

RESEARCH

Open Access



# Human microbiota dysbiosis after SARS-CoV-2 infection have the potential to predict disease prognosis

Jie Zhou<sup>1,2†</sup>, Xiping Yang<sup>1,2†</sup>, Yuecong Yang<sup>1,2</sup>, Yiru Wei<sup>1,2</sup>, Dongjia Lu<sup>1,2</sup>, Yulan Xie<sup>1,2</sup>, Hao Liang<sup>1,2,3,4</sup>, Ping Cui<sup>1,3</sup>, Li Ye<sup>1,2,4\*</sup> and Jiegang Huang<sup>1,2\*</sup>

## Abstract

**Background** The studies on SARS-CoV-2 and human microbiota have yielded inconsistent results regarding microbiota  $\alpha$ -diversity and key microbiota. To address these issues and explore the predictive ability of human microbiota for the prognosis of SARS-CoV-2 infection, we conducted a reanalysis of existing studies.

**Methods** We reviewed the existing studies on SARS-CoV-2 and human microbiota in the Pubmed and Bioproject databases (from inception through October 29, 2021) and extracted the available raw 16S rRNA sequencing data of human microbiota. Firstly, we used meta-analysis and bioinformatics methods to reanalyze the raw data and evaluate the impact of SARS-CoV-2 on human microbial  $\alpha$ -diversity. Secondly, machine learning (ML) was employed to assess the ability of microbiota to predict the prognosis of SARS-CoV-2 infection. Finally, we aimed to identify the key microbiota associated with SARS-CoV-2 infection.

**Results** A total of 20 studies related to SARS-CoV-2 and human microbiota were included, involving gut ( $n=9$ ), respiratory ( $n=11$ ), oral ( $n=3$ ), and skin ( $n=1$ ) microbiota. Meta-analysis showed that in gut studies, when limiting factors were studies ruled out the effect of antibiotics, cross-sectional and case-control studies, Chinese studies, American studies, and Illumina MiSeq sequencing studies, SARS-CoV-2 infection was associated with down-regulation of microbiota  $\alpha$ -diversity ( $P < 0.05$ ). In respiratory studies, SARS-CoV-2 infection was associated with down-regulation of  $\alpha$ -diversity when the limiting factor was V4 sequencing region ( $P < 0.05$ ). Additionally, the  $\alpha$ -diversity of skin microbiota was down-regulated at multiple time points following SARS-CoV-2 infection ( $P < 0.05$ ). However, no significant difference in oral microbiota  $\alpha$ -diversity was observed after SARS-CoV-2 infection. ML models based on baseline respiratory (oropharynx) microbiota profiles exhibited the ability to predict outcomes (survival and death, Random Forest, AUC=0.847, Sensitivity=0.833, Specificity=0.750) after SARS-CoV-2 infection. The shared differential *Prevotella* and *Streptococcus* in the gut, respiratory tract, and oral cavity was associated with the severity and recovery of SARS-CoV-2 infection.

<sup>†</sup>Jie Zhou and Xiping Yang contributed equally to this work.

\*Correspondence:

Li Ye

yeli@gxmu.edu.cn

Jiegang Huang

jieganghuang@gxmu.edu.cn

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



**Conclusions** SARS-CoV-2 infection was related to the down-regulation of  $\alpha$ -diversity in the human gut and respiratory microbiota. The respiratory microbiota had the potential to predict the prognosis of individuals infected with SARS-CoV-2. *Prevotella* and *Streptococcus* might be key microbiota in SARS-CoV-2 infection.

**Keywords** SARS-CoV-2/COVID-19, Human microbiota,  $\alpha$ -diversity, Key microbiota, Machine learning

## Background

SARS-CoV-2 infection can manifest as asymptomatic, mild symptoms, acute respiratory symptoms, and even respiratory failure [1]. Elderly individuals and those with weakened immune function are prone to experiencing severe complications from SARS-CoV-2 [2]. Studies have suggested that auxiliary proteins encoded by SARS-CoV-2, such as ORF9c and ORF10, may play a crucial role in virus replication and immune evasion, contributing to the pathogenic mechanism of SARS-CoV-2 [3]. Additionally, research indicates that respiratory microbiota could influence the progression of COVID-19 by promoting local mucosal inflammation, modifying pathogen-associated molecular patterns, and depleting beneficial bacteria [4–6]. These findings suggest that human microbiota might have an important role in SARS-CoV-2 infection.

So far, there have been many studies exploring the relationships between SARS-CoV-2 infection and human microbiota. However, whether SARS-CoV-2 infection affects human microbiota is still controversial, especially regarding the  $\alpha$ -diversity index used to evaluate the integrity of the microbiota. One study showed a significant up-regulation of the  $\alpha$ -diversity of the gut microbiota in SARS-CoV-2 infected individuals [7]. Conversely, several other studies suggested a significant down-regulation of the  $\alpha$ -diversity of the gut microbiota following SARS-CoV-2 infection [8–12], while some studies found no change in the  $\alpha$ -diversity of the gut microbiota after infection [13, 14]. Similar inconsistencies can be observed in studies on SARS-CoV-2 infection and respiratory microbiota [15–17]. The alteration of human microbiota caused by SARS-CoV-2 infection has been utilized for diagnosing the severity of the disease [18], but there are limited research reports on whether it can predict disease prognosis. Studies have shown that human microbiota plays an important role in immune regulation following SARS-CoV-2 infection and can serve as biomarkers of the infection [19]. Different studies have identified different key human microbiota associated with SARS-CoV-2 [9, 20, 21]. Therefore, further investigation is necessary to determine the key human microbiota associated with SARS-CoV-2.

To address the aforementioned issues, we reviewed the existing studies on SARS-CoV-2 and human microbiota in the Pubmed and Bioproject databases and extracted the available raw 16S rRNA sequencing data of human

microbiota. Firstly, we utilized meta-analysis and bioinformatics methods to assess the impact of SARS-CoV-2 on the  $\alpha$ -diversity of human microbiota. Secondly, ML was employed to evaluate the ability of the microbiota to predict the prognosis of SARS-CoV-2 infection. Finally, we aimed to identify the key microbiota associated with SARS-CoV-2 infection.

## Methods

### Data sources and search strategy

PubMed and BioProject were searched using keywords, medical subject headings (MeSH) terms, and synonyms for SARS-CoV-2 and human microbiota from inception to Oct 29, 2021. Two independent reviewers evaluated each study, and an independent reviewer reviewed all studies. The exclusion criteria for studies were as follows: 1. Non-microbiota studies; 2. Non-SARS-CoV-2 studies; 3. No population studies; 4. Reviews, Commentaries, etc.; 5. Non-16S rRNA gene sequence studies; 6. Raw sequences not open access; 7. The raw data of the studies cannot distinguish between cases and controls; 8. Studies without negative control samples.

