Microb Health Dis 2023; 5: e845

NOT TO BE SNIFFED AT: REPURPOSING EXISTING RNA-SEQ DATA TO EXAMINE CHANGES TO HUMAN NASAL MICROBIOME COMPOSITION IN COVID-19 INFECTION, COMPARED TO INFLUENZA AND HEALTHY CONTROLS

F. Wainwright[1], G. Misirli[1], P. Andras[2]

[1]Keele University School of Computing and Mathematics, The Covert, Keele, Newcastle under Lyme, UK
[2]Edinburgh Napier University School of Computing and School of Engineering and the Built Environment, Merchiston Campus, Edinburgh, UK

Corresponding Author: Felicity Wainwright, MSc; email: f.g.wainwright@keele.ac.uk

Abstract – *Objective:* The human microbiome is essential in maintaining healthy physiology; compositional changes have been implicated in numerous physical and mental diseases. Thus far, COVID-19 microbiome research has focused primarily upon gut and lung bacterial communities. However, the early stages of COVID-19 infection and immune response occur in the nasal epithelium. Therefore, investigating nasal microbiome changes in early-stage COVID-19 may yield key insights into the immune system mechanisms involved in progression from mild/no symptoms to systemic organ failure/death, why this occurs in certain individuals, and how it may be prevented with early warning.

Patients and Methods: Here we repurposed existing RNA-seq data to characterise the human nasal microbiome in COVID-19 infected samples and compared the taxonomic profile to healthy control and influenza-infected control samples, to identify COVID-19 specific nasal microbiome changes and attempt to rationalise these in the context of what is already known regarding mechanisms of the immune response to COVID-19.

Results: We demonstrate that existing RNA-seq reads from human nasal swabs can be repurposed to characterise the human nasal microbiome robustly and accurately in health, early-stage COVID-19, and influenza. We observe that nasal microbiome composition (presence and abundance of phyla, genera, and species) significantly differs between health and disease, and between COVID-19 and influenza.

Conclusions: Our observed healthy nasal microbial profiles match the findings of previous research, demonstrating that repurposing existing RNA-seq data is as accurate as targeted methods for taxonomic classification. We also observed many differential changes in the nasal microbiome profile to be disease specific. This will be key to enabling the potential for differential diagnosis based upon nasal microbiome profiles in the future.

Keywords: Computational biology, RNA-seq, Classification, Microbiota, COVID-19, Influenza, Human.

INTRODUCTION

Background

Research has thoroughly characterised the human microbiome and its role in health and disease[1]. Microbiome research in COVID-19 *vs.* health has primarily focused on gut and lung communities. There is a strong rationale for this in the gut-lung axis, where gut microbiome composition affects lung infection susceptibility, and respiratory infection can alter gut microbiome composition towards proinflammatory dysbiosis[2]. However, early-stage COVID-19 groundwork is laid in the nasal epithelium; SARS-CoV-2 enters the body *via* nasal epithelial cells, and infection begins with viral upper airway penetration; the highest viral load is found in nasal swabs[3]. Disease escalation results from a loss of immune regulation between protective and altered responses due to an exacerbation of inflammatory components[4]. Identifying indicators of this during early-stage disease would be beneficial so that preventative de-escalation measures could be taken.

We propose that investigating nasal microbiome changes in early-stage COVID-19 will generate key insight into microbiome and immune system interplay, how this influences disease progression, why it occurs in certain patients, and whether it can be prevented.

To date, limited COVID-19 nasal microbiome research has been conducted. Targeted 16S rRNA-seq methods have observed that nasal/oropharyngeal (NOP) microbiome composition in 21 COVID-19 patients (paucisymptomatic or admitted to ICU) was altered compared to 10 negative control participants and 8 participants with a different human coronavirus (HKU/NL63/OC43). Several taxa were differentially expressed in disease *vs.* health. The phylum *Deinococcus-Thermus* was only observed in health. *Candidatus Saccharibacteria* was significantly decreased in ICU COVID-19 patients compared to paucisymptomatic patients and healthy controls. A complete depletion of the genera *Bifidobacterium* and *Clostridium* was observed in ICU patients, while *Salmonella*, *Scardovia*, *Serratia* and *Pectobacteriaceae* were observed only in this condition[5].

The rationale for examining the nasal microbiome in COVID-19 is clear; nasal cavity samples are more easily obtained than lung samples, benefiting both researchers and participants. Furthermore, the localised initial immune response (pre-systemic disease), means that nasal swab data enables investigation into microbiome changes in the earliest detectable disease stages.

Traditionally, targeted microbiome characterisation methods (microbial culture, 16S amplicon sequencing etc.) have been used. However, by repurposing existing RNA-sequencing (RNA-seq) data deposited in online databases such as the Sequence Read Archive (SRA), the human microbiome can be easily and inexpensively characterised as accurately as with targeted methods. Source studies typically focus on host reads, with the unmapped fraction considered a by-product and discarded, but research has begun to demonstrate RNA-seq data repurposing for microbiome characterisation[6].

Here, we characterised the nasal microbiome in COVID-19, health, and influenza by reusing existing RNA-seq data. We have compared COVID-19 nasal samples to those from healthy controls, and from influenza patients. Including influenza data enabled us to ascertain whether compositional deviations from the healthy profile are COVID-19-specific or are part of a more generalised inflammatory response.

Microbial communities were successfully identified using existing RNA-seq data, following read sample quality assessment and mapping to the human reference genome. We taxonomically classified the unmapped reads to characterise the microbiome in each state at phylum, genus, and species levels. As expected, we observed a distinct and characteristic taxonomic profile in each state, and our most abundant taxa in the healthy samples are 'expected' according to previous nasal microbiome studies.

The COVID-19 Pandemic

The COVID-19 pandemic began in Wuhan, China, in December 2019. It has since spread to ~200 countries and had a devastating effect due to its contagiousness and the high proportion of patients requiring ICU care. Viral RNA is detectable in infected individuals 1-3 days before symptom onset; viral load peaks at symptom onset, gradually decreasing over 1-2 weeks[7]. Clinical manifestation ranges from an unnoticed asymptomatic infection to severe pneumonia, multiple organ failure, and death[3]. The global death toll has exceeded 6.4 million people.

The Immune System and COVID-19

COVID-19 is caused by Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2), an enveloped RNA virus (diameter 60-140 nm), spherical with a fatty outer layer and a surface crown ("corona") of club-shaped spikes with a large binding surface area. The spike S1 protein has high affinity with the N-terminal helix of angiotensin-converting enzyme 2 (ACE2), an enzyme attached to cell membranes in the lungs, arteries, heart, kidney, and intestines that has been identified as a key cell entry point[8]. Viral particles infect the lungs by binding to and entering respiratory epithelial cells (type II alveolar pneumocytes)[7]. Upon viral spike protein binding to host cell ACE2 receptors, the enzyme furin enables the virus to enter the host cell; the spike protein must be cleaved by furin or furin-like proteases to become fully functional[9].

Viral propagation in the host cell induces a limited innate immune response, now detectable with nasal swabs[3]. The virus proliferates in the respiratory tract, where a more forceful immune response leads to clinical manifestation. Subsequent disease development and severity can be predicted by innate response cytokines; high cytokine (IL-6, IL-10) serum levels can indicate greater disease severity[10]. In >85% of COVID-19 cases, a proportionate immune response eradicates the virus, and the patient has mild/no symptoms, but in 10-15% of cases, the immune response is disproportionately forceful: an immunopathological phase occurs and the patient develops more severe disease requiring hospitalisation, typically due to hypoxic pneumonia[7].

The Complement System and COVID-19

The complement system (part of the innate immune system) enhances antibodies' and phagocytic cells' ability to eradicate pathogens and damaged host cells, instigate inflammation, and attack pathogen cell membranes[11]. It is a bridge between innate and acquired immunity, comprising >30 proteins (e.g., serum proteins, cell membrane receptors) synthesised by the liver and circulating in the blood as inactive precursors[11]. However, excessive complement activation can contribute to destructive host cell/tissue inflammation and COVID-19 pathogenesis[12]. There is therapeutic promise in immunotherapies during the immunopathological phase of COVID-19; C5a-C5aR axis blockade may prevent acute respiratory distress syndrome (ARDS) from worsening or causing death[7].

