Lost microbes of COVID-19: Bifidobacterium, Faecalibacterium depletion and decreased microbiome diversity associated with SARS-CoV-2 infection severity

Sabine Hazan, Neil Stollman, Huseyin S Bozkurt, Sonya Dave, Andreas J Papoutsis, Jordan Daniels, Brad D Barrows, Eamonn MM Quigley, Thomas J Borody

ABSTRACT

Objective The study objective was to compare gut microbiome diversity and composition in SARS-CoV-2 PCR-positive patients whose symptoms ranged from asymptomatic to severe versus PCR-negative exposed controls.

Design Using a cross-sectional design, we performed shotgun next-generation sequencing on stool samples to evaluate gut microbiome composition and diversity in both patients with SARS-CoV-2 PCR-confirmed infections, which had presented to Ventura Clinical Trials for care from March 2020 through October 2021 and SARS-CoV-2 PCR-negative exposed controls. Patients were classified as being asymptomatic or having mild, moderate or severe symptoms based on National Institute of Health criteria. Exposed controls were individuals with prolonged or repeated close contact with patients with SARS-CoV-2 infection or their samples, for example, household members of patients or frontline healthcare workers. Microbiome diversity and composition were compared between patients and exposed controls at all taxonomic levels.

Results Compared with controls (n=20), severely symptomatic SARS-CoV-2-infected patients (n=28) had significantly less bacterial diversity (Shannon Index, p=0.0499; Simpson Index, p=0.0581), and positive patients overall had lower relative abundances of Bifidobacterium (p<0.0001), Faecalibacterium (p=0.0077) and Roseburium (p=0.0327), while having increased Bacteroides (p=0.0075). Interestingly, there was an inverse association between disease severity and abundance of the same bacteria.

Conclusion We hypothesise that low bacterial diversity and depletion of Bifidobacterium genera either before or after infection led to reduced proinflammatory function, thereby allowing SARS-CoV-2 infection to become symptomatic. This particular dysbiosis pattern may be a susceptibility marker for symptomatic severity from SARS-CoV-2 infection and may be amenable to preinfection, intrainfection or postinfection intervention.

Trial registration number NCT04031469 (PCR−) and 04359836 (PCR+).

INTRODUCTION

The abundance of Bifidobacterium decreases with increasing age and body mass index (BMI) and Bifidobacterium is the active ingredient of many probiotics. In vitro studies have demonstrated the benefits of these...
Gram-positive bacteria to include enhanced ATP production, immune modulation and competence, mucosal barrier integrity, restriction of bacterial adherence to and invasion of the intestinal epithelium and modulation of central nervous system activity. Additionally, *Bifidobacterium* have anti-inflammatory properties: *Bifidobacterium animalis*, *B. longum* and *B. bifidum* decrease the function of the ‘master switch’ proinflammatory tumour necrosis factor-α (TNF-α), increase the anti-inflammatory cytokine IL-10 and promote the Th1 while inhibiting the Th2 immune response. In a mouse model of inflammatory bowel disease (IBD), *B. bifidum* and *B. animalis* reduced proinflammatory cytokines and restored intestinal barrier integrity.

With respect to SARS-CoV-2 infection, there is immunologic coordination between the gut and lungs. Numerous studies have suggested that a healthy gut microbiome may be associated with a decrease in SARS-CoV-2-related mortality, and that probiotics should be considered for prophylaxis and/or treatment of SARS-CoV-2 or its associated secondary infections. However, as of February 2022, despite the publication of nearly 8000 studies on SARS-CoV-2 infection, few ongoing studies (clinicaltrials.gov: NCT04443075 and NCT04486482) and only five publications to date have examined gut microbiome changes in SARS-CoV-2 infected patients. Nevertheless, an association between the status of the gut microbiome and outcome from this infection has been suggested. Accordingly, increased abundances of the *Streptococcus*, *Rothia*, *Veillonella* and *Actinomyces* genera were associated with inflammation, whereas increased abundances of *Collinsella aerofaciens*, *Collinsella tanakaei*, *Streptococcus infantis* and *Morganella morgani* were associated with faecal samples with high SARS-CoV-2 infectivity, and increased *Lachnospiraceae* and *Enterobacteriaceae* abundances were associated with increased mortality and need for artificial ventilation. Species potentially protective against SARS-CoV-2 infection include *Parabacteroides merdae*, *Bacteroides stercoreus*, *Allistipes onderdonkii*, *Lachnospiraceae bacterium* and *F. prausnitzii*, while vulnerability to infection and increased severity were associated with decreased abundance of *B. bifidum*. A recent study correlated aspects of the gut microbiome with ‘Long-COVID’, including reduced levels of *F. prausnitzii* on admission. In short, there is still a compelling need to elucidate changes in the human gut microbiome due to SARS-CoV-2 and their relationship with clinical outcomes.

The scientific community and lay public are increasingly interested in the therapeutic potential of probiotics. Bifidobacteria have potential to improve clinical conditions ranging from IBD to *Clostridiods difficile* infections. Treatment with specific strains of *Bifidobacterium in vitro* has been shown to reduce toxins from *C. difficile*. In vivo, *Bifidobacterium* can restore colonic integrity and *B. longum* administered intranasally in mice prior to exposure to influenza has been associated with reduction in mortality. Given that *Bifidobacterium* are common component of several probiotic products and appear to be associated with SARC-CoV-2 infections, one could ask if probiotics might have a role in SARS-CoV-2 therapy or prevention.