### Processing of raw data and calculation of $\alpha$ -diversity

Raw 16S rRNA sequencing sequences download and conversion using the sratoolkit tool developed by NCBI. The Fastp software for quality control of the original sequencing data and then imported the sequences into QIIME 2 (version 2021.11) [22]. Quality filtering of raw sequences using the q2-demux plugin, followed by DADA2 [23] (via q2-dada2) for denoising and generation of amplicon sequence variants (ASV) table. All ASV were aligned with Mafft [24] (q2-alignment) and used to construct a phylogenetic tree with fasttree2 [25] (q2-phylogeny). Species annotation of ASV was performed with silva-138 99% OTU reference sequences pre-trained with classify-sklearn Naive Bayes classifier [26] and q2-feature-classifier plugin [27].

The generated ASV table, phylogenetic tree, and species annotation files were imported into R 4.2.2., and the vegan package was used to normalize the sample sequence according to the minimum sample sequence number so that each sample could be compared under the same sequence number. Microbial  $\alpha$ -diversity indexes [Observed, Chao1, ACE, Fisher, Shannon, Simpson, Invsimpson, and Phylogenetic diversity (PD) indexes] were

calculated by the R microeco ([https://chiliubio.github.io/microeco\\_tutorial/](https://chiliubio.github.io/microeco_tutorial/)), [https://chiliubio.github.io/microeco\\_tutorial/https://chiliubio.github.io/microeco\\_tutorial/](https://chiliubio.github.io/microeco_tutorial/https://chiliubio.github.io/microeco_tutorial/) [28] package. According to the above steps, all available studies were processed respectively, and then quantitative consolidation of meta-analysis was performed.

### Meta-analysis

Quantitative synthetic analysis of  $\alpha$ -diversity indices from different sources using Review Manager 5.3 software (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2014.). For continuous variables, mean and standard deviation (*SD*) were used as analytical statistics, and each effect size was represented by a 95% *CI* value. The Fixed-effect models were used to quantitatively combine the  $\alpha$ -diversity of each study. The heterogeneity of the included studies was tested by the *Q* test and  $I^2$  test. If there was significant heterogeneity among studies (*Q* test of  $P < 0.05$  or  $I^2 > 50\%$ ), subgroup analyses were performed for included studies to assess the sources of heterogeneity. Limiting factors for subgroup analysis included antibiotic, study type (Cross-sectional studies and case–control studies were grouped into one category, and longitudinal studies into another), country (Chinese or American studies), sequencing platform (Illumina MiSeq), and sequencing region (V3–V4 or V4). Sensitivity analyses were performed by removing the studies with the largest sample size. Funnel plots were drawn to describe publication bias.

### Machine learning (ML) strategy

We used the microbiological feature screening tool linear discriminant analysis effect size (LEfSe, linear discriminant analysis [LDA] score  $> 2$ ,  $P < 0.05$ ) in the microeco package to identify the different ASV between groups and then construct the ML models. We used the caret package in R to build and evaluate ML models. We considered eight models related to kernel learning (support vector machine [SVM]), ensemble (random forest [RF] and extreme gradient boosting [XGB]), instance-based (k-nearest neighbor [KNN]), regulation (logistic regression [LR]), Bayesian (naiveBayes [NB]), decision tree (DT), neural networks (NN) algorithms to train the ASV table of the gut microbiota. We randomly divided the ASV tables into training and test datasets in a 7:3 ratio before training. We evaluated the training effects of different ML models by using tenfold cross-validation, and we repeated the process 10 times to obtain optimal parametric modeling. We evaluated the performance of the models in the test dataset from three perspectives: area under the curve [AUC], sensitivity, and specificity.

### Identification of key microbiota

The microeco package in R was employed to identify differential microbiota through the linear discriminant analysis effect size (LEfSe) method (LDA score  $> 2$ ,  $P < 0.05$ ). Cytoscape 3.7.1 was used to plot venn networks to identify shared differential microbiota across studies or different body parts.

### Statistical analysis

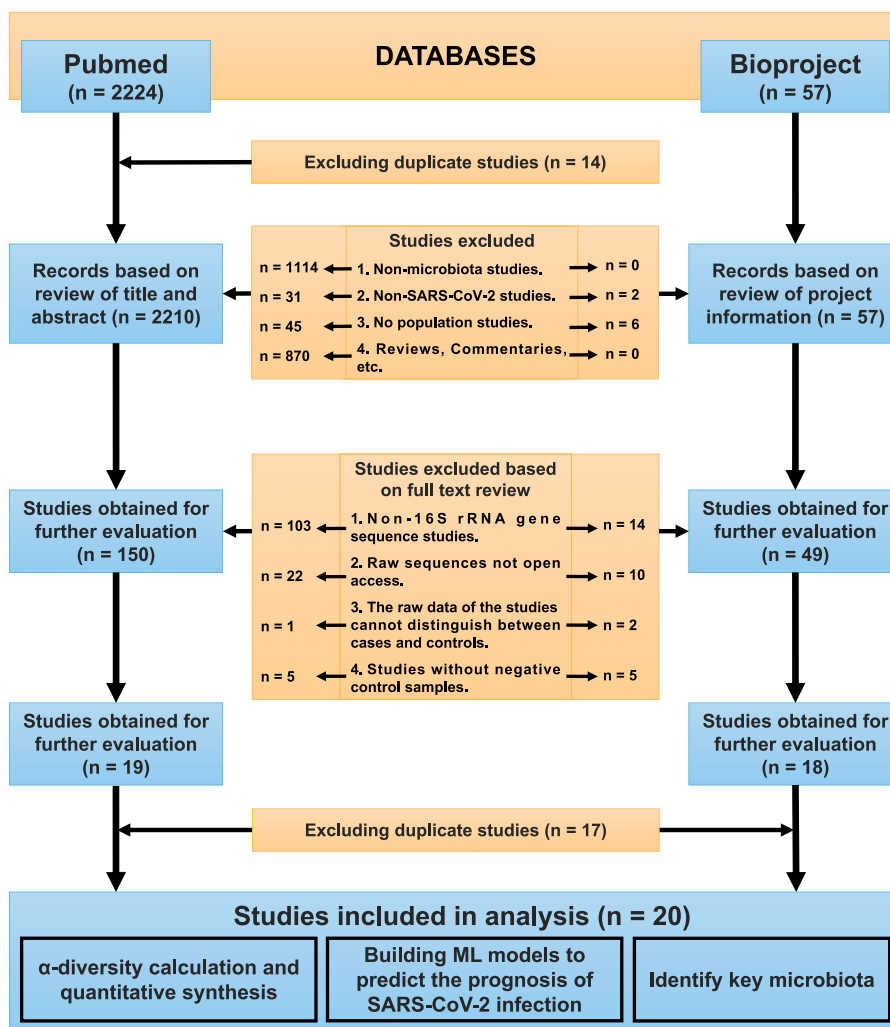
For data that meet normality, we analyzed them using the Student's *t* test. For data that did not meet normality, we analyzed them using the Mann–Whitney *U* test. All tests were two-sided, and  $P < 0.05$  was considered statistically significant.

