SARS-CoV-2 Infection Reduces Human Nasopharyngeal Commensal Microbiome with Inclusion of Pathobionts

Md. Nazmul Hoque
Bangabandhu Sheikh Mujibur Rahman Agricultural University https://orcid.org/0000-0002-4861-0030
Md. Murshed Hasan Sarkar
Bangladesh Council of Scientific and Industrial Research
M. Shaminur Rahman
University of Dhaka Faculty of Biological Sciences
Shahina Akter
Bangladesh Council of Scientific and Industrial Research
Tanjina Akhtar Banu
Bangladesh Council of Scientific and Industrial Research
Barna Goswami
Bangladesh Council of Scientific and Industrial Research
Iffat Jahan
Bangladesh Council of Scientific and Industrial Research
M. Saddam Hossain
Bangladesh Council of Scientific and Industrial Research
A. K. Mohammad Shamsuzzaman
Bangladesh Council of Scientific and Industrial Research
Tasnim Nafisa
Bangladesh Council of Scientific and Industrial Research
M. Maruf Ahmed Molla
Bangladesh Council of Scientific and Industrial Research
Mahmuda Yeasmin
Bangladesh Council of Scientific and Industrial Research
Asish Kumar Ghosh
Bangladesh Council of Scientific and Industrial Research
Eshrar Osman
Bangladesh Council of Scientific and Industrial Research
Mohammad Samir Uzzaman
Bangladesh Council of Scientific and Industrial Research
Md Ahashan Habib
Bangladesh Council of Scientific and Industrial Research
Abu Sayeed Mohammad Mahmud
Bangladesh Council of Scientific and Industrial Research
Keith A Crandall
The George Washington University
Tofazzal Islam

Bangabandhu Sheikh Mujibur Rahman Agricultural University
M. Salim Khan (k2salim@yahoo.com)
https://orcid.org/0000-0003-0603-6728

Research

Keywords: SARS-CoV-2, nasopharyngeal microbiome, RNA-seq analyses

Posted Date: May 6th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-480069/v1

Version of Record: A version of this preprint was published at Scientific Reports on December 1st, 2021. See the published version at https://doi.org/10.1038/s41598-021-03245-4.

Abstract

Background

The SARS-CoV-2 primarily enters into the human body through nasopharyngeal tract (NT) and is the etiological agent of COVID-19. The microbiota of the NT may play a role in host immunity against respiratory infectious diseases. However, scant information is available on interactions of SARS-CoV-2 with the nasopharyngeal microbiome. To shed light on the effects and consequences of SARS-CoV-2 infection on microbiome diversity and composition, we conducted a high throughput RNA-Seq metagenomic investigation of 22 NT swab samples (including COVID-19 = 8, Recovered = 7, and Healthy = 7).

Results

This study for the first time demonstrates the association of microbiome diversity and their concomitant genomic features in the NT of COVID-19 and Recovered patients compared to Healthy humans, and discusses the role of the altered microbiomes in the pathophysiology of the SARS-CoV-2 infections. Our RNA-Seq metagenomic analyses detected 2281 bacterial species (including 1477, 919 and 676 in samples of Healthy human, COVID-19 and Recovered patients, respectively) indicating a distinct microbiome dysbiosis (COVID-19>Recovered>Healthy). The samples from COVID-19 patients and Recovered individuals had inclusion of 67% (including *Streptococcus salivarius*, *S. mitis*, *Neisseria subflava*, *Veillonella dispar*, *Acinetobacter junii*, *Prevotella melaninogenica* etc.) and 77% (including *Pseudomonas stutzeri*, *Staphylococcus capitis*, *S. epidermidis*, *P. mendocina*, *Moraxella osloensis*, *A. indicus*, *Escherichia coli* etc.) opportunistic pathogenic bacteria, respectively compared to Healthy individuals. Notably, 79% commensal bacteria (e.g., *Pseudomonas* sp. LPH1, *Brevundimonas* sp. Bb-A, *P. oleovorans*, *Pseudomonas* sp. phDV1, *Brevundimonas* sp. DS20, *Idiomarinaceae bacterium* HL-53, *Alishewanella* sp. 205, *Sphingobacterium psychroaquaticum* etc.) were found in healthy individuals but not detected in COVID-19 patients and Recovered individuals. Similar dysbiosis was also found in viral and archaeal fraction of the microbiomes. Although, 55 viral and 48 archaeal genera were detected, only 16% viral and 27% archaeal genera were shared across three metagenomes. We also detected altered metabolic pathways and functional genes including resistance to antibiotics and toxic compounds in the pathophysiology of COVID-19.

Conclusions

The nasopharyngeal microbiome dysbiosis and their genomic features in COVID-19, Recovered and Healthy individuals determined by our RNA-Seq analyses shed light on early interactions of SARS-CoV-2 with the nasopharyngeal resident microbiota that might be helpful for developing microbiome-based diagnostics and therapeutics for this novel pandemic disease.

Introduction

Immediately after the first emergence of SARS-CoV-2 in Wuhan province of China in December 2019, this novel Coronavirus disease 2019 (COVID-19) spread across 217 countries and/or territories of the globe as a pandemic [1, 2]. SARS-CoV-2 emerged as one of the deadliest human pathogens in the last hundred years after the Spanish Flu in 1918–1920 [3]. The virus primarily enters the human body mainly through the ACE2 and TMPRSS2 receptors, and nasal epithelial cells of the nasopharyngeal tract (NT) of the respiratory system [4], and then gradually move towards the lung to initiate infection. The pathophysiology of SARS-CoV-2 infections can be attributed to aberrant immune responses in clearing the virus [5, 6]. Given the unequivocal association between viral and bacterial co-infections and

respiratory disease severity, there is a pressing need to better understand how interactions of SARS-CoV-2 with the host microbiome in the respiratory tracts correlate with viral infections that facilitate opportunistic co-infections [7].

Human microbiota play a crucial role in immunity and health of individual hosts. Any alterations of the diversity and population of the resident microbiota are associated with various chronic and acute human diseases [5, 8]. In fact, the host-microbiota symbiotic equilibrium is highly complex and associated with a number of intrinsic and environmental factors. Disruption of the homeostasis in composition of human microbiota leads to a state of "dysbiosis"[9]. Therefore, microbiome dysbiosis in the respiratory tract by the pathogenic respiratory virus can increase the mortality rate in patients [10, 11]. Clinical trials and high throughput sequencing (metagenomic and RNA-Seq)-based investigations revealed the co-presence of diverse viruses, bacteria, archaea and fungi in the respiratory tracts of SARS-CoV-2 infected patients [12, 13]. Recent studies of bronchoalveolar lavage fluid showed about 50% of the patients who died of COVID-19 had secondary bacterial infections [12, 14].

Human health is the outcome of the complex interactions between the inhabiting microbiome and its human host [15]. The inhaled SARS-CoV-2 virus particle likely binds to epithelial cells in the nasal cavity, replicates and then migrates down the respiratory tract along the conducting airways which triggers a robust innate immune response [4]. Therefore, it can assume that during this propagation, migration and immune response, the resident microbiomes in the respiratory airways could be altered or changed, and inclusion of some of the pathobionts might aggravate the progression and lethality of the disease caused by SARS-CoV-2. Although several lines of evidence suggest that development of COVID-19 disease modulates the population and diversity of resident commensal microbiota of humans, little is known about the outcome of the interactions of SARS-CoV-2 with nasal commensal microbiome which is thought to be critical for transmission, modulation, and progression of COVID-19 [16, 17]. It is important to understand the interactions of SARS-CoV-2 on the composition and diversity of microbiomes in the NT of COVID-19 and Recovered patients compared to the Healthy individuals. To shed light on the effects and consequences of SARS-CoV-2 infection on the NT microbiome, we conducted a high throughput RNA-Seq analysis of the nasopharyngeal swabs of randomly selected Healthy humans, COVID-19 and Recovered patients. Furthermore, we conducted a functional analysis to identify potential biological mechanisms linking the shift of microbiome, SARS-CoV-2 genomic diversity and host disease in COVID-19, Recovered and Healthy states. This report for the first time demonstrates the association of microbiome diversity and composition (Graphical abstract), and their concomitant genomic features in the nasal cavity of COVID-19 and Recovered patients compared to Healthy humans, and discusses the role of the altered microbiome in the pathophysiology of the SARS-CoV-2 infections.

Results

SARS-CoV-2 infection modulates the community composition and diversity of nasopharyngeal microbiomes

To shed light on the effects of SARS-CoV-2 infections in the diversity and composition of human nasal microbiome, we analyzed RNA-Seq data of nasopharyngeal samples of randomly selected Healthy human, COVID-19 and Recovered patients. The alpha-diversity (within sample diversity through Shannon and Simpson indices) assessment identified significant differences in microbial species richness among the COVID-19, Recovered and Healthy controls regardless of the method used to tabulate microbial abundances i.e., either PathoScope 2.0 (PS) or MG-RAST 4.13 (MR), showing higher diversity in the microbial niche of COVID-19 samples (COVID-19, $p = 0.0065$; Recovered, $p = 0.0105$; Healthy, $p = 0.0401$; Kruskal-Wallis rank sum test) (Fig. 1A). The beta diversity (between sample or metagenome diversity) through principal coordinate analysis (PCoA), as measured on the Bray-Curtis distance method using PS and MR data, showed distinct discrimination across the metagenomes and separated samples by microbial population structure (Fig. 1B). Therefore, significant variation in microbiome diversity and composition

across these three metagenomes (p = 0.0059, Kruskal-Wallis test) was evident. The diversity of microbiomes between samples (PCoA) however did not vary significantly (p = 0.651, Kruskal–Wallis rank sum test) according to the gender (male or female) of the study population (Fig. 1B).

Microbiome composition at the domain level was numerically dominated by bacteria, with a relative abundance of 81.58%, followed by viruses (18.22%) and archaea (0.20%) (Data S1). The unique and shared distribution of microbes found in the three participant groups are represented by comprehensive Venn diagrams (Fig. 2). Our analyses detected 532 bacterial genera, including 486, 421 and 509 in COVID-19, Recovered and Healthy nasopharyngeal samples, respectively, and of them, 72.74% genera were common in three sample groups (Fig. 2A, Data S1). Notably, we detected 2281 bacterial species through PS analysis, of which 919, 676 and 1477 species were found in COVID-19, Recovered, and Healthy samples, respectively (Fig. 2B, Data S1). In this study, compared to COVID-19 and Recovered samples, the Healthy samples had unique or sole association of 1166 bacterial species which underwent dysbiosis during the pathogenesis and subsequent recovery phase of COVID-19 (Data S2). Of the identified bacterial species, 67.27% and 76.78% had sole association in COVID-19 and Recovered samples, respectively indicating their opportunistic inclusion in COVID-19 patients and re-establishment beneficial commensal flora with the recovery of SARS-CoV-2 infections (Fig. 2B, Table 1, Data S2).

The MR pipeline detected 55 viral and 48 archaeal genera in three metagenomes, and of them, 16.37% viral (Fig. 2C), and 27.08% archaeal (Fig. 2D) genera were found to be shared among these metagenomes. By comparing these genera across the sample categories, we identified 35, 31 and 23 viral, and 42, 20 and 35 archaeal genera in COVID-19, Recovered and Healthy-swabs, respectively (Table S1). We found that 5.45% and 16.36% viral (Fig. 2B), and 37.5%, and 8.33% archaeal genera in Healthy and Recovered samples, respectively shared with those of COVID-19 samples. The COVID-19 patients had sole association of 14 (25.45%) viral, and 7 (14.58%) viral genera (Fig. 2C-D, Table S1, Data S1). Moreover, 213, 52 and 71 viral species (bacteriophages mostly) were identified in COVID-19, Recovered and Healthy metagenomes, respectively. Among the detected viral species, 97.65% and 98.07% had sole association with COVID-19 and Recovered samples, respectively (Data S2).

SARS-CoV-2 infection reduces the diversity and composition of nasopharyngeal commensal bacteria

The present microbiome study demonstrated that both the composition and the relative abundances of bacterial taxa at the phylum, order, family, genus, and species-level differed significantly (p = 0.031, Kruskal Wallis test) among COVID-19, Recovered and Healthy controls. *Firmicutes* was found to be the most predominant bacterial phylum in COVID-19 and Recovered metagenomes with a relative abundance of 76.31% and 39.28%, respectively (Data S1). The other predominant bacterial phyla were *Bacteroidetes* (7.07%), *Proteobacteria* (6.85%), *Fusobacteria* (5.37%), *Actinobacteria* (1.45%), and *Cyanobacteria* (1.23%) in COVID-19 metagenome, and *Proteobacteria* (27.42%), *Actinobacteria* (16.61%), *Bacteroidetes* (12.86%), and *Cyanobacteria* (1.71%) in Recovered metagenome (Data S1). On the other hand, *Proteobacteria* (50.19%), *Bacteroidetes* (41.60%), and *Firmicutes* (6.42%) were the most abundant phyla in Healthy control samples (Data S1).

