

SHORT REPORT

Open Access



Nasopharyngeal Microbiota in SARS-CoV-2 Positive and Negative Patients

Phillip A. Engen¹, Ankur Naqib¹, Cheryl Jennings², Stefan J. Green³, Alan Landay⁴, Ali Keshavarzian^{1,5,6†} and Robin M. Voigt^{1,6*†} 

Abstract

We investigated nasopharyngeal microbial community structure in COVID-19-positive and -negative patients. High-throughput 16S ribosomal RNA gene amplicon sequencing revealed significant microbial community structure differences between COVID-19-positive and -negative patients. This proof-of-concept study demonstrates that: (1) nasopharyngeal microbiome communities can be assessed using collection samples already collected for SARS-CoV-2 testing (viral transport media) and (2) SARS-CoV-2 infection is associated with altered dysbiotic microbial profiles which could be a biomarker for disease progression and prognosis in SARS-CoV-2.

Keywords: SARS-CoV-2, COVID-19, Nasopharyngeal, Microbiota, Dysbiosis

Introduction

Since the appearance of severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infections in 2019, cases of coronavirus disease 2019 (COVID-19) have spread rapidly around the world infecting over 152 million people and claiming over 3.2 million lives, as of May 3, 2021 (<https://covid19.who.int/>). COVID-19 is primarily transmitted through the respiratory tract via aerosolized droplets containing viral particles [1]. Host-associated microorganisms can influence viral infectivity [2] and are the major players mediating immune-inflammation [3]. Conversely, studies have shown that viruses can modulate microbiota community in the oropharyngeal and respiratory tract [4]. However, few studies have examined the nasopharyngeal microbiota in COVID-19 patients. To date, two studies found no significant differences in microbial diversity between COVID-19-positive and negative patients [5, 6] whereas two studies did note

differences in community structure [7, 8]. One of these studies reports a significant decrease in microbial diversity, differences between microbial communities, and a higher abundance of *Propionibacteriaceae* and a reduction in *Corynebacterium accolens* in COVID-19-positive relative to -negative patients [7]. The other study observed differences between microbial communities, with a significantly lower abundance of *Fusobacterium periodonticum* in COVID-19 positive compared to negative patients [8]. These mixed results highlight the critical need for more information.

This proof-of-concept study utilized viral transport media (VTM) used for SARS-CoV-2 nasopharyngeal swab sample collection/testing to assess nasopharyngeal microbiota communities in COVID-19-positive and negative patients. Given the critical importance of understanding the nasopharyngeal microbiome in SARS-CoV-2 infection and COVID-19, use of alternate sources, like VTM, is essential.

Methods

Sample Cohort, Collection, Extraction and Sequencing

We examined COVID-19-positive ($n=9$) and -negative ($n=10$) subjects recruited from Rush University

*Correspondence: robin_voigt@rush.edu

†Ali Keshavarzian and Robin M. Voigt co-senior authors.

¹ Rush Medical College, Rush Center for Integrated Microbiome and Chronobiology Research, Rush University Medical Center, 1725 W. Harrison St. STE 206, Chicago, IL 60612, USA

Full list of author information is available at the end of the article



© The Author(s) 2021. **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 Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Medical Center (RUMC) located in Chicago, IL in the early stage of the pandemic (April 2020). Based on the Centers for Disease Control and Prevention (CDC) classification, the nine COVID-19-positive subjects had mild COVID-19 with no hospitalizations or deaths reported (See Supplementary Table 1, [Additional File 1](#)). The study was approved by the RUMC Institutional Review Board [COVID-19 Biorepository (ORA #20,032,309)] to use random de-identified remnant specimens from standard care testing of COVID-19. There were no a priori exclusion criteria.

The feasibility of using nasopharyngeal swab (NPS) VTM collection for SARS-CoV-2 detection and microbiome analysis has been previously reported [9]. The NPS were collected according to CDC guidelines. Sterile synthetic-head, plastic-shaft swabs were used to collect specimens for diagnostic testing and were placed into properly labeled collection tubes containing 3 mL of VTM [REMELE Micro Test™ M4RT®]. The NPS were sent to the RUMC clinical microbiology laboratory and heat inactivated [65°C for (30) minutes] prior to RT-PCR testing on an Abbott *m*2000 device [Abbott Laboratories] [10]. Due to the surge in testing and specimen processing backlog in the early stage of the pandemic the NPS tubes were stored at 4°C for up to nine days prior to processing for storage. The NPS + VTM tubes were vortexed briefly for (3) seconds and the VTM was aliquoted into multiple 250 µl aliquots, and stored at -80°C.

