

1 **Characterization of Microbial Co-infections in the Respiratory Tract of** 2 **hospitalized COVID-19 patients**

3 Huanzi Zhong^{1,3*}, Yanqun Wang^{2*}, Zhun Shi^{1*}, Lu Zhang^{4*}, Huahui Ren^{1,3*}, Weiqun He^{2*}, Zhaoyong Zhang^{2*}, Airu Zhu^{2*},
4 Jingxian Zhao^{2*}, Fei Xiao^{5*}, Fangming Yang^{1,6}, Tianzhu Liang^{1,7}, Feng Ye², Bei Zhong⁸, Shicong Ruan⁹, Mian Gan², Jiahui
5 Zhu^{1,10}, Fang Li², Fuqiang Li^{1,11}, Daxi Wang^{1,7}, Jiandong Li^{1,7,12}, Peidi Ren^{1,7}, Shida Zhu^{1,13}, Huanming Yang^{1,14,15}, Jian Wang^{1,14},
6 Karsten Kristiansen^{1,3}, Hein Min Tun¹⁶, Weijun Chen^{17,12}, Nanshan Zhong², Xun Xu^{1,18#}, Yi-min Li^{2#}, Junhua Li^{1,7,19#}, Jincun
7 Zhao^{2,4#}

8 ¹BGI-Shenzhen, Shenzhen, 518083, China.

9 ²State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of
10 Respiratory Health, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong, 510120, China.

11 ³Laboratory of Genomics and Molecular Biomedicine, Department of Biology, University of Copenhagen, 2100 Copenhagen,
12 Denmark.

13 ⁴Institute of Infectious disease, Guangzhou Eighth People's Hospital of Guangzhou Medical University, Guangzhou, Guangdong,
14 510060, China.

15 ⁵Department of Infectious Diseases, Guangdong Provincial Key Laboratory of Biomedical Imaging, Guangdong Provincial
16 Engineering Research Center of Molecular Imaging, The Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, Guangdong,
17 519000, China.

18 ⁶School of Future Technology, University of Chinese Academy of Sciences, Beijing 101408, China.

19 ⁷Shenzhen Key Laboratory of Unknown Pathogen Identification, BGI-Shenzhen, Shenzhen, 518083, China.

20 ⁸The Sixth Affiliated Hospital of Guangzhou Medical University, Qingyuan People's Hospital, Qingyuan, Guangdong, China.

21 ⁹Yangjiang People's Hospital, Yangjiang, Guangdong, China.

22 ¹⁰State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing
23 210096, China.

24 ¹¹Guangdong Provincial Key Laboratory of Human Disease Genomics, Shenzhen Key Laboratory of Genomics, BGI-Shenzhen,
25 Shenzhen 518083, China.

26 ¹²BGI Education Center, University of Chinese Academy of Sciences, Shenzhen, 518083, China.

27 ¹³Shenzhen Engineering Laboratory for Innovative Molecular Diagnostics, BGI-Shenzhen, Shenzhen, 518120, China.

28 ¹⁴James D. Watson Institute of Genome Science, Hangzhou, 310008, China.

29 ¹⁵Guangdong Provincial Academician Workstation of BGI Synthetic Genomics, BGI-Shenzhen, Shenzhen, 518120, China.

30 ¹⁶HKU-Pasteur Research Pole, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong
31 Kong SAR, China.

32 ¹⁷BGI PathoGenesis Pharmaceutical Technology Co., Ltd, BGI-Shenzhen, Shenzhen 518083, China.

33 ¹⁸Guangdong Provincial Key Laboratory of Genome Read and Write, BGI-Shenzhen, Shenzhen, 518120, China.

34 ¹⁹School of Biology and Biological Engineering, South China University of Technology, Guangzhou, China.

35

36 Correspondence should be addressed to **Jincun Zhao** (zhaojincun@gird.cn), **Junhua Li** (lijunhua@genomics.cn), **Yi-min Li**
37 (dryiminli@vip.163.com), and **Xun Xu** (xuxun@genomics.cn).

38 *These authors contributed equally to this work.

39 #These authors jointly supervised this work.

40

11 Summary

12 **Background** Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused a global pandemic of Coronavirus
13 disease 2019 (COVID-19). However, microbial composition of the respiratory tract and other infected tissues, as well as their
14 possible pathogenic contributions to varying degrees of disease severity in COVID-19 patients remain unclear.

15 **Method** Between January 27 and February 26, 2020, serial clinical specimens (sputum, nasal and throat swab, anal swab and
16 feces) were collected from a cohort of hospitalized COVID-19 patients, including 8 mildly and 15 severely ill patients (requiring
17 ICU admission and mechanical ventilation), in the Guangdong province, China. Total RNA was extracted and ultra-deep
18 metatranscriptomic sequencing was performed in combination with laboratory diagnostic assays. Co-infection rates, the
19 prevalence and abundance of microbial communities in these COVID-19 patients were determined.

20 **Findings** Notably, respiratory microbial co-infections were exclusively found in 84.6% of severely ill patients (11/13), among
21 which viral and bacterial co-infections were detected by sequencing in 30.8% (4/13) and 69.2% (9/13) of the patients,
22 respectively. In addition, for 23.1% (3/13) of the patients, bacterial co-infections with *Burkholderia cepacia* complex (BCC) and
23 *Staphylococcus epidermidis* were also confirmed by bacterial culture. Further, a time-dependent, secondary infection of *B.*
24 *cenocepacia* with expressions of multiple virulence genes in one severely ill patient was demonstrated, which might be the
25 primary cause of his disease deterioration and death one month after ICU admission.

26 **Interpretation** Our findings identified distinct patterns of co-infections with SARS-CoV-2 and various respiratory pathogenic
27 microbes in hospitalized COVID-19 patients in relation to disease severity. Detection and tracking of BCC-associated nosocomial
28 infections are recommended to improve the pre-emptive treatment regimen and reduce fatal outcomes of hospitalized patients
29 infected with SARS-CoV-2.

30 **Funding** National Science and Technology Major Project of China, National Major Project for Control and Prevention of
31 Infectious Disease in China, the emergency grants for prevention and control of SARS-CoV-2 of Ministry of Science and
32 Technology and Guangdong province, Guangdong Provincial Key Laboratory of Genome Read and Write, Guangdong Provincial
33 Academician Workstation of BGI Synthetic Genomics, and Shenzhen Engineering Laboratory for Innovative Molecular
34 Diagnostics.

35

36 Introduction

37 As of July 1st, 2020, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has infected over 10 million and resulted in
38 more than 500 thousand deaths worldwide¹. The pandemic poses a significant threat to public health and global economy.

39 Respiratory viruses, such as coronaviruses and influenza, can lead to acute damage of the epithelial barrier and facilitate invasions
40 of other pathogens^{2,3}. For instance, secondary infections by *Stenotrophomonas maltophilia*, *Klebsiella pneumoniae* or *Escherichia*
41 *coli* were reported to cause serious complications in patients with SARS, such as bacteremia, sepsis and nosocomial pneumonia
42 (NP)⁴. In addition, *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Staphylococcus aureus* were frequently associated
43 with NP and mortality in influenza pandemics⁵. It was estimated that approximately 29%-55% of the total 300,000 deaths in the
44 2009 H1N1 pandemic were caused by secondary bacterial NP⁶⁻⁸.

45 Concerns about co-infections of SARS-CoV-2 with known viruses, bacteria and fungi have also been raised. In severely ill
46 patients, acute respiratory distress syndrome (ARDS) tends to rapidly deteriorate the conditions of the patient, and
47 mechanical ventilation is generally required^{9,10}. Such invasive procedures can further increase the risks of ventilator-associated
48 pneumonia (VAP) in these patients¹¹. In 99 confirmed Wuhan patients enrolled in January 2020, one (1%) had positive cultures of
49 *Acinetobacter baumannii*, *K. pneumoniae* and *Aspergillus flavus*, and four (4%) were diagnosed with infection by *Candida*, but no
50 influenza viruses were detected⁹. A later retrospective study in Wuhan patients further demonstrated that half of the deceased
51 patients (27 out of 54) had experienced secondary infections¹². By using real-time reverse transcriptase-polymerase chain reaction
52 tests, Kim et al. recently reported a 20.7% (24 out of 116 specimens) co-infection rate with SARS-CoV-2 and other respiratory
53 viruses in Northern California, including rhinovirus (6.9%) and orthopneumovirus (5.2%)¹³. However, microbial co-infections as
54 well as their possible effects on clinical outcomes of SARS-CoV-2 infected patients remain largely unknown.

55 Here, by simultaneously applied diagnostic technologies (cultures and colorimetric assays) and metatranscriptomic sequencing,
56 microbial co-infections in a Guangdong cohort of 23 patients hospitalized with SARS-CoV-2 infection were comprehensively
57 evaluated. High co-infection rate (84.6%) of at least one type of respiratory pathogens among severely ill patients was
58 documented, whereas none of the patients with mild symptoms presented evidence of co-infection. Our results further
59 demonstrated that BCC bacteria, *Staphylococcus epidermidis* and *Mycoplasma spp.* were the most prevalent pathogenic bacteria
60 associated with culture-confirmed and/or metatranscriptomic-detected co-infections, which might increase risk for prolonged ICU
61 stay or even mortality. Our findings demonstrate the value of metatranscriptomics for unbiased co-detection of respiratory
62 pathogens together with SARS-CoV-2 and provide in-time information and useful practical suggestions regarding the adequate
63 monitoring and management of bacterial co-infections in severely ill COVID-19 patients.

