Article @ Virology

Metagenomic Analysis of Blood Samples from Ten COVID-19 Patients in Wuhan, China

Jiwei Ding[1] †, Hui Long[2] †, Wei Wang[3] †, Baoli Zhu[4,5,6,7], Nicholas R. Waterfield[8], Qiang Wang[9], Yun Wu[10]*, Shan Cen1[10]*, Qingming Wu[2,9,10]**

1. *Institute of Medicinal Biotechnology, Chinese Academy of Medical Science & Peking Union Medical College, 100050, P.R. China.*
2. *Tianyou Hospital, Wuhan University of Science and Technology, Wuhan, Hubei, 430064, P.R. China.*
3. *Department of respiratory medicine, Beijing Friendship Hospital, Capital Medical University, 100050, P.R. China.*
4. *CAS Key Laboratory of Pathogenic Microbiology and Immunology, Institute of Microbiology, Chinese Academy of Sciences, Beijing 100101, P.R. China.*
5. *Savaid Medical School, University of Chinese Academy of Sciences, Beijing 100049, P.R. China.*
6. *Beijing Key Laboratory of Antimicrobial Resistance and Pathogen Genomics, Beijing 100101, P.R. China.*
7. *Department of Pathogenic Biology, School of Basic Medical Sciences, Southwest Medical University, Luzhou, Sichuan 646000, P.R. China.*
8. *Warwick Medical School, Warwick University, Coventry CV4 7AL, UK*
9. *Institute of Infection, Immunology and Tumor Microenvironment, Hubei Province, Key Laboratory of Occupational Hazard Identification and Control, Medical College, Wuhan University of Science and Technology, Wuhan 430065, P.R. China.*
10. *Beijing Institute of Tropical Medicine, Beijing Friendship Hospital, Capital Medical University, 100050, P.R.China.*

Article History: Submitted: 21/06/2023; Revised: 18/07/2023; Accepted: 21/07/2023
DOI:10.21092/jav.v12i3.110
Key Words: Metagenomics, SARS-CoV-2, COVID-19, Coinfection, LEfSe Analysis

† These authors contributed equally to this work and share first authorship.
* Corresponding author: Qingming Wu, wuhe9224@sina.com,86-18802705686;
Shan Cen, shancen@imb.pumc.edu.cn, 86-10-63037279;
Yun Wu, wuyun@ccmu.edu.cn, 86-13601028461;
Qiang Wang, wangqiang@wust.edu.cn, 86-15971509779

ABSTRACT

Metagenomic next-generation sequencing of samples from 10 patients with documented SARS-CoV-2 in Wuhan, China from Feb 3rd to Mar 13th identified atypical coinfection of *chlamydophilla* and *avocado sunblotch viroid* in 1 of 10 blood samples from a critically ill patient. Furthermore, LEfSe analysis revealed that SARS-CoV-2 infection significantly reshaped the blood microbial community. The patients showed an increase of bacteria belonging to *Campylobacterales* and *Rickettsiales* order whereas healthy volunteers were highly colonized by *Proteobacteria* and *Actinobacteria*. Collectively, our study highlights the importance of microbial sampling and judicious use of antimicrobials in severe and critically ill COVID-19 patients, and provides insights into the changes in microbial community composition by SARS-CoV-2 infection.

Abbreviations:

SARS-CoV-2, Severe Acute Coronavirus-2; COVID-19, Coronavirus Disease 2019 ; ARDS, Acute Respiratory Distress Syndrome; RT-PCR, Reverse Transcription Polymerase Chain Reaction; ELISA, Enzyme-linked Immunosorbent Assay;
LDA, Linear Discriminant Analysis; LEfSe, Linear discriminant analysis Effect Size.

Author contributions

JWD, SC, QMW designed the studies and wrote the manuscript. NRW, SC and YW supervised the study and revised the manuscript. QW and YW collected patients' data and serum samples. HL and WW were responsible for DNA extraction and sequencing. JWD and HL performed metagenomics analysis. BLZ provided guidance in data analysis and R software.

Introduction

An outbreak of pneumonia associated with the infection of severe acute coronavirus 2 (SARS-CoV-2) started in December, 2019 and spread rapidly to most places of the world. As of Nov 20th, 2020, the total number of infected individuals has risen sharply to 57,137,661 worldwide. The clinical spectrum of this acute respiratory disease denoted as coronavirus disease 2019 (COVID-19) ranges from asymptomatic to critically ill cases. Reported clinical complications included sepsis [1, 2] and acute respiratory distress syndrome (ARDS) [3, 4], which are the main causes of requirement for mechanical ventilation and other intensive medical treatment. Risk factors associated with the development of ARDS and progression from ARDS to death includes older age, neutrophilia, comorbidities such as chronic pulmonary disease，hypertension, diabetes and cardiovascular diseases [3, 5, 6].

The epidemiological and clinical characteristics of SARS-CoV-2 pneumonia patients are intensively studied[7-12]. However, the co-infection with other bacteria or viruses and whether the co-infection will strengthen the illness has remained unclear. Bacterial/fungal coinfection research conducted by Goyal and colleagues [13] reported 19 of 338 (6%) cases of bacteria during hospital admission. In another study conducted by Kim D et al, 24 (20.7%) were positive for one or more other pathogens in SARS-CoV-2 positive patients, compared with 294 of the 1101 specimens (26.7%) negative for SARS-CoV-2, the main predominant co-infection were rhinovirus/ enterovirus (6.9%), respiratory syncytial virus (5.2%), and non-SARS-CoV-2 Coronaviridae (4.3%) [14]. Several lines of evidences indicates higher incidence of co-infection in critically ill patients and non-survivors compared to patients with mild or moderate symptoms[15]. In a retrospective report, co-infection was identified 3.7% of the patients and 41% of patients receiving intensive care ($p<0.005$) [16]. Christopher J. Lehmann reported that co-infection could range from 14-31% in intensive care units (ICU) patients and occurred in approx. 50% of non-survivors[16] . Zhou et al showed that in the current 2019 coronavirus (COVID-19) pandemic, half of patients with COVID-19 who have died had secondary bacterial infections[17]. In a recent letter, Cox et al. highlighted a need to monitor coinfections in SARS-CoV-2 infected patients to understand whether coinfection shapes disease progression, and to enable antimicrobial stewardship [18].

