Acute SARS-CoV-2 infection is associated with an expansion of bacteria pathogens in the nose including *Pseudomonas Aeruginosa*

Nicholas S. Rhoades¹, Amanda Pinski¹, Alisha N. Monsibais¹, Allen Jankeel¹, Brianna M. Doratt¹, Isaac R. Cinco¹, Izabela Ibraim¹ Ilhem Messaoudi^{1*}

¹Department of Molecular Biology and Biochemistry, University of California Irvine, Irvine, CA, USA

* To whom correspondence should be addressed

Ilhem Messaoudi, PhD

Molecular Biology and Biochemistry

University of California Irvine

2400 Biological Sciences III

Irvine, CA 92697

Phone: 949-824-3078

Email: imessaou@uci.edu

Keywords

SARS-CoV-2, nasal microbiome, inflammation, *Pseudomonas Aeruginosa*

This PDF file includes:

Main Text

Figures 1 to 5

Supplemental figures 1 to 1

ABSTRACT

Much of the research conducted on SARS-CoV-2 and COVID-19 has focused on the systemic host response, especially that generated by severely ill patients. Very few studies have investigated the impact of acute SARS-CoV-2 within the nasopharynx, the site of initial infection and viral replication. In this study we profiled changes in the nasal microbial communities as well as in host transcriptional profile during acute SARS-CoV-2 infection using 16S amplicon sequencing and RNA sequencing. These analyses were coupled to viral genome sequencing. Our microbiome analysis revealed that the nasal microbiome of COVID patients was unique and was marked by an expansion of bacterial pathogens. Some of these microbes (i.e. *Acinetobacter*) were shared with COVID negative health care providers from the same medical center but absent in COVID negative outpatients seeking care at the same institutions suggesting acquisition of nosocomial respiratory pathogens. Specifically, we report a distinct increase in the prevalence and abundance of the pathogen *Pseudomonas aeruginosa* in COVID patients that correlated with viral RNA load. These data suggest that the inflammatory environment caused by SARS-CoV-2 infection and potentially exposure to the hospital environment leads to an expansion of bacterial pathogens in the nasal cavity that could contribute to increased incidence of secondary bacterial infections. Additionally, we observed a robust host transcriptional response in the nasal epithelia of COVID patients, indicative of an antiviral innate immune response and neuronal damage. Finally, analysis of viral genomes did not reveal an association between viral loads and viral sequences.

INTRODUCTION

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) is the causative agent of coronavirus disease 2019 (COVID-19) and the cause of the ongoing global pandemic(1, 2). COVID-19 is a respiratory disease ranging in clinical presentation from an asymptomatic (~80% of cases) to a fatal infection (~1-2%) accompanied by cytokine storm, acute respiratory distress syndrome and coagulopathy (3, 4). Long-term consequences of COVID-19 range from anosmia and ageusia(3, 5, 6) to respiratory (e.g., dyspnea) and nervous system (e.g., headache, anosmia) complications (7). Numerous studies have identified advanced age, comorbidities (notably, obesity and diabetes), and biological sex as risk factors for severe COVID-19 (4, 5, 8, 9). Other factors modulating disease, including mucosal microbial defenses and viral burden at the site of pathogenesis have been under-explored.

SARS-CoV-2 is primarily spread through the inhalation of virus-laden respiratory droplets(5). Initial replication occurs in the upper respiratory tract (URT) following the interaction of the viral spike protein and host ACE2, and subsequent entry into cells (10-12). The nasal cavity is believed to be the initial site of viral replication rather than the oral cavity due to higher expression of ACE2 (13-15) and greater levels of viral RNA (vRNA) in this site compared to matched oral specimens (16). Rapid and uncontrolled replication can lead to infection of the lower respiratory tract (LRT) and severe disease. However, several studies have demonstrated a lack of correlation between disease severity and viral RNA load in nasal cavities, suggesting that other factors contribute to disease at the site of infection (17-19).

Recent research has demonstrated an increasingly critical role of the mucosal microbiome in modulating viral infection (20-23). Disruption of microbial communities by viral infection can exacerbate inflammation, facilitate coinfection, and deregulate the adaptive immune response (20-23). Furthermore, severe inflammation in olfactory epithelium can lead to short- and long-term symptoms like anosmia and ageusia(24, 25). However, while numerous studies have investigated the longitudinal dynamics of COVID-19 in blood, a limited number of studies have examined the respiratory microbiome in acute patients(26-31). Studies of nasopharyngeal samples are limited

to small cohorts of mild and severe COVID-19 patients and show variable differences in diversity and increased abundance of select phyla without accounting for the potential impact of viral burden(32-35). Other studies of bronchoalveolar lavage and pharyngeal samples in severe patients have identified a general dysbiosis of the lower respiratory microbiome similar to that of pneumonia patients(36, 37).

Studies aimed at understanding the impact of COVID-19 and viral burden on nasal microbial communities is clearly lacking but urgently needed to characterize both acute infection and provide insight into long-term symptoms of COVID-19. In this study, we determined the impact of acute SARS-CoV-2 infection with a spectrum of viral RNA loads on nasal microbial communities, local host transcriptional responses and potential association with viral genome sequence diversity. Our cross-sectional study of hospitalized COVID-19 patients, and uninfected outpatients and healthcare workers indicated a distinct shift in the nasal microbiome of COVID-19 patients featuring the expansion of pathobionts such as *Rothia*, *Acinetobacter*, and *Pseudomonas*. This was accompanied by the upregulation of host antiviral immune genes and the downregulation of genes with critical roles in mucosal and neuronal cell homeostasis. However, no association between viral loads and viral genome sequences was detected. Overall, our data suggests SARS-CoV-2 infection is associated with an expansion of bacterial pathogens in the nasal cavity and virus-induced host dysregulation at the site of infection.

