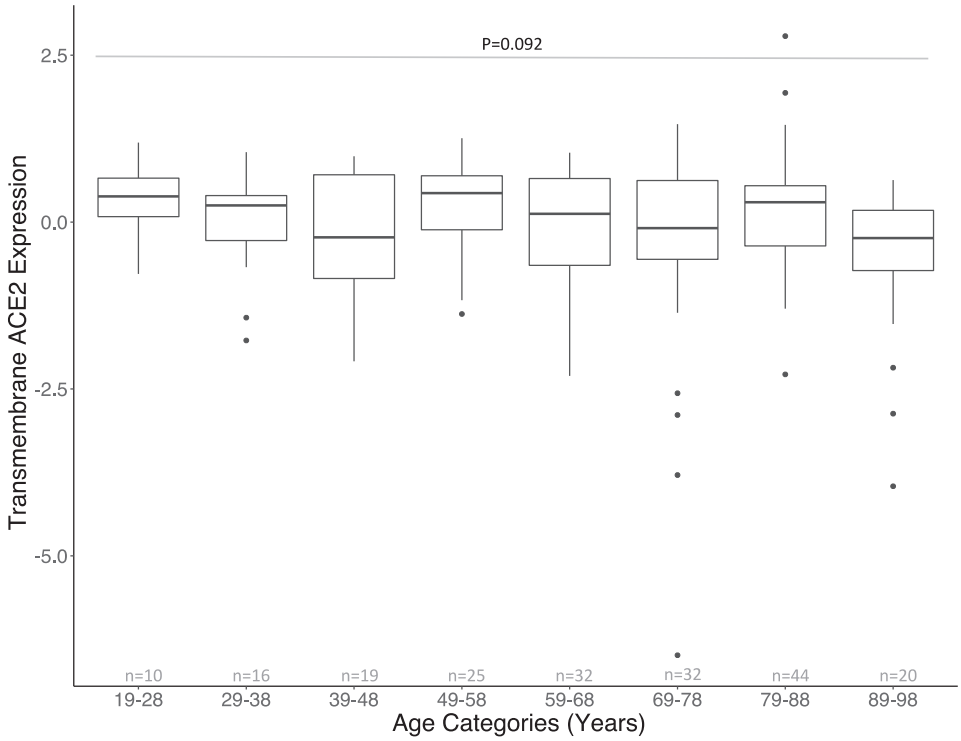

Fig. 1. Relationship between age and nasopharyngeal transmembrane *ACE2* transcription in unmatched COVID-19-negative participants. Boxplots of transmembrane *ACE2* transcription by 10-year age categories in unmatched COVID-19-negative participants between the ages of 19 and 98 (n = 198); boxes represent the Q1-Q3 interquartile range, whiskers represent 1.5x the Q1 or Q3 and horizontal lines the median transmembrane *ACE2* transcription by age category. 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* transcription among age categories (ANOVA, P = 0•092).