We found significant differences (p = 0.021, Kruskal-Wallis test) in the diversity and relative abundance of the bacteria at the genus level. The individual and inter-individual microbiome variability showed that Healthy individuals had higher number of bacterial genera (average 318.57/person) compared to COVID-19 (234.50/patient) and Recovered (156.29/person) patients (Fig. S1A). However, the bacterial genera detected in the metagenome COVID-19 patient mapped to the highest number of reads per genus (476.95 reads/genus) compared to the healthy (278.89 reads/genus) and COVID-19 recovered (37.36 reads/genus) individuals (Fig. S1B). In COVID-19 metagenome, *Streptococcus* was the most abundant bacterial pathogen with a relative abundance of 37.16% followed by

Veillonella (24.25%), *Prevotella* (4.97%), *Staphylococcus* (3.49%), *Fusobacterium* (2.89%), *Clostridium* (2.59%), *Leptotrichia* (2.26%), and *Coprobacillus* (2.24%) (Fig. 3, Data S1). Likewise, top abundant bacterial genera in the metagenome of recovered individuals were *Staphylococcus* (28.82%), *Streptomyces* (9.11%), *Acinetobacter* (8.96%), *Corynebacterium* (5.18%), *Streptococcus* (2.48%), and *Helicobacter* (2.42%) (Fig. 3, Data S1). Conversely, the Healthy control metagenome was predominated by *Pedobacter* (11.69%), *Sphingobacterium* (6.35%), *Pseudomonas* (5.03%), *Enterobacter* (4.79%), *Flavobacterium* (4.62%), *Pseudoalteromonas* (3.82%), *Escherichia* (3.50%), *Exiguobacterium* (3.38%), *Shewanella* (3.16%), *Chryseobacterium* (2.58%), *Aeromonas* (2.52%), *Klebsiella* (2.50%), and *Vibrio* (2.03%). The rest of the genera detected in these metagenomes had relatively lower abundances (<2.0%) (Fig. 3, Data S1).

We further investigated the species-level differences of microbial communities across these three metagenomes through PS analysis, which revealed significant variations ($p = 0.011$, Kruskal-Wallis test) in microbiome composition, diversity and relative abundances (Fig. 4, Table 1, Data S1 and S2). The COVID-19 metagenome was dominated by 73 species (7.94%) of *Streptococcus* genus while *Clostridium*, *Chrysobacterium*, *Paenibacillus*, *Neissaria*, *Acinetobacter*, *Staphylococcus*, and *Corynebacterium* genera were represented by 38, 26, 25, 23, 22, 22 and 17 different species, respectively. Similarly, the bacteriome of the Recovered patients was dominated by different species of *Corynebacterium* (39), *Acinetobacter* (36), *Chryseobacterium* (30), *Sphingobacterium* (26), *Staphylococcus* (24), *Pseudomonas* (18), and *Flavobacterium* (16) (Fig. 4, Table 1, Data S1-S2). In contrast, the bacteriome of the Healthy control was predominantly represented by 145 species (9.82%) of *Pseudomonas* genus followed by 49, 41, 38, 38, 36, 36, 33, and 26 species of *Vibrio*, *Shewanella*, *Acinetobacter*, *Flavobacterium*, *Chryseobacterium*, *Enterobacter*, *Pseudoalteromonas*, and *Sphingobacterium* genera, respectively (Fig. 4, Table 1). Remarkably, 78.94% (1166/1477) bacterial species were solely associated with the Healthy nasopharyngeal samples, and not detected in SARS-CoV-2 infected (COVID-19) and Recovered patients. Among these depleted commensal species *Pseudomonas* sp. LPH1 (25.32%), *Brevundimonas* sp. Bb-A (4.85%), *P.oleovorans* (3.25%), *Pseudomonas* sp. phDV1 (1.75%) and *Brevundimonas* sp. DS20 (1.38%) were top abundant (Fig. 4, Table 1, Data S1-S2). Conversely, the COVID-19 samples had sole association of 609 opportunistic bacterial species, and of them, *Streptococcussalivarius* K12 (19.13%), *S. mitis* (18.13%), *Neisseria subflava* (13.77%), *Veillonella dispar* (11.03%), *Acinetobacter junii 64.*5 (3.31%), *V. parvula* (2.35%), *Prevotella melaninogenica* (2.26%), *S. parasanguinis* (2.22%), *Streptococcus* sp. LPB0220 (1.98%), *N. flavescens* (1.26%), and *V. atypica* (1.11%) were top abundant (Fig. 4, Table 2, Data S2). Likewise, the Recovered patients of COVID-19 had sole association of 519 species with comparatively higher relative abundances of *Pseudomonas stutzeri* DSM 4166 (8.48%), *Staphylococcus capitis* (5.89%), *S. epidermidis* RP62A (5.13%), *P. mendocina* NK-01 (3.12%), *Moraxellaosloensis* A1920 (2.60%), *A. indicus* A648 (2.57%), *Escherichia coli* (2.41%), *Sphingobacterium* sp. G1-14 (1.96%), *A. junii* 64.5 (1.85%), *S. pneumoniae* (1.60%), *Ralstonia pickettii* (1.55%), *Micrococcus luteus* (1.46%), *Rheinheimera* sp. D18 (1.33%), *Corynebacterium segmentosum* (1.26%), *Elizabethkingia anopheles* (1.25%), and *Cutibacterium acnes* (1.08%). The rest of the species in these metagenomes had relatively lower (< 1.0%) abundances (Fig. 4, Table 1, Data S2).

SARS-CoV-2 infection associated dysbiosis of viruses and archaea in the nasopharyngeal tract

The COVID-19 and Recovered patients had higher number of viral genera and/or species compared to Healthy controls. The COVID-19 and Healthy samples were predominated by the genus of *Betacoronavirus* with a relative abundance of 97.95% and 83.57%, respectively. The other abundant viral genera in COVID-19 samples were *Alphacoronavirus* (0.90%) and *Siphovirus* (0.54%) while *Alphacoronavirus* (3.54%), *T1-like viruses* (3.25%), *Cystovirus* (3.23%), *Siphovirus* (1.91%), *Myovirus* (1.09%), *Lambda-like viruses* (0.78%), and *N4-like viruses* (0.72%) (Fig. 5, Data S1). The Recovered sample were mostly dominated by *Alphacoronavirus* (92.54%) followed by *Betacoronavirus* (1.23%), *Siphovirus* (1.16%), *T1-like viruses* (0.81%), *Macavirus* (0.67%), and *SP6-like viruses* (0.56%) (Fig. 5, Data S1).

Most of these viral genera had lower relative abundances in Healthy controls. In addition, the SARS-CoV-2 was found as the predominantly abundant viral strain in COVID-19 (99.44%) patients whereas *Human coronavirus* NL63 remained as the top abundant viral strain in Recovered (90.89%) patients. The rest of the viral strains in three metagenomes were mostly different strains of bacteriophages (Data S1).

In this study, top abundant archaeal genera COVID-19 samples were *Halogeometricum* (19.57%), *Haloquadratum* (10.53%), *Natrialba* (8.06%), *Methanosarcina* (6.25%), *Halorhabdus* (5.76%), *Methanocaldococcus* (5.76%), *Haloterrigena* (5.59%), *Methanobrevibacter* (4.11%), *Halorubrum* (3.95%), *Methanococcoides* (3.95%), *Methanococcus* (2.96%), and *Methanocorpusculum* (2.96%). Correspondingly, *Haloterrigena* (36.47%), *Methanocaldococcus* (35.29%), *Halogeometricum* (6.47%), *Thermococcus* (5.29%), *Haloquadratum* (2.94%), *Methanosarcina* (2.35%), and *Pyrococcus* (2.80%) were the predominating archaeal genera in the Recovered metagenome (Fig. 6, Data S1). The Healthy samples, however, possessed *Methanospirillum* (15.34%), *Methanoregula* (10.58%), *Methanocaldococcus* (10.05%), *Methanosarcina* (9.00%), *Thermococcus* (7.94%), *Haloterrigena* (6.88%), *Methanococcus* (3.18%), *Methanoculleus* (3.18%), *Methanosphaera* (3.18%), *Euryarchaeota* (3.18%), *Methanobrevibacter* (2.65%), *Sulfolobus* (2.65%), and *Haloquadratum* (2.12%) genera with relatively higher abundances. The rest of the archaeal genera had lower relative abundances (<2.0%) across these metagenomes (Fig. 6, Data S1). Notably, seven archaeal genera (14.58%) had inclusion in COVID-19 metagenome compared to those of Healthy controls, and of them, *Methanopyrus* (0.82%) and *Metallosphaera* (0.66%) were predominantly abundant (Data S2). The Recovered and Healthy samples also had sole association of 2 (4.17%) and 3 (6.25%) archaeal genera, and the relative abundances of these genera remained much lower (0.59%) in both metagenomes (Data S2).

SARS-CoV-2 infection associated changes in genomic potentials of the microbiomes

In our current RNA-Seq data set, there was a broad variation in resistance to antibiotics and toxic compounds (RATCs) diversity and composition across three metagenomes (Fig. 7, Data S3). The categories and relative abundances of the RATC were significantly correlated ($p = 0.027$, Kruskal-Wallis test) with the relative abundance of the associated bacteria found in COVID-19, Recovered and Healthy control metagenomes (Data S3). MG-RAST identified 49 RATCs distributed in the bacterial genomes of the three metagenomes (Fig. 7, Data S3). The COVID-19 associated microbiomes harbored the 40 different RATC groups while 32 and 44 RATC gene families were detected in the microbiomes of Recovered and Healthy control samples. Of the detected RATCs, genes associated with biofilm formation in *Staphylococcus* had the highest relative abundances in Recovered samples (80.0%) followed by COVID-19 samples (45.95%), and Healthy controls (3.57%). Moreover, genes associated with acriflavin resistance were also predominantly abundant in COVID-19 (58.86%) and Recovered (66.67%) metagenomes. The other predominantly abundant RATC functional groups in COVID-19 metagenome were quorum sensing: autoinducer-2 synthesis (29.73%), multidrug resistance cluster; *mdtABCD* (21.0%), cobalt-zinc-cadmium resistance (20.64%), BlaR1 regulatory family (15.56%), multidrug resistance efflux pump; *pmr*A (11.39%), resistance to fluoroquinolones (10.83%), *lsrACDBFGE* operon (10.81%) and multidrug resistance efflux pumps (10.25%) (Fig. 7, Data S3). Simultaneously, the microbiomes of the Recovered samples harbored higher abundance of genes encoding biofilm adhesin biosynthesis (20.0%), cobalt-zinc-cadmium resistance (18.97%), multiple antibiotic resistance (MAR) locus (17.50%), methicillin resistance in *Staphylococci* (14.66%), macrolide-specific efflux protein; *macA* (13.33%), arsenic resistance (12.93%), multidrug resistance efflux pumps (12.93%), and *YjgK* cluster linked to biofilm formation (11.0%) (Fig. 7, Data S3). Conversely, beta-lactamase resistance (35.71%), multidrug resistance efflux pump; *pmrA* (29.10%), biofilm formation in *Staphylococcus* (22.73%), quorum sensing in *Yersinia* (21.43%) and *Pseudomonas* (17.86%), and teicoplanin-resistance in *Staphylococcus* (12.82%) were top abundant RATC functional groups in the microbiomes of the Healthy

controls. The rest of RATC groups also varied in their expression levels across the three metagenomes, being more prevalent in the COVID-19 and Recovered microbiomes (Fig. 7, Data S3).

By examining the correlation between the different gene families of the same KEGG pathway for COVID-19, Recovered and Healthy controls microbiomes, we found significant differences ($p = 0.034$, Kruskal-Wallis test) in their relative compositions and abundances. Our analysis revealed that genes coding for pyruvate carboxylase (*pyc*) had several-fold overexpression in the microbiomes of the COVID-19 (26.31%) and Recovered (4.11%) groups compared to Healthy controls (1.80%) (Fig. 8A, Data S3). In addition, genes encoding for adherent junction (25.17%), tight junction (24.48%), environmental information processing (15.03%), carbohydrates metabolism (11.51%) and oxidative phosphorylation (7.26%) had higher relative abundances in the COVID-19 associated nasopharyngeal microbiomes (Fig. 8A, Data S3). The Recovered microbiomes however had higher relative abundances of genes coding for focal adhesion (53.08%), transport and catabolism (40.44%), cell adhesion molecules (40.00%), genetic information and processing (35.12%), lysosome activity (28.68%), endocytosis (27.13%), and cell-to-cell communication (20.38%) (Fig. 8A, Data S3). Conversely, the gene families associated with bacterial secretion system (40.09%), cell motility (39.83%), succinyl-CoA synthetase subunit C and D (34.35%), gap junction (27.27%), protein metabolism (19.49%), TCA cycle (18.72%) and citrate synthase (16.89%) remained overexpressed in Healthy nasopharyngeal microbiomes (Fig. 8A, Data S3).

We also sought to gain further insight into the SEED hierarchical protein functions represented by different genes, and found 37 statistically different ($p = 0.013$, Kruskal-Wallis test) SEED functions in COVID (COVID-19 and Recovered) and Healthy control metagenomes. Overall, the COVID-19 and Recovered samples associated microbiomes showed a higher relative abundance of these SEED functions compared to those of Healthy controls. For instance, cytokine-cytokine receptor interaction (50.0%), regulation of oxidative stress response (26.82%), phage integration and excision (24.92%), toxin-antitoxin regulation system (17.24%), protection from reactive oxygen species (14.96%), and phage regulation (6.06%) related gene expression remained more than two-times overexpressed in Healthy commensal microbiomes compared two COVID-19 and Recovered sample-associated microbiomes (Fig. 8B, Data S3). The COVID-19 microbiomes were enriched in genes coding for cell growth and death (55.60%), *Streptococcuspyogenes* virulence regulators (42.44%), fratricide in *Streptococcus* (32.67%), neuroactive ligand receptor interaction (28.21%), epithelial cell signaling (20.35%), clustering-based subsystems (14.65%), proteolytic pathways (9.2%), prophage lysogenic conversion modules (3.69%) and bacterial invasion of epithelial cells (1.93%) compared to rest of the two metagenomes (Fig. 8B, Data S3). The Recovered microbiomes however had a higher abundance of SEED functions involved in glutathione: non-redox reactions (30.30%) and redox cycle (6.06%), coagulase cascade (20.0%), osmotic stress (16.07%), membrane transport (14.55%), MT1-MMP pericellular network (10.0%), BarA-UvrY(*SirA*) two-component regulatory system (9.29%), and oxidative stress (5.54%). In contrast, most of these SEED modules had relatively lower abundance in Healthy sample microbiomes (Fig. 8B, Data S3).

