



OPEN

## SARS-CoV-2 does not have a strong effect on the nasopharyngeal microbial composition

Tzipi Braun<sup>1</sup>, Shiraz Halevi<sup>2,3</sup>, Rotem Hadar<sup>1</sup>, Gilate Efroni<sup>1</sup>, Efrat Glick Saar<sup>1</sup>, Natan Keller<sup>2,3</sup>, Amnon Amir<sup>1</sup>, Sharon Amit<sup>2,3</sup> & Yael Haberman<sup>1,4</sup>✉

The coronavirus disease 2019 (COVID-19) has rapidly spread around the world, impacting the lives of many individuals. Growing evidence suggests that the nasopharyngeal and respiratory tract microbiome are influenced by various health and disease conditions, including the presence and the severity of different viral disease. To evaluate the potential interactions between Severe Acute Respiratory Syndrome Corona 2 (SARS-CoV-2) and the nasopharyngeal microbiome. Microbial composition of nasopharyngeal swab samples submitted to the clinical microbiology lab for suspected SARS-CoV-2 infections was assessed using 16S amplicon sequencing. The study included a total of 55 nasopharyngeal samples from 33 subjects, with longitudinal sampling available for 12 out of the 33 subjects. 21 of the 33 subjects had at least one positive COVID-19 PCR results as determined by the clinical microbiology lab. Inter-personal variation was the strongest factor explaining > 75% of the microbial variation, irrespective of the SARS-CoV-2 status. No significant effect of SARS-CoV-2 on the nasopharyngeal microbial community was observed using multiple analysis methods. These results indicate that unlike some other viruses, for which an effect on the microbial composition was noted, SARS-CoV-2 does not have a strong effect on the nasopharynx microbial habitants.

Severe Acute Respiratory Syndrome Corona 2 (SARS-CoV-2)<sup>1,2</sup> is the coronavirus responsible for the 2019–2020 pandemic. It infects the upper respiratory tract nasopharynx, and is transmitted mostly by tiny droplets and aerosols<sup>2</sup>, with some evidence suggesting also a fecal mode of transmission<sup>3</sup>. Due to low cross-protective immunity from related viral infections, SARS-CoV-2 transmissibility is high, hence facilitating widespread person-to-person transmission<sup>4,5</sup>. In fact, this SARS-CoV-2 has the highest transmissibility when compared to previous pandemics<sup>6,7</sup>. However, its clinical course is highly variable between subjects, and factors contributing to this heterogeneity are not yet entirely known.

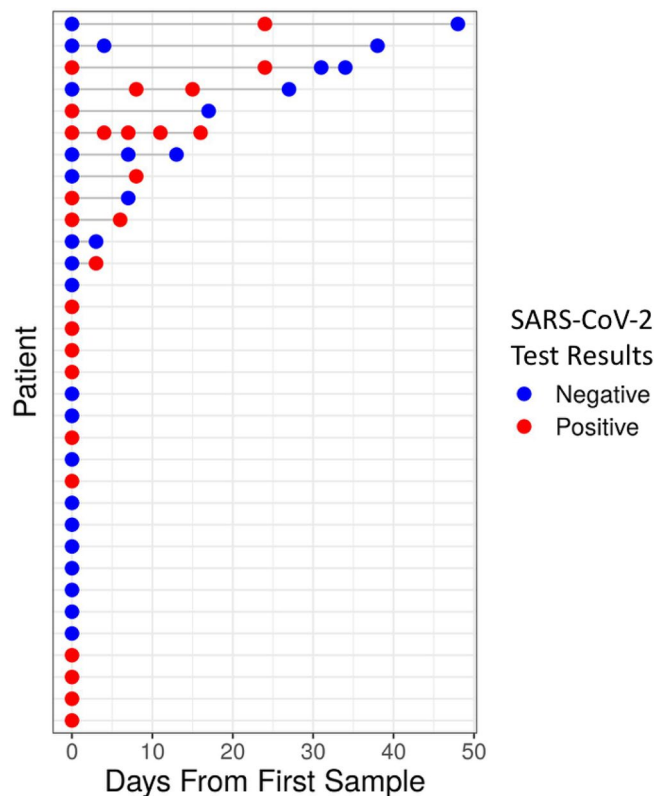
Previous studies have shown that different environmental factors, such as smoking, affect the upper respiratory microbial composition<sup>8</sup>. Additionally, studies have shown a correlation between respiratory tract microbial composition and the severity of various viral diseases, including respiratory syncytial virus in children<sup>9</sup> and influenza<sup>10</sup>. Others have observed reduced alpha diversity in the nasopharynx of patients during viral diseases<sup>11</sup>.