As a new emerging infectious disease, metagenomic sequencing provides a clue for recovering the whole viral genome to meet an epidemiological purpose and also to determine the co-infection which may be potentially relevant to an increase in morbidity and mortality[12, 19-21]. In this study, we performed metagenomic sequencing on 10 blood samples from patients who have

been tested positive for SARS-CoV-2 diagnostic testing aiming to elucidate the microbial community in COVID-19 patients and identify the key co-infection that shaped the disease development. Our results suggested there are discrepancies in microbial compositions between the healthy individuals and COVID-19 patients. More importantly, we identified that *chlamy--dophilla* and *Avocado sunblotch viroid* might be probable risk factors for severe illness from COVID-19.

Material and Methods

1. Study design and participants

Ten patients who are diagnosed as SARS-CoV-2 pneumonia, according to WHO interim guidance, from Feb 3rd to 20th, 2020 in Tianyou Hospital of Wuhan University of Science and Technology are involved in our study (see table 1), which was approved by Ethical Review Board of Medical College of Wuhan Science and Technology University. All patients are clinically typed as ordinary patients except for a 37-year old female, who was typed as critically ill patient. We collected clinical data for age, sex, hospitalized/leave date, exposure history, symptoms from onset to hospital admission (fever, cough, chest pain, choking, dyspnoea, malaise, myalgia, nausea, vomiting, diarrhoea, hemoptysis), chronic medical history (hypertension, diabetes, malignancy, etc). Laboratory values taken on admission included blood routine, inflamma--tory index such as hsCRP, SAA, PCT, live and kidney function, myocardial enzymo--gram, coagulation function, arterial blood gas analysis. Computed tomography scan, medical treatment (antiviral agents, antibacterial agents, corticosteroids, oxygen therapy, immunoglobulin, Traditional Chinese medicine and treatment outcomes were also collected. As a routine, electronic medical data were archived onto a local server, from which we retrieved these data.

2.SARS-CoV-2-associated blood sample collection and IgG/IgM antibody testing

Written consent from all patients were obtained under a study protocol approved by National Health Commission, P.R. China. Peripheral blood samples were obtained from 10 patients in Tianyou Hosipital who was given a diagnosis of COVID-19 upon the qualitative reverse transcription (RT-PCR) testing for SARS-CoV-2 RNA. The samples were subjected to RNA extraction using the QIAamp viral RNA minikit (Qiagen), and RNA was reverse-transcribed by Superscript reverse transcriptase kit (Invitrogen), followed by qRT-PCR testing for SARS-CoV-2 using primers. Samples were also subjected to anti-IgG/IgM antibody detection using an enzyme-linked immuno--sorbent assay (ELISA).

3. RNA sequencing

Ten milliliters of peripheral blood samples were collected into Lakebio cell-free RNA preservation tube (Lakebio). Blood samples were first centrifuged at 1600g for 10 min,

and the supernatant was centrifuged again at 16000*g* for 10 min to remove residual cells. The plasma was inactivated at 56°C for 30min and then stored at -80°C. Each plasma sample was frozen and thawed only once. Viral RNA was extracted using the QIAamp Viral RNA Mini Kit (Qiagen). The sequencing library was constructed using the NEBNext Multiplex Small RNA Library Prep Set (New England Biolabs). Approximately 10 million 100 bp paired-end reads were collected for each sample on the NextSeq 500 platform (Illumina).

4. Raw data filtering and rapid identification of pathogen species

4.1 Quality control

It is necessary to first remove low-quality sequencing data prior to assembly of the reads in order to improve the accuracy of reads for the follow-up analyses. To this end, raw reads were filtered using Trim Galore version 0.6.4 (https://github.com/FelixKrueger/TrimGalore) (setting: -q 20 --length 20 -e 0.1 --stringency 3) to remove adaptors, low-quality reads and reads less than 20bp.

4.2 Remove host contamination

To eliminate the confusion caused by host sequences, all the clean data that passed QC was mapped to the host reference genome utilizing BWA version 0.7.17[22], only the unmapped sequences were used in subsequent analysis.

4.3 Read Assembly and Species identification

Clean reads were de novo assembled by MEGAHIT version 1.0[23]. Clean reads are aligned to assembled contigs using BWA version 0.7.17. A host sequence was determined based on BLAST version 2.7.1 and was removed by satisfying one of the following conditions: (1) Length of matched area ≥ 500 bp, alignment similarity ≥ 90%; (2) Length of matched area accounts for more than 80% of the total length of contigs, and alignment similarity ≥ 90%. Then, (3) Cdhit version 4.6[24] was used to cluster the assembled contigs from each sample. Contigs were then classified by BLASTx against the NT database using alignment similarity ≥ 80%, length of matched area ≥ 500 bp and e-value ≤ 10^{-5}. Contigs with significant BLASTx hits were confirmed as virus sequences.

5. Linear Discriminant Analysis Effect Size (LEfSe)

The LEfSe[25] was calculated using the online version of Galaxy3. For OTUs with an average abundance in all samples that was greater than 0.1%, abundances were normalized to the sum of the values per sample in 1 million and then subjected to linear discriminant analysis (LDA). The LDA was performed using a one-against-all strategy, and OTUs showing a score higher than 2.0 were selected.

Table 1. Clinical characteristics of the patients and healthy volunteers involved in the study

