

ARTICLE

Open Access

Characterization of respiratory microbial dysbiosis in hospitalized COVID-19 patients

Huanzi Zhong^{1,2}, Yanqun Wang³, Zhun Shi¹, Lu Zhang^{4,5}, Huahui Ren^{1,2}, Weiqun He³, Zhaoyong Zhang³, Airu Zhu³, Jingxian Zhao³, Fei Xiao⁶, Fangming Yang^{1,7}, Tianzhu Liang^{1,8}, Feng Ye³, Bei Zhong⁹, Shicong Ruan¹⁰, Mian Gan³, Jiahui Zhu^{1,11}, Fang Li³, Fuqiang Li^{1,12}, Daxi Wang^{1,8}, Jiandong Li^{1,8,13}, Peidi Ren^{1,8}, Shida Zhu^{1,14}, Huanming Yang^{1,15,16}, Jian Wang^{1,15}, Karsten Kristiansen^{1,2}, Hein Min Tun¹⁷, Weijun Chen^{13,18}, Nanshan Zhong³, Xun Xu^{1,19}, Yi-min Li³, Junhua Li^{1,8,20} and Jincun Zhao^{3,4}

Abstract

Severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has caused a global pandemic of Coronavirus disease 2019 (COVID-19). However, the microbial composition of the respiratory tract and other infected tissues as well as their possible pathogenic contributions to varying degrees of disease severity in COVID-19 patients remain unclear. Between 27 January and 26 February 2020, serial clinical specimens (sputum, nasal and throat swab, anal swab and feces) were collected from a cohort of hospitalized COVID-19 patients, including 8 mildly and 15 severely ill patients in Guangdong province, China. Total RNA was extracted and ultra-deep metatranscriptomic sequencing was performed in combination with laboratory diagnostic assays. We identified distinct signatures of microbial dysbiosis among severely ill COVID-19 patients on broad spectrum antimicrobial therapy. Co-detection of other human respiratory viruses (including human alphaherpesvirus 1, rhinovirus B, and human orthopneumovirus) was demonstrated in 30.8% (4/13) of the severely ill patients, but not in any of the mildly affected patients. Notably, the predominant respiratory microbial taxa of severely ill patients were *Burkholderia cepacia* complex (BCC), *Staphylococcus epidermidis*, or *Mycoplasma spp.* (including *M. hominis* and *M. orale*). The presence of the former two bacterial taxa was also confirmed by clinical cultures of respiratory specimens (expectorated sputum or nasal secretions) in 23.1% (3/13) of the severe cases. Finally, a time-dependent, secondary infection of *B. cenocepacia* with expressions of multiple virulence genes was demonstrated in one severely ill patient, which might accelerate his disease deterioration and death occurring one month after ICU admission. Our findings point to SARS-CoV-2-related microbial dysbiosis and various antibiotic-resistant respiratory microbes/pathogens in hospitalized COVID-19 patients in relation to disease severity. Detection and tracking strategies are needed to prevent the spread of antimicrobial resistance, improve the treatment regimen and clinical outcomes of hospitalized, severely ill COVID-19 patients.

Correspondence: Xun Xu (xuxun@genomics.cn) or Yi-min Li (dryiminli@vip.163.com) or Junhua Li (lijunhua@genomics.cn) or Jincun Zhao (zhaojincun@gird.cn)

¹BGI-Shenzhen, Shenzhen 518083, China

²Laboratory of Genomics and Molecular Biomedicine, Department of Biology, University of Copenhagen, 2100 Copenhagen, Denmark

Full list of author information is available at the end of the article

These authors contributed equally: Huanzi Zhong, Yanqun Wang, Zhun Shi, Lu Zhang, Huahui Ren, Weiqun He, Zhaoyong Zhang, Airu Zhu, Jingxian Zhao, Fei Xiao

These authors jointly supervised this work: Xun Xu, Yi-Min Li, Junhua Li, Jincun Zhao

Introduction

As of 31 January 2021, severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) has infected more than 102 million and resulted in more than 2.2 million deaths worldwide¹. The pandemic poses a significant threat to public health and the global economy.

Respiratory viruses, such as coronaviruses and influenza virus, can lead to acute damage of the epithelial barrier and facilitate invasions of other pathogens^{2,3}. For instance,

© The Author(s) 2021



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

secondary infections by *Stenotrophomonas maltophilia*, *Klebsiella pneumoniae*, or *Escherichia coli* were reported to cause serious complications in patients with SARS, such as bacteremia, sepsis, and nosocomial pneumonia (NP)⁴. In addition, *Streptococcus pneumoniae*, *Haemophilus influenzae*, and *Staphylococcus aureus* were frequently associated with NP and mortality in influenza pandemics⁵. It was estimated that approximately 29%–55% of the total 300,000 deaths in the 2009 H1N1 pandemic were caused by secondary bacterial NP^{6–8}.

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

Here, combining diagnostic technologies (cultures and colorimetric assays) and metatranscriptomic sequencing, microbial coinfections in a Guangdong cohort of 23 patients hospitalized with SARS-CoV-2 infection were comprehensively evaluated. Our results revealed distinct differences in microbial composition between mildly and severely ill patients in both the respiratory and gastrointestinal tract. We further demonstrate that *Burkholderia cepacia* complex (BCC) bacteria, *Staphylococcus epidermidis* and *Mycoplasma spp.* were most related to opportunistic pathogens of the respiratory tract in severe cases, possibly exhibiting resistance towards multiple antibiotics, increasing the risk for prolonged intensive care unit (ICU) stay or even associated with increased mortality. By contrast, *Veillonella*, *Neisseria*, *Streptococcus*, and *Prevotella* were identified as the dominant active microbes in the respiratory tract of patients with mild symptoms, similar to what has been reported for healthy adults without infection¹⁴. Our findings demonstrate the value of metatranscriptomics for an unbiased evaluation of the respiratory microbiota associated with

SARS-CoV-2 and provide useful information and suggestions regarding the adequate monitoring and management of the multidrug-resistant (MDR) bacteria in the COVID-19 pandemic.