Discussion

In the current study, we demonstrated a remarkable shift in the diversity and population of the nasopharyngeal microbiomes (bacteria and archaea) of COVID-19 and recovered patients of COVID-19 compared to the healthy humans using a cutting-edge RNA-Seq technology. Our metagenomic analyses of nasopharyngeal samples from COVID-19 patients, Recovered and Healthy individuals also revealed that SARS-CoV-2 infection reduces commensal bacteria and archaea but increases the inclusion of pathobionts in the nasopharyngeal tract of human. Furthermore, we identified a number of microbial genomic features, altered metabolic pathways, and functional genes associated with COVID-19 pathogenesis. Although the dysbiosis of human gut microbiome by the infection of SARS-CoV-2 has

been reported in several studies [5, 18, 19], our study for the first time determined the interactions and consequences of SARS-CoV-2 infection with resident nasopharyngeal microbiome of human.

Since some of the COVID-19 patients showed persistent symptoms after recovery and/or subsequently develop multisystem inflammation [5], we hypothesized that the altered nasopharyngeal microbiomes of the SARS-CoV-2 infections could lead to abnormal inflammatory reactions to worsen the symptoms and treatment outcomes of COVID-19. The microbiome diversity (alpha and beta diversity) measures suggest that microbial dysbiosis is closely linked to SARS-CoV-2 infections. Our analysis also found a substantial microbial disparity between COVID-19 and Healthy-controls swab samples keeping the closest relationship between COVID-19 and Recovered nasopharyngeal samples, and related microbial signatures. The COVID-19-associated microbiomes had a greater variance in diversity than those of Healthy human nasopharyngeal microbiomes, agreeing with several recent studies [20, 21]. In a recent study, Zou et al. reported that loss of salutary species in COVID-19 persisted in most patients despite clearance of SARS-CoV-2 virus, and is associated with a more long-lasting detrimental effect to the gut microbiome [21].

The nasopharyngeal microbiomes predominantly identified in this study are bacteria (> 81.50%), however, other concomitant microbial players like viruses and archaea were also detected, highlighting the novel insights of diverse microbial association with SARS-CoV-2 to exacerbating the course, common symptoms and treatment outcomes of the COVID-19 [13, 19, 22]. In COVID-19 samples, the relative abundances of *Firmicutes*, *Bacteroidetes* and *Proteobacteria* phyla, and their related genera such as *Streptococcus*, *Veillonella*, *Prevotella*, *Acinetobacter*, *Staphylococcus*, *Clostridium*, *Coprobacillus*, and *Neisseria* remained much higher compared to those of Recovered and Healthy samples (Table 1). For instance, two bacterial genera such as *Streptococcus* and *Veillonella* had more than 5-fold higher relative abundances in COVID-19 associated samples than the Recovered and Healthy samples. *Prevotella*, *Veillonella*, and *Streptococcus* (Table 1) are the most common genera that reside in the lungs and nasopharynx [23, 24], and recently these genera have been reported to play an opportunistic role in the progression of lung infections [24, 25]. Despite having almost homogeneous genetic backgrounds and living in the same region, there were vast differences in microbiome signatures in nasal cavities of Healthy, COVID-19 patients and Recovered individuals. The Healthy people had higher number of bacterial genera compared to COVID-19 patients and Recovered individuals. Moreover, the inter-individual microbiome signature (bacterial genera) of participants was also observed, and was largely stable in COVID-19 patients. The human microbiota shows a remarkable amount of diversity among different individuals, and there are ample of emerging evidences that the microbial communities can be altered to change pathophysiological state or even cure of disease is a compelling drivers of microbiome variation [26].

Recent advances in next generation sequencing technology provided overwhelming evidence that microbial traits, for instance, pathogenicity, virulence, antibiotic resistance, and metabolic potentials are linked with species or strain specific genomic characteristics [19, 20, 27]. We demonstrated that the Healthy samples possessed highest number of commensal bacterial species while SARS-CoV-2 infections and subsequent therapy reduced 37.78% and 54.23% bacterial species in COVID-19 and Recovered patient's nasopharyngeal samples, respectively indicating the perturbations of beneficial microbes amid COVID-19. The COVID-19 and Recovered patients had sole association of 609 and 519 bacterial species, respectively. These results are in line with several previous studies who reported that cross-talk and/or interactions between SARS-CoV-2 and oral microbiomes [22], and SARS-CoV-2 and gut microbiomes [28] is associated with the pathophysiology of lung diseases. Therefore, majority of the bacterial species identified in COVID-19 (66.26%) and Recovered (76.78%) samples had an opportunistic inclusion by the depletion of beneficial commensal microbes during the course of SARS-CoV-2 pathogenesis and recovery phase. The composition of the bacterial communities in the nasopharynx is more diverse than any other parts of the upper respiratory tract, which

are characterized by different *Streptococcal* species, *Staphylococcus* spp., *Haemophilus* spp., *Neisseria* spp., *Rothia* spp., and anaerobes, including *Veillonella* spp., *Prevotella* spp., and *Leptotrichia* spp. [15, 29, 30]. Despite some preliminary evidence, it is still unclear whether similar synergies exist between SARS-CoV-2 and specific bacterial infections as exist with influenza [31, 32]. Surprisingly, 78.94% bacterial species had sole association with Healthy NT samples, which were not detected in COVID-19 and Recovered patients indicating the potential dysbiosis of commensal microbiomes during pathophysiology of SARS-CoV-2 infections. In this study, different species of *Pedobacter*, *Sphingobacterium*, *Pseudomonas*, *Enterobacter*, and *Flavobacterium* genera remained top abundant in Healthy human NT swabs. Our current findings are corroborated with recent clinical trials reporting that COVID-19 patients who received medical care had a relatively higher abundance of *Streptococcus, Clostridium, Lactobacillus*, and *Bifidobacterium* and lower levels of *Bacteroidetes, Roseburia, Faecalibacterium, Coprococcus* and *Parabacteroides* compared with healthy controls corroborating our current findings[33, 34]. Though, the nature of nasopharyngeal commensal bacteria that exert antiviral effect in the lungs is still elusive, some members of these commensal species (desaminotyrosine producing anaerobes) are important in the priming of the pulmonary innate immune system [19]. Despite the beneficial role of these commensal bacterial species are largely unknown, reduction in the composition and relative abundances of these commensal bacteria from the NT, play an important role in microbiome formation, which can significantly affect the immune response against SARS-CoV-2 infections [35].

This state-of-the-art RNASeq approach also provided an exciting opportunity for investigating the integrated cross-kingdom interactions of 'multibiome' such as viruses (other than SARS-CoV-2) and archaea (< 20.0% of total microbiome). Unlike bacteria, the diversity and composition of viruses and archaea always remained much lower in both COVID-19 and Healthy samples. SARS-CoV-2 was the predominantly abundant (> 99.0% relative abundances) viral strain in COVID-19 metagenome and rest of the viral strains in this metagenome were associated with *Alphacoronavirus* and *Siphoviruses.* On the contrary, *Human coronavirus* NL63 was the most abundant viral strain (90.89%) in COVID-Recovered metagenome followed by SARS-CoV-2 (2.76%). Other viral strains in three metagenomes were mostly different strains of bacterial phages. Recently, stronger relationship among other viruses, bacteria, fungi, and SARS-CoV-2 have been reported from different countries [36-38]. The endemic human coronaviruses (hCoVs) can cause co-infections, sequential infections or can be co-detected with each other or with other respiratory viruses [39]. The archaeal fraction of the humans NT microbiome across three metagenomes were predominated by different methanogenic and thermophilic genera. Though the role of these accompanying microbiomes in the pathophysiology of COVID-19 patients remains a mystery, it may include virus-induced airway damage, cell loss, goblet cell hyperplasia, altered mucus secretion, reduced ciliary beat frequency, function and clearance, reduced oxygen exchange, and damage to the immune system [40].

One of the important findings of our research is the detection of various homologs of RATCs belonging to different protein families among the microbiome of three metagenomes. Interestingly, the composition and abundances of the RATCs remained significantly correlated with the abundances of the related bacteria found in the nasal cavities of the healthy human, COVID-19 and COVID-19 recovered patients. Furthermore, the RATC also varied greatly in COVID-19, Recovered and Healthy controls supporting the dynamic dysbiosis of microbiomes in the corresponding metagenomes, their genetic diversity, and selective pressure for the maintenance of antimicrobial resistance. The microbiomes of COVID-19 and Recovered patients' nasopharyngeal swabs were enriched with genes coding for biofilm formation in *Staphylococcus*, quorum sensing: autoinducer-2 synthesis, *mdtABCD*, cobalt-zinc-cadmium, acriflavine, arsenic, fluoroquinolones and methicillin resistance, BlaR1 regulatory family, *mac*A, MAR locus, *pmr*A, *YjgK* cluster, and *lsrACDBFGE* operon. Conversely, beta-lactamase resistance, quorum sensing in *Yersinia* and *Pseudomonas*, and teicoplanin-resistance in *Staphylococcus* encoding genes remained predominantly abundant in

nasopharyngeal microbiomes of the Healthy controls. Remarkably, since the beginning of the COVID-19 pandemic, there has been growing concern for a potential rise in AMR secondary to increased antibiotic prescription for COVID-19 patients [41]. Moreover, severe COVID-19, which particularly affects elderly patients with multiple comorbidities, may be an important factor in determining changes in colonization pressure and multidrug resistance [42].

The functional analysis of the microbiomes revealed that genes coding for pyruvate carboxylase (*pyc*), adherent junction, tight junction, environmental information processing, carbohydrates metabolism, and oxidative phosphorylation had higher relative abundances in the COVID-19 associated nasopharyngeal microbiomes compared to those of Recovered and Healthy control microbiomes. These metabolic functional changes further lead to increased cytokine-cytokine receptor interaction, regulation of oxidative stress response, phage integration and excision, toxin-antitoxin regulation, protection from reactive oxygen species, and phage regulated gene expression in COVID-19 associated microbiomes as also reported previously in other viral diseases [43, 44]. The Recovered microbiomes were enriched with metabolic genes coding for focal adhesion, transport and catabolism, cell adhesion molecules, genetic information and processing, lysosome activity, endocytosis, cell-to-cell communication and glycolysis. These genomic potentials of the COVID-19 and Recovered nasopharyngeal microbiomes were also evidenced by the higher expression of genes involved in glutathione: non-redox reactions, redox cycle, coagulase cascade, osmotic stress, membrane transport, MT1-MMP pericellular network, and BarA-UvrY (*SirA*) two-component regulatory activities. Conversely, the Healthy nasopharyngeal microbiomes showed a lower abundance and expression of these metabolic functional genes. The metabolic health of an individual is represented by the proper functioning of organismal metabolic processes coordinated by multiple physiological systems [44]. The differentially abundant functions and pathways identified in this study corroborates with the findings from previous reports, relating to decreased lipid and glycan metabolism, increased carbohydrate metabolism, and other characteristics of the microbiome linked to COVID-19 [45]. Several predicted functional pathways differed between COVID-19 and Healthy controls, perhaps reflecting metabolic changes associated with the progression of COVID-19 pathogenesis, and novel host-microbiome interactions in SARS-CoV-2 infected patients.

However as of yet, no study has identified the microbial species that interact with SARS-CoV-2, one hypothesis regarding microbiome interactions with SARS-CoV-2 is relevant to the microbiomes' impacts on cytokines[28]. SARS-CoV-2 induces excessive and prolonged inflammatory cytokine storm, one of the potential causes of extrapulmonary tissue damage [46] and subsequent NT microbiome dysbiosis. Therefore, dysregulated and/or exaggerated cytokine responses and NT tissue damage could play an important role in creating a favorable environment for the NT commensal microbiota to become potential opportunists to further aggravating the disease condition. Medical interventions that the COVID-19 patients received might also be associated with microbiome modulations by interacting with the intestinal and NT microbiota. Previous experience from SARS and MERS shows that antiviral therapy can reduce viral load at in the early stages of the disease[46-48]. Previous studies showed that over expression of IL-18 and cytokine activities were found to be strongly correlating with several bacteria genera such as *Peptostreptococcus*, *Fusobacterium* and *Citrobacter*[33, 34].