Herein, we evaluate the relationships between gut microbiome diversity and composition compared with clinical outcome in cross-sectional groups of SARS-CoV-2 PCR-confirmed positive patients (ranging from asymptomatic to severely symptomatic) versus SARS-CoV-2 PCR-confirmed negative controls. Our controls are SARS-CoV-2 exposed persons who remained PCR negative and asymptomatic. The controls likely had similar viral exposures, but appeared protected against infection, and our data suggest that some protection may reside in the microbiome.

**METHODS**

**Study design and patients**

Individuals who were tested for SARS-CoV-2 infection either because they were symptomatic or had been exposed to a ‘case’ were eligible for enrolment the week following testing if either they or a household member was positive. Controls eligible for enrolment were PCR negative for SARS-CoV-2, remained antibody negative for 3–6 months and asymptomatic for 6–12 months. Additionally, controls had to either share a household with at least one symptomatic SARS-CoV-2-positive individual or be a healthcare worker who had been repeatedly exposed to symptomatic SARS-CoV-2-positive patients or numerous SARS-CoV-2-positive samples. All exposed controls were ones that, despite exposure to SARS-CoV-2, chose not to quarantine or take prophylaxis for SARS-CoV-2 infection and none had yet been vaccinated. Patients did not wear Personal Protective Equipment (PPE) inside their homes and staff did not wear full PPE (ie, did not wear masks) at the office because of its scarcity during this global pandemic. Patients undergoing treatment with total parenteral nutrition, or those with a history of significant gastrointestinal surgery (eg, bariatric surgery, total colectomy with ileorectal anastomosis, proctocolectomy, postoperative stoma, ostomy or ileoanal pouch) were excluded.

This study was performed between 1 March 2020 and 31 October 2021, with all but one subject recruited prior to 1 June 2021. During that time, alpha and epsilon variants predominated in the USA.
SARS-CoV-2-infected patients were categorised as asymptomatic carriers. Patients and controls were classified as underweight, normal weight, overweight, obese or severely obese based on BMI criteria of the Center for Disease Control and Prevention.\textsuperscript{31}

Stool sample collection and processing

Patients and controls within the same household collected stool samples within a week of the index case being positive. Patients had stool samples collected at baseline, prior to any treatment, and within 48 hours of symptom onset. No subjects had been commenced on antibiotics, SARS-CoV-2 infection treatments, over the counter (OTC) remedies (e.g., vitamins, antipyretics, analgesics) or supplemental oxygen between the time of symptom onset or demonstration of PCR positivity and stool collections. Subjects were instructed (and educated on the procedure and sterile methods) to collect 1 mL of fresh stool and place it directly in a DNA/RNA Shield Fecal Collection Tube (Zymo Research, Tustin, California) and then mix sample thoroughly. Here, 1 mL of faeces is more than sufficient to capture the microflora of the gut accurately and consistently. This method is chosen to eliminate the need for whole stool mixing and aliquoting. The solution in the Fecal Collection Tube is designed to preserve samples at ambient temperature (4°C–25°C) for $\geq$2 years, or below –20°C indefinitely. Once samples reached our laboratory, they were immediately frozen at –20°C.

Following faecal collection, each individual sample DNA was extracted and purified with the Qiagen PowerFecal Pro DNA extraction kit. The isolated DNA was then quantitated using the Quantus Fluorometer with the QuantFluor ONE dsDNA kit. After DNA quantification, the DNA was normalised, that is, all samples begin library preparation (following DNA extraction and purification) with 100 ng of input DNA. Libraries were then prepared using shotgun methodology with Illumina’s Nextera Flex kit. Samples then underwent the shotgun metagenomic processing procedure of tagmentation, amplification, indexing and purification. Following completion of this shotgun metagenomic standard protocol, purified libraries were again normalised to standardise sequencing depth during the next-generation sequencing (NGS) run on the NextSeq 500/550. We achieved consistency of sequencing depth (ie, number of reads) by normalising the samples’ pooling concentrations (ie, molarity), loading the same number of samples per sequencing run, consistently using the same NextSeq High Output kits.

After completion of sequencing on the Illumina NextSeq with 500/550 High-Output Kits V.2.5 (300 cycles), the raw data were streamed in real time to Illumina’s BaseSpace cloud for FASTQ (Fast Quality, a standard text file type for storing biological sequence information) conversion. The FASTQ files were then sent through One Codex’s bioinformatics pipeline for metagenomic annotation and analyses to elucidate the microflora composition and relative abundances of the top genera and species for all patients and controls.

Data analysis

We assessed differences in relative abundance across taxa between the gut microbiome of SARS-CoV-2-infected patients and exposed controls and calculated Shannon and Simpson alpha diversity indices with One Codex’s bioinformatics analysis pipeline using Jupyter Notebook in Python. Specifically, the One Codex Database consists of $\sim$114K complete microbial genomes (One Codex, San Francisco, California). During processing, reads were first screened against the human genome and then mapped to the microbial reference database using a k-mer-based classification. Individual sequences (NGS read or contig) were compared against the One Codex Database (One Codex) by exact alignment using k-mers, where $k=31$. Based on the relative frequency, unique k-mers were filtered to control for sequencing or reference genome artefacts.

The sequencing depth followed ProgenaBiome’s standard operating procedures and was $8.29\times10^9$ reads on average for this study. One should note that shallow metagenomic sequencing is typically only 0.5 million reads but is still considered sufficient for taxonomic phyla level analysis (and even genera for the most abundant bacteria).