34 **Methods**

35 **Enrollment of hospitalized patients with SARS-CoV-2 infection, collection of clinical specimens**

36 Twenty-three patients admitted to hospital in the period January 10-March 31, 2020 with confirmed SARS-CoV-2 infection based
37 on a positive SARS-CoV-2 test were included. A total of 67 clinical specimens from the respiratory and gastrointestinal tract
38 (including throat swab, nasal swab, sputum, anal swab and feces) were collected from the above patients at the First Affiliated
39 Hospital of Guangzhou Medical University (thirteen patients), the Fifth Affiliated Hospital of Sun Yat-sen University (two
40 patients), Yangjiang People's Hospital (five patients) and Qingyuan People's Hospital (three patients) during the period (between
41 January 27 and February 26, 2020). Patients were classified into mild (n=8, without intensive care unit admission) and severe
42 (n=15, with ICU admission) cases based on their severity of SARS-CoV-2 symptoms. Detailed de-identified information for
43 patients and clinical specimens are presented in **appendix 2 p 2: TableS1** and **appendix 2 p 3: TableS2**, respectively.

44 **Ethics statement.** The study was reviewed and approved by the ethics committees of all the four hospitals and the institutional
45 review board of BGI-Shenzhen.

36 **Laboratory diagnosis of nosocomial bacterial and fungal infections**

37 All severely ill COVID-19 patients admitted to the ICU for more than 48 hours were monitored for nosocomial infections, which
38 were defined according to the definitions of the US Centers for Disease Control and Prevention (CDC)¹⁴. Culturing of sputum
39 and nasal secretions was conducted according to standard protocols for the diagnosis of nosocomial microbial infections
40 (hospital-acquired and ventilator-associated) in all severely ill COVID-19 patients with ICU admission¹⁵. Blood samples of
41 patients (P01 and P05) hospitalized in the Fifth Affiliated Hospital of Sun Yat-sen University were collected for
42 colorimetric assay-based (1-3)- β -D-glucan test to diagnose fungal infection (Dynamiker Biotechnology, Tianjin, China, Catalogue
43 number: DNK-1401-1).

14 **RNA extraction, metatranscriptomic library preparation and sequencing**

15 For each sample, total RNA was extracted (QiAamp RNeasy Mini Kit, Qiagen, Germany) and the concentration was quantified
16 (Qubit RNA HS Assay Kit, Thermo Fisher Scientific, Waltham, MA, USA). Purified RNA samples were then subjected to
17 pre-processing, DNA nanoball-based libraries construction and high-throughput metatranscriptomic sequencing on the
18 DNBSEQ-T7 platform (100nt paired-end reads, MGI, Shenzhen, China) as described previously¹⁶.

19 **Identification and removal of human RNA reads from metatranscriptomic data**

20 For each sample, the raw metatranscriptomic reads were processed using Fastp (v0.19.5, default settings)¹⁷ to filter low-quality
21 data and adapter contaminations and generate the clean reads for further analyses. Human-derived reads were identified with
22 the following steps: 1) identification of human ribosomal RNA (rRNA) by aligning clean reads to human rRNA sequences
23 (28S, 18S, 5.8S, 45S, 5S, U5 small nuclear RNA, as well as mitochondrial mt12S) using BWA-MEM 0.7.17-r1188¹⁸; 2)
24 identification of human transcripts by mapping reads to the hg19 reference genome using the RNA-seq aligner HISAT2 (version
25 2.1.0, default settings)¹⁹; and 3) a second-round identification of human reads by aligning remaining reads to hg38 using
26 Kraken2 (version 2.0.8-beta, default settings)²⁰. All human RNA reads were then removed to generate qualified non-human
27 RNA-seq data. The number of human RNA-seq reads identified at each step is presented in **appendix 2 p 4: TableS3**.

28 **Characterization of viral communities in hospitalized patients with SARS-CoV-2 infection**

29 Before identification of virome and microbiota, SortMeRNA version 4.2.0²¹ (default settings) was applied to filter microbial
30 rRNA from non-human metatranscriptomic data. The remaining non-human non-rRNA reads were processed by Kraken2X v2.0.8
31 beta (default parameters)²⁰ with a self-built viral protein database by extracting protein sequences from all complete viral genomes
32 deposited in the NCBI RefSeq database (8,872 genomes downloaded on March 1st, 2020 including the SARS-CoV-2 complete
33 genome reference sequence, GCF_009858895.2). The number of reads annotated to each viral family was summarized based on
34 the read alignment results of Kraken 2X, and all RNA reads annotated to family *Coronaviridae* were considered as
35 SARS-CoV-2-like reads. For each sample, the ratio of SARS-CoV-2-like reads to total clean reads (referred to the SARS-CoV-2
36 viral load), and the ratio of SARS-CoV-2-like reads to total viral reads were calculated accordingly (**appendix 2 p 5: TableS4**).

37 After ranking the aligned reads of all detected viral species in each sample, non-*Coronaviridae* viral species with more than
38 10,000 aligned RNA reads were presented (**appendix 2 p 6: TableS5**) and selected for co-detection with known respiratory
39 viruses in hospitalized patients with SARS-CoV-2 infection. Three human respiratory viral species (including human
40 alphaherpesvirus 1, human orthopneumovirus and rhinovirus B) detected in respiratory samples from 4 severe cases (P01, P05,
41 P09 and P13) met the above criterion. Representative genomes of each species including human herpesvirus 1 strain 17
42 (NC_001806.2), human orthopneumovirus subgroup A (NC_038235.1) and rhinovirus B isolate 3039 (KF958308.1) were
43 downloaded from NCBI. For each patient, one representative sample (P01N201, P05S207, P09N205 and P12T211) with the
44 highest number of reads assigned to the targeted species was used for coverage analysis. Reads assigned to a given species were
45 aligned against the corresponding reference genome by bowtie2 v2.3.0 (the '-sensitive' mode)²². Sequencing depth and genome
46 coverage of each reference genome were determined with BEDTools coverage v2.27.1 (genomcov -ibam sort.bam -bg)²³.
47 Reliable co-detection with known respiratory viruses was defined when >50% of the genome was covered.

48 **Characterization of non-viral microbial communities in hospitalized patients with SARS-CoV-2** 49 **infection**

50 Non-viral microbial taxon assignment of the non-human non-rRNA reads was performed using clade-specific marker gene-based
51 MetaPhlan2 with the default parameter options for non-viral microbial composition (--ignore-viruses)²⁴. The presence

52 and relative abundances of non-viral microbial taxa at the phylum, genus and species were estimated (**appendix 2 p 7: Table S6**).
53 Mono-dominance of a given microbial taxa (genus or species) was defined if a taxon had a relative abundance >60% in one
54 sample as suggested by a previous study ²⁵. Most of the RNA reads of the two predominant bacterial genera *Burkholderia* and
55 *Parabacteroides*, respectively, identified in the respiratory and gastrointestinal tract of severe cases could hardly be assigned to
56 species level by MetaPhlan2, which might reflect that the two genera contain closely related species that are difficult to
57 differentiate by marker genes. In order to determine which species and how abundant species were in samples mono-dominated by
58 *Burkholderia* or *Parabacteroides*, we downloaded four reference genome sequences of the most frequently isolated BCC species
59 (*B. cenocepacia* J2315, *B. multivorans* ATCC BAA-247, *B. cepacia* ATCC 25416 and *B. dolosa* AU0158) and two gut
60 *Parabacteroides* species (*P. distasonis* ATCC 8503 and *P. merdae* NCTC13052). For each sample, reads were mapped against
61 corresponding references by bowtie2 v2.3.0, and the sequencing depth and genome coverage were estimated by BEDTools
62 coverage v2.27.1 as described above. The summary of coverage and depth of reference genomes for selected samples are
63 presented in **appendix 2 p 8: Table S7**. Likewise, metatranscriptomics-detected bacterial or fungal respiratory co-infections were
64 defined if the respiratory specimens (at least one sample) of severe patients were mono-dominated (relative abundance >60%) by
65 pathogenic microbes known to cause nosocomial infections (**appendix 2 p 9: Tables8**).

36 **Identification of expressed virulence factors (VFs) in *B. cenocepacia***

37 To identify the presence and expression patterns of potential virulence factors in *B. cenocepacia* identified in P01, we collected
38 multiple functional categories of virulence genes previously studied and verified by gene mutation analysis in *B. cenocepacia*
39 strains as well as corresponding gene ID in the annotated J2315 genome ²⁶, including 1) resistance to stress conditions, 2)
70 antimicrobial resistance, 3) quorum sensing, 4) iron uptake, 5) flagella and cable pilus, 6) lipopolysaccharide and 7)
71 exopolysaccharide. In addition, a pathogenicity island identified on chromosome 2 (BCAM0233-BCAM0281) of *B. cenocepacia*
72 J2315 by using comparative genomics was included ²⁷. Non-human non-rRNA microbial reads of all samples from P01 were
73 mapped against the reference genome of *B. cenocepacia* J2315 using bowtie2 v2.3.0 as described above and identified by the gene
74 IDs in the J2315 genome. For each sample, only virulence genes with more than 10 mapped reads were retained. A total of 50
75 expressed virulence genes were identified in clinical samples collected from P01 and presented in **appendix 2 p 10: Tables9**.