COVID-19 Nasal Epithelium Immune Response

The mucosal immune response, incorporating innate and adaptive immune system components, is a key disease defence. Central to the cranial pharyngeal mucosal immune response is the nasal cavity and nasopharynx-associated lymphoid tissue (NALT), which helps maintain immune homeostasis between commensal microbiota and pathogenic invaders[13]. The NALT contains dendritic cells, macrophages, and lymphocytes, of which ~50% are immunoglobulin-producing B-lymphocytes[13]. Lymphoid tissues harbour M cells: these use trans-epithelial transport to relocate micro-organisms from the apical surface toward immune cells at the basolateral site[13]. NALT-associated cells (sinonasal solitary chemosensory cells) release chemokines and cytokines, activating downstream immune cascades[13]. The nose and NALT generate Th1- and Th2- polarised lymphocytes and IgA-committed B cells[13]. Nasal epithelial ciliated and goblet cells, alongside dendritic cells, microfold cells and macrophages, form a gateway to local and systemic immune response initiation[13]. Nasal epithelial cells express the highest levels of ACE2 and the cellular serine protease TMPRSS2[13].

If the immune system cannot eliminate localised SARS-CoV-2 from the epithelial barrier, the virus progresses to the endothelium. The resulting inflammation is associated with microthromboses, macrovascular thromboses, and pulmonary embolism in COVID-19[14].

The Healthy Nasal Microbiome

The human microbiome contains the collective genomes of commensal, symbiotic, and pathogenic microorganisms at multiple body habitats, each harbouring a distinctive taxonomic profile[15]. The nose is a transition zone between dry skin and moist mucoid airways, typically with a similar microbial pattern to the skin; prominent genera include *Propionibacterium*, *Corynebacterium*, and *Staphylococcus*[16].

Healthy nasal microbiome signature data is well-documented; in The Human Microbiome Project (HMP)[17], healthy nasal swab samples underwent 16S rRNA sequencing and taxonomic classification. Examination of the anterior nares microbiome observed a broad distribution and near-ubiquity of opportunistic pathogens, including *S. aureus*, *Escherichia coli, S. epidermidis*, *Propionibacterium acnes*, and *Klebsiella pneumoniae*. The most abundant healthy nasal microbiome species were *P. acnes* (42.5%), *S. epidermidis* (12.7%) and *S. aureus* (5.0%). Koskinen et al[18] characterised the human nasal microbiome with Illumina MiSeq next generation sequencing, identifying 23 bacterial phyla. Abundant phyla were *Actinobacteria* (50% abundance), *Firmicutes* (28%), *Proteobacteria* (14%), and *Bacteroidetes* (1.5%). Abundant genera were *Corynebacterium* (43%), a human skin bacterium found frequently in the nose; *Staphylococcus* (15%), known nasal microbiome inhabitants; *Dolosigranulum* (4%), commensal upper respiratory tract inhabitants associated with infection and health; and *Peptoniphilus* (4%), a known abundant nasal, gut and vaginal inhabitant[18].

The Human Microbiome in COVID-19

COVID-19 microbiome research has focused primarily on the gut and lung. Alteration of gut microbial metabolites and related species can incur lung inflammation responses and disease development[19]. In viral infections, intestinal injury is caused by lymphocyte migration *via* the CCL25-CCR9 axis from the respiratory tract to the intestinal mucosa[20]. The link between gut microbiome dysbiosis and lung health is clear. The gut/lung communication mechanism remains unclear, but microbiome/dysbiosis regulation suggests that metabolites (short chain fatty acids), colonisation-induced immunity, and other protective functions, influence viral invasion and lung colonisation[21]. Gut and lung microbiome profiles have been found to be significantly altered in COVID-19, impacting disease severity. Understanding the "gut-lung axis" is vital for addressing COVID-19 disease progression, the relevance of pre-existing conditions, and complication risk.

Some uncommon COVID-19 symptoms are gastro-intestinal (GI): Schmulson et al[22] found that GI symptom frequency in 2,800 COVID-19 patients ranged from 3.0% to 39.6%, concluding that COVID-19 can manifest initially with GI symptoms. SARS-CoV-2 is detectable in faeces; intestinal epithelial cells, particularly small intestine enterocytes, express ACE2 receptors. SARS-CoV-2 also uses receptors for transmembrane protease serine 2 (TMRPSS2), an enzyme expressed in the small intestinal epithelial cells, to enter cells[23]. SARS-CoV-2 activity may cause gut ACE2 modifications leading to increased intestinal inflammation and diarrhea susceptibility. ACE2 and TMPRSS2 have high co-expression in enterocytes, the oesophagus, and lungs, greatest in the small intestine; 20% expressed in enterocytes and 5% in colon cells[23]. ACE2 has a significant role in intestinal inflammation and microbiome composition; a link has been established between ACE2 amino acid transport and gut microbial ecology during SARS-COV-2 infection[23].

COVID-19 significantly decreases gut microbiome diversity, incurs higher relative abundance of certain taxa (*Streptococcus*, *Rothia*, *Veillonella*, *Actinomyces*) *vs.* healthy controls, and lowers beneficial symbiont abundance[24]. Several potentially immunomodulatory commensal gut species (*Faecalibacterium prausnitzii*, *Eubacterium rectale*, *Bifidobacteria*) are under-represented in COVID-19 patients, with abundances remaining low up to 30 days post-recovery[5]. There is also a link between gut microbiome composition and healthy individuals' COVID-19 susceptibility; increased *Lactobacillus* abundance correlated with raised anti-inflammatory IL-10 levels and improved disease prognosis, while increased proinflammatory bacteria (*Klebsiella*, *Streptococcus*, *Ruminococcus gnavus*) abundance correlated with raised proinflammatory cytokine levels and increased disease severity[25].

SARS-CoV-2 infection also results in lung bacterial dysbiosis; COVID-19 patients' lung taxonomy differs from healthy control samples (p = .001)[26]. The most abundant lung bacterial genera in COVID-19 patients are *Acinetobacter* (80.70% of total sequences), *Chryseobacterium* (2.68%), *Burkholderia* (2.00%), *Brevundimonas* (1.18%), *Sphingobium* (0.93%), and *Enterobacteriaceae* (0.68%)[27].

The Human Microbiome in Influenza

Research has investigated human respiratory microbiome changes during influenza infection. US adults with influenza had increased *Streptococcus pneumoniae* and *Staphylococcus aureus* in the nose and throat microbiomes[28]. The first human population study to investigate the relationship

between the nose/throat microbiome and influenza hypothesised that the nose/throat microbiome is a factor in influenza susceptibility, observing taxonomic associations between the nose/throat microbiome and influenza; relative abundances of *Alloprevotella*, *Prevotella* and *Bacteroides* oligotypes were differentially expressed in influenza compared to healthy controls[29].

Moving from Targeted to Non-Targeted Microbiome Characterisation Methods

Traditional microbiome characterisation studies implement targeted classification methods such as microbial culture, shotgun sequencing, and 16S rRNA sequencing. The advent of widely available, comprehensive microbial reference genomes has led to research into repurposing unmapped sequencing reads (RNA-seq/DNA-seq data) for analysing human microbiome communities in healthy and diseased human tissues; Mangul et al[6] proposed the "lost and found pipeline", a process of taxonomically classifying RNA-seq reads that do not map to the human genome (the microbial read fraction). We previously repurposed SRA-derived RNA-seq reads to demonstrate that this non-targeted approach can characterise the human gut, lung, skin, and blood microbiome as effectively as targeted methods [under review at time of publication]. This principle underpins our current research, in which we implement our non-targeted microbiome characterisation workflow to investigate the human nasal microbiome in COVID-19 compared to healthy and influenza samples.

PATIENTS AND METHODS

Sample Criteria

We identified a suitable set of data to demonstrate distinct bacterial signatures from the human nasal microbiome in COVID-19 infection, healthy controls, and influenza-infected controls. We considered several criteria to determine the samples' read suitability. Reads must:

1. Not map to human reference genome hg38.
2. Not be low quality (low quality reads/errors would account for a lack of mapping to hg38).
3. Not contain artificial sequences (e.g., sequence adapters).
4. Not be repetitive (this would hinder identification).
5. Be of sufficient length (short reads decrease classification accuracy).

We identified the following variables for consideration:

1. Origin: All RNA-seq datasets must be single-species origin (Homo sapiens) and from the nose/nasal cavity.
2. Health/disease: We included healthy controls to demonstrate that COVID-19 alters the nasal microbiome, and influenza controls to demonstrate that this alteration is disease specific, not a generic inflammation response.
3. Sequencing technology: We used samples sequenced with Illumina here; we may consider the suitability/mapping rates of different platforms in future experiments.
4. Mapping tool: We compared TopHat and HISAT2 and found that HISAT2 maps RNA-seq reads to hg38 more comprehensively, so we used HISAT2 here.