METHODS

Sample collection and extraction

RNA was extracted from remnant de-identified nasal swab samples collected in viral transport medium (VTM) from patients and healthcare workers (HCW) at the University of California, Irvine Medical Center (UCIMC) using the Quick-RNA™ Viral 96 Kit (Zymo, Cat#R1041). SARS-CoV-2 viral loads were determined using qPCR with primers specific to the nucleoprotein (NP). A one-step reaction was prepared using 5ul of extracted RNA or standard, 500nM of forward (5'-GGGGAAGTCTCTCCTGCTAGAAT-3') and reverse (5'-CAGACATTTTGCTCTCAAGCTG-3') SARS-CoV-2-nucleocapsid primers, 125 nM of SARS2-nucleocapsid probe (5'-FAM-

TTGCTGCTGCTTGACAGATT-BHQ1-3'), and TaqPath™ 1-Step RT-qPCR Master Mix (Applied Biosystems, Foster City, CA, USA). PCR cycling conditions were 2 min at 95 °C, 15 min at 50 °C, denaturation for 2 min at 95 °C followed by 45 cycles of 3 s at 95 °C and 30 s 60 °C on the StepOnePlus™ Real-Time PCR System (Applied Biosystems, Foster City, CA, USA). All samples were run in triplicate. Viral load concentration for the target samples was determined by plotting the average Ct values against the log concentration of the standards in the dilution series. Positivity was determined as a cycle threshold (Ct) value less than or equal to 40 for the RBD gene. Samples were categorized into three groups: healthcare workers (HCW; n=45), patients testing negative for SARS-CoV-2 (CoV-; n=21), and patients who tested positive for SARS-CoV-2 (CoV+; n=68). The latter group was further sub-divided into three groups based on Ct values: High Ct value (>34.6; n=23), Mid Ct value (32.9-25.0; n=23), and Low Ct values (<23.5; n=22).

16S amplicon libraries construction and data analysis

DNA that co-eluted with extracted RNA was used as the template to amplify the hypervariable V4 region of the 16S rRNA gene using PCR primers (515F/806R with the forward primer containing a 12-bp barcode) in duplicate reactions containing: 12.5 ul GoTaq master mix, 9.5 ul nuclease-free H₂O, 1 ul template DNA, and 1 ul 10uM primer mix. Thermal cycling parameters were 94°C for 3 minutes; 35 cycles of 94°C for 45 seconds, 50°C for 1 minute, and 72°C for 1 minute and 30 seconds; followed by 72°C for 10 minutes. PCR products were purified using a MinElute 96 UF PCR Purification Kit (Qiagen, Valencia, CA, USA). Libraries were sequenced (2 x 300 bases) using Illumina MiSeq.

Raw FASTQ 16S rRNA gene amplicon sequences were uploaded and processed using the QIIME 2 version 2019.10 (38) analysis pipeline as we have previously described (39). Briefly, sequences were demultiplexed and quality filtered using the DADA2 plugin for QIIME 2 (40), which filters chimeric sequences. The generated sequence variants were then aligned using MAFFT (41), and a phylogenetic tree was constructed using FastTree 2 (42). Taxonomy was assigned to sequence variants using q2-feature-classifier against the SILVA Database (release:

(43). After taxonomic classification all sequences not assigned to a known phyla were removed along with any sequence assigned to mitochondria or chloroplast. After removal of low quality and contaminate sequences taxonomy was reassigned to sequence variants using q2-feature-classifier against the expanded Human Oral Microbiome Database (eHOMD: release version 15.21) (44). This curated and site specific database was used to improve species level resolution. To prevent sequencing depth bias, samples were rarified to 8,000 sequences per sample before α and β diversity analysis. QIIME 2 was also used to generate the following α diversity metrics: richness (as observed taxonomic units), Shannon evenness, and phylogenetic diversity. β diversity was estimated in QIIME 2 using weighted and unweighted UniFrac distances (45).

All statistical analyses were conducted using PRISM (V8) or the R package Vegan (46). PERMANOVAs were performed using the Vegan function ADONIS. 1-way, non-parametric Kruskal-Wallis ANOVA were implemented using PRISM to generate p-values and utilizing the corresponding post-hoc-test when the initial ANOVA was significant. The LEfSe algorithm was used to identify differentially abundant taxa and pathways between groups with a logarithmic Linear discriminant analysis (LDA) score cutoff of 2 (47).

RNAseq library prep and analysis

Quantity and quality of RNA extracted from 8 nasal swabs of (4 CoV+ and 4 HCW) was determined using an Agilent 2100 BioAnalyzer. cDNA libraries were constructed using the NEB Next Ultra II Directional RNA Library kit (Thermo Fischer). Briefly, RNA was treated with RNAseH and DNase I after depletion of ribosomal rRNA. Adapters were ligated to cDNA products. The ~300 bp amplicons were PCR-amplified and indexed with a unique molecular identifier. cDNA libraries were assessed for quality and quantity on the BioAnalyzer prior to single-end sequencing (x100bp) using the Illumina NovaSeq platform.

Bioinformatic analysis was performed using the systemPipeR RNA-Seq workflow (48). RNA-Seq reads were demultiplexed, quality-filtered and trimmed for quality and adapter removal

using Trim Galore (average Phred Score cut-off = 30, minimum length of 50 basepairs). Tophat was used to align filtered reads to the reference genome *Homo sapiens* (GRCh38). The file was used for annotation. Raw expression values (gene-level read counts) were generated using the summarizeOverlaps function and normalized (read per kilobase of transcript per million mapped reads, rpkm) using the edgeR package. Statistical analysis with edgeR was used to determine differentially expressed genes (DEGs) meeting the following criteria: genes with median rpkm of ≥ 1 , a false discovery rate (FDR) corrected p-value ≤ 0.05 and a $\log_2$ fold change ≥ 1 compared to CoV- samples. Functional enrichment of DEGs was performed using Metascape to identify relevant Gene Ontology (GO) biological process terms (49). Heatmaps, bar graphs and volcano plots were generated using R package ggplot2.

Viral genome library construction and analysis

Enrichment for viral RNA was performed using the QIASeq™ SARS-CoV-2 Primer Panel V2 (Qiagen) Panel V2) followed by library construction with the Qiagen FX DNA library preparation kit. Samples were indexed, pooled and validated with 2100 Agilent BioAn. Prior to Illumina sequencing (2x100bp, 1 M reads).