76 To compare the expression levels between different genes, we performed normalization of target gene expression levels among all
77 detected virulence genes using the following equation:

78
$$\text{Gene expression level}_i = \frac{\frac{N_i}{Lg_i}}{\sum_{i=1}^k \frac{N_i}{Lg_i}}$$

79 where i (1,2, ... k) refers to a given virulence gene identified in *B. cenocepacia* J2315; Lg_i is the length of gene i ; N_i is the reads
80 number that mapped to gene i .

31 **Statistical analyses**

32 Non-metric multidimensional scaling (NMDS) ordination of respiratory microbial community was conducted using the Manhattan
33 distances based on a presence/absence matrix of genus profiles of 47 respiratory specimens (7 from mild cases and 40 from severe
34 cases) (R version 3.6.1, vegan package). Kruskal–Wallis test and Wilcoxon rank-sum test were performed to compare differences
35 in SARS-CoV-2 viral loads among five different types of clinical specimens (throat swab, nasal swab, sputum, anal swab and
36 feces) and between overall respiratory specimens from mild and severe cases, respectively (R version 3.6.1, coin package).

37 **Results**

38 **Demographic information of patients and clinical specimens used in the study**

39 23 patients with COVID-19 hospitalized in the period January 10–March 31, 2020 in four hospitals in the Guangdong Province,
40 China, were enrolled into this study. 15 infected patients (40–80 years) admitted to the ICU and received mechanical ventilation
41 were defined as having severe COVID-19, and the remaining 8 patients (2–65 years) were mild cases (**appendix 2 p 2: TableS1**).
42 60 % (9 out of 15) of severely ill patients received invasive mechanical ventilation (**appendix 2 p 2: TableS1**), and 95.7% of the
43 patients (22 out of 23) received antiviral medications. To prevent and control nosocomial infections, all severe cases received
44 broad-spectrum antibiotics, and simultaneously 93.35% (14 out of 15) received antifungal agents (**appendix 2 p 2: TableS1**). By
45 contrast, none of the mild cases were treated with antibacterial or antifungal drugs. Up to March 31, 2020, 53.3% (8 out of 15) of
46 the severely ill patients had been transferred out of ICU or discharged from hospitals, and all mild cases had been discharged,
47 whereas an elderly male patient died one-month after admission to ICU (P01) (**appendix 2 p 2: TableS2**). Sixty-seven serial
48 clinical specimens from the respiratory tract (RT) (n=47, sputum, nasal and throat swab) and gastrointestinal tract (GIT) (n=20,
49 anal swab and feces) of these patients were obtained during the same above period for comprehensive assessment of microbial
50 characteristics after SARS-CoV-2 infection. A detailed timeline of specimen collections and clinical events for the 23 COVID-19
51 cases are shown in **appendix 2 p 3: TableS2**.

52 **Work flow of ultra-deep metatranscriptomic sequencing**

53 After quality control, an average of 268.3 Gb metatranscriptomic data were generated per sample (**appendix 2 p 4: TableS3**). We
54 applied an integrated bioinformatics pipeline to detect human, viral and non-viral microbial reads in total RNA-seq data
55 (**appendix 1 p 1: Supplementary figure 1 and Methods**). The percentage of human RNA reads (including ribosomal RNA and
56 non-rRNA transcripts) varied between different types of specimens, constituting a relatively high fraction of total high-quality
57 reads among RT specimens and a low fraction among GIT specimens (**appendix 1 p 1: Supplementary figure 1 and appendix 2**
58 **p 4: TableS3**). After removing host data, SortMeRNA was applied²¹ to filter microbial rRNA from the metatranscriptomic data.
59 The final remaining non-human non-rRNA data (ranged from 386Mb to 145Gb) were then used to assess viral and non-viral

10 microbial composition by Kraken2X²⁰ and MetaPhlan2²⁴, respectively(**Methods**). Detailed data statistics for each processing step
11 are provided in **appendix 2 p 4: TableS3**.

12 **Co-detection of viruses in clinical specimens**

13 First, viral RNA reads belonging to family *Coronaviridae* were defined as SARS-CoV-2-like reads. As expected, *Coronaviridae*
14 was the most abundant virus and was detected in all clinical specimens and varied between 0.01 and 286,418 mapped reads per
15 million (RPM) (**figure 1A**). No statistically significant differences were observed in SARS-CoV-2 viral loads in various types of
16 specimens (**appendix 1 p 2: Supplementary figure 2A**, Kruskal–Wallis test, $P=0.561$), or respiratory specimens
17 between mild and severe cases (**appendix 1 p 2: Supplementary figure 2B**, Wilcoxon rank-sum test, $P=0.455$). Although
18 SARS-CoV-2 viral loads in both RT and GIT gradually decreased at later time points of infection (**appendix 1 p 2:**
19 **Supplementary figure 2C-D**), viral shedding varied in different patients. For instance, RT specimens had consistently lower
20 SARS-CoV-2 viral loads than GIT specimens in P01, while specimens from the two sites showed comparable viral levels in P05
21 and P10 across all sampled time points (**appendix 1 p 2: Supplementary figure 2E**). Previous studies also reported inconsistent
22 temporal patterns of viral shedding in different types of specimens and patients^{28–30}, which might be affected by sample sizes,
23 specimens collected at different stages of infection as well as treatments with a cocktail of antimicrobial agents³¹.

24 Except for *Coronaviridae*, RNA-seq analysis also revealed a great diversity of viral composition in clinical samples from
25 infected patients. Natural hosts of the highly abundant viruses differed, including but not limited to animals (e.g., *Picornaviridae*,
26 *Pneumoviridae* and *Herpesviridae*), bacteria (e.g., *Podoviridae*, *Siphoviridae* and *Myoviridae*) and plants (*Virgaviridae*) (**figure**
27 **1B** and **appendix 2 p 5: Table S4**). The co-infections of known human respiratory viruses were further confirmed (viral
28 genome coverage>50%) in four out of thirteen severely ill patients with metatranscriptomic data of respiratory samples (30.8%),
29 including human alphaherpesvirus 1 in P01 and P05, rhinovirus B in P09, and human orthopneumovirus in P13 (**figure 1B** and
30 **appendix 1 p 4: Supplementary figure 3A-B**). The shedding patterns of human alphaherpesvirus 1 and human
31 orthopneumovirus in throat samples of P01 and P13 were similar to those of *Coronaviridae* in the same patient (**appendix 1 p 4:**
32 **Supplementary figure 3C-D**). By contrast, co-detection of other respiratory viruses was not observed in mild cases without ICU
33 admission (**appendix 2 p 5: Table S4**). Although some case studies reported the co-infection of SARS-CoV-2 and influenza
34 viruses^{32,33}, this was not observed in this Guangdong cohort.

35 Plant viruses belonging to the family *Virgaviridae*, especially Pepper mild mottle virus (PMMoV) and Tomato mosaic
36 virus (ToMV), were found to be the most dominant viruses in the feces (P53F0203 and P14F226) or in the throat swabs
37 (P62T0207) (**figure 2B** and **appendix 2 p 5: Table S4**). Some pioneer studies have also reported strong evidence supporting the
38 presence of PMMoV and ToMV in human-associated samples^{34–36}. All these findings have suggested the presence of plant-human

39 virus transmission. However, the potential pathogenicity of plant viruses in humans or in other vertebrates remains largely
40 unknown^{37,38}.

41 **Detection of non-viral microbial co-infections in clinical specimens**

42 We next evaluated the occurrence of bacterial/fungal co-infections that have been shown to be associated with worse clinical
43 outcomes than viral co-infections in hospitalized patients. Notably, results of hospital-laboratory based cultures and assays on
44 specimens demonstrated the presence of nosocomial fungal (n=1) and bacterial infections (n=3) in severely ill COVID-19 patients
45 (**appendix 2 p 2: Table S1**). In detail, one patient (P01) tested positive for (1-3)- β -d-glucan (a common component of the fungal
46 cell wall) in blood samples. Two patients (P04 and P20) had positive sputum cultures for *Burkholderia cepacia* complex (BCC)
47 species, the most common respiratory pathogens causing NP in cystic fibrosis (CF) patients^{39,40}. *Staphylococcus epidermidis*, a
48 typical skin bacterium that has been increasingly recognized as a multidrug-resistant nosocomial pathogen^{41,42}, was identified by
49 culturing of nasal secretions of one patient (P06).

50 Next, the non-viral microbial RNA composition of all 67 clinical specimens was analyzed to fully assess possible nosocomial
51 infections. As none of mild cases was admitted to ICU or received antibacterial/antifungal agents, we compared the respiratory
52 microbial communities between mild (n=7) and severe (n=13) cases to determine microbial dysbiosis, co-infection and their
53 associations with clinical outcomes. Remarkable microbial differences in RT specimens between mild and severe cases were
54 observed (**figure 2**). Although the specimens of the mild cases were collected from three different hospitals (Guangzhou,
55 Yangjiang and Qingyuan), their RT samples (six throat swabs and one sputum) consistently exhibited a larger number of detected
56 microbial taxa (**figure 2A** and **appendix 1 p 5: Supplementary figure 4A**) and similar microbial RNA community compositions
57 (**figure 2B** and **appendix 1 p 5: Supplementary figure 4B**). On average, 28 genera and 48 species were detected per RT sample
58 in mild cases while only 6 genera and 4 species were detected per RT sample in severe cases (**appendix 1 p 5: Supplementary**
59 **figure 4C**). The predominant RT bacteria in mild cases were *Veillonella*, *Neisseria*, *Streptococcus* and *Prevotella* (individual
60 occurrence >80% and mean relative abundance>5%) (**figure 2B,C**), which is consistent with common microbial communities
61 found in the human nasal and oral cavity⁴³. However, except *Veillonella*, each of the latter three genera enriched in mild cases was
62 only detected in few severe cases (n \leq 3, **figure 2A**) and had a mean abundance of less than 0.05% (**figure 2C**).