Sample	Clinical type	Gender	Age	Time from onset to hospitalization (days)	Clinical symptoms at addimision	Nucleic acid testing before admission	Comorbidities	Duration of hospitalization (days)	Medication before admission	Antiviral and antibiotics therapy	Oxygen therapies (days)	Other therapies	Clinical outcomes
SmR2D200315_2A	ordinary	M	63	7	Fever, cough, loss of appetite, diarrhea, muscle ache	/	hypertension	15	/	Abidor 0.2g tid for 6 days;Moxifloxacin 0.4g, qd for 10 days	16		recover and discharge
SmR3D200315_3A	ordinary	F	48	3	Dry cough	/	/	18	/	ganciclovir 0.5g qd for 11 days;Levofloxacin 0.4g, qd for 12 days		immunoglobin 10g qd for 3 days	recover and discharge
SmR4D200315_4A	ordinary	M	80	13	Fever, cough, muscle ache, fatigue, palpitation, shortness of breath	/	hypertension、Postoperative bladder cancer	19	Oseltamivir, Lianhua Qingwen	Abidor 0.2g tid for 10 days;ganciclovir 0.5g qd for 13 days;Moxifloxacin 0.4g, qd for 15 days	19	immunoglobin 10g qd for 7 days	recover and discharge
SmR5D200315_5A	ordinary	M	46	8	Fever, dry cough, chest distress and dyspnea	/	bronchitis	20	γ-globulin, methylprednisolone, clindamycin, abidol, Jinyebaidu, oseltamivir, Asmet	Abidor 0.2g tid for 10 days; ribavirin 0.5g bid for 5 days;ganciclovir 0.5g qd for 8 days;Moxifloxacin 0.4g, qd for 8 days	20	immunoglobin 10g qd for 7 days	recover and discharge
SmR6D200315_6A	critically ill	F	37	12	Fever, cough, chest pain, dyspnea, diarrhea	negative	/	36	methylprednisolone	Abidor 0.2g tid for 7 days; ribavirin 0.5g bid for 3 days; Lopinavir/ritonavir, 2 tablets for 7 days;Moxifloxacin 0.4g, qd for 12 days	36	immunoglobin 10g qd for 11 days;Miresone 40mg for 8 days	recover and discharge
SmR7D200315_7A	ordinary	M	69	10	fever	positive	hypertension	11	amoxicillin, Kewei, moxifloxacin	ganciclovir 0.5g qd for 3 days;Levofloxacin 0.4g, qd for 10 days			recover and discharge
SmR8D200315_8A	ordinary	F	44	4	Diarrhea, fever, cough	/	/	18	Moxifloxacin, oseltamivir and Lianhua Qingwen	Abidor 0.2g tid for 9 days;ribavirin 0.5g bid for 3 days; ganciclovir 0.5g qd for 10 days;Levofloxacin 0.4g, qd for 8 days	19		recover and discharge
SmR9D200315_9A	ordinary	M	60	14	Fever, cough, dyspnea, hemoptysis, general fatigue, soreness, chest pain, chest distress	positive	/	17	/	Abidor 0.2g tid for 5 days; ribavirin 0.5g bid for 5 days;ganciclovir 0.5g qd for 9 days;Moxifloxacin 0.4g, qd for 12 days	17	immunoglobin 10g qd for 8 days;methylprednisolone succinate 40mg qd for 3 days	recover and discharge
SmR10D200315_10A	ordinary	M	38	10	Fever, cough, chest distress, shortness of breath, muscle ache	positive	hypertension	15	Oseltamivir, Lianhua Qingwen, moxifloxacin	Abidor 0.2g tid for 8 days; ganciclovir 0.5g qd for 10 days;Cefoperazone tazobacta 2g bid for 7 days	12	immunoglobin 10g qd for 4 days;methylprednisolone succinate 40mg for 3 days	recover and discharge
SmR11D200315_11A	ordinary	M	46	4	Fever, nausea, vomiting, fatigue	negative	/	14	antivirals and antibiotics	Abidor 0.2g tid for 7 days; ribavirin 0.5g bid for 7 days;Moxifloxacin 0.4g, qd for 7 days		immunoglobin 10g qd for 5 days	recover and discharge
SmR11D200315_11A	healthy	M	38										
SmR1D200224_4A	healthy	F	34										

6. Statistical analysis

All statistical analyses were processed with SPSS software Version 14.0. Continuous variables expressed as mean±SD were compared by unpaired Student t test and categorical data presented as number (%) were compared by $\chi 2$ test or Fisher's exact test. A two-sided p value of <0.05 was considered statistically significant.

Results

1. Clinical characteristics of the patients involved in the study

A total of 10 COVID-19 patients and two healthy vonlunteers were involved in this study with an average age of 50.3±13.9 years (see table 1). Among them, 7 (70%) were males and 3 (30%) were females. Three (~30%) had been in close contact with diagnosed patients or visited fever clinic. Nine (90%) of people were non-severe. Additionally, five (50%) of the cohort have chronic diseases, of which four were reported to have chronic hypertension, one has bronchitis, and one suffered from bladder cancer combined with hypertension.

The critically ill patient was confirmed to have a decreased lymphocyte count of $3.0x10^9$/L, which returned to a normal level after treatment. Chest CT examinations indicated infections in bilateral lungs for all patients. The patient with chronic bronchitis also exhibited interstitial fibrosis. During the infection, the critically ill patient exhibited significant renal and heart function impairment which indicated the kidney and heart are organs that SARS-CoV-2 can target. All patients received oxygen therapy and a combined use of ganciclovir and either moxifloxacin or levofloxacin during treatment. Three patients including the critically ill patient received an injection of methylprednisolone sodium succinate at 40mg qd for successive 3~8 days. All patients improved and were discharged before Mar 13, 2020 (see figure 1A).

2. Analysis of metagenomics sequencing results

To assess the respiratory microbiome, 88, 960, 887 total reads were reconstructed with MEGAHIT. Nucleic acid mapping to the spike-in was identified in each sample and was correctly identified by the mapping software, providing confidence in the approach (data not shown). 4, 869, 768–13, 426, 153 (mean 7, 413, 407 ± 2, 659, 810) analyzed sequences were obtained from each sample. Classification of microbial taxa using Kraken resulted in the detected bacteria being classified into 5 kingdoms, 44 phyla, 83 classes, 184 orders, 416 families, and 1304 genera.

3. Differences of microbiota diversity between COVID-19 patients and healthy individuals

Alpha-diversity indices (Chao 1 and Shannon) (see figure 1B, C) and beta--diversity indices (bray-curtis and jaccard index) (see figure 1D, E) of healthy individuals and patients were calculated. Comparison between patients and healthy

group showed significant differences in Chao 1 (3180.2 ± 336.2, 3306.3 ± 554.6), and Shannon (5.7 ± 0.3, 5.1 ± 2.2) indices. Beta diversity is used to evaluate the temporal and spatial changes in species composition, reflecting whether there is difference in bacterial communities between groups. Inter-group and Intra-group Beta distance was shown in 3D plot (see figure 1D, E), with the results showing a clear separation between healthy and patient groups. Likewise, the PCA plot showing the dissimilarity of microbial community and also revealed a distinct microbiota composition between healthy and patient groups (see figure 1F).