Adapters and primers were removed from demultiplexed sequencing readings using TrimGalore and MaskPrimers.py (pRESTO). Merged reads were aligned to SARS-CoV-2 Wuhan isolate NC_045512.2 with BWA-mem software version 0.7.17. Genomes with greater than 90% coverage and 10X depth were retained for phylogenetic analysis with Nextstrain. Amino acid changes and clade assignments were identified using Nextclade. A set of comparator sequences representing the original Wuhan isolate (NC_045512.2), variants of concern (B.1.351, EPI_ISL_960123; P.1, EPI_ISL_833167; B.1.1.7, EPI_ISL_659057), isolates from January 2020 New York (EPI_ISL_1293138) and Washington (EPI_ISL_404895), and representative isolates from California January-November 2020 (EPI_ISL_406036, EPI_ISL_411954, EPI_ISL_429875, EPI_ISL_436642, EPI_ISL_444023, EPI_ISL_569672, EPI_ISL_548382, EPI_ISL_548612, EPI_ISL_582897, EPI_ISL_625601, EPI_ISL_753324) were used for phylogenetic analysis. All

sequences from Orange County, California from January 2020 to December 2020 were also retrieved from GISAID.

RESULTS

SARS-CoV-2 infection is associated with expansion of pathobionts in the nasal microbiome

To determine if SARS-CoV-2 infection is associated with a shift in the composition of the nasal microbiome, we utilized 16S rRNA gene amplicon sequencing to profile nasal samples from hospitalized COVID-19 patients (CoV+ inpatient), and SARS-CoV-2 negative out-patients (CoV- outpatient) and healthcare workers (CoV- HCW) (**Figure 1A**). At the community level, the nasal microbiome of CoV+ patients was distinct from both CoV- HCW and CoV- individuals with host status accounting for 9.3% of the total variation of unweighted Unifrac distance (**Figure 1B**). Additionally, nasal microbial communities in CoV+ inpatients showed the highest intra-group variability followed by communities from CoV- HCW and CoV- outpatients (**Figure 1C**). These same trends held for abundance based weighted Unifrac distance (**Supp Figure 1A, B**). We also observed an increased richness in nasal community from COV+ patients and HCW as indicated by a higher number of observed Amplicon Sequencing Variants (ASVs) compared to Cov- outpatients (**Figure 1D**). As expected, all three groups had similarly low community Shannon diversity values, indicative of a low-complexity community dominated by a few microbes (**Figure 1E**). We next probed the taxonomic landscape of the nasal microbiome to identify microbial taxa that were driving the shift in the overall community. At the phyla level, samples from all groups were dominated by Actinobacteria, Firmicutes, and Proteobacteria (**Figure 1F, Supp. Figure 1C**). Common nasal genera such as *Corynebacterium*, *Staphylococcus*, *Streptococcus*, *Dososigranulum*, and *Neisseria* were highly abundant in all three groups (**Figure 1F, Supp Figure 1C**).

We next explored differences between the three groups. Analysis using LefSe revealed enrichment of specific genera in each group (**Figure 2A, Supp Table 1**). Specifically,

Anaerococcus was enriched in CoV- outpatients, while *Lawsonella* and *Burkholderia* were most abundant in CoV- HCW (**Figure 2A, Supp Table 1**). Finally, several pathogenic bacteria such as *Rothia*, *Acinetobacter*, and *Pseudomonas* were most abundant in COV+ in-patients (**Figure 2A, Supp Table 1**). To further explore the distribution of specific taxonomic groups, we plotted the abundance of select taxa across all three groups. The two most abundant genera, *Corynebacterium* and *Staphylococcus* were equally abundant in the groups (**Figure 2B,C**). Interestingly, relative abundance of *Anoxybacillus* was higher in both outpatients CoV- and CoV+ inpatients compared to HCW (**Figure 2D**). Conversely, *Burkholderia cepacia* was significantly more abundant in CoV- HCW (**Figure 1F, 2E**). On the other hand, the genera *Acinetobacter*, and *Pseudomonas* were most abundant in the COV+ inpatients (**Figure 2 F, G**). When probed at the species level, *Pseudomonas aeruginosa* was highly enriched in CoV+ inpatients (**Figure 2H**). Taken together, these data suggest SARS-CoV-2 infection is associated with a shift in the nasal microbiome and acquisition of pathobionts.

Associations between SARS-CoV-2 vRNA load and the nasal microbiome

We next sought to determine whether SARS-CoV-2 viral RNA (vRNA) load is associated with changes in the nasal microbiome. To this end, we divided CoV+ inpatients into three groups based on cycle threshold (Ct) values at admission: High Ct value (>34.6; n=23), Mid Ct value (32.9-25.0; n=23), and Low Ct values (<23.5; n=22) (**Supp Figure 1 D**). We compared the number of observed ASVs, and community composition based on unweighted and weighted Unifrac distance, as well as intra-group variability across the 3 Ct groups (**Supp Figure 1 E-I**). vRNA load did not impact the diversity of the nasal microbial community (**Supp Figure 1 E-G**) However, when the relative abundance of taxa was taken into account, viral loads explained 10% of the overall community composition (**Supp Figure 1H**, PERMANOVA p value<0.001). Moreover, the nasal communities of patients with low and mid Ct groups were more similar to each other than either group was to the high Ct (**Supp Figure 1I**), suggesting that shifts in the nasal microbiome associated with SARS-CoV-2 infection may be affected by vRNA load.

We next assessed differentially abundant genera between the 3 Ct groups (**Figure 3**). LefSe analysis showed that nasal microbial communities in patients with low vRNA loads were had higher abundance of *Streptococcus*, while those of patients with moderate vRNA loads were denoted by abundance of *Corynebacterium*, and those of patients with high vRNA loads by *Cutibacterium*, *Neisseria*, and *Pseudomonas* (**Figure 3A**). The relative abundance of the genus *Streptococcus* was highest in the high CT group (**Figure 3B**), while that of *Corynebacterium* was more abundant in the Mid and Low CT groups (**Figure 3C**). Finally, *Neisseria* and *Pseudomonas* were more abundant in the low Ct group (**Figure 3D,E**) and *Pseudomonas areuginosa* was positively associated with viral loads (**Figure 3F**).