Conclusions

Host-microbiomes interactions are exceptionally complex in the case of SARS-CoV-2 infections. Our RNA-Seq metagenomic analyses revealed that SARS-CoV-2 infection had significant effect on the diversity and composition of nasopharyngeal microbiomes of human. The identifiable changes in the microbiome diversity, composition and associated genomic features demonstrated in this study might be associated with the development, treatment, and resolution of COVID-19 disease. The SARS-CoV-2 infection results in remarkable depletion of nasopharyngeal

commensal microbiomes of Healthy human with inclusion of different opportunistic pathogens (bacteria and archaea) in COVID-19 and Recovered samples. Several predicted functional pathways differed between COVID-19 and COVID-19 recovered patients compared to Healthy individuals which reflect the metabolic changes associated with the progression of SARS-CoV-2 pathogenesis. These interactions are further complicated by the common co-existence of bacteria and archaea that interact with both host and SARS-CoV-2. These findings may serve as potential for microbiome-based diagnostic markers and therapeutics for this pandemic disease. However, future time-course studies with a larger sample size are needed to elucidate the dynamic changes in composition and diversity of commensal microbiomes and the inclusion of pathobionts in the whole respiratory system and gut during the progression of COVID-19 caused by SARS-CoV-2 infection.

Methods

Subject recruitment and sample collection

This study was conducted in accordance with the guidelines of the Director General of Health Services (DGHS) of Bangladesh during May to July, 2020. The study participants provided written informed consent consistent with the experiment. Twenty-two (n=22) nasopharyngeal samples (including COVID-19 = 8, Recovered = 7, and Healthy = 7) were collected from Dhaka city of Bangladesh. The average age of the study people was 41.86 (range 22-72) years, and of them, 15 (68.18%) were male and 7 (38.82%) females (Data S1). Collected samples were preserved at − 20 °C until further use for RNA extraction and RT-qPCR assay. The RT-qPCR was performed for *ORF1ab* and *N* genes of SARS-CoV-2 using novel Coronavirus (2019-nCoV) Nucleic Acid Diagnostic Kit (PCR-Fluorescence Probing, Sansure Biotech Inc.) according to the manufacturer's instructions. Viral RNA was extracted using a PureLink viral RNA/DNA minikit (Thermo Fisher Scientific, USA). Thermal cycling was performed at 50 °C for 30 min for reverse transcription, followed by 95 °C for 1 min, and then 45 cycles of 95 °C for 15 s, 60 °C for 30 s on an Analytik-Jena qTOWER instrument (Analytik Jena, Germany).

RNA sequencing

We utilized total RNA-Seq approach for the metagenomics component of the study. The cDNA of all 22 samples was used to prepare paired-end libraries with the Nextera DNA Flex library preparation kit (Illumina, Inc., San Diego, CA) according to the manufacturer's instructions. Paired-end (2 x 150 bp reads) sequencing of the prepared library pool of the samples was performed using a NextSeq high throughput kit with an Illumina NextSeq 550 sequencer at the Genomic Research Laboratory, Bangladesh Council of Scientific and Industrial Research (BCSIR), Dhaka, Bangladesh.

Taxonomic mapping, classification, diversity and community analysis

The paired-end sequences of COVID-19, Recovered and Healthy samples (n = 22) were analyzed using both mapping-based and assembly-based hybrid methods of PathoScope 2.0 (PS) [49] and MG-RAST (release version 4.1) (MR) [50]. The sequencing reads were filtered through BBDuk (with options k = 21, mink = 6, ktrim = r, ftm = 5, qtrim = rl, trimq = 20, minlen = 30, overwrite = true) to remove Illumina adapters, known Illumina artifacts, and phiX. Any sequence below these thresholds or reads containing more than one 'N' were discarded [27]. In the current study, 120.50 million reads were generated from the samples of three metagenomes with an average of 5.0 million reads per sample passed the quality control steps, and of these quality reads, an average of 3.22 million aligned reads per sample mapped to the ribosomal (rRNA) reference gene libraries (Data S1). Alpha diversity (diversity within samples) was estimated using the Shannon and Simpson diversity indices. To visualize differences in bacterial diversity, a principal coordinate analysis (PCoA; at genus level) was performed based on the Bray-Curtis distance method. The

comparative Venn diagrams for unique and shared microbial taxa were produced by FunRich (released version 3.1.4) [51]. Taxonomic abundance was determined by applying the "Best Hit Classification" option using the NCBI database as a reference with the following settings: maximum *e*-value of $1x10^{-30}$; minimum identity of 95% for bacteria, and 60% for viruses and archaea using a minimum alignment length of 20 as the set parameters. Microbial taxa that were detected in one group of sample but not detected in rest of the two groups are denoted as solely (unique) associated microbiomes [27].

Functional profiling of the microbiomes

We performed the taxonomic functional classification through mapping the reads onto the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [52], and SEED subsystem identifiers [50] on the MR server using the partially modified set parameters (*e*-value cutoff: $1x10^{-30}$, min. % identity cutoff: 60%, and min. alignment length cutoff: 20) [53].

Statistical analysis

The non-parametric test Kruskal-Wallis rank sum test was used to evaluate differences in the relative percent abundance of taxa in COVID-19, Recovered and Healthy sequence data. The gene counts were normalized by dividing the number of gene hits to individual taxa/function by total number of gene hits in each metagenome to remove bias due to differences in sequencing efforts. In addition, statistical tests were also applied with non-parametric test Kruskal-Wallis rank sum test at different KEGG and SEED subsystems levels through IBM SPSS (SPSS, Version 23.0, IBM Corp., NY, USA) [27].

Declarations

Author contributions

M.N.H. conceived and designed the study, performed bioinformatics analysis, visualized figures, interpreted results and drafted the original manuscript. M.M.H.S. collected samples, extracted RNA and performed sequencing. M.S.R. curated the data and performed bioinformatics analysis. S.A., T.A.B., B.W., I.J., M.S.H., A.K.M.S., T.N., M.M.A.M., M.Y., A.K.G., E.O., M.S.U., M.A.H. and A.S.M.M. edited the manuscript. K.A.C., T.I. and M.S.K. conceived the study and critically reviewed the drafted manuscript.

Funding

This project is financed by the Ministry of Science and Technology, Government of the People's Republic of Bangladesh (Grant No. GOB 224125200).

Competing interests

The authors declare no competing interests.

Availability of data and materials

All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. The sequence data reported in this article has been deposited in the National Center for Biotechnology Information (NCBI) under BioProject accession number PRJNA720904. Supplementary materials supporting the results of the study are available in this article as Fig. S1, Table S1, Data S1, Data S2 and Data S3.

Ethical statement

The protocol for sample collection from COVID-19, Recovered and Healthy humans, sample processing, transport, and RNA extraction was approved by the National Institute of Laboratory Medicine and Referral Center of Bangladesh.

Acknowledgements

This project is financed by the Ministry of Science and Technology, Government of the People's Republic of Bangladesh. We are thankful to those who provided us the samples.

Abbreviations

NT: Nasopharyngeal tract

SARS-CoV-2: Severe acute respiratory syndrome coronavirus-2

PCoA: Principal coordinate analysis

PS: PathoScope

MR: MG-RAST

RATC: Resistance to antibiotics and toxic compounds

References

1. He X, Lau EH, Wu P, Deng X, Wang J, Hao X, Lau YC, Wong JY, Guan Y, Tan X: **Temporal dynamics in viral shedding and transmissibility of COVID-19.***Nature medicine* 2020, **26:**672-675.
2. Hoque MN, Chaudhury A, Akanda MAM, Hossain MA, Islam MT: **Genomic diversity and evolution, diagnosis, prevention, and therapeutics of the pandemic COVID-19 disease.***PeerJ* 2020, **8:**e9689.
3. Reid AH, Fanning TG, Hultin JV, Taubenberger JK: **Origin and evolution of the 1918 "Spanish" influenza virus hemagglutinin gene.***Proceedings of the National Academy of Sciences* 1999, **96:**1651-1656.
4. Gallo O, Locatello LG, Mazzoni A, Novelli L, Annunziato F: **The central role of the nasal microenvironment in the transmission, modulation, and clinical progression of SARS-CoV-2 infection.***Mucosal immunology* 2020**:**1-12.
5. Yeoh YK, Zuo T, Lui GC-Y, Zhang F, Liu Q, Li AY, Chung AC, Cheung CP, Tso EY, Fung KS: **Gut microbiota composition reflects disease severity and dysfunctional immune responses in patients with COVID-19.***Gut* 2021, **70:**698-706.
6. Mostafa HH, Fissel JA, Fanelli B, Bergman Y, Gniazdowski V, Dadlani M, Carroll KC, Colwell RR, Simner PJ: **Metagenomic next-generation sequencing of nasopharyngeal specimens collected from confirmed and suspect covid-19 patients.***Mbio* 2020, **11**.
7. Kaul D, Rathnasinghe R, Ferres M, Tan GS, Barrera A, Pickett BE, Methe BA, Das SR, Budnik I, Halpin RA: **Microbiome disturbance and resilience dynamics of the upper respiratory tract during influenza A virus infection.***Nature communications* 2020, **11:**1-12.
8. Ren L, Wang Y, Zhong J, Zhang D, Xiao Y, Yang J, Fan G, Guo L, Shen Z, Liu W: **Dynamics of the upper respiratory tract microbiota and its association with fatality in COVID-19 patients.** 2020.

9. Dumas A, Bernard L, Poquet Y, Lugo-Villarino G, Neyrolles O: **The role of the lung microbiota and the gut–lung axis in respiratory infectious diseases.***Cellular microbiology* 2018, **20:**e12966.
10. Mirzaei R, Goodarzi P, Asadi M, Soltani A, Aljanabi HAA, Jeda AS, Dashtbin S, Jalalifar S, Mohammadzadeh R, Teimoori A: **Bacterial co-infections with SARS-CoV-2.***IUBMB life* 2020, **72:**2097-2111.
11. Quah J, Jiang B, Tan PC, Siau C, Tan TY: **Impact of microbial Aetiology on mortality in severe community-acquired pneumonia.***BMC infectious diseases* 2018, **18:**1-9.
12. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, Qiu Y, Wang J, Liu Y, Wei Y: **Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study.***The lancet* 2020, **395:**507-513.
13. Hoque MN, Rahman MS, Ahmed R, Hossain MS, Islam MS, Crandall KA, Islam T, Hossain MA, Siddiki AZ: **Diversity and genomic determinants of the microbiomes associated with COVID-19 and non-COVID respiratory diseases.***bioRxiv* 2020.
14. Zhou F, Yu T, Du R, Fan G, Liu Y, Liu Z, Xiang J, Wang Y, Song B, Gu X: **Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study.***The lancet* 2020.
15. Kumpitsch C, Koskinen K, Schöpf V, Moissl-Eichinger C: **The microbiome of the upper respiratory tract in health and disease.***BMC biology* 2019, **17:**1-20.
16. Wypych TP, Wickramasinghe LC, Marsland BJ: **The influence of the microbiome on respiratory health.***Nature immunology* 2019, **20:**1279-1290.
17. Mathieu E, Escribano-Vazquez U, Descamps D, Cherbuy C, Langella P, Riffault S, Remot A, Thomas M: **Paradigms of lung microbiota functions in health and disease, particularly, in asthma.***Frontiers in physiology* 2018, **9:**1168.
18. Dhar D, Mohanty A: **Gut microbiota and Covid-19-possible link and implications.***Virus research* 2020**:**198018.
19. Sencio V, Machado MG, Trottein F: **The lung–gut axis during viral respiratory infections: The impact of gut dysbiosis on secondary disease outcomes.***Mucosal Immunology* 2021**:**1-9.
20. Hoque MN, Istiaq A, Rahman MS, Islam MR, Anwar A, Siddiki AZ, Sultana M, Crandall KA, Hossain MA: **Microbiome dynamics and genomic determinants of bovine mastitis.***Genomics* 2020, **112:**5188-5203.
21. Zuo T, Liu Q, Zhang F, Lui GC-Y, Tso EY, Yeoh YK, Chen Z, Boon SS, Chan FK, Chan PK: **Depicting SARS-CoV-2 faecal viral activity in association with gut microbiota composition in patients with COVID-19.***Gut* 2021, **70:**276-284.
22. Bao L, Zhang C, Dong J, Zhao L, Li Y, Sun J: **Oral microbiome and SARS-CoV-2: Beware of lung co-infection.***Frontiers in Microbiology* 2020, **11:**1840.
23. Segal LN, Alekseyenko AV, Clemente JC, Kulkarni R, Wu B, Chen H, Berger KI, Goldring RM, Rom WN, Blaser MJ: **Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation.***Microbiome* 2013, **1:**1-12.
24. Pragman AA, Lyu T, Baller JA, Gould TJ, Kelly RF, Reilly CS, Isaacson RE, Wendt CH: **The lung tissue microbiota of mild and moderate chronic obstructive pulmonary disease.***Microbiome* 2018, **6:**1-19.
25. Zeppa SD, Agostini D, Piccoli G, Stocchi V, Sestili P: **Gut Microbiota Status in COVID-19: An Unrecognized Player?** *Frontiers in cellular and infection microbiology* 2020, **10**.
26. Scepanovic P, Hodel F, Mondot S, Partula V, Byrd A, Hammer C, Alanio C, Bergstedt J, Patin E, Touvier M: **A comprehensive assessment of demographic, environmental, and host genetic associations with gut microbiome diversity in healthy individuals.***Microbiome* 2019, **7:**1-15.