Isolation of viral nucleic acids and bacterial DNA from nasal swab VTM (200ul) samples were performed using the NucleoMag Pathogen manufacturer's protocol (Macherey–Nagel, Duran, Germany). Microbiome characterization was performed using a PCR-next-generation sequencing (NGS) approach with a two-stage PCR protocol, as described previously [11]. The V4 variable region of microbial 16S rRNA genes was amplified with the 515F/806R primer set (515F:GTGYCAGCMGCCGCGGTAA; 806R:GGACTACNVGGGTWTCTAAT) [12] and using Fluidigm Access Array primers for Illumina sequencers. Negative controls (*i.e.*, PCR reagent blanks; $n=5$) were amplified and sequenced with samples. Sequencing was performed using an Illumina MiniSeq with a mid-output kit and paired-end 153 base reads.

Analysis of Microbial 16S rRNA Gene Amplicon Sequences

Raw sequences obtained from the sequencer were merged using the PEAR (Paired-End read merger) algorithm (v0.9.11) [13]. Merged sequences were then quality filtered and denoised using the DADA2 algorithm within the QIIME2 (v 2020.8.0) workflow [14, 15]. Amplicon sequence variants (ASVs) were generated and utilized for all downstream analyses. Taxonomy was assigned to ASVs by using the naïve Bayes taxonomy classifier

trained with the SILVA 138 99% OTU database [16, 17]. A total of 527,628 sequencing clusters were generated, with an average depth of 27,770 sequences per sample (median = 15,318; min = 5,164; max = 145,270). Four reagent contaminant ASVs were identified and removed using decontam package based on the prevalence of the ASVs in the reagent negative blank controls, using default parameters [18]. Unassigned ASVs and chloroplast and mitochondrial ASVs were removed from statistical analyses [19].

Statistical Analysis

Analyses of alpha- and beta-diversity were used to compare nasopharyngeal microbial communities in COVID-19-positive and -negative patients; all analyses were performed on feature (ASV) counts. Alpha-diversity metrics were calculated on rarefied data (5,000 sequences/sample). Differences in alpha-diversity indices and bacterial ratios were assessed for significance using unpaired t-test or Mann–Whitney test, based on the outcome of Shapiro–Wilks normality test. Significance levels were set at $p < 0.05$. Permutation Multivariate Analysis of Variance (PERMANOVA) was used to assess global differences in microbial communities between groups [20]. Significance of PERMANOVA values were determined using 9,999 permutations and adjustment for multiple testing was conducted using the Benjamini–Hochberg false-discovery rate correction. Visualization of data was performed using principal coordinates analysis (PCoA) based on a Bray–Curtis dissimilarity distance matrix within the software package QIIME2.

Random forest models (number of runs = 1,023) were used to predict featured taxa of importance using the R version of the Boruta algorithm [21]. Analysis of composition of microbiomes (ANCOM) was performed on microbial communities to identify differentially abundant taxa between groups [22]. Additionally, differential abundances of individual taxa between groups were determined using differential abundance analysis (DESeq2) generating a q -value [23], as DESeq2 has increased sensitivity on smaller datasets (< 20 samples).

Results

Analysis of communities using 16S ribosomal RNA gene amplicon sequencing revealed that microbial alpha-diversity indices were not significantly different between patient groups, but taxonomic feature richness was lower in COVID-19-positive compared to -negative patients (See Supplementary Table 2, [Additional File 1](#)). Significant differences (PERMANOVA: $q = 0.016$) in nasopharyngeal microbial community structure were observed between positive and negative patients in beta diversity analyses conducted on microbial features