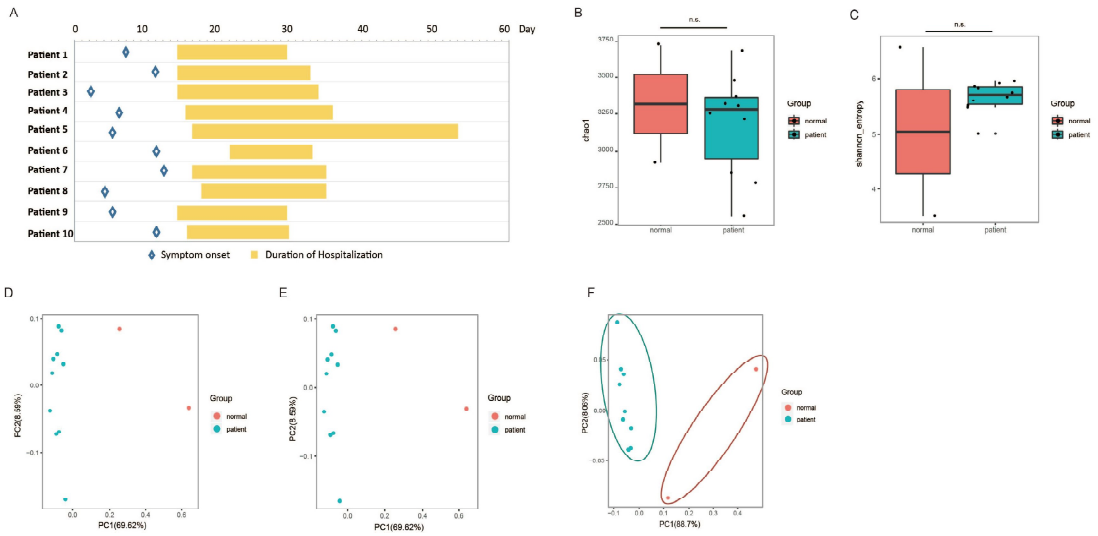

Figure 1. Boxplot of Alpha diversity indices.

Alpha diversity indices are composite indices reflecting species abundance and uniformity. (A) Timeline of patient symptom onset, hospitalization, and discharge, through the course of disease for 10 patients hospitalized. (B) Chao 1 indice reflects the OTU abundance in samples. (C) Shannon indice reflects the diversity of OTU in samples. The greater the Chao index, the higher the expected species richness of the microbiota, and the greater the Shannon index, the higher the diversity of the microbiota. Boxes indicates the interquartile range (IQR) between the first and third quartiles (25th and 75th percentiles, respectively), and the horizontal line inside the box represents the median. NS, non-significant (Student's t-test). (D) (E) Beta diversity analysis of patient and healthy samples. Three dimensional principal coordinates analysis plots showing the association of each sample using either the Bray-Curtis (D) or jaccard (E) metric. Individual samples are represented at spheres which are colored by group. (F) PCA score plot of the microbiota in different groups. Blue: COVID-19 patients, Red: healthy individuals. Each point in the plot indicates one sample, and different colors represent different groups. The distance between points reflects the level of difference. The closer the samples in the plot,

4. Viroids are predominant on the level of Phylum in critical patient

In respect to the nucleic acid mapping analysis of bacteria, viral or other pathogens we also identified various hits which we then mapped at both the phylum and genus level. The five most dominant (top 5) phyla identified were viruses, bacteria, viroids, archaea, and eukaryota (see figure 2A). The results indicated that all of patients were also co-infected with other bacteria and/ or viruses. This accounted for >75% of the blood microbial composition for all patients. One patient detected low frequency of Archaea and one patient detected the existence of Eukaryota (less than 1%).

Notably, viroids characterized the co-infection of the critical patient, showed ~25% of relative abundance, and ~27% and 10% of relative abundance in other two ordinary patients. Heatmap analysis gives a visualization of relative changes in microbial abundance at the phylum level (see figure 2B).

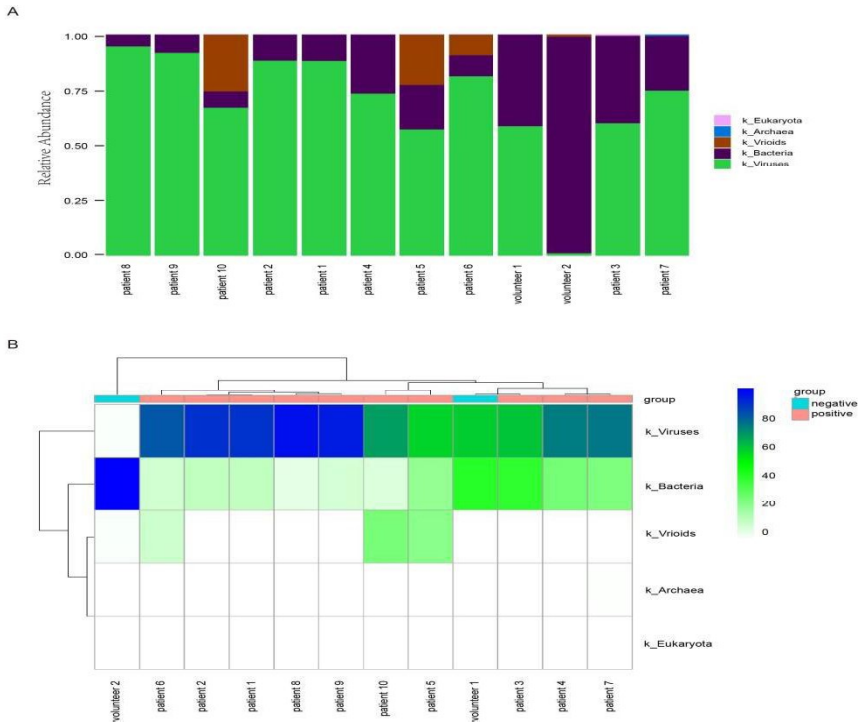

Figure 2. viroids predominate the co-infected microbes of the critically ill patient.

(A) Stacked bar-plot representation of dominant microbial community member groups and their relative abundance at the level of phylum based on mNGS sequences reconstructed by MGHIT. (B) Heatmap showing the top5 phyla with significant differences of relative abundance in 12 samples. The color code indicates relative abundance, ranging from white (low abundance) to green to blue (high abundance).

5. Avocado sunblotch viroid is detected from blood sample of critically ill patient

In order to further dissect which types of viroids are potentially associated with the critically ill patient, the relative abundance of these agents was then analyzed further. The top 15 taxa identified were ascribed to *burkholderia, malvastrum leaf curl Pillippines betasatellite, melegrivirus A, Diuris virus B, Murine osteosarcoma virus, ageratum yellow vein Singapore alpha--satellite, pseudomonas, dasshen mosaic virus, coconut foliar decay virus, grapevine yellow speckle viroid 1, tomato marchitez virus, Avocado sunblotch viroid, peanut stunt virus satellite RNA, vicia cryptic virus and pelargonium necrotic spot virus* (Figure 3A).

