SARS-CoV-2 infection and viral load are associated with the upper respiratory tract microbiome

Christian Rosas-Salazar, MD, MPH,[a]* Kyle S. Kimura, MD,[b]* Meghan H. Shilts, MHS, MS,[c] Britton A. Strickland, BS,[c,d] Michael H. Freeman, MD,[b] Bronson C. Wessinger, BE,[e] Veerain Gupta, MS,[e] Hunter M. Brown, MS,[c] Seesandra V. Rajagopala, PhD,[c] Justin H. Turner, MD, PhD,[b] and Suman R. Das, PhD[b,c]   *Nashville, Tenn*

Background: Little is known about the relationships between severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the respiratory virus responsible for the ongoing coronavirus disease 2019 (COVID-19) pandemic, and the upper respiratory tract (URT) microbiome.
Objective: We sought to compare the URT microbiome between SARS-CoV-2–infected and –uninfected adults and to examine the association of SARS-CoV-2 viral load with the URT microbiome during COVID-19.
Methods: We characterized the URT microbiome using 16S ribosomal RNA sequencing in 59 adults (38 with confirmed, symptomatic, mild to moderate COVID-19 and 21 asymptomatic, uninfected controls). In those with COVID-19, we measured SARS-CoV-2 viral load using quantitative reverse transcription PCR. We then examined the association of SARS-CoV-2 infection status and its viral load with the α-diversity, β-diversity, and abundance of bacterial taxa of the URT microbiome. Our main models were all adjusted for age and sex.

Results: The observed species index was significantly higher in SARS-CoV-2–infected than in –uninfected adults (β linear regression coefficient = 7.53; 95% CI, 0.17-14.89; *P* = .045). In differential abundance testing, 9 amplicon sequence variants were significantly different in both of our comparisons, with *Peptoniphilus lacrimalis*, *Campylobacter hominis*, *Prevotella 9 copri*, and an *Anaerococcus* unclassified amplicon sequence variant being more abundant in those with SARS-CoV-2 infection and in those with high viral load during COVID-19, whereas *Corynebacterium* unclassified, *Staphylococcus haemolyticus*, *Prevotella disiens*, and 2 *Corynebacterium_1* unclassified amplicon sequence variants were more abundant in those without SARS-CoV-2 infection and in those with low viral load during COVID-19.
Conclusions: Our findings suggest complex associations between SARS-CoV-2 and the URT microbiome in adults. Future studies are needed to examine how these viral-bacterial interactions can impact the clinical progression, severity, and recovery of COVID-19. (J Allergy Clin Immunol 2021;147:1226-33.)

Key words: 16S rRNA sequencing, coronavirus, COVID-19, microbiome, nasal, nasopharynx, respiratory, SARS-CoV-2

INTRODUCTION

The body of research suggests that interactions between common respiratory viruses and the upper respiratory tract (URT) microbiome can impact respiratory health.[1] In this context, we and others have shown that viral-bacterial interactions can influence viral load,[2] host transcriptome patterns,[3,4] acute severity,[3,4] and even long-term outcomes of common respiratory viruses (such as respiratory syncytial virus),[5] as well as the acute immune response to these infectious agents.[2,6] However, little is known about the relationship between severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2), the respiratory virus responsible for the ongoing coronavirus disease 2019 (COVID-19) pandemic, and the URT microbiome. To start filling this gap in knowledge, we (1) compared the URT microbiome between SARS-CoV-2–infected and –uninfected adults and (2) examined the association of SARS-CoV-2 viral load (an independent predictor of disease severity[7,8]) with the URT microbiome during COVID-19. Full details of the methods are available in this article's Methods section in the Online Repository at www.jacionline.org.

Abbreviations used
ASV: Amplicon sequence variant
COVID-19: Coronavirus disease 2019
LRT: Lower respiratory tract
SARS-CoV-2: Severe acute respiratory syndrome coronavirus 2
URT: Upper respiratory tract

RESULTS AND DISCUSSION

Fifty-nine adults were included in the current study (38 with confirmed, symptomatic, mild to moderate COVID-19 [based on criteria from the World Health Organization[9]] enrolled as part of a clinical trial and 21 asymptomatic, uninfected adults enrolled to serve as a control group[10]). Their baseline characteristics are presented in Table I. The median (interquartile range) age was 30 (27-45) years. None of the participants had used antibiotics in the previous 2 weeks or were using intranasal medications at the time of sampling. There were no significant differences in baseline characteristics between those with and without SARS-CoV-2 infection, although those infected were more likely to have at least 1 other comorbidity (Table I).

The URT microbiome was characterized in midturbinate swabs from all 59 participants using 16S ribosomal RNA sequencing of the V4 region in an Illumina MiSeq platform with 2×250 base-pair reads as previously described.[11-13] Following quality control and initial data processing steps, the median (interquartile range) sequence count per sample was 8,142 (3,423-15,233) among all samples. The most abundant genera in SARS-CoV-2–uninfected samples included *Staphylococcus* (41.56%), *Corynebacterium_1* (28.09%), *Moraxella* (8.48%), *Dolosigranulum* (3.56%), and *Neisseria* unclassified (1.98%), whereas the most abundant genera in SARS-CoV-2–infected samples included *Corynebacterium_1* (33.66%), *Staphylococcus* (29.34%), *Dolosigranulum* (5.29%), *Peptoniphilus* (3.91%), and *Lawsonella* (3.22%) (Fig 1).

The α-diversity of the URT microbiome was overall higher in SARS-CoV-2–infected than in –uninfected adults, although only the observed species index was significantly different (β linear regression coefficient = 7.53; 95% CI, 0.17-14.89; *P* = .045) (Fig 2, *A*). There were no significant differences in any of the measured β-diversity metrics between groups (*P* > .05 for the Bray-Curtis and Jaccard indices using permutational multivariate ANOVA[14]) (Fig 2, *B*). In differential abundance testing using DE-Seq2,[15] 21 amplicon sequence variants (ASVs) were significantly different between groups (Fig 3, *A*), with 13 being more abundant (including *Brevundimonas*, *Corynebacterium*, *Granilucatella*, *Anaerococcus*, and *Peptoniphulus* ASVs) and 8 being less abundant (including *Corynebacterium_1*, *Prevotella*, *Staphylococcus*, *Anaerostipes*, and *Neisseria* ASVs) in SARS-CoV-2–infected *versus* –uninfected adults (Fig 3, *B*).