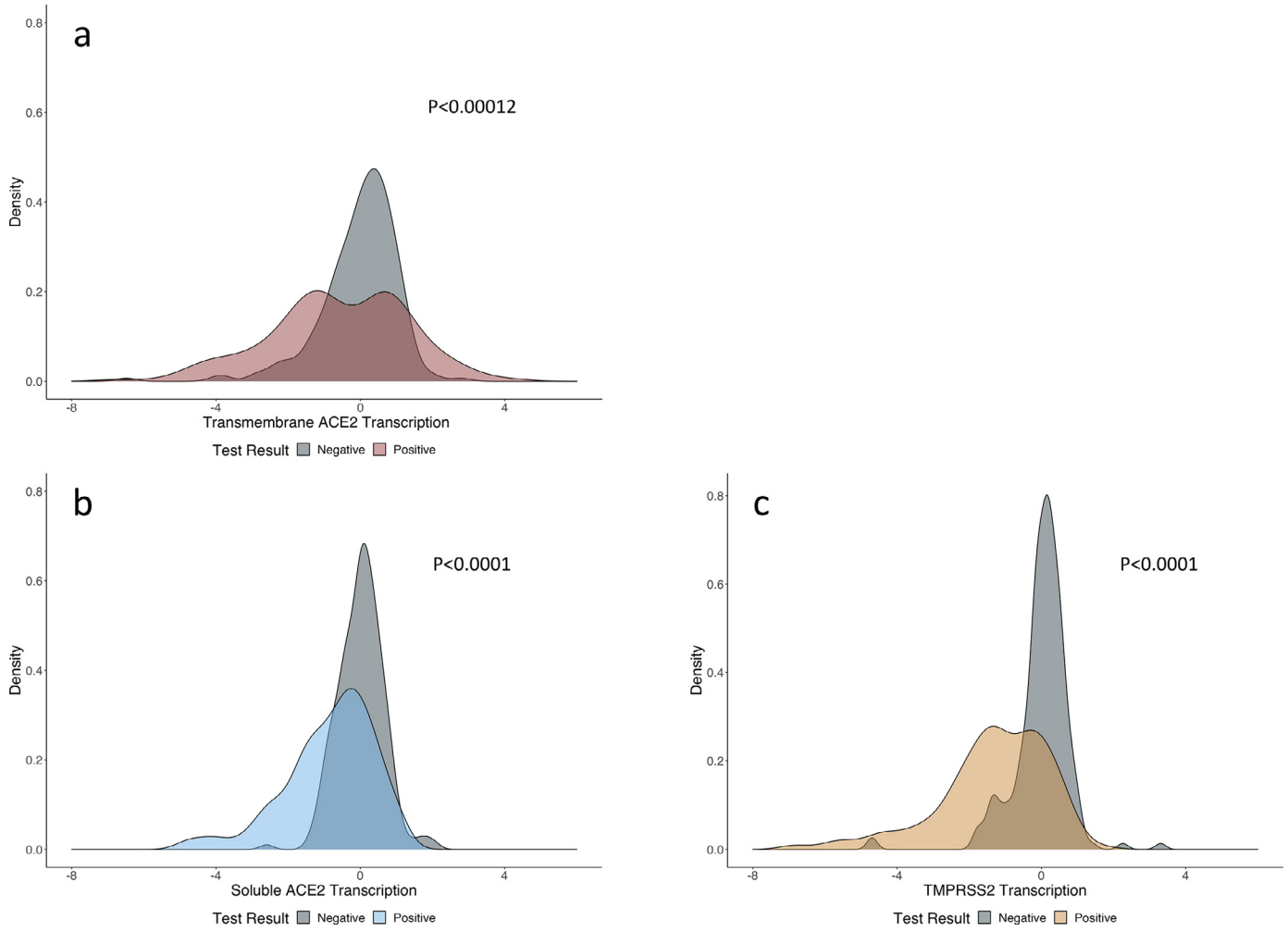

Fig. 2. Relative nasopharyngeal transcription of targeted host genes by COVID-19 test result. Gene transcription is portrayed in kernel density plots stratified by COVID-19 test result. Probability densities of relative host gene transcription are shown by positive (red, transmembrane *ACE2*; blue, soluble *ACE2*; yellow, *TMPRSS2*) and negative COVID-19 test results (gray). Levene's test was used to detect non-equal variance in gene transcription for all host targets between COVID-19-negative and -positive participants. A two-tailed, paired *t*-test was used to examine mean difference in host gene transcription by COVID-19 test result assuming unequal variance: (a) transmembrane *ACE2* (P = 0•00,012), (b) soluble *ACE2* (P<0•0001) and (c) *TMPRSS2* (P<0•0001).

Fig. 2b) and *TMPRSS2* (P<0•0001, Fig. 2c) in participants who tested COVID-19-positive in comparison to -negative (Fig. 2).