Interestingly, Burkholderia was ubiqui--tously present across most of the samples. Compared with healthy individuals, the abundance of Burkholderia was lower in SARS-CoV-2 infected individuals. It is unusual and a little intriguing that plant viruses and viroids predominate the microbial community. Avocado sunblotch viroid was only identified from the sample of critically patient and occupied approx. 20% of the abundance for the patient. Interestingly, this organism is not normally recognized as a pathological organism to humans and it is possible that its infection may facilitate the exacerbation of COVID-19, which of course needs further research.

Heatmaps provided do allow for the visualization of relative changes in microbial abundance at the genera level (Figure 3B).

6. Chlamydophilla is a probable causative bacterial for severe development of COVID-19

In order to further dissect which bacteria may be associated with the critically ill patient, the relative abundance of “at least” the genus level was analyzed specifically for the potential for any bacterial community isolates. The top 15 taxa identified were *burkholderia, pseudomonas, propioni--bacteriaceae, acinetobacter, rhodobiaceae, comamonas, sinobacteraceae, microcoleus, succinivibrionaceae, cupriavidus, rhodo--spirllum, capnocytophaga, chlamydophilla, rhodopirellula and subdoligranulum* (Figure 4A).

Burkholderia and *pseudomonas* was the major two genera among all the bacteria. It is worthy to note that *chlamydophilla* was only identified from the sample of critically ill patient and occupied approx. 10% of the relative abundance of the microbial community for the patient. To substantiate the relative abundance at genus level, we further analyzed the relative abundance at species level, unclassified *chlamydophilla* species accounted for ~12% of total species (Figure 4B). Unexpectedly, unclassified *granullicella* species was also enriched and accounted for ~5%.

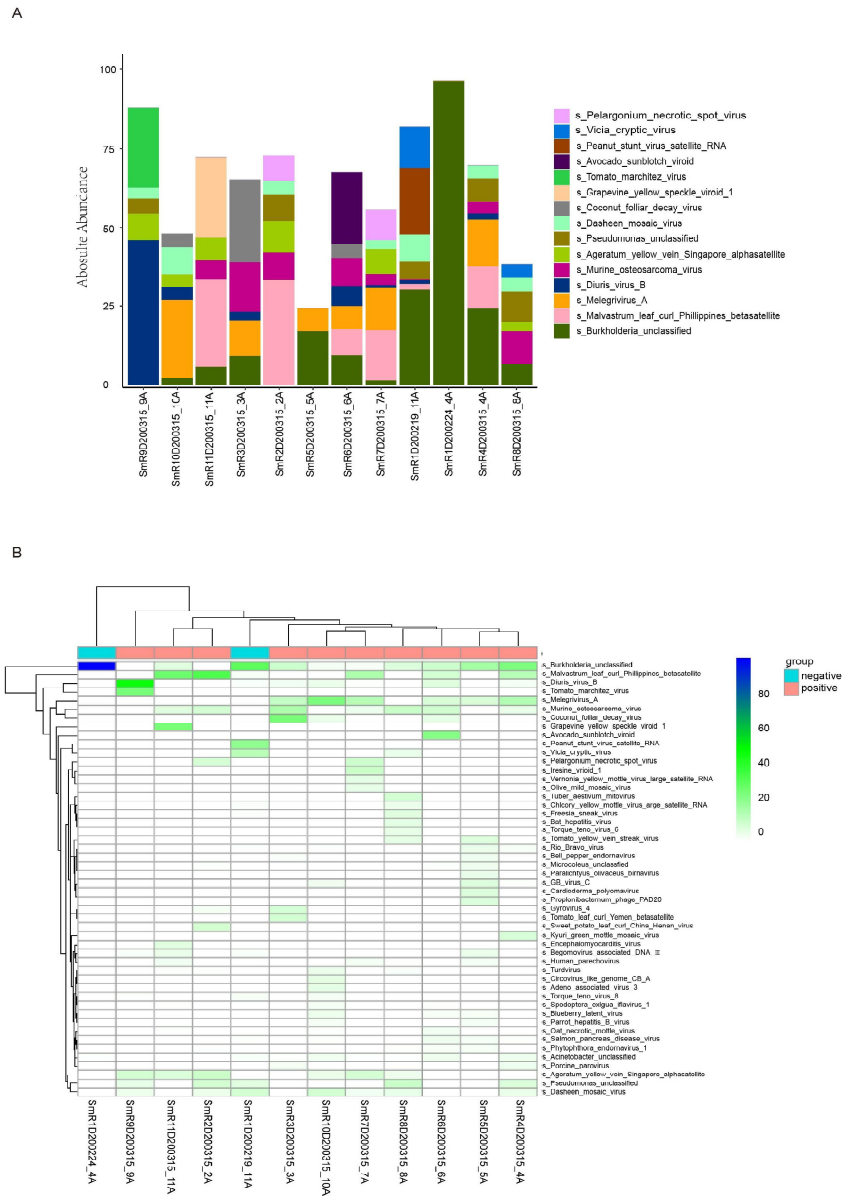

Figure 3. Avocado sunblotch viroid characterized the microbial community of the critically ill patient.

(A) Stacked bar-plot representation of dominant microbial community member groups and their relative abundance at the level of species based on mNGS sequences reconstructed by MGHIT. (B) Heatmap showing the species with significant differences of relative abundance in 12 samples.