(i.e., amplicon sequence variants, ASVs) (Fig. 1A; See Supplementary Table 3, [Additional File 1](#)). Microbial communities in both groups were dominated by the phyla Proteobacteria, Actinobacteria and Firmicutes, and the genera *Corynebacterium*, *Morganella*, *Moraxella*, *Escherichia-Shigella*, *Proteus*, and *Staphylococcus* (> 50% of all sequences; Fig. 1B-C; See Supplementary Table 4, [Additional File 1](#)). The microbial community in COVID-19-positive patients can be characterized as pro-inflammatory as exemplified by significantly higher (Mann–Whitney test: $p=0.0002$) Proteobacteria-to-Actinobacteria ratio relative to COVID-19-negative patients (Fig. 1D). Between-group differences in taxon abundances were identified using ANCOM (Table 1) and DESeq2 (See Supplementary Tables 5–6, [Additional File 1](#)). ANCOM identified the genus *Streptococcus* as significantly ($q < 0.05$) less abundant in COVID-19-positive relative to -negative patients, with nine other genera trending ($p < 0.05$) towards significance (Table 1; Fig. 1E). DESeq2 analysis identified 25 significantly differentially abundant genera, including lower abundances of *Rothia* and *Prevotella* in COVID-19-positive patients (See Supplementary Table 6, [Additional File 1](#)). A machine learning approach (Boruta) was employed for feature selection to identify taxa driving differences between COVID-19-positive and -negative patients. This analysis identified nine genera, including *Anaerococcus*, *Streptococcus*, *Enterococcus*, and *Bacillus*, that were also identified in DESeq2 and ANCOM analyses (See Supplementary Fig. 1, [Additional File 2](#)).

Discussion

This proof-of-concept study provides evidence that current biorepositories with nasopharyngeal VTM can be utilized to assess microbiome communities. Importantly, this study is the first to show that COVID-19-positive patients have a dysbiotic nasopharyngeal microbial community characterized by loss of putative nasal commensal bacteria and an increase in putative pro-inflammatory bacteria. Specifically, this study revealed a characteristic nasal microbiome in COVID-19-positive individuals with (1) a trend toward reduced microbial alpha diversity, (2) an alteration in the relative abundance of multiple taxa within the phylum Proteobacteria and a significantly increased ratio of Proteobacteria:Actinobacteria, and (3)

decreased relative abundance of nasal commensal organisms such as *Corynebacterium* (Actinobacteria) [24] and *Streptococcus* (Firmicutes), previously shown to be affected by influenza virus infection [25]. These results need to be further evaluated in future studies using metagenomic and metatranscriptomic techniques, but may reflect baseline differences prior to infection leading to susceptibility to SARS-CoV-2 infection or may be a result of SARS-CoV-2 infection, leading to post-infection microbiome alteration.

Recent cross-sectional studies have attempted to distinguish nasal microbial profiles collected from COVID-19-positive and negative subjects using nasopharyngeal swab [5–8]. Of the four prior studies using nasopharyngeal swabs, our results are most similar to those of *Mostafa* et al. [7]. *Mostafa* et al., using non-targeted shotgun sequencing approaches, found significant decreases in microbial alpha diversity, significantly different in microbial community structures, and a higher relative abundance of *Propionibacteriaceae* and lower *Corynebacterium accolens* in COVID-19-positive relative to -negative individuals. In comparison, VTM data in this study indicated a similarly reduced (though not significant) microbial diversity, a significant difference between microbial communities, and a markedly decreased abundance of genera *Corynebacterium* in COVID-19-positive relative to -negative patients. Due to mismatches between the 16S rRNA gene sequence of bacteria from the genus *Propionibacterium* and the V4 primers employed in this study (e.g., *Meisel* et al. [26]), *Propionibacterium* were not well represented and no direct comparison with the results of *Mostafa* et al. could be made. Elsewhere, *Nardelli* and colleagues targeted the V1-V3 variable regions of 16S rRNA genes using amplicon sequencing [8], while *De Maio* et al. targeted the V5-V6 variable regions [5]. Differences in PCR conditions can make direct comparison of results generated with different primer sets challenging, as PCR conditions can introduce substantial bias into observed microbial communities. Finally, *Braun* et al. performed a cross-sectional 16S rRNA sequencing study, with longitudinal sampling from a subset of subjects, and targeting the V4 variable region [6], as performed in this study. However, *Braun* et al. reported no significant effect of SARS-CoV-2 on the nasopharyngeal microbial community using

(See figure on next page.)

