Nasopharyngeal microbial communities of patients infected with SARS-COV-2 that developed COVID-19.

Maria Paz Ventero[1*], Rafael Ricardo Castro Cuadrat[2*], Inmaculada Vidal[1], Bruno Gabriel N. Andrade[3,4], Carmen Molina-Pardines[1], Jose M. Haro-Moreno[5], Felipe H. Coutinho[5], Esperanza Merino[6], Luciana CA Regitano[4], Cynthia B Silveira[7], Haithem Afli[4], Mario López-Pérez[1,5], Juan Carlos Rodríguez[1,5§.]

[1]Microbiology Department, Alicante University General Hospital - Alicante Institute of Sanitary and Biomedical Research (ISABIAL), Alicante, Spain

[2]Department of Molecular Epidemiology, German Institute of Human Nutrition Potsdam-Rehbrücke, Nuthetal, Germany

[3]Embrapa Pecuária Sudeste, São Carlos, Brazil.

[4]Department of computer science, Cork Institute of Technology, Ireland.

[5]Evolutionary Genomics Group, División de Microbiología, Universidad Miguel Hernández, San Juan de Alicante, Spain.

[6]Infectious Diseases Unit, Alicante University General Hospital, Alicante Institute for Health and Biomedical Research (ISABIAL), Alicante, Spain.

[7]Department of Biology, University of Miami, Miami, USA

* These authors contributed equally to this work

[§]Corresponding author:

Juan Carlos Rodríguez (rodriguez_juadia@gva.es)

Running title: Nasopharyngeal microbiome associated with COVID-19

Keywords: COVID19 / SARS-CoV-2 / Microbiome

E-mails:

Maria Paz Ventero. maripazvm@gmail.com

Rafael Ricardo Castro Cuadrat. rafaelcuadrat@gmail.com

Inmaculada Vidal. vidal_inm@gva.es

Bruno Gabriel N. Andrade. bgabriel.andrade@gmail.com

Carmen Molina-Pardines. molina_carpar@gva.es

Jose M. Haro-Moreno. moreno.jose@umh.es

Felipe H. Coutinho. felipehcoutinho@gmail.com

Esperanza Merino. merinoluc@gmail.com

Luciana CA Regitano. luciana.regitano@embrapa.br

Cynthia B Silveira. cynthiabsilveira@gmail.com

Haithem Afli. haithem.afli@adaptcentre.ie

Mario López-Pérez. mario.lopezp@umh.es

Juan Carlos Rodríguez. rodriguez_juadia@gva.es

ABSTRACT

Background: SARS-CoV-2 is an RNA virus causing COVID-19. The clinical characteristics and epidemiology of COVID-19 have been extensively investigated, however studies focused on the patient's microbiota are still lacking. In this study, we investigated the nasopharyngeal microbiome composition of patients who developed different severity levels of COVID-19. We performed Rdna-SSU (16S) sequencing from nasopharyngeal swab samples obtained from SARS-CoV-2 positive (56) and negative (18) patients in the province of Alicante (Spain) in their

first visit to the hospital. Positive SARS-CoV-2 patients were observed and later categorized in mild (symptomatic without hospitalization), moderate (hospitalization) and severe (admission to ICU). We compared the microbiome diversity and OTU composition among severity groups using Similarity Percentage (SIMPER) analysis and Maaslin2. We also built bacterial co-abundance networks for each group using Fastpar.

Results: Statistical analysis indicated differences in the nasopharyngeal microbiome of COVID19 patients. 62 OTUs were found exclusively in SARS-CoV-2 positive patients, mostly classified as members of the phylum Bacteroidetes (18) and Firmicutes (25). OTUs classified as *Prevotella* were found to be significantly more abundant in patients that developed more severe COVID-19. Furthemore, co-abundance analysis indicated a loss of network complexity among samples from patients that later developed more severe symptoms.

Conclusions: Our preliminary study shows that the nasopharyngeal microbiome of COVID-19 patients showed differences in the composition of specific OTUs and complexity of co-abundance networks. These microbes with differential abundances among groups could serve as biomarkers for COVID-19 severity. Nevertheless, further studies with larger sample sizes should be conducted to validate these results.

IMPORTANCE

This work has studied the microbiota of the nasopharyngeal tract in COVID19 patients using advanced techniques of molecular microbiology. Diverse microorganisms, most of which are harmless or even beneficial to the host, colonize the nasopharyngeal tract. These microorganisms are the microbiota, and they are present in every people. However, changes in this microbiota could be related to different diseases as cancer, gastrointestinal pathologies or even COVID19. This study has been performed to investigate the microbiota from patients with COVID19, in order to determinate its implication in the pathology severity. The results

obtained showed that it is possible that several specific microorganisms are present only in patients with severe COVID19. These data, could be used as a prognostic biomarker to early detect whose patients will develop a severe COVID19 and improve their clinical management.

BACKGROUND

Severe Acute Respiratory Syndrome Coronavirus-2 (SARS-CoV-2) is a positive-sense single-stranded RNA virus causing Coronavirus Disease 2019 (COVID-19) [1]. On January 30, 2020, the World Health Organization (WHO) declared the COVID-19 outbreak as “public health emergency of international concern” and two months later on March 11th as a pandemic. The SARS-CoV-2 virus was first reported in central city of Wuhan, Hubei province, China, and presented 70% of similarity with the SARS-CoV-1 virus [2] and 96% similarity with a bat coronavirus, which is an evidence of the original host of this zoonosis [1], although the exact source has yet to be elucidated. While the most common symptoms are fever, cough and dyspnoea, the disease can cause other less frequent clinical manifestations such as myalgia, headaches, breathlessness, fatigue and nausea [3].

Viruses and bacteria are often present in the respiratory tract of healthy and asymptomatic individuals [4]. Microaspiration of aerosols and direct mucosal dispersal is responsible for a constant inflow of microbes and viruses towards lower airways [4]. Disease and inflammatory processes that lead to the emergence of anaerobic zones, or mucus accumulation in the alveoli can drastically change the microbial community of the airways [4]. For example, in diseased individuals, the lung microbiome composition undergoes a decrease in diversity [7] accompanied by a shift in the dominant taxa: from Bacteroidetes to Gammaproteobacteria, a class that includes many respiratory pathogens.