## Results

### Meta-analysis of human microbiota $\alpha$ -diversity in SARS-CoV-2 infected individuals

A total of 2224 records from PubMed and 57 records from BioProject were retrieved (Fig. 1). Through a review of existing studies, we found inconsistent results regarding changes in human microbiota  $\alpha$ -diversity following SARS-CoV-2 infection across the studies (Supplementary Table 1).

Twenty studies, including raw sequences of the gut, respiratory, oral, and skin microbiota, were included in the meta-analysis of  $\alpha$ -diversity (Supplementary Table 1). Among them, nine studies were related to the gut microbiota, including 265 samples from SARS-CoV-2 infected individuals and 237 samples from non-SARS-CoV-2 infected individuals. Our results showed that the  $\alpha$ -diversity of gut microbiota was significantly down-regulated in SARS-CoV-2 infected individuals compared with non-SARS-CoV-2 individuals (Fixed-effect models; ACE, Chao1, Observed, InvSimpson, Fisher, Shannon, Simpson, and PD indexes,  $P < 0.05$ ), but there were significant heterogeneity among studies (Fig. 2A). Therefore, we used subgroup analysis to analyze the sources of heterogeneity. When the limiting factors were studies ruled out the effect of antibiotics (Simpson index,  $P < 0.001$ , Supplementary Fig. 1A), cross-sectional and case–control studies of published articles (ACE and Fisher indexes,  $P < 0.001$ ; Supplementary Fig. 1B), Chinese studies (Simpson index,  $P < 0.001$ , Supplementary Fig. 1C), American studies (InvSimpson index,  $P < 0.01$ , Supplementary Fig. 1D), and Illumina MiSeq studies (ACE, Chao1, Observed, Fisher, and InvSimpson indexes,  $P < 0.001$ , Supplementary Fig. 1E), and there were no significant heterogeneity among included studies. These results remained consistent when the studies with the largest sample size were removed.



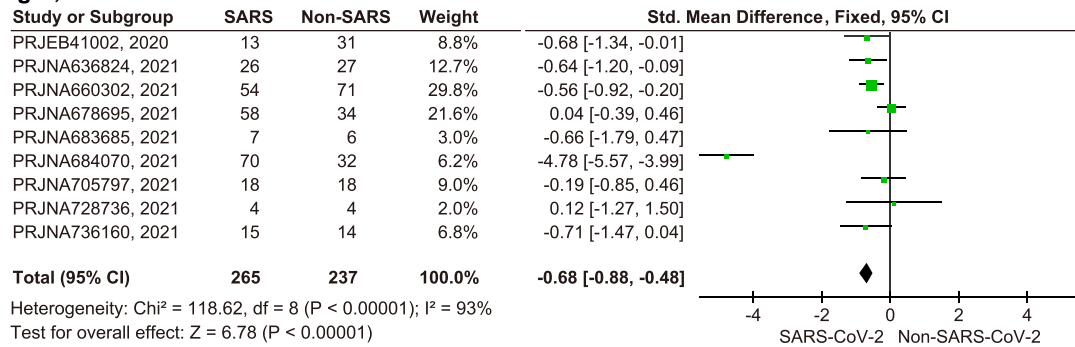
**Fig. 1** Data mining pipeline

There were eleven studies related to the respiratory microbiota, including 1,076 samples from SARS-CoV-2 infected individuals, and 283 samples from non-SARS-CoV-2 infected individuals. Analysis based on fixed-effects models showed that the  $\alpha$ -diversity of respiratory microbiota was significantly down-regulated in SARS-CoV-2 infected individuals compared with non-SARS-CoV-2 infected individuals (ACE, Chao1, PD, and Simpson indexes,  $P < 0.01$ ; significant heterogeneity; Fig. 2B). Subgroup analyses that included 5 identified cross-sectional and case-control studies still showed that SARS-CoV-2 infection was associated with significant down-regulation of respiratory microbiota  $\alpha$ -diversity (PD index,  $P = 0.03$ , no significant heterogeneity, Supplementary Fig. 2A). However, SARS-CoV-2 infection was not associated with  $\alpha$ -diversity when the limiting factor was the V4 sequencing region (Simpson

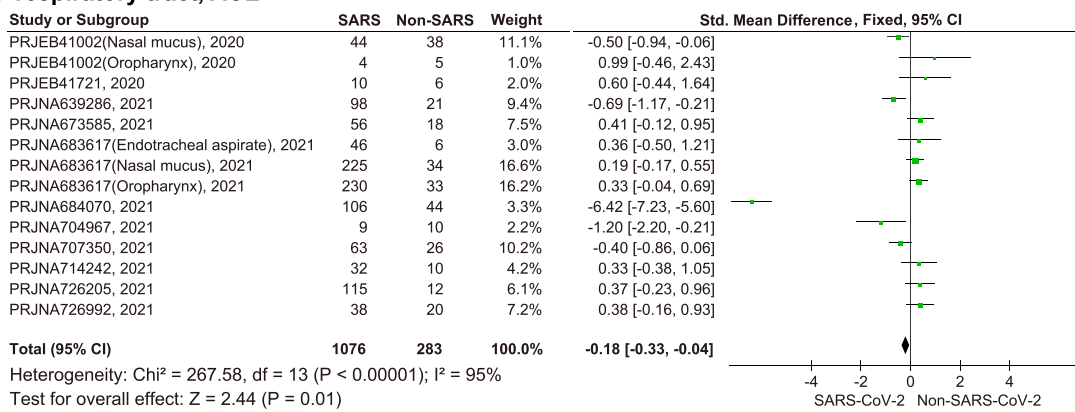
index,  $P = 0.47$ , no significant heterogeneity, Supplementary Fig. 2B). These results remained consistent when the studies with the largest sample size were removed. The results of other limiting factors were still heterogeneous or the results of sensitivity analysis were unstable, or the number of included studies was less than three.