Fig. 1 Microbial profiles of the COVID-19-positive and -negative patients. **A** Principal coordinate analysis (PCoA) of nasopharyngeal microbial communities in patients with and without SARS-CoV-2. Communities were significantly different between groups, as assessed by PERMANOVA ($q=0.016$). **B** Stacked column plots depicting the average relative abundance of bacterial phyla in COVID-19-positive and -negative patients. **C** Stacked column plots depicting the average relative abundance of bacterial genera in COVID-19-positive and -negative patients. **D** Ratio of Proteobacteria-to-Actinobacteria in COVID-19-positive and -negative patients (Mann–Whitney test; $p=0.0002$). **E** Relative abundance and standard deviation of nasopharyngeal genera identified in ANCOM analyses in COVID-19-positive and -negative patients

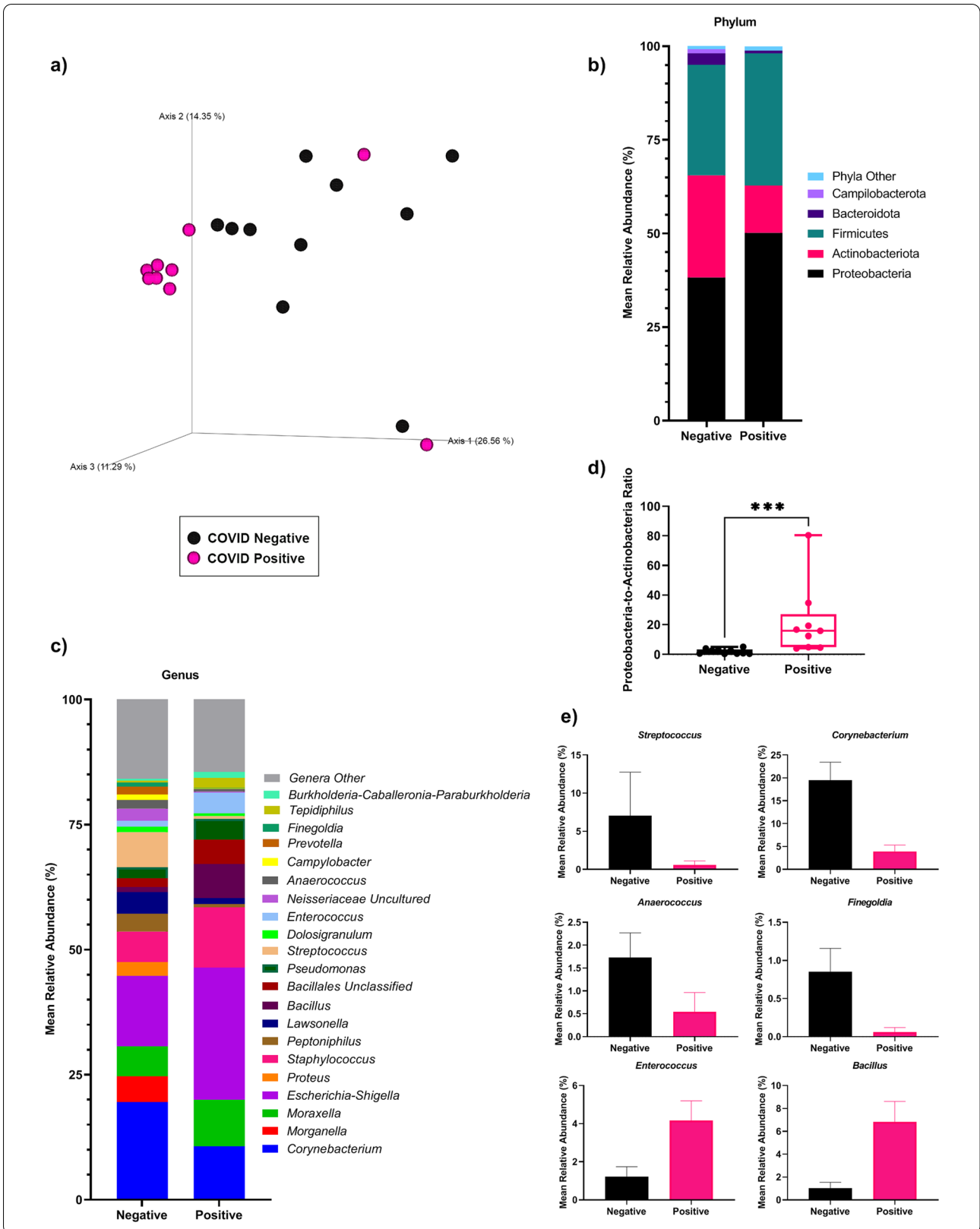


Table 1 Results of analysis of compositions (ANCOM) between COVID-19-positive and -negative nasopharyngeal samples