27. Hoque MN, Istiaq A, Clement RA, Sultana M, Crandall KA, Siddiki AZ, Hossain MA: **Metagenomic deep sequencing reveals association of microbiome signature with functional biases in bovine mastitis.***Scientific reports* 2019, **9:**1-14.
28. Kalantar-Zadeh K, Ward SA, Kalantar-Zadeh K, El-Omar EM: **Considering the effects of microbiome and diet on SARS-CoV-2 infection: nanotechnology roles.***ACS nano* 2020, **14:**5179-5182.
29. de Steenhuijsen Piters WA, Huijskens EG, Wyllie AL, Biesbroek G, Van Den Bergh MR, Veenhoven RH, Wang X, Trzciński K, Bonten MJ, Rossen JW: **Dysbiosis of upper respiratory tract microbiota in elderly pneumonia patients.***The ISME journal* 2016, **10:**97-108.
30. Man WH, de Steenhuijsen Piters WA, Bogaert D: **The microbiota of the respiratory tract: gatekeeper to respiratory health.***Nature Reviews Microbiology* 2017, **15:**259-270.
31. Rowe HM, Meliopoulos VA, Iverson A, Bomme P, Schultz-Cherry S, Rosch JW: **Direct interactions with influenza promote bacterial adherence during respiratory infections.***Nature microbiology* 2019, **4:**1328-1336.
32. Petersen E, Koopmans M, Go U, Hamer DH, Petrosillo N, Castelli F, Storgaard M, Al Khalili S, Simonsen L: **Comparing SARS-CoV-2 with SARS-CoV and influenza pandemics.***The Lancet infectious diseases* 2020.
33. Guo M, Tao W, Flavell RA, Zhu S: **Potential intestinal infection and faecal–oral transmission of SARS-CoV-2.***Nature Reviews Gastroenterology & Hepatology* 2021**:**1-15.
34. Yildiz S, Mazel-Sanchez B, Kandasamy M, Manicassamy B, Schmolke M: **Influenza A virus infection impacts systemic microbiota dynamics and causes quantitative enteric dysbiosis.***Microbiome* 2018, **6:**1-17.
35. Janda L, Mihalčin M, Šťastná M: **Is a healthy microbiome responsible for lower mortality in COVID-19?***Biologia* 2021, **76:**819-829.
36. Wu X, Cai Y, Huang X, Yu X, Zhao L, Wang F, Li Q, Gu S, Xu T, Li Y: **Co-infection with SARS-CoV-2 and influenza A virus in patient with pneumonia, China.***Emerging infectious diseases* 2020, **26:**1324.
37. Calcagno A, Ghisetti V, Burdino E, Trunfio M, Allice T, Boglione L, Bonora S, Di Perri G: **Co-infection with other respiratory pathogens in COVID-19 patients.***Clinical Microbiology and Infection* 2021, **27:**297-298.
38. Singh B, Kaur P, Reid R-J, Shamoon F, Bikkina M: **COVID-19 and Influenza Co-Infection: Report of Three Cases.***Cureus* 2020, **12**.
39. Chaung J, Chan D, Pada S, Tambyah PA: **Coinfection with COVID-19 and coronavirus HKU1—the critical need for repeat testing if clinically indicated.***Journal of medical virology* 2020, **92:**1785-1786.
40. Eckburg PB, Lepp PW, Relman DA: **Archaea and their potential role in human disease.***Infection and immunity* 2003, **71:**591-596.
41. Rawson TM, Ming D, Ahmad R, Moore LS, Holmes AH: **Antimicrobial use, drug-resistant infections and COVID-19.***Nature Reviews Microbiology* 2020, **18:**409-410.
42. Donà D, Di Chiara C, Sharland M: **Multi-drug-resistant infections in the COVID-19 era: a framework for considering the potential impact.***Journal of Hospital Infection* 2020, **106:**198-199.
43. Wu J, Wang K, Wang X, Pang Y, Jiang C: **The role of the gut microbiome and its metabolites in metabolic diseases.***Protein & Cell* 2020**:**1-14.
44. Ayres JS: **A metabolic handbook for the COVID-19 pandemic.***Nature metabolism* 2020, **2:**572-585.
45. Breiman A, Ruvën-Clouet N, Le Pendu J: **Harnessing the natural anti-glycan immune response to limit the transmission of enveloped viruses such as SARS-CoV-2.***PLoS Pathogens* 2020, **16:**e1008556.
46. Ye Q, Wang B, Mao J: **The pathogenesis and treatment of theCytokine Storm'in COVID-19.***Journal of infection* 2020, **80:**607-613.

47. Falzarano D, De Wit E, Rasmussen AL, Feldmann F, Okumura A, Scott DP, Brining D, Bushmaker T, Martellaro C, Baseler L: **Treatment with interferon-α2b and ribavirin improves outcome in MERS-CoV–infected rhesus macaques.***Nature medicine* 2013, **19:**1313-1317.
48. Stockman LJ, Bellamy R, Garner P: **SARS: systematic review of treatment effects.***PLoS med* 2006, **3:**e343.
49. Hong C, Manimaran S, Shen Y, Perez-Rogers JF, Byrd AL, Castro-Nallar E, Crandall KA, Johnson WEJM: **PathoScope 2.0: a complete computational framework for strain identification in environmental or clinical sequencing samples.** 2014, **2:**1-15.
50. Glass EM, Wilkening J, Wilke A, Antonopoulos D, Meyer F: **Using the metagenomics RAST server (MG-RAST) for analyzing shotgun metagenomes.***Cold Spring Harbor Protocols* 2010, **2010:**pdb. prot5368.
51. Fonseka P, Pathan M, Chitti SV, Kang T, Mathivanan S: **FunRich enables enrichment analysis of OMICs datasets.***Journal of Molecular Biology* 2020**:**166747.
52. Kanehisa M, Sato Y, Furumichi M, Morishima K, Tanabe M: **New approach for understanding genome variations in KEGG.***Nucleic acids research* 2019, **47:**D590-D595.
53. Hoque MN, Istiaq A, Clement RA, Gibson KM, Saha O, Islam OK, Abir RA, Sultana M, Siddiki A, Crandall KA: **Insights into the resistome of bovine clinical mastitis microbiome, a key factor in disease complication.***Frontiers in Microbiology* 2020, **11:**860.
54. Santacroce L, Charitos IA, Ballini A, Inchingolo F, Luperto P, De Nitto E, Topi S: **The human respiratory system and its microbiome at a glimpse.***Biology* 2020, **9:**318.
55. Huttenhower C, Gevers D, Knight R, Abubucker S, Badger JH, Chinwalla AT, Creasy HH, Earl AM, FitzGerald MG, Fulton RS: **Structure, function and diversity of the healthy human microbiome.***nature* 2012, **486:**207.
56. Gomes S, Cavadas B, Ferreira JC, Marques PI, Monteiro C, Sucena M, Sousa C, Rodrigues LV, Teixeira G, Pinto P: **Profiling of lung microbiota discloses differences in adenocarcinoma and squamous cell carcinoma.***Scientific reports* 2019, **9:**1-11.
57. Igartua C, Davenport ER, Gilad Y, Nicolae DL, Pinto J, Ober C: **Host genetic variation in mucosal immunity pathways influences the upper airway microbiome.***Microbiome* 2017, **5:**1-17.
58. Ryan MP, Pembroke JT: **Brevundimonas spp: Emerging global opportunistic pathogens.***Virulence* 2018, **9:**480-493.
59. O'hara C, Steigerwalt A, Hill B, Farmer J, Fanning G, Brenner D: **Enterobacter hormaechei, a new species of the family Enterobacteriaceae formerly known as enteric group 75.***Journal of clinical microbiology* 1989, **27:**2046-2049.
60. Ma T, Fu J, Xie N, Ma S, Lei L, Zhai W, Shen Y, Sun C, Wang S, Shen Z: **Fitness cost of blaNDM-5-carrying p3R-IncX3 plasmids in wild-type NDM-free Enterobacteriaceae.***Microorganisms* 2020, **8:**377.
61. Gautam L, Kaur R, Kumar S, Bansal A, Gautam V, Singh M, Ray P: **Pseudomonas oleovorans sepsis in a child: the first reported case in India.***Japanese journal of infectious diseases* 2015**:**JJID. 2014.2174.
62. Agard MJ, Ozer EA, Morris AR, Piseaux R, Hauser AR: **A genomic approach to identify Klebsiella pneumoniae and Acinetobacter baumannii strains with enhanced competitive fitness in the lungs during multistrain pneumonia.***Infection and immunity* 2019, **87**.
63. Renom F, Yáñez A, Garau M, Rubí M, Centeno M-J, Gorriz M-T, Medinas M, Ramis F, Soriano JB, Agustí À: **Prognosis of COPD patients requiring frequent hospitalization: role of airway infection.***Respiratory medicine* 2010, **104:**840-848.
64. Bokhari S, Abbas N, Singh M, Cindrich RB, Zeana C: **Empedobacter brevis bacteremia in a patient infected with HIV: case report and review of literature.***Case reports in infectious diseases* 2015, **2015**.

65. Kaci G, Goudercourt D, Dennin V, Pot B, Doré J, Ehrlich SD, Renault P, Blottière HM, Daniel C, Delorme C: **Anti-inflammatory properties of Streptococcus salivarius, a commensal bacterium of the oral cavity and digestive tract.***Applied and environmental microbiology* 2014, **80:**928-934.
66. Hols P, Ledesma-García L, Gabant P, Mignolet J: **Mobilization of microbiota commensals and their bacteriocins for therapeutics.***Trends in microbiology* 2019, **27:**690-702.
67. Engen SA, Rørvik GH, Schreurs O, Blix IJ, Schenck K: **The oral commensal Streptococcus mitis activates the aryl hydrocarbon receptor in human oral epithelial cells.***International journal of oral science* 2017, **9:**145-150.
68. van den Munckhof EH, Hafkamp HC, de Kluijver J, Kuijper EJ, de Koning MN, Quint WG, Knetsch CW: **Nasal microbiota dominated by Moraxella spp. is associated with respiratory health in the elderly population: a case control study.***Respiratory Research* 2020, **21:**1-9.
69. Faner R, Sibila O, Agustí A, Bernasconi E, Chalmers JD, Huffnagle GB, Manichanh C, Molyneaux PL, Paredes R, Brocal VP: **The microbiome in respiratory medicine: current challenges and future perspectives.***European Respiratory Journal* 2017, **49**.
70. Liu G, Tang CM, Exley RM: **Non-pathogenic Neisseria: members of an abundant, multi-habitat, diverse genus.***Microbiology* 2015, **161:**1297-1312.
71. Weyand NJ: **Neisseria models of infection and persistence in the upper respiratory tract.***Pathogens and disease* 2017, **75**.
72. Ibironke O, McGuinness LR, Lu S-E, Wang Y, Hussain S, Weisel CP, Kerkhof LJ: **Species-level evaluation of the human respiratory microbiome.***GigaScience* 2020, **9:**giaa038.
73. Weiser JN, Ferreira DM, Paton JC: **Streptococcus pneumoniae: transmission, colonization and invasion.***Nature Reviews Microbiology* 2018, **16:**355-367.
74. Kosikowska U, Biernasiuk A, Rybojad P, Łoś R, Malm A: **Haemophilus parainfluenzae as a marker of the upper respiratory tract microbiota changes under the influence of preoperative prophylaxis with or without postoperative treatment in patients with lung cancer.***BMC microbiology* 2016, **16:**1-9.
75. Yamamoto S, Saito M, Tamura A, Prawisuda D, Mizutani T, Yotsuyanagi H: **The Human Microbiome and COVID-19: A Systematic Review.***Available at SSRN 3750675* 2021.
76. Bogaert D, de Groot R, Hermans P: **Streptococcus pneumoniae colonisation: the key to pneumococcal disease.***The Lancet infectious diseases* 2004, **4:**144-154.
77. Cameron D, Jiang J-H, Hassan K, Elbourne L, Tuck K, Paulsen I, Peleg A: **Insights on virulence from the complete genome of Staphylococcus capitis.***Frontiers in microbiology* 2015, **6:**980.
78. Chen H-W, Liu P-F, Liu Y-T, Kuo S, Zhang X-Q, Schooley RT, Rohde H, Gallo RL, Huang C-M: **Nasal commensal Staphylococcus epidermidis counteracts influenza virus.***Scientific reports* 2016, **6:**1-12.
79. Zhou Y, Lin F, Cui Z, Zhang X, Hu C, Shen T, Chen C, Zhang X, Guo X: **Correlation between either Cupriavidus or Porphyromonas and primary pulmonary tuberculosis found by analysing the microbiota in patients' bronchoalveolar lavage fluid.***PLoS One* 2015, **10:**e0124194.
80. Mora M, Wink L, Kögler I, Mahnert A, Rettberg P, Schwendner P, Demets R, Cockell C, Alekhova T, Klingl A: **Space Station conditions are selective but do not alter microbial characteristics relevant to human health.***Nature communications* 2019, **10:**1-18.
81. Kanafani Z, Kanj S: **Acinetobacter infection: Epidemiology, microbiology, pathogenesis, clinical features, and diagnosis.***Wolters Kluwer* 2013, **2:**21-33.
82. Krawczyk B, Michalik M, Fordon M, Wysocka M, Samet A, Nowicki B: **Escherichia coli Strains with Virulent Factors Typical for Uropathogens were Isolated from Sinuses from Patients with Chronic Rhinosinusitis—Case**

Report.*Pathogens* 2020, **9:**318.

83. Lambiase A, Rossano F, Del Pezzo M, Raia V, Sepe A, de Gregorio F, Catania MR: **Sphingobacterium respiratory tract infection in patients with cystic fibrosis.***BMC research notes* 2009, **2:**1-5.
84. Coenye T, Vandamme P, LiPuma JJ: **Infection by Ralstonia species in cystic fibrosis patients: identification of R. pickettii and R. mannitolilytica by polymerase chain reaction.***Emerging infectious diseases* 2002, **8:**692.