The relative abundance of each microbial taxonomic classification was estimated based on the depth and coverage of sequencing across every available reference genome. Beta-diversity was calculated as weighted UniFrac distance visualised in a distance matrix using the phylum-level relative abundance obtained from One Codex. Thirteen genera were selected based on our experience and knowledge of critical players in the gut microbiome as well as similarity to other studies.\textsuperscript{18,29}

To compare patients across subgroups and patients to exposed controls, Analysis of Variation (ANOVA), Mann-Whitney U, Kruskal-Wallis tests and $\chi^2$ test statistics were conducted using GraphPad V.8 with p values $\leq$0.05 considered as significant. Dunn’s $\textit{post-hoc}$ was used for Kruskal-Wallis test, with correction for multiple comparisons in all situations.

All authors had access to study data and reviewed and approved the final manuscript.

RESULTS

Patient characteristics

Demographic and clinical characteristics of patients ($n=50$) and exposed controls ($n=20$) are presented in online supplemental tables 1 and 2, and summarised in table 1. All patients were resident of USA, with states indicated in online supplemental table 1. Twenty-four of 50 (48%) patients and 7 of 20 (35%) of exposed controls were men. The mean±SEM age in years was 50.0±2.5 for patients and 44.4±3.6 for exposed controls. Forty-four of 50 (88%) patients were non-Hispanic white; 5 of 50
Table 1  Summary of demographics along with clinical and dietary characteristics of subjects

<table>
<thead>
<tr>
<th></th>
<th>USA resident</th>
<th>Male</th>
<th>Mean age ±SEM</th>
<th>Median age</th>
<th>Race</th>
<th>Severe COVID-19 comorbidities*</th>
<th>Normal stool frequency</th>
<th>Lack of appetite</th>
<th>Omnivore diet</th>
</tr>
</thead>
<tbody>
<tr>
<td>Demographics</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>SARS-CoV-2 negative;</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>exposed control</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n=20)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Asymptomatic (n=4)</td>
<td>20 (100.00%)</td>
<td>7 (35.00%)</td>
<td>44.40±3.62</td>
<td>48</td>
<td>White 17 (85.00%)</td>
<td>12 (60.00%)</td>
<td>20 (100.00%)</td>
<td>0 (0.00%)</td>
<td>19† (100.00%)</td>
</tr>
<tr>
<td>Mild (n=6)</td>
<td>4 (100.00%)</td>
<td>3 (75.00%)</td>
<td>47.50±9.40</td>
<td>50.5</td>
<td>Hispanic 2 (10.00%)</td>
<td>3 (75.00%)</td>
<td>4 (100.00%)</td>
<td>0 (0.00%)</td>
<td>4 (100.00%)</td>
</tr>
<tr>
<td>Moderate (n=12)</td>
<td>6 (100.00%)</td>
<td>2 (33.33%)</td>
<td>37.67±4.65</td>
<td>37</td>
<td>Black 1 (5.00%)</td>
<td>5 (83.33%)</td>
<td>6 (100.00%)</td>
<td>0 (0.00%)</td>
<td>6 (100.00%)</td>
</tr>
<tr>
<td>Severe (n=28)</td>
<td>12 (100.00%)</td>
<td>5 (41.66%)</td>
<td>50.5±5.92</td>
<td>60</td>
<td>Other 0 (0.00%)</td>
<td>9 (75.00%)</td>
<td>10 (83.33%)</td>
<td>0 (0.00%)</td>
<td>10 (100.00%)</td>
</tr>
<tr>
<td>Total (n=50)</td>
<td>20 (100.00%)</td>
<td>7 (35.00%)</td>
<td>44.40±3.62</td>
<td>50.5</td>
<td></td>
<td>32 (64.00%)</td>
<td>44 (88.00%)</td>
<td>3 (6.00%)</td>
<td>50 (100.00%)</td>
</tr>
<tr>
<td>P value</td>
<td>ns</td>
<td>0.5706</td>
<td>0.5221</td>
<td>0.3826</td>
<td></td>
<td>0.5099</td>
<td>0.2891</td>
<td>0.6663</td>
<td></td>
</tr>
</tbody>
</table>

Clinical/dietary characteristics

|                        |              |      |               |            |                 |                               |                       |                 |              |
|------------------------|--------------|------|---------------|------------|-----------------|-------------------------------|                       |                 |              |
| SARS-CoV-2 positive    |              |      |               |            |                 |                               |                       |                 |              |
| Asymptomatic           |              |      |               |            |                 |                               |                       |                 |              |
| (n=4)                  |              |      |               |            |                 |                               |                       |                 |              |
| Mild                   |              |      |               |            |                 |                               |                       |                 |              |
| (n=6)                  |              |      |               |            |                 |                               |                       |                 |              |
| Moderate               |              |      |               |            |                 |                               |                       |                 |              |
| (n=12)                 |              |      |               |            |                 |                               |                       |                 |              |
| Severe                 |              |      |               |            |                 |                               |                       |                 |              |
| (n=28)                 |              |      |               |            |                 |                               |                       |                 |              |
| Total                  |              |      |               |            |                 |                               |                       |                 |              |

*Numbers in cells indicate number of subjects with percentage in categories, except for age and BMI, which indicate value. P values calculated via one-way ANOVA (age and BMI) or χ² (others). Normal stool frequency is defined as absence of diarrhoea. Total refers to sum of SARS-CoV-2-positive subjects, and it is not used in statistics.