The median (interquartile range) cycle threshold value for the detection of SARS-CoV-2 nucleocapsid gene region 1 (N1) using quantitative reverse transcription PCR was 21.38 (18.64-23.27). The most abundant genera in SARS-CoV-2–infected samples with high viral load (defined as N1 cycle threshold values by quantitative reverse transcription PCR below the median) included *Corynebacterium_1* (35.69%), *Staphylococcus* (28.83%), *Peptoniphilus* (6.67%%), *Anaerococcus* (4.79%%), and *Bacteroides* (3.83%), whereas the most abundant genera in SARS-CoV-2–infected samples with low viral load included *Corynebacterium_1* (41.44%), *Staphylococcus* (20.75%),

Dolosigranulum (12.30%), *Lawsonella* (4.50%), and *Peptoniphilus* (2.76%).

In the adults with COVID-19, there were no significant associations between high *versus* low SARS-CoV-2 viral load and any of the α-diversity or β-diversity metrics of the URT microbiome (*P* > .05 for all comparisons). In differential abundance testing using DESeq2, 21 ASVs were significantly different between groups (Fig 4, *A*), with 9 being more abundant (including *Neisseriacea*, *Anaerococcus*, *Peptoniphulus*, *Campylobacter*, and *Enterococcus* ASVs) and 12 being less abundant (including *Corynebacterium_1*, *Staphylococcus*, *Granilucatella*, *Neisseria*, and *Prevotella* ASVs) in those with high viral load when compared with those with low viral load (Fig 4, *B*). The abundance of 14 of these 21 ASVs was significantly different and had a consistent direction of association using a similar definition of high viral load but based on N2 cycle threshold values by quantitative reverse transcription PCR (Fig 4, *B*). Furthermore, the abundance of 9 of these 21 ASVs was significantly different and had a consistent direction of association between adults with and without SARS-CoV-2 infection, with *Peptoniphilus lacrimalis*, *Campylobacter hominis*, *Prevotella 9 copri*, and an *Anaerococcus* unclassified ASV being more abundant in those with SARS-CoV-2 infection and in those with high viral load during COVID-19, whereas *Corynebacterium* unclassified, *Staphylococcus haemolyticus*, *Prevotella disiens*, and 2 *Corynebacterium_1* unclassified ASVs were more abundant in those without SARS-CoV-2 infection and in those with low viral loads during COVID-19 (Fig 3 and Fig 4, *B*).

To our knowledge, only 3 other published studies have compared the URT or lower respiratory tract (LRT) microbiome between SARS-CoV-2–infected and –uninfected adults,[16-18] and no previous published studies have examined the association of SARS-CoV-2 viral load with the URT or LRT microbiome during COVID-19. In a study using 16S ribosomal RNA sequencing, De Maio et al[16] found no differences in the α-diversity, β-diversity, or abundance of taxa of the URT microbiome between adults with and without SARS-CoV-2. In another study using metatranscriptomics, Zhang et al[17] also found no differences in the α-diversity of the URT microbiome between groups, although the α-diversity of the LRT was lower and 18 species (of both the URT and the LRT) were less abundant in adults with COVID-19, none of which overlapped with the ones we found to be differentially abundant in adults with and without SARS-CoV-2 in our study. In addition to the methodological differences between these studies and ours, neither of them included a group of asymptomatic, uninfected controls, as the comparisons group for both these studies included adults with other acute respiratory infections. In one study using metatranscriptomics that did include a group of uninfected, asymptomatic controls, Shen et al[18] found differences in the β-diversity of the LRT microbiome between adults with and without SARS-CoV-2 infection, but no α-diversity or differential abundance analyses were performed and this study was smaller than ours (including only 8 SARS-CoV-2–infected adults). Furthermore, none of the 3 aforementioned studies adjusted for potential confounders in their statistical analyses, which can also explain the discrepant results, because all our statistical analyses included age and sex as covariates. To prevent overfitting and maximize sample size, we did not include other covariates in our main models. However, we obtained similar results in sensitivity analyses including other potential confounders. For example, 14 (66.67%) and 18 (85.71%) of the 21 ASVs that were significantly

TABLE I. Baseline characteristics of study participants by SARS-CoV-2 infection status*†

Baseline characteristic	All (n = 59)	Uninfected (n = 21)	Infected (n = 38)	P value‡
Age (y)	30.00 (27.00- 45.00)	30.00 (29.00- 37.00)	30.50 (25.25- 50.00)	.61
Male sex	33 (55.93)	12 (57.14)	21 (55.26)	.89
Use of antibiotics in the last 2 wk	0	0	0	—
Current use of intranasal medications	0	0	0	—
Nasal congestion	—	—	32 (91.43)	—
Loss of taste or smell	—	—	18 (50.00)	
Cough	—	—	33 (91.67)	—
Fever	—	—	0	—
Shortness of breath	—	—	18 (50.00)	—
Current smoker	2 (3.51)	1 (5.00)	1 (2.70)	.65
Obese	16 (35.56)	4 (23.53)	12 (42.86)	.19
Diabetes	3 (5.26)	0 (0.00)	3 (8.11)	.19
Hypertension	9 (15.79)	2 (10.00)	7 (18.92)	.38
Lung disease	4 (7.02)	0 (0.00)	4 (10.81)	.13
Heart disease	2 (3.51)	1 (5.00)	1 (2.70)	.65