Two recent studies have attempted to identify differences between the nasopharyngeal microbiota of SARS-CoV-2 patients compared to healthy controls, with mixed results. One cohort, used 16S amplicon sequencing and included 40 patients, reported no significant differences between SARS-CoV-2 positive and negative subjects<sup>4</sup>. The second study used the direct Oxford Nanopore long-read third generation sequencing and included 50 patients suspected for COVID-19, and found some differences between those subjects that were positive and negative for SARS-CoV-2<sup>12</sup>. Here, we used 16S amplicon sequencing to define differences also by using longitudinal samples, which supplement the other two previous studies that used cross-sectional sampling.

### Results

**Cohort characteristics.** 55 Nasopharyngeal samples from 33 subjects, confirmed or suspect for COVID-19, were processed by 16S rRNA amplicon sequencing in parallel to SARS-CoV-2 testing by RT-PCR in the clinical microbiology lab. Twelve of the 33 patients had longitudinal samples with 2–5 samples each (Fig. 1). Overall,

<sup>1</sup>Pediatric Gastroenterology, Hepatology and Nutrition Unit, The Edmond and Lily Safra Children's Hospital, Sheba Medical Center, Affiliated with the Tel-Aviv University, Tel Hashomer, Israel. <sup>2</sup>Clinical Microbiology, The Chaim Sheba Medical Centre, Ramat-Gan, Israel. <sup>3</sup>Sakler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel. <sup>4</sup>Cincinnati Children's Hospital Medical Center and the University of Cincinnati College of Medicine, Cincinnati, OH, USA. ✉email: Yael.Haberman@sheba.health.gov.il



**Figure 1.** Longitudinal cohort of 55 samples from 33 subjects confirmed or suspected as being SARS-CoV-2 positive. Each row corresponds to an enrolled subject, over the number of days since first sample included in this study. Colors represent SARS-CoV-2 test results for a specific sample, with red and blue indicating SARS-CoV-2 positive and negative samples respectively.

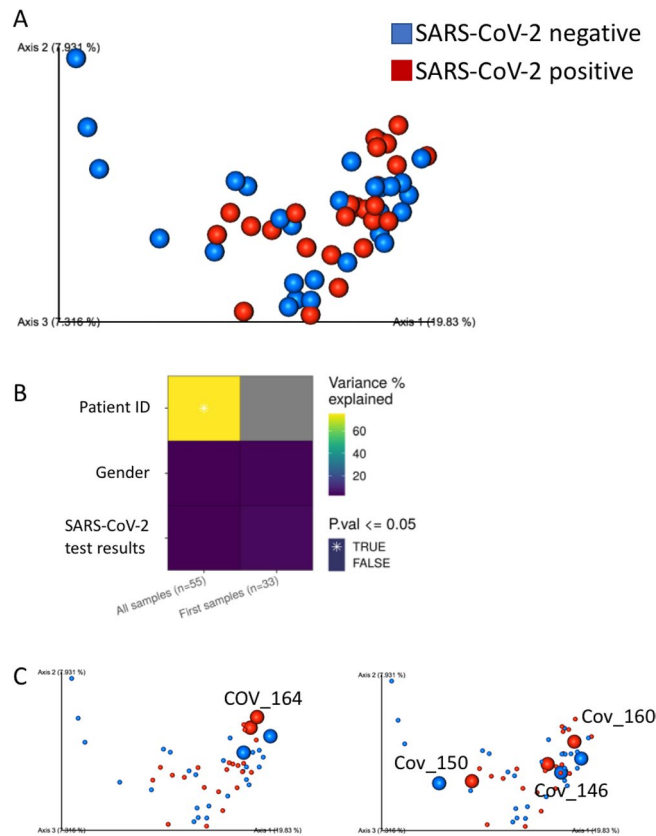
	Subjects (n = 33) with total 55 samples	
Age median (IQR)	n = 32	52 (30, 68)
Gender (male) %	n = 32	18 (56%)
SARS-CoV-2 positive $\geq 1$ sample	n = 33	21 (64%)

**Table 1.** Cohort characteristics.

we had 29 SARS-CoV-2 negative samples and 26 SARS-CoV-2 positive samples, with 21 of the 33 patients having at least one positive SARS-CoV-2 sample. Patients had a median age of 52 years and 56% were male (Table 1).