Identification of Suitable Samples

We focused on the nasal microbiome, to investigate changes in early-stage COVID-19 before the virus proliferates in the respiratory tract and manifests clinically. The healthy nasal microbiome has well-documented signature data[18]. By including healthy samples, we can check for the expected taxonomic profile by comparing our results to those from previous research, further demonstrating our non-targeted microbiome classification technique [under review] is robust. We can then compare our healthy profile to our COVID-19 profile to ascertain COVID-19's effect on the nasal microbiome. Similarly, we included influenza-infected samples to demonstrate that the differential COVID-19 microbiome profile is COVID-19 specific, not a generic inflammation response. We demonstrate this by observing different profiles in COVID-19 and influenza samples. We compiled a set of SRA data to include each state (COVID-19, healthy, influenza). Study metadata was re-

corded to ensure the key variables were sufficiently represented where possible. We imported a maximum of 10 read samples per state from the SRA into Galaxy. The imported samples were then quality assessed with FastQC prior to mapping. The original researchers obtained the relevant ethics approval, so further approval need not be sought.

Selecting the Most Effective Mapping Method

We mapped the reads to the human reference genome hg38, ensuring that the highest possible percentage of host reads was identified and removed before classifying the unmapped fraction. We mapped the samples to hg38 using the mapping tool HISAT2; justification for choosing this mapping tool is as follows. We compared the mapping tools HISAT2 and TopHat; existing RNA-seq data (12 human blood control samples from 4 studies) was FastQC quality assessed then mapped to hg38 with TopHat[30] and HISAT2[31]. We achieved mean mapping rates of 77.25% with TopHat and 89.45% with HISAT2. Therefore, HISAT2 was the chosen mapping tool moving forward.

Quality Assessment of the Unaligned Reads

The unmapped reads output by HISAT2 were quality assessed using FastQC. For paired-end data, forward and reverse reads were quality assessed. We recorded each sample's per sequence quality score; this reports the probability an incorrect base call has occurred. The FastQC per base sequence content graphs were checked to verify the samples were biological reads, not long repeating base sequences (errors). The FastQC adapter content graphs were checked to ensure that human reads with sequencing adapters were not contaminating the unmapped read pool, which would lower the mapping rate and contaminate the "unmapped" samples, reducing taxonomic classification accuracy. The results of the HISAT2 mapping and FASTQC quality assessment are in Table 1.

In summary, quality assessing unmapped reads was important, to verify that reads failed to map to hg38 because they were not human, not because of poor quality. Quality assessing twice (total reads then unmapped reads) ensured that high read quality remained so following mapping.

Taxonomic Classification and Data Analysis

The unmapped read samples were then subsampled to 1 million reads to allow Kraken[32] classification to run quickly and efficiently, while ensuring a sufficiently large, classified read pool to analyse. We conducted Kraken classification on the subsampled unmapped reads to assign taxonomic labels, using the Kraken bacterial database to ensure Kraken only identified bacterial RNA in the samples. One Kraken-mpa-report was conducted on all Kraken classification outputs and the taxonomic data was recorded and compiled for statistical analysis.

The dataset was filtered to phylum, genus, and species, then normalised by calculating each taxon's percentage out of the total data. Normalisation corrects the bias from varying library sizes and is essential in microbiome sequencing experiments[33]. Each taxon's average percentage was used to identify the 20 most abundant phyla, genera, and species in each state. Prior research into the human nasal microbiome in health, COVID-19, and influenza was reviewed, to verify that we observed the 'expected' taxa.

RESULTS

Mapping and Quality Assessment

On the SRA, we identified 9 suitable COVID-19 infected human nasal samples, then selected 10 healthy samples and 10 influenza samples. Each study's means and standard deviations (SD) for mapping rate, read count, sequence length and phred score are in Table 1, from which we have two conclusions.

TABLE 1. A SUMMARY (MEAN & SD) OF THE GALAXY RESULTS (10 HEALTHY, 9 COVID-19 AND 10 INFLUENZA SAMPLES), FOLLOWING HISAT2 MAPPING AND FASTQC UNMAPPED READ ANALYSIS.

Study Accession:	**Total Reads:**		**Unmapped Reads:**		
PRJNA… **(no. of samples)**	**Phred Score**	**HISAT2 Mapping Rate**	**Unmapped Read Count**	**Sequence Length**	**Phred Score**
Nasal RNA Illumina, Healthy:					
637909 (6 single)	38 (0)	96.71% (0.82%)	1,274,990 (311,542.63)	100 (0)	37.5 (0.55)
430406 (4 paired)	38 (0)	97.56% (0.55%)	6,777,450.5 (3,277,334.11)	100 (0)	37 (0)
Nasal RNA Illumina, COVID-19:					
691164 (4 paired)	38.88 (0.25)	29.94% (1.32%)	863,123,848.5 (69,847,160.48)	151 (0)	38.88 (0.25)
648499 (1 single)	40 (N/A)	89.38% (N/A)	37,888,644 (N/A)	91 (N/A)	40 (N/A)
645534 (4 paired)	34.25 (0.29)	72.07% (16.67%)	5,297,584.5 (2,632,099.01)	35-151 (N/A)	34.25 (0.29)
Nasal RNA Illumina, Flu:					
482564 (10 paired)	40 (0)	97.62% (0.17%)	2,080,771.8 (491,180.99)	100 (0)	40 (0)

First, all reads are sufficient quality; although there is no formally defined lower limit, a phred score of ≥33 should be aimed for[34]. All studies have a mean phred score >33 before and after mapping. Second, mean mapping rates range from 29.94% to 97.62%, visualised in Figure 1, a bar graph of each study's mean HISAT2 mapping rates, with SDs included as error bars.

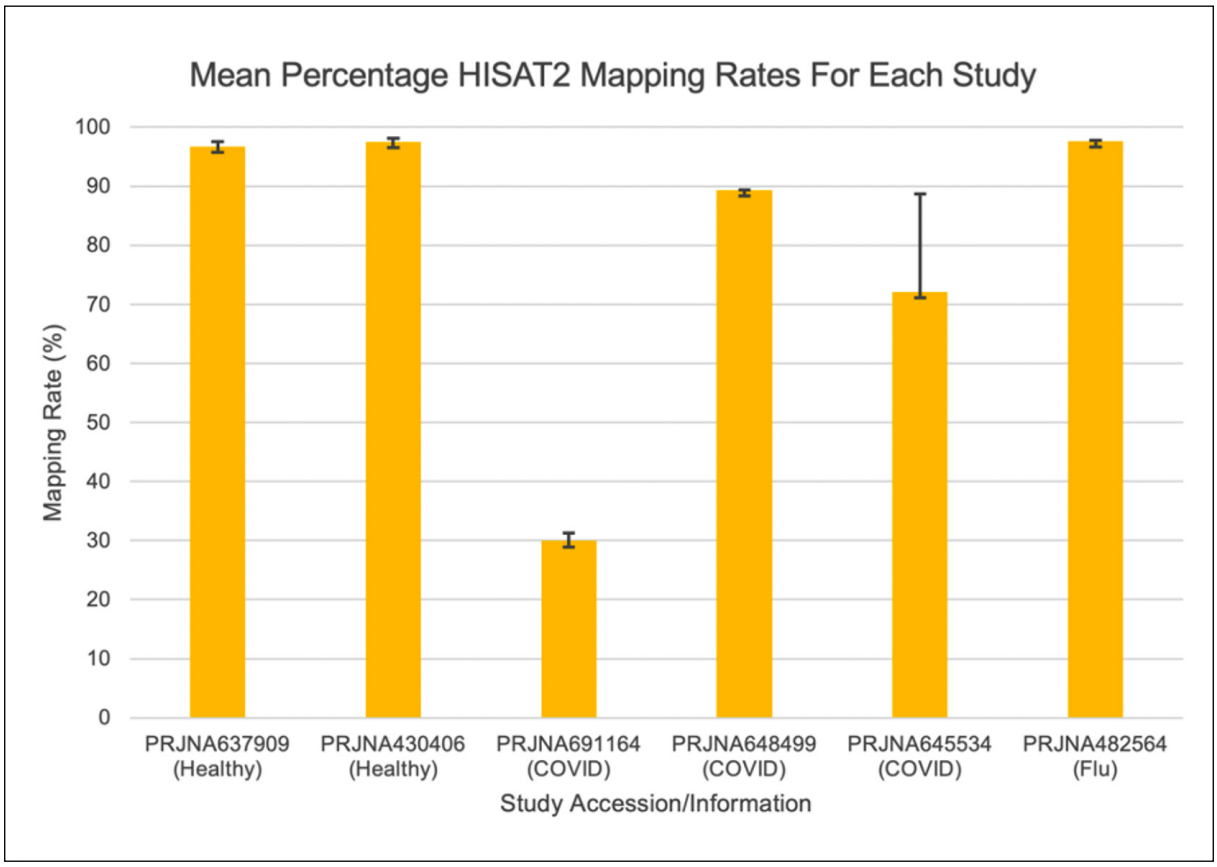

Figure 1. Mean % mapping rates for each study following HISAT2 mapping to hg38. The SDs for each mean % mapping rate are included as error bars.