3.3. Host gene transcription and viral RNA load in COVID-19-positive participants

The relationship between transmembrane *ACE2* transcription and viral load was assessed in nasopharyngeal specimens from COVID-19-positive participants. We estimated viral load by qRT-PCR quantification of viral genome copies, a commonly-used surrogate, which has been associated with the presence of culturable SARS-CoV-2 from nasopharyngeal specimens [32–36]. Nasopharyngeal transcription of transmembrane *ACE2* positively correlated with viral RNA (linear regression, B = 0•49, R^2=0•14, P<0•0001). Transcription of soluble *ACE2* negatively correlated with viral RNA (linear regression, B= −0•85, R^2= 0•26, P<0•0001). No correlation was found between TMPRSS2 transcription and viral RNA (linear regression, B= −0•042, R^2=<0•10, P = 0•69) (Figure S2) by bivariate analysis. Multiple linear regression estimated that a one-unit change in transmembrane *ACE2* transcription increases viral RNA load by 0•89 Log_{10} GE/ml (95%CI: 0•59 to 1•18) adjusting for age, biological sex, transcription of *TMPRSS2*, and soluble *ACE2* (Table 2). Biological sex could have been dropped from the model as evident from comparison between nested models (partial F-test, P = 0•76), but was kept for validity. A partial F-test indicated effect modification between transmembrane and soluble *ACE2* transcription (P = 0•010) (Table 2). No effect modification was observed between transcription of transmembrane *ACE2* and *TMPRSS2* (partial F-test, P = 0•23). Transcription of soluble *ACE2* decreases viral load (B= −0•099, 95%CI: [−0•18 to −0•022]) (Table 2[g]). The association between transmembrane *ACE2* transcription and viral RNA in nasopharyngeal tissue differs by the concomitant level of soluble *ACE2* transcribed (Fig. 3). Effect modification was visualized by categorizing soluble *ACE2* by the relative mean transcription +/- 1 | 2 standard deviations of all study participants (low= −1•66 | −2•86, mean= −0•44 | −0•44, high=0•77 | 1•98, SD=1•21) (Table 2, Fig. 3 & Figure S4).

4. Discussion

In the described study, we measured nasopharyngeal *ACE2* and *TMPRRS2* gene transcription 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* transcription and age in unmatched COVID-19-negative participants, differences in host gene transcription between matched COVID-19-negative and -positive participants, and the role of soluble *ACE2* transcription in SARS-CoV-2 infection in those who tested positive for COVID-19.

Table 2
Unadjusted, adjusted and effect modification inclusive linear regression models of SARS-CoV-2 viral RNA load.

Model	Effect Estimate (B)	95% CI for B	P-Value[f]
Transmembrane *ACE2* Transcription [a]	0.49	(0.33 to 0.66)	<0.0001
Age and Sex Adjusted [b]	0.36	(0.19 to 0.53)	<0.0001
Covariate Adjusted [c]	0.89	(0.59 to 1.18)	<0.0001
Effect Modification [d,e]	–	–	0.010[g]
Modification was Evaluated at +/- 1SD and 2SD	1SD | 2SD	1SD | 2SD	–
Low Soluble *ACE2* Transcription	0.79 | 0.91	(0.49 to 1.08) | (0.62 to 1.19)	–
Mean Soluble *ACE2* Transcription	0.67 | 0.67	(0.34 to 1.00) | (0.34 to 1.00)	–
High Soluble *ACE2* Transcription	0.55 | 0.43	(0.16 to 0.94) | (−0.028 to 0.88)	–

Regression of SARS-CoV-2 viral RNA load by transmembrane *ACE2*, soluble *ACE2*, *TMPRSS2* transcription, age, and biological sex. Columns report beta coefficients (B), 95% confidence intervals (95% CI) and P-values for transmembrane *ACE2*.

[a] the unadjusted effect of transmembrane *ACE2* transcription on SARS-CoV-2 viral load (Log_{10} GE/ml).

[b] the effect of transmembrane *ACE2* transcription adjusted by age and biological sex on SARS-CoV-2 viral load (Log_{10} GE/ml).

[c] the effect of transmembrane *ACE2* transcription adjusted by: age, biological sex, soluble *ACE2*, and *TMPRSS2* expression on SARS-CoV-2 viral load (Log_{10} GE/ml).