SAR-CoV-2 infection remodels the nasal epithelium transcriptome

We compared the nasal transcriptomes of CoV- HCW and CoV+ inpatients using RNA-Seq to understand changes in host responses to infection (**Figure 4**). We identified a total of 692 differentially expressed genes (DEGs) between the two groups (n=377 upregulated, n=315 downregulated) in the CoV+ group (**Figure 4A**). Functional enrichment showed that upregulated DEGs enriched to multiple Gene Ontology (GO) terms associated with innate and adaptive host defense pathways such as “response to virus”, and “inflammatory response”, “response to bacterium” and “lymphocyte activation” (**Figure 4B**). A large number of interferon-stimulated genes (ISGs; e.g., *HERC5*, *IFITM3*, *ISG15*, *OAS2*, *RSAD2*) and those related to type I interferon signaling (e.g., *IRF7*, *STAT1*) mapped to these GO terms “response to virus” (**Figure 4C**). Genes with roles in NFkB and JUN-ATP-1 signaling, such as *BCL3*, *JUN* and *MYD88*, mapped to GO terms “inflammatory response”. In terms of cell-mediated immunity, expression of genes involved in leukocyte chemotaxis (e.g., *CCL2*, *CCL4*, *CXCL9*, *CCR1*), hematopoiesis (e.g., *CD53*, *IKZF1*, *IKZF3*, *JAK3*) and B cell activation (e.g., *IGHA*, *LCK*, *ZAP70*) were upregulated in CoV+ patients. Interestingly, genes that encode inhibitory receptors (e.g. *CD300E* and *LAGLS9*) as well as genes involved in cell death (e.g. *BCL2*, *NUPR1* and *ZBP1*) were also highly upregulated. Additionally,

upregulated DEGs (e.g. *SNCA*, *EGR1* and *MDK*) belonged to GO terms associated with cell death, including “neuron death” (**Figure 4B**).

Downregulated DEGs mapped to GO terms related to maintenance of homeostasis (e.g., “response to wounding”, “), cellular organization (e.g., “microtubule-based movement”) and neuronal tissue processes (e.g., “sensory organ development”) (**Figure 4C**). Notable downregulated DEGs encoded for integrins (e.g., *ITGA2/3/V*, *ITGB1/6*) laminins (e.g., *LAMA2/B2/B3/C2*) and microtubules (e.g., *DNAH3/7/10/11*, *DYNC2H1*, *WDR19*) comprising intracellular and extracellular structures like cilia and cell-cell adhesion junctions. We also observed a downregulation of genes associated with mucin production in nasal passages (e.g., *MUC20*, *MUC5AC*). Of the ion channels downregulated (e.g., *ATP1A1*, *SLC39A*), several were involved in chloride ion balance and neuron homeostasis (e.g., *CLCN3*, *CLCA1*, *SLC1A1*, *SLC4A8*) (**Figure 4A, E**).

Viral genome sequences do not correlate with viral loads

To identify potential relationships between vRNA load and viral evolution, we assembled and characterized SARS-CoV-2 genomes from each CoV+ in-patient (**Figure 5**). Ct value was negatively associated with viral genome coverage, with 100%, 74% and 13% of genomes from Ct groups 10-20, 21-30 and 31-40 achieving greater than 90% genome coverage, respectively (**Figure 5A**). All genomes with adequate coverage harbored the spike D614G and nsp12 P4715L mutations (**Figure 5B**). One synonymous mutation in nsp3 (F924) was also present in all genomes. The nucleoprotein T205I mutation associated with the 501.V2 variant was found in one individual (Ct 31-40). Genomes from each Ct group belonged to three main clades 20A, 20B, and 20C with no distinct relation to current variants of concern (P.1, B.1.1.7, B.1.351, B.1.427, B.1.429) or partitioning according to Ct value (**Figure 5C**). Compared to all genomes collected during 2020 in the local county (Orange County, California from January to December 2020), very few sequences belonging to the 20H clade, which emerged in October, were detected (**Fig.S3**).

DISCUSSION

In this study, we explored connections between acute SARS-CoV-2 infection, the nasal microbiome, and the local host transcriptional response. While a wealth of studies have focused on the systemic host response to SARS-CoV-2 and its association with disease severity, few studies have investigated the impact of acute SARS-CoV-2 on the nasal epithelium. SARS-CoV-2 infection occurs primarily via respiratory droplets with the majority of initial viral replication taking place in the nasal epithelia (11, 14). The majority of individual who contract SARS-CoV-2 are able to clear infection within the upper respiratory tract resulting in mild-moderate disease (3). However, in some individuals, the initial response is not sufficient, and SARS-CoV-2 migrates into the lower respiratory tract leading to severe COVID19 often characterized by acute respiratory distress syndrome, secondary pneumonias, and cytokine storms (3, 4, 27, 50). There is a clear need to understand both the host and microbial factors in the nasal cavity during acute infection that may have implications for both viral clearance and disease severity.

The nasal microbiome of a healthy individual is dominated by *Corynebacterium*, *Staphylococcus*, *Streptococcus*, *Dolosigranulum*, and *Moraxella* (51-53). The composition of the nasal microbiome can impact host susceptibility and disease course following respiratory infection (54, 55). Additionally, acute viral infections have been shown to modulate bacterial communities potentially favoring the expansion of opportunistic pathogens (56, 57). In this study we show that nasal microbiome of CoV+ patients was enriched in pathogenic bacteria such *Acinetobacter*, *Moraxella* and *Pseudomonas aeruginosa*. An expansion of *Pseudomonas* was also observed following Influenza A infection and in throat samples collected from COVID-19 patients (58, 59). These data suggest that expansion of *Pseudomonas* is not SARS-CoV-2 specific but rather represents a response to the inflammatory responses associated with acute viral infections. Indeed, *Pseudomonas aeruginosa* has been shown to thrive in hyperinflammatory environments such as the cystic fibrosis lung (60, 61). Another potential explanation, is that changes in the mucosal environment such as accumulation of mucus favor the growth of this organism. Our findings are in line with increased risk of secondary bacterial infections in COVID-19 patients,

especially those who are severely ill despite high rates of antibiotic use (62-64). Patients with secondary infections have a poor prognosis, lower discharge rates, and higher mortality rates (65). Further longitudinal studies should focus on determining if these pathobionts cause secondary pneumonias associated with COVID-19. Our findings suggest that the composition of the nasal maybe provide a powerful biomarker of an individual's risk for secondary pneumonias in viral infections.