We FastQC quality assessed the total and unmapped samples. We checked the per base sequence content graphs to verify that the samples were human genome reads, not sequences of repeated bases; each total sample had approximately 25% of each base, the expected biological read distribution. We also checked the adapter content graphs to ensure that sequencing adapters were not contaminating the unmapped read pool.

In summary, FastQC quality assessment demonstrates our unmapped read samples are of sufficient phred quality, are genuine human genome reads, and are not contaminated with sequence read adapters.

Taxonomic Classification with Kraken

We conducted Kraken taxonomic classification then exported and normalised the dataset. Table 2 shows the 5 most abundant phyla (top section), genera (middle section), and species (bottom section) respectively, as a percentage of the total classified bacterial reads in each state (more than 5 taxa at each level were identified, but only the most abundant taxa are presented, for conciseness).

TABLE 2. THE 5 MOST ABUNDANT PHYLA, GENERA, AND SPECIES BY PERCENTAGE OF TOTAL BACTERIAL READS IN EACH OF THE 3 CONDITIONS – MEAN (M) AND STANDARD DEVIATION (SD) VALUES ARE GIVEN.

	Healthy (%)		COVID-19 (%)		Flu (%)	
	M	***SD***	**M**	***SD***	**M**	***SD***
1. Phylum						
Proteobacteria	43.87	24.73	19.92	6.21	31.52	2.04
Tenericutes	12.64	6.08	16.09	15.74	19.62	1.78
Actinobacteria	9.32	15.97	10.28	14.39	3.80	0.76
Firmicutes	7.87	4.04	18.04	11.56	7.33	0.35
Bacteroidetes	6.11	2.99	8.40	8.16	7.12	0.73
2. Genus						
Healthy:						
Alteromonas	25.46	30.30	0.42	0.76	0.86	0.94
Mycoplasma	12.49	6.04	15.60	15.33	18.28	1.57
Mycobacterium	5.37	16.67	0.43	0.49	1.55	0.63
Ehrlichia	2.13	1.27	1.26	1.34	2.35	0.35
Propionibacterium	2.01	2.32	0.26	0.37	0.35	0.15
COVID-19:						
Mycoplasma	12.49	6.04	15.60	15.33	18.28	1.57
Thermoanaerobacter	1.47	0.57	4.13	4.69	1.36	0.13
Corynebacterium	0.38	0.35	3.85	10.02	0.14	0.06
Streptococcus	0.73	0.41	3.59	4.31	1.73	0.30
Shigella	0.49	0.24	2.59	2.97	1.52	0.27
Influenza:						
Mycoplasma	12.49	6.04	15.60	15.33	18.28	1.57
Salmonella	0.34	0.33	0.30	0.85	3.57	0.96
Escherichia	1.19	0.40	1.55	1.42	2.81	0.14
Candidatus Carsonella	1.26	0.44	1.23	1.23	2.55	0.23
Ehrlichia	2.13	1.27	1.26	1.34	2.35	0.35
3. Species						
Healthy:						
A. macleodii	25.46	30.30	0.42	0.76	0.86	0.94
M. hyopneumoniae	8.77	4.12	10.89	10.66	13.12	1.05
M. tuberculosis	5.32	16.68	0.32	0.45	1.48	0.63
E. canis	1.90	1.21	1.12	1.23	2.14	0.35
P. acnes	1.79	2.07	0.11	0.16	0.33	0.14
COVID-19:						
M. hyopneumoniae	8.77	4.12	10.89	10.66	13.12	1.05
T. wiegelii	1.09	0.36	3.82	4.89	1.17	0.11
S. flexneri	0.40	0.20	2.55	2.94	1.48	0.28
M. pulmonis	1.59	0.95	2.01	2.25	1.84	0.21
M. hyorhinis	1.51	0.67	1.96	1.97	2.43	0.18
Influenza:						
M. hyopneumoniae	8.77	4.12	10.89	10.66	13.12	1.05
S. enterica	0.34	0.33	0.30	0.85	3.57	0.96
E. coli	1.19	0.40	1.55	1.42	2.81	0.14
C. Carsonella ruddii	1.26	0.44	1.23	1.23	2.55	0.23
M. hyorhinis	1.51	0.67	1.96	1.97	2.43	0.18

Phylum-level observations

The phylum classification data is shown in Table 2 (top section) and as stacked bars in Figure 2, which includes the full list of 23 phyla, sorted in order of most abundant phyla based on average abundance in the healthy samples. Each state has a characteristic microbiome signature, though *Proteobacteria* dominates all three; 43.87% abundance in healthy samples, decreasing to 19.92% in COVID-19 and 31.52% in influenza. Further noticeable differential phylum abundances depending on health/disease states are as follows:

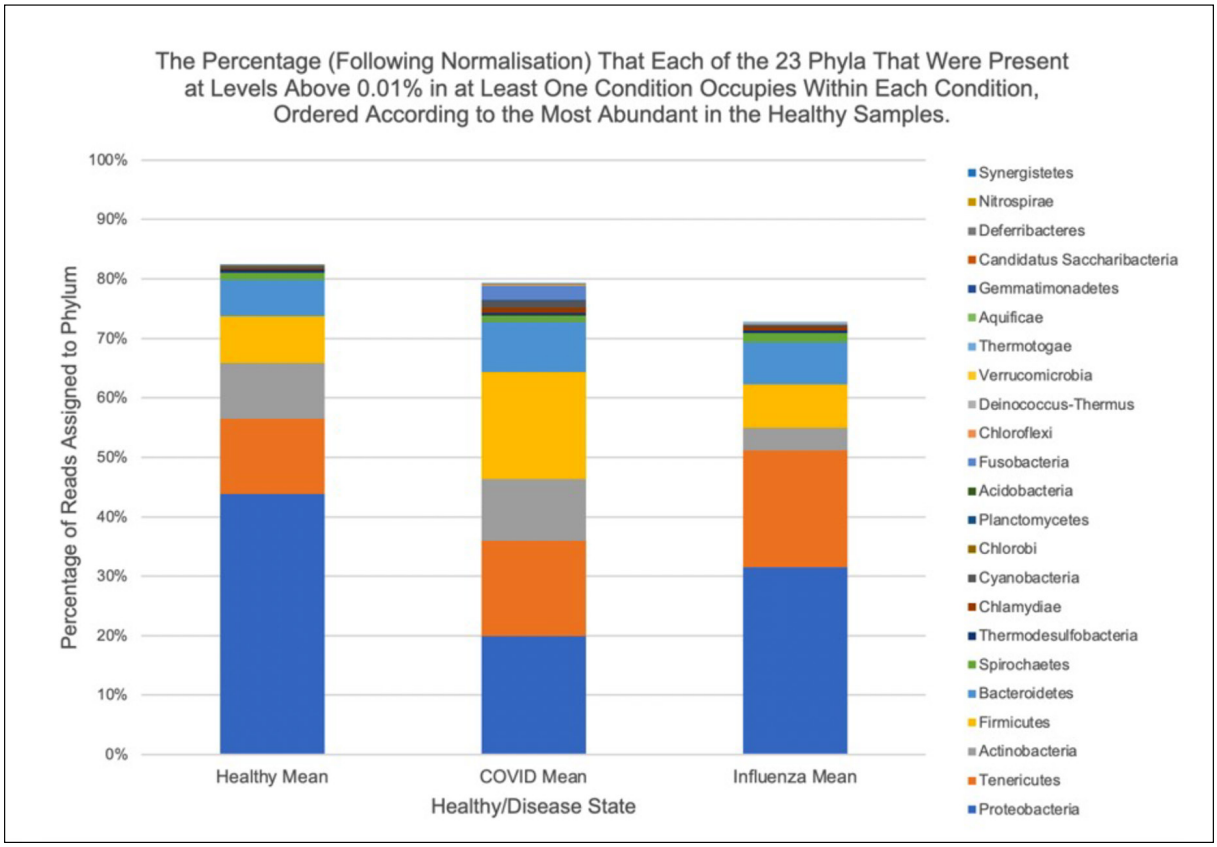

Figure 2. The percentage of each microbiome condition occupied by each of the 23 classified phyla, with each colour representing a different phylum.