**Factors effecting nasopharyngeal microbial composition.** Unweighted unfrac based PCoA was used to visually explore samples similarity and variations. As can be seen, SARS-CoV-2 testing result does not seem to have a strong effect, as samples do not cluster by COVID-19 test results (Fig. 2A). The contribution of the various factors to the microbial composition was quantified using a PERMANOVA test (see “Methods”), either using all of the samples, or by using only single sample per subject to avoid personal bias. Inter-personal variation, indicated by the patient ID, was highly important, explaining 76% of microbial variation ( $p$  value = 0.001), while SARS-CoV-2 test results, and gender did not have a significant effect, with a  $p$  value  $> 0.2$  (Fig. 2B) using either all samples or only one sample per patient. The strong effect of patient identity on the microbial composition can be observed in the PCoA, where samples from the same patient clustered close to each other, regardless of COVID-19 test results (Fig. 2C).

**Inability to detect a strong contribution of COVID-19 results on microbial composition.** Since the microbial compositional analysis showed no difference between COVID-19 positive and negative samples, other methods were used to try and detect such differences. To avoid personal bias, we used only one sample per patient. A heatmap visualization showed no overt difference between the two groups (Fig. 3A). There was also no significant difference between the groups in any of the alpha diversity measures used including Faith’s

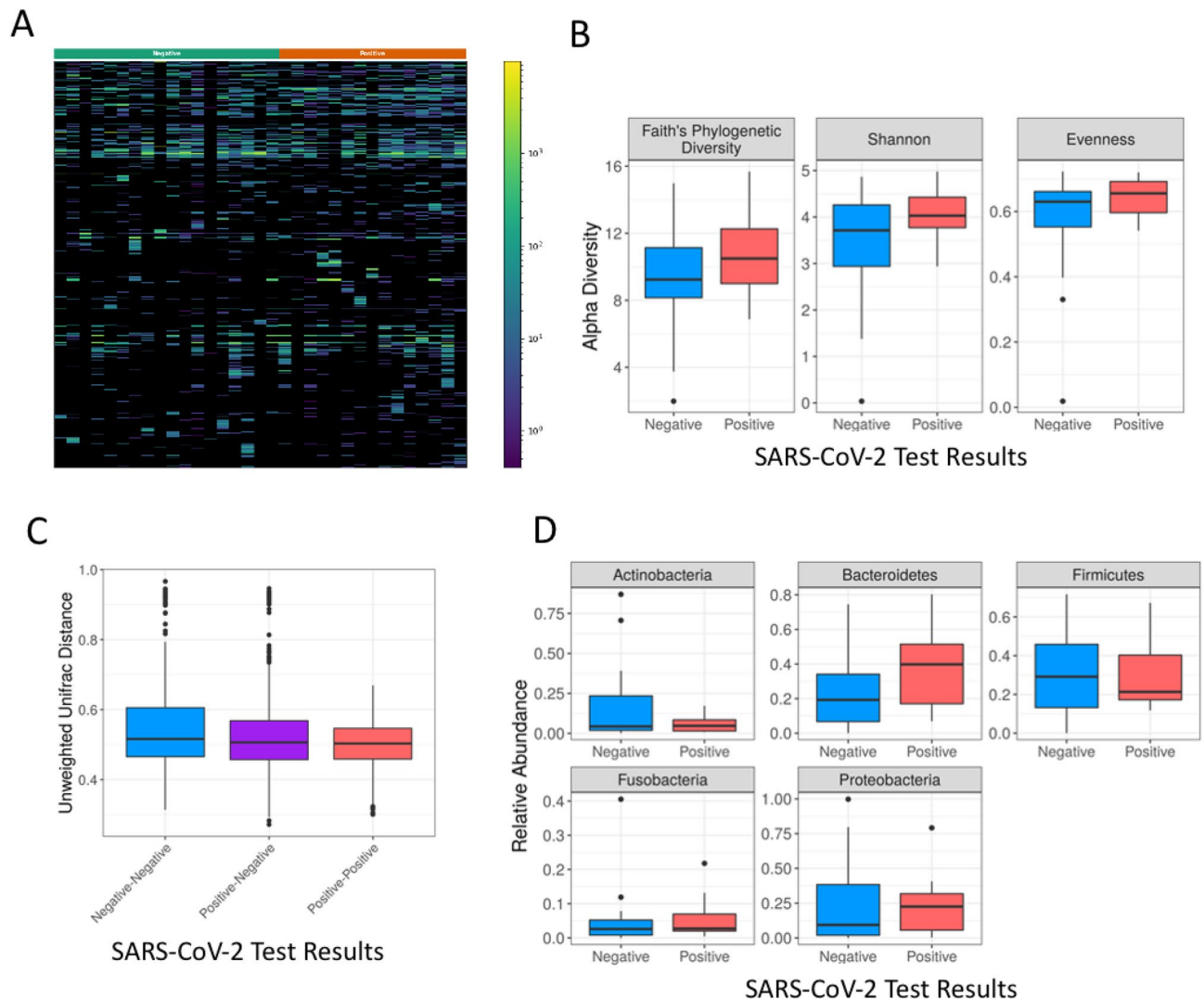


**Figure 2.** Personal variation has the strongest effect on nasopharyngeal microbial composition. (A) Unweighted UniFrac PCoA plot of all samples, colored by SARS-CoV-2 test results. (B) PERMANOVA analysis of microbial variance explained by subject (patient ID), gender and SARS-CoV-2 test result, using all longitudinal samples, as well as the first samples from each subject. \* indicates statistical significance with  $p \leq 0.05$ , and n is shown in brackets. (C) Unweighted UniFrac PCoA plot of all samples, colored by SARS-CoV-2 test results. SARS-CoV-2 positive and negative samples from specific subjects with longitude sampling are marked with increased size and labeled with the subject ID, showing clustering is driven by subject ID rather than SARS-CoV-2 test results.