†Nineteen total subjects had data available for diet within exposed control.
(10%) were Hispanic and 1 of 50 (2%) Native American and 17 of 20 (85.0%) of exposed controls were non-Hispanic white; 2 of 20 (10.0%), Hispanic and 1 of 20 (5.0%), Black. Of patients, 28 of 50 (56%) had severe, 12 of 50 (24%) had moderate and 6 of 50 (12%) had mild disease and 4 of 50 (8%) were asymptomatic. Thirty-two of 50 (64%) patients and 12 of 20 (60.0%) exposed controls had underlying comorbidities considered risk factors for increased severity of SARS-CoV-2 infection by the Center for Disease Control (CDC). The mean±SEM BMI of the 46 patients with available data was 27.1±0.98 compared with 25.1±0.96 for the 20 exposed controls. There was no significant difference (p>0.2) in gender, age, racial demographics, loss of appetite, change in stool frequency, diet or presence of underlying comorbidities.

Of the exposed controls, 16 were household contacts of SARS-CoV-2-positive patients in the study, 2 were healthcare workers with extensive, non-protected, exposure to SARS-CoV-2-positive patients and 2 were laboratory personnel exposed to thousands of SARS-CoV-2 samples (healthcare workers and laboratory personnel did not wear full PPE, that is, did not wear a face mask, due its scarcity; see the Methods section). During the timeframe of the study, none of the patients or controls was on SARS-CoV-2 prophylaxis or treatment, and none had yet been vaccinated. No patients or exposed control were positive for SARS-CoV-2 prior to the study.

**Gut microbiome diversity and composition**

Figure 1 depicts pie charts of the composition of the gut microbiome for the exposed control at the phylum level (figure 1A) and genus level (figure 1B). At phylum level, *Firmicutes* and *Bacteroides* dominated, comprising 59.6% (exposed control) and 54.7% (SARS-CoV-2 positive) and 29.1% (exposed control) and 40.4% (SARS-CoV-2 positive) of all phyla, respectively. At the level of genus, *Bacteroides* contributed 12.4% (exposed control) and 21.8% (SARS-CoV-2 positive), *Alistipes* 6.4% (exposed control) and 7.2% (SARS-CoV-2 positive) and *Bifidobacterium* 7.6% (exposed control) and 1.5% (SARS-CoV-2 positive).

Figure 2 shows two diversity indices for all subgroups studied, namely, Shannon diversity (figure 2A) and Simpson diversity index (figure 2B). The overall p value for Shannon index (richness of bacterial composition) demonstrated a significant (p=0.0499) decrease in diversity with increased severity, and significance was seen for exposed control versus severely symptomatic (p=0.0201), analysed via Kruskal-Wallis test. The Simpson (evenness of bacterial composition) indexes showed a trend (p=0.0581) of a decrease in diversity with increased SARS-CoV-2 severity.

Further metagenomic analysis comparing SARS-CoV-2 patients and controls revealed significant differences in relative abundance of specific bacteria. The relative abundance of SARS-CoV-2 positive (exposed control) versus negative subjects is presented in table 2, along with comparative p values via Mann-Whitney U test. Patients with SARS-CoV-2 infection showed a significantly decreased relative abundance of *Bifidobacterium* and *Faecalibacterium*, and significantly increased relative abundance of *Bacteroides* (table 2).

Table 3 lists the genera/species relative abundances (mean±SEM) for various levels of severity of SARS-CoV-2-positive patients versus exposed control. Analysed via Kruskal-Wallis test, the main effect (ie, overall p value) of these changes are shown in the left column. Table 3 proceeds to compare, correcting for multiple comparison, the three levels of severity in infected patients versus exposed control and asymptomatic groups. Specifically, increased disease severity was associated with decreased relative abundance of *Bifidobacterium*, *Faecalibacterium*, *F*
Table 2. Relative abundances of Bacteroides increase and of Bifidobacterium, Faecalibacterium and Roseburium decrease in SARS-CoV-2 positive subjects versus SARS-CoV-2 negative exposed controls

<table>
<thead>
<tr>
<th>Genus (±species)</th>
<th>Exposed controls (mean±SEM)</th>
<th>SARS-CoV-2 positive (mean±SEM)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Allstipes</td>
<td>0.0639±0.0095</td>
<td>0.0721±0.0100</td>
<td>0.8709</td>
</tr>
<tr>
<td>Bacteroides</td>
<td>0.1235±0.0178</td>
<td>0.2183±0.0191</td>
<td>0.0025</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>0.0755±0.0219</td>
<td>0.0147±0.0051</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>Blautia</td>
<td>0.0261±0.0040</td>
<td>0.0524±0.0088</td>
<td>0.1349</td>
</tr>
<tr>
<td>Clostridium</td>
<td>0.0431±0.0075</td>
<td>0.0309±0.0039</td>
<td>0.9948</td>
</tr>
<tr>
<td>Collinsella</td>
<td>0.0146±0.0045</td>
<td>0.0158±0.0029</td>
<td>0.9948</td>
</tr>
<tr>
<td>Dorea</td>
<td>0.0137±0.0024</td>
<td>0.0185±0.0022</td>
<td>0.2777</td>
</tr>
<tr>
<td>Eubacterium</td>
<td>0.0441±0.0063</td>
<td>0.0402±0.0043</td>
<td>0.4786</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>0.0550±0.0086</td>
<td>0.0310±0.0039</td>
<td>0.0137</td>
</tr>
<tr>
<td>F. prausnitzii</td>
<td>0.0542±0.0085</td>
<td>0.0313±0.0039</td>
<td>0.0153</td>
</tr>
<tr>
<td>Prevotella</td>
<td>0.0110±0.0086</td>
<td>0.0091±0.0066</td>
<td>0.6538</td>
</tr>
<tr>
<td>Roseburium</td>
<td>0.0329±0.0056</td>
<td>0.0195±0.0032</td>
<td>0.0097</td>
</tr>
<tr>
<td>Ruminococcus</td>
<td>0.0376±0.0079</td>
<td>0.0391±0.0056</td>
<td>0.9844</td>
</tr>
</tbody>
</table>