63 Of note, RT microbial features of severe cases were identified to be patient-specific. Among 40 respiratory samples from severe
64 patients, over 60% were mono-dominated (relative abundance>60%) by bacterial genus *Burkholderia* (11 samples from P01, P04
65 and P20), *Staphylococcus* (6 samples from P10 and P19) or *Mycoplasma* (7 samples from P05, P06, P14, and P18) (**figure 2B,C**
66 and **appendix 1 p 5: Supplementary figure 4D**). Each genus was detected in 69.2% of (9 out of 13) severe patients and 92.3%
67 (12 out of 13) of severe patients were positive for at least one of the three genera (**figure 2A**), indicating their prevalence in RT of
68 patients hospitalized with severe COVID-19. By mapping RNA reads to the reference genomes of BCC species (**Methods**), we
69 further confirmed the predominant expression of *B. cenocepacia* in the respiratory tract of P01, and *B. multivorans* in P04 and P20

70 who had positive sputum cultures of BCC (**appendix 1 p 6: Supplementary figure 5A** and **appendix 2 p 8: Table S7**). All
71 *Staphylococcus* RNA reads of RT samples from P06 (who also had positive *S. epidermidis* culture), P10 and P19 were also
72 assigned to *S. epidermidis* (**figure 2B** and **appendix 1 p 6: Supplementary figure 5B**). However, *S. aureus*, a major
73 hospital-acquired pathogen⁴⁴, was not detected in metatranscriptomic data of any of the sequenced RT samples. *Mycoplasma orale*
74 and *M. hominis*, rather than *M. pneumoniae*, were the two identified pathogenic *Mycoplasma* members (**figure 2B** and **appendix 1**
75 **p 6: Supplementary figure 5C**). *Propionibacterium* and *Escherichia* were also frequently detected in RT samples of severe cases
76 (individual occurrence >80%) but were less abundant than the former three genera (mean relative abundance <3%) (**figure 2C**).
77 All the five dominant bacterial genera in severe cases have been frequently associated with nosocomial infections, while they
78 were not detected or present in extremely low abundance in mild cases (relative abundance <0.15%) (**figure 2A,C**).

79 Ascomycetic transcripts (mainly from *Debaryomycetacea*) were identified in all five throat swabs collected from P13 (**figure 2B**
30 and **appendix 1 p 5 : Supplementary figure 4D**). Interestingly, the ascomycetes constituted only 5.7% of the total non-viral
31 microbes at the first time point (8 February) and increased to 18.2% ~87.4% at later points (11-20 February) (**appendix 1 p 5:**
32 **Supplementary figure 4D**). P13, an elderly man in his 70s, had onset of COVID-19 symptoms on January 30 and had been
33 admitted to the ICU since February 5 (**appendix 2 p 2: Table S1**). These observations suggested a rapid succession from
34 bacteria to fungi had occurred in the respiratory microbiota of P13 three days after ICU admission. These finding indicated that
35 serial monitoring to track other respiratory viruses (human orthopneumovirus, **appendix 1 p 4: Supplementary figure 3**) and
36 secondary fungal infections are required to avoid delayed treatment for such patients. Using the same criterion of microbial
37 monodominance (relative abundance >60%), we determined that a metatranscriptomics-based bacterial co-infection rate among
38 severely ill COVID-19 patients was 69.2% (9 out of 13), and 23.1% (3 out of 13, including P04, P06 and P20) of them had the
39 same results based on laboratory-confirmed bacterial cultures (**appendix 2 p 9: Table S8**). Considering all results from clinical
40 laboratories and metatranscriptomic analyses, we report that 84.6% (11 out of 13) of severely ill patients were co-infected with at
41 least one type of respiratory pathogenic microbes (**appendix 2 p 9: Table S8**).

42 Additionally, three severe cases (P05, P07 and P11) had high expression levels of *Veillonella* but low levels (or no detection) of
43 the above pathogenic genera in their RT samples (**figure 2A,B**). These individuals had been transferred out of the ICU or
44 discharged from hospitals (**appendix 2 p 2: Table S1**). Most bacterial pathogens in RT infections, including those we identified in
45 patients with COVID-19, are aerobic or facultative organisms, while *Veillonella spp.* are strictly anaerobic and have been reported
46 to be part of normal oral cavities and rarely isolated in nosocomial infections^{45,46}. Of note, all three patients did not receive
47 invasive ventilation during ICU admission (**appendix 2 p 2: Table S1**), and two of them (P07 and P11) were negative for
48 laboratory-confirmed or sequencing-detected co-infections (**appendix 2 p 9: Table S8**).

49 Similarly, severe cases appeared to have a distinct gut metatranscriptome compared to mild cases. Both of the two mild cases
50 having fecal specimens showed highly abundant *Campylobacter* in the gut metatranscriptome (**appendix 1 p 7: Supplementary**

31 **figure 6A**), different from previous results of healthy adults⁴⁷. While *Parabacteroides* (including *P. distasonis* and *P. merdae*)
32 mono-dominated the gut microbial transcripts in five samples from four severe cases and displayed a much higher abundance in
33 severe cases than that in mild cases (**appendix 1 p 7: Supplementary figure 6B-D** and **appendix 2 p 8: Table S7**). The extreme
34 bloom of *P. distasonis*, a low abundant but common taxa in the human gut, has been reported after beta-lactam ceftriaxone
35 treatment²⁵. Taken together, metatranscriptomic findings have not only complemented and enhanced the laboratory-diagnosed
36 microbial co-infections but also provided comprehensive information of microbial dysbiosis on SARS-CoV-2 infected patients.

37 **Secondary *B. cenocepacia* infection-associated death of a severely ill COVID-19 patient**

38 In order to identify factors associated with the in-hospital death of patient P01, serial clinical specimens were collected and
39 analyzed. Notably, a distinct timeline for secondary *B. cenocepacia* infection in this patient was clearly observed. On the first
40 day of sampling (January 27 2020), up to 99.9% of non-viral microbial transcripts in his throat swab were assigned to *B.*
41 *cenocepacia* (P01T127) while most of the transcripts in his nasal swab collected at the same day were from *E.coli* (P01N127)
42 (**figure 3A**). *B. cenocepacia* was predominantly present in both throat and nasal swabs for all the following sampling time points
43 (29 January–07 February 2020) (**figure 3A**). In addition, *B. cenocepacia* was detected in all his GIT samples, and the percentage
44 of *B. cenocepacia* to non-viral transcripts gradually increased from 9.5% to 64% (from January 29 to February 07) (**figure 3A**).
45 The time-dependent dynamics of transcript levels of *B. cenocepacia* suggested that transfer from the upper respiratory tract to the
46 lower gastrointestinal tract had caused a secondary systemic infection in P01. Our findings were also consistent with the patient's
47 death certificate record indicating septic shock as the cause of death (**appendix 2 p 2: Table S1**). Indeed, several studies have
48 pointed out that among BCC-infected CF patients, infection with *B. cenocepacia*, rather than other common isolated BCC
49 members (such as *B. multivorans* and *B. cepacia*), constituted the highest risk factor of death^{26,40}.

50 Next, virulence factors (VF) expressed by *B. cenocepacia* in P01 were analyzed in order to better understand the pathogenic
51 mechanisms of this possible lethal pathogen in this severe COVID-19 case (**Methods**). The gene *rpoE* (a member of
52 the extracytoplasmic function subfamily of sigma factors) was the most abundantly expressed VF during the entire sampling
53 period in P01 with SARS-CoV-2 infection (**figure 3B**). RpoE, as a stress response regulator, has been demonstrated to be essential
54 for the growth of *B. cenocepacia* and the delay of phagolysosomal fusion in macrophages during infection⁴⁸.
55 A delay in phagolysosomal fusion has also been reported to be an important host immune escape strategy for several bacterial
56 pathogens⁴⁹. Other VFs in response to oxidative stress conditions in the host environment, such as those encoding superoxide
57 dismutase, peroxidase or catalase (*sodC* and *katB*), were also expressed (**figure 3B, appendix 2 p 10: Table S9**). A panel of genes
58 belonging to RND (Resistance-Nodulation-Division) family transporters that confer multidrug resistance to *B. cenocepacia*⁵⁰
59 were also expressed in large amount in all types of specimens. We also detected expressions of genes encoding flagella and cable
60 pilus (*fliC*, *fliI*, *fliG* and *adhA*), which can facilitate the bacterial adhesion to host cells and mucin⁵¹ (**figure 3B**). Besides,
61 expressions of genes involved in quorum sensing, iron uptake (by competing with the host for iron), biosynthesis of

lipopolysaccharide (LPS) and exopolysaccharide (EPS) were also detected (**figure 3B**), indicating their active roles in the regulation of bacterial cell aggregation, biofilm formation and toxin production during infection.