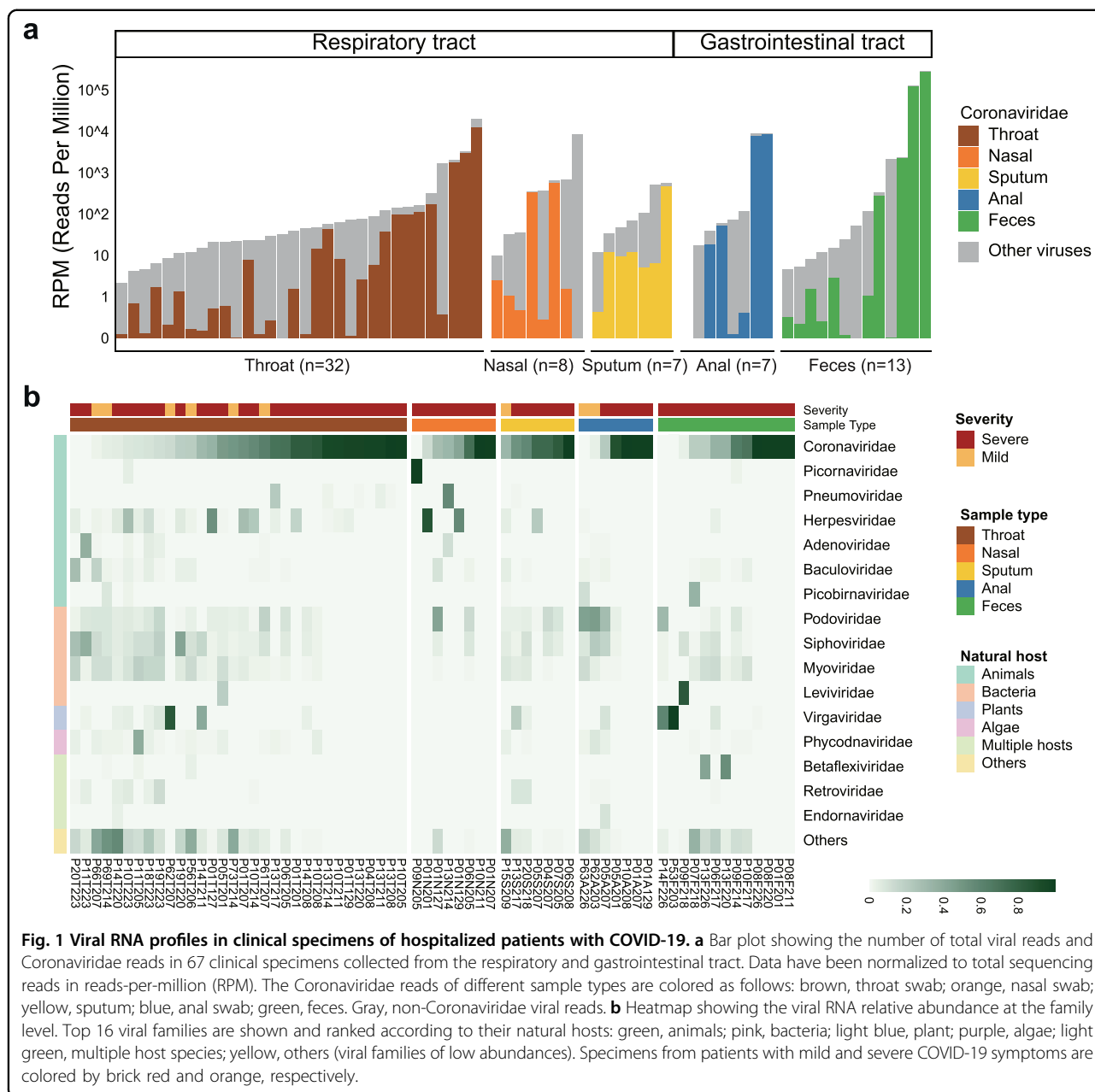
Results

Demographic information of patients and clinical specimens used in the study

Twenty-three patients with COVID-19 hospitalized in the period 10 January–31 March 2020, in four hospitals in the Guangdong Province, China, were enrolled in this study. Fifteen infected patients (41–79 year old) admitted to the ICU and receiving mechanical ventilation were defined as having severe COVID-19, and the remaining eight patients (2–65 year old) were mild cases (Supplementary Table S1). Briefly, 95.7% of the patients (22 out of 23) received antiviral medications. All severe cases received broad-spectrum antibiotics to prevent and control nosocomial infections, and simultaneously 93.35% (14 out of 15) received antifungal agents (Supplementary Table S1). Also, 60% (9 out of 15) of severely ill patients received invasive mechanical ventilation. By contrast, none of the mild cases were treated with antibacterial or antifungal drugs. Up to 31 March, 2020, 53.3% (8 out of 15) of the severe ill patients had been transferred out of ICU or discharged from hospitals, and all mild cases had been discharged, whereas a 79-year-old patient died one month after admission to ICU (P01) (Supplementary Table S2). Sixty-seven serial clinical specimens from the respiratory tract (RT) ($n = 47$, sputum, nasal and throat swab) and gastrointestinal tract (GIT) ($n = 20$, anal swab and feces) of these patients were obtained between 27 January and 26 February 2020 for a comprehensive assessment of microbial characteristics after SARS-CoV-2 infection. A detailed timeline of specimen collections and clinical events for the 23 COVID-19 cases are shown in Supplementary Table S2.

Workflow of ultra-deep metatranscriptomic sequencing

After quality control, an average of 268.3 Gb metatranscriptomic data were generated per sample (Supplementary Table S3). We applied an integrated bioinformatics pipeline to detect human, viral, and nonviral microbial reads in the total RNA-seq data (Supplementary Fig. S1 and Materials and methods). The percentage of human RNA reads (including human rRNA and non-rRNA human transcripts) varied between different types of specimens, constituting a high fraction of total reads among RT specimens (average percentage of 64.56%) and a low fraction among GIT specimens (average percentage of 22.56%) (Supplementary Table S3 and Fig. S1). After removing host data, SortMeRNA was applied¹⁵ to filter microbial rRNA from the metatranscriptomic data. The final remaining nonhuman nonmicrobial rRNA data (ranged from 386 Mb



to 145 Gb) were then used to assess viral and nonviral microbial composition by Kraken2X¹⁶ and MetaPhlan2¹⁷, respectively (Materials and methods). Detailed data statistics for each processing step are provided in Supplementary Table S3.

Co-detection of viruses in clinical specimens of COVID-19 patients

We first assessed the viral composition in RT and GIT specimens. As expected, *Coronaviridae* (mostly contributed by reads assigned to SARS-related coronavirus, Supplementary Table S3) was the most abundant virus,

and was detected in all clinical specimens and varied between 0.01 and 286,418 mapped reads count per million (RPM) (Fig. 1a). Given the presence of confounding factors including days post symptom onset and various treatments on severe cases¹⁸, no comparison of the temporal abundance of SARS-CoV-2-like virus was conducted between mild and severe groups or between types of specimens. Although the SARS-CoV-2 RPM in samples of RT and GIT decreased consistently at later time points of infection (Supplementary Fig. S2a, b), it varied in different severely ill patients. For instance, RT specimens had consistently lower SARS-CoV-2 RPM than GIT

specimens in P01, while specimens from the two sites showed comparable viral levels in P05 and P10 across all sampled time points (Supplementary Fig. S2c).

Besides *Coronaviridae*, RNA-seq analysis also revealed a great diversity of viral composition in clinical samples from infected patients. Natural hosts of the highly abundant viruses differed, including but not limited to animals (e.g., *Picornaviridae*, *Pneumoviridae*, and *Herpesviridae*), bacteria (e.g., *Podoviridae*, *Siphoviridae*, and *Myoviridae*), and plants (*Virgaviridae*) (Fig. 1b and Supplementary Table S4). The co-detection of other high-titer, known human respiratory viruses (genome coverage for representative viral genomes > 50%) was further confirmed in four out of thirteen severely ill patients with metatranscriptomic data of samples from the respiratory tract (30.8%), including human alphaherpesvirus 1 in P01 and P05, rhinovirus B in P09, and human orthopneumovirus in P13 (Fig. 1b and Supplementary Fig. S3a, b). The changes in relative abundance (presented in the unit RPM) of human alphaherpesvirus 1 and human orthopneumovirus in throat samples of P01 and P13 were similar to that of SARS-CoV-2 in the same patient (Supplementary Fig. S3c, d). Although some case studies reported the coinfection of SARS-CoV-2 and influenza viruses^{19,20}, this was not observed in this Guangdong cohort. By contrast, none of these common human respiratory viruses were consistently detected in mild cases without ICU admission (Supplementary Table S4).

Additionally, plant viruses belonging to the family *Virgaviridae*, especially Pepper mild mottle virus (PMMoV) and Tomato mosaic virus (ToMV), were found to be the most dominant viruses in two fecal specimens (P53F203 and P14F226) (Fig. 2b and Supplementary Table S4). Although some pioneer studies have also reported strong evidence supporting the presence of PMMoV and ToMV in human-associated samples^{21–23}, the presence of plant viruses in the fecal samples might also be obtained from food. The extent of possible virus transmission between plants and humans or other vertebrates remains largely unknown.

Characterization of microbial dysbiosis in clinical specimens of COVID-19 patients

We next analyzed the hospital-laboratory-based results as well as the metatranscriptomic sequencing-based nonviral microbial composition to identify key active bacterial/fungal members that might be associated with clinical outcomes in hospitalized patients. Notably, results of cultures and laboratory assays on clinical specimens demonstrated the presence of potential nosocomial fungal ($n = 1$) and bacterial coinfections ($n = 3$) in severely ill COVID-19 patients (Supplementary Table S1). In detail, one patient (P01) tested positive for (1–3)- β -D-glucan (a common component of the fungal cell wall) in blood

samples. Two patients (P04 and P20) had positive sputum cultures for *Burkholderia cepacia* complex (BCC) species, the most common respiratory pathogens causing NP in cystic fibrosis (CF) patients^{24,25}. *S. epidermidis*, a typical skin bacterium that has been increasingly recognized as a MDR nosocomial pathogen^{26,27}, was identified by culturing multiple nasal secretions of one patient (P06).