Tables

Table 1. Top abundant commensal and pathogenic (mostly opportunistic) nasopharyngeal bacteria identified in current and previous studies.

Bacterial species	Sample groups			Virulence status		Previous reports
	Healthy	COVID-19	Recovered	Commensal	Pathogenic	
Pseudomonas sp. LPH1	+	-	+	Y	Y	[54-56]
Rheinheimera sp. D18	+	+	+	NR	Y	[57]
Pseudomonas mendocina	+	+	+	Y	Y	[54-56]
Brevundimonas sp. Bb-A	+	-	+	Y	NR	[56, 58]
Enterobacter hormaechei	+	+	+	NR	Y	[59, 60]
Pseudomonas oleovorans	+	-	+	Y	Y	[56, 61]
Klebsiella pneumoniae	+	+	+	NR	Y	[62, 63]
Orbus sp. IPMB12	+	+	-	NR	NR	NR
Rheinheimera sp. LHK132	+	+	+	NR	Y	[57]
Pseudomonas sp. phDV1	+	-	-	Y	Y	[54]
Empedobacter brevis	+	+	+	NR	Y	[64]
Brevundimonas sp. DS20	+	-	+	NR	Y	[58]
Streptococcus salivarius	-	+	+	Y	NR	[65, 66]
Streptococcus mitis	-	+	+	Y	Y	[56, 67-69]
Neisseria subflava	-	+	+	Y	NR	[56, 70, 71]
Veillonella dispar	-	+	+	Y	Y	[56, 69, 72]
Veillonella parvula	-	+	+	Y	Y	[56, 69, 72]
Prevotella melaninogenica	-	+	+	Y	Y	[56, 69, 72]
Streptococcus parasanguinis	-	+	-	Y	Y	[56, 66, 68, 73]
Streptococcus sp. LPB0220	-	+	-	Y	Y	[56, 66, 69, 73]
Haemophilus parainfluenzae	+	+	+	Y	Y	[56, 68, 69, 74, 75]
Streptococcus pneumoniae	+	+	+	NR	Y	[56, 69, 73, 76]
Pseudomonas stutzeri	+	+	+	Y	Y	[54-56]
Staphylococcus capitis	-	+	+	NR	Y	[68, 77]
Staphylococcus epidermidis	+	+	+	Y	Y	[15, 68, 78]
Cupriavidus metallidurans	+	-	+	NR	Y	[79, 80]
Moraxella osloensis	+	+	+	NR	Y	[63, 68]
Acinetobacter indicus	+	+	+	Y	Y	[81]
Escherichia coli	+	+	+	Y	Y	[69, 82]
Sphingobacterium sp. G1-14	+	+	+	NR	Y	[83]
Acinetobacter junii	+	+	+	NR	Y	[81]

Ralstonia pickettii	-	+	+	NR	Y	[84]

Here: '+' refers to particular species present in the corresponding sample group while '-' refers to the absence, Y= yes, either commensal or opportunistic, NR= not reported.

Species	Phylum	Relative abundance (%)	
		COVID-19	Recovered
Streptococcus salivarius K12	*Firmicutes*	19.13	0.03
Streptococcus mitis	*Firmicutes*	18.13	0.11
Neisseria subflava	*Proteobacteria*	13.77	0.06
Veillonella dispar	*Firmicutes*	11.03	0.01
Acinetobacter junii 64.5	*Proteobacteria*	3.31	1.85
Veillonella parvula	*Firmicutes*	2.35	0.05
Prevotella melaninogenica	*Firmicutes*	2.26	0.08
Streptococcus parasanguinis	*Firmicutes*	2.22	0.0
Streptococcus sp. LPB0220	*Firmicutes*	1.98	0.02
Neisseria flavescens	*Proteobacteria*	1.26	0.0
Veillonella atypica	*Firmicutes*	1.10	0.02
Pseudomonas stutzeri DSM 4166	*Proteobacteria*	0.01	8.48
Staphylococcus capitis	*Firmicutes*	0.005	5.89
Staphylococcus epidermidis RP62A	*Firmicutes*	0.20	5.13
Pseudomonas mendocina NK-01	*Proteobacteria*	0.15	3.12
Moraxella osloensis A1920	*Proteobacteria*	0.13	2.60
Acinetobacter indicus A648	*Proteobacteria*	0.05	2.57
Escherichia coli	*Proteobacteria*	0.12	2.41
Sphingobacterium sp. G1-14	*Bacteroidetes*	0.043	1.96
Streptococcus pneumoniae	*Firmicutes*	0.005	1.60
Ralstonia pickettii	*Proteobacteria*	0.002	1.55
Micrococcus luteus	*Actinobacteria*	0.035	1.46

Table 2. Top abundant opportunistic bacterial species in COVID-19 and Recovered patients (not detected in Healthy controls).

Figures

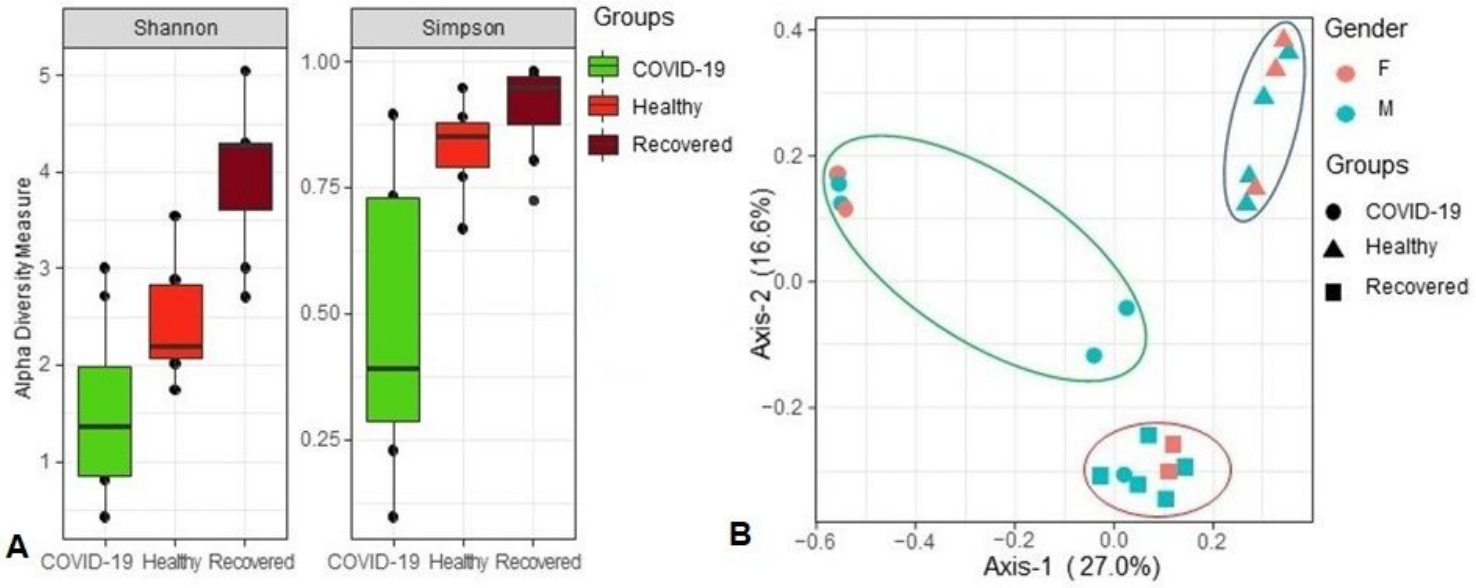

Figure 1

Differences in microbiome diversity and community structure in COVID-19, Recovered and Healthy nasopharyngeal sample metagenomes. A Box plots showing significantly higher microbial species richness in COVID-19 compared to Recovered and Healthy control samples (COVID-19, p = 0.0065; Recovered, p = 0.0105; Healthy, p = 0.0401; Wilcoxon rank sum test) (P=0.036, Kruskal-Wallis test) in both Shannon and Simpson estimated alpha diversity calculation. B Principal coordinates analysis (PCoA) measured on the Bray-Curtis distance method separated samples by microbial population structure. Each color represents an individual, and shapes indicate the population in three metagenomes. Statistical analysis using Kruskal–Wallis tests showed significant microbial diversity variations across the three metagenomes (p = 0.0059, Kruskal-Wallis test).

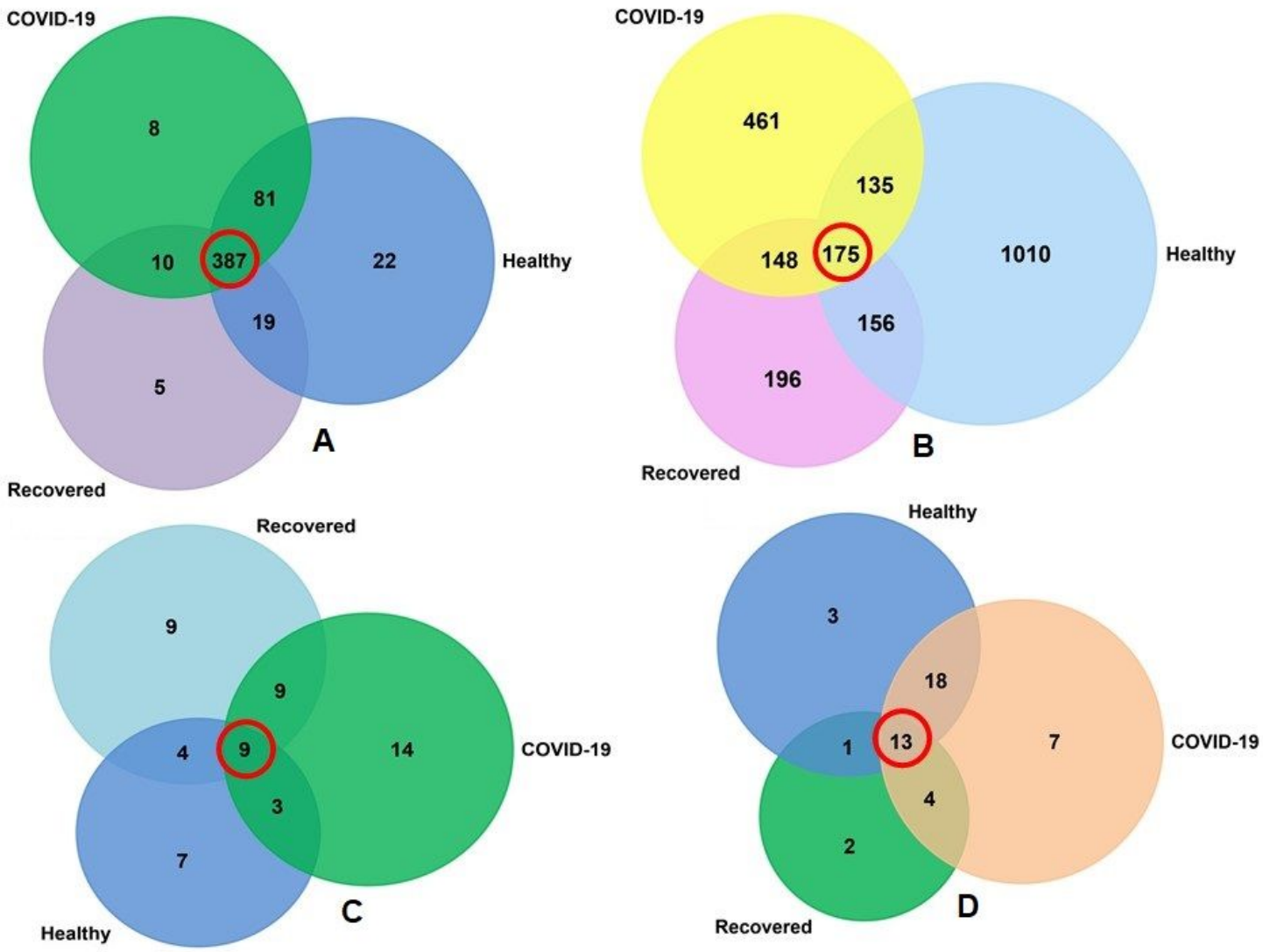

Figure 2

Taxonomic composition of COVID-19, Recovered and Healthy control metagenomes. Venn diagrams representing the core unique and shared microbiomes in COVID-19, Recovered and Healthy nasopharyngeal samples. A Venn diagram showing unique and shared bacterial genera. Out of 532 detected genera, only 387 genera (highlighted in red) were found to be shared across the given conditions. B Venn diagram comparison of unique and shared bacterial species where only 175 (highlighted in red) species shared among the conditions. C Venn diagrams representing unique and shared viral genera identified in three metagenomes. Of the detected archaeal genera (n=55), 14,9 and 7 genera had unique association with COVID-19, Recovered and Healthy samples, respectively, and 9 genera (highlighted in red) were found to be shared across three metagenomes. D Venn diagrams representing unique and shared archaeal genera identified in three metagenomes. Of the detected archaeal genera (n=48), 7, 2 and 3 genera had unique association with COVID-19, Recovered and Healthy samples, respectively, and 13 genera (highlighted in red) were found to be shared across three metagenomes. More information on the taxonomic result is also available in Data S1.