phylogenetic diversity, Shannon and evenness (Fig. 3B), with Wilcoxon rank sum test  $p$  values  $> 0.1$  for all alpha diversity measures. An unweighted unifracs distance analysis also did not detect significant difference within or between positive and negative samples (Fig. 3C), with a PERMANOVA  $p$  value of 0.21. Similarly, we did not detect a significant difference in the relative abundance of the five most abundant phyla (Fig. 3D), (Wilcoxon rank sum test,  $p$  value  $> 0.06$ ). We were also not able to detect any significant differences between specific amplicon sequence variants (ASVs) with an FDR of 0.25 using two different platforms—calour permutation rank mean test with dsFDR multiple hypothesis correction<sup>13</sup>, and maaslin2<sup>14</sup> using all samples and controlling for patient ID as a random effect.

## Discussion

We were unable to detect a significant microbial pattern in the nasopharynx of SARS-CoV-2 positive subjects in comparison to SARS-CoV-2 negative subjects. This may imply that unlike what was seen with Rhinovirus infection, where a significantly higher diversity was observed in non-infected individuals compared to infected individuals<sup>11</sup>, COVID-19 does not have a strong effect on the nasopharyngeal microbial composition. In contrast, we have noticed a strong personal microbial signature, with samples from the same individual clustering together irrespective of whether samples were positive or negative for SARS-CoV-2, further implying that SARS-CoV-2 does not have strong effect on the nasopharynx microbiome. Two recent studies analyzed COVID-19 positive and negative nasopharyngeal samples with mixed results. The first used cross-sectional 16S amplicon sequencing<sup>4</sup> and found no significant difference between infected and uninfected patient, similar to the results reported here. The second study also used a cross-sectional design and the direct Oxford Nanopore long-read third generation sequencing<sup>12</sup>. This study identified reduced microbial diversity and some differences in microbial communities. This significant difference was observed at the species level but not at the genus or family level. Therefore, the inability to detect differences in our current study may stem from the lower phylogenetic resolution of 16S rRNA amplicon sequencing compared to Oxford Nanopore derived long-reads, differences in populations, or sample sizes. It is also possible that the severity of certain viral diseases increases upon antibiotics administration, which