Next, the nonviral RNA data of all 67 clinical specimens were analyzed to fully assess the active microbial composition using MetaPhlan2. As none of the mild cases were admitted to ICU or received antibacterial/antifungal agents, we compared the RT specimens between mild ($n = 7$) and severe ($n = 13$) cases to examine the microbial dysbiosis in patients exhibiting differential disease severity. Remarkable differences in RT microbial richness (number of detected taxa) and composition between mild and severe cases were observed (Fig. 2). All RT specimens (including six throat swabs and one sputum) of mildly ill patients who were admitted to three hospitals (located in Guangzhou, Yangjiang and Qingyuan, Supplementary Table S1) consistently exhibited a larger number of detected microbial taxa (Fig. 2a and Supplementary Fig. S4a) and similar microbial RNA community compositions (Fig. 2b and Supplementary Fig. S4b). Notably, the number of respiratory microbial taxa (at the genus and species level) detected in mild cases was significantly higher than that in severe cases (Supplementary Fig. S4c, $P < 0.001$, Wilcoxon rank-sum test). On average, 28 genera and 48 species were detected per RT sample in mild cases, while only six genera and four species were detected per RT sample in severe cases (Supplementary Fig. S4c and Table S7), which might reflect the pronounced effects of the administration of broad-spectrum antimicrobial agents in these severely ill patients. The predominant RT bacteria in mild cases were *Veillonella*, *Neisseria*, *Streptococcus*, and *Prevotella* (occurrence > 80% individuals and mean relative abundance > 5%) (Fig. 2b, c), which is similar to microbial communities commonly reported in the nasal and oral cavity of healthy human adults^{14,28}. However, except for *Veillonella*, each of the latter three genera enriched in mild cases was only detected in few severe cases ($n \leq 3$, Fig. 2a). The four genera associated with mild cases also showed significantly higher mean abundance in mild than in severe cases (Fig. 2c, $P < 0.05$, Wilcoxon rank-sum test).

Of note, several prevalent RT microbial features in severe cases were identified to be patient-specific. Among 40 respiratory samples from severe patients, over 60% were mono-dominated (relative abundance > 60%, as suggested by Hildebrandt et al.²⁹) by the bacterial genus *Burkholderia* (11 samples from P01, P04, and P20), *Staphylococcus* (6 samples from P10 and P19) or *Mycoplasma* (7 samples from P05, P06, P14, and P18) (Fig. 2b, c and Supplementary Table S2). Each genus was detected

experimental procedures were performed simultaneously at the same time on all clinical specimens from both mild and severe cases to minimize the batch effects and possible microbial contaminations. By mapping RNA reads to the reference genomes of BCC species (Materials and methods), we further confirmed the predominant expression of *B. cenocepacia* in the respiratory tract of P01, and *B. multivorans* in P04 and P20, who also provided positive sputum cultures of BCC (Supplementary Fig. S5a and Table S8). All *Staphylococcus* RNA reads of RT samples from P06 (who also provided positive *S. epidermidis* culture), P10 and P19 were assigned to *S. epidermidis* (Fig. 2b and Supplementary Fig. S5b). However, *S. aureus*, a major hospital-acquired pathogen³⁰, was not detected in metatranscriptomic data of any sequenced RT samples. *Mycoplasma orale* and *M. hominis*, rather than *M. pneumoniae*, were the two dominating *Mycoplasma* members (Fig. 2b and Supplementary Fig. S5c). *Propionibacterium* and *Escherichia* were also frequently detected in RT samples from severe cases (occurrence > 80% individuals) but were less abundant than the former three genera (mean relative abundance < 3%) (Fig. 2c). Moreover, all the five prevalent genera in severe cases have been reported to be antibiotic-resistant bacteria and/or associated with nosocomial infections, while they were not detected or present in extremely low abundance in mild cases (relative abundance < 0.15%) (Fig. 2a, c and Supplementary Table S6).

Consistently, we detected both positive results for the blood (1–3)- β -D-glucan levels (Supplementary Table S9) as well as GIT expression of ascomycetic transcripts (mainly from the genus *Saccharomycetaceae*) in the P01 (Supplementary Table S6 and Fig. 3a), who died one month after ICU admission. Ascomycetic transcripts (mainly from *Debaryomycetacea*) were also identified in all five throat swabs collected from P13 (Fig. 2b). Interestingly, the ascomycetes constituted only 5.7% of the total nonviral microbes at the first time point (8 February) and increased to 18.2%–87.4% at later points (11–20 February) (Fig. 2b). P13, a 79-year-old man, had onset of COVID-19 symptoms on 30 January and had been admitted to the ICU since 5 February (Supplementary Table S1). Although the patient received daily antimicrobial treatment with a combination of antibiotic (meropenem, targocid, polymyxin b sulfate, amikacin, or sulperazone), antifungal (candidas and/or amphotericin B) and antiviral drugs (ribavirin) during the entire sampling period (6 February to 26 February), these observations suggested that a rapid succession from bacteria to fungi had occurred in the microbiota of the respiratory tract in P13 three days after ICU admission. These findings collectively indicated that serial monitoring to track respiratory fungi and possibly secondary fungal infections is required to avoid delayed treatment for such patients.

In addition, three severe cases (P05, P07, and P11), despite receiving multi-agent antimicrobial therapy (Supplementary Table S1), had high expression levels of *Veillonella* but low levels (or no detection) of the above potential opportunistic pathogens (or pathogens) in their RT samples (Fig. 2a, b). Most well-known respiratory bacterial pathogens, as well as potential high-abundant pathogenic candidates we identified/isolated in clinical specimens of COVID-19 patients with severe symptoms, are aerobic or facultative organisms. In contrast, *Veillonella spp.* are strictly anaerobic and have been reported to be part of normal oral cavities and rarely isolated in nosocomial infections^{31,32}. Of note, all three patients (P05, P07, and P11) had been transferred out of the ICU or discharged from hospitals and none of them received invasive ventilation during ICU admission (Supplementary Table S1).

In addition, severe cases also appeared to have a distinct gut metatranscriptome compared to mild cases. GIT specimens (anal swabs) from the two mild patients with no antimicrobial treatments consistently showed high abundances of Proteobacteria (e.g., *Campylobacter*) and *Streptococcus* in their gut metatranscriptome (Supplementary Fig. S6a). Similarly, a recent study using 16S rRNA gene-based amplicon sequencing (fecal samples) also reported a significantly higher relative abundance of *Streptococcus* in antibiotic treatment-naïve COVID-19 patients than in age-, sex-, and body mass index-matched healthy controls³³. Our mild case-related GIT microbial transcripts using anal swabs also differed from the previous metatranscriptomic study of healthy adults using fecal samples, whose gut microbiota was dominated by Firmicutes and Bacteroidetes³⁴. Longitudinal samples from both patients and healthy controls are needed to characterize the COVID-19-related gut microbial dysbiosis using the same sampling sites and sequencing strategy.

On the other hand, *Parabacteroides* constituted one of the major active GIT bacteria of the severe cases contrasting mild cases. For instance, *Parabacteroides* (including *P. distasonis* and *P. merdae*) mono-dominated the gut microbial transcripts in five fecal samples from four severe cases (relative abundance > 60%, P07, P09, P10, and P13) (Supplementary Fig. S6b–d). The genus also displayed a relatively high abundance in several other fecal samples and anal swabs of severe cases (relative abundance > 20%) (Supplementary Table S8). Interestingly, an extreme bloom of *P. distasonis*, a low-abundant but common taxa in the human gut, has been reported after beta-lactam ceftriaxone treatment²⁹. Thus, metatranscriptomic findings have not only complemented and enhanced the laboratory-based detection of candidate pathogens but also provided comprehensive information on microbial dysbiosis in COVID-19 patients.