Hospital acquired infections (HAI) are a major public health concern. A recent multi-center study found that 46% of patients hospitalized with COVID-19 suffered from a HAI (50). HAI are often attributed to contamination associated with healthcare workers or hospital environments (66). One of the major reservoirs for opportunistic pathogens is the nasal cavity which can act as the entry point and contribute to the spread of bacterial infections to other sections of the respiratory tract (67). It has been well established that HCW are commonly carriers of pathogens associated with nosocomial infection such as MRSA given their extended exposure to the hospital environment (68). Our analysis found an enrichment of pathobionts such as *Escherichia*, *Klebsiella*, and *Burkholderia cepacia* in the nasal communities of HCW. *B. cepacia* is a highly prevalent nosocomial infection in patients with underlying conditions such as cystic fibrosis (69, 70) and the elderly (71). The relative abundance of *Acinetobacter* was significantly higher in both HCW and CoV+ patients suggesting potential transfer between these two groups. Further studies should explore the role of HCW in the spread and acquisition of these pathogens in the hospital environment beyond the known pattern of MRSA.

In this study we also profiled the host transcriptional profile of the nasal epithelium in a subset of patients. Our findings highlight an upregulation of genes associated with innate immune cell activation, antiviral defense, and inflammation as recently described (72, 73). We also observed a transcriptional signature associated with cell death which has been proposed as a key biosignature of COVID-19 infection (74, 75). Furthermore, the robust upregulation of genes associated with neuronal death and downregulation of genes influencing epithelial integrity and sensory organ development supports a mechanism for infection-induced anosmia and ageusia

(24, 25, 76). Finally, we did not detect overrepresentation of viral genomes to specific clades not well-established in California in late 2020 or partitioning by Ct value (77, 78). These findings suggest that viral genetic variation may have limited impact on viral replication in mucosal tissues, although studies utilizing larger cohorts are needed to confirm this hypothesis.

In summary, data presented in this manuscript show that SARS-CoV-2 infection lead to a distinct shift in the composition of the nasal microbiome including an expansion of known bacterial pathogens such as *Pseudomonas aeruginosa*. Additionally, HCW harbor nasal enriched in pathogens known to cause nosocomial infections such as *Acinetobacter* and *Burkholderia cepacia*. Our host transcriptional profiling illustrated a robust local immune response to SARS-CoV-2 infection and provided support for neuronal damage in the URT leading to anosmia. This study had some limitations. The samples used in this study were initially collected for SARS-CoV-2 diagnosis and were therefore not processed in the typical fashion for either host transcriptomics or microbiome analysis. Nevertheless, we were able to obtain enough DNA and RNA for the studies shown here. Additionally, our study consisted of only 1 time point. A longitudinal study should be performed to gain much needed insight into the dynamic changes in microbial communities and host responses within the nasal cavity. Finally, analysis of samples from individuals with asymptomatic infection would provide additional valuable insight into determinants of disease.

Funding

This study was supported by the National Center for Research Resources and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant UL1 TR001414, 1R01AI152258-02, and 3R01AA028735-01S1. N.R. is supported by NIH T32 AI007319. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Author Contributions

N.S.R., and I.M. conceived and designed the experiments. N.S.R., A.P., A.M. B.D., A.J. and I.C. performed the experiments. N.S.R, A.P., and A.M. analyzed the data. N.S.R, A.P., and I.M. wrote the paper. All authors have read and approved the final draft of the manuscript.

Acknowledgements

Aspects of experimental design figures were generated using graphics from Biorender.com. We wish to acknowledge the support of the Chao Family Comprehensive Cancer Center resources, award number P30CA062203 for providing the de-identified biospecimens used in this study.

Competing interests

The authors declare that they have no competing interests.

FIGURE LEGENDS

Figure 1: The nasal microbiome of SARS-CoV-2 infected patients is distinct. (A) Study design schematic. (B) Principal-coordinate analysis of nasal microbial communities unweighted UniFrac distance colored by host status. The contribution of host status the total variance in the unweighted UniFrac dissimilarity matrices was measured using PERMANOVA (Adonis with 10,000 permutations). (C, D, E) Violin plot illustrating (D) average unweighted UniFrac distances (C) number of observed amplicon sequencing variants and (D) shannon diversity (E) split by host status. Significance for panels C-E was determined using Kruskal Wallis non-parametric ANOVA (p-values inset at the bottom of each panel), with Dunn's multiple comparison * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$. (F) Bubble plots of bacterial genera found at greater than 1% average abundance across the entire study population ordered from left to right in descending average abundance. The size of each circle indicates the average relative abundance for each taxa in the denoted group and the color of each circle denotes which bacterial Phyla each genera belongs to.

Figure 2: The nasal microbiome of SARS-CoV-2 infected patients and healthcare workers are enriched in opportunistic bacterial pathogens. (A) Select differentially abundant genera in the nasal microbiome between CoV-, HCW, and CoV+ individuals. Differential abundance was determined using LEfSe (Log10 LDA score > 2). (B-H) Scatter plots of bacterial genera and species of interest identified by LEfSe analysis plotted as $\log_{10}$ relative abundance + 0.01 and colored by host status. Horizontal black lines represent the mean of each group with whisker the SEM. Significance for panels B-H was determined using Kruskal Wallis non-parametric ANOVA, with Dunn's multiple comparison * = $p < 0.05$, ** = $p < 0.01$, *** = $p < 0.001$, **** = $p < 0.0001$.

Figure 3: The nasal microbiome of SARS-CoV-2 infected patients is associated with viral RNA load. (A) Differentially abundant genera in the nasal microbiome between CoV+ individuals

stratified by viral RNA load into High, Mid, and Low Ct. Differential abundance was determined using LEFSE (Log10 LDA score > 2). (B-F) Scatter plots of bacterial genera and species of interest identified by LEFSE analysis plotted as log₁₀ relative abundance + 0.01 and across the three groups. Horizontal black bars represent the mean of each group with whisker the SEM. Significance for panels B-H was determined using Kruskal Wallis non-parametric ANOVA, with Dunn's multiple comparison * = p < 0.05, ** = p < 0.01, *** = p < 0.001, **** = p < 0.0001.