Mean±SEM relative abundances, as well as p value via Mann-Whitney U test are indicated, with bold marking p<0.05.

Depicted in figure 3 are the 12 most abundant families and the 12 most abundant genera for patients, stratified by disease severity and in comparison to exposed controls. Distinguished by colour, the bars represent the relative per cent bacterial families and genera abundance. Note the reduced diversity of the microbiome of SARS-CoV-2-positive patients shown in column B.

Figure 4 summarises the microbiome changes according to SARS-CoV-2 positivity and severity, with green boxes depicting significant elevation and red boxes indicating significant depletion in genera/species abundance associated with SARS-CoV-2.

Figure 5A, B exhibit the relative abundance of Bifidobacterium for each subject, grouped by SARS-CoV-2 infection severity. This diagram, with subject groups ordered by severity (from severe on left, to exposed controls on the right), depicts how Bifidobacterium abundance increases as severity decreases.

Analysis of the beta-diversity of subjects demonstrated that the diversity of control subjects cluster separately from that of patients. Figure 6A shows the beta-diversity-weighted (quantitative) UniFrac analysis featuring phyla bacterial profiles for all individuals in the study (n=70). Figure 6B reveals that, although there is a range of dissimilarity, the SARS-CoV-2-negative individuals are more similar to one another than they are to SARS-CoV-2-positive patients. The matrix also highlights clusters of similarity among SARS-CoV-2-positive patients, and darker quadrants of dissimilarity where positive and negative patients intersect. At a more granular level, figure 6B used principal component (PC) analysis of genera, where the axes depict the per cent of variance. In PC analysis, points closer together are more similar (less divergence with axis representing directions of divergence). Herein, the PC1 accounts for 43.16% of the variation, whereas PC2 accounts for 12.78%. This analysis reveals a clear divergence of a subset of SARS-CoV-2-positive patients clustering on the right side tracking along the x-axis (PC1), highlighting microbiota divergence as a function of disease. Thus, figure 6 shows that exposed controls cluster similar separately from SARS-CoV-2 patients; that is, patients are more similar in terms of their microbiome to each other than to controls.

**DISCUSSION**

**Immune function and health could be enhanced by bacterial abundance**

Interactions between the host and gut microbiota are complex, numerous and bidirectional. Gut microbiota regulate the development and function of the innate and adaptive immune systems, potentially allowing them to protect against infections and infection severity. The primary findings of our study are that SARS-CoV-2 positivity and infection severity are associated with decreased levels of the protective Bifidobacterium and Faecalibacterium genera and with decreased bacterial diversity, as exemplified by the Shannon and Simpson indices. This accords with studies showing bacterial diversity inversely relates to the presence of various common disorders. Uniquely, our study compared SARS-CoV-2-exposed SARS-CoV-2-negative persons (ie, controls) with symptomatic and asymptomatic SARS-CoV-2-positive patients. Thus, we controlled for SARS-CoV-2 exposure.

The genus Bifidobacterium has important immune functions, is a major component of the microbiome and is frequently used in probiotics. Bifidobacterium increase Treg responses and reduce cell damage by inhibiting
### Table 3 Relative abundance of *Bacteroides* increases and those of *Bifidobacterium*, *Faecalibacterium* and *Roseburium* decrease with increasing severity of disease