Although the clinical characteristics and epidemiology of COVID-19 have been described [8,9,10], studies focused on the associations between the patient's microbiota and the onset

of the disease are still limited. This pilot study aims to characterize the nasopharyngeal mucosal microbial communities of SARS-CoV-2 infected patients. We investigated samples from a control group of SARS-CoV-2 negative patients and three groups of SARS-CoV-2 positive patients, divided according to disease severity: one group of symptomatic patients that did not require hospitalization, a second group of patients that were admitted to conventional hospitalization facilities, and a third group of patients that required admission to the ICU.

METHODS

Patients and experimental design

56 nasopharyngeal microbiome samples from SARS-CoV-2 positive patients and 18 samples from SARS-CoV-2 negative patients were collected during March and April of 2020 in the Emergency Service of Hospital General Universitario de Alicante (HGUA). Cobas SARS-CoV-2 PCR Test for the Cobas 6800 System (Roche Molecular Systems, Branchburg, NJ, USA) was used to detect the presence of SARS-CoV-2 [11].

Patients were first classified based on SARS-CoV-2 presence, and then regarding their later developments (hospital admission and severity). All samples were obtained before the onset of severe symptoms, and before any treatment was administered to the patients. Following these criteria, four groups were established: group 0: SARS-Cov-2 negative patients (n=18); group 1: mild COVID19 symptoms but no later hospital admission (n= 19); group 2: severe COVID19 symptoms followed by hospital admission (n=18); and group 3: patients with severe COVID19 symptoms which were eventually admitted into intensive care units (ICU) (n=19). Protocols were developed in accordance with the national ethical and legal standards, and following the guidelines established in the Declaration of Helsinki (2000). The research project was conducted under the written approval of the Ethic Committee of Clinical Research with

Drug (In Spanish, CEIm) of the "Hospital General Universitario de Alicante (Spain)", and in collaboration with the Biobank of Clinical and Biomedical Research Institute of Alicante (ISABIAL), which are included in the Valencian Network of Biobanks.

DNA isolation and Sequencing

DNA from nasopharyngeal samples was isolated using the QIAamp DNA Mini Kit (QIAgen) following the protocol recommended by the manufacturer. Sequencing libraries were prepared according to the 16S Metagenomic Sequencing Library Preparation protocol distributed by Illumina. Briefly, the sequence spanning the hypervariable regions V3 and V4 of the 16S rRNA gene was amplified through PCR and amplicons were quantified using a Qubit 4 Fluorometer (Qubit dsDNA HS Assay Kit) and validated by 4200 TapeStation (company). Amplicons were sequenced with Illumina MiSeq System using the 2x300bp cartridge. The quality of raw sequences was assessed by FastQC software.

Taxonomic classification of amplicon sequences

Paired end reads of 300 bp were generated with an average overlap of 140 bp. Sequences were trimmed using trimmomatic [12] and the resulting paired reads were merged using casper [13], generating individual fragments of about 460 bp. Given the uneven coverage between samples, the number of individual reads was standardized to 20,000 per sample, removing samples that did not reach this sequencing depth. Merged amplicon sequences were grouped in operational taxonomic units (OTUs) using cd-hit [14] with an identity of 97%. Sequences were queried against small subunits (16S) rRNA genes by the SILVA database [15] for taxonomic classification. Sequences with low identity (< 70%) to any reference 16S rRNA gene or classified as eukaryotic were excluded from further analysis.

Testing for differences in taxonomic composition among patient groups

We sought to determine how different samples were grouped according to their OTU composition. To that end, non-metric multidimensional scaling (NMDS) analysis was performed based on Bray-Curtis dissimilarity measures were calculated among samples based on relative OTU abundances (i.e. percentages) through the Vegan (v 2.5-6) package in R (v 3.6.3). The relative abundances of OTUs were also used to test for statistically significant differences among severity groups. Group OTU compositions were compared through ANOSIM. Next, Similarity Percentage (SIMPER) analysis was used to determine which OTUs were responsible for driving the differences in community composition among groups. For this analysis, all six possible pairwise combinations of severity groups were tested.

OTU association with COVID-19 severity

To infer associations between the severity of COVID-19 and the airways microbiome, general linear models (GLM) were built using the R package MaAsLin2 with centred log-transformed (CLR) OTUs counts as the dependent variable and the severity group (with group 0 and group 1 as references), adjusted by gender and age, as the independent variable. Only OTUs that presented a prevalence of 20% over the sample space were considered. The resulting p-values were adjusted for multiple testing using the Benjamini-Hochberg method (BH).

Co-abundance networks for COVID-19 severity groups

Fastpar [16], a multi-thread implementation of the SparCC algorithm [17], was used to generate co-abundance networks among OTUs of each of the four severity groups with default parameters (50 iterations and correlation threshold of 0.2) and 1,000 bootstrap iterations to infer significance. Results were processed using an in-house ipython notebook to generate network matrices for visualization with Cytoscape 3.8 [18]. The network matrices

were loaded in the Cytoscape 3.8 software, and connections filtered by p-value (≤ 0.05) and correlation (≤ -0.6 or ≥ 0.6).

RESULTS

Study Set

Seventy-four patients were included in this pilot study to assess associations between the nasopharyngeal microbiome composition and the severity of the COVID19 disease. However, only 65 samples remained after quality coverage control (see Material and Methods). Data including age, sex, diagnosis, hospital admission, and disease severity were registered (Table S1). Sixteen patients belonged to the negative control (Group 0, no-SARS-CoV-2), whereas the remaining patients were classified into three groups (Group 1, 2 and 3) according to the severity (see methods). The average age of the patients was *ca*. 60 years old and around 49% of them were diagnosed with pneumonia.