[d] the effect of transmembrane *ACE2* transcription adjusted by: age, biological sex, soluble *ACE2*, and *TMPRSS2* transcription; including effect modification by soluble *ACE2* transcription or how the effect of transmembrane *ACE2* transcription changes by the level of soluble *ACE2* transcription.

[e] soluble *ACE2* transcription was categorized as: low, mean and high and reported for +/- 1 or 2 SD or greater, simple slopes were calculated across transmembrane *ACE2* transcription from: −7 to 5 by 0.5.

[f] P-values are reported from the results of a partial F-test to evaluate nested models.

[g] soluble *ACE2* effect modification was also assessed as a continuous value using multiple linear regression (B= −0.099, 95%CI: [−0.18 to −0.022], $P = 0.011$).

Among participants who tested negative for COVID-19, we found no relationship between nasopharyngeal *ACE2* transcription and age in a group of adults between 19 and 98 years of age. This result agrees with findings from other efforts to characterize the relationship between age and nasopharyngeal *ACE2* expression at the population level. Bunyavanich et al. reported a difference in nasal *ACE2* expression levels between children younger than ten years old and young adults from eighteen to twenty-four years old [37]. Children younger than ten years old were found to have on average lower nasal *ACE2* expression [37]. No difference in nasal *ACE2* expression was evident between young adults, and adults twenty-five or more years in age which agrees with the findings of our study [37]. Lack of a relationship between nasopharyngeal *ACE2* transcription and progressive aging suggests that expression of ACE2 protein in the upper airway may not correlate with expression of ACE2 protein in the RAAS [38]. The physiological importance of ACE2 protein in the RAAS suggests that its expression would increase in response to age as older people have higher prevalence of cardiovascular disease or similar