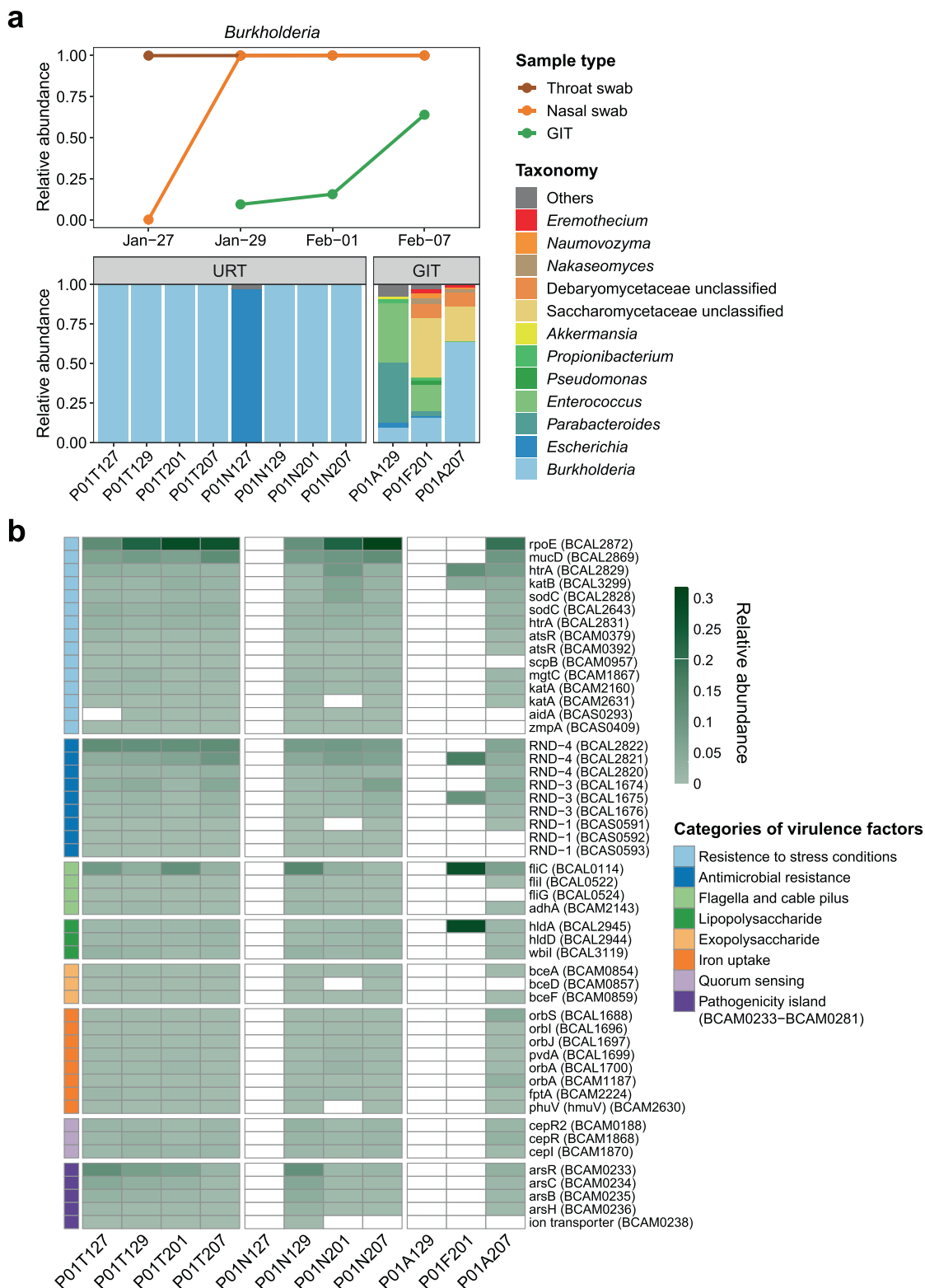


Fig. 3 (See legend on next page.)

(see figure on previous page)

Fig. 3 Identification of the potential secondary *B. cenocepacia* infection in P01. **a** Bar plot showing the relative expression levels of nonviral microbes in all specimens from P01. A timeline chart showing the corresponding changes in abundance of *B. cenocepacia* transcripts from 27 January to 07 February 2020 in P01. Brown indicates throat swabs; orange indicates nasal swabs and green indicates GIT samples including anal swabs and feces. A total of 11 specimens from the respiratory tract (RT) and gastrointestinal tract (GIT) are shown. Orange, fungi; blue, bacteria. **b** Heatmap showing the relative expression levels of virulence factors of *B. cenocepacia*. A total of 50 identified virulence genes are shown and ranked according to their functional categories: light blue, resistance to stress conditions; blue, antimicrobial resistance; light green, flagella and cable pilus; green, lipopolysaccharide; pink, exopolysaccharide; orange, iron uptake; light purple, quorum sensing; purple, genes located in a pathogenicity island.

Identification of secondary coinfection with *B. cenocepacia*

In order to retrospectively investigate possible non-COVID-19 risk factors associated with the in-hospital death of patient P01, serial clinical specimens were collected and analyzed. Notably, a distinct co-detection of transcripts belonging to *B. cenocepacia* was clearly observed in a time-dependent manner in this patient. On the first day of sampling (27 January 2020), up to 99.9% of nonviral microbial transcripts in his throat swab were assigned to *B. cenocepacia* (P01T127), while most of the transcripts in his nasal swab collected on the same day were from *Escherichia* (mainly from *E. coli*) (P01N127) (Fig. 3a). *B. cenocepacia* was subsequently predominantly present in both throat and nasal swabs for all the following sampling time points (29 January–07 February 2020) (Fig. 3a). In addition, *B. cenocepacia* was detected in all GIT samples from this patient, and the percentage of *B. cenocepacia* to nonviral transcripts gradually increased from 9.5% to 64% (from 29 January to 7 February) (Fig. 3a). The time-dependent dynamics of transcript levels of *B. cenocepacia* suggested that transfer from the upper respiratory tract to the lower gastrointestinal tract had caused a secondary systemic infection in P01. Our findings were also consistent with the patient's death certificate record, indicating bacteremic sepsis as one of the leading causes of his death (Supplementary Table S1). Indeed, several retrospective studies have pointed out that among BCC-infected CF patients, infection with *B. cenocepacia*, rather than other commonly isolated BCC members (such as *B. multivorans* and *B. cepacia*), constituted the highest risk factor of death^{25,35}.

Next, virulence factors (VF) expressed by *B. cenocepacia* in P01 were analyzed to understand better the pathogenic mechanisms of this possible lethal pathogen in this severe COVID-19 case (Materials and methods). The gene *rpoE* (a member of the extracytoplasmic function subfamily of sigma factors) was the most abundantly expressed VF during the entire sampling period in P01 with SARS-CoV-2 infection (Fig. 3b). RpoE, as a stress response regulator, has been demonstrated to be essential for the growth of *B. cenocepacia* and the delay of phagolysosomal fusion in macrophages during infection³⁶. A delay in phagolysosomal fusion has also been an important host immune escape strategy for

several bacterial pathogens³⁷. Other VFs in response to oxidative stress conditions in the host environment, such as those encoding superoxide dismutase, peroxidase or catalase (*sodC* and *katB*) were also expressed (Fig. 3b and Supplementary Table S9). A panel of genes belonging to resistance-nodulation-division (RND) family transporters that confer multidrug resistance to *B. cenocepacia*³⁸ were also highly expressed in all types of specimens. We also detected expressions of genes encoding flagella and cable pilus (*fliC*, *fliI*, *fliG*, and *adhA*), which can facilitate the bacterial adhesion to host cells and mucin³⁹ (Fig. 3b). In addition, 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 (Fig. 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 combined with clinical laboratory diagnosis, including cultures and colorimetric assays, identified key characteristics of the microbial dysbiosis associated with hospitalized patients infected with SARS-CoV-2. The most prevalent respiratory bacteria in our severely ill COVID-19 patients were BCC bacteria, *S. epidermidis* and *Mycoplasma spp.* (including *M. hominis* and *M. orale*). These organisms are distinct from prior results on pathogenic bacteria identified in previous coronavirus outbreak and influenza pandemics (e.g., *S. pneumoniae*, *S. aureus*, *K. pneumoniae*, and *M. pneumoniae*)^{4–8} or those associated with ICU-acquired bloodstream infections (BSIs) (*Acinetobacter baumannii*, *K. pneumoniae*, *Enterococcus spp.*, *Candida albicans*, and *C. parapsilosis*) in a cohort of 50 severely ill COVID-19 patients in Athens, Greece⁴⁰. A strong confounding factor in this study is the markedly different treatment instigated for the mild and severe cases, where multiple and broad-spectrum antimicrobial agents were given to severe cases but not mild cases. Another limitation, however, is the relatively small sample size of patients ($n = 23$) as well as the highly biased number of specimens from severe and mild patients (58 vs 9). Also, uninfected controls were not

enrolled in the study, though the microbial composition of the respiratory and gastrointestinal tract has been reported in healthy individuals^{14,28,34}. All these inherent limitations, therefore, prevented our microbial observations in treatment-naïve mildly ill patients from being associated with COVID-19 infection, but provided important information of microbial dysbiosis and related multidrug resistance in severely ill patients receiving antimicrobial treatments.