Discussion

In this study, ultra-deep metatranscriptomic sequencing in combination with clinical laboratory diagnosis, including cultures, and colorimetric assays provided a characteristic spectrum of co-detected respiratory pathogenic microbes in a Guangdong hospitalized patient cohort infected with SARS-CoV-2, which is distinct from prior results on the most common bacterial co-infections identified in previous coronavirus outbreak and influenza pandemics⁴⁻⁸ or common pathogens associated with nosocomial respiratory infections in China^{52,53}.

In particular, respiratory co-infections with BCC were detected in 23.1% of severe cases with evidence from both laboratory cultures and metatranscriptomic results. The serial metatranscriptomic data of all specimens from P01 have revealed the timeline of a secondary infection with *B. cenocepacia* alongside the expression of various virulence genes, which could confer the abilities of the lethal pathogen to evade host defenses (e.g., *rpoE*), adhere target tissues (e.g., flagella-coding genes and *adhA*), produce toxins (e.g. genes encoding biosynthetic enzymes for production of LPS and EPS) and resist the effects of multiple antibiotics (RND family), which eventually may have led to the fatal outcome of the patient. Interestingly, some severely ill COVID-19 patients have been reported to display CF-like pulmonary pathological features including mucus plugs, bronchial wall thickening, and a subsequently impaired mucociliary clearance⁵⁴⁻⁵⁷, which could promote the adherence and colonization of mucin-degrading pathogens in the respiratory tract. Furthermore, BCC bacteria, a major threat to hospitalized CF patients, were found to be predominantly localized in the phagocytes and mucus of CF patients⁵⁸. Even though the small sample size in our study limited our ability to assess the general prevalence of BCC infections among severely ill SARS-CoV-2 patients, our findings highlight the need for monitoring and controlling nosocomial BCC/SARS-CoV-2 co-infection by using rapid diagnostic technologies such as PCR-based multilocus sequence typing and enzyme-linked immunosorbent assay. Metatranscriptomic data also revealed that 30.8% of severely ill patients had respiratory tract co-infections associated with *Mycoplasma spp.* (including *M. hominis* and *M. orale*), which lack a cell wall and have inherent resistance to commonly administered beta-lactam antibiotics. Although *Mycoplasma* usually causes mild illness, it has also been associated with serious infections in seniors and immunocompromised individuals^{59,60}, who have a high risk for developing severe symptoms from COVID-19. Moreover, we did observe super-high expression levels of *M. orale* in the RT of two severely ill patients (P14 and P18) with prolonged ICU stay (>30 days). Thus, our results suggest that the possibility of *Mycoplasma*-associated co-infection in severely ill COVID-19 patients warrants increased attention.

The SARS-CoV-2 infection is often accompanied by GI symptoms such as vomiting, diarrhea and abdominal pain^{61,62}, and the virus has been recently been proven to infect the human gastrointestinal tract⁶³⁻⁶⁵. The dominating Proteobacteria transcripts in

32 the gut of antibiotic treatment naïve mild COVID-19 cases in this study as well as the marked gut microbial changes during
33 respiratory viral infections in prediabetes reported in a recent longitudinal multi-omics study⁶⁶ both support interactions between
34 host, respiratory viruses and the commensal microbiota, possibly via the mucosal immune system^{67,68}.

35 Currently, no effective drugs have been licensed for human use against SARS-CoV-2 infection, but a preliminary study has shown
36 an effect of early treatment with Remdesivir⁶⁹. Instead, the widespread use of antimicrobial agents (including broad-spectrum
37 antibiotics) has been documented in many studies (including ours)^{9,12} for the prevention and treatment of secondary infections in
38 COVID-19 patients, especially in those who needed mechanical ventilation. Without initial specimens before any treatments, it is
39 difficult to distinguish to what extent differences in the respiratory microbial patterns associated with SARS-CoV-2 infections
40 reflect the disease or the antimicrobial treatment, or both. Still, our results suggest that the distinct bacterial communities detected
41 in the respiratory and gastrointestinal tract of the severely ill patients might relate to the significant disruption of the normal
42 human microbiota caused by the treatment with antibiotics, allowing colonization by pathogenic antibiotic-resistant bacteria.
43 Worldwide attention should be drawn to potential future threats from a blooming reservoir of antimicrobial-resistant organisms
44 and associated genes after the COVID-19 pandemic.

75 **Contributors**

76 J.Z., J.L., Y.L and X.X. conceived the project. Y.W., L.Z., Z.Z., A.Z., J. Zhao., W.H., F.X., F.Ye., B.Z. and S.R. collected clinical
77 specimens and clinical information from patients, processed of RNA extraction, and performed in-hospital diagnostic tests. Ji.L.
78 and P.R. performed metatranscriptomic library construction. H.Z., Z.S., H.R., F.Yang., T.L., J.Z. and F.L. processed the
79 sequencing data and conducted bioinformatic analyses. H.Z., Z.S., H.R., and J.L. interpreted the data. H.Z. and J.L. wrote the first
80 version of the manuscript. J.Z., K.K. and H.M.T. contributed substantially to the revisions of the manuscript. All other authors
81 participated in discussions, provided useful comments and revised the manuscript. All authors approved the final version.

32 **Declaration of interests**

33 The authors declare no competing interests.

34 **Acknowledgments**

35 This work was funded by National Science and Technology Major Project of China (No:2017ZX10303406), the National Major
36 Project for Control and Prevention of Infectious Disease in China (2018ZX10301101-004), the emergency grants for prevention
37 and control of SARS-CoV-2 of Ministry of Science and Technology (2020YFC0841400) and Guangdong province
38 (2020B1111320003, 2020B111108001, 2018B020207013), Guangdong Provincial Key Laboratory of Genome Read and Write

39 (No. 2017B030301011), Guangdong Provincial Academician Workstation of BGI Synthetic Genomics (No. 2017B090904014),
30 and Shenzhen Engineering Laboratory for Innovative Molecular Diagnostics (DRC-SZ[2016]884). Metatranscriptomic
31 sequencing of this work was supported by China National GeneBank.

32 **Data and software availability**

33 The non-human metatranscriptomic sequencing data from 67 clinical specimens that support the findings of this study have been
34 deposited into CNSA (CNGB Nucleotide Sequence Archive) of CNGBdb with accession number CNP0001066
35 (<https://db.cngb.org/cnsa/>). Details of software, code and parameters used for data analyses of the current study are provided and
36 publicly available from GitHub (<https://github.com/BGI-IORI/RNA-seq-2019nCov>).