There were three cross-sectional studies related to oral microbiota, including 239 samples from confirmed and convalescent SARS-CoV-2 infected individuals and 218 samples from non-SARS-CoV-2 infected individuals. All three studies ruled out the effect of antibiotics. Fixed-effects models showed that the  $\alpha$ -diversity of oral microbiota were not significantly different in SARS-CoV-2 infected individuals compared with non-SARS-CoV-2 infected individuals (ACE and Chao1 indexes,  $P > 0.05$ ; no significant heterogeneity; Fig. 2C). These results remained

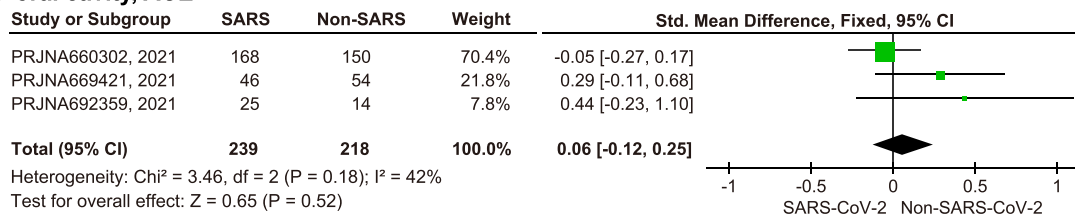
**A gut, ACE**



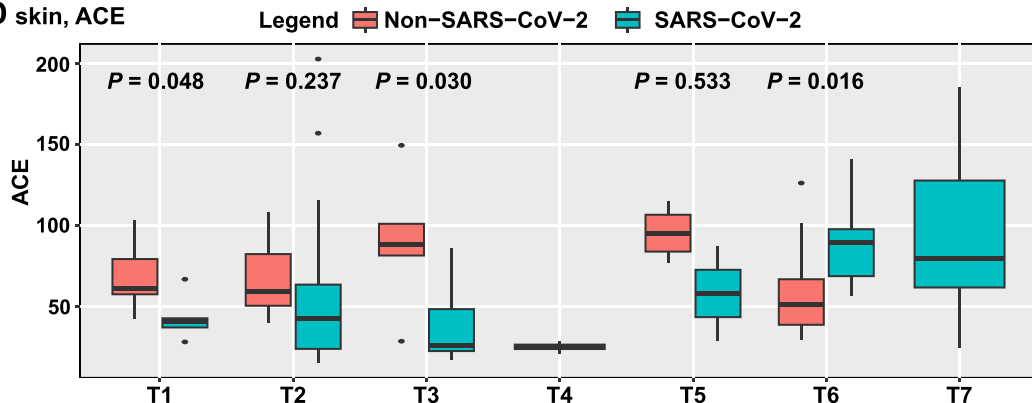
**B respiratory tract, ACE**



**C oral cavity, ACE**



**D skin, ACE**



**Fig. 2** α-diversity analysis (ACE index) of human microbiota between SARS-CoV-2 and non-SARS-CoV-2 infected individuals. The meta-analysis based on the fixed-effects models quantitatively merged studies on gut (A), respiratory (B), and oral (C) studies, respectively. Changes in α-diversity of skin microbiota from observation time point 1 to time point 7 (T1-T7) in SARS-CoV-2 and non-SARS-CoV-2 infected individuals (D)

consistent when the studies with the largest sample size were removed. Due to the small number of studies, the potential confounding factors were not further analyzed.

Moreover, in a longitudinal study of 81 skin samples, the skin microbiota  $\alpha$ -diversity of SARS-CoV-2 infected individuals was significantly lower than that of non-SARS-CoV-2 infected individuals at sampling points 1 ( $P=0.048$ ) and 3 ( $P=0.03$ ). The skin microbiota  $\alpha$ -diversity of SARS-CoV-2 infected individuals decreased first and then increased (Fig. 2D).

#### ML models based on the human respiratory microbiota could predict the prognosis of SARS-CoV-2 infection

Respiratory (nasopharynx and oropharynx) microbiota study PRJNA683617 provides the outcomes, including death and survival, of hospitalized SARS-CoV-2 infected individuals. Based on LEfSe analysis, we identified differential microbial ASV features between dead and surviving SARS-CoV-2 infected individuals at baseline sampling time, and constructed eight ML models based on these features (Table 1). We found that both the nasopharyngeal and oropharyngeal microbiota have predictive potential for the prognosis of SARS-CoV-2 infected individuals. In the ML models constructed by the nasopharyngeal microbiota, SVM showed the best predictive performance, with AUC, sensitivity, and specificity of 0.781, 0.947, and 0.667, respectively. In the ML models constructed by the oropharyngeal microbiota, RF showed the best predictive performance, with AUC, sensitivity, and specificity of 0.847, 0.833, and 0.750, respectively.

#### *Prevotella* and *Streptococcus* were shared differential bacteria in gut, respiratory tract, and oral cavity after SARS-CoV-2 infection

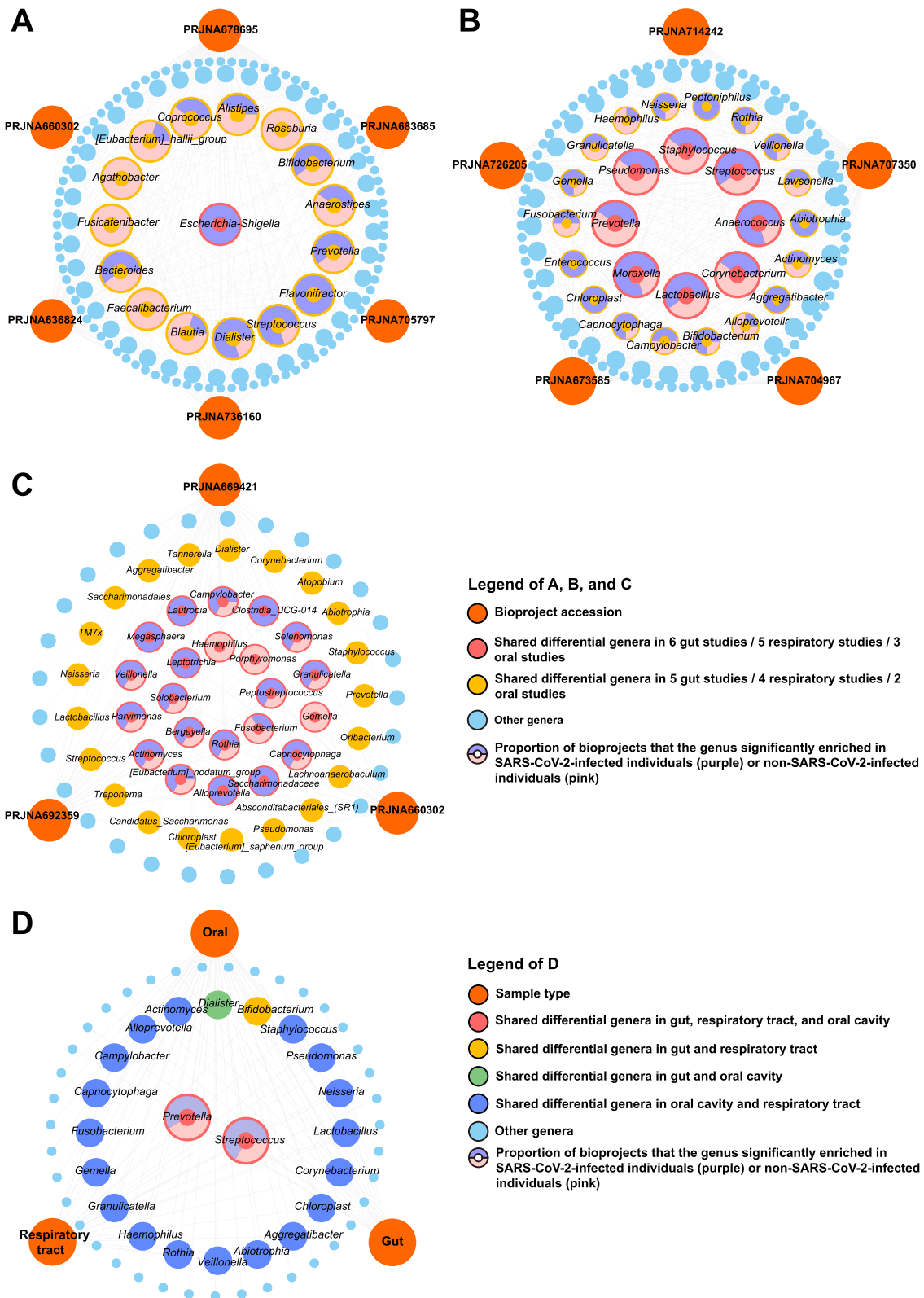
The above studies suggested that the type of study was one of the sources of heterogeneity. Therefore, we first identified the shared differential bacteria in multiple cross-sectional and case-control studies, and then observed their changing characteristics in longitudinal studies. Through LEfSe and Venn network analysis, we found that the abundance of *Escherichia-Shigella* were significantly different between SARS-CoV-2 and non-SARS-CoV-2 infected individuals in six cross-sectional and case-control gut microbiota studies (Fig. 3A). *Anaerococcus*, *Corynebacterium*, *Lactobacillus*, *Moraxella*, *Prevotella*, *Pseudomonas*, *Staphylococcus*, and *Streptococcus* were shared differential genera in five cross-sectional and case-control respiratory microbiota studies (Fig. 3B). There were 22 shared differential genera such as *Alloprevotella*, *Actinomyces*, and *Bergeyella* in three cross-sectional oral microbiota studies (Fig. 3C).