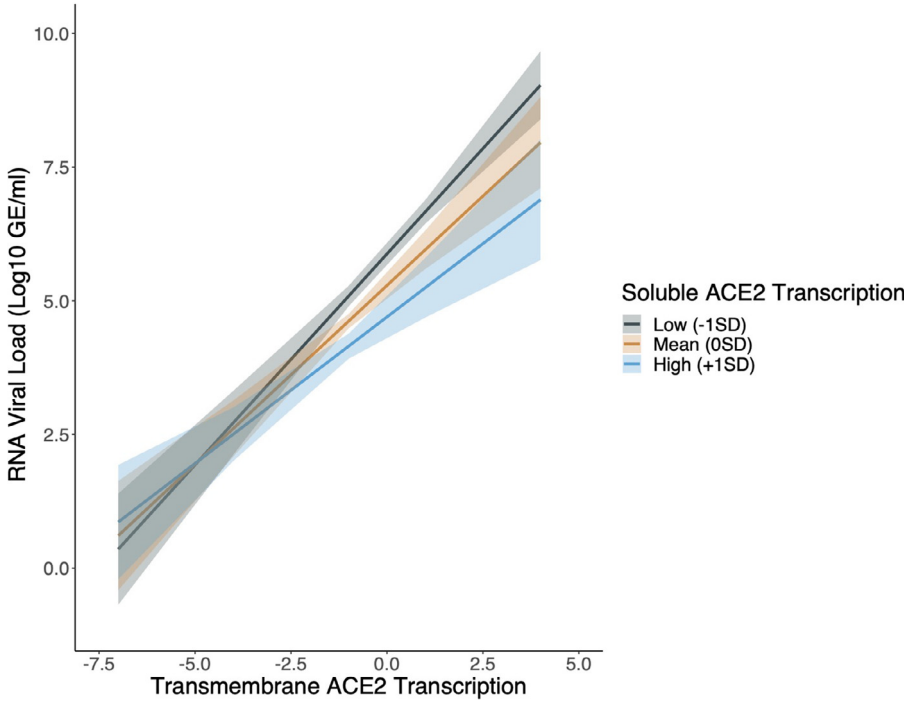

Fig. 3. The association between transmembrane *ACE2* transcription and SARS-CoV-2 RNA load in nasopharyngeal tissue differs by the amount of soluble *ACE2* transcription. Soluble *ACE2* transcription was categorized into low, mean and high levels to demonstrate the relationship. The low category (gray) represents soluble *ACE2* transcription one standard deviation or greater below the mean transcription ($n = 74$). The mean category (orange) codes for the mean soluble *ACE2* transcription at zero standard deviations. The high category (blue) indicates soluble *ACE2* transcription one standard deviation or greater above mean transcription ($n = 80$). Shaded areas represent 95% confidence intervals, solid lines represent B-coefficients for simple slopes derived from multiple linear regression. Figure S4 visualizes the corresponding effect of soluble *ACE2* transcription +/- 2 SD or greater from the mean; estimates are reported in Table 2.

comorbidities such as type 2 diabetes [39]. If nasopharyngeal expression of ACE2 protein occurs independently of the RAAS pathway, future studies should ascertain what factors, if any, regulate expression of ACE2 protein 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 protein expression, which protect against SARS-CoV-2 viral infection [37,40].

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. The matched analysis demonstrated that mean transcription of transmembrane *ACE2*, 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 transcription or host gene transcription 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 [41]. Numerous studies have demonstrated that coronavirus replication disrupts cellular transcription, as resources required by the cell to produce mRNA are instead sequestered by viral genome replication [42]. 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, type 2 diabetes, and severe COVID-19 [43]. Answering these questions, 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* transcription and SARS-CoV-2 RNA load. Bivariate and multivariable analysis both suggest that *ACE2* plays a dual role in SARS-CoV-2 infection. Transmembrane *ACE2* positively correlates with viral RNA load, while soluble *ACE2* may limit viral infection. Interestingly, effect modification was observed between transmembrane and soluble *ACE2* transcription, implying that the proportion between these molecules may have more importance than absolute expression of either transcript 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. These findings have interesting potential for clinical application, considering RT-qPCR quantification of *ACE2* transcripts can be easily integrated into existing COVID-19 molecular diagnostics. Routine clinical measurement of *ACE2* may provide a useful predictive marker of peak viral load in patients, disease severity, or SARS-CoV-2 transmissibility after the onset of symptoms.

Though we are not equipped to characterize the mechanism by which transcription of soluble ACE2 protein plays a protective role in SARS-CoV-2 infection, identification of this effect at the population level warrants further investigation of the underlying molecular process. Previous work has postulated that soluble ACE2 protein restricts SARS-CoV/SARS-CoV-2 infection by acting as a decoy substrate [44]. In a study of virus-ACE2 dynamics in engineered human tissue, virus particles bound to soluble ACE2 protein instead of binding to transmembrane ACE2 protein and, consequently were unable to infect susceptible cells [44]. Moreover, recombinant soluble ACE2 protein was recently administered to a COVID-19 patient requiring ventilation [11]. In the case report, administration of soluble ACE2 protein decreased viral RNA load in the patient's plasma, tracheal aspirate and nasopharyngeal specimens. Additionally, soluble ACE2 protein did not interfere with production of immunoglobulin G and assumed its cardiovascular protective role in the RAAS [11].

Soluble ACE2 protein may also influence the viral use of Neuropilin-1 as a co-receptor or alternative receptor for SARS-CoV-2. Given that the role of Neuropilin-1 in cell entry and infectivity may depend upon high SARS-CoV-2 viral load [45], it is conceivable that accessory expression of soluble ACE2 protein may limit SARS-CoV-2 tropism.

5. Strengths and limitations

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 19 years of age. The difference in mean host gene expression observed between participants who tested negative or positive for COVID-19 may be due to reverse causality, which cannot be resolved in a cross-sectional study design. Unmeasured confounding from comorbidities or patient drug prescriptions may bias the estimated effect of transmembrane *ACE2* transcription on viral RNA load. Peters et al. reported increased expression of transmembrane *ACE2* in asthma patients with type 2 diabetes and decreased expression in those prescribed inhaled corticosteroids [38]. We describe associations between transmembrane, soluble *ACE2*, and SARS-CoV-2 RNA load, which suggest the dual role of *ACE2* transcripts in viral infection. However, we cannot provide evidence of the mechanism responsible for this association.