(Phylum) Genus	Median Percentile Abundance: COVID-19-Negative	Median Percentile Abundance: COVID-19-Positive	W Score
(Firmicutes) <i>Streptococcus</i>	213.92	1.0	19
Genera trending towards significance			
(Firmicutes) <i>Enterococcus</i>	37.25	292	8
(Proteobacteria) <i>Burkholderia-Caballeronia-Paraburkholderia</i>	1.0	91.33	8
(Proteobacteria) <i>Gulbenkiania</i>	1.0	33.67	6
(Firmicutes) <i>Finegoldia</i>	121.83	1.0	5
(Firmicutes) <i>Anaerococcus</i>	173.42	1.0	4
(Proteobacteria) <i>Neisseriaceae Uncultured</i>	371.42	1.0	4
(Firmicutes) <i>Bacillus</i>	47.58	624.67	4
(Bacteroidetes) <i>Prevotella</i>	70.58	1.0	3
(Actinobacteria) <i>Corynebacterium</i>	3,434.08	236.67	3

ANCOM: samples with fewer than 5,000 sequences were removed; features that were not present in at least 4 samples were removed; features with an overall count of less than 500 [10% of feature biom – due to small sample size ≤ 10 per group] were also removed. All listed significant features [genus] rejected the null hypothesis. Median percentile abundance (average of the 25th, 50th, 75th percentiles) and analysis W scores determined using ANCOM. ANCOM values were corrected for multiple testing using the Benjamini–Hochberg method (q -value < 0.05; bold). Taxa also identified as significant in DESeq2 analyses are bolded

multiple analysis methods. This proof-of-concept study using VTM revealed significant differences microbiome between COVID-19-positive and COVID-19-negative samples in this study. We note, however, that future analyses will need to also determine the genome sequence of the SARS-CoV-2 virus, as different variants may differentially alter microbial community structure. At the time of sampling for this study, currently circulating variants of concern had not yet arisen.

Limitations include the relatively small sample size that may have constrained our ability to identify additional significant difference between groups. We acknowledge a lack of inclusion of in-depth clinical medical information, as only basic demographic data was collected from patients at the COVID-19 rapid COVID-19 testing sites (i.e., drive-thru or patient drop-off) from where these samples were obtained. Additionally, internal household spousal or live-in close members as COVID-19 negative controls would be an ideal comparison cohort to help exclude environmental effects. We suggest that future research studies should attempt to include these data.

In conclusion, this proof-of concept study demonstrates the feasibility of using VTM from nasopharyngeal swab collections to study nasopharyngeal microbiota in COVID-19. As hospitals have collected and stored COVID-19 testing swabs and VTM over the entirety of the pandemic, this opens a huge potential dataset for interrogating the nasal microbiome. Secondly, this proof-of-concept study provides preliminary data suggesting the presence of a dysbiotic and pro-inflammatory nasopharyngeal microbiota in

COVID-19-positive patients. Our study provides a strong scientific rationale for future studies to investigate the relationship between nasal microbiome and SARS-CoV-2 infection and COVID-19 severity, and also the relationship to the long-lasting effects of COVID-19.

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12575-021-00148-6>.

Additional file 1: Supplementary Table 1. Demographics and clinical characteristics of COVID-19-positive individuals. **Supplementary Table 2.** Alpha-diversity values in nasopharyngeal samples. **Supplementary Table 3.** Global community analysis of nasopharyngeal microbial community structure in COVID-19-positive and -negative patients, as assessed by Permutational Multivariate Analyses of Variance (PERMANOVA). **Supplementary Table 4.** Relative abundances of bacterial taxa in COVID-19-positive and -negative nasopharyngeal samples. **Supplementary Table 5.** Differential relative abundance of nasopharyngeal bacterial phyla in COVID-19-positive and -negative samples identified by DESeq2 analysis. **Supplementary Table 6.** Genus taxonomic level differential abundance DeSeq2 analysis between COVID-19-positive and -negative patient's nasopharyngeal samples.