**Table 1** ML models based on the human respiratory microbiota were used to predict the prognosis of SARS-CoV-2 infection

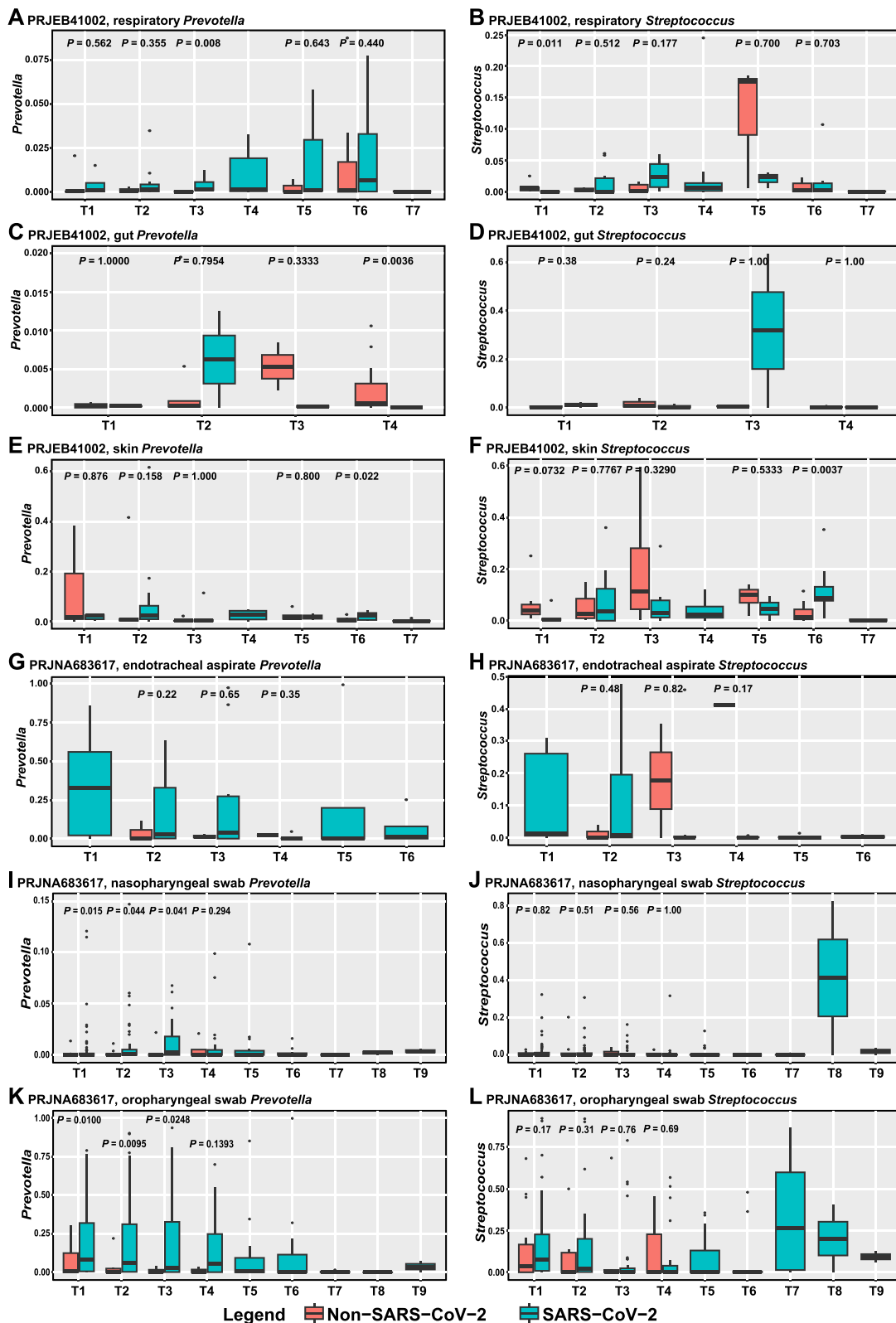
	Models	AUC	Sensitivity	Specificity
Data: the nasopharyngeal microbiota of SARS-CoV-2 infected individuals; Predicting outcomes: survival ( $n=56$ ) and death ( $n=17$ ).	SVM	0.781	0.947	0.667
	RF	0.772	0.947	0.667
	KNN	0.754	0.842	0.667
	LR	0.518	1.000	0.333
	NB	0.632	0.000	1.000
	DT	0.500	1.000	0.000
	NNET	0.509	1.000	0.333
	XGB	0.737	0.842	0.667
Data: the oropharyngeal microbiota of SARS-CoV-2 infected individuals; Predicting outcomes: survival ( $n=56$ ) and death ( $n=16$ ).	SVM	0.688	0.778	0.750
	RF	0.847	0.833	0.750
	KNN	0.75	0.833	0.500
	LR	0.542	0.000	1.000
	NB	0.604	1.000	0.250
	DT	0.5	1.000	0.000
	NNET	0.451	0.667	0.500
	XGB	0.556	0.944	0.500

Further, genera with significant differences in more than half of the studies in each sample type were included in the Venn network. It could be seen that the *Prevotella* and *Streptococcus*, which were shared in the gut, respiratory tract, and oral cavity, were enriched in SARS-CoV-2 infected individuals in more than half of the studies (Fig. 3D).