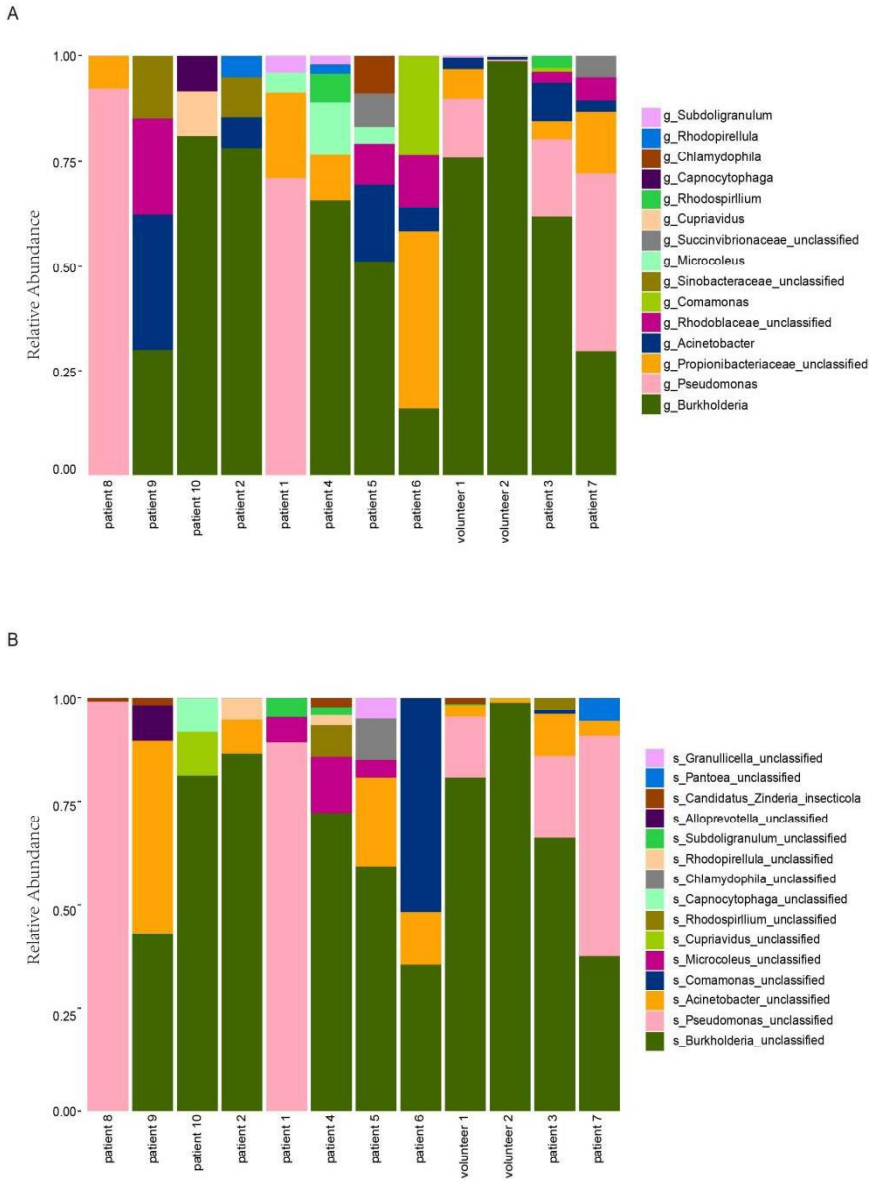

Figure 4. Chlamydophilla characterized the microbial community of the critically ill patient.

(A) Stacked bar-plot representation of dominant bacterial community member groups and their relative abundance at the level of genus based on mNGS sequences reconstructed by MGHIT. (B) Stacked bar-plot representation of dominant bacterial community member groups and their relative abundance at the level of species.

6. Differences in bacterial composition between COVID-19 patients and healthy individuals

To further investigate the information gleaned between COVID-19 patients and healthy individuals, a line discriminant analysis (LDA) effect size (LEfSe) analysis were performed using phylum to genus-level data to identify high-dimensional biomarkers for each group. The cladogram is shown in Figure 5A and LDA score distribution map is shown in Figure 5B. Interestingly, the microbial community of COVID-19 patients are predominated by phylum Cyanobacteria, which was not seen in healthy individuals (Figure 5). The patients showed an increase of bacteria belonging to Campylobacterales and Rickettsiales order. In contrast, healthy people were highly colonized by Proteo--bacteria and Actinobacteria, compared with COVID-19 group.

Healthy groups showed an increase of bacteria belonging to Alteromonadales and Burkholderiales order, Especially, Jiangellales, Propionibacteriales and Strepto--mycetales orders were only found in healthy groups. Eight representative species (top8) which featured COVID-19 patients were sp. PCC 7407 (p=0.0021), sp. NIES-3708 (p=0.0053), sp. NH-16 (p=0.0015), sp. JS92-SW72 (p=0.0012), pollinicola (p=0.0007), nigro-viridis (p=0.0055), minutus (p=0.0059) and Lachnospiraceae bacterium oral taxon 500 (p=0.0043).

Discussion

Pathogen co-infections are commonly identified in viral respiratory tract infection and often lead to an unfortunate increase of morbidity and mortality[26,27], necessitating timely diagnosis and proper use of antibiotics. For example, the prevalence of bacterial infection in patients with severe influenza is estimated to be around 20~30%, which is associated with an increased severity of illness and a higher risk of death. Most fatalities in 1918 influenza pandemic has been ascribed to the subsequent increase in bacterial infection levels, A good example includes *Streptococcus pneumonia*[28]. The characteristics of co-infection in patients infected with SARS-CoV-2 as well as its consequences remain elusive and become an important knowledge gap for better therapeutic strategies for COVID-19.

In our study, metagenomic methods have been employed to elucidate the obvious concern of the threat of co-infection with pathogens which critically ill patients face given the diverse characteristics of SARS-CoV-2-infected or indeed healthy individuals. We included one critically ill patient, 9 ordinary patients and 2 healthy individuals. We found that there exists remarkable divergence in species composition between healthy individuals and SARS-CoV-2 infected patients. Strikingly, two atypical co-infections were identified in critically ill patient, namely *chlamydophilla*