In particular, respiratory *BCC* mono-dominated 23.1% of severe cases (relative abundance > 60%), showing co-detection evidence from both laboratory cultures and metatranscriptomic results in P04 and P20. The serial metatranscriptomic data of all specimens from P01 revealed the timeline of secondary nosocomial 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 the production of LPS and EPS), and resist the effects of multiple antibiotics (RND family), which eventually may have led to the life-threatening bacteremic sepsis of the patient. In addition to our study, two recent studies also reported nosocomial BCC-associated deaths of COVID-19 patients^{41,42}. However, all three studies are single case reports and could hardly be used to draw inferences about an undetected BCC-associated nosocomial outbreak whose primary sources have been identified to be contaminated medical products/devices (mainly the disinfectant products)⁴³. BCC bacteria, a major threat to hospitalized CF patients, are predominantly localized in the phagocytes and mucus in the CF lung⁴⁴ and might accelerate the decline in pulmonary function of COVID-19 patients. Although with a low reported incidence, given the lethal outcomes in relation to the BCC coinfections, management strategies should be developed for these hospitals to track the potentially infectious source of various medical products, the frequency of the infections occurred, and finally, to prevent nosocomial BCC-coinfections in hospitalized COVID-19 patients.

Except for *B. cenocepacia*, the other prevalent respiratory bacteria found in our severely ill patients (*B. multivorans*, *M. hominis*, *M. orale* and *S. epidermidis*) usually cause mild or no symptoms; however, they are reported to be widespread in hospital environments and may act as reservoirs for antibiotic resistance genes (ARGs)^{45–47}. *Mycoplasma spp.* lack a cell wall and have inherent resistance to commonly administered beta-lactam antibiotics. Several studies further indicated escalating antibiotic resistance levels in mycoplasmas^{45,46}, including macrolide, tetracycline, or fluoroquinolone classes of

antibiotics. Super-high expression levels of *M. orale* genes were found in the RT of two severe ill patients (P14 and P18) with prolonged ICU stay (> 30 days), though both received antimicrobial therapy during this period. Similarly, among coagulase-negative staphylococci, *S. epidermidis* has been reported to cause the greatest number of nosocomial infections, particularly the chronic infections associated with indwelling medical devices⁴⁷. Although *S. epidermidis* does not usually produce aggressive toxins, several hospital-adapted lineages have increasingly acquired clinically relevant resistance determinants and formed an unignorable challenge among nosocomial infections²⁷. Additional evidence has demonstrated the horizontal transfer of ARGs, including methicillin resistance, from *S. epidermidis* to *S. aureus*^{48–50}, a known formidable, virulent nosocomial pathogen. Our findings also agree with two independent studies that consistently reported a high incidence of ICU-acquired BSIs (51.2%–54%) in severely ill COVID-19 patients, mostly due to MDR pathogens^{40,51}.

Currently, no effective drugs have been licensed for human use against SARS-CoV-2 infection. Instead, the widespread use of antimicrobial agents (including broad-spectrum antibiotics) has been documented in many studies (including ours)^{9,12} to prevent and treat possible secondary infections in COVID-19 patients, especially in those who needed mechanical ventilation. Without clinical specimens before any treatments and specimens from non-infected healthy controls, it is difficult to distinguish to what extent differences in the respiratory microbial patterns associated with SARS-CoV-2 infections reflect the disease or the antimicrobial treatment, or both. Still, the distinct bacterial communities as well as their dramatic and rapid shifts in the respiratory and gastrointestinal tract of the severely ill patients might be related to not only the excessive antimicrobial therapy in many cases but also the significant disruption of the normal human microbiota caused by the antimicrobial therapy allowing colonization by MDR organisms including opportunistic pathogens. Only a single patient died from secondary bacterial sepsis in our study, which recommends narrower and more targeted treatment therapy for secondary or coinfections with other organisms. Stepwise strategies are needed to monitor hospital/ICU acquired MDR organisms, control the spread of ARGs, optimize antimicrobial therapy for COVID-19 patients, and prevent potential future threats from a blooming reservoir of MDR organisms after the global pandemic, by using rapid diagnostic technologies such as antigen and antibody testing (e.g., immunofluorescence assays and enzyme-linked immunosorbent assays) and nucleic acid-based molecular testing (e.g., polymerase chain reaction and DNA microarray)⁵².

Materials and methods

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

Twenty-three patients admitted to hospital in January 10–March 31, 2020, with confirmed SARS-CoV-2 infection based on a positive SARS-CoV-2 test were included. A total of 67 clinical specimens from the respiratory and gastrointestinal tract (including throat swab, nasal swab, sputum, anal swab and feces) were collected from the above patients at the First Affiliated Hospital of Guangzhou Medical University (thirteen patients), the Fifth Affiliated Hospital of Sun Yat-sen University (two patients), Yangjiang People's Hospital (five patients) and Qingyuan People's Hospital (three patients) between 27 January and 26 February 2020. All specimens were stored at -80°C before nucleic acid extraction. Patients were classified into mild ($n = 8$, without intensive care unit admission) and severe ($n = 15$, with ICU admission) cases based on their severity of SARS-CoV-2 symptoms. Detailed de-identified information for patients and clinical specimens are presented in Supplementary Tables S1 and S2, respectively.

Laboratory diagnosis of nosocomial bacterial and fungal infections

All severely ill COVID-19 patients admitted to the ICU for more than 48 h were monitored for nosocomial infections, which were defined according to the definitions of the US Centers for Disease Control and Prevention⁵³. Culture of sputum and nasal secretions was conducted according to standard protocols for diagnosing nosocomial microbial infections (hospital-acquired and ventilator-associated) in all severely ill COVID-19 patients with ICU admission⁵⁴. Blood samples of patients (P01 and P05) hospitalized in the Fifth Affiliated Hospital of Sun Yat-sen University were collected for colorimetric assay-based (1–3)- β -D-glucan test to diagnose fungal infection (Dynamiker Biotechnology, Tianjin, China, Catalog number: DNK-1401-1).

RNA extraction, metatranscriptomic library preparation, and sequencing

For each clinical sample, total RNA was extracted (QiAamp RNeasy Mini Kit, Qiagen, Germany). DNA from human and microbes was then removed from RNA using DNase I and the concentration was quantified (Qubit RNA HS Assay Kit, Thermo Fisher Scientific, Waltham, MA, USA). Purified RNA samples were regularly shipped on dry ice to BGI-Shenzhen and subjected to preprocessing, DNA nanoball-based library construction, and high-throughput metatranscriptomic sequencing on the DNBSEQ-T7 platform (100 nt paired-end reads, MGI, Shenzhen, China)⁵⁵. Three negative controls (NCs) from nuclease-free water were prepared for library construction

and metatranscriptomic sequencing in parallel with the clinical samples as described previously in our pilot study⁵⁵ while all the NCs failed to yield any sequencing data due to relatively low biomass.