At multiple time points in the longitudinal study of PRJEB41002, the abundance of *Prevotella* and *Streptococcus* in respiratory tract (Fig. 4A, B), gut (Fig. 4C, D), and skin (Fig. 4E, F) of SARS-CoV-2 infected individuals was higher than that of non-SARS-CoV-2 infected individuals. Similarly, in the respiratory longitudinal study PRJNA683617 (Fig. 4G-L), the trend of *Prevotella* and *Streptococcus* was still to enrich SARS-CoV-2 infected individuals. In addition, we found that *Prevotella* and *Streptococcus* were up-regulated in the oral cavity of confirmed SARS-CoV-2 infected individuals, but down-regulated in confirmed SARS-CoV-2 recovery individuals (Fig. 5A, B). In respiratory study PRJNA714242, *Prevotella* ( $P=0.056$ ) and *Streptococcus* ( $P=0.43$ ) were upregulated in critical SARS-CoV-2 infected individuals (Fig. 5C, D). In respiratory study PRJNA673585, *Prevotella* ( $P=0.029$ ) and *Streptococcus* ( $P=0.2$ ) were also enriched in severe SARS-CoV-2 infected individuals (Fig. 5E, F). However, in respiratory study PRJNA707350, *Prevotella* was significantly down-regulated in symptomatic ( $P<0.01$ ) and asymptomatic ( $P=0.018$ ) SARS-CoV-2 infected individuals (Fig. 5G, H).

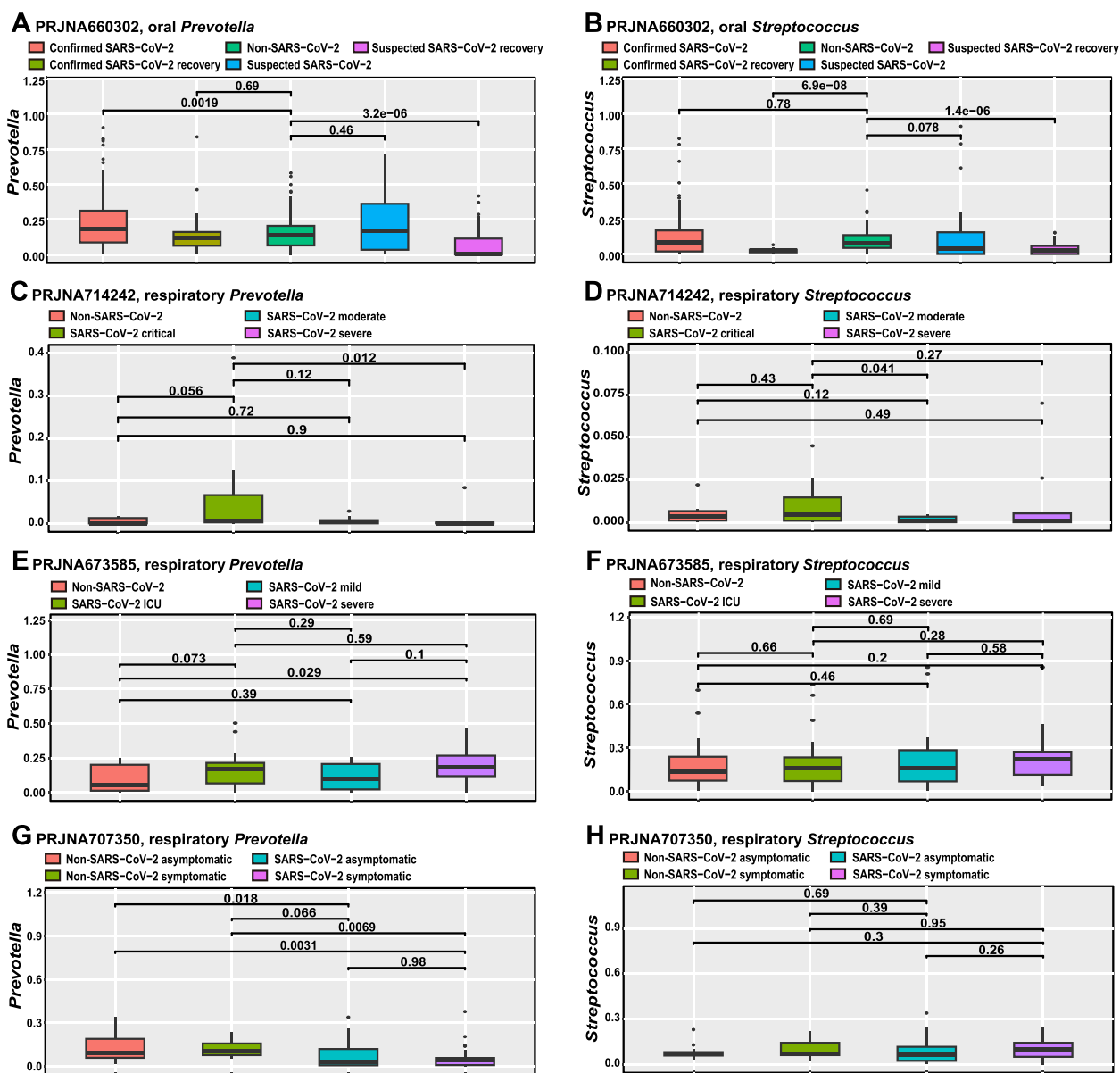


**Fig. 3** Identification of key genera altered after SARS-CoV-2 infection based on LEfSe and venn networks. **A** Venn network of six cross-sectional and case-control studies of gut genera. **B** Venn network of five cross-sectional and case-control studies of respiratory genera. **C** Venn network of three cross-sectional studies of oral genera. **D** Venn network of gut, respiratory, and oral genera



**Fig. 4** The boxplots showed the changes of *Prevotella* and *Streptococcus* in the gut, respiratory tract, and skin of SARS-CoV-2 and non-SARS-CoV-2 infected individuals at different observation time points (A-L)





**Fig. 5** Boxplots of the changes of *Prevotella* and *Streptococcus* in SARS-CoV-2 infected individuals with different infection status (A-H)

**Discussion**

Here, we reviewed the results of studies on the  $\alpha$ -diversity of the human gut, respiratory, oral, and skin microbiota associated with SARS-CoV-2. We assembled the largest dataset available to date in order to assess the relationship between SARS-CoV-2 infection and the  $\alpha$ -diversity of the human microbiota. Our meta-analysis revealed a significant down-regulation of the microbiota  $\alpha$ -diversity in the gut and respiratory systems among individuals with SARS-CoV-2 infection, which is consistent with numerous current findings [8–12, 16, 29–34]. It should be noted that although state-of-the-art and widely used

microbiological analysis software such as QIIME2 and the DADA2 algorithm were employed to mitigate heterogeneity in the processing and analysis of raw sequences from different studies, notable heterogeneity still existed among the included studies. For gut and respiratory studies, we observed that study type was one of the sources of heterogeneity. Cross-sectional and case-control studies typically involved samples collected at a single time point, while longitudinal studies consisted of samples collected at multiple time points. In gut studies, other sources of heterogeneity included antibiotics, country, and sequencing platform. In respiratory studies, sources

of heterogeneity also encompassed sequencing regions. Previous studies have demonstrated that regional factors [35], antibiotic [36], gender [37], age [38], and diet [39] can influence the composition of the human microbiota. Due to limited access to open information included in this study, we were unable to analyze the sources of heterogeneity from additional perspectives. In summary, our study clarified that the  $\alpha$ -diversity of gut and respiratory microbiota is downregulated after SARS-CoV-2 infection, providing readers with an understanding of the microbial characteristics of different human body sites after SARS-CoV-2 infection.