Microbiome taxonomic composition differs among severity groups

The bacterial phylum Firmicutes was the most abundant in the nasopharynx microbiome among patients from all severity groups (52.94% ± 4.04%), followed by Bacteroidota (22.06% ± 6.07%), Proteobacteria (12.75% ± 7.28%) and Actinobacteria (5.4% ± 0.6%). At the genus level, *Streptococcus* was the most abundant taxon (25.23% ± 2.03%), followed by *Prevotella* (16.20% ± 5.66%), *Veillonella* (14.45% ± 2.20%), *Haemophilus* (5.28% ± 4.76%) and *Moraxella* (3.24% ± 3.6%) (Figure S1 and Table S2). A total of 62 OTUs were found exclusively in SARS-CoV-2 positive patients (at a minimum of three samples). Most of these OTUs were classified as members of the phylum Bacteroidetes (18) and Firmicutes (25). Notably, the most common genera among the OTUs found exclusively on COVID-19 positive patients were *Prevotella* (13), followed by *Leptotrichia* (4) and *Streptococcus* (4). Samples were compared based on the

192 relative abundances of OTUs. This analysis revealed that samples did not cluster according to
193 the severity group neither by hierarchical clustering (Figure S2A and 2B) or NMDS (Figure S2C).
194 Nevertheless, the differences in OTU composition among severity groups were significant
195 according to ANOSIM (R = 0.046, p = 0.036).
196 SIMPER analysis revealed that 25 OTUs were responsible for approximately 70% (p-value
197 0.04) of the differences in community composition between severity groups 1 and 3 (Table
198 S3). These OTUs were classified as members of the phyla Bacteroidota, Firmicutes,
199 Fusobacteriota and Proteobacteria. Eleven OTUs had higher average abundance among
200 samples from severity group 1, among which were included three OTUs classified as members
201 of the genus *Veillonella*. On the other hand, 14 OTUs were more abundant among samples
202 from severity group 3, among which were included four OTUs classified as *Prevotella*.
203 **Multiple OTUs display differential abundance according to COVID-19 severity**
204 Using group 0 as a reference, we identified a total of 10 significant associations between
205 bacterial OTUs and patient severity (p-value < 0.05, q-value < 0.25), corrected for age and sex.
206 Among those, 9 were positively associated (8 in group 2 and 1 in group 3 when contrasted
207 with group 0) and 1 negatively associated (in group 3 contrasted with group 0) (Table S4,
208 Figure 1A). Of the OTUs positively associated with severity, 3 were classified as members of
209 the genus *Prevotella* (OTUs 4, 14 and 16). Due to the heterogeneity of group 0, we decided to
210 investigate also the differences within the SARS-CoV-2 positive patients, using group 1 as
211 reference. The GLM model showed just 1 significant OTU (OTU 16), a *Prevotella* also found to
212 be significantly associated with severity in the first model (Table S4, Figure 1B). We did not
213 find any OTUs significantly different between groups 1 and 0. Figure 1A shows the coefficients
214 for all the significant OTUs found by both GLMs and figure 1B shows OTU 16 CLR transformed
215 counts for all severity groups.

Co-abundance networks for COVID-19 severity groups

In order to investigate how OTUs correlate in the different groups, we generated a total of 4 co-abundance networks, one for each severity group. For the severity group 0, the SARS-CoV-2 negative group, the network displayed 118 nodes with 179 edges. Regarding the other three severity groups, ranging from mild to high severity, the complexity of the network decreased with the increase of severity. The network for patients with mild symptoms (group 1) has 137 nodes with 457 edges, while the network for patients with severe symptoms but not admitted in ICU (group 2) had 129 nodes with 171 edges and the network for severe patients admitted in ICU (group 3) had 100 nodes and 148 edges. In the network of severity group 1, OTU 16 (*Prevotella*, associated with severity in two GLMs) displayed 18 co-abundant OTUs connected in the network in first degree (Figure 2). Among these connections, ten were negative associations while eight were positive. Most of these connections with OTU 16 were absent from networks of severity groups 2 and 3. Only 3 and 2 first degree connections remained in each of these networks respectively (Figure S3).

DISCUSSION

In this preliminary study, analysis of the taxonomic composition of the samples showed differences between patients that developed different onsets of COVID-19. These changes in nasopharyngeal community composition are subtle, meaning that they are restricted to few taxa out of the complete meta-community. Nevertheless, there are detectable and significant changes among OTU abundances. These changes could be linked to the different severity groups, as we identified both taxa that were present exclusively among COVID-19 positive patients as well as those whose abundance was significantly higher or lower among different severity groups. Not only this, but also the complexity of co-abundance networks (which can

be taken as a proxy for potential interactions between taxa), was decreased among patients that developed more severe cases of COVID-19. Below we discuss the mechanisms by which specific microbes might play a role in either enhancing or decreasing the severity of COVID-19. Those results suggest potential biomarkers for the onset of the disease.

Potential associations between bacterial taxa and COVID-19 severity.

Among the OTUs positively associated with COVID-19 severity, three were classified as members of the genus *Prevotella*, and one to a closely related genus, *Alloprevotella*. A recent study showed that *Prevotella* proteins can promote viral infection through multiple interactions with NF-κB signalling pathway, which is also involved in COVID-19 severity [19]. The genus *Prevotella* is usually considered commensal and, as such, rarely involved in infections. However, some strains have been identified as opportunistic pathogens in chronic infections, abscesses and anaerobic pneumonia [20,21,22,23]. The role of some strains of *Prevotella* in chronic mucosal inflammation has been demonstrated. They are involved with augmented T helper type 17 (Th17)-mediated mucosal inflammation, through activation of Toll-like receptor 2, followed by production of cytokines by antigen-presenting cells, including interleukin-23 (IL-23) and IL-1 [23]. The severe symptoms of COVID-19 are associated with cytokine storms, many of which are involved in TH17 type responses [24]. The significant association of *Prevotella* sp. and disease severity observed here suggests a possible link between *Prevotella* sp. and the COVID-19 through the activation of immunity signaling pathways that modulate inflammation, and this link should be further explored.

Reduced network complexity among patients who later developed more severe COVID-19.

Several studies demonstrated the usefulness of co-abundance networks to elucidate changes in the microbiome associated with human diseases [26,27,28,29]. By switching from