We also acknowledge the limitations of using quantitative polymerase chain reaction (qPCR) -based methods as a surrogate for gene expression and viral load. The measure of viral RNA load may overestimate infectious viral titre and viral RNA may be isolated in the absence of infectious virus; however, laboratory studies have demonstrated the presence of culturable virus in nasopharyngeal specimens across a broad range of Ct values [32–36]. Similarly, we acknowledge that the abundance of host gene transcripts does not reflect post-transcriptional regulation of expression. However, experimental studies have found a strong concordance between *ACE2* transcript expression and protein concentration in upper airway tissues [46]. Notwithstanding these limitations, qPCR-based methods are both sensitive and, practical for high-throughput use in clinical laboratories with pandemic-imposed resource constraints, while easily integrated into existing COVID-19 molecular diagnostics.

Future studies are required to validate our findings at the molecular level. Importantly, the correlation between alternatively spliced *ACE2* transcript and soluble ACE2 protein should be investigated in nasopharyngeal tissue of COVID-19 -negative and -positive specimens.

We have characterized nasopharyngeal *ACE2* transcription in a cross-sectional sample of people tested for COVID-19 in British Columbia, Canada. Analysis shows no relationship between age and nasopharyngeal *ACE2* transcription in COVID-19-negative participants 19–98 years old. Transcription of nasopharyngeal *ACE2* and *TMPRSS2* differs between COVID-19-positive and -negative groups, with lower mean values and greater variation observed in COVID-19-positive participants. The role of nasopharyngeal ACE2 in SARS-CoV-2 infection of the upper-airway may be differentiated by gene isoform expression; transmembrane *ACE2* isoform transcription positively correlates with viral RNA load, while soluble *ACE2* isoform transcription shows a negative association. The contrasting role of transmembrane and soluble *ACE2* transcription is positioned to be of importance for prevention and care of COVID-19. Individuals with high endogenous soluble *ACE2* expression may be less likely to transmit the virus to others or develop a lower respiratory infection as they clinically present with lower viral load. Similarly, we would expect treatment with recombinant soluble ACE2 protein to decrease

viral load and, thereby, the severity and transmissibility of COVID-19 disease.

Contributions

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 Dr. Karen Chu. AMN, DDWT and KSK performed the experiments. CDL, HS and AMN 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. AMN, DDWT, KSK and IS verified the raw data. All authors interpreted the data, read the manuscript and provided their approval to publish.

Declaration of Competing 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 Center 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 (COV-55). We would like to thank Michael Donoghue for managing this multi-stakeholder project and Dr. Karen Chu for assisting in medical copyediting.

Data Sharing Statement

The de-identified participant data that support the findings of this study are available from the British Columbia Ministry of Health. Restrictions apply to the availability of these data, which were used under agreement at the British Columbia Center for Disease Control for this study.

Supplementary materials

Supplementary material associated with this article can be found in the online version at doi:10.1016/j.ebiom.2021.103316.