Identification and removal of human RNA reads from metatranscriptomic data

For each sample, the raw metatranscriptomic reads were processed using Fastp (v0.19.5, default settings)⁵⁶ to filter low-quality data and adapter contaminations and generate the clean reads for further analyses. Human-derived reads were identified with the following steps: (1) identification of human ribosomal RNA (rRNA) by aligning clean reads to human rRNA sequences (28S, 18S, 5.8S, 45S, 5S, U5 small nuclear RNA, as well as mitochondrial mt12S) using BWA-MEM (0.7.17-r1188)⁵⁷; (2) identification of human transcripts by mapping reads to the hg19 reference genome using the RNA-seq aligner HISAT2 (version 2.1.0, default settings)⁵⁸; and (3) a second-round identification of human reads by aligning remaining reads to hg38 using Kraken2 (version 2.0.8-beta, default settings)¹⁶. All human RNA reads were then removed to generate qualified nonhuman RNA-seq data. The number of human RNA-seq reads identified at each step is presented in Supplementary Table S3.

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

Before the identification of virome and microbiota, SortMeRNA version 4.2.0¹⁵ (default settings) was applied to filter microbial rRNA (28S, 18S, 5.8S; and 23S, 16, and 5S rRNA from the SILVA database) from nonhuman metatranscriptomic data. Given the nucleotide-based methods exhibit lower accuracy on the viral read, we used Kraken2X which uses translated search against a protein database for viral classification¹⁶. The remaining nonhuman non-microbial rRNA reads were processed by Kraken2X v2.0.8 beta (default parameters)¹⁶ with a self-built viral protein database by extracting protein sequences from all complete viral genomes deposited in the NCBI RefSeq database (8872 genomes downloaded on 1 March 2020 including the SARS-CoV-2 complete genome reference sequence, GCF_009858895.2). The number of reads annotated to each viral family was summarized based on the read alignment results of Kraken2X. By ranking the number of *Coronaviridae* reads at the species level, we found that most of the *Coronaviridae* reads were annotated to SARS-related Coronavirus (Spearman's $\rho > 0.996$ between the number of RNA reads annotated to *Coronaviridae* and that annotated to SARS-related coronavirus) whereas only a tiny fraction of RNA reads (median number = 4) mapped to the common human coronaviruses (Human coronavirus NL63, 229E, and HKU1) (Supplementary Table S3). The remaining

Coronaviridae reads were also mostly mapped to reference genomes of bat coronaviruses, which might result from a misclassification of SARS-related reads, as they are the closest relatives of SARS-CoV-2 (Supplementary Table S3). Based on the above observation, the Kraken2X-annotated *Coronaviridae* reads were considered as SARS-CoV-2-like reads in this study. For each sample, the ratio of SARS-CoV-2-like reads to total clean reads and the ratio of SARS-CoV-2-like reads to total viral reads were calculated accordingly (Supplementary Table S4).

After ranking the aligned reads of all detected viral species in each sample, highly abundant non-*Coronaviridae* viral species (> 10,000 aligned RNA reads per species) were identified (Supplementary Table S5) and selected for robust co-detection with known respiratory viruses in hospitalized patients with SARS-CoV-2 infection. Three human respiratory viral species (human alphaherpesvirus 1, human orthopneumovirus and rhinovirus B) co-detected in respiratory samples from four severe cases (P01, P05, P09, and P13) met the above criterion. Representative genomes of each species, including human herpesvirus 1 strain 17 (NC_001806.2), human orthopneumovirus subgroup A (NC_038235.1), and rhinovirus B isolate 3039 (KF958308.1) were downloaded from NCBI. One representative sample (P01N201, P05S207, P09N205, and P12T211) with the highest number of reads assigned to the targeted species was used for coverage analysis for each patient. Reads assigned to a given species were aligned against the corresponding reference genome by bowtie2 v2.3.0 (the '-sensitive' mode, local alignment)⁵⁹. Sequencing depth and genome coverage of each reference genome were determined with BEDTools v2.27.1 (genomecov -ibam sort.bam -bg)⁶⁰. Robust co-detection with known respiratory viruses was defined when > 50% of the genome was covered. Considering the reported relatively low classification accuracy at the species level using the Kmer-based classification algorithm of Kraken, we did not focus on the possible low-titer, human viral species with low coverage and a small number of mapped reads.

Characterization of nonviral microbial communities in hospitalized patients with SARS-CoV-2 infection

MetaPhlan2, a clade-specific marker gene-based alignment method, requires considerable genome coverage and depth to detect marker genes and the presence of a given microbial taxon¹⁷. To minimize the detection of potential low-level microbial contaminants from environmental sources (e.g., reagents and kits), we used non-human non-rRNA reads (input RNA data) and MetaPhlan2 (version 2.7.0) (default parameter options, except for -ignore-viruses)¹⁷. Importantly, the default parameter "stat-q" was set as 0.1¹⁷ and it excluded the 10% of markers with the highest and the 10% with the

lowest abundance for calculating the robust abundance of a given taxon. The MetaPhlan2-based relative abundance profiles of nonviral microbial taxa at the genus level were presented in Supplementary Table S6. The robust presence-absence profiles of respiratory microbial taxa at both genus and species levels are presented in Supplementary Table S7. Mono-dominance of a given microbial taxon (genus or species) was defined if a taxon had a relative abundance > 60% in one sample as suggested by a previous study²⁹.

Most of the RNA reads of the two predominant bacterial genera *Burkholderia* and *Parabacteroides*, respectively, identified in the respiratory and gastrointestinal tract of severe cases could hardly be assigned to species level by MetaPhlan2, which might reflect that the two genera contain closely related species that are difficult to differentiate by marker genes. In order to determine which species and how abundant species were in samples mono-dominated by *Burkholderia* or *Parabacteroides*, we downloaded four reference genome sequences of the most frequently isolated BCC species (*B. cenocepacia* J2315, *B. multivorans* ATCC BAA-247, *B. cepacia* ATCC 25416 and *B. dolosa* AU0158) and two gut *Parabacteroides* species (*P. distasonis* ATCC 8503 and *P. merdae* NCTC13052). For each sample, reads were mapped against corresponding references by bowtie2 v2.3.0, and the sequencing depth and genome coverage were estimated by BEDTools v2.27.1 as described above. The summary of coverage and depth of reference genomes for selected samples are presented in Supplementary Table S8. Prevalent SARS-CoV-2-associated respiratory bacteria or fungi were considered to be present if the respiratory specimens (at least one sample) of patients were mono-dominated (relative abundance > 60%) by a given taxon (metatranscriptomic sequencing) (Supplementary Table S9).

Identification of expressed VF in *B. cenocepacia*

To identify the presence and expression patterns of potential VF in *B. cenocepacia* identified in P01, we collected multiple functional categories of virulence genes previously studied and verified by gene mutation analysis in *B. cenocepacia* strains as well as corresponding gene ID in the annotated J2315 genome³⁵, including (1) resistance to stress conditions, (2) antimicrobial resistance, (3) quorum sensing, (4) iron uptake, (5) flagella and cable pilus, (6) LPS, and (7) EPS. In addition, a pathogenicity island identified on chromosome 2 (BCAM0233-BCAM0281) of *B. cenocepacia* J2315 by using comparative genomics was included⁶¹. Nonhuman non-rRNA microbial reads of all samples from P01 were mapped against the reference genome of *B. cenocepacia* J2315 using bowtie2 v2.3.0 as described above and identified by the gene IDs in the J2315 genome. For each sample, only

virulence genes with more than 10 mapped reads were retained. A total of 50 expressed virulence genes were identified in clinical samples collected from P01 and presented in Supplementary Table S10.

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

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

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 number that mapped to gene i .

Statistical analyses

Non-metric multidimensional scaling ordination of respiratory microbial community was conducted using the Manhattan distances based on a presence/absence matrix of genus profiles of 47 respiratory specimens (7 from mild cases and 40 from severe cases) (R version 3.6.1, vegan package). Wilcoxon rank-sum tests were performed to compare differences in the (mean) value of bacterial relative abundance at the genus level between the overall respiratory specimens of mild and severe cases, respectively (R version 3.6.1, coin package).