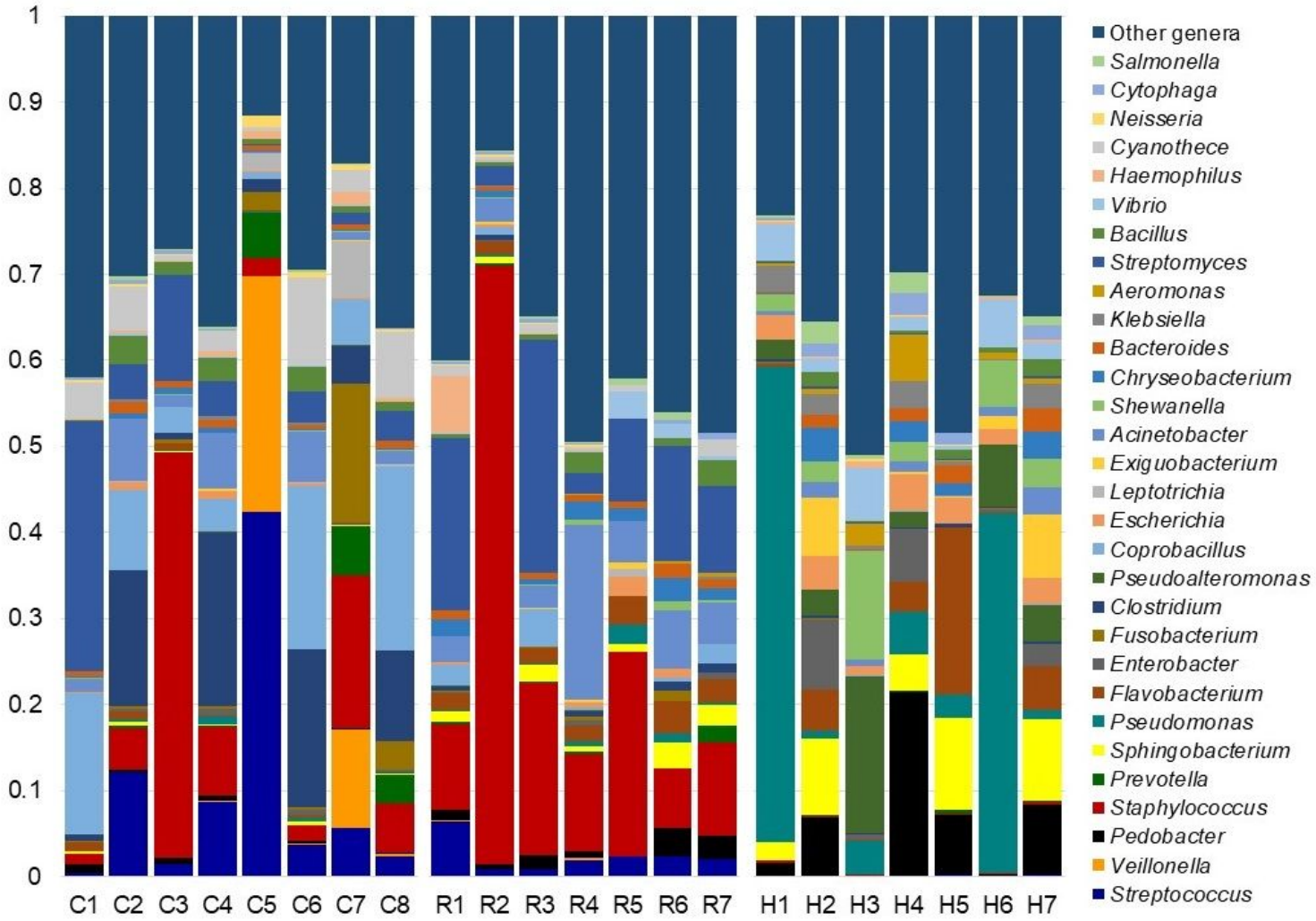

Figure 3

The genus level taxonomic profile of bacteria in COVID-19 (C1-C8), Recovered (R1-R7) and Healthy (H1-H7) nasopharyngeal samples. The relative abundance of 30 most abundant bacterial genera are sorted from bottom to top by their deceasing proportion, with the remaining genera keeping as 'Other genera'. Each stacked bar plot represents the abundance of bacterial genera in each sample of the corresponding category. Notable differences in bacterial populations are those where the taxon is abundant in COVID-19 and Recovered samples, and effectively undetected in the Healthy controls. The distribution and relative abundance of the bacterial genera in the study metagenomes are also available in Data S1 and S2.

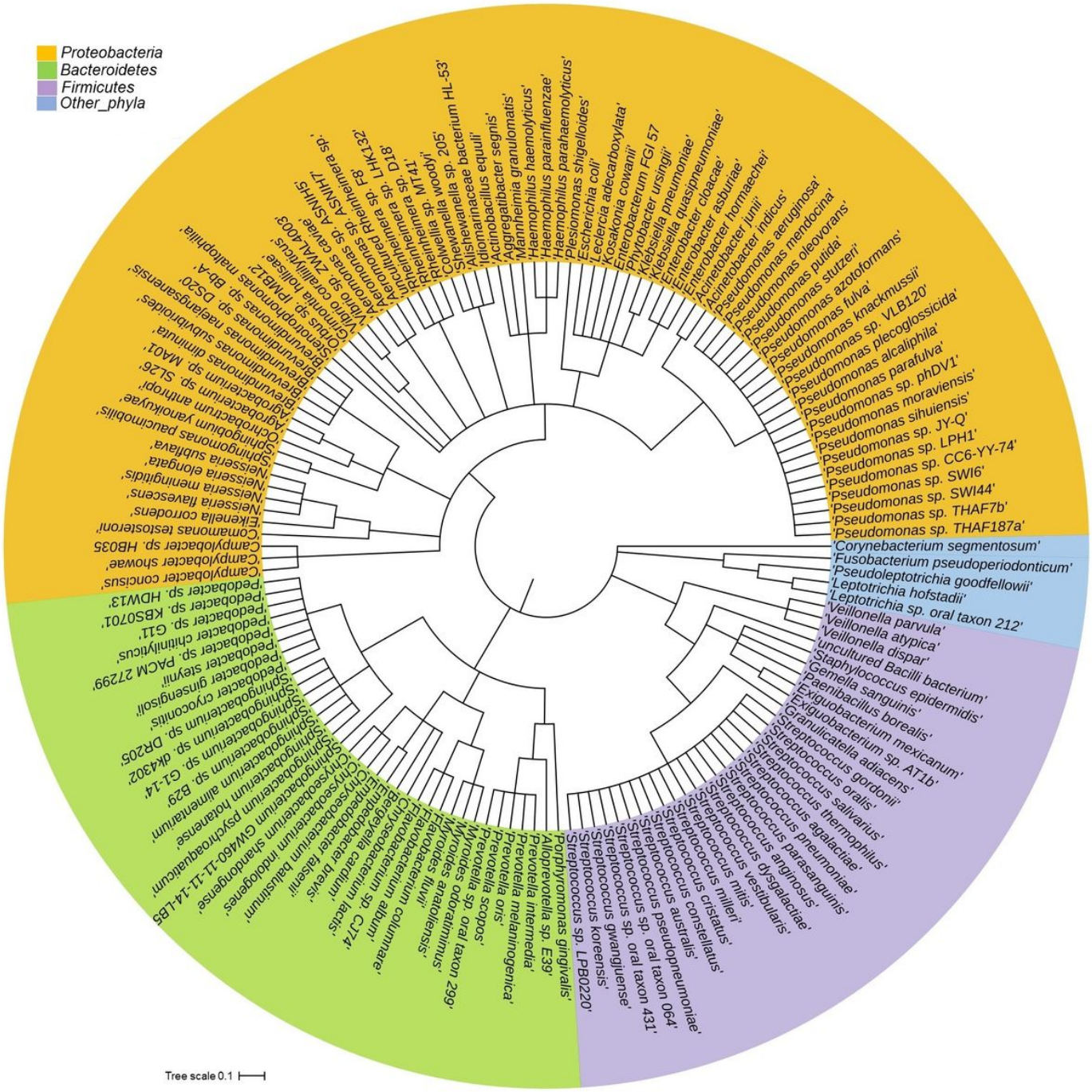

Figure 4

The species and/or strain level taxonomic representation of microbiome in COVID-19, Recovered and Healthy metagenomes. Taxonomic dendrogram in the midpoint rooted phylogenetic tree was generated with top abundant 200 species of bacteria found in three metagenomes. The tree was made based on the maximum likelihood method in Clustal W and displayed with iTOL (interactive Tree Of Life). Color ranges identify different strains within the tree. Strains and/or species are color-coded by different phyla of bacteria present in > 95% of samples. The strains in the phylogenetic tree are also available in Data S1 and S2.

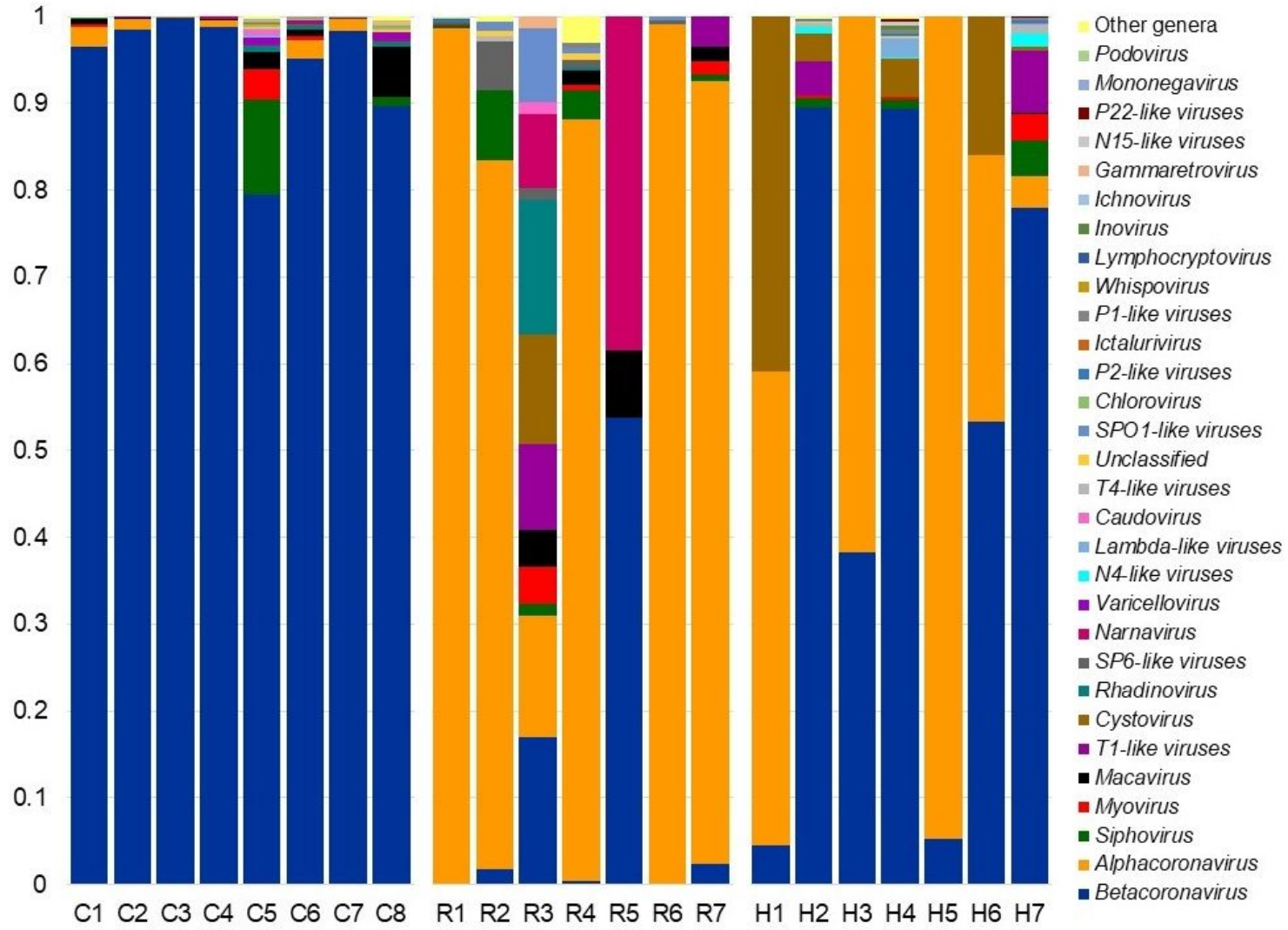

Figure 5

The genus level taxonomic profile of viruses in COVID-19 (C1-C8), Recovered (R1-R7) and Healthy (H1-H7) nasopharyngeal samples. Stacked bar plots showing the relative abundance and distribution of the 30 most top abundant viral genera, with ranks ordered from bottom to top by their decreasing proportion, with the remaining genera keeping as 'Other genera'. Each stacked bar plot represents the abundance of viral genera in each sample of the corresponding category. Notable differences in viral populations are those where the taxon is abundant in COVID-19 and Recovered samples, and effectively undetected in the Healthy controls. The distribution and relative abundance of the archaeal genera in the study metagenomes are also available in Data S1 and S2.