**Figure 3.** SARS-CoV-2 test result does not seem to have a strong effect on the nasopharyngeal microbial composition. **(A)** Heatmap showing all observed ASVs using the first sample from each subject. Each row represents a different ASV and each column a different sample. SARS-CoV-2 test results are noted in the color bar above (green for negative and orange for positive). **(B)** Boxplots of alpha diversity values, using the first sample from each subject, by SARS-CoV-2 test result. Three different alpha diversity measures were used—Faith's phylogenetic diversity, Shannon, and evenness, as indicated. There were no significant differences in any of the measures tested ( $p$  value  $> 0.1$ ). **(C)** Boxplot of unweighted unifrac distances between positive-positive, negative-negative and positive-negative sample pairs, using only the first sample per subject. No significant difference was detected, (PERMANOVA  $p$  value of 0.21). **(D)** Boxplots of the relative abundance of the five most abundant phyla by SARS-CoV-2 test result, using only the first sample per subject (Wilcoxon rank sum test  $p$  value  $> 0.05$ ).

reduce the microbial diversity, and potentially facilitate more rapid viral replication. To test for this possibility, larger studies that will take into account antibiotics administration during COVID-19 are required.

The longitudinal subset in our study included samples from different stages of the SARS-CoV-2 infection including pre- and post- SARS-CoV-2 positive results, corresponding to early and late infection stages, as well as negative tests obtained following SARS-CoV-2 infection. We did not observe any prominent effects of the infection during any of these stages. In contrast, we observed a strong personal effect, as was previously described in other studies of nasopharyngeal microbiome<sup>4</sup>. This result supports the validity of this cohort, showing that the lack of SARS-CoV-2 infection effect on the nasopharyngeal microbiome was not due to technical problems.

Our work has several strengths as it includes both cross sectional and longitudinal samples, and was based on samples obtained during routine screening for COVID19. Limitations include the limited cohort size, the use of 16S rRNA amplicon sequencing, and the limited clinical data. While there is a possibility that a cohort of patients with more severe cases will show a stronger effect, in such a cohort, it would be hard to assess how much of the difference is attributed to the presence of the virus itself or is due to the COVID-19 disease severity, treatment-related conditions, and medications. It is also possible that emerging new variants<sup>15</sup> may alter the nasopharyngeal microbial composition, and more research is required to address such concern. Further research

is needed to validate if and to what extent, COVID-19 infection influences the nasopharyngeal microbial composition, and disease outcome.

## Methods

**Study design and sample processing.** Ethical approval for the study was granted by the Sheba Local Research Ethics Committee and all methods were performed in accordance with the relevant guidelines and regulations. Since this study used nasopharyngeal specimens already submitted to the microbiology core as part of clinical workup and without identifiable patient information other than age, gender, and viral results, an exemption from patient consent was granted from the Sheba Local Research Ethics Committee. The primary goal of this study was to characterize the microbial composition and diversity between individuals with positive and negative SARS-CoV-2 results, and to characterize the microbial dynamics within individuals for which we have longitudinal sampling.

We randomly included samples with no specific selection or exclusion criteria other than, when possible, obtaining several samples per subject as implicated (Fig. 1 and Table 1). Sterile swabs were used to collect the nasopharyngeal samples. Nasopharyngeal samples specimens were analyzed for SARS-CoV-2 presence at the Clinical Microbiology Lab at the Sheba Medical Center in Israel, between April and May 2020. Nasopharyngeal samples were collected into UTM (Copan) using sterile swabs and lysed with Universal LB lysis buffer (Seegene). Nucleic acids were extracted using StarMag universal cartridge kit (Seegene) on a Microlab Starlet (Hamilton) extraction robot, and RT-PCR of E, RdRp and N genes of the SARS CoV-2 virus was done using Allplex 2019 nCoV assay (Seegene) according to manufacturer's protocol. Nucleic acid extracts from samples were used in parallel also for broad-range high throughput 16S rRNA amplicon sequencing (16S-seq). Negative controls including swab blanks (sterile swabs), extraction blanks (reagents), and PCR controls were also included in the sampling and analyses. 55 unique specimens for 33 individuals passed processing, quality control, and filtering, and were included in this study.