264 individual OTU associations to a community interaction approach it is possible to attain a
265 better understanding of the dynamic of microbiome/phenotype associations, revealing
266 microbial consortia (and not only an OTU) that might be collectively influencing the host
267 phenotype. Our linear models showed OTU 16 (*Prevotella* sp.) as an important OTU associated
268 with severity. This OTU had the highest number of connections in the network, followed by
269 OTU 9 (*Veillonella* sp.). Of the four networks generated, the severity group 1 network showed
270 the higher number of interactions with this OTU. Ecological networking, *in vitro* and clinical
271 studies showed that *Prevotella* sp. and *Veillonella* sp. are keystone species in microbiomes
272 during airway disease progression, especially in diseases associated with mucus accumulation
273 such as cystic fibrosis [30-32]. These anaerobes are efficient at degrading mucin molecules on
274 the airway mucosa, releasing byproducts that enable the colonization and growth of
275 pathogenic bacteria that are poor at degrading mucus for growth [33]. In COVID patients,
276 *Prevotella* sp. and *Veillonella* sp. could have a similar role due to the decreased mucociliary
277 clearance caused by the viral infection [34]. Lower rates of clearing increase the residence
278 time of *Prevotella* sp. and *Veillonella* sp. in the airways, likely increasing their mucus
279 metabolism and enabling further colonization by pathogenic bacteria that may cause
280 pneumonia.
281 OTU 96, classified as *Dolosigranulum sp.*, was identified in the group 1 network by having a
282 negative relationship with OTU 16 (*Prevotella* sp.) as first-degree neighbor (Figure 3). OTU 96
283 did not pass the q-value threshold established for the GLMs but shows significant p-value
284 (0.003 in the model comparing group 2 and group 0, and 0.02 in the model comparing group
285 2 and group 1 as reference). The only species currently described in this genus is
286 *Dolosigranulum pigrum*, which is commonly found in the nasopharynx microbiome and is
287 predicted to benefit the host through protection against pneumococcal colonization [35-36]

and through protection against inflammation damage [37]. One study also found a lower abundance of *Dolosigranulum* in children with Influenza A Virus compared to healthy children [38]. In addition, a study reported that patients with their airway microbiota dominated by *Corynebacterium* and *Dolosigranulum* experienced the lowest rates of early loss of asthma control and have a longer time to develop at least 2 episodes [39]. We did not identify Corynebacterium directly connected to OTU16 (*Prevotella* sp.), but OTU 78, classified as Corynebacterium is positively associated with OTU 96 (0.7479, p-value 0.001) in the co-abundance network from group 1 (Figure 3), indicating that in asymptomatic patients those two taxa are forming a consortium that might protect from disease development. This "consortium" was also implicated in resistance to recurrent ear infections and it was proposed as a probiotic candidate for upper respiratory tract infections [40]. The reason that we did not have lower q-value in our GLM for those two taxa could be the lack of power due to the small size of our study. Thus, these associations warrant further investigation.

LIMITATIONS

The major limitation of our study is the small sample size. With only about 15 samples per severity group it is difficult to find statistically significant associations between microbiome composition and disease severity. Nevertheless, this limitation is more likely to lead to false negatives than to false positives. We also cannot rule out confounding factors that might explain the differences between groups. Another important limitation is the fact that we performed amplicon rather than whole genome shotgun sequencing. This leads to three issues. First, some of the bacterial diversity is lost due to the fact that the selected primers do not amplify the entirety of bacterial diversity. Second, some genomes have more than a single copy of the 16S operon, which can lead to an overestimation of their abundance in the

312 samples. Third, without metagenomes (and metagenome assembled genomes) we could not
313 make inferences about the presence of virulence factors and other features of the genomes
314 of the microbes in our samples. We resorted to 16S amplification because our non-invasive
315 approach to collect samples yields low DNA amounts that are inadequate for sequencing.
316 However, as far as we know, this is a unique pilot study in the field. The aim is to be able to
317 transfer the first useful results to help clinical practice in the fight against the virus and to
318 optimize all the protocols and analyses for a second analysis in which the sample size will be
319 much larger. We are currently working on collecting more samples and optimizing protocols
320 that will allow us to obtain whole genome shotgun sequencing from them.
321

322 **CONCLUSION**

323 Our data provides preliminary evidence of significant differences in the composition of the
324 upper airway microbiome according to COVID-19 severity, suggesting potential biomarkers of
325 disease severity. While the richness indexes did not show significant differences among
326 groups, specific taxa were significantly associated with disease development. We also
327 demonstrated that the complexity of the co-abundance network is decreased in patients who
328 came to develop severe cases of the disease, indicating that the interactions between the
329 taxa are also relevant to this process. Further studies will be necessary to shed light on the
330 molecular mechanisms that give rise to these associations. Finally, we make no claim that the
331 differences in microbiome composition reported here are the cause of of COVID-19 severity.
332 Nevertheless, the significant associations found between these variables suggests that the
333 role of the microbiome on the onset of disease severity warrants further investigation.
334

335 **DECLARATIONS**

Ethics approval and consent to participate

This study has the written approval of the Ethic Committee of Clinical Research of Alicante University General Hospital (Ref. CEIm: PI2020-052). The samples used in this work are from clinical nasopharyngeal aspirates used to diagnose the COVID19 pathology during the first emergency state in Spain and stored in the Alicante University General Hospital Biobank. In this period, it was allowed to collected samples by Biobanks without obtaining the Informed Consent (Dictamen COVID19 D.20200327/2 CEI DGSP-CSISP).

Consent for publication

Not applicable

Availability of data and materials

Raw data was deposited to the National Center for Biotechnology Information Sequence Read Archive under BioProject accession number PRJNA673585.

All data generated during this study are included in this published article [and its supplementary information files].

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported by a grant from Instituto de Salud Carlos III (ISCIII; grant number COV20/00236).

Authors' contributions

JCR conceived the study. MPV, IV, CM and EM collected the data. RC, BA, CS, JHN and FH analysed the data. MPV, RC, MLP, FH, BA and JCR wrote the paper. All authors reviewed and approved the final version of the manuscript.

Acknowledgments

Not applicable

REFERENCES

1. Andersen KG, Rambaut A, Lipkin WI, Holmes EC, Garry RF. The proximal origin of SARS-CoV-2. Nat Med. 2020;26:450–2.

2. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, et al. A Novel Coronavirus from Patients with Pneumonia in China, 2019. N Engl J Med. 2020;382:727–33.

3. Guan W, Ni Z, Hu Y, Liang W, Ou C, He J, et al. Clinical Characteristics of Coronavirus Disease 2019 in China. N Engl J Med. 2020;382:1708–20.

4. Huffnagle GB, Dickson RP, Lukacs NW. The respiratory tract microbiome and lung inflammation: a two-way street. Mucosal Immunol. 2017;10:299–306.

5. Kostric M, Milger K, Krauss-Etschmann S, Engel M, Vestergaard G, Schloter M, et al. Development of a Stable Lung Microbiome in Healthy Neonatal Mice. Microb Ecol. 2018;75:529–42.

6. Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Falkowski NR, Huffnagle GB, et al. Bacterial Topography of the Healthy Human Lower Respiratory Tract. Clemente JC, editor. mBio. 2017;8:e02287-16, /mbio/8/1/e02287-16.atom.