*The data are presented as median (interquartile range) for continuous variables or number (%) for categorical variables.
†The estimates were calculated for participants with complete data.
‡P value for the comparison between groups using a Mann-Whitney U or Pearson χ^2 test, as appropriate.
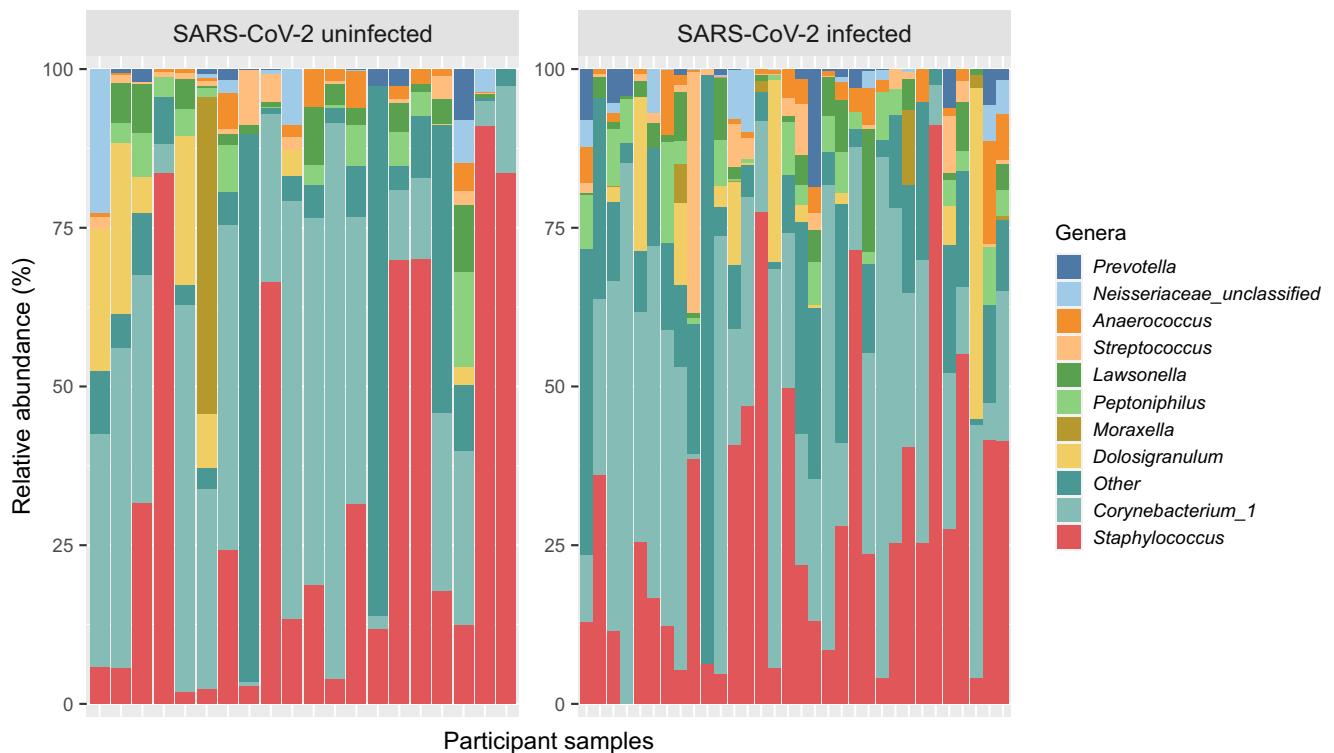
FIG 1. Stacked bar chart of the relative abundance of the most common genera of the URT microbiome in adults with and without SARS-CoV-2 infection. The bars represent individual participant samples. Only the top 10 most abundant genera across all samples are shown. The other genera were collapsed into the "Other" category. The genera were ordered according to their relative abundance across all samples.

different between SARS-CoV-2–infected and –uninfected adults in our initial analyses remained differentially abundant when adding the presence of obesity or having at least 1 other comorbidity, respectively, as covariates to the models (data not shown).

The data on bacterial coinfections in patients with COVID-19 are rapidly emerging.[19] Interestingly, 1 previous report found a high abundance of *Brevundimonas* spp in the lungs of 20 deceased

adults with COVID-19 and 1 case report described a SARS-CoV-2 + *Granulicatella adiacens* coinfection, both of which were differentially abundant between SARS-CoV-2–infected and –uninfected adults in our study and are overall rare taxa of the URT or LRT microbiome in the general population.[20,21] Furthermore, we found multiple taxa to be associated not only with SARS-CoV-2 infection status but also with a higher viral load during COVID-19 (such as
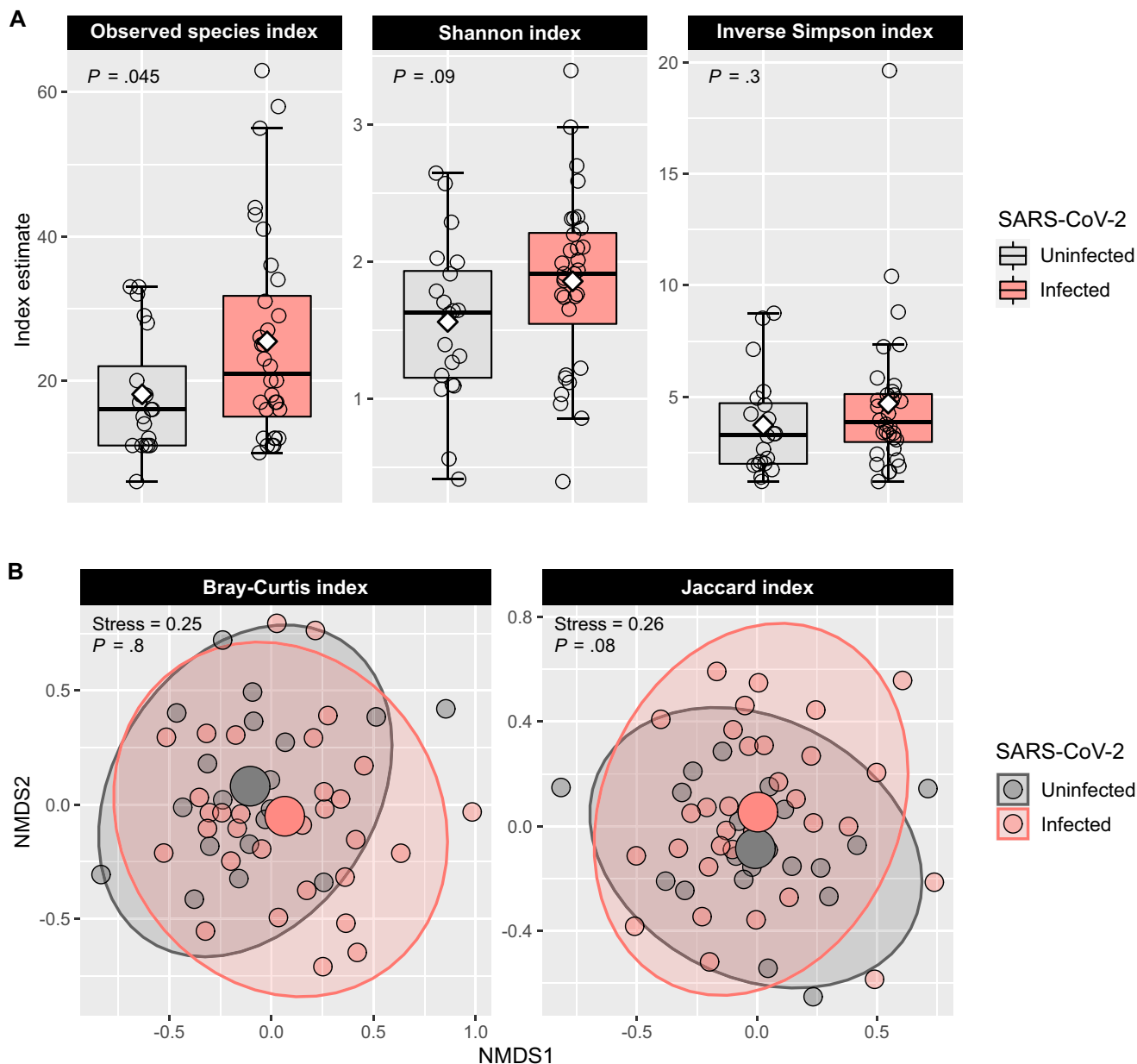
FIG 2. The α- and β-diversity of the URT microbiome in adults with and without SARS-CoV-2 infection. **A,** The box-and-whisker plots show the mean (diamond), median (middle bar), first quartile (lower bar), third quartile (upper bar), minimum observation above the lowest fence (lower whisker), and maximum observation below the upper fence (upper whisker) of common α-diversity metrics for each group. The P values for the comparison between groups using linear regression models including age and sex as covariates are also shown. **B,** The scatter plots show each participant's microbial community composition (small circles) by group, as well as each group's centroid (large circles) and 95% CI ellipses. The scatter plots were generated using nonmetric-multidimensional scaling (NMDS) ordination based on common β-diversity metrics. For ease of visualization, only 2 dimensions were used. The NMDS stress values and the P values for the comparison between groups using permutational multivariate ANOVA models including age and sex as covariates are also shown.