- Significantly increased abundance of *Firmicutes* in COVID-19 (18.04%) but slightly decreased in influenza (7.33%) compared to healthy (7.87%).
- Slightly increased *Actinobacteria* in COVID-19 (10.28%) and significantly decreased in influenza (3.80%) compared to healthy (9.32%).
- Increased *Tenericutes* in COVID-19 (16.09%) and influenza (19.62%) compared to healthy samples (12.64%).
- Increased *Bacteroidetes* in COVID-19 (8.40%) and influenza (7.12%) compared to healthy (6.11%).

Genus-level observations

Our genera results are shown in Table 2 (middle section) and in Figure 3, by compiling the three lists of 20 most abundant genera into one list of 34 and sorting it into the order of most abundant genera based on each genus' average abundance within the healthy samples.

We observe disease-dependent differential expression of numerous nasal microbiome genera – this section summarises the headline results. Firstly, several genera decreased in one or both diseases compared to healthy samples:

- *Alteromonas* is dominant in healthy samples (25.46%), but strikingly diminished in COVID-19 (0.42%) and influenza (0.86%).
- *Mycobacterium* abundance was diminished in both diseases (more significantly in COVID-19 (0.43%) than in influenza (1.55%)) compared to healthy samples (5.37%).
- *Propionibacterium* was reduced in COVID-19 (0.26%) and influenza (0.35%) compared to healthy (2.01%).

Meanwhile, several genera increased in abundance in COVID-19 compared to both healthy and influenza samples:

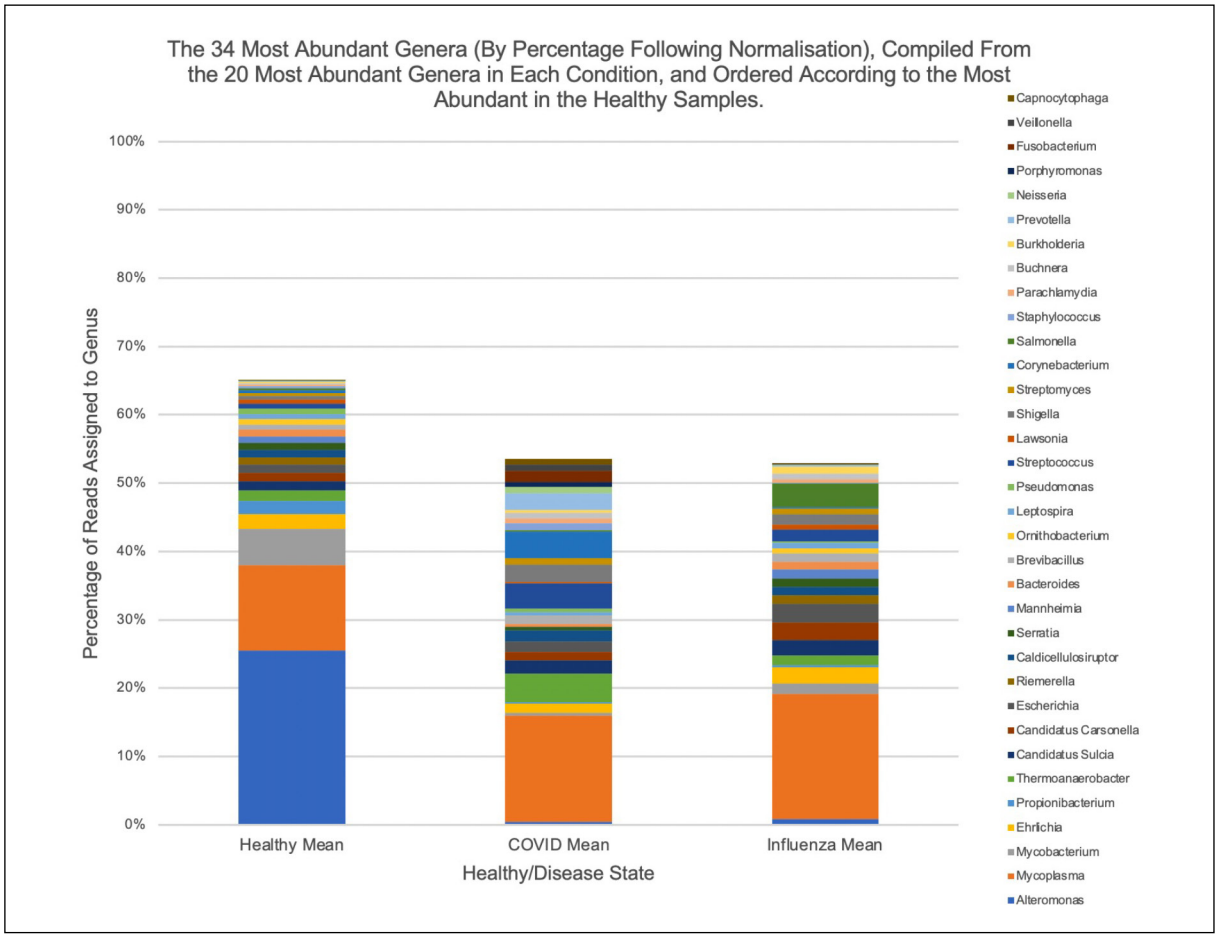

Figure 3. The percentage of each condition occupied by each of the 34 most abundant genera (compiled from the 20 most abundant genera from each microbiome condition), with each colour representing a different genus.

- *Thermoanaerobacter* was significantly increased in COVID-19 (4.13%), but slightly decreased in influenza (1.36%), compared to healthy (1.47%).
- *Corynebacterium* was increased in COVID-19 (3.85%) and slightly decreased in influenza (0.14%) compared to healthy (0.38%).
- *Lactobacillus* was slightly increased in COVID-19 (0.74%) and decreased in influenza (0.38%) compared to healthy (0.65%).
- *Streptococcus* was increased in COVID-19 (3.59%) and increased less so in influenza (1.73%), compared to healthy (0.73%).
- *Prevotella* had raised abundance of 2.44% in COVID-19, compared to 0.20% in influenza and 0.12% in healthy.
- *Salmonella* abundance was raised in influenza (3.57%) but not in COVID-19 (0.30%) compared to healthy (0.34%).
- *Escherichia* was raised in influenza (2.81%) compared to COVID-19 (1.55%) and healthy (1.19%), as was *Candidatus Carsonella* (2.55% in influenza, *vs.* 1.23% in COVID-19 and 1.26% in healthy).

Species-level observations

Our species results are shown in Table 2 (bottom section) and in Figure 4, by compiling the three lists of 20 most abundant species into one list of 32 and sorting it into the order of most abundant species based on the average abundance of each species within the healthy samples.