Figure 4: Transcriptional profiling of the nasal passages reveal robust immune activation.

(A) Volcano plot of global gene expression changes in CoV+ patients relative to CoV- HCWs. Upregulated differentially expressed genes (DEGs) are indicated in red; downregulated in blue. (B) Functional enrichment of upregulated. Horizontal bars represent the number of genes enriching to each GO term with color intensity representing the negative log of the FDR-adjusted p-value [-log(q-value)]. Range of colors based on the GO terms with the lowest and highest -log(q-value) values. (C) Heatmap of upregulated. Exemplar DEGs are highlighted. Columns of all heatmaps represent the rpkm of one individual. Range of colors per each heatmap is based on scaled and centered rpkm values of the represented DEGs. Red represents upregulation; blue, downregulation. (D) Functional enrichment of downregulated DEGs as described in (C). (E) Heatmap of downregulated DEGs with exemplar DEG highlighted. See (D) for additional details

Figure 5: Viral RNA load and genetic variation. (A) Genome coverage (%) for each Ct group. (B) Amino acids mutations occurring in genomes with coverage > 90%. Each row represents an individual subject. Horizontal axis is order 3' to 5'. (C) Nextstrain phylogenetic tree of genomes in (B) the original Wuhan isolate, variants of concern (B.1.1351, P.1, B.1.1.7), January 2020 Washington and New York isolates, and representative isolates from California January-September 2020 are used for comparison.

Supplemental Figure 1: Additional features of the nasal microbiome associated with SARS-CoV-2 infection.

(A) Principal-coordinate analysis of nasal microbiome weighted UniFrac distance colored by host status. The contribution of host status to the total variance in the weighted UniFrac dissimilarity matrices was measured using PERMANOVA (Adonis with 10,000 permutations). (B) Violin plot illustrating average weighted UniFrac distances within each group. (C) Stacked bar of bacterial genera found at greater than 1% average abundance across the entire study population. Each vertical bar represents an individual sample and are ordered horizontally by host status. (D) Violin plot of Ct values generated via reverse transcription quantitative PCR of the SARS-CoV-2 Spike gene showing the distribution of samples within each CT group. (E) Violin plot of observed amplicon sequencing variants split by host Ct group within CoV+ patients. (F) Principal-coordinate analysis of nasal microbiome unweighted UniFrac distance colored by Ct group. The contribution of Ct value to the total variance in the weighted UniFrac dissimilarity matrices was measured using PERMANOVA (Adonis with 10,000 permutations). (G) Violin plot illustrating average unweighted UniFrac distances within each group. (H) Principal-coordinate analysis of nasal microbiome weighted UniFrac distance colored by Ct group. The contribution of the Ct to the total variance in the weighted UniFrac dissimilarity matrices was measured using PERMANOVA (Adonis with 10,000 permutations). (I) Violin plot illustrating average weighted UniFrac distances between each CT group. Significance for panels B,E,G,I was determined using Kruskal Wallis non-parametric ANOVA, with Dunn's multiple comparison * = p < 0.05, ** = p < 0.01, *** = p < 0.001, **** = p < 0.0001.

Supplemental Figure 2: Nextstrain phylogenetic tree of genomes in Figure 5B, variants of concern (P.1, B.1.1.7, B.1.351, B.1.427, B.1.429) and all genomes collected from Orange County, California January 2020 to December 2020.

REFERENCES

1. P. Zhou *et al.*, A pneumonia outbreak associated with a new coronavirus of probable bat origin. *Nature* **579**, 270-273 (2020).
2. N. Zhu *et al.*, A Novel Coronavirus from Patients with Pneumonia in China, 2019. *The New England Journal of Medicine* **382**, 727-733 (2020).
3. W.-J. Guan *et al.*, Clinical Characteristics of Coronavirus Disease 2019 in China. *The New England Journal of Medicine* **382**, 1708-1720 (2020).
4. B. Hu, H. Guo, P. Zhou, Z.-L. Shi, Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*, 1-14 (2020).
5. A. G. Harrison, T. Lin, P. Wang, Mechanisms of SARS-CoV-2 transmission and pathogenesis. *Trends in Immunology* **0**, (2020).
6. D. Hornuss *et al.*, Anosmia in COVID-19 patients. *Clinical Microbiology and Infection* **26**, 1426-1427 (2020).
7. S. Lopez-Leon *et al.*, More Than 50 Long-Term Effects of COVID-19: A Systematic Review and Meta-Analysis. *Research Square*, (2021).
8. L. A. Callender *et al.*, The Impact of Pre-existing Comorbidities and Therapeutic Interventions on COVID-19. *Frontiers in Immunology* **11**, (2020).
9. H. Ejaz *et al.*, COVID-19 and comorbidities: Deleterious impact on infected patients. *J Infect Public Health* **13**, 1833-1839 (2020).
10. M. Hoffmann *et al.*, SARS-CoV-2 Cell Entry Depends on ACE2 and TMPRSS2 and Is Blocked by a Clinically Proven Protease Inhibitor. *Cell* **181**, 271-280.e278 (2020).
11. N. Murgolo *et al.*, SARS-CoV-2 tropism, entry, replication, and propagation: Considerations for drug discovery and development. *PLOS Pathogens* **17**, e1009225 (2021).
12. J. Shang *et al.*, Cell entry mechanisms of SARS-CoV-2. *Proceedings of the National Academy of Sciences* **117**, 11727-11734 (2020).
13. M. Chen *et al.*, Elevated ACE2 expression in the olfactory neuroepithelium: implications for anosmia and upper respiratory SARS-CoV-2 entry and replication. *European Respiratory Journal*, (2020).
14. W. Sungnak *et al.*, SARS-CoV-2 entry factors are highly expressed in nasal epithelial cells together with innate immune genes. *Nature Medicine* **26**, 681-687 (2020).
15. X. Zou *et al.*, Single-cell RNA-seq data analysis on the receptor ACE2 expression reveals the potential risk of different human organs vulnerable to 2019-nCoV infection. *Front. Med.* **14**, 185-192 (2020).
16. L. Zou *et al.*, SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. *N Engl J Med* **382**, 1177-1179 (2020).
17. G. Lui *et al.*, Viral dynamics of SARS-CoV-2 across a spectrum of disease severity in COVID-19. *The Journal of Infection* **81**, 318-356 (2020).
18. A. Yilmaz *et al.*, Upper Respiratory Tract Levels of Severe Acute Respiratory Syndrome Coronavirus 2 RNA and Duration of Viral RNA Shedding Do Not Differ Between Patients With Mild and Severe/Critical Coronavirus Disease 2019. *The Journal of Infectious Diseases*, (2020).
19. L. Zou *et al.*, SARS-CoV-2 Viral Load in Upper Respiratory Specimens of Infected Patients. *New England Journal of Medicine* **382**, 1177-1179 (2020).
20. O. Gallo, L. G. Locatello, A. Mazzone, L. Novelli, F. Annunziato, The central role of the nasal microenvironment in the transmission, modulation, and clinical progression of SARS-CoV-2 infection. *Mucosal Immunology* **14**, 305-316 (2021).
21. Y. He, J. Wang, F. Li, Y. Shi, Main Clinical Features of COVID-19 and Potential Prognostic and Therapeutic Value of the Microbiota in SARS-CoV-2 Infections. *Front. Microbiol.* **11**, (2020).