7. Kalantar KL, Moazed F, Christenson SC, Wilson J, Deiss T, Belzer A, et al. Metagenomic comparison of tracheal aspirate and mini-bronchial alveolar lavage for assessment of respiratory microbiota. American Journal of Physiology-Lung Cellular and Molecular Physiology. 2019;316:L578–84.

8. Rothan HA, Byrareddy SN. The epidemiology and pathogenesis of coronavirus disease (COVID-19) outbreak. Journal of Autoimmunity. 2020;109:102433.

9. Wu F, Zhao S, Yu B, Chen Y-M, Wang W, Song Z-G, et al. A new coronavirus associated with human respiratory disease in China. Nature. 2020;579:265–9.

10. Chen N, Zhou M, Dong X, Qu J, Gong F, Han Y, et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. The Lancet. 2020;395:507–13.

11. Poljak M, Korva M, Knap Gašper N, Fujs Komloš K, Sagadin M, Uršič T, et al. Clinical Evaluation of the cobas SARS-CoV-2 Test and a Diagnostic Platform Switch during 48 Hours in the Midst of the COVID-19 Pandemic. McAdam AJ, editor. J Clin Microbiol. 2020;58:e00599-20, /jcm/58/6/JCM.00599-20.atom.

12. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114–20.

13. Kwon S, Lee B, Yoon S. CASPER: context-aware scheme for paired-end reads from high-throughput amplicon sequencing. BMC Bioinformatics. 2014;15:S10.

14. Fu L, Niu B, Zhu Z, Wu S, Li W. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics. 2012;28:3150–2.

15. Quast C, Pruesse E, Yilmaz P, Gerken J, Schweer T, Yarza P, Peplies J, Glöckner FO. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. Nucl. Acids Res. 41 (D1): D590-D596. 2013

16. Watts SC, Ritchie SC, Inouye M, Holt KE. FastSpar: rapid and scalable correlation estimation for compositional data. Stegle O, editor. Bioinformatics. 2019;35:1064–6.

17. Friedman J, Alm EJ. Inferring Correlation Networks from Genomic Survey Data. von Mering C, editor. PLoS Comput Biol. 2012;8:e1002687.

18. Shannon P. Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Research. 2003;13:2498–504.

19. Khan AA, Khan Z. COVID-2019-associated overexpressed Prevotella proteins mediated host–pathogen interactions and their role in coronavirus outbreak. Elofsson A, editor. Bioinformatics. 2020;36:4065–9.

20. Nagy E. Anaerobic Infections: Update on Treatment Considerations. Drugs. 2010;70:841–58.

21. Brook I. Anaerobic Pulmonary Infections in Children: Pediatric Emergency Care. 2004;20:636–40.

22. Brook I. MICROBIOLOGY OF COMMON INFECTIONS IN THE UPPER RESPIRATORY TRACT. Primary Care: Clinics in Office Practice. 1998;25:633–48.

23. Larsen JM. The immune response to *Prevotella* bacteria in chronic inflammatory disease. Immunology. 2017;151:363–74.

24. Wu D, Yang XO. TH17 responses in cytokine storm of COVID-19: An emerging target of JAK2 inhibitor Fedratinib. Journal of Microbiology, Immunology and Infection. 2020;53:368–70.

25. de Carvalho Baptista IM, Martinho FC, Nascimento GG, da Rocha Santos CE, Prado RF do, Valera MC. Colonization of oropharynx and lower respiratory tract in critical patients: Risk of ventilator-associated pneumonia. Archives of Oral Biology. 2018;85:64–9.

26. Chen L, Collij V, Jaeger M, van den Munckhof ICL, Vich Vila A, Kurilshikov A, et al. Gut microbial co-abundance networks show specificity in inflammatory bowel disease and obesity. Nat Commun. 2020;11:4018.

27. Wang H, Li Y, Feng X, Li Y, Wang W, Qiu C, et al. Dysfunctional gut microbiota and relative co-abundance network in infantile eczema. Gut Pathog. 2016;8:36.

28. Faust K, Sathirapongsasuti JF, Izard J, Segata N, Gevers D, Raes J, et al. Microbial Co-occurrence Relationships in the Human Microbiome. Ouzounis CA, editor. PLoS Comput Biol. 2012;8:e1002606.

29. Greenblum S, Turnbaugh PJ, Borenstein E. Metagenomic systems biology of the human gut microbiome reveals topological shifts associated with obesity and inflammatory bowel disease. Proceedings of the National Academy of Sciences. 2012;109:594–9.

30. Silveira, Cynthia, Ana Georgina Cobian-Guemes, Carla Uranga, Jonathon Baker, Anna Edlund, Forest Rohwer, and Douglas Conrad. "Multi-omics study of keystone species in the cystic fibrosis lung microbiome." (2020).

31. Quinn, Robert A., Katrine Whiteson, Yan Wei Lim, Jiangchao Zhao, Douglas Conrad, John J. LiPuma, Forest Rohwer, and Stefanie Widder. "Ecological networking of cystic fibrosis lung infections." *NPJ biofilms and microbiomes* 2, no. 1 (2016): 1-11.

32. Flynn, Jeffrey M., David Niccum, Jordan M. Dunitz, and Ryan C. Hunter. "Evidence and role for bacterial mucin degradation in cystic fibrosis airway disease." *PLoS pathogens* 12, no. 8 (2016): e1005846.

33. Flynn, Jeffrey M., Lydia C. Cameron, Talia D. Wiggen, Jordan M. Dunitz, William R. Harcombe, and Ryan C. Hunter. "Disruption of Cross-Feeding Inhibits Pathogen Growth in the Sputa of Patients with Cystic Fibrosis." *mSphere* 5, no. 2 (2020).