P lacrimalis, C hominis, Prevotella 9 copri, and an *Anaerococcus* unclassified ASV). Taken together, these findings suggest that SARS-CoV-2 can directly impact the abundance of certain URT taxa. Within a given genus, associations appeared to be taxon-specific, because SARS-CoV-2 infection and high viral load during COVID-19 were associated with an increase in certain

Corynebacterium ASVs but with a decrease in others. Our findings of both SARS-CoV-2 infection status and its viral load being associated with various taxa of the URT microbiome could also indicate that the microbiome patterns we observed are unique to this respiratory virus. This is supported by the results of the study by Zhang et al[17] described above, which showed substantial

A

B
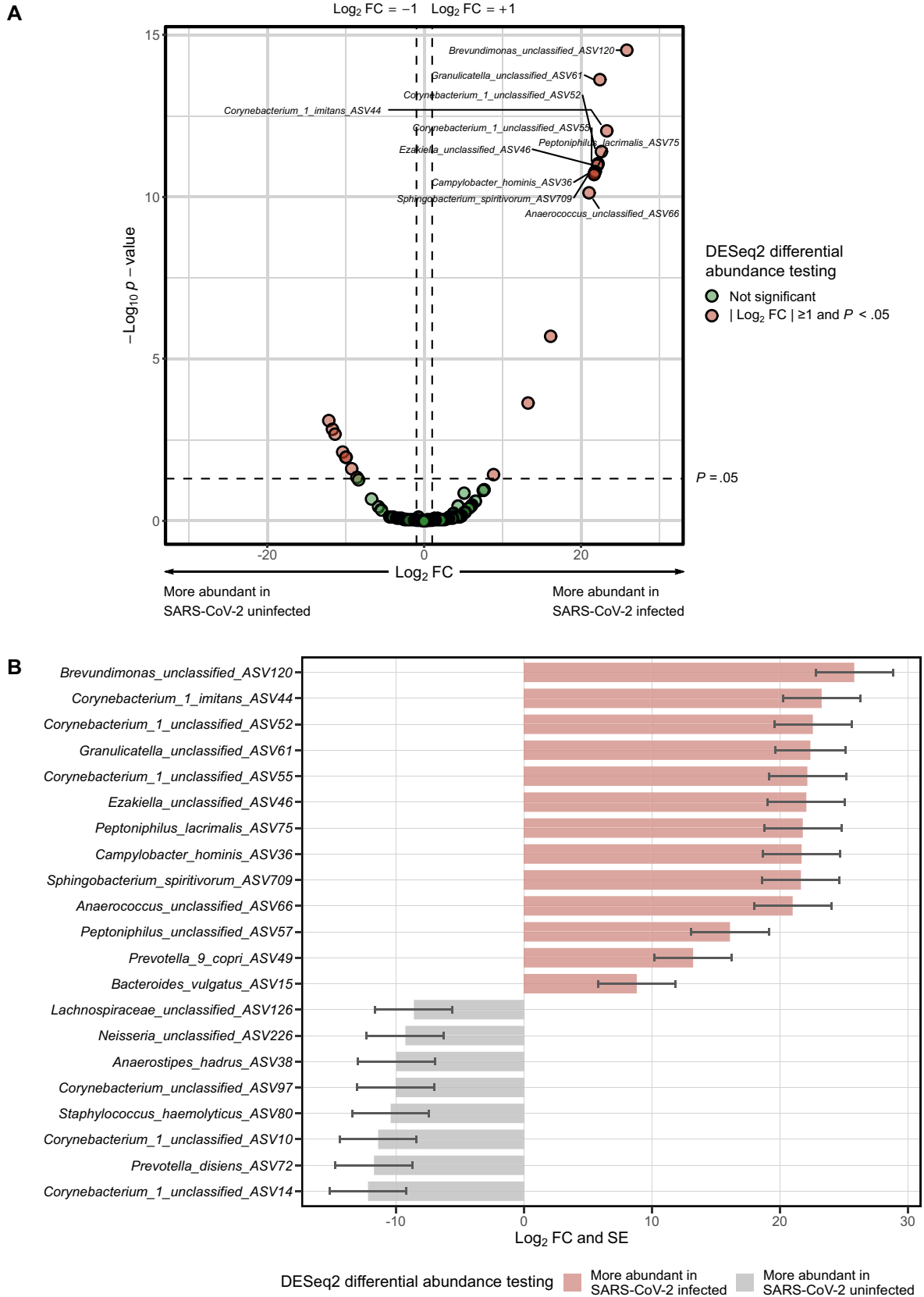
FIG 3. Differences in the abundance of taxa of the URT microbiome between adults with and without SARS-CoV-2 infection. Differential abundance testing was conducted using DESeq2 models at the ASV level including age and sex as covariates. **A,** Volcano plot of $\log_2$ fold change (FC) *vs* statistical significance. The red circles indicate ASVs that were significantly different between groups. Only the top 10 most significantly different ASVs are labeled. **B,** Bar plot depicting the $\log_2$ FCs and SEs for ASVs that were significantly different between groups.