ML based on the human microbiota has been applied to predict various diseases and identify biomarkers. For example, it has been used to predict *Vibrio cholerae* infection [40], ulcerative colitis [41], and more. Similarly, the gut microbiota has shown promise in distinguishing the severity of COVID-19 [18] and effectively predicting protein markers for severe cases [42]. However, it remains to be explored whether the microbiota altered by SARS-CoV-2 infection can predict disease prognosis, including survival and death. In our study, we found that in the early stages of SARS-CoV-2 infection, alterations in the nasopharyngeal and oropharyngeal microbiota had the potential to predict patient survival and death. We observed that the predictive performance differed between the nasopharynx and oropharynx, as well as among different ML models. This suggests that when utilizing human microbiota to predict disease prognosis, we should consider the results from different body parts and ML models comprehensively. In the models we constructed, the AUC of the optimal model was only 0.847. This might be due to the small sample size and changes in microbiota characteristics following treatment for SARS-CoV-2 infection. Nonetheless, our study demonstrated the potential of ML based on human microbiota in predicting the prognosis of SARS-CoV-2-infected individuals, which may help in providing targeted treatment for severely SARS-CoV-2-infected individuals.

A study found that *Dialister invisus* ASV represents a unique case of overlap between the oral and gut microbiota in healthy individuals. Normally, the oral and gut microbiota differ under physiological conditions, and the presence of overlapping microbiota may indicate a certain pathological state [43]. In our study, we discovered genera such as *Prevotella* and *Streptococcus* that overlapped in the gut, respiratory tract, and oral cavity. Several studies have also demonstrated a significant up-regulation of *Prevotella* [44–47], *Streptococcus* [9, 44–46] and *Veillonella* [9, 44–46] following SARS-CoV-2 infection. *Prevotella*, a strictly anaerobic gram-negative bacillus, is known to be a major genus found in human

skin, oral cavity, vagina, and gut [48]. It is frequently associated with respiratory tract infections, such as inhalation pneumonia [49] and pulmonary empyema [50]. Additionally, studies [51] have shown an increased abundance of *Prevotella* in the presence of viral infections associated with Human Immunodeficiency Virus, Papillomavirus, Herpesviridae, and respiratory viruses. Our study confirmed dysregulation of *Prevotella* in the human skin, oral cavity, gut, and respiratory tract after SARS-CoV-2 infection. Furthermore, *Prevotella* was found to be related to the severity and recovery of SARS-CoV-2 infection. Other studies have reported a correlation between long-lasting COVID-19 symptoms and elevated expression of oral *Prevotella* [52], which may be due to the ability of *Prevotella* to produce proteins that promote SARS-CoV-2 infection [53]. The precise mechanism by which *Prevotella* affects COVID-19 is not yet clear. However, previous research [54] has revealed that certain *Prevotella* strains can produce virulence factors that increase inflammatory response by activating Toll-like receptor 2 and inducing Th17-polarizing cytokines in antigen-presenting cells (such as IL-23 and IL-1), or stimulating epithelial cells to produce IL-8, IL-6, and CCL20. In summary, when individuals become ill due to the invasion of foreign pathogens, the normal human microbiota may be translocated and transformed into pathogenic bacteria, exacerbating the disease. The key bacteria *Prevotella* and *Streptococcus* proposed by us provide clues for future animal and in vitro experiments on SARS-CoV-2 infection intervention.

Our study has the following limitations: Firstly, due to the continuous updating of databases, our research may not reflect the latest research status. Secondly, the limited number of studies included in our analysis leads to certain limitations in generalizing our findings. Thirdly, there is a lack of animal and in vitro experimental validation for the key microbial communities we identified.

## Conclusions

Our study showed that SARS-CoV-2 infection was related to the down-regulation of  $\alpha$ -diversity in the human gut and respiratory microbiota. The respiratory microbiota had the potential to predict the prognosis of individuals infected with SARS-CoV-2. *Prevotella* and *Streptococcus* might be key microbiota in SARS-CoV-2 infection.

## Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12879-023-08784-x>.

**Additional file 1: Supplementary Table 1.** Characteristics of included studies.

**Additional file 2: Supplementary Figure 1.** Subgroup analysis of SARS-CoV-2 infection-associated gut microbiota studies. Forest plots for limiting

factors were studies ruled out the effect of antibiotics (A), cross-sectional and case-control studies (B), Chinese studies (C), American studies (D), and Illumina MiSeq studies (E). **Supplementary Figure 2.** Subgroup analysis of SARS-CoV-2 infection-associated respiratory microbiota studies. Forest plots for limiting factors were cross-sectional and case-control studies (A), V4 sequencing studies (B).

### Acknowledgements

Not applicable.

### Authors' contributions

Conceptualization, J.H. and L.Y.; methodology, J.H., J.Z., and X.Y.; software, J.Z., X.Y.; validation, J.H., L.Y., and P.C.; formal analysis, J.Z., X.Y.; investigation, Y.Y., Y.W., D.L., and Y. X.; resources, network database; data curation, J.Z., X.Y.; writing—original draft preparation, J.Z., X.Y., Y.Y., Y.W., and D.L.; writing—review and editing, J.Z., X.Y., J.H., L.Y., P.C., and H.L.; visualization, J.Z., X.Y.; supervision, J.H.; project administration, J.H. and L.Y.; funding acquisition, J.H., L.Y., P.C. and J.Z. All authors have read and agreed to the published version of the manuscript.

### Funding

This research was funded by the National Natural Science Foundation of China (NSFC, 82060366, 82002134, 31970167) and the Innovation Project of Guangxi Graduate Education (No. YCBZ2022098).

### Availability of data and materials

The datasets used in the current study can be obtained from NCBI according to the Bioproject accession number mentioned in Supplementary Table 1.

### Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

#### Author details

<sup>1</sup>Guangxi Key Laboratory of AIDS Prevention and Treatment & School of Public Health, Guangxi Medical University, Nanning, Guangxi, China. <sup>2</sup>Guangxi Universities Key Laboratory of Prevention and Control of Highly Prevalent Disease, Nanning, Guangxi, China. <sup>3</sup>Life Science Institute, Guangxi Medical University, Nanning, Guangxi, China. <sup>4</sup>Collaborative Innovation Centre of Regenerative Medicine and Medical BioResource Development and Application Co-Constructed By the Province and Ministry, Nanning, Guangxi, China.