37

38

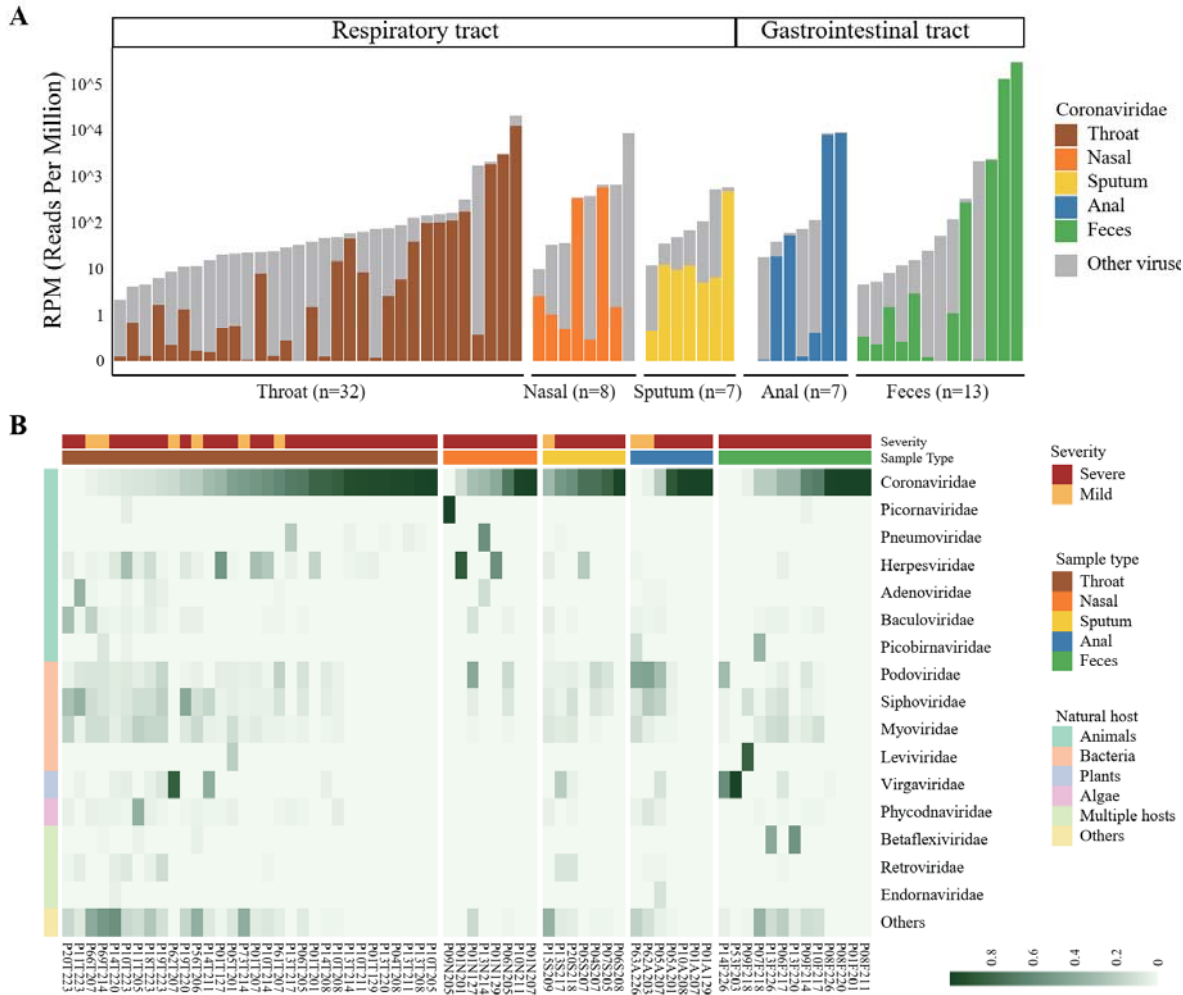
39 References

- 40 1 World Health Organization. Coronavirus disease (COVID-19) Situation Report–163. 2020
41 [https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200701-covid-19-sitrep-163.pdf?sfvrsn=c202f0](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200701-covid-19-sitrep-163.pdf?sfvrsn=c202f05b_2)
42 [5b_2](https://www.who.int/docs/default-source/coronaviruse/situation-reports/20200701-covid-19-sitrep-163.pdf?sfvrsn=c202f05b_2).
- 43 2 Short KR, Kroeze EJBV, Fouchier RAM, Kuiken T. Pathogenesis of influenza-induced acute respiratory distress
44 syndrome. *Lancet Infect. Dis.* 2014. DOI:10.1016/S1473-3099(13)70286-X.
- 45 3 De Wit E, Van Doremalen N, Falzarano D, Munster VJ. SARS and MERS: Recent insights into emerging coronaviruses.
46 *Nat. Rev. Microbiol.* 2016. DOI:10.1038/nrmicro.2016.81.
- 47 4 Peiris JSM, Chu CM, Cheng VCC, *et al.* Clinical progression and viral load in a community outbreak of
48 coronavirus-associated SARS pneumonia: A prospective study. *Lancet* 2003. DOI:10.1016/S0140-6736(03)13412-5.
- 49 5 Morris DE, Cleary DW, Clarke SC. Secondary bacterial infections associated with influenza pandemics. *Front. Microbiol.*
50 2017. DOI:10.3389/fmicb.2017.01041.
- 51 6 Bacterial coinfections in lung tissue specimens from fatal cases of 2009 pandemic influenza a (H1N1) - United States,
52 May-August 2009. *Morb Mortal Wkly Rep* 2009. DOI:10.1097/01.inf.0000366838.82857.78.
- 53 7 Gill JR, Sheng ZM, Ely SF, *et al.* Pulmonary pathologic findings of fatal 2009 Pandemic influenza A/H1N1 viral
54 infections. *Arch Pathol Lab Med* 2010. DOI:10.1043/1543-2165-134.2.235.
- 55 8 Weinberger DM, Simonsen L, Jordan R, Steiner C, Miller M, Viboud C. Impact of the 2009 influenza pandemic on
56 pneumococcal pneumonia hospitalizations in the United States. *J Infect Dis* 2012. DOI:10.1093/infdis/jir749.
- 57 9 Chen N, Zhou M, Dong X, *et al.* Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus
58 pneumonia in Wuhan, China: a descriptive study. *Lancet* 2020. DOI:10.1016/S0140-6736(20)30211-7.
- 59 10 Meng L, Qiu H, Wan L, *et al.* Intubation and Ventilation amid the COVID-19 Outbreak: Wuhan's Experience.
60 *Anesthesiology* 2020. DOI:10.1097/ALN.0000000000003296.
- 61 11 Bonten MJM, Kollef MH, Hall JB. Risk Factors for Ventilator-associated Pneumonia: From Epidemiology to Patient
62 Management. *Clin Infect Dis* 2004. DOI:10.1086/383039.
- 63 12 Zhou F, Yu T, Du R, *et al.* Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan,
64 China: a retrospective cohort study. *Lancet* 2020. DOI:10.1016/S0140-6736(20)30566-3.
- 65 13 Kim D, Quinn J, Pinsky B, Shah NH, Brown I. Rates of Co-infection between SARS-CoV-2 and Other Respiratory
66 Pathogens. *JAMA - J. Am. Med. Assoc.* 2020. DOI:10.1001/jama.2020.6266.
- 67 14 Garner JS, Jarvis WR, Emori TG, Horan TC, Hughes JM. CDC definitions for nosocomial infections, 1988. *AJIC Am J*
68 *Infect Control* 1988. DOI:10.1016/0196-6553(88)90053-3.
- 69 15 Kalil AC, Metersky ML, Klompas M, *et al.* Management of Adults With Hospital-acquired and Ventilator-associated
70 Pneumonia: 2016 Clinical Practice Guidelines by the Infectious Diseases Society of America and the American Thoracic
71 Society. *Clin. Infect. Dis.* 2016. DOI:10.1093/cid/ciw353.
- 72 16 Xiao M, Liu X, Ji J, *et al.* Multiple approaches for massively parallel sequencing of HCoV-19 genomes directly from
73 clinical samples 2. *bioRxiv* 2020. DOI:10.1101/2020.03.16.993584.
- 74 17 Chen S, Zhou Y, Chen Y, Gu J. Fastp: An ultra-fast all-in-one FASTQ preprocessor. In: *Bioinformatics*. 2018.
75 DOI:10.1093/bioinformatics/bty560.
- 76 18 Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 2009.
77 DOI:10.1093/bioinformatics/btp324.
- 78 19 Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and
79 HISAT-genotype. *Nat Biotechnol* 2019. DOI:10.1038/s41587-019-0201-4.
- 80 20 Wood DE, Lu J, Langmead B. Improved metagenomic analysis with Kraken 2. *Genome Biol* 2019.
81 DOI:10.1186/s13059-019-1891-0.
- 82 21 Kopylova E, Noé L, Touzet H. SortMeRNA: Fast and accurate filtering of ribosomal RNAs in metatranscriptomic data.
83 *Bioinformatics* 2012. DOI:10.1093/bioinformatics/bts611.
- 84 22 Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 2012. DOI:10.1038/nmeth.1923.
- 85 23 Quinlan AR, Hall IM. BEDTools: A flexible suite of utilities for comparing genomic features. *Bioinformatics* 2010.
86 DOI:10.1093/bioinformatics/btq033.

- 17 24 Truong DT, Franzosa EA, Tickle TL, *et al.* MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nat Methods*
18 2015; **12**: 902–3.
- 19 25 Hildebrand F, Moitinho-Silva L, Blasche S, *et al.* Antibiotics-induced monodominance of a novel gut bacterial order. *Gut*
20 2019. DOI:10.1136/gutjnl-2018-317715.
- 21 26 Drevinek P, Mahenthiralingam E. Burkholderia cenocepacia in cystic fibrosis: Epidemiology and molecular mechanisms
22 of virulence. *Clin. Microbiol. Infect.* 2010. DOI:10.1111/j.1469-0691.2010.03237.x.
- 23 27 Holden MTG, Seth-Smith HMB, Crossman LC, *et al.* The genome of Burkholderia cenocepacia J2315, an epidemic
24 pathogen of cystic fibrosis patients. *J Bacteriol* 2009. DOI:10.1128/JB.01230-08.
- 25 28 Zou L, Ruan F, Huang M, *et al.* SARS-CoV-2 viral load in upper respiratory specimens of infected patients. *N. Engl. J.*
26 *Med.* 2020. DOI:10.1056/NEJMc2001737.
- 27 29 He X, Lau EHY, Wu P, *et al.* Temporal dynamics in viral shedding and transmissibility of COVID-19. *Nat Med* 2020.
28 DOI:10.1038/s41591-020-0869-5.
- 29 30 Liu Y, Yan LM, Wan L, *et al.* Viral dynamics in mild and severe cases of COVID-19. *Lancet Infect. Dis.* 2020.
30 DOI:10.1016/S1473-3099(20)30232-2.
- 31 31 Zheng S, Fan J, Yu F, *et al.* Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang
32 province, China, January-March 2020: retrospective cohort study. *BMJ* 2020. DOI:10.1136/bmj.m1443.
- 33 32 Azekawa S, Namkoong H, Mitamura K, Kawaoka Y, Saito F. Co-infection with SARS-CoV-2 and influenza A virus.
34 *IDCases.* 2020; **20**: e00775.
- 35 33 Wu X, Cai Y, Huang X, *et al.* Co-infection with SARS-CoV-2 and Influenza A Virus in Patient with Pneumonia, China.
36 *Emerg Infect Dis* 2020. DOI:10.3201/eid2606.200299.
- 37 34 Zhang T, Breitbart M, Lee WH, *et al.* RNA viral community in human feces: Prevalence of plant pathogenic viruses.
38 *PLoS Biol* 2006. DOI:10.1371/journal.pbio.0040003.
- 39 35 Liu R, Vaishnav RA, Roberts AM, Friedland RP. Humans Have Antibodies against a Plant Virus: Evidence from Tobacco
40 Mosaic Virus. *PLoS One* 2013. DOI:10.1371/journal.pone.0060621.
- 41 36 Colson P, Richet H, Desnues C, *et al.* Pepper mild mottle virus, a plant virus associated with specific immune responses,
42 fever, abdominal pains, and pruritus in humans. *PLoS One* 2010. DOI:10.1371/journal.pone.0010041.
- 43 37 Balique F, Lecoq H, Raoult D, Colson P. Can plant viruses cross the kingdom border and be pathogenic to humans?
44 *Viruses.* 2015. DOI:10.3390/v7042074.
- 45 38 Jiwaji M, Matcher GF, de Bruyn MM, *et al.* Providence virus: An animal virus that replicates in plants or a plant virus that
46 infects and replicates in animal cells? *PLoS One* 2019. DOI:10.1371/journal.pone.0217494.
- 47 39 Courtney JM, Dunbar KEA, McDowell A, *et al.* Clinical outcome of Burkholderia cepacia complex infection in cystic
48 fibrosis adults. *J Cyst Fibros* 2004. DOI:10.1016/j.jcf.2004.01.005.
- 49 40 Zlosnik JEA, Zhou G, Brant R, *et al.* Burkholderia species infections in patients with cystic fibrosis in British Columbia,
50 Canada: 30 years' experience. *Ann Am Thorac Soc* 2015. DOI:10.1513/AnnalsATS.201408-395OC.
- 51 41 Miragaia M, Couto I, Pereira SFF, *et al.* Molecular characterization of methicillin-resistant Staphylococcus epidermidis
52 clones: Evidence of geographic dissemination. *J Clin Microbiol* 2002. DOI:10.1128/JCM.40.2.430-438.2002.
- 53 42 Lee JYH, Monk IR, Gonçalves da Silva A, *et al.* Global spread of three multidrug-resistant lineages of Staphylococcus
54 epidermidis. *Nat Microbiol* 2018. DOI:10.1038/s41564-018-0230-7.
- 55 43 Huffnagle GB, Dickson RP, Lukacs NW. The respiratory tract microbiome and lung inflammation: A two-way street.
56 *Mucosal Immunol.* 2017. DOI:10.1038/mi.2016.108.
- 57 44 Klein E, Smith DL, Laxminarayan R. Hospitalizations and deaths caused by methicillin-resistant Staphylococcus aureus,
58 United States, 1999-2005. *Emerg Infect Dis* 2007. DOI:10.3201/eid1312.070629.
- 59 45 Mager DL, Ximenez-Fyvie LA, Haffajee AD, Socransky SS. Distribution of selected bacterial species on intraoral
60 surfaces. *J Clin Periodontol* 2003. DOI:10.1034/j.1600-051X.2003.00376.x.
- 61 46 Takeshita T, Kageyama S, Furuta M, *et al.* Bacterial diversity in saliva and oral health-related conditions: The Hisayama
62 Study. *Sci Rep* 2016. DOI:10.1038/srep22164.
- 63 47 Abu-Ali GS, Mehta RS, Lloyd-Price J, *et al.* Metatranscriptome of human faecal microbial communities in a cohort of
64 adult men. *Nat Microbiol* 2018. DOI:10.1038/s41564-017-0084-4.
- 65 48 Flanagan RS, Valvano MA. Burkholderia cenocepacia requires RpoE for growth under stress conditions and delay of