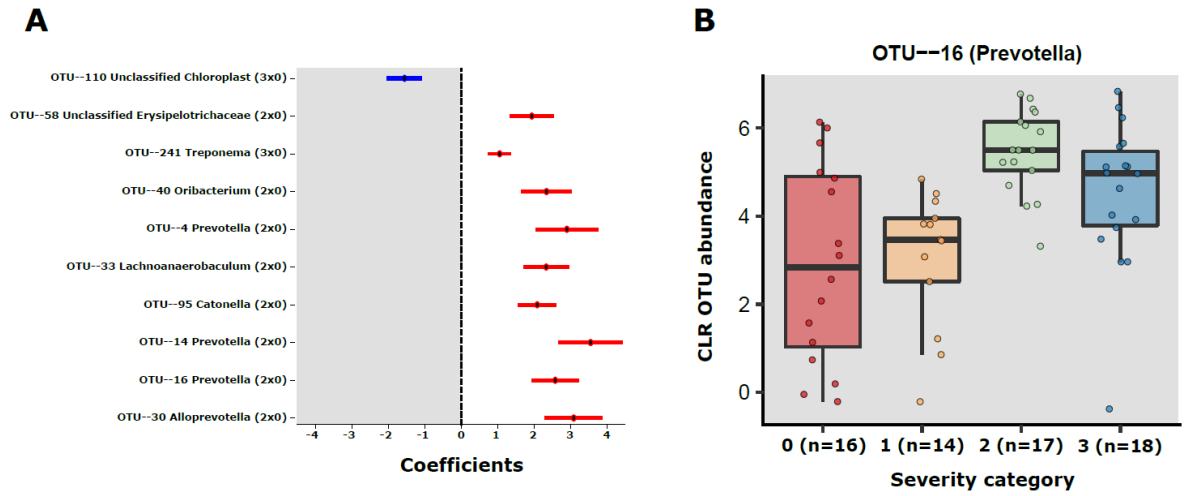
34. Robinot, Rémy, Mathieu Hubert, Guilherme Dias de Melo, Françoise Lazarini, Timothée Bruel, Nikaïa Smith, Sylvain Levallois et al. "SARS-CoV-2 infection damages airway motile cilia and impairs mucociliary clearance." *bioRxiv* (2020).

35. Biesbroek G, Bosch AATM, Wang X, Keijser BJF, Veenhoven RH, Sanders EAM, et al. The Impact of Breastfeeding on Nasopharyngeal Microbial Communities in Infants. Am J Respir Crit Care Med. 2014;140612135546007.

454 36. Bomar L, Brugger SD, Yost BH, Davies SS, Lemon KP. Corynebacterium accolens Releases
455 Antipneumococcal Free Fatty Acids from Human Nostril and Skin Surface Triacylglycerols.
456 mBio. 2016;7:e01725-15, /mbio/7/1/e01725-15.atom.
457 37. Ortiz Moyano R, Raya Tonetti F, Tomokiyo M, Kanmani P, Vizoso-Pinto MG, Kim H, et al.
458 The Ability of Respiratory Commensal Bacteria to Beneficially Modulate the Lung Innate
459 Immune Response Is a Strain Dependent Characteristic. Microorganisms. 2020;8:727.
460 38. Wen Z, Xie G, Zhou Q, Qiu C, Li J, Hu Q, et al. Distinct Nasopharyngeal and Oropharyngeal
461 Microbiota of Children with Influenza A Virus Compared with Healthy Children. BioMed
462 Research International. 2018;2018:1–9.
463 39. Zhou Y, Jackson D, Bacharier LB, Mauger D, Boushey H, Castro M, et al. The upper-airway
464 microbiota and loss of asthma control among asthmatic children. Nat Commun.
465 2019;10:5714.
466 40. Lappan R, S Peacock C. Corynebacterium and Dolosigranulum: future probiotic candidates
467 for upper respiratory tract infections. Microbiol Aust [Internet]. 2019 [cited 2020 Sep 3];
468 Available from: http://microbiology.publish.csiro.au/?paper=MA19051

469

470 **FIGURES**



471

472 **Figure 1**. **A**) Error bar plot of the GLM coefficients interval for each OTU. The Y-axis shows the

473 OTU number and Genus classification. The X-axis represents the CLR abundance. In red the

474 positively associated OTUs and in blue the negatively associated OTUS. **B**) OTU 16 (*Prevotella*)

475 center log transformed (CLR) abundance in the severity groups 0-3.