**DNA extraction, PCR amplification, and sequencing.** DNA extraction and PCR amplification of the variable region 4 (V4) of the 16S rRNA gene using Illumina adapted universal primers 515F/806R39 was conducted using the direct PCR protocol [Extract-N-Amp Plant PCR kit (Sigma-Aldrich, Inc.)] as previously described<sup>16</sup>. Briefly, PCRs were conducted in 96 wells plate [denaturation for 3 min at 94 °C; 35 cycles (98 °C, 60 s; 55 °C, 60 s; 72 °C, 60 s) followed by elongation for 10 min at 72 °C]. Positive amplicons were pooled in equimolar concentrations into a composite sample that was size selected (300–500 bp) using agarose gel to reduce non-specific products from host DNA. Sequencing was performed on the Illumina MiSeq platform with the addition of 20% PhiX, and generating paired-end reads of 175b in length in each direction.

**Microbiome data processing and analysis.** Reads were processed in a data curation pipeline implemented in QIIME 2 version 2019.4<sup>17</sup>. Reads were demultiplexed according to sample specific barcodes. Quality control was performed by truncating reads after three consecutive Phred scores lower than 20. Reads with ambiguous base calls or shorter than 150 bp after quality truncation were discarded. Amplicon Sequence variant (ASV) detection was performed using Deblur<sup>18</sup>. Unweighted UniFrac was used as a measure of beta-diversity = between sample diversity<sup>19</sup>, using a phylogenetic tree generated by SEPP<sup>20</sup>. Faith's phylogenetic diversity, Shannon diversity and evenness were calculated in QIIME 2 as measures of alpha diversity. All samples were rarefied to 2000 reads for alpha and beta diversity analysis, to avoid sample size affect. The resulting distance matrix was used to perform a principal coordinate analysis (PCoA). heatmaps were generated using Calour version 2018.10.1 with default parameters<sup>21</sup>. Differentially expressed ASVs between positive and negative results were detected using one sample per subject, with a non-parametric rank mean test as implemented in Calour<sup>21</sup> with dsFDR multiple hypothesis correction<sup>13</sup> (FDR < 0.25). As a second approach to test for ASVs significantly associated with COVID-19 positive test, we used MaAsLin2 (Multivariate Association with Linear Models) R package version 1.0.0<sup>14</sup> with an FDR of 0.25. This approach used all samples (including multiple samples per subject when available), controlling for age, gender, and patient ID as random effects.

PERMANOVA: Quantifications of variance were calculated using PERMANOVA with the adonis function in the R package Vegan<sup>22</sup>, using 999 iterations, on the rarefied Unweighted UniFrac distance values. The total variance explained by each variable was calculated independently of other variables (that is, as the sole variable in the model).

**Data availability.** The study datasets were deposited at the National Center for Biotechnology Information as BioProject PRJNA688646.

Received: 5 January 2021; Accepted: 13 April 2021

Published online: 26 April 2021

## References

1. Wiersinga, W. J., Rhodes, A., Cheng, A. C., Peacock, S. J. & Prescott, H. C. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review. *JAMA* **324**(8), 782–793 (2020).
2. Wang, J. & Du, G. COVID-19 may transmit through aerosol. *Irish J. Med. Sci.* **189**, 1143–1144 (2020).
3. Karia, R., Gupta, I., Khandait, H., Yadav, A. & Yadav, A. COVID-19 and its modes of transmission. *SN Compr. Clin. Med.* <https://doi.org/10.1007/s42399-020-00498-4> (2020).
4. De Maio, F., Posteraro, B., Ponziani, F. R., Cattani, P., Gasbarrini, A. & Sanguinetti, M. Nasopharyngeal microbiota profiling of SARS-CoV-2 infected patients (2020).
5. Fan, J. *et al.* The lung tissue microbiota features of 20 deceased patients with COVID-19. *J. Infect.* **81**(3), e64–e67 (2020).