<table>
<thead>
<tr>
<th>Genus (species)</th>
<th>Main effect</th>
<th>Severe</th>
<th>Moderate</th>
<th>Mild</th>
<th>Asymptomatic</th>
<th>Exposed control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alistipes</td>
<td>0.8119</td>
<td>0.0793±0.0166</td>
<td>0.0520±0.0084</td>
<td>0.0664±0.0196</td>
<td>0.0901±0.0272</td>
<td>0.0839±0.0095</td>
</tr>
<tr>
<td>Bacteroides</td>
<td>0.0075</td>
<td>0.2187±0.0272</td>
<td>0.2849±0.0305</td>
<td>0.0912±0.0226</td>
<td>0.2058±0.0606</td>
<td>0.1235±0.0178</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>&lt;0.0001</td>
<td>0.0018±0.0006</td>
<td>0.0037±0.0015</td>
<td>0.0507±0.0244</td>
<td>0.0840±0.0308</td>
<td>0.0755±0.0219</td>
</tr>
<tr>
<td>Blautia</td>
<td>0.2098</td>
<td>0.0495±0.0095</td>
<td>0.0426±0.0132</td>
<td>0.1037±0.0510</td>
<td>0.0260±0.0067</td>
<td>0.0261±0.0040</td>
</tr>
<tr>
<td>Clostridium</td>
<td>0.2721</td>
<td>0.0331±0.0062</td>
<td>0.0334±0.0055</td>
<td>0.0265±0.0083</td>
<td>0.0145±0.0025</td>
<td>0.0431±0.0075</td>
</tr>
<tr>
<td>Collinsella</td>
<td>0.7476</td>
<td>0.0145±0.0035</td>
<td>0.0107±0.0038</td>
<td>0.0203±0.0126</td>
<td>0.0334±0.0174</td>
<td>0.0146±0.0045</td>
</tr>
<tr>
<td>Dorea</td>
<td>0.4820</td>
<td>0.0218±0.0035</td>
<td>0.0126±0.0032</td>
<td>0.0174±0.0046</td>
<td>0.0152±0.0015</td>
<td>0.0137±0.0024</td>
</tr>
<tr>
<td>Eubacterium</td>
<td>0.9619</td>
<td>0.0436±0.0070</td>
<td>0.0367±0.0064</td>
<td>0.0366±0.0085</td>
<td>0.0332±0.0058</td>
<td>0.0441±0.0063</td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>0.0077</td>
<td>0.0209±0.0037</td>
<td>0.0428±0.0093</td>
<td>0.0359±0.0139</td>
<td>0.0597±0.0098</td>
<td>0.0550±0.0086</td>
</tr>
<tr>
<td><em>F. prausnitzii</em></td>
<td>0.0054</td>
<td>0.0220±0.0041</td>
<td>0.0417±0.0092</td>
<td>0.0356±0.0137</td>
<td>0.0589±0.0095</td>
<td>0.0542±0.0085</td>
</tr>
<tr>
<td>Prevotella</td>
<td>0.1687</td>
<td>0.0149±0.0118</td>
<td>0.0000±0.0000</td>
<td>0.0050±0.0046</td>
<td>0.0023±0.0020</td>
<td>0.0110±0.0086</td>
</tr>
<tr>
<td>Roseburium</td>
<td>0.0327</td>
<td>0.0146±0.0037</td>
<td>0.0245±0.0075</td>
<td>0.0268±0.0138</td>
<td>0.0274±0.0081</td>
<td>0.0329±0.0056</td>
</tr>
<tr>
<td>Ruminococcus</td>
<td>0.8033</td>
<td>0.0384±0.0076</td>
<td>0.0293±0.0073</td>
<td>0.0620±0.0250</td>
<td>0.0304±0.0139</td>
<td>0.0376±0.0079</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genus (species)</th>
<th>P value vs exposed control</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
<th>Asymptomatic</th>
<th>P value vs asymptomatic</th>
<th>Mild</th>
<th>Moderate</th>
<th>Severe</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alistipes</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td></td>
</tr>
<tr>
<td>Bacteroides</td>
<td>&gt;0.9999</td>
<td>0.0078</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td></td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>&gt;0.9999</td>
<td>0.0006</td>
<td>0.0026</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>0.0006</td>
<td>&lt;0.0001</td>
<td></td>
</tr>
<tr>
<td>Blautia</td>
<td>0.4116</td>
<td>&gt;0.9999</td>
<td>0.6373</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td></td>
</tr>
<tr>
<td>Clostridium</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>0.4944</td>
<td>&gt;0.9999</td>
<td>0.9841</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td></td>
</tr>
<tr>
<td>Collinsella</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td></td>
</tr>
<tr>
<td>Dorea</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td></td>
</tr>
<tr>
<td>Eubacterium</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td></td>
</tr>
<tr>
<td>Faecalibacterium</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>0.0082</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>0.1422</td>
<td></td>
</tr>
<tr>
<td><em>F. prausnitzii</em></td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>0.0109</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>0.1568</td>
<td></td>
</tr>
<tr>
<td>Prevotella</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>0.8916</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td></td>
</tr>
<tr>
<td>Roseburium</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>0.0196</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td></td>
</tr>
<tr>
<td>Ruminococcus</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td>&gt;0.9999</td>
<td></td>
</tr>
</tbody>
</table>

For A and B, bold values indicate p<0.05. Note, no apparently significant (p<0.05) p values were observed in post-hocs, with main effects non-significant.
TNF-α and macrophages. Bifidobacterium also protects against intestinal epithelial cell damage independently from their effects on TNF-α production. The exopolysaccharide coat, which is a feature of some Bifidobacterium, plays a significant role in this protective effect. These immune functions of Bifidobacterium could be critical in relation to its SARS-CoV-2 infection-prevention effects.

Evidence has accumulated to support a beneficial effect from supplementation with Bifidobacterium in numerous disease states. The numbers of commensal Bifidobacterium have been shown to decrease with age and obesity, major SARS-CoV-2 infection risk factors. We demonstrate that patients with a more severe course of viral infection had decreased abundance of Bifidobacterium. However, it should be noted that there are no definitive studies concerning what constitutes a normal baseline abundance of Bifidobacterium in a ‘healthy’ individual.