Received: 26 March 2023 Accepted: 2 November 2023

Published online: 29 November 2023

### References

- Wu Z, McGoogan JM. Characteristics of and Important Lessons From the Coronavirus Disease 2019 (COVID-19) Outbreak in China: Summary of a Report of 72314 Cases From the Chinese Center for Disease Control and Prevention. *JAMA*. 2020;323(13):1239–42.
- Kesheh MM, Hosseini P, Soltani S, Zandi M. An overview on the seven pathogenic human coronaviruses. *Rev Med Virol*. 2022;32(2): e2282.
- Zandi M. ORF9c and ORF10 as accessory proteins of SARS-CoV-2 in immune evasion. *Nat Rev Immunol*. 2022;22(5):331–331.
- Hernandez-Teran A, Mejia-Nepomuceno F, Herrera MT, Barreto O, Garcia E, Castillejos M, Boukadida C, Matias-Florentino M, Rincon-Rubio A, Avila-Rios S, et al. Dysbiosis and structural disruption of the respiratory microbiota in COVID-19 patients with severe and fatal outcomes. *Sci Rep*. 2021;11(1):21297.
- Mizutani T, Ishizaka A, Koga M, Ikeuchi K, Saito M, Adachi E, Yamayoshi S, Iwatsuki-Horimoto K, Yasuhara A, Kiyono H et al. Correlation Analysis between Gut Microbiota Alterations and the Cytokine Response in Patients with Coronavirus Disease during Hospitalization. *Microbiol Spectr*. 2022;10(2):e0168921.
- Buttenschon J, Vogt S, Mattner J. Compartmentalized immune responses and the local microbiota determine mucosal and systemic immunity against SARS-CoV-2. *Cell Mol Immunol*. 2022;19(2):130–2.
- Xu R, Liu P, Zhang T, Wu Q, Zeng M, Ma Y, Jin X, Xu J, Zhang Z, Zhang C. Progressive deterioration of the upper respiratory tract and the gut microbiomes in children during the early infection stages of COVID-19. *J Genet Genomics*. 2021;48(9):803–14.
- Mazzarelli A, Giancola ML, Farina A, Marchioni L, Rueca M, Gruber CEM, Bartolini B, Ascoli Bartoli T, Maffongelli G, Capobianchi MR, et al. 16S rRNA gene sequencing of rectal swab in patients affected by COVID-19. *PLoS ONE*. 2021;16(2): e0247041.
- Gu S, Chen Y, Wu Z, Chen Y, Gao H, Lv L, Guo F, Zhang X, Luo R, Huang C, et al. Alterations of the Gut Microbiota in Patients With Coronavirus Disease 2019 or H1N1 Influenza. *Clin Infect Dis*. 2020;71(10):2669–78.
- Tao W, Zhang G, Wang X, Guo M, Zeng W, Xu Z, Cao D, Pan A, Wang Y, Zhang K, et al. Analysis of the intestinal microbiota in COVID-19 patients and its correlation with the inflammatory factor IL-18. *Med Microecol*. 2020;5: 100023.
- Lv LX, Gu SL, Jiang HY, Yan R, Chen YF, Chen YB, Luo R, Huang CJ, Lu HF, Zheng BW et al. Gut mycobiota alterations in patients with COVID-19 and H1N1 infections and their associations with clinical features. *Commun Biol*. 2021;4(1):480.
- Chen Y, Gu S, Chen Y, Lu H, Shi D, Guo J, Wu WR, Yang Y, Li Y, Xu KJ, et al. Six-month follow-up of gut microbiota richness in patients with COVID-19. *Gut*. 2022;71(1):222–5.
- Zhou Y, Zhang J, Zhang D, Ma WL, Wang X. Linking the gut microbiota to persistent symptoms in survivors of COVID-19 after discharge. *J Microbiol*. 2021;59(10):941–8.
- Newsome RC, Gauthier J, Hernandez MC, Abraham GE, Robinson TO, Williams HB, Sloan M, Owings A, Laird H, Christian T, et al. The gut microbiome of COVID-19 recovered patients returns to uninfected status in a minority-dominated United States cohort. *Gut Microbes*. 2021;13(1):1–15.
- Hurst JH, McCumber AW, Aquino JN, Rodriguez J, Heston SM, Lugo DJ, Rotta AT, Turner NA, Pfeiffer TS, Gurley TC, et al. Age-Related Changes in the Nasopharyngeal Microbiome Are Associated With Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV-2) Infection and Symptoms Among Children, Adolescents, and Young Adults. *Clin Infect Dis*. 2022;75(1):e928–37.
- Rueca M, Fontana A, Bartolini B, Piselli P, Mazzarelli A, Copetti M, Binda E, Perri F, Gruber CEM, Nicastrì E et al. Investigation of Nasal/Oropharyngeal Microbial Community of COVID-19 Patients by 16S rDNA Sequencing. *Int J Environ Res Public Health*. 2021;18(4):2174.
- Ventero MP, Cuadrat RRC, Vidal I, Andrade BGN, Molina-Pardines C, Haro-Moreno JM, Coutinho FH, Merino E, Regitano LCA, Silveira CB, et al. Nasopharyngeal Microbial Communities of Patients Infected With SARS-CoV-2 That Developed COVID-19. *Front Microbiol*. 2021;12: 637430.
- Nagata N, Takeuchi T, Masuoka H, Aoki R, Ishikane M, Iwamoto N, Sugiyama M, Suda W, Nakanishi Y, Terada-Hirashima J, et al. Human Gut Microbiota and Its Metabolites Impact Immune Responses in COVID-19 and Its Complications. *Gastroenterology*. 2023;164(2):272–88.
- Sarkar A, Harty S, Moeller AH, Klein SL, Erdman SE, Friston KJ, Carmody RN. The gut microbiome as a biomarker of differential susceptibility to SARS-CoV-2. *Trends Mol Med*. 2021;27(12):1115–34.
- Khan M, Mathew BJ, Gupta P, Garg G, Khadanga S, Vyas AK, Singh AK. Gut Dysbiosis and IL-21 Response in Patients with Severe COVID-19. *Microorganisms*. 2021;9(6):1292.
- Moreira-Rosario A, Marques C, Pinheiro H, Araujo JR, Ribeiro P, Rocha R, Mota I, Pestana D, Ribeiro R, Pereira A, et al. Gut Microbiota Diversity and C-Reactive Protein Are Predictors of Disease Severity in COVID-19 Patients. *Front Microbiol*. 2021;12: 705020.
- Estaki M, Jiang L, Bokulich NA, McDonald D, Gonzalez A, Kosciolk T, Martino C, Zhu Q, Birmingham A, Vazquez-Baeza Y