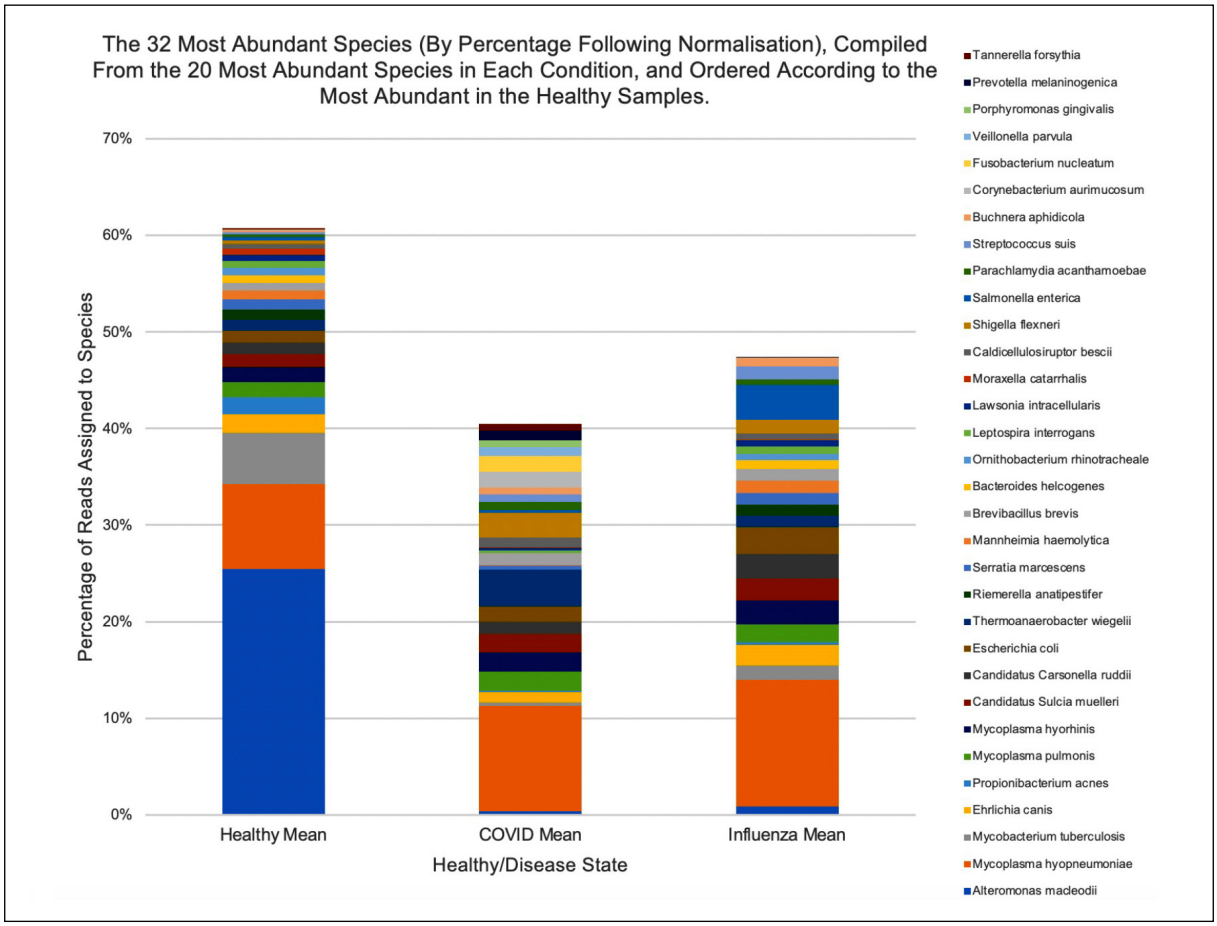

Figure 4. The percentage of each condition occupied by each of the 32 most abundant species (compiled from the 20 most abundant species from each microbiome condition), with each colour representing a different species.

The following species were diminished in both diseases, all more significantly so in COVID-19 than in influenza:

- *Alteromonas macleodii* was the most abundant species in healthy samples (25.46%); this significantly decreased to 0.42% in COVID-19, and 0.86% in influenza.
- Decreased *Propionibacterium acnes* abundance was observed in both diseases. Mean healthy abundance was 1.79%, decreasing to 0.11% in COVID-19 and 0.33% in influenza.
- The third most abundant species in healthy samples, *Mycobacterium tuberculosis* (5.32%), also decreased more significantly in COVID-19 (0.32%) than in influenza (1.48%).

In contrast, some species demonstrated increased abundance in both diseases:

- *Mycoplasma hyopneumoniae* had 8.77% abundance in healthy samples, slightly increasing to 10.89% in COVID-19 and 13.12% in influenza.
- *Shigella flexneri* increased from 0.40% (healthy) to 2.55% (COVID-19) and 1.48% (influenza).

Finally, some species' abundances were increased in just one disease:

- Compared to 1.09% healthy abundance, *Thermoanaerobacter wiegelii* was significantly increased in COVID-19 (3.82%) but not influenza (1.17%).
- *Staphylococcus aureus* abundance was 0.01% (healthy and influenza), increasing to 0.20% (COVID-19).

DISCUSSION

Mapping and Quality Assessment

Mean mapping rates to hg38 range from 29.94% to 97.62%. We have previously demonstrated more complete hg38 mapping in Illumina-sequenced datasets than Ion Torrent-sequenced [under review

at time of publication]. Here we used Illumina datasets, observing varied mean mapping rates. PRJNA691164 (COVID-19) deviates from the other studies (29.94%). PRJNA645534 (COVID-19) has the greatest SD (16.67%). The nasal microbiome is considered low microbial biomass[35], so high mapping rates were expected. Our COVID-19 studies had the lowest mean and greatest SD; this may indicate that COVID-19 increases nasal microbial biomass, causing lower mapping rates in COVID-19 samples.

Phylum-Level Observations

Although each state has a characteristic microbiome signature, *Proteobacteria* dominates all three; 43.87% abundance in healthy samples, decreasing to 19.92% in COVID-19 and 31.52% in influenza. According to research, *Firmicutes*, *Actinobacteria*, *Bacteroidetes* and *Proteobacteria* are the basis of the healthy human nasal microbiome[18]; all were among our 5 most abundant healthy phyla, demonstrating that our microbiome characterisation method yields the 'expected' nasal microbiome taxa. Previously, chronic rhinosinusitis (CRS) patients had increased nasal *Proteobacteria* and decreased *Bacteroidetes*[36], while nasal *Proteobacteria* abundance was significantly higher in asthma patients who had ≥1 asthma-related ER visit compared to those who did not[37]. This suggests that increased *Proteobacteria* is observed in infection and inflammatory states. We observe significantly increased abundance of *Firmicutes* (a known prominent nasal phylum[19]) in COVID-19 (18.04%) but slightly decreased in influenza (7.33%) compared to healthy (7.87%). We observe slightly increased *Actinobacteria* in COVID-19 (10.28%) and significantly decreased in influenza (3.80%) compared to healthy (9.32%). Previously, chronic rhinosinusitis (CRS) patients showed significantly lower *Actinobacteria* abundance than healthy controls[36]. We observe increased *Tenericutes* in COVID-19 (16.09%) and influenza (19.62%) compared to healthy samples (12.64%). *Tenericutes* was previously observed as the 2nd most abundant phylum in COVID-19 (18.09%) but not in healthy[38]. We observe increased *Bacteroidetes* in COVID-19 (8.40%) and influenza (7.12%) compared to healthy (6.11%). Previously, 4 *Bacteroidetes* species (*B. dorei*, *B. thetaiotaomicron*, *B. massilensis* and *B. ovatus*) were inversely correlated with faecal SARS-CoV-2 load (Spearman Rho<-0.2, $p<.05$)[39]. These species are associated with murine colon ACE2 downregulation[40], suggesting that *Bacteroidetes* abundance correlates to COVID-19 severity and can hinder SARS-CoV-2 host entry.

Genus-Level Observations

We observe disease-dependent differential expression of several nasal microbiome genera. *Alteromonas* is dominant in healthy samples (25.46%), but strikingly diminished in COVID-19 (0.42%) and influenza (0.86%). Significant *Alteromonas* depletion in both diseases suggests that this is not disease-specific, but a generalised immune/inflammatory response. There are two potential explanations for *Alteromonas* eradication in disease. First, inflammation creates a hostile nasal microbiome environment for *Alteromonas*. Second, inflammation creates a favourable environment for a competitor whose growth impedes *Alteromonas*. Many microorganisms depend on the activity of others to grow and reproduce successfully[41].

Several other genera decreased in one/both diseases compared to healthy samples. *Mycobacterium* abundance was diminished in both diseases, more significantly in COVID-19 (0.43%) than in influenza (1.55%) compared to healthy samples (5.37%). *Propionibacterium* was also reduced in COVID-19 (0.26%) and influenza (0.35%) compared to healthy (2.01%). Meanwhile, several genera increased in abundance in COVID-19 compared to both healthy and influenza. *Thermoanaerobacter* was significantly increased in COVID-19 (4.13%), but slightly decreased in influenza (1.36%), compared to healthy (1.47%). *Corynebacterium* was also increased in COVID-19 (3.85%) and slightly decreased in influenza (0.14%) compared to healthy (0.38%). *Lactobacillus* was slightly increased in COVID-19 (0.74%) and decreased in influenza (0.38%) compared to healthy (0.65%). *Streptococcus* was increased in COVID-19 (3.59%) and increased less so in influenza (1.73%), compared to healthy (0.73%). *Prevotella* had raised abundance of 2.44% in COVID-19, compared to 0.20% in influenza and 0.12% in healthy. In contrast, *Salmonella* abundance was raised in influenza (3.57%) but not in COVID-19 (0.30%) compared to healthy (0.34%). *Escherichia* was also raised in influenza (2.81%) compared to COVID-19 (1.55%) and healthy (1.19%), as was *Candidatus Carsonella* (2.55% in influenza *vs.* 1.23% in COVID-19 and 1.26% in healthy).

The differential expression of genera in both diseases compared to healthy controls (*Alteromonas*, *Mycobacterium* and *Propionibacterium*) suggests that some nasal microbiome changes occur due to a generalised (not disease specific) inflammatory response. Meanwhile, the differential expression of other genera in just one disease indicates that other nasal microbiome changes are disease specific.