Additional file 2: Supplementary Figure 1. Microbial features driving differentiation of COVID-19-positive and -negative patients. A predictive model based on the genus-level relative abundance was generated using Boruta. Green boxes are bacterial genera that are strongly associated with differentiating the groups using the Boruta feature selection algorithm. Blue boxes are the shadow genera introduced into Random Forest classifier to act as benchmarks. Red boxes are bacterial genera that were not associated with differentiating the groups.

Acknowledgements

We would like to thank the RUMC COVID-19 Biorepository investigators Drs. Mary Hayden and Nagarjuna Tippireddy for providing clinical data.

Authors' contributions

A.K. and R.M.V. conceived and designed the experiments and contributed to writing the manuscript. P.A.E. wrote the manuscript, performed the experiments, and analyzed the data. A.N., S.J.G., and C.J. analyzed the data and contributed to writing the manuscript. A.L. designed the RUMC COVID-19 repository. A.L. and C.J. provided the biological samples for experimentation. All authors reviewed the manuscript. The author(s) read and approved the final manuscript.

Funding

None to report.

Availability of data and materials

Raw sequence data (FASTQ files) were deposited in the National Center for Biotechnology Information (NCBI) Sequence Read Archive (SRA), under the BioProject identifier PRJNA704967.

Declarations

Ethics approval and consent to participate

The study was approved by Rush University Medical Center Institutional Review Board [COVID-19 Biorepository (ORA #20032309)].

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Rush Medical College, Rush Center for Integrated Microbiome and Chronobiology Research, Rush University Medical Center, 1725 W. Harrison St. STE 206, Chicago, IL 60612, USA. ²Department of Molecular Pathogens and Immunity, Rush University Medical Center, Chicago, IL, USA. ³Genomics and Microbiome Core Facility, Rush University Medical Center, Chicago, IL, USA. ⁴Department of Internal Medicine, Division of Geriatrics and Palliative Medicine, Rush Medical College, Chicago, IL, USA. ⁵Department of Physiology, Rush University Medical Center, Chicago, IL, USA. ⁶Department of Medicine, Rush University Medical Center, IL, Chicago, USA.