476

477

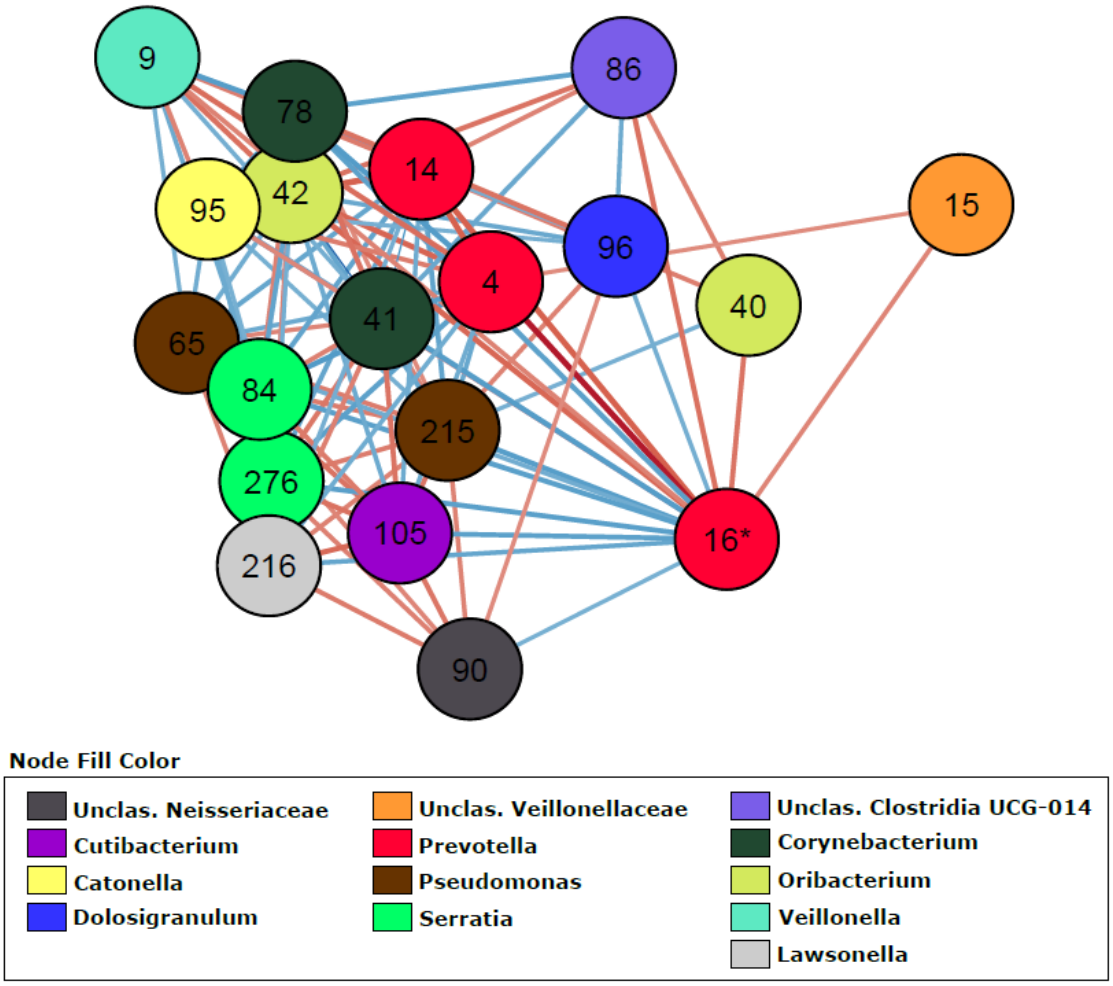


Figure 2. Co-abundance network (severity group 1) showing only first-degree neighbours of OTU 16 (*Prevotella* sp.). OTUs are represented by nodes and significant correlations by edges. Blue edges represent negative associations and red, positive associations. The colour of nodes was defined by the taxonomic classification of the OTU at Genus rank.

SUPPLEMENTARY MATERIAL

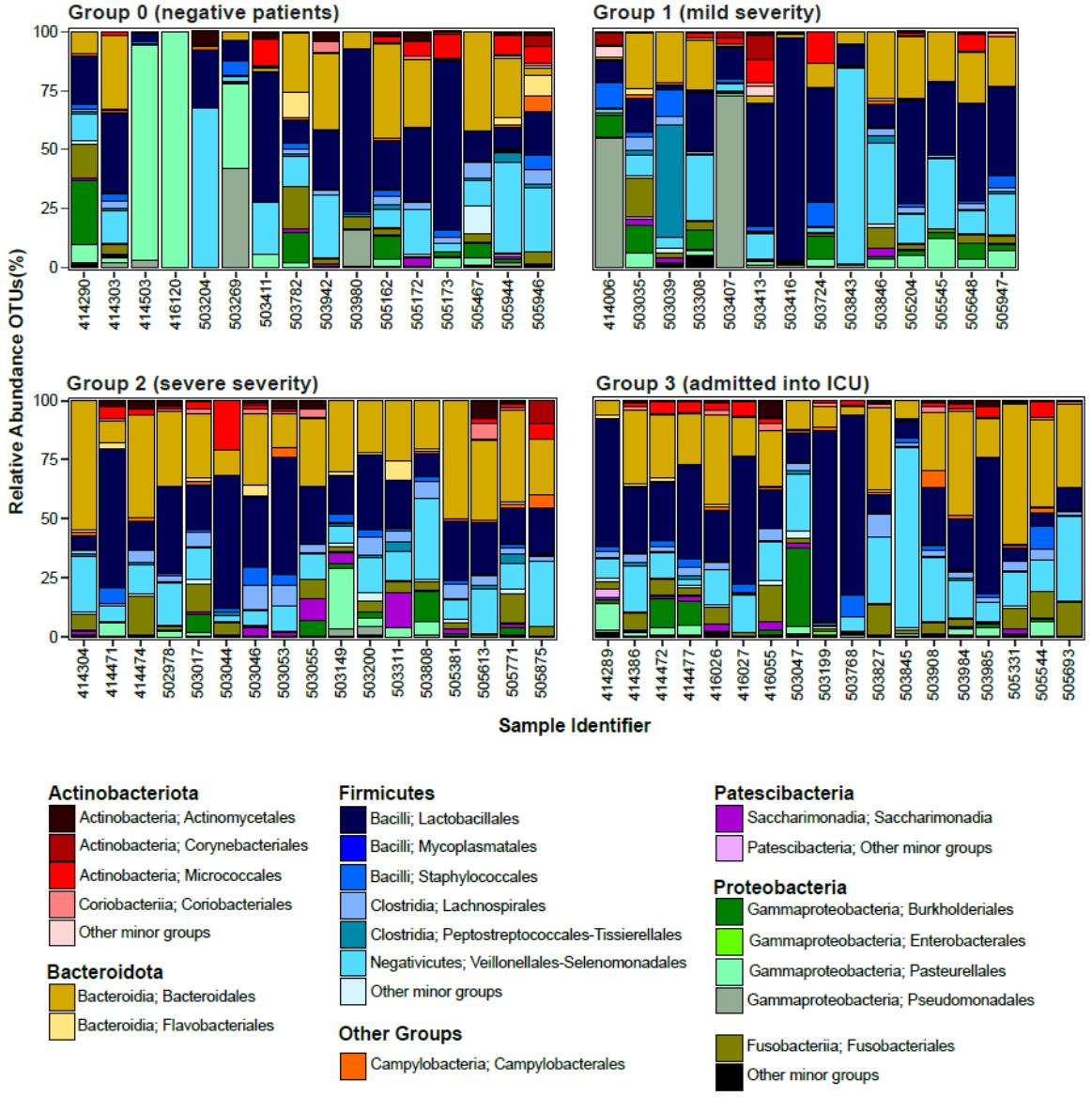
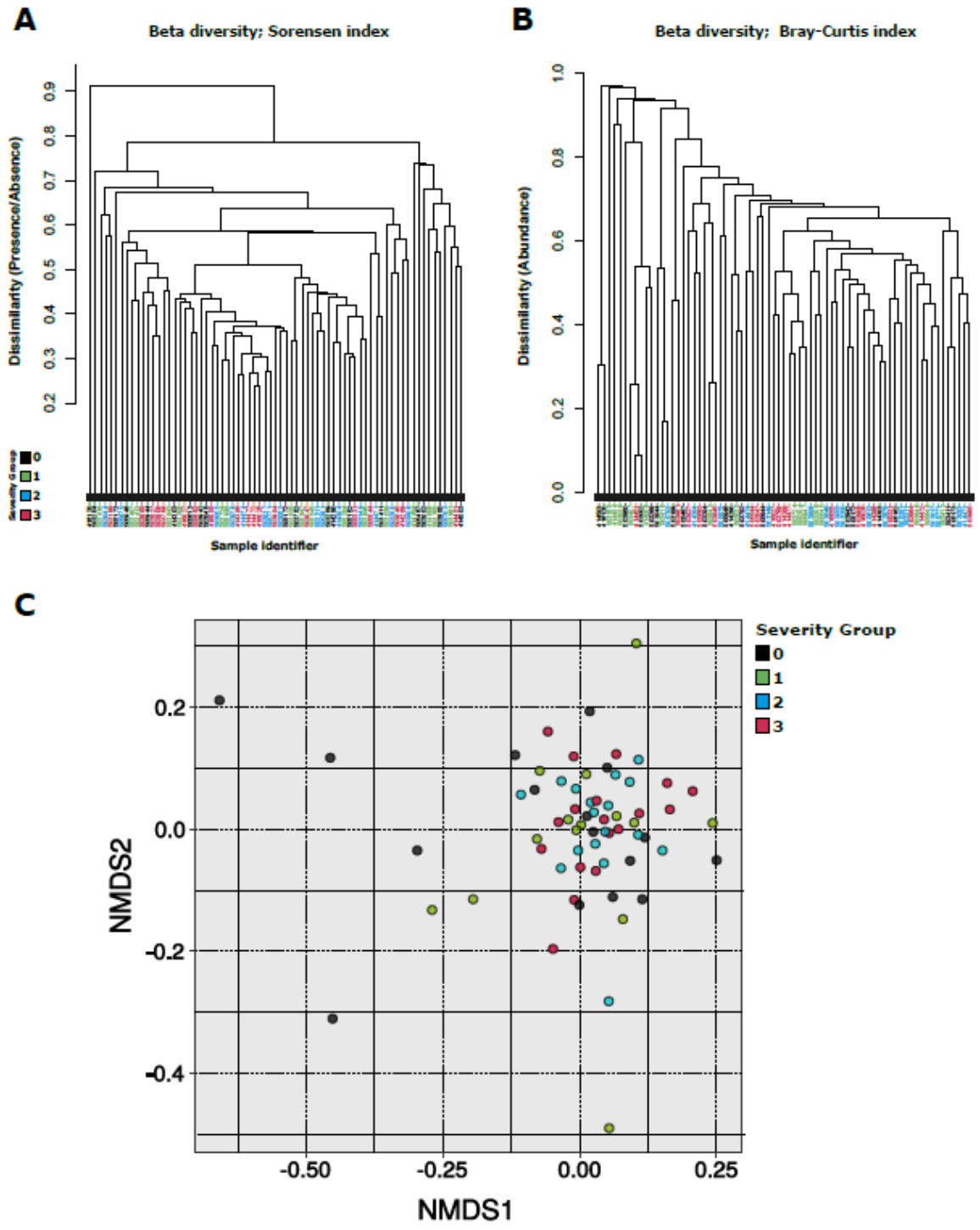