**Figure 3** Graphic of relative abundance of the 12 most common (A) families and (B) genera. The top group represents the SARS-CoV-2 positive samples (n=50), stratified by severity. The bottom group represents the exposed control samples (n=20). The coloured boxes represent the fraction of the entire rectangle composed of the given family/genera of bacteria.
The abundance of *Faecalibacterium* genus and *F. prausnitzii* species was also inversely related to SARS-CoV-2 positivity and SARS-CoV-2 infection severity in this analysis. Age and diabetes are risk factors for SARS-CoV-2 infection, and *F. prausnitzii* levels decline markedly in elder and diabetic populations. In fact, *Faecalibacterium* levels have been considered an indirect ‘indicator’ of overall human health. The abundance of *F. prausnitzii* is reduced by the ‘Western’ diet (consumption of more meat, animal fat, sugar, processed foods and low fibre), while it is enhanced by the high-fibre containing ‘Mediterranean’ diet of vegetables and fruits with low meat intake. Preliminary studies showed that reduced ingestion of a Mediterranean diet within the same country is associated with increased SARS-CoV-2-related death rates. In short, we show that *F. prausnitzii* levels negatively correlated to SARS-CoV-2 infection severity and prior studies show that reduced *F. prausnitzii* is associated with SARS-CoV-2 infection vulnerabilities such as age, diabetes, obesity and possibly diet.

SARS-CoV-2 positivity and severity were also associated with decreased abundance of *Roseburium* and increased abundance of *Bacteroides*. The implications of these changes remain unclear.

**Innate immunity could be enhanced by increased level of beneficial bacteria**

The pathological impact of SARS-CoV-2 infection includes both direct effects from viral invasion and complex immunological responses including, in its most severe form, the ‘cytokine storm’. The cytokine storm is the result of a sudden increase in circulating levels of proinflammatory cytokines produced by activated macrophages, mast cells, endothelial cells and epithelial cells during innate immune responses, which appear to be modulated by the abundance of *Bifidobacterium* and *Faecalibacterium* and bacterial diversity (5, 23, 25). Steroid treatment has situational success in SARS-CoV-2 infection, based on suppressing this over activation of the innate immune system, reviewed by Tang et al

Zhao reported that elevated serum levels of proinflammatory cytokines such as IL-16 and IL-17 predict poor prognoses in patients with SARS-CoV-2 infection. Also, Tao et al showed that changes in gut microbiota composition might contribute to SARS-CoV-2-induced
production of inflammatory cytokines in the intestine, which may lead to cytokine storm onset. Both authors report significantly reduced gut microbiota diversity and increased opportunistic pathogens in patients with SARS-CoV-2. Interestingly, the bloom of opportunistic pathogens positively correlated with the number of Th17 cells. Bozkurt and Quigley reported that IL-6 and IL-17 promote viral persistence by immune interactions through cellular autophagy via the inositol-requiring enzyme 1 pathway. Additionally, some species of Bifidobacterium are likely to prevent the replication of coronaviruses by reducing endoplasmic reticulum stress, also through the inositol-requiring enzyme 1 pathway. Reduced Bifidobacterium abundance has been observed in the gut microbiome of patients with IBD, which has mechanisms involving IL-17. Furthermore, the direct endoscopic delivery of Bifidobacterium has been shown to be effective in promoting symptom resolution and

Figure 6 SARS-CoV-2 positive patients' microbiome is more similar to each other than to that of exposed controls. (A) Weighted UniFrac distance matrix of phylum level SARS-CoV-2 positive (n=50) and exposed negative control samples (n=20). Distance of microbiome differences increases with increasing blue colour intensity (see legend top right). The centre of the diagram consists of negative subjects on both axis and is yellow indicative of less distance (ie, less difference in microbiome). The central area of the left as well as central-top side of diagram, consists of negative subjects on one axis and positive on the other, and are darker blue, indicative of more distance (more difference in microbiome). (B) Principal component analysis of microbiota from SARS-CoV-2 positive (n=50) and exposed negative controls (n=20). Dots closer in distance are more similar in microbiome composition. Axes depict the per cent of variance explained by principal component (PC) 1 and 2. Plots are based on bacterial genera relative abundance profiles.
mucosal healing in IBD—an effect likely to be associated with the anti-Th17 effect of *Bifidobacterium*. (8) Figure 7 demonstrates how *Bifidobacterium* might hypothetically quell a heightened immune response by dampening the effect of the master switch TNF-α.

CONCLUSIONS
Given our cross-sectional study design, it is not possible to determine whether the differences in *Bifidobacterium* levels observed between patients and exposed controls preceded or followed infection. If preceding infection, they could be a marker of susceptibility, and boosting *Bifidobacterium* levels might decrease the risk or severity of SARS-CoV-2 infection. If these changes followed infection, alteration of the gut microbiome (such as through faecal microbiota transplantation or possibly probiotic supplementation) to increase *Bifidobacterium* could be an area worth exploring for improved outcomes. If future studies can demonstrate improved outcomes, such therapy can be considered for complex cases of SARS-CoV-2 infection. Developing outbreaks within tightly closed communities such as nursing homes might be a good setting in which to assess susceptibility; faecal samples could be collected during the outbreak and run post hoc on ‘cases’ and ‘controls’. Future studies of individuals with baseline prepanademic microbiome data would be highly valuable, although acquiring such baseline preinfection microbiome data is still costly.

With the lack of data on the gut microbiome prior to onset of SARS-CoV-2 infection, we cannot completely rule out the confounding effect of illness on the microbiome. Nonetheless, we eliminated effects of treatment on the gut microbiome by sampling prior to administration of SARS-CoV-2 infection therapeutics of any kind and within 48 hours of symptom onset. Specifically, no subjects were given antibiotics, antivirals, anti-inflammatory medicines, oxygen or any other therapeutic agent between symptom onset or PCR positivity and stool sampling. We also note that the prevalence of appetite changes, alterations of stool frequency and gastrointestinal (GI) symptoms, in general, were not significantly different between any of the severity groups or controls (table 1), although the small sample sizes for some groups should be considered in evaluating these statistics.