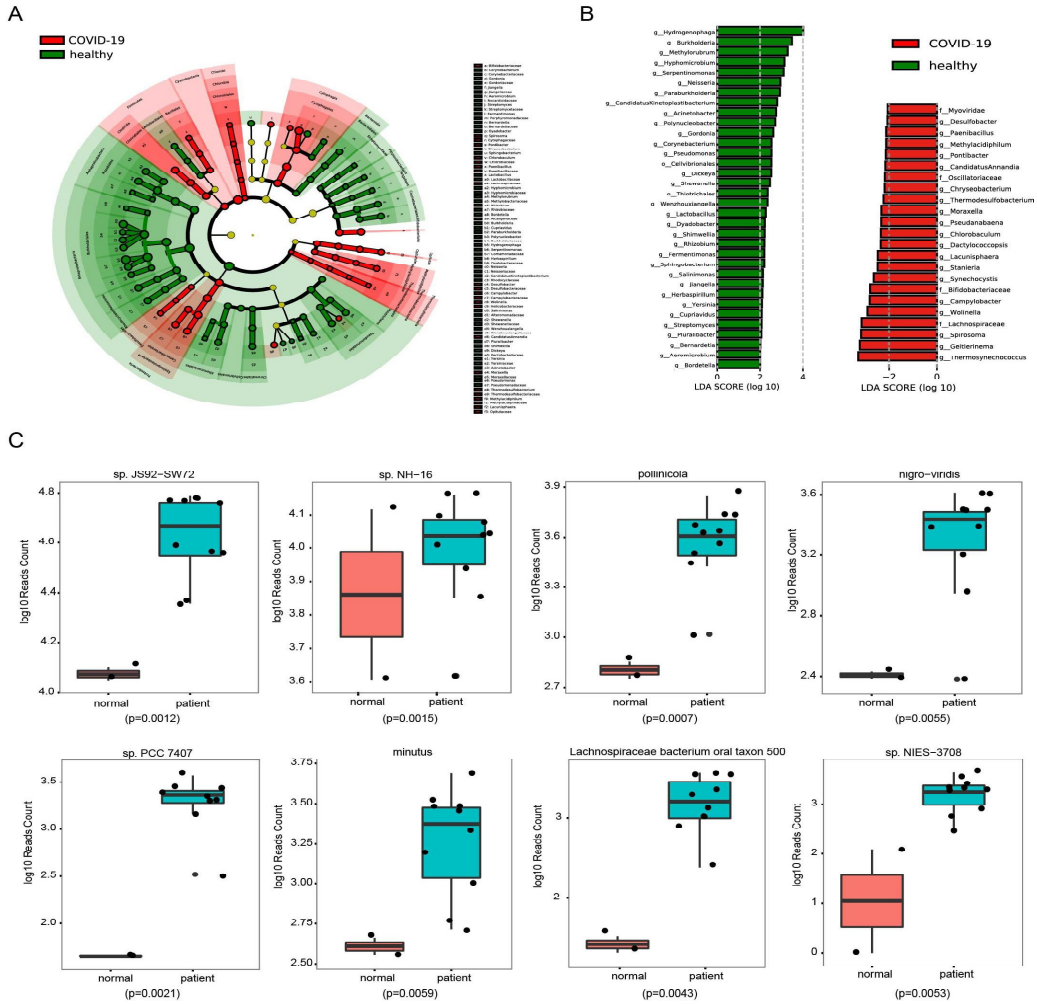

Figure 5. Differential abundances of microbial communities between COVID-19 group and healthy individual group.

(A) Taxonomic cladogram generated from LEfSe analysis showing significant difference in microbiota profile of the two groups, red represented the enriched taxa in COVID-19 patients' microbial community and green represented the enriched taxa in healthy individuals' microbial community. (B) Differently abundant taxa detected with cut-off value of linear discriminant analysis (LDA) score >2.0. Enriched taxa in COVID-19 patients (green) were indicated with positive LDA score, while taxa enriched in healthy individuals (red) have negative LDA score. (C) Comparison of relative abundance levels of microbiota at species level in COVID-19 patients and healthy individuals was evaluated. The boxplot shows 8 species of significant differences between COVID-19 patients and healthy individuals. The quartiles above and below the median with dark line at center of the box denoting median, black dots showing each sample. The respective p value for each family group is reported using Wilcoxon rank sum test.

and Avocado sunblotch viroid, which we propose was likely associated to the severity of illness. To the best of our knowledge, the identification of Avocado sunblotch viroid is the first time to be identified in the blood of SARS-CoV-2 positive patients while the presences of the members of the *chlamydophilla* genus is also quite rare in SARS-CoV-2 infected people. The genus of *chlamydophilla* comprises obligate intracellular pathogens that can occasionally cause septic infection. Notably, *Chlamy--dophila pneumoniae* causes acute and chronic respiratory tract infection with the clinical manifestation including sinusitis, asthma, pharyngitis, bronchitis severe pneumonia and cardiac diseases[29].

Co-infection with *C. pneumonia* was reported only in two cases in a large US study involving 5700 COVID-19 patients and in those cases no associated clinical outcomes are reported[30]. Prior evidences suggests that C. pneumonia infection induces proinflammatory cytokines including IL-1β, IL-6, IL-8, IL-12, TNF-α, IFN-γ, and intercellular adhesion molecule-1 in various systems including human peripheral blood mononuclear cells, alveolar macrophages, and in several mouse cell models[31]. It is plausible to speculate that co-infection with C. pneumonia might aggravate the occurrence and clinical symptom of cytokine storm which may be lethal to COVID-19 patients.

However, due to the limited number of patients involved in our study, we could not deduce a definite causal relationship between C. pneumonia co-infection and the severity of the illness. A larger cohort study is needed to decipher the relationship between these pathogenic microorganisms and the disease progression of COVID-19.

Strikingly, COVID-19 patients were highly colonized by cyanobacteria. Four of eight species (sp. PCC 7407, sp. NIES-3708, sp. JS92-SW72, nigro-viridis and minutus) which exhibit the most significant divergence in relative abundance between COVID-19 and healthy group belong to this taxa. Cyanobacterial LPS is associated with a range of pathological effects for humans, from gastro-intestinal illness, cutaneous signs and symptoms, allergy, respiratory disease, headache and fever[32]. Thus cyanobacterial co-infection could also lead to severer clinical symptoms. Moreover, an increase of bacteria belonging to Campylo--bacterales and Rickettsiales order were observed in patients whereas a high number of reads associated with Alteromonadales, Burkholderiales, Jiangellales, Propioni--bacteriales and Streptomycetales orders were detected in the healthy group. This suggested that Sars-CoV-2 infection alters the bacterial composition significantly, although the physiological significance of this alteration still needs further study.

Taken together, our study provides the first evidences that there is possible linkage between two atypical microorganisms and

the severity of illness in SARS-CoV-2 patient. We also found that distinct microbial compositions in the SARS-CoV-2 infected cohort. Our studies highlight that co-infecting microbial sampling should be encouraged for severe and critically ill patients. This would allow the judicious use of antimicrobials to potentially improve clinical outcomes through treatment of the co-infection. Nevertheless, we accept that more experimental data are needed to support our hypothesis regarding co-infections during Sars-CoV-2 pandemic.

Competing interests

The authors declare all financial and non-financial competing interests.