A

B
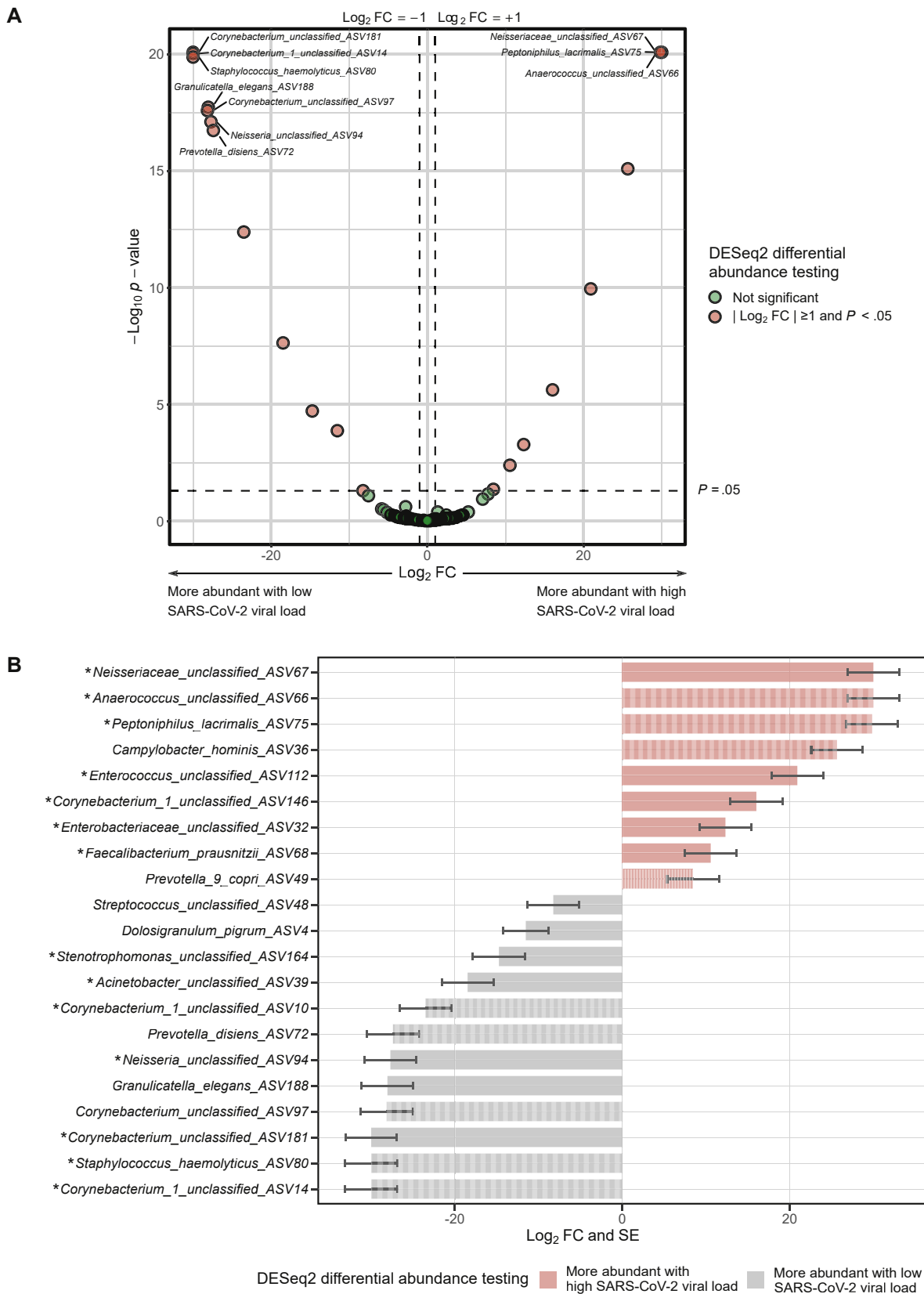
FIG 4. Differences in the abundance of taxa of the URT microbiome between SARS-CoV-2–infected adults with and without high viral load (defined as a quantitative reverse transcription PCR cycle threshold value below the median for the detection of SARS-CoV-2 nucleocapsid gene region 1 [N1]). Differential abundance testing was conducted using DESeq2 models at the ASV level including age and sex as covariates. **A,** Volcano plot of log2 fold change (FC) *vs* statistical significance. The red circles indicate ASVs

differences in the URT and LRT microbiome between adults with COVID-19 pneumonia and those with pneumonia due to other infectious agents (many of which were likely viral), as well as those of a recent study showing distinct gut microbiome signatures in adults with SARS-CoV-2 *versus* those with influenza H1N1.[22] In the same context, we have previously shown that URT microbiome profiles in children with other common acute respiratory infections are virus-specific.[23]