22. S. Khatiwada, A. Subedi, Lung microbiome and coronavirus disease 2019 (COVID-19): Possible link and implications. *Human Microbiome Journal* **17**, 100073 (2020).
23. J. Wilks, T. Golovkina, Influence of Microbiota on Viral Infections. *PLOS Pathogens* **8**, e1002681 (2012).
24. G. D. d. Melo *et al.*, COVID-19-related anosmia is associated with viral persistence and inflammation in human olfactory epithelium and brain infection in hamsters. *Science Translational Medicine*, (2021).
25. N. E. Rawson, L. Huang, Impact of Oronasal Inflammation on Taste and Smell: an Introduction. *Ann N Y Acad Sci* **1170**, 581-584 (2009).
26. J. Hadjadj *et al.*, Impaired type I interferon activity and inflammatory responses in severe COVID-19 patients. *Science* **369**, 718-724 (2020).
27. L. Kuri-Cervantes *et al.*, Comprehensive mapping of immune perturbations associated with severe COVID-19. *Sci Immunol* **5**, (2020).
28. D. Mathew *et al.*, Deep immune profiling of COVID-19 patients reveals distinct immunotypes with therapeutic implications. *Science* **369**, (2020).
29. C. R. Moderbacher *et al.*, Antigen-Specific Adaptive Immunity to SARS-CoV-2 in Acute COVID-19 and Associations with Age and Disease Severity. *Cell* **0**, (2020).
30. A. T. Tan *et al.*, Early induction of functional SARS-CoV-2-specific T cells associates with rapid viral clearance and mild disease in COVID-19 patients. *Cell Reports* **34**, (2021).
31. A. J. Wilk *et al.*, A single-cell atlas of the peripheral immune response in patients with severe COVID-19. *Nature Medicine* **26**, 1070-1076 (2020).
32. D. Butler *et al.*, Shotgun transcriptome, spatial omics, and isothermal profiling of SARS-CoV-2 infection reveals unique host responses, viral diversification, and drug interactions. *Nature Communications* **12**, 1660 (2021).
33. F. De Maio *et al.*, Nasopharyngeal Microbiota Profiling of SARS-CoV-2 Infected Patients. *Biological Procedures Online* **22**, 18 (2020).
34. C. Nardelli *et al.*, Nasopharyngeal Microbiome Signature in COVID-19 Positive Patients: Can We Definitely Get a Role to *Fusobacterium periodonticum*? *Front. Cell. Infect. Microbiol.* **11**, (2021).
35. M. Rueca *et al.*, Investigation of Nasal/Oropharyngeal Microbial Community of COVID-19 Patients by 16S rDNA Sequencing. *Int J Environ Res Public Health* **18**, (2021).
36. A. Budding, E. Sieswerda, B. Wintermans, M. Bos, "An Age Dependent Pharyngeal Microbiota Signature Associated with SARS-CoV-2 Infection," (Social Science Research Network, Rochester, NY, 2020).
37. Z. Shen *et al.*, Genomic Diversity of Severe Acute Respiratory Syndrome–Coronavirus 2 in Patients With Coronavirus Disease 2019. *Clinical Infectious Diseases* **71**, 713-720 (2020).
38. E. Bolyen *et al.*, Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. *Nat Biotechnol* **37**, 852-857 (2019).
39. N. Rhoades *et al.*, Maturation of the infant rhesus macaque gut microbiome and its role in the development of diarrheal disease. *Genome Biol* **20**, 173 (2019).
40. B. J. Callahan *et al.*, DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods* **13**, 581-583 (2016).
41. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol Biol Evol* **30**, 772-780 (2013).
42. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS One* **5**, e9490 (2010).
43. C. Quast *et al.*, The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res* **41**, D590-596 (2013).
44. I. F. Escapa *et al.*, New Insights into Human Nostril Microbiome from the Expanded Human Oral Microbiome Database (eHOMD): a Resource for the Microbiome of the Human Aerodigestive Tract. *mSystems* **3**, (2018).