Received: 16 March 2021 Accepted: 17 May 2021

Published online: 01 June 2021

References

- Chan JF, Yuan S, Kok KH, To KK, Chu H, Yang J, Xing F, Liu J, Yip CC, Poon RW, et al. A familial cluster of pneumonia associated with the 2019 novel coronavirus indicating person-to-person transmission: a study of a family cluster. *Lancet*. 2020;395(10223):514–23.
- Li N, Ma WT, Pang M, Fan QL, Hua JL. The Commensal Microbiota and Viral Infection: A Comprehensive Review. *Front Immunol*. 2019;10:1551.
- Belkaid Y, Hand TW. Role of the microbiota in immunity and inflammation. *Cell*. 2014;157(1):121–41.
- Man WH, de Steenhuijsen Pitsers WA, Bogaert D. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat Rev Microbiol*. 2017;15(5):259–70.
- De Maio F, Posteraro B, Ponziani FR, Cattani P, Gasbarrini A, Sanguinetti M. Nasopharyngeal Microbiota Profiling of SARS-CoV-2 Infected Patients. *Biological procedures online*. 2020;22:18.
- Braun T, Halevi S, Hadar R, Efroni G, Glick Saar E, Keller N, Amir A, Amit S, Haberman Y. SARS-CoV-2 does not have a strong effect on the nasopharyngeal microbial composition. *Sci Rep*. 2021;11(1):8922.
- Mostafa HH, Fissel JA, Fanelli B, Bergman Y, Gnizdowski V, Dadlani M, Carroll KC, Colwell RR, Simmer PJ. Metagenomic Next-Generation Sequencing of Nasopharyngeal Specimens Collected from Confirmed and Suspect COVID-19 Patients. *mBio* 2020;11(6):e01969-20.
- Nardelli C, Gentile I, Setaro M, Di Domenico C, Pinchera B, Buonomo AR, Zappulo E, Scotto R, Scaglione GL, Castaldo G, et al. Nasopharyngeal Microbiome Signature in COVID-19 Positive Patients: Can We Definitively Get a Role to *Fusobacterium periodonticum*? *Front Cell Infect Microbiol*. 2021;11:625581.
- Minich J, Ali F, Marotz C, Belda-Ferre P, Chiang L, Shaffer JP, Carpenter CS, McDonald D, Gilbert J, Allard SM, et al. Feasibility of using alternative swabs and storage solutions for paired SARS-CoV-2 detection and microbiome analysis in the hospital environment. *Res Square*. 2020;19:128.
- Moore NM, Li H, Schejbal D, Lindsley J, Hayden MK: Comparison of Two Commercial Molecular Tests and a Laboratory-Developed Modification of the CDC 2019-nCoV Reverse Transcriptase PCR Assay for the Detection of SARS-CoV-2. *J Clin Microbiol*. 2020;58(8):e00938-20.
- Naqib A, Poggi S, Wang W, Hyde M, Kunstman K, Green SJ. Making and Sequencing Heavily Multiplexed, High-Throughput 16S Ribosomal RNA Gene Amplicon Libraries Using a Flexible Two-Stage PCR Protocol. *Methods Mol Biol*. 2018;1783:149–69.
- Caporaso JG, Lauber CL, Walters WA, Berg-lyons D, Huntley J, Fierer N, Owens SM, Betley J, Fraser L, Bauer M, et al. Ultra-high-throughput microbial community analysis on the Illumina HiSeq and MiSeq platforms. *ISME J*. 2012;6(8):1621–4.
- Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. *Bioinformatics*. 2014;30(5):614–20.
- 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(7):581–3.
- Estaki M, Jiang L, Bokulich NA, McDonald D, Gonzalez A, Kosciolk T, Martino C, Zhu Q, Birmingham A, Vazquez-Baeza Y, et al. QIIME 2 Enables Comprehensive End-to-End Analysis of Diverse Microbiome Data and Comparative Studies with Publicly Available Data. *Curr Protoc Bioinformatics*. 2020;70(1):e100.
- Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Gregory Caporaso J. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2's q2-feature-classifier plugin. *Microbiome*. 2018;6(1):90.
- Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glockner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41(Database issue):D590-596.
- 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(1):226.
- Hanshew AS, Mason CJ, Raffa KF, Currie CR. Minimization of chloroplast contamination in 16S rRNA gene pyrosequencing of insect herbivore bacterial communities. *J Microbiol Methods*. 2013;95(2):149-55.
- Kelly BJ, Gross R, Bittinger K, Sherrill-Mix S, Lewis JD, Collman RG, Bushman FD, Li H. Power and sample-size estimation for microbiome studies using pairwise distances and PERMANOVA. *Bioinformatics*. 2015;31(15):2461–8.
- Rudnicki MKaW. Feature Selection with the Boruta Package. *J Stat Softw*. 2010;36(1):1–13.
- Mandal S, Van Treuren W, White RA, Eggesbo M, Knight R, Peddada SD. Analysis of composition of microbiomes: a novel method for studying microbial composition. *Microb Ecol Health Dis*. 2015;26:27663.
- Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol*. 2014;15(12):550.
- Hang J, Zavaljevski N, Yang Y, Desai V, Ruck RC, Macareo LR, Jarman RG, Reifman J, Kuschner RA, Keiser PB. Composition and variation of respiratory microbiota in healthy military personnel. *PLoS ONE*. 2017;12(12):e0188461.
- Ding T, Song T, Zhou B, Geber A, Ma Y, Zhang L, Volk M, Kapadia SN, Jenkins SG, Salvatore M et al: Microbial Composition of the Human Nasopharynx Varies According to Influenza Virus Type and Vaccination Status. *mBio* 2019;10(4):e01296-19.
- Meisel JS, Hannigan GD, Tyldsley AS, SanMiguel AJ, Hodkinson BP, Zheng Q, Grice EA. Skin Microbiome Surveys Are Strongly Influenced by Experimental Design. *J Invest Dermatol*. 2016;136(5):947–56.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.