- 36 phagolysosomal fusion in macrophages. *Microbiology* 2008. DOI:10.1099/mic.0.2007/013714-0.
- 37 49 Uribe-Quero E, Rosales C. Control of phagocytosis by microbial pathogens. *Front. Immunol.* 2017.
38 DOI:10.3389/fimmu.2017.01368.
- 39 50 Bazzini S, Udine C, Sass A, *et al.* Deciphering the role of rnd efflux transporters in burkholderia cenocepacia. *PLoS One*
40 2011. DOI:10.1371/journal.pone.0018902.
- 41 51 Tomich M, Mohr CD. Adherence and autoaggregation phenotypes of a Burkholderia cenocepacia cable pilus mutant.
42 *FEMS Microbiol Lett* 2003. DOI:10.1016/S0378-1097(03)00785-7.
- 43 52 Ding JG, Sun QF, Li KC, *et al.* Retrospective analysis of nosocomial infections in the intensive care unit of a tertiary
44 hospital in China during 2003 and 2007. *BMC Infect Dis* 2009. DOI:10.1186/1471-2334-9-115.
- 45 53 Tao L, Hu B, Rosenthal VD, Gao X, He L. Device-associated infection rates in 398 intensive care units in Shanghai,
46 China: International Nosocomial Infection Control Consortium (INICC) findings. *Int J Infect Dis* 2011.
47 DOI:10.1016/j.ijid.2011.06.009.
- 48 54 Lyczak JB, Cannon CL, Pier GB. Lung infections associated with cystic fibrosis. *Clin. Microbiol. Rev.* 2002.
49 DOI:10.1128/CMR.15.2.194-222.2002.
- 50 55 Bernheim A, Mei X, Huang M, *et al.* Chest CT Findings in Coronavirus Disease-19 (COVID-19): Relationship to
51 Duration of Infection. *Radiology* 2020. DOI:10.1148/radiol.2020200463.
- 52 56 Li K, Wu J, Wu F, *et al.* The Clinical and Chest CT Features Associated with Severe and Critical COVID-19 Pneumonia.
53 *Invest Radiol* 2020. DOI:10.1097/RLI.0000000000000672.
- 54 57 Wu J, Wu X, Zeng W, *et al.* Chest CT Findings in Patients With Coronavirus Disease 2019 and Its Relationship With
55 Clinical Features. *Invest Radiol* 2020. DOI:10.1097/RLI.0000000000000670.
- 56 58 Schwab U, Abdullah LH, Perlmutter OS, *et al.* Localization of Burkholderia cepacia complex bacteria in cystic fibrosis
57 lungs and interactions with Pseudomonas aeruginosa in hypoxic mucus. *Infect Immun* 2014. DOI:10.1128/IAI.01876-14.
- 58 59 Paessler M, Levinson A, Patel JB, Schuster M, Minda M, Nachamkin I. Disseminated Mycoplasma orale infection in a
59 patient with common variable immunodeficiency syndrome. *Diagn Microbiol Infect Dis* 2002.
60 DOI:10.1016/S0732-8893(02)00429-7.
- 61 60 García C, Ugalde E, Monteagudo I, *et al.* Isolation of Mycoplasma hominis in critically ill patients with pulmonary
62 infections: Clinical and microbiological analysis in an intensive care unit. *Intensive Care Med* 2007.
63 DOI:10.1007/s00134-006-0435-8.
- 64 61 Pan L, Mu M, Yang P, *et al.* Clinical Characteristics of COVID-19 Patients With Digestive Symptoms in Hubei, China: A
65 Descriptive, Cross-Sectional, Multicenter Study. *Am J Gastroenterol* 2020. DOI:10.14309/ajg.0000000000000620.
- 66 62 Wong SH, Lui RNS, Sung JJY. Covid-19 and the digestive system. *J. Gastroenterol. Hepatol.* 2020.
67 DOI:10.1111/jgh.15047.
- 68 63 Lamers MM, Beumer J, van der Vaart J, *et al.* SARS-CoV-2 Productively Infects Human Gut Enterocytes. *bioRxiv* 2020.
69 DOI:10.1101/2020.04.25.060350.
- 70 64 Xiao F, Tang M, Zheng X, Liu Y, Li X, Shan H. Evidence for Gastrointestinal Infection of SARS-CoV-2.
71 *Gastroenterology* 2020. DOI:10.1053/j.gastro.2020.02.055.
- 72 65 Xiao F, Sun J, Xu Y, *et al.* Infectious SARS-CoV-2 in Feces of Patient with Severe COVID-19. *Emerg Infect Dis J* 2020;
73 **26**. DOI:10.3201/eid2608.200681.
- 74 66 Zhou W, Sailani MR, Contrepois K, *et al.* Longitudinal multi-omics of host-microbe dynamics in prediabetes. *Nature*
75 2019. DOI:10.1038/s41586-019-1236-x.
- 76 67 Wells JM, Rossia O, Meijerink M, Van Baarlen P. Epithelial crosstalk at the microbiota-mucosal interface. *Proc Natl*
77 *Acad Sci U S A* 2011. DOI:10.1073/pnas.1000092107.
- 78 68 Marsland BJ, Trompette A, Gollwitzer ES. The gut-lung axis in respiratory disease. In: *Annals of the American Thoracic*
79 *Society*. 2015. DOI:10.1513/AnnalsATS.201503-133AW.
- 80 69 Beigel JH, Tomashek KM, Dodd LE, *et al.* Remdesivir for the Treatment of Covid-19 — Preliminary Report. *N Engl J*
81 *Med* 2020; published online May. DOI:10.1056/NEJMoa2007764.