- 576 45. C. Lozupone, M. E. Lladser, D. Knights, J. Stombaugh, R. Knight, UniFrac: an effective
577 distance metric for microbial community comparison. *ISME J* **5**, 169-172 (2011).
- 578 46. P. Dixon, VEGAN, a package of R functions for community ecology. **14**, 927-930 (2003).
- 579 47. N. Segata *et al.*, Metagenomic biomarker discovery and explanation. *Genome Biol* **12**,
580 R60 (2011).
- 581 48. H. B. TW, T. Girke, systemPipeR: NGS workflow and report generation environment.
582 *BMC Bioinformatics* **17**, 388 (2016).
- 583 49. Y. Zhou *et al.*, Metascape provides a biologist-oriented resource for the analysis of
584 systems-level datasets. *Nat Commun* **10**, 1523 (2019).
- 585 50. G. Grasselli *et al.*, Hospital-Acquired Infections in Critically Ill Patients With COVID-19.
586 *Chest*, (2021).
- 587 51. F. J. Whelan *et al.*, The loss of topography in the microbial communities of the upper
588 respiratory tract in the elderly. *Ann Am Thorac Soc* **11**, 513-521 (2014).
- 589 52. C. M. Bassis, A. L. Tang, V. B. Young, M. A. Pynnonen, The nasal cavity microbiota of
590 healthy adults. *Microbiome* **2**, 27 (2014).
- 591 53. J. C. Stearns *et al.*, Culture and molecular-based profiles show shifts in bacterial
592 communities of the upper respiratory tract that occur with age. *ISME J* **9**, 1246-1259
593 (2015).
- 594 54. S. M. Teo *et al.*, The infant nasopharyngeal microbiome impacts severity of lower
595 respiratory infection and risk of asthma development. *Cell Host Microbe* **17**, 704-715
596 (2015).
- 597 55. K. H. Lee *et al.*, The respiratory microbiome and susceptibility to influenza virus
598 infection. *PLoS One* **14**, e0207898 (2019).
- 599 56. S. Edouard *et al.*, The nasopharyngeal microbiota in patients with viral respiratory tract
600 infections is enriched in bacterial pathogens. *Eur J Clin Microbiol Infect Dis* **37**, 1725-
601 1733 (2018).
- 602 57. N. Wolter *et al.*, High nasopharyngeal pneumococcal density, increased by viral
603 coinfection, is associated with invasive pneumococcal pneumonia. *J Infect Dis* **210**,
604 1649-1657 (2014).
- 605 58. R. Xu *et al.*, Temporal association between human upper respiratory and gut bacterial
606 microbiomes during the course of COVID-19 in adults. *Commun Biol* **4**, 240 (2021).
- 607 59. D. Kaul *et al.*, Microbiome disturbance and resilience dynamics of the upper respiratory
608 tract during influenza A virus infection. *Nat Commun* **11**, 2537 (2020).
- 609 60. C. K. Lin, B. I. Kazmierczak, Inflammation: A Double-Edged Sword in the Response to
610 *Pseudomonas aeruginosa* Infection. *J Innate Immun* **9**, 250-261 (2017).
- 611 61. S. A. Riquelme, A. Prince, Airway immunometabolites fuel *Pseudomonas aeruginosa*
612 infection. *Respir Res* **21**, 326 (2020).
- 613 62. Y. Feng *et al.*, COVID-19 with Different Severities: A Multicenter Study of Clinical
614 Features. *Am J Respir Crit Care Med* **201**, 1380-1388 (2020).
- 615 63. M. Vaillancourt, P. Jorth, The Unrecognized Threat of Secondary Bacterial Infections
616 with COVID-19. *mBio* **11**, (2020).
- 617 64. F. Zhou *et al.*, Clinical course and risk factors for mortality of adult inpatients with
618 COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* **395**, 1054-1062
619 (2020).
- 620 65. H. Zhang *et al.*, Risks and features of secondary infections in severe and critical ill
621 COVID-19 patients. *Emerg Microbes Infect* **9**, 1958-1964 (2020).
- 622 66. E. Tajeddin *et al.*, The role of the intensive care unit environment and health-care
623 workers in the transmission of bacteria associated with hospital acquired infections. *J*
624 *Infect Public Health* **9**, 13-23 (2016).
- 625 67. S. Dimitri-Pinheiro, R. Soares, P. Barata, The Microbiome of the Nose-Friend or Foe?
626 *Allergy Rhinol (Providence)* **11**, 2152656720911605 (2020).

68. N. A. El Aila, N. A. Al Laham, B. M. Ayesh, Nasal carriage of methicillin resistant Staphylococcus aureus among health care workers at Al Shifa hospital in Gaza Strip. *BMC Infect Dis* **17**, 28 (2017).
69. C. P. Coutinho *et al.*, Long-term colonization of the cystic fibrosis lung by Burkholderia cepacia complex bacteria: epidemiology, clonal variation, and genome-wide expression alterations. *Front Cell Infect Microbiol* **1**, 12 (2011).
70. L. A. Kalish *et al.*, Impact of Burkholderia dolosa on lung function and survival in cystic fibrosis. *Am J Respir Crit Care Med* **173**, 421-425 (2006).
71. N. G. El Chakhtoura *et al.*, A 17-Year Nationwide Study of Burkholderia cepacia Complex Bloodstream Infections Among Patients in the United States Veterans Health Administration. *Clin Infect Dis* **65**, 1253-1259 (2017).
72. D. Butler *et al.*, Shotgun transcriptome, spatial omics, and isothermal profiling of SARS-CoV-2 infection reveals unique host responses, viral diversification, and drug interactions. *Nat Commun* **12**, 1660 (2021).
73. A. Islam *et al.*, Transcriptome of nasopharyngeal samples from COVID-19 patients and a comparative analysis with other SARS-CoV-2 infection models reveal disparate host responses against SARS-CoV-2. *J Transl Med* **19**, 32 (2021).
74. N. A. P. Lieberman *et al.*, In vivo antiviral host transcriptional response to SARS-CoV-2 by viral load, sex, and age. *PLoS Biol* **18**, e3000849 (2020).
75. D. L. Ng *et al.*, A diagnostic host response biosignature for COVID-19 from RNA profiling of nasal swabs and blood. *Sci Adv* **7**, (2021).
76. G. D. de Melo *et al.*, COVID-19-related anosmia is associated with viral persistence and inflammation in human olfactory epithelium and brain infection in hamsters. *Sci Transl Med*, (2021).
77. C. E. Gomez, B. Perdiguero, M. Esteban, Emerging SARS-CoV-2 Variants and Impact in Global Vaccination Programs against SARS-CoV-2/COVID-19. *Vaccines (Basel)* **9**, (2021).
78. W. Zhang *et al.*, Emergence of a Novel SARS-CoV-2 Variant in Southern California. *JAMA* **325**, 1324-1326 (2021).

Figure 1

Figure 2

Figure 3

Figure S1

Figure S2