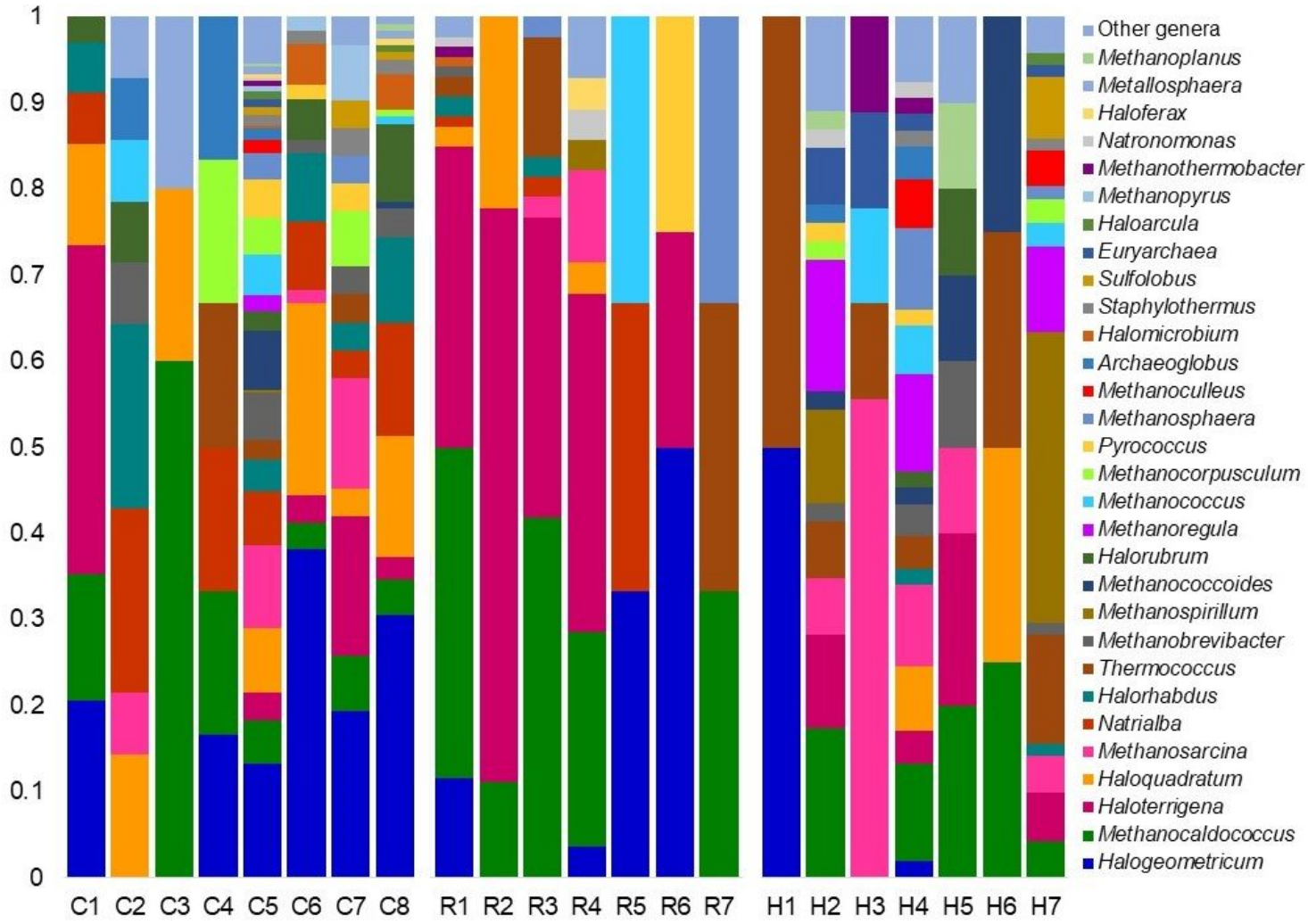

Figure 6

The genus level taxonomic profile of archaea in COVID-19 (C1-C8), Recovered (R1-R7) and Healthy (H1-H7) nasopharyngeal samples. Stacked bar plots showing the relative abundance and distribution of the 30 most top abundant archaeal genera, with ranks ordered from bottom to top by their decreasing proportion, with the remaining genera keeping as ‘Other genera’. Each stacked bar plot represents the abundance of archaeal genera in each sample of the corresponding category. Notable differences in archaeal populations are those where the taxon is abundant in COVID-19 and Recovered samples, and effectively undetected in the Healthy controls. The distribution and relative abundance of the archaeal genera in the study metagenomes are also available in Data S1 and S2.

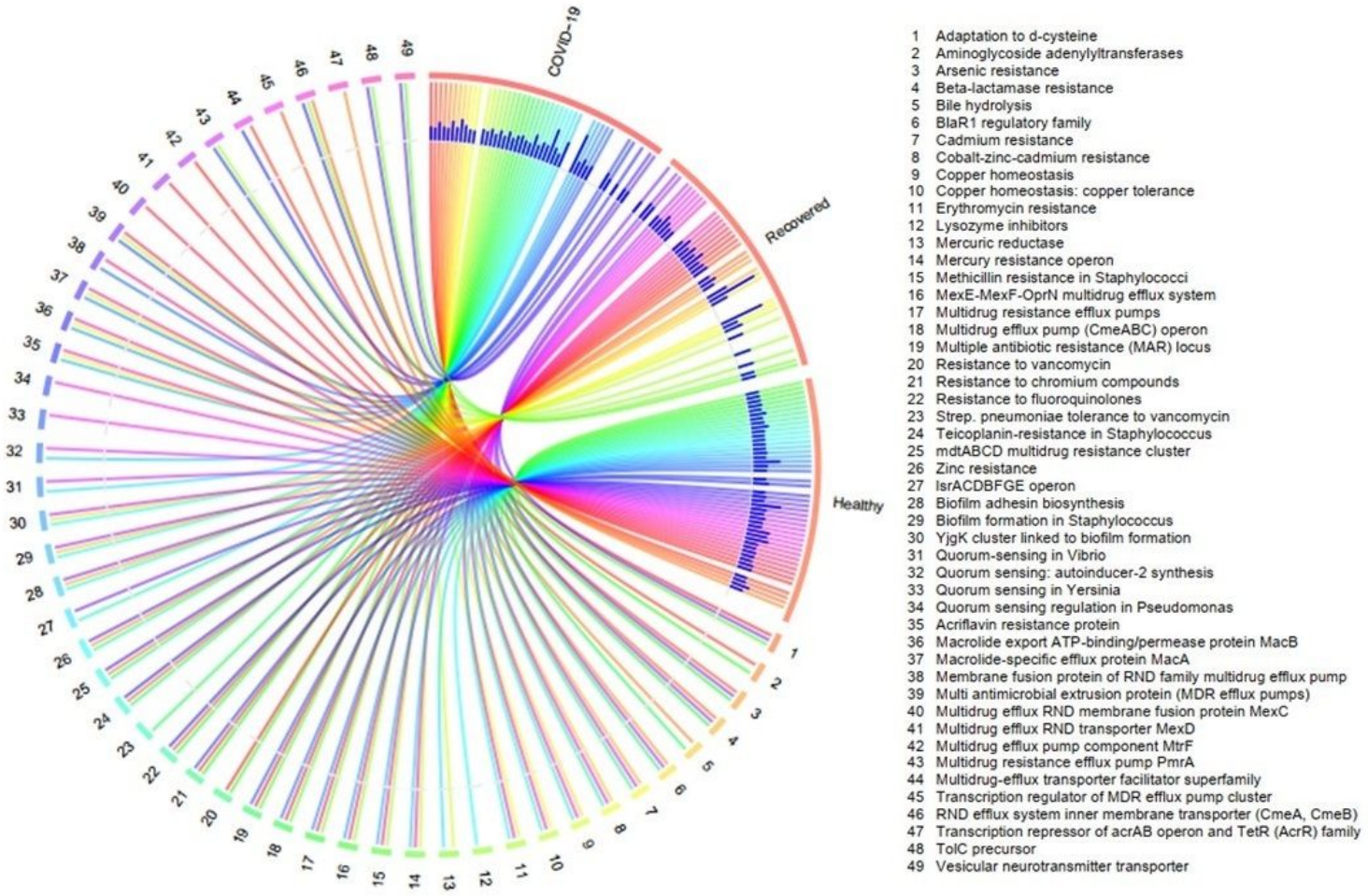

Figure 7

Distribution of the resistance to antibiotic and toxic compounds (RATC) genes in COVID-19, Recovered and Healthy metagenomes. The circular plot illustrates the diversity and relative abundance of the RATC genes detected among the microbiomes of the three metagenomes through SEED subsystems analysis. The association of the RATC genes according to metagenome is shown by different colored ribbons and the relative abundances these genes are represented by inner blue colored bars. Some of the RATC functional groups shared among microbes of the three metagenomes (COVID-19, Recovered and Healthy), and rest are effectively undetected in the microbiomes of the other metagenomes. The numerical values 1 to 49 represent different RATC functional groups found across the study metagenomes. More data on RATC functional groups and their relative abundances are also available in Data S3.

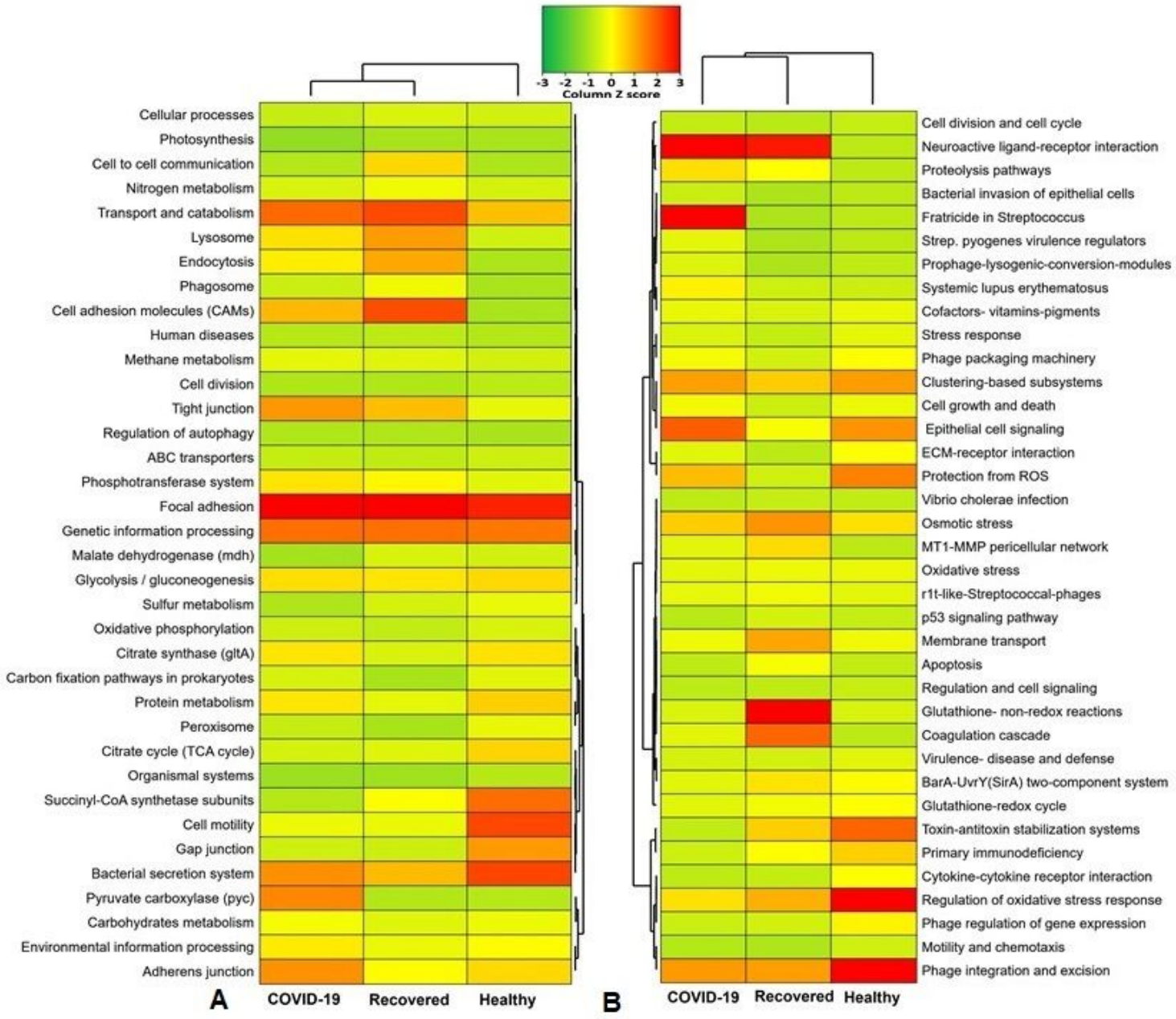

Figure 8

Functional annotation of the COVID-19, Recovered and Healthy nasopharyngeal sample related sequences. A Heatmap representing the average relative abundance hierarchical clustering of the predicted KEGG Orthologs (KOs) functional pathways of the microbiome across four metagenome groups. B Heatmap showing the average relative abundance hierarchical clustering of the predicted SEED functions in different levels among the microbiomes of four metagenomes. The color bars (column Z score) at the top represent the relative abundance of putative genes. The color codes indicate the presence and completeness of each KEGG and SEED module, expressed as a value between -3 (lowest abundance) and 3 (highest abundance). The red color indicates the more abundant patterns, whilst green cells account for less abundant putative genes in that particular metagenome.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- DataS1.xlsx
- DataS2.xlsx
- DataS3.xlsx
- Fig.S1.jpg
- FigureswithLegends.docx
- TableS1.docx