Our study has numerous strengths, such as the inclusion of an overall young population with a limited number of comorbidities, no recent use of antibiotics or current use of intranasal medications, and a true control group. We should also acknowledge several limitations. First, our study is cross-sectional and we lacked longitudinal samples. Thus, there is a possibility of reverse causation (eg, that URT microbiome patterns associated with SARS-CoV-2 infection status that we found preceded the actual infection). Several of the taxa we found to be differentially abundant between groups are overall uncommon in the general population, which makes this unlikely, but viral-bacterial interactions are likely complex and associations could be bidirectional (eg, with SARS-CoV-2 impacting certain URT taxa and other URT taxa in turn influencing SARS-CoV-2 replication). Second, because it is inherent to 16S ribosomal RNA sequencing, we were unable to accurately classify some taxa at the species level. Third, there is a possibility of residual confounding. For instance, we only captured data on antibiotic use in the 2 weeks before enrollment and it is possible that some participants had used antibiotics before that time frame, which could have impacted our results. In addition, we lacked data on the participants' atopic status and did not test for the presence of other respiratory viruses. Because our samples were all obtained in the southern United States between April and June of 2020, coinfections with influenza or respiratory syncytial virus are unlikely, but other respiratory viruses (particularly human rhinovirus or other coronaviruses[24]) could have been present. Fourth, we did not have LRT samples. However, the URT is the portal of entry and an active site of replication of SARS-CoV-2, as well as a common harboring site for potential pathogens, thus of critical importance in the pathogenesis of this respiratory virus.[25] Last, our results cannot be extended to adults with asymptomatic, severe, or critical COVID-19,[9] because only those with symptomatic, mild to moderate COVID-19 were included in our study. In spite of these limitations, our study is a stepping stone in examining the role of the URT microbiome in SARS-CoV-2–related outcomes and in understanding the development of bacterial coinfections during COVID-19.

In summary, we found substantial differences in the URT microbiome between SARS-CoV-2–infected and –uninfected adults. Furthermore, we show that the SARS-CoV-2 viral load is associated with the URT microbiome during COVID-19. Future studies with larger sample sizes and serial sample collection will be needed to examine how SARS-CoV-2 interacts with the URT

microbiome and how these viral-bacterial interactions can impact the clinical progression, severity, and recovery of COVID-19.

Key messages

- **There are substantial differences in the URT microbiome between SARS-CoV-2–infected and –uninfected adults.**
- **The SARS-CoV-2 viral load is associated with the URT microbiome during COVID-19.**

REFERENCES

1. Man WH, de Steenhuijsen Piters WA, Bogaert D. The microbiota of the respiratory tract: gatekeeper to respiratory health. Nat Rev Microbiol 2017;15:259-70.
2. Ederveen THA, Ferwerda G, Ahout IM, Vissers M, de Groot R, Boekhorst J, et al. Haemophilus is overrepresented in the nasopharynx of infants hospitalized with RSV infection and associated with increased viral load and enhanced mucosal CXCL8 responses. Microbiome 2018;6:10.
3. de Steenhuijsen Piters WA, Heinonen S, Hasrat R, Bunsow E, Smith B, Suarez-Arrabal MC, et al. Nasopharyngeal microbiota, host transcriptome, and disease severity in children with respiratory syncytial virus infection. Am J Respir Crit Care Med 2016;194:1104-15.
4. Sonawane AR, Tian L, Chu CY, Qiu X, Wang L, Holden-Wiltse J, et al. Microbiome-transcriptome interactions related to severity of respiratory syncytial virus infection. Sci Rep 2019;9:13824.
5. Rosas-Salazar C, Shilts MH, Tovchigrechko A, Schobel S, Chappell JD, Larkin EK, et al. Nasopharyngeal lactobacillus is associated with childhood wheezing illnesses following respiratory syncytial virus infection in infancy. J Allergy Clin Immunol 2018;142:1447-56.
6. Shilts MH, Rosas-Salazar C, Turi KN, Rajan D, Rajagopala SV, Patterson MF, et al. Nasopharyngeal haemophilus and local immune response during infant respiratory syncytial virus infection [published online ahead of print July 3, 2020]. J Allergy Clin Immunol. https://doi.org/10.1016/j.jaci.2020.06.023.
7. Liu Y, Yan LM, Wan L, Xiang TX, Le A, Liu JM, et al. Viral dynamics in mild and severe cases of COVID-19. Lancet Infect Dis 2020;20:656-7.
8. Pujadas E, Chaudhry F, McBride R, Richter F, Zhao S, Wajnberg A, et al. SARS-CoV-2 viral load predicts COVID-19 mortality. Lancet Respir Med 2020;8:e70.
9. World Health Organization. Clinical management of COVID-19: interim guidance, 27 May 2020. Available at: https://apps.who.int/iris/handle/10665/332196. Accessed February 23, 2021.
10. Kimura KS, Freeman MH, Wessinger BC, Gupta V, Sheng Q, Huang LC, et al. Interim analysis of an open-label randomized controlled trial evaluating nasal irrigations in non-hospitalized patients with COVID-19. Int Forum Allergy Rhinol 2020;10:1325-8.
11. Singh K, Gobert AP, Coburn LA, Barry DP, Allaman M, Asim M, et al. Dietary arginine regulates severity of experimental colitis and affects the colonic microbiome. Front Cell Infect Microbiol 2019;9:66.
12. Hiremath G, Shilts MH, Boone HH, Correa H, Acra S, Tovchigrechko A, et al. The salivary microbiome is altered in children with eosinophilic esophagitis and correlates with disease activity. Clin Transl Gastroenterol 2019;10:e00039.
13. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 2013;79:5112-20.
14. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community Ecology Package, 2014. Available at: http://CRAN.R-project.org/package=vegan. Accessed December 15, 2020.
15. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

◀─────────

that were significantly different between groups. Only the top 10 most significantly different ASVs are labeled. **B,** Bar plot depicting the $\log_2$ FCs and SEs for ASVs that were significantly different between groups. The asterisks indicate ASVs that were significantly different between groups and had a consistent direction of association in similar DESeq2 analyses that used a definition of high viral load based on a quantitative reverse transcription PCR cycle threshold value below the median for the detection of SARS-CoV-2 nucleocapsid gene region 2 (N2). The striped bars indicate ASVs that were significantly different between groups and had a consistent direction of association in similar DESeq2 analyses comparing adults with and without SARS-CoV-2 infection.

16. De Maio F, Posteraro B, Ponziani FR, Cattani P, Gasbarrini A, Sanguinetti M. Nasopharyngeal microbiota profiling of SARS-CoV-2 infected patients. Biol Proced Online 2020;22:18.

17. Zhang H, Ai JW, Yang W, Zhou X, He F, Xie S, et al. Metatranscriptomic characterization of COVID-19 identified a host transcriptional classifier associated with immune signaling [published online ahead of print May 28, 2020]. Clin Infect Dis. https://doi.org/10.1093/cid/ciaa663.