6. Sahu, K. K., Mishra, A. K. & Lal, A. Trajectory of the COVID-19 pandemic: chasing a moving target. *Ann. Transl. Med.* **8**(11), 694. <https://doi.org/10.21037/atm-20-2793> (2020).
7. Sahu, K. K., Mishra, A. K. & Lal, A. COVID-19 disease: tackling a pandemic in 21st century. *QJM* **113**(7), 519–520. <https://doi.org/10.1093/qjmed/hcaa117> (2020).
8. De Boeck, I. *et al.* Comparing the healthy nose and nasopharynx microbiota reveals continuity as well as niche-specificity. *Front. Microbiol.* **8**, 2372 (2017).
9. de Steenhuisen Piters, W. A. A. *et al.* Nasopharyngeal microbiota, host transcriptome, and disease severity in children with respiratory syncytial virus infection. *Am. J. Respir. Crit. Care Med.* **194**(9), 1104–1115 (2016).
10. Lee, K. H. *et al.* The respiratory microbiota: associations with influenza symptomatology and viral shedding. *Ann. Epidemiol.* **37**, 51–56 (2019).
11. Allen, E. K. *et al.* Characterization of the nasopharyngeal microbiota in health and during rhinovirus challenge. *Microbiome* **2**(1), 22 (2014).
12. Mostafa, H. H., Fissel, J. A., Fanelli, B. *et al.* Metagenomic next-generation sequencing of nasopharyngeal specimens collected from confirmed and suspect COVID-19 patients. *MBio* **11**(6) (2020).
13. Jiang, L., Amir, A., Morton, J. T., Heller, R., Arias-Castro, E. & Knight, R. Discrete false-discovery rate improves identification of differentially abundant microbes. *MSystems* **2**(6) (2017).
14. Mallick, H., McIver, L. J., Rahnavard, A. *et al.* Multivariable association in population-scale meta-omics studies.
15. Quéromès, G. *et al.* Characterization of SARS-CoV-2 ORF6 deletion variants detected in a nosocomial cluster during routine genomic surveillance, Lyon, France. *Emerg Microbes Infect.* **10**(1), 167–177. <https://doi.org/10.1080/22221751.2021.1872351> (2021).
16. Di Segni, A. *et al.* Guided protocol for fecal microbial characterization by 16S rRNA-amplicon sequencing. *J. Vis. Exp.* <https://doi.org/10.3791/56845> (2018).
17. Bolyen, E. *et al.* Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat. Biotechnol.* **37**(8), 852–857. <https://doi.org/10.1038/s41587-019-0209-9> (2019).
18. Amir, A. *et al.* Deblur rapidly resolves single-nucleotide community sequence patterns. *mSystems* <https://doi.org/10.1128/mSystems.00191-16> (2017).
19. Lozupone, C., Lladser, M. E., Knights, D., Stombaugh, J. & Knight, R. UniFrac: an effective distance metric for microbial community comparison. *ISME J.* **5**(2), 169–172. <https://doi.org/10.1038/ismej.2010.133> (2011).
20. Mirarab, S., Nguyen, N. & Warnow, T. S. E. P. P. SATé-enabled phylogenetic placement. *Pac. Symp. Biocomput.* [https://doi.org/10.1142/9789814366496\\_0024](https://doi.org/10.1142/9789814366496_0024) (2012).
21. Xu, Z. Z. *et al.* Calour: An interactive, microbe-centric analysis tool. *mSystems* <https://doi.org/10.1128/mSystems.00269-18> (2019).
22. Oksanen, J., Blanchet, F. G., Friendly, M. *et al.* vegan: Community Ecology Package (2019).

## Acknowledgements

We thank the subjects who participated in our study. We thank the Sheba Medical Center support of the Sheba Microbiome Center. YH is also supported by the Israel Science Foundation (908/15), the I-CORE program (41/11), Bill and Melinda Gates Foundation (OPP1144149), the Helmsley Charitable Trust, and the ERC (758313).

## Author contributions

Study concept and design: T.B., R.H., A.A., Y.H. Acquisition of data: S.H., R.H., G.E., E.G.S., N.K., S.A., Y.H. Analysis and interpretation of data: T.B., R.H., A.A., Y.H. Drafting of the manuscript and critical review: T.B., S.H., R.H., G.E., E.G.S., N.K., A.A., S.A., Y.H.

## Funding

The authors do not have any financial relationships to disclose.

## Competing interests

The authors declare no competing interests.

## Additional information

**Correspondence** and requests for materials should be addressed to Y.H.

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

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



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

© The Author(s) 2021