15 **Figures and Figure legends**

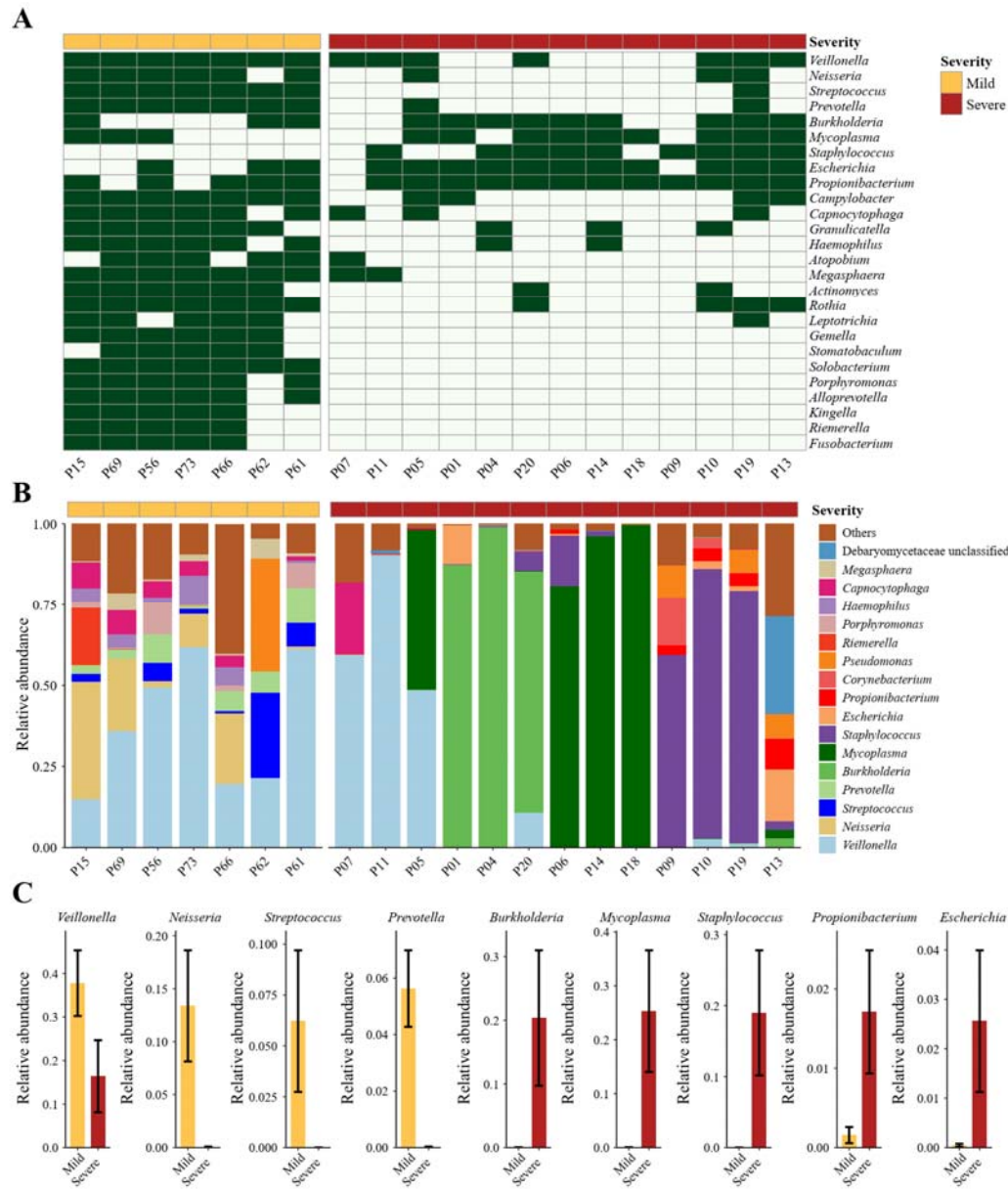


16
17 **Figure 1: Viral RNA profiles in clinical specimens of hospitalized patients with COVID-19**

18 (A) Bar plot showing the number of total viral reads and *Coronaviridae* reads in 67 clinical specimens collected from the
19 respiratory and gastrointestinal tract. Data have been normalized to total sequencing reads in reads-per-million (RPM). The
20 number of *Coronaviridae* reads of different sample types is colored as follows: brown, throat swab; orange, nasal swab; yellow,
21 sputum; blue, anal swab; green, feces. Gray, number of non-*Coronaviridae* viral reads.

22 (B) Heatmap showing the viral RNA abundance at the family level. Top16 viral families are shown and ranked according to their
23 natural hosts: green, animals; pink, bacteria; light blue, plant; purple, algae; light green, multiple host species; yellow, others (viral
24 families of low abundances). Specimens from patients with mild and severe COVID-19 symptoms are colored by brick red and
25 orange, respectively.

26



57

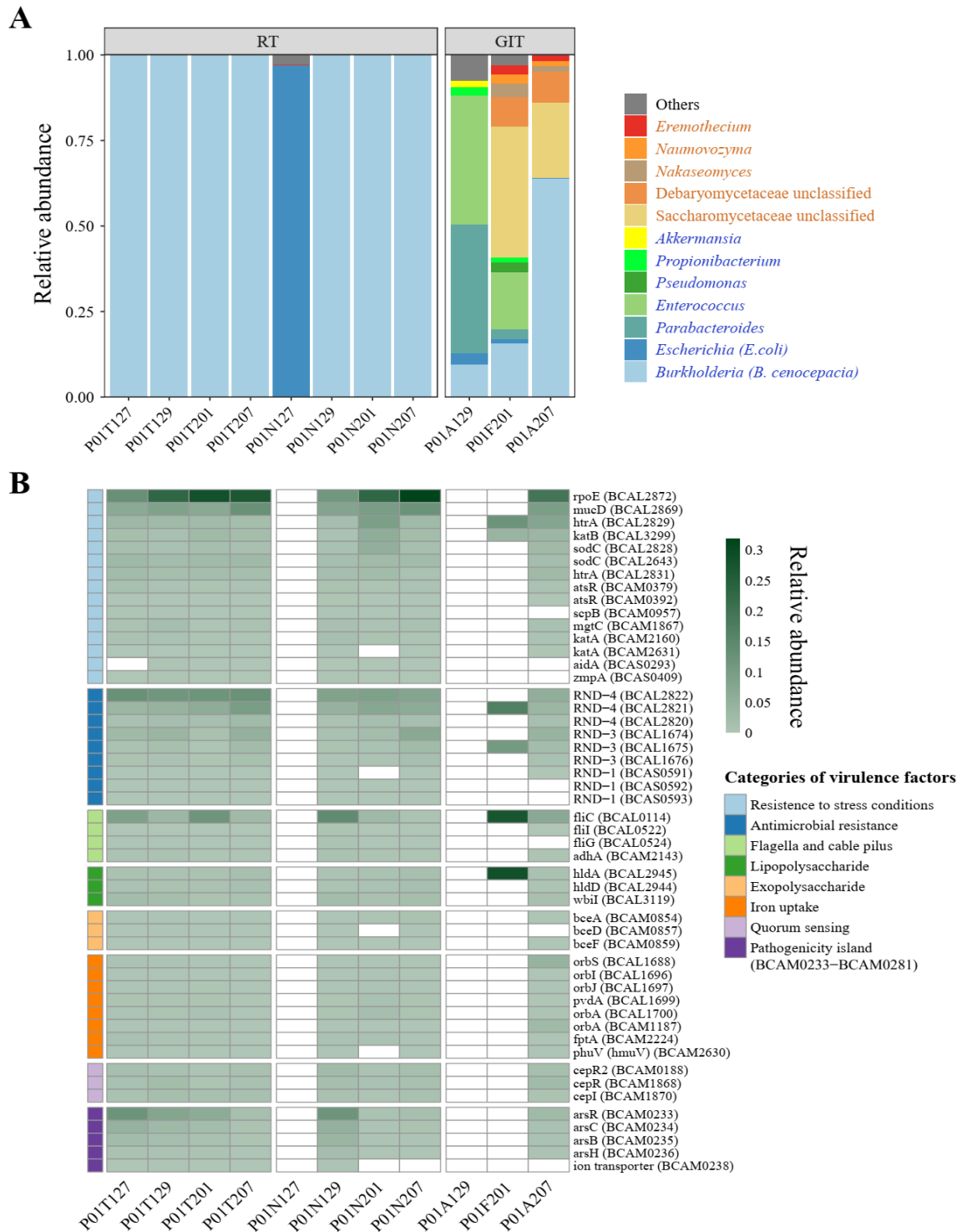
58 **Figure 2: Distinct respiratory microbial signatures in mild and severe cases**

59 (A) Presence/absence profile of non-viral microbial genera in mild and severe cases. Orange, mild; brick red, severe. Only
 60 common genera detected in over 60% of patients in the mild cases (n>4) or severe cases (n>7) are shown.

61 (B) Bar plot showing the relative expression levels of non-viral microbes in mild and severe cases.

62 (C) Relative expression levels of selected genera differing between mild and severe cases. For each genus, the bar chart and black
 63 error bars denote the mean and standard error values of expression levels in mild (orange) and severe (brick red) cases.

64 For severe cases with more than one sequenced respiratory clinical specimens: the presence of a given genus is considered if any
 65 of samples from this patient had positive detection of this genus (relative abundance >0) (A), and the mean relative abundance of
 66 each detected genus across all samples is calculated to represent its expression level in this patient (B-C).



37

38 **Figure 3: Identification of secondary *B. cenocepacia* infection in P01**

39 (A) Bar plot showing the relative expression levels of non-viral microbes in all specimens from P01.

70 A total of 11 specimens from the respiratory tract (RT) and gastrointestinal tract (GIT) are shown. Orange, fungi; blue, bacteria.

71 (B) Heatmap showing the relative expression levels of virulence factors of *B. cenocepacia*. A total of 50 identified virulence genes

72 are shown and ranked according to their functional categories: light blue, resistance to stress conditions; blue, antimicrobial

73 resistance; light green, flagella and cable pilus; green, lipopolysaccharide; pink, exopolysaccharide; orange, iron uptake; light

74 purple, quorum sensing; purple, genes located in a pathogenicity island.