Species-Level Observations

We observed several species with differential expression in health/disease. Most differentially expressed species' abundances increased in one/both diseases compared to healthy samples, although there were exceptions. For example, *Alteromonas macleodii* was the most abundant species in healthy samples (25.46%); this significantly decreased to 0.42% in COVID-19, and 0.86% in influenza. Previously, Vaziri et al[42] compared microbial DNA from stool samples of end-stage renal disease (ESRD) patients to healthy participants, finding that in healthy control samples, average *A. macleodii* intensity was 6,602, increasing to 7,674 in ESRD samples. We previously observed significant *A. macleodii* dominance in the gut, lung, and blood microbiomes of healthy human participants [under review at time of publication]. However, this dominance was mitigated following an experimental trial of the effect of metaSPAdes assembly of the read samples prior to taxonomic classification. This suggests that metaSPAdes assembly may mitigate single taxa dominating samples, which potentially arises from the misassignment of some short (unassembled) fragments due to taxonomic similarities between species, and the short fragments being insufficiently long to be unique to a single taxon. Thus, moving forward, we will consider the possibility of assembling read samples before classification.

We also observed decreased *Propionibacterium acnes* in both diseases. Mean healthy abundance was 1.79%, decreasing to 0.11% in COVID-19 and 0.33% in influenza. Previously, *P. acnes* was one of the most abundant microorganisms in samples from healthy middle meatus swabs, alongside *Staphylococcus aureus* and *Staphylococcus epidermidis*[43]. Our third most abundant species in healthy samples was *Mycobacterium tuberculosis* (5.32%), decreasing more significantly in COVID-19 (0.32%) than in influenza (1.48%). Two unusual cases in which a patient had triple infections of SARS-CoV-2, *M. tuberculosis* and HIV were reported, although this co-infection did not appear to worsen COVID-19 outcomes[44]. A report from Liaoning Province, China, suggested that 36% of COVID-19 cases were infected with tuberculosis, or had been previously[45].

We observe that all 3 species with decreased abundance in both diseases decreased more greatly in COVID-19 than influenza. However, some species demonstrated increased abundance in both diseases. *Mycoplasma hyopneumoniae* had 8.77% abundance in healthy samples, slightly increasing to 10.89% in COVID-19 and 13.12% in influenza. *Shigella flexneri* increased from 0.40% (healthy) to 2.55% (COVID-19) and 1.48% (influenza). This human faecal microbiome species has been implicated as an infection-causing pathogen in a compromised gut microbiome[46].

Some species were increased in just one disease. Compared to 1.09% healthy abundance, *Thermoanaerobacter wiegelii* was significantly increased in COVID-19 (3.82%) but not influenza (1.17%). Meanwhile, our *S. aureus* abundance was 0.01% (healthy and influenza), increasing to 0.20% (COVID-19). Prior research associates *S. aureus* with the human nasal microbiome; the region between the anterior nasal vestibule and the posterior nasopharyngeal cavity is the favoured *S. aureus* colonization site. Furthermore, although it is commensal to the healthy nasal microbiome in ~30% of humans, it can also be a dangerous and life-threatening pathogen[16].

We observed the most disease-specific species increases in influenza. *Salmonella enterica* abundance was 0.34% in healthy samples and 0.30% in COVID-19, increasing to 3.57% in influenza. *Escherichia coli* increased from 1.19% in healthy samples and 1.55% in COVID-19 to 2.81% in influenza. *Candidatus Carsonella ruddii* abundance was 1.26% in healthy samples and 1.23% in COVID-19, increasing to 2.55% in influenza. *S. enterica* serotypes are associated with bacteriaemia[47], but a COVID-19 link has not yet been made. Likewise, to date no research has implicated *Ca. C. ruddii* in COVID-19. However, *E. coli* was differentially abundant in healthy and COVID-19 samples in both bronchoalveolar lavage fluid (BALF) and peripheral blood mononuclear cell (PBMC) tissues[48].

Our graphs enable nasal microbiome compositional differentiation between diseases, at several taxonomic levels. Taxonomic abundances are less differential at the phylum level but are still significant, demonstrating that a change in health status can be observed in the nasal microbiome's compositional profile, as expected from previous research.

We have demonstrated that RNA-seq data from prior studies can be repurposed for the secondary use of characterising the human nasal microbiome in health, COVID-19, and influenza, and that the results from these characterisations concur with disease-specific microbiome analyses reported in the literature.

CONCLUSIONS

In conclusion, we demonstrate that existing RNA-seq reads from human nasal swabs can be repurposed to characterise the human nasal microbiome in health, COVID-19 and influenza, by taxonomically classifying the reads that remain unmapped to human reference genome hg38. Our results show that human nasal microbiome composition (relative abundances) differs between health/disease, and COVID-19/influenza. This was expected, based on prior human microbiome research in COVID-19[26] and influenza[28]. Furthermore, our results concur with those from previous microbiome characterisation studies, confirming the validity of our hypothesis that accurate microbiome composition data can be generated by the repurposing and secondary analysis of RNA-seq data.

Our research further demonstrates that our non-targeted microbiome characterisation technique can be implemented as accurately and effectively as the established targeted methods. This benefits future human microbiome research, demonstrating that findings can be generated from pre-existing RNA-seq datasets, which can be exploited with a fraction of the resources and costs of targeted, sequencing-based methodologies. In the context of COVID-19, our research is beneficial as it demonstrates that the early-stage COVID-19 nasal microbiome can be taxonomically differentiated firstly from the healthy microbiome, which may indicate a potential method of early warning in the disease, and secondly from the influenza-infected microbiome, indicating that these compositional changes are disease-specific, as opposed to a generalised inflammatory response; this is key in terms of differential diagnosis.

Conflict of Interest

All authors of this work have no conflicts of interest (actual or potential), financial or otherwise, to declare.

Acknowledgements

Felicity Wainwright would like to thank both Göksel Mısırlı and Peter Andras for their continued support and guidance in both this research and Felicity's PhD project as a whole. All authors would like to thank the original researchers who generated and deposited the data that was utilised and repurposed in this research.

Informed Consent

This research utilised and repurposed existing RNA-seq data that was deposited in online databases (the Sequence Read Archive) and made freely accessible – therefore, informed consent was obtained from the participants/patients in the first instance by the researchers who generated this existing data.

Authors' Contribution

Research design was a collaboration between all three authors. Research and data analysis was undertaken predominantly by Felicity Wainwright, with guidance and support by both Göksel Mısırlı and Peter Andras. The manuscript was written by Felicity Wainwright, with guidance, support, and revision/editing contributions by both Göksel Mısırlı and Peter Andras.

Funding

This research was undertaken as part of an ongoing self-funded PhD (Felicity Wainwright).

ORCID ID

Felicity Wainwright: https://orcid.org/0000-0002-6427-4903; Göksel Mısırlı: https://orcid.org/0000-0002-2454-7188; Peter Andras: https://orcid.org/0000-0002-9321-3296.