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;181:271–80 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] 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.
[5] Bestle D, Heindl MR, Limburg H, et al. TMPRSS2 and furin are both essential for proteolytic activation of SARS-CoV-2 in human airway cells. Life Sci Alliance 2020;3:e202000786.
[6] South AM, Diz DI, Chappell MC. COVID-19, ACE2, and the cardiovascular consequences. Am J Physiol Circ Physiol 2020;318:H1084–90.
[7] 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.
[8] ur Rehman S, M Tabish. Alternative splicing of ACE2 possibly generates variants that may limit the entry of SARS-CoV-2: a potential therapeutic approach using SSOs. Clin Sci 2020;134:1143–50.
[9] National Center for Biotechnology Information. Homo sapiens complex locus ACE2, encoding angiotensin I converting enzyme (peptidyl-dipeptidase A) 2. 2020. https://www.ncbi.nlm.nih.gov/IEB/Research/Acembly/av.cgi?db=human&l=ACE2.
[10] Clark HF. The secreted protein discovery initiative (SPDI), a large-scale effort to identify novel human secreted and transmembrane proteins: a bioinformatics assessment. Genome Res 2003;13:2265–70.
[11] 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.
[12] 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 2020;367(80−):1444–8.
[13] Jia HP, Look DC, Tan P, et al. Ectodomain shedding of angiotensin converting enzyme 2 in human airway epithelia. Am J Physiol Cell Mol Physiol 2009;297:L84–96.
[14] Amati F, Vancheri C, Latini A, et al. Expression profiles of the SARS-CoV-2 host invasion genes in nasopharyngeal and oropharyngeal swabs of COVID-19 patients. Heliyon 2020;6:e05143.
[15] 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.
[16] 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.
[17] 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.
[18] 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.
[19] Hologic. Instructions for using the Aptima Swab specimen collection kit for patient-collected specimens. 2020.
[20] 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.
[21] Martínez-Martínez M, Diez-Valcarce M, Hernández M, Rodríguez-Lázaro D. Design and application of nucleic acid standards for quantitative detection of enteric viruses by real-time PCR. Food Environ Virol 2011;3:92–8.
[22] 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.
[23] Cohen J. A power primer. Psychol Bull 1992;112:155–9.
[24] 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.
[25] Faraone S V. Interpreting estimates of treatment effects: implications for managed care. P T 2008;33:700–11.
[26] Kwak SG, Kim JH. Central limit theorem: the cornerstone of modern statistics. Korean J Anesthesiol 2017;70:144.
[27] 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.
[28] 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.
[29] Fox J, Monette G. Generalized collinearity diagnostics. J Am Stat Assoc 1992;87:178.
[30] VanderWeele TJ. Principles of confounder selection. Eur J Epidemiol 2019;34:211–9.
[31] R Core Team (2020). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
[32] Case JB, Bailey AL, Kim AS, Chen RE, Diamond MS. Growth, detection, quantification, and inactivation of SARS-CoV-2. Virology 2020;548:39–48.
[33] 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.
[34] La Scola B, Le Bideau M, Andreani J, et al. Viral RNA load as determined by cell culture as a management tool for discharge of SARS-CoV-2 patients from infectious disease wards. Eur J Clin Microbiol Infect Dis 2020;39:1059–61.
[35] Ladhani SN, Chow JY, Janarthanan R, et al. Investigation of SARS-CoV-2 outbreaks in six care homes in London. EClinicalMedicine 2020 April 2020;26:100533.
[36] Cevik M, Tate M, Lloyd O, Maraolo AE, Schafers J, Ho A. SARS-CoV-2, SARS-CoV, and MERS-CoV viral load dynamics, duration of viral shedding, and infectiousness: a systematic review and meta-analysis. The Lancet Microbe 2021;2:e13–22.
[37] 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.
[38] Peters MC, Sajuthi S, Deford P, et al. COVID-19−related genes in sputum cells in asthma. Relationship to demographic features and corticosteroids. Am J Respir Crit Care Med 2020;202:83–90.
[39] Ni W, Yang X, Yang D, et al. Role of angiotensin-converting enzyme 2 (ACE2) in COVID-19. Crit Care 2020;24:422.
[40] Dong Y, Mo X, Hu Y, et al. Epidemiology of COVID-19 among children in China. Pediatrics 2020;145:e20200702.
[41] 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.
[42] Nakagawa K., Lokugamage K.G., Makino S. Viral and cellular mRNA translation in coronavirus-infected cells. 2016: 165−92.
[43] 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.

[44] 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;181:905–13 e7.
[45] Cantuti-Castelvetri L, Ojha R, Pedro LD, et al. Neuropilin-1 facilitates SARS-CoV-2 cell entry and infectivity. Science 2020(80−):eabd2985.
[46] Ortiz ME, Thurman A, Pezzulo AA, et al. Heterogeneous expression of the SARS-Coronavirus-2 receptor ACE2 in the human respiratory tract. EBioMedicine 2020;60:102976.
[47] Geneious Prime 2021.3. 2020. https://www.geneious.com/.