Acknowledgements

This work was funded by the National Major Project for Control and Prevention of Infectious Disease in China (2018ZX10301101-004), the emergency grants for prevention and control of SARS-CoV-2 of Ministry of Science and Technology (2020YFC0841400) and Guangdong province (2020B1111320003, 2020B111107001, 2020B111108001, and 2018B020207013), Guangdong Provincial Key Laboratory of Genome Read and Write (2017B030301011), Guangdong Provincial Academician Workstation of BGI Synthetic Genomics (2017B090904014), and Shenzhen Engineering Laboratory for Innovative Molecular Diagnostics (DRC-SZ[2016]884). Metatranscriptomic sequencing of this work was supported by China National GeneBank.

Author details

¹BGI-Shenzhen, Shenzhen 518083, China. ²Laboratory of Genomics and Molecular Biomedicine, Department of Biology, University of Copenhagen, 2100 Copenhagen, Denmark. ³State Key Laboratory of Respiratory Disease, National Clinical Research Center for Respiratory Disease, Guangzhou Institute of Respiratory Health, the First Affiliated Hospital of Guangzhou Medical University, Guangzhou, Guangdong 510120, China. ⁴Institute of Infectious disease, Guangzhou Eighth People's Hospital of Guangzhou Medical University, Guangzhou, Guangdong 510060, China. ⁵Guangzhou Customs District Technology Center, Guangzhou 510700, China. ⁶Department of Infectious Diseases, Guangdong Provincial Key Laboratory of Biomedical Imaging, Guangdong Provincial Engineering Research Center of Molecular Imaging, The Fifth Affiliated Hospital, Sun Yat-sen University, Zhuhai, Guangdong 519000, China. ⁷School of Future Technology, University of Chinese Academy of Sciences, Beijing 101408, China. ⁸Shenzhen Key Laboratory of Unknown Pathogen Identification, BGI-Shenzhen, Shenzhen 518083, China. ⁹The Sixth Affiliated Hospital of Guangzhou Medical University, Qingyuan People's Hospital, Qingyuan, Guangdong, China. ¹⁰Yangjiang People's Hospital, Yangjiang, Guangdong, China. ¹¹State Key Laboratory of Bioelectronics, School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China. ¹²Guangdong Provincial Key Laboratory of Human Disease Genomics, Shenzhen Key Laboratory of Genomics, BGI-Shenzhen, Shenzhen

518083, China. ¹³BGI Education Center, University of Chinese Academy of Sciences, Shenzhen 518083, China. ¹⁴Shenzhen Engineering Laboratory for Innovative Molecular Diagnostics, BGI-Shenzhen, Shenzhen 518120, China. ¹⁵James D. Watson Institute of Genome Science, Hangzhou 310008, China. ¹⁶Guangdong Provincial Academician Workstation of BGI Synthetic Genomics, BGI-Shenzhen, Shenzhen 518120, China. ¹⁷HKU-Pasteur Research Pole, School of Public Health, Li Ka Shing Faculty of Medicine, The University of Hong Kong, Hong Kong SAR, China. ¹⁸BGI PathoGenesis Pharmaceutical Technology Co., Ltd., BGI-Shenzhen, Shenzhen 518083, China. ¹⁹Guangdong Provincial Key Laboratory of Genome Read and Write, BGI-Shenzhen, Shenzhen 518120, China. ²⁰School of Biology and Biological Engineering, South China University of Technology, Guangzhou, China

Author contributions

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 specimens and clinical information from patients, processed of RNA extraction, and performed in-hospital diagnostic tests. Ji L. and P.R. performed metatranscriptomic library construction. H.Z., Z.S., H.R., F. Yang., T.L., J.Z., and F.L. processed the 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 version of the paper. J.Z., K.K., and H.M.T. contributed substantially to the revisions of the paper. All other authors participated in discussions, provided useful comments and revised the paper. All authors approved the final version.

Data availability

The nonhuman metatranscriptomic sequencing data from 67 clinical specimens that support the findings of this study have been deposited into CNSA (CNGB Nucleotide Sequence Archive) of CNGBdb with accession number [CNP0001066](https://doi.org/10.55577/CNP0001066). Details of software, code, and parameters used for data analyses of the current study are provided and publicly available from [GitHub](https://github.com).

Ethics approval and consent to participate

Informed consent was obtained from all enrolled patients from the First Affiliated Hospital of Guangzhou Medical University, the Fifth Affiliated Hospital of Sun Yat-Sen University, Yangjiang People's Hospital and Qingyuan People's Hospital. The overall study was reviewed and approved by the ethics committees of all the four hospitals. The institutional review board of BGI-Shenzhen approved the sequencing and downstream analyses of samples collected by the aforementioned hospitals under the ethical clearance no. BGHRB 20008.

Conflict of interest

The authors declare no competing interests.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41421-021-00257-2>.

Received: 30 September 2020 Accepted: 26 February 2021

Published online: 13 April 2021

References

- World Health Organisation. COVID-19 Weekly Epidemiological Update. 2 February 2021. <https://www.who.int/publications/m/item/weekly-epidemiological-update-2-february-2021>.
- Short, K. R., Kroeze, E. J. B. V., Fouchier, R. A. M. & Kuiken, T. Pathogenesis of influenza-induced acute respiratory distress syndrome. *Lancet Infect. Dis.* **14**, 57–69 (2014).
- De Wit, E., Van Doremalen, N., Falzarano, D. & Munster, V. J. SARS and MERS: recent insights into emerging coronaviruses. *Nat. Rev. Microbiol.* **14**, 523–534 (2016).
- Peiris, J. S. M. et al. Clinical progression and viral load in a community outbreak of coronavirus-associated SARS pneumonia: a prospective study. *Lancet* **361**, 1767–1772 (2003).