18. Shen Z, Xiao Y, Kang L, Ma W, Shi L, Zhang L, et al. Genomic diversity of severe acute respiratory syndrome-coronavirus 2 in patients with coronavirus disease 2019. Clin Infect Dis 2020;71:713-20.

19. Lansbury L, Lim B, Baskaran V, Lim WS. Co-infections in people with COVID-19: a systematic review and meta-analysis. J Infect 2020;81:266-75.

20. Fan J, Li X, Gao Y, Zhou J, Wang S, Huang B, et al. The lung tissue microbiota features of 20 deceased patients with COVID-19. J Infect 2020;81:e64-7.

21. Campanella T. COVID-19 complicates an already difficult presentation of infective endocarditis. Contagion Live.

22. Gu S, Chen Y, Wu Z, Chen Y, Gao H, Lv L, et al. Alterations of the gut microbiota in patients with COVID-19 or H1N1 influenza. Clin Infect Dis 2020;5.

23. Rosas-Salazar C, Shilts MH, Tovchigrechko A, Schobel S, Chappell JD, Larkin EK, et al. Differences in the nasopharyngeal microbiome during acute respiratory tract infection with human rhinovirus and respiratory syncytial virus in infancy. J Infect Dis 2016;214:1924-8.

24. Kim D, Quinn J, Pinsky B, Shah NH, Brown I. Rates of co-infection between SARS-CoV-2 and other respiratory pathogens. JAMA 2020;323: 2085-6.

25. Sungnak W, Huang N, Becavin C, Berg M, Queen R, Litvinukova M, 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.

METHODS
Overview of the study design

For the current study, we included nonhospitalized patients 18 years or older who were diagnosed with SARS-CoV-2 infection (confirmed by qualitative PCR) at Vanderbilt University Medical Center or one of its affiliated centers in Nashville, Tenn. These patients were enrolled as part of a clinical trial examining the effect of several types of nasal irrigations on URT symptoms and viral load during COVID-19. The detailed methods for this clinical trial have been previously reported.[E1] Exclusion criteria for these patients included current use of nasal saline irrigations or other intra-nasal medications, inability to perform nasal irrigations or to collect URT samples in a separate house bathroom or away from household contacts, or need for hospitalization related to SARS-CoV-2 infection. Thus, only patients with asymptomatic, mild, or moderate COVID-19 (based on criteria from the World Health Organization[E2]) were included in the clinical trial. Eligible patients were contacted and enrolled in the study within 24 hours of initial diagnosis. In parallel, we enrolled asymptomatic adults within the Vanderbilt University community (including employees, students, and faculty, among others) to serve as a control group. Exclusion criteria for these asymptomatic adults were similar to those used for patients enrolled in the clinical trial.

Following adequate training, all participants were asked to obtain a midturbinate swab on the day of enrollment (ie, before any study intervention for those enrolled in the clinical trial) using a self-collection kit (FLOQSwabs; Copan Diagnostics, Inc, Murrieta, Calif). These enrollment samples were used for the current study. The collection of all samples included in the current study occurred between April and June of 2020. Each adult provided informed consent for their participation. The Institutional Review Board of Vanderbilt University approved this study.

SARS-CoV-2 testing by quantitative RT-PCR

To rule out asymptomatic infection in controls and measure viral load in SARS-CoV-2–infected patients, we performed quantitative reverse transcription PCR (RT-qPCR) in the midturbinate swabs. Total RNA was extracted from the swabs using a phenol-chloroform–based method. The swabs were placed in Red 1.5 mL RINO screw-cap tubes (NextAdvance, Troy, NY) prefilled with RNase-free zirconium oxide beads and QIAzol Lysis Reagent (Qiagen, Germantown, Md) was added. Samples were then homogenized in a Bullet Blender 24 Gold (NextAdvance) for 3 minutes at maximum speed. Following homogenization, genomic DNA was eliminated with gDNA Eliminator columns (Qiagen) and RNA purified using the RNeasy Mini Plus Kit (Qiagen) following the manufacturer's protocol. The RNA quality was measured using an Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, Calif). The United States Centers for Disease Control and Prevention primers and probes designed for the detection of SARS-CoV-2 (2019-nCoV) were purchased from Integrated DNA Technologies (Coralville, Iowa).[E3] Both the SARS-CoV-2 nucleocapside gene region 1 (N1) and nucleocapside gene region 2 (N2) were targeted for the detection of SARS-CoV-2. RNase P was also examined as a measure of RNA quality and quantity. RT-qPCR was performed using SuperScript III One-Step RT-PCR System with Platinum Taq DNA Polymerase (Invitrogen, Carlsbad, Calif) as per manufacturer's instructions on a Bio-Rad CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, Calif). Plasmid controls for 2 SARS-CoV-2 nucleocapsid and RNase P were also ordered from Integrated DNA Technologies at a concentration of 66,666 copies/reaction. No template controls and an extraction negative were used as negative controls. Reactions were prepared using 12.5 μL Super-Script III Master Mix (ThermoFisher, Waltham, Mass), 1 μL each 400 nm forward and reverse primer, 1 μL 400 nM FAM-labeled probe, 1 μL Platinum Taq polymerase, 3 μL template RNA, and 7.25 μL PCR Certified Water (Teknova, Hollister, Calif). RNA was reverse transcribed at 50°C for 15 minutes, and PCR conditions were run at 95°C denaturation step for 2 minutes, followed by 40 cycles of 95°C for 15 seconds and 55°C for 30 seconds. The cycle threshold values were captured and calculated by the CFX Maestro (Bio-Rad) software and used as a measure of viral load.