SARS-CoV-2 infection presentation variability correlates with colon microbiome bacterial composition and overall diversity. The same changes we observe due to SARS-CoV-2 infection, namely reduced *Bifidobacterium* and/or *Faecalibacterium* abundance, are associated with SARS-CoV-2 infection risk factors, including old age, obesity and diabetes. (9 37 39 43) Thus, colon microbiome diversity and relative abundance of *Bifidobacterium* and *Faecalibacterium* should be explored as potential markers for predicting SARS-CoV-2 infection severity.

In summary, we demonstrate in a study of PCR-positive and PCR-negative SARS-CoV-2-exposed subjects, reduced bacterial diversity and reduced levels of various genus/species are highly associated with both SARS-CoV-2

---

**Figure 7** Potential mechanism for cytokine storm and immune hyper-response in SARS-CoV-2 positive patients. In individuals infected with SARS-CoV-2, the macrophages become activated; these in turn activate T-cells, additional macrophages, and neutrophils—all of which release cytokines, including TNF-α. *Bifidobacterium*, when present in sufficient numbers, can bind to TNF-α and prevent the subsequent cytokine storm. Therefore, patients with a bifidobacterial dysbiosis characterised by low levels of *Bifidobacterium* lack this line of defense, which may lead to a cytokine storm.
positivity and SARS-CoV-2 infection severity. These findings suggest that probiotic supplementation or faecal microbiota transplantation should be explored as a potential therapeutic avenue for patients with SARS-CoV-2. Additionally, individual colon microbiome evaluation may predict vulnerability to the development of severe SARS-CoV-2 infection. Finally, our data suggest a new area for exploration: if SARS-CoV-2 severity is found to be dependent on the microbiome, then accounting for microbiome differences could reduce variability in outcomes for SARS-CoV-2.

**Author affiliations**

1N/A, ProgenaBiome LLC, Ventura, California, USA
2Division of Gastroenterology, Alta Bates Summit Medical Center, Berkeley, California, USA
3Clinic of Gastroenterology, Istanbul Maltepe University, Istanbul, Turkey
4N/A, Microbiome Research, Inc, Ventura, California, USA
5Medical Writing and Biostatistics, North End Advisory, Smyrna, Georgia, USA
6Division of Gastroenterology and Hepatology, The Methodist Hospital, Weill Cornell Medical College, Houston, Texas, USA
7N/A, Centre for Digestive Diseases, Five Dock, New South Wales, Australia

**Acknowledgements** Medical writing assistance was provided by Sonya Dave, PhD (an author on the publication) and was funded by ProgenaBiome. The authors thank all clinicians for their involvement and contribution to the study. The authors thank Kate Hendricks, MD, MPH for many helpful editorial suggestions. Finally, the authors owe a depth of gratitude to the late Sydney M Finegold, MD for mentorship that sparked the interest in the microbiome to many scientists, including authors of this paper.

**Contributors** All authors (SH, NS, HSB, SD, AJP, JD, BDB, EMMQ and TJB) participated in the drafting, critical revision, and approval of the final version of the manuscript. SH led study design. SH and AJP conducted the bioinformatic analysis. SD conducted the statistical analysis and was a major contributor to writing the paper. SH was primarily responsible for interpretation of the study results, with contributions from all authors. EMMQ and TJB are senior authors who provided overall direction and advice. SH acts as the guarantor and accepts full responsibility for the work, has access to all data, and controlled the decision to publish.

**Funding** The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

**Competing interests** SH declares that she has pecuniary interest in Topelia Pty Ltd in Australia, and Topelia Therapeutics Inc. in USA developing COVID-19 preventative/treatment options is being pursued. She has also filed patents relevant to Coronavirus treatments. She is the founder and owner of Microbiome research foundation, Progenabiome and Ventura Clinical Trials. TJB declares that he has pecuniary interest in Topelia Pty Ltd in Australia, and Topelia Therapeutics Inc. in USA developing COVID-19 preventative/treatment medications. He has also filed patents relevant to COVID-19 treatments. SD declares she has corporate affiliation to McKesson Specialty Health/Ontara and North End Advisory, LLC. SD is unaware of SARS-CoV-2 and microbiome projects and not directly involved in COVID-19 relevant projects at McKesson, but they may exist. AJP and BDB have corporate affiliations to Progenabiome. EMMQ serves as a consultant to Precisionbiotics, Novavaxes, Salix, Biocodex and Axon Pharma and has received research support from 4D Pharma.

**Patient consent for publication** Not applicable.

**Ethics approval** The study was conducted in accordance with ethical principles of the Declaration of Helsinki, the International Council for Harmonisation Harmonised Tripartite Guideline for Good Clinical Practice and the Ethical and Independent Review Board. This study involves human participants and was approved by the ‘Ethical and Independent Review Board’ (IRB, IRB00007807). Participants gave informed consent to participate in the study before taking part.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data availability statement** Data are available upon reasonable request. Data available upon reasonable request from corresponding author, Dr. Sabine Hazan.

**Supplemental material** This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

**Open access** This is an open access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited, appropriate credit is given, any changes made indicated, and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc/4.0/.

**ORCID iDs**

Sonya Dave http://orcid.org/0000-0002-7653-0388
Eamonn MM Quigley http://orcid.org/0000-0003-4151-7180

**REFERENCES**


Overview of variants in countries. Available: https://covariants.org/