REFERENCES

1) Bresalier RS, Chapkin RS. Human microbiome in health and disease: the good, the bad, and the bugly. Dig Dis Sci 2020; 65: 671-673.
2) 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. Cell Microbiol 2018; 20: e12966.
3) 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 Immunol 2021; 14: 305-316.
4) García LF. Immune response, inflammation, and the clinical spectrum of COVID-19. Front Immunol 2020; 11: 1441.
5) Rueca M, Fontana A, Bartolini B, Piselli P, Mazzarelli A, Copetti M, Binda E, Perri F, Gruber CE, Nicastri E, Marchioni L. Investigation of nasal/oropharyngeal microbial community of COVID-19 patients by 16S rDNA sequencing. Int J Environ Res Public Health 2021; 18: 2174.
6) Mangul S, Yang HT, Strauli N, Gruhl F, Porath HT, Hsieh K, Chen L, Daley T, Christenson S, Wesolowska-Andersen A, Spreafico R. ROP: dumpster diving in RNA-sequencing to find the source of 1 trillion reads across diverse adult human tissues. Genome Biol 2018; 19: 1-12.
7) Chouaki Benmansour N, Carvelli J, Vivier E. Complement cascade in severe forms of COVID-19: Recent advances in therapy. Eur J Immunol 2021; 51: 1652-1659.
8) Hamming I, Timens W, Bulthuis MLC, Lely AT, Navis GV, van Goor H. Tissue distribution of ACE2 protein, the functional receptor for SARS coronavirus. A first step in understanding SARS pathogenesis. J Pathol 2004; 203: 631-637.
9) Hoffmann M, Kleine-Weber H, Schroeder S, Krüger N, Herrler T, Erichsen S, Schiergens TS, Herrler G, Wu NH, Nitsche A, Müller MA. SARS-CoV-2 cell entry depends on ACE2 and TMPRSS2 and is blocked by a clinically proven protease inhibitor. Cell 2020; 181: 271-280.
10) Han H, Ma Q, Li C, Liu R, Zhao L, Wang W, Zhang P, Liu X, Gao G, Liu F, Jiang Y. Profiling serum cytokines in COVID-19 patients reveals IL-6 and IL-10 are disease severity predictors. Emerg Microbes & Infect 2020; 9: 1123-1130.
11) Dunkelberger JR, Song WC. Complement and its role in innate and adaptive immune responses. Cell Res 2010; 20: 34-50.
12) Magro C, Mulvey JJ, Berlin D, Nuovo G, Salvatore S, Harp J, Baxter-Stoltzfus A, Laurence J. Complement associated microvascular injury and thrombosis in the pathogenesis of severe COVID-19 infection: a report of five cases. Transl Res 2020; 220: 1-13.
13) 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 Immunol 2021; 14: 305-316.
14) Llitjos JF, Leclerc M, Chochois C, Monsallier JM, Ramakers M, Auvray M, Merouani K. High incidence of venous thromboembolic events in anticoagulated severe COVID-19 patients. J Thromb Haemost 2020; 18: 1743-1746.
15) Costello EK, Lauber CL, Hamady M, Fierer N, Gordon JI, Knight R. Bacterial community variation in human body habitats across space and time. Science 2009; 326: 1694-1697.
16) Laux C, Peschel A, Krismer B. Staphylococcus aureus colonization of the human nose and interaction with other microbiome members. Microbiol Spectr 2019; 7.
17) The Human Microbiome Project Consortium. Structure, function and diversity of the healthy human microbiome. Nature 2012; 486: 207-214.
18) Koskinen K, Reichert JL, Hoier S, Schachenreiter J, Duller S, Moissl-Eichinger C, Schöpf V. The nasal microbiome mirrors and potentially shapes olfactory function. Sci Rep 2018; 8: 1-11.
19) Zhang D, Li S, Wang N, Tan HY, Zhang Z, Feng Y. The cross-talk between gut microbiota and lungs in common lung diseases. Front Microbiol 2020; 11: 301.
20) Dilantika C, Sedyaningsih ER, Kasper MR, Agtini M, Listiyaningsih E, Uyeki TM, Burgess TH, Blair PJ, Putnam SD. Influenza virus infection among pediatric patients reporting diarrhea and influenza-like illness. BMC Infect Dis 2010; 10: 1-5.
21) Barcik W, Boutin RC, Sokolowska M, Finlay BB. The role of lung and gut microbiota in the pathology of asthma. Immunity 2020; 52: 241-255.
22) Schmulson M, Dávalos MF, Berumen J. Beware: Gastrointestinal symptoms can be a manifestation of COVID-19. Rev Gastroenterol Méx (English Edition) 2020; 85: 282-287.
23) Villapol S. Gastrointestinal symptoms associated with COVID-19: impact on the gut microbiome. Transl Res 2020; 226: 57-59.
24) Gu S, Chen Y, Wu Z, Chen Y, Gao H, Lv L, Guo F, Zhang X, Luo R, Huang C, Lu H. Alterations of the gut microbiota in patients with coronavirus disease 2019 or H1N1 influenza. Clin Infect Dis 2020; 71: 2669-2678.
25) Gou W, Fu Y, Yue L, Chen GD, Cai X, Shuai M, Xu F, Yi X, Chen H, Zhu Y, Xiao ML. Gut microbiota, inflammation, and molecular signatures of host response to infection. J Genet Genomics 2020; 48: 792-802.
26) Shen Z, Xiao Y, Kang L, Ma W, Shi L, Zhang L, Zhou Z, Yang J, Zhong J, Yang D, Guo L. Genomic diversity of severe acute respiratory syndrome–coronavirus 2 in patients with coronavirus disease 2019. Clin Infect Dis 2020; 71: 713-720.
27) Fan J, Li X, Gao Y, Zhou J, Wang S, Huang B, Wu J, Cao Q, Chen Y, Wang Z, Luo D. The lung tissue microbiota features of 20 deceased patients with COVID-19. J Infect 2020; 81: e64-e67.
28) De Lastours V, Malosh R, Ramadugu K, Srinivasan U, Dawid S, Ohmit S, Foxman B. Co-colonization by Streptococcus pneumoniae and Staphylococcus aureus in the throat during acute respiratory illnesses. Epidemiol Infect 2016; 144: 3507-3519.
29) Lee KH, Gordon A, Shedden K, Kuan G, Ng S, Balmaseda A, Foxman B. The respiratory microbiome and susceptibility to influenza virus infection. PLoS One 2019; 14: e0207898.

30) Trapnell C, Pachter L, Salzberg SL. TopHat: discovering splice junctions with RNA-Seq. Bioinformatics 2009; 25: 1105-1111.
31) Kim D, Langmead B, Salzberg SL. HISAT: a fast spliced aligner with low memory requirements. Nat Methods 2015; 12: 357-360.
32) Wood DE, Salzberg SL. Kraken: ultrafast metagenomic sequence classification using exact alignments. Genome Biol 2014; 15: 1-12.
33) Chen L, Reeve J, Zhang L, Huang S, Wang X, Chen J. GMPR: A robust normalization method for zero-inflated count data with application to microbiome sequencing data. PeerJ 2018; 6: e4600.
34) Olsen RH, Christensen H. Transcriptomics, in Christensen H. (eds) Introduction to bioinformatics in microbiology. Springer, Cham, 2018, pp 177-188.
35) Biesbroek G, Sanders EA, Roeselers G, Wang X, Caspers MP, Trzciński K, Bogaert D, Keijser BJ. Deep sequencing analyses of low density microbial communities: working at the boundary of accurate microbiota detection. PLoS One 2012; 7: e32942.
36) Choi EB, Hong SW, Kim DK, Jeon SG, Kim KR, Cho SH, Gho YS, Jee YK, Kim YK. Decreased diversity of nasal microbiota and their secreted extracellular vesicles in patients with chronic rhinosinusitis based on a metagenomic analysis. Allergy 2014; 69: 517-526.
37) Yang HJ, LoSavio PS, Engen PA, Naqib A, Mehta A, Kota R, Khan RJ, Tobin MC, Green SJ, Schleimer RP, Keshavarzian A. Association of nasal microbiome and asthma control in patients with chronic rhinosinusitis. Clin Exp Allergy 2018; 48: 1744-1747.
38) Hoque MN, Rahman MS, Ahmed R, Hossain MS, Islam MS, Islam T, Hossain MA, Siddiki AZ. Diversity and genomic determinants of the microbiomes associated with COVID-19 and non-COVID respiratory diseases. Gene Rep 2021; 23: 101200.
39) Zuo T, Zhang F, Lui GC, Yeoh YK, Li AY, Zhan H, Wan Y, Chung AC, Cheung CP, Chen N, Lai CK. Alterations in gut microbiota of patients with COVID-19 during time of hospitalization. Gastroenterology 2020; 159: 944-955.
40) Geva-Zatorsky N, Sefik E, Kua L, Pasman L, Tan TG, Ortiz-Lopez A, Yanortsang TB, Yang L, Jupp R, Mathis D, Benoist C. Mining the human gut microbiota for immunomodulatory organisms. Cell 2017; 168: 928-943.
41) Schink B. Synergistic interactions in the microbial world. Antonie Van Leeuwenhoek 2002; 81: 257-261.
42) Vaziri ND, Wong J, Pahl M, Piceno YM, Yuan J, DeSantis TZ, Ni Z, Nguyen TH, Andersen GL. Chronic kidney disease alters intestinal microbial flora. Kidney Int 2013; 83: 308-315.
43) Ramakrishnan VR, Feazel LM, Gitomer SA, Ir D, Robertson CE, Frank DN. The microbiome of the middle meatus in healthy adults. PloS one 2013; 8: e85507.
44) Rivas N, Espinoza M, Loban A, Luque O, Jurado J, Henry-Hurtado N, Goodridge A. Case report: COVID-19 recovery from triple infection with Mycobacterium tuberculosis, HIV, and SARS-CoV-2. Am J Trop Med Hyg 2020; 103(4): 1597.
45) Chen Y, Wang Y, Fleming J, Yu Y, Gu Y, Liu C, Fan L, Wang X, Cheng M, Bi L, Liu Y. Active or latent tuberculosis increases susceptibility to COVID-19 and disease severity. MedRxiv 2020, ahead of print, DOI: https://doi.org/10.1101/2020.03.10.20033795.
46) Stecher B, Hardt WD. The role of microbiota in infectious disease. Trends Microbiol 2008; 16: 107-114.
47) Tamber S, Dougherty B, Nguy K. Responding to new clusters of COVID-19: Salmonella enterica serovars associated with bacteremia in Canada, 2006–2019. Can Commun Dis Rep 2021; 47: 259.
48) Dereschuk K, Apostol L, Ranjan I, Chakladar J, Li WT, Rajasekaran M, Chang EY, Ongkeko WM. Identification of lung and blood microbiota implicated in COVID-19 prognosis. Cells 2021; 10: 1452.