Characterization of the URT microbiome

The methods used to characterize the URT microbiome have been previously described in detail.[E4-E6] In brief, we extracted bacterial DNA from the midturbinate swabs using the DNeasy PowerSoil kit (Qiagen). Dual-indexed universal primers appended with Illumina-compatible adapters were used to amplify the hypervariable V4 region of the 16S ribosomal RNA (rRNA) gene. Amplicons targeting the V4 region of the bacterial 16S rRNA gene were generated by combining 7 μL template, 12.5 μL MyTaq HS Mix (Bioline, Memphis, Tenn), 0.75 μL dimethyl sulfoxide (Sigma, St Louis, Mo), 3 μL PCR Certified Water (Teknova), and 1 μL of each 10 μM primer. DNA was denatured at 95°C for 2 minutes, followed by 30 cycles of 95°C for 20 seconds, 55°C for 15 seconds, and 72°C for 5 minutes, and a final extension at 72°C for 10 minutes. Each amplified sample was run on a 1% agarose gel to confirm reaction success. Amplicons were cleaned and normalized with the SequalPrep Normalization Kit (ThermoFisher). Normalized amplicons were pooled and cleaned with 1X AMPure XP beads (Beckman Coulter, Brea, Calif). The pool was then sequenced on an Illumina MiSeq platform with 2 × 250 base-pair reads. Negative controls (2 PCR and 1 extraction negative control) were amplified and sequenced concurrently with the participant samples. Following the sequencing procedure, only a small fraction of 16S rRNA sequences were found in the negative controls compared with participant samples and the bacterial sequences recovered had little overlap with the participant samples.

Next, we processed the 16S rRNA sequences using the R *dada2* package by following its standard operating procedure (available at: https://benjjneb. github.io/dada2/tutorial.html).[E7] To this end, sequences were grouped into ASVs and taxonomy was assigned using the SILVA reference database.[E8] Low-quality sequences, chimeras, and nonbacterial sequences were discarded as part of the *dada2* pipeline. To remove any suspected contaminants that were found in the negative controls, we processed the remaining sequences using the R package *decontam*.[E9] We then retained ASVs that were present in more than 1 sample and discarded samples with less than 1000 sequences (n = 7) using the R *phyloseq* package.[E10] Last, we calculated the relative abundances of individual taxa using simple proportions.

Statistical analyses

Statistical analyses were conducted in R version 3.1.10.[E11] First, we rarefied the processed data set to the lowest library size of all samples (n = 1063). This rarefaction process was repeated multiple times (n = 400) and the results were averaged. We then calculated common α-diversity (ie, observed species, Shannon, and inverse Simpson indices) and β-diversity (ie, Bray-Curtis and Jaccard indices) metrics using the R *phyloseq* package.[E10] To examine the differences in α-diversity metrics between groups, we used linear regression. To evaluate the differences in β-diversity metrics between groups, we used nonmetric-multidimensional scaling (using 999 iterations) and permutational multivariate ANOVA (using 999 permutations), as implemented in the *metamds* and *adonis2* functions of the R *vegan* package, respectively.[E12,E13] We next performed differential abundance testing with the R *DESeq2* package using a nonrarified data set.[E14] The statistical analyses were all conducted at the ASV level, and our main models all included age and sex as covariates. For ASVs unclassified at the species level, their identities at higher levels were used.

For SARS-CoV-2–positive samples, we defined high viral load as an RT-qPCR cycle threshold value below the median for the detection of SARS-CoV-2 nucleocapside gene region 1 (N1). We also conducted secondary analyses using a definition of high viral load based on an RT-qPCR cycle threshold value below the median for the detection of SARS-CoV-2 nucleocapside gene region 2 (N2).

Statistical significance was defined as P less than .05 after controlling for multiple comparisons using the Benjamini-Hochberg procedure, when appropriate.[E15] For *DESeq2* analyses, we additionally used an absolute fold change greater than or equal to 2 (ie, an absolute fold change ≥1 in the log$_2$ scale) to define statistical significance. Figures were created in R version 3.1.10.[E11] Minor esthetic edits to figures (eg, paneling, text insertion, or label formatting) were made with Inkscape version 1.0.1 (available at: https://inkscape.org/).

REFERENCES

E1. Kimura KS, Freeman MH, Wessinger BC, Gupta V, Sheng Q, Huang LC, et al. Interim analysis of an open-label randomized controlled trial evaluating nasal irrigations in non-hospitalized patients with COVID-19. Int Forum Allergy Rhinol 2020;10:1325-8.

E2. World Health Organization. Clinical management of COVID-19: interim guidance, 27 May 2020. World Health Organization. Available at: https://apps.who.int/iris/handle/10665/332196. Accessed February 23, 2021.

E3. Centers for Disease Control and Prevention. CDC 2019-Novel Coronavirus (2019-nCoV) Real-Time RT-PCR Diagnostic Panel. Available at: https://www.fda.gov/media/134922/download. Accessed December 15, 2020..

E4. Singh K, Gobert AP, Coburn LA, Barry DP, Allaman M, Asim M, et al. Dietary arginine regulates severity of experimental colitis and affects the colonic microbiome. Front Cell Infect Microbiol 2019;9:66.

E5. Hiremath G, Shilts MH, Boone HH, Correa H, Acra S, Tovchigrechko A, et al. The salivary microbiome is altered in children with eosinophilic esophagitis and correlates with disease activity. Clin Transl Gastroenterol 2019;10:e00039.

E6. Kozich JJ, Westcott SL, Baxter NT, Highlander SK, Schloss PD. Development of a dual-index sequencing strategy and curation pipeline for analyzing amplicon sequence data on the MiSeq Illumina sequencing platform. Appl Environ Microbiol 2013;79:5112-20.

E7. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. Nat Methods 2016;13:581-3.

E8. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, et al. SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB. Nucleic Acids Res 2007;35:7188-96.

E9. Davis NM, Proctor DM, Holmes SP, Relman DA, Callahan BJ. Simple statistical identification and removal of contaminant sequences in marker-gene and metagenomics data. Microbiome 2018;6:226.

E10. McMurdie PJ, Holmes S. phyloseq: an R package for reproducible interactive analysis and graphics of microbiome census data. PLoS One 2013;8:e61217.

E11. R Development Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2006.

E12. Anderson MJ. A new method for non-parametric multivariate analysis of variance. Austral Ecol 2001;26:32-46.

E13. Oksanen J, Blanchet FG, Kindt R, Legendre P, Minchin PR, O'Hara RB, et al. vegan: Community Ecology Package, 2014. Available at: http://CRAN.R-project.org/package=vegan. Accessed December 15, 2020.

E14. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol 2014;15:550.

E15. Benjamini Y, Hochberg Y. Controlling the false discovery rate—a practical and powerful approach to multiple testing. J Royal Stat Soc Series B-Methodol 1995;57:289-300.