References

[1] Wiersinga W J, Rhodes A, Cheng A C, et al. Pathophysiology, transmission, diagnosis, and treatment of coronavirus disease 2019 (COVID-19): a review[J]. Jama, 2020, 324(8): 782-793.

[2] Colantuoni A, Martini R, Caprari P, et al. COVID-19 sepsis and microcirculation dysfunction[J]. Frontiers in physiology, 2020, 11: 747.

[3] Wu C, Chen X, Cai Y, et al. Risk factors associated with acute respiratory distress syndrome and death in patients with coronavirus disease 2019 pneumonia in Wuhan, China[J]. JAMA internal medicine, 2020, 180(7): 934-943.

[4] Takemoto M L S, Menezes M O, Andreucci C B, et al. Clinical characteristics and risk factors for mortality in obstetric patients with severe COVID‐19 in Brazil: a surveillance database analysis[J]. BJOG: An International Journal of Obstetrics & Gynaecology, 2020, 127(13): 1618-1626.

[5] Kuderer N M, Choueiri T K, Shah D P, et al. Clinical impact of COVID-19 on patients with cancer (CCC19): a cohort study[J]. The Lancet, 2020, 395(10241): 1907-1918.

[6] Cummings M J, Baldwin M R, Abrams D, et al. Epidemiology, clinical course, and outcomes of critically ill adults with COVID-19 in New York City: a prospective cohort study[J]. The lancet, 2020, 395(10239): 1763-1770.

[7] Whittaker E, Bamford A, Kenny J, et al. Clinical characteristics of 58 children with a pediatric inflammatory multisystem syndrome temporally associated with SARS-CoV-2[J]. Jama, 2020, 324(3): 259-269.

[8] Yue H, Zhang M, Xing L, et al. The epidemiology and clinical characteristics of co‐infection of SARS‐CoV‐2 and influenza viruses in patients during COVID‐19 outbreak[J]. Journal of medical virology, 2020, 92(11): 2870-2873.

[9] Chen H, Guo J, Wang C, et al. Clinical characteristics and intrauterine vertical transmission potential of COVID-19 infection in nine pregnant women: a retrospective review of medical records[J]. The lancet, 2020, 395(10226): 809-815..

[10] Chen L, Huang S, Yang J, et al. Clinical characteristics in patients with SARS‐CoV‐2/HBV co‐infection[J]. Journal of viral hepatitis, 2020, 27(12): 1504-1507.

[11] Pan A, Liu L, Wang C, et al. Association of public health interventions with the epidemiology of the COVID-19 outbreak in Wuhan, China[J]. Jama, 2020, 323(19): 1915-1923.

[12] Lu J, Du Plessis L, Liu Z, et al. Genomic epidemiology of SARS-CoV-2 in Guangdong province, China[J]. Cell, 2020, 181(5): 997-1003. e9.

[13] Grasselli G, Zanella A. Critically ill patients with COVID-19 in New York City[J]. Lancet (London, England), 2020, 395(10239): 1740.

[14] Kim D, Quinn J, Pinsky B, et al. Rates of co-infection between SARS-CoV-2 and other respiratory pathogens[J]. jama, 2020, 323(20): 2085-2086.

[15] Langford B J, So M, Raybardhan S, et al. Bacterial co-infection and secondary infection in patients with COVID-19: a living rapid review and meta-analysis[J]. Clinical microbiology and infection, 2020, 26(12): 1622-1629.

[16] Lehmann C J, Pho M T, Pitrak D, et al. Community-acquired coinfection in coronavirus disease 2019: a retrospective observational experience[J]. Clinical Infectious Diseases, 2021, 72(8): 1450-1452.

[17] Zhou F, Yu T, Du R, et al. Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study[J]. The lancet, 2020, 395(10229): 1054-1062.

[18] Cox M J, Loman N, Bogaert D, et al. Co-infections: potentially lethal and unexplored in COVID-19[J]. The Lancet Microbe, 2020, 1(1): e11.

[19] Lam T T Y, Jia N, Zhang Y W, et al. Identifying SARS-CoV-2-related coronaviruses in Malayan pangolins[J]. Nature, 2020, 583(7815): 282-285.

[20] Wu F, Zhao S, Yu B, et al. A new coronavirus associated with human respiratory disease in China[J]. Nature, 2020, 579(7798): 265-269.

[21] Zuo T, Liu Q, Zhang F, et al. Depicting SARS-CoV-2 faecal viral activity in association with gut microbiota composition in patients with COVID-19[J]. Gut, 2021, 70(2): 276-284.

[22] Chertow D S, Memoli M J. Bacterial coinfection in influenza: a grand rounds review[J]. Jama, 2013, 309(3): 275-282.

[23] Meskill S D, O’Bryant S C. Respiratory virus co-infection in acute respiratory infections in children[J]. Current infectious disease reports, 2020, 22: 1-8.

[24] Morens D M, Taubenberger J K, Fauci A S. Predominant role of bacterial pneumonia as a cause of death in pandemic influenza: implications for pandemic influenza preparedness[J]. The Journal of infectious diseases, 2008, 198(7): 962-970.

[25] Honarmand H. Q Fever: an old but still a poorly understood disease[J]. Interdisciplinary perspectives on infectious diseases, 2012, 2012(1): 131932.

[26] Richardson S, Hirsch J S, Narasimhan M, et al. Presenting characteristics, comorbidities, and outcomes among 5700 patients hospitalized with COVID-19 in the New York City area[J]. JAMA, 2020, 323(20): 2052-2059.

[27] Yang J, Hooper W C, Phillips D J, et al. Induction of proinflammatory cytokines in human lung epithelial cells during Chlamydia

pneumoniae infection[J]. Infection and immunity, 2003, 71(2): 614-620.

[28] Facciponte D N, Bough M W, Seidler D, et al. Identifying aerosolized cyanobacteria in the human respiratory tract: A proposed mechanism for cyanotoxin-associated diseases[J]. Science of the Total Environment, 2018, 645: 1003-1013.

[29] Li H, Durbin R. Fast and accurate short read alignment with Burrows－Wheeler transform[J]. bioinformatics, 2009, 25(14): 1754-1760.

[30] Li D, Liu C M, Luo R, et al. MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph[J]. Bioinformatics, 2015, 31(10): 1674-1676.

[31] Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences[J]. Bioinformatics, 2006, 22(13): 1658-1659.

[32] Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation[J]. Genome biology, 2011, 12: 1-18.