Figure S1. Relative abundance of bacterial populations, at genus level, in the microbiome of patients within COVID-19 severity groups. Only microorganisms with a relative abundance greater than 0.5% are shown in the legend.



494

495 **Figure S2. Beta diversity.** Dendrogram based on **A**) Bray-Curtis dissimilarity and **B**) Sørensen

496 dissimilarity values. **C) Comparison of sample taxonomic profiles by severity group.**

497 Nonmetric multidimensional scaling was applied to determine the clustering patterns of

498 samples according to their OTU abundance patterns. Each dot represents a sample color

499 coded according to the severity group it belongs to. The closer the samples are, the more

similar was their OTU abundance composition. No clear clustering of samples by severity group was observed.

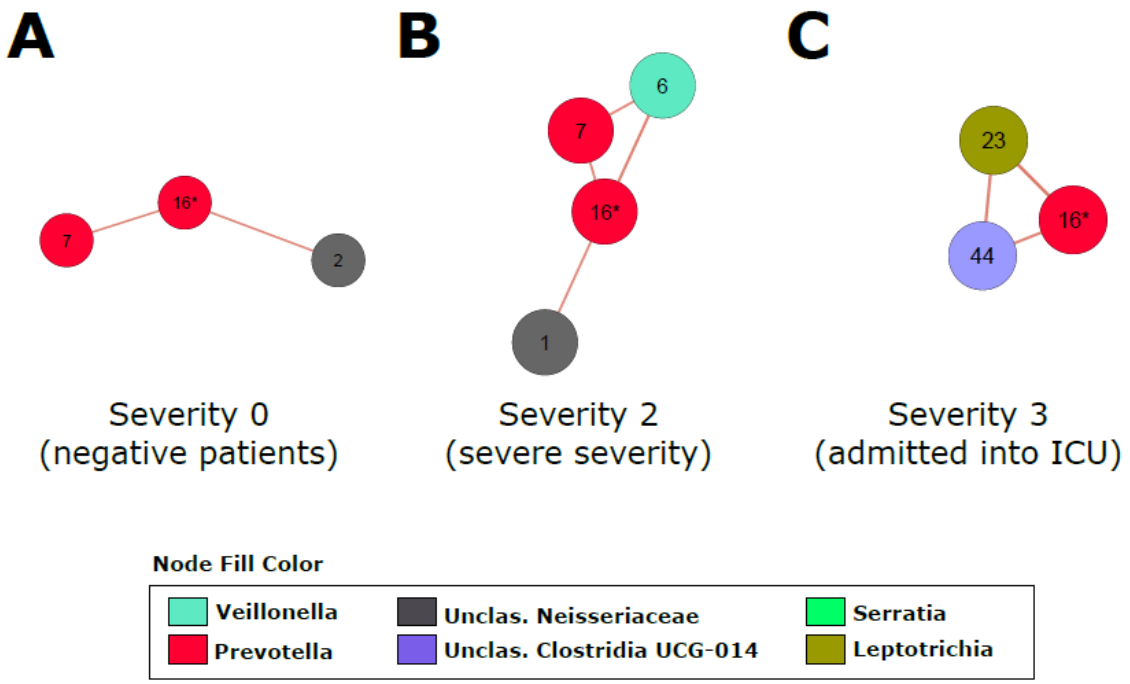


Figure S3. Co-abundance network showing only first-degree neighbours of OTU 16 (*Prevotella* sp.). **A)** Severity group 0 **B)** Severity group 2 and **C)** Severity group 3.

Table S1. Clinical features of patients.

Table S2. OTUs taxonomic classification.

Table S3. OTUs showing approximately 70% of the differences in community composition between severity groups 1 and 3 (SIMPER).

Table S4. Maaslin2 results (GLM) for OTUs associations (q-value < 0.25) for both models.