5. Morris, D. E., Cleary, D. W. & Clarke, S. C. Secondary bacterial infections associated with influenza pandemics. *Front. Microbiol.* **8**, 1041 (2017).
6. Centers for Disease Control and Prevention (CDC). Bacterial coinfections in lung tissue specimens from fatal cases of 2009 pandemic influenza A (H1N1) —United States, May–August 2009. *Morb. Mortal. Wkly. Rep.* **58**, 1071–1074 (2009).
7. Gill, J. R. et al. Pulmonary pathologic findings of fatal 2009 Pandemic influenza A/H1N1 viral infections. *Arch. Pathol. Lab. Med.* **134**, 235–243 (2010).
8. Weinberger, D. M. et al. Impact of the 2009 influenza pandemic on pneumococcal pneumonia hospitalizations in the United States. *J. Infect. Dis.* **205**, 458–465 (2012).
9. Chen, N. et al. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. *Lancet* **395**, 507–513 (2020).
10. Meng, L. et al. Intubation and ventilation amid the COVID-19 outbreak: Wuhan's experience. *Anesthesiology* **132**, 1317–1332 (2020).
11. Bonten, M. J. M., Kollef, M. H. & Hall, J. B. Risk factors for ventilator-associated pneumonia: from epidemiology to patient management. *Clin. Infect. Dis.* **38**, 1141–1149 (2004).
12. Zhou, F. et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *Lancet* **395**, 1054–1062 (2020).
13. Kim, D., Quinn, J., Pinsky, B., Shah, N. H. & Brown, I. Rates of co-infection between SARS-CoV-2 and other respiratory pathogens. *J. Am. Med. Assoc.* **323**, 2085–2086 (2020).
14. Man, W. H., De Steenhuijsen Piters, W. A. A. & Bogaert, D. The microbiota of the respiratory tract: gatekeeper to respiratory health. *Nat. Rev. Microbiol.* **15**, 259–270 (2017).
15. Kopylova, E., Noé, L. & Touzet, H. SortMeRNA: fast and accurate filtering of ribosomal RNAs in metatranscriptomic data. *Bioinformatics* **28**, 3211–3217 (2012).
16. Wood, D. E., Lu, J. & Langmead, B. Improved metagenomic analysis with Kraken 2. *Genome Biol.* **20**, 257 (2019).
17. Truong, D. T. et al. MetaPhlan2 for enhanced metagenomic taxonomic profiling. *Nat. Methods* **12**, 902–903 (2015).
18. Zheng, S. et al. Viral load dynamics and disease severity in patients infected with SARS-CoV-2 in Zhejiang province, China, January–March 2020: retrospective cohort study. *BMJ* **369**, m1443 (2020).
19. Azekawa, S., Namkoong, H., Mitamura, K., Kawaoka, Y. & Saito, F. Co-infection with SARS-CoV-2 and influenza A virus. *IDCases* **20**, e00775 (2020).
20. Wu, X. et al. Co-infection with SARS-CoV-2 and influenza A virus in patient with pneumonia, China. *Emerg. Infect. Dis.* **26**, 1324–1326 (2020).
21. Zhang, T. et al. RNA viral community in human feces: prevalence of plant pathogenic viruses. *PLoS Biol.* **4**, e3 (2006).
22. Liu, R., Vaishnav, R. A., Roberts, A. M. & Friedland, R. P. Humans have antibodies against a plant virus: evidence from tobacco mosaic virus. *PLoS ONE* **8**, e60621 (2013).
23. Colson, P. et al. Pepper mild mottle virus, a plant virus associated with specific immune responses, fever, abdominal pains, and pruritus in humans. *PLoS ONE* **5**, e10041 (2010).
24. Courtney, J. M. et al. Clinical outcome of *Burkholderia cepacia* complex infection in cystic fibrosis adults. *J. Cyst. Fibros.* **3**, 93–98 (2004).
25. Zlosnik, J. E. A. et al. *Burkholderia* species infections in patients with cystic fibrosis in British Columbia, Canada: 30 years' experience. *Ann. Am. Thorac. Soc.* **12**, 70–78 (2015).
26. Miragaia, M. et al. Molecular characterization of methicillin-resistant *Staphylococcus epidermidis* clones: evidence of geographic dissemination. *J. Clin. Microbiol.* **40**, 430–438 (2002).
27. Lee, J. Y. H. et al. Global spread of three multidrug-resistant lineages of *Staphylococcus epidermidis*. *Nat. Microbiol.* **3**, 1175–1185 (2018).
28. Huffnagle, G. B., Dickson, R. P. & Lukacs, N. W. The respiratory tract microbiome and lung inflammation: a two-way street. *Mucosal Immunol.* **10**, 299–306 (2017).
29. Hildebrand, F. et al. Antibiotics-induced monodominance of a novel gut bacterial order. *Gut* **68**, 1781–1790 (2019).
30. Klein, E., Smith, D. L. & Laxminarayan, R. Hospitalizations and deaths caused by methicillin-resistant *Staphylococcus aureus*, United States, 1999–2005. *Emerg. Infect. Dis.* **13**, 1840–1846 (2007).
31. Mager, D. L., Ximenez-Fyvie, L. A., Haffajee, A. D. & Socransky, S. S. Distribution of selected bacterial species on intraoral surfaces. *J. Clin. Periodontol.* **30**, 644–654 (2003).
32. Takeshita, T. et al. Bacterial diversity in saliva and oral health-related conditions: the Hisayama Study. *Sci. Rep.* **6**, 22164 (2016).
33. Gu, S. et al. Alterations of the gut microbiota in patients with coronavirus disease 2019 or H1N1 influenza. *Clin. Infect. Dis.* **71**, 2669–2678 (2020).
34. Abu-Ali, G. S. et al. Metatranscriptome of human faecal microbial communities in a cohort of adult men. *Nat. Microbiol.* **3**, 356–366 (2018).
35. Drevinek, P. & Mahenthiralingam, E. *Burkholderia cenocepacia* in cystic fibrosis: epidemiology and molecular mechanisms of virulence. *Clin. Microbiol. Infect.* **16**, 821–830 (2010).
36. Flannagan, R. S. & Valvano, M. A. *Burkholderia cenocepacia* requires RpoE for growth under stress conditions and delay of phagolysosomal fusion in macrophages. *Microbiology* **154** (Pt 2), 643–653 (2008).
37. Uribe-Quero, E. & Rosales, C. Control of phagocytosis by microbial pathogens. *Front. Immunol.* **8**, 1368 (2017).
38. Bazzini, S. et al. Deciphering the role of *md* efflux transporters in *Burkholderia cenocepacia*. *PLoS ONE* **6**, e18902 (2011).
39. Tomich, M. & Mohr, C. D. Adherence and autoaggregation phenotypes of a *Burkholderia cenocepacia* cable pilus mutant. *FEMS Microbiol. Lett.* **228**, 287–297 (2003).
40. Kokkoris, S. et al. ICU-acquired bloodstream infections in critically ill patients with COVID-19. *J. Hosp. Infect.* **107**, 95–97 (2020).
41. Osman, H., Nguyen, P. & FIRST, C. A. S. E. First case of covid-19 complicated with *Burkholderia cepacia* pneumonia and bacteremia. *Chest* **158**, A544 (2020).
42. Meys, I. et al. Coronavirus disease 2019 in patients with inborn errors of immunity: an international study. *J. Allergy Clin. Immunol.* **147**, 520–531 (2020).
43. Häfliger, E., Atkinson, A. & Marschall, J. Systematic review of healthcare-associated *Burkholderia cepacia* complex outbreaks: presentation, causes and outbreak control. *Infect. Prev. Pract.* **2**, 100082 (2020).
44. Schwab, U. et al. Localization of *Burkholderia cepacia* complex bacteria in cystic fibrosis lungs and interactions with *Pseudomonas aeruginosa* in hypoxic murine. *Infect. Immun.* **82**, 4729–4745 (2014).
45. Valentine-King, M. A., Cisneros, K., James, M. O., Huigens, R. W. & Brown, M. B. Turning the tide against antibiotic resistance by evaluating novel, halogenated phenazine, quinoline, and NH125 compounds against ureaplasma species clinical isolates and Mycoplasma type strains. *Antimicrob. Agents Chemother.* **63**, e02265–18 (2019).
46. Krausse, R. & Schubert, S. In-vitro activities of tetracyclines, macrolides, fluoroquinolones and clindamycin against *Mycoplasma hominis* and *Ureaplasma ssp.* isolated in Germany over 20 years. *Clin. Microbiol. Infect.* **16**, 1649–1655 (2009).
47. Otto, M. *Staphylococcus epidermidis*—the 'accidental' pathogen. *Nat. Rev. Microbiol.* **7**, 555–567 (2009).
48. Méric, G. et al. Ecological overlap and horizontal gene transfer in *Staphylococcus aureus* and *Staphylococcus epidermidis*. *Genome Biol. Evol.* **7**, 1313–1328 (2015).
49. Rossi, C. C. et al. Transfer of mupirocin resistance from *Staphylococcus haemolyticus* clinical strains to *Staphylococcus aureus* through conjugative and mobilizable plasmids. *FEMS Microbiol. Lett.* **363**, fnw121 (2016).
50. Cafini, F. et al. Horizontal gene transmission of the *cf* gene to MRSA and *Enterococcus*: role of *Staphylococcus epidermidis* as a reservoir and alternative pathway for the spread of linezolid resistance. *J. Antimicrob. Chemother.* **71**, 587–592 (2016).
51. Soriano, M. C. et al. Low incidence of co-infection, but high incidence of ICU-acquired infections in critically ill patients with COVID-19. *J. Infect.* **82**, e20–e21 (2020).
52. Zumla, A. et al. Rapid point of care diagnostic tests for viral and bacterial respiratory tract infections—needs, advances, and future prospects. *Lancet Infect. Dis.* **14**, 1123–1135 (2014).
53. Garner, J. S., Jarvis, W. R., Emori, T. G., Horan, T. C. & Hughes, J. M. CDC definitions for nosocomial infections, 1988. *Z. Arztl. Fortbild.* **85**, 818–827 (1991).
54. Kalil, A. C. et al. Management of adults with hospital-acquired and ventilator-associated pneumonia: 2016 Clinical Practice Guidelines by the Infectious Diseases Society of America and the American Thoracic Society. *Clin. Infect. Dis.* **63**, e61–e111 (2016).
55. Xiao, M. et al. Multiple approaches for massively parallel sequencing of SARS-CoV-2 genomes directly from clinical samples. *Genome Med.* **12**, 57 (2020).
56. Chen, S., Zhou, Y., Chen, Y. & Gu, J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890 (2018).
57. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).

58. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
59. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
60. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
61. Holden, M. T. G. et al. The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. *J. Bacteriol.* **191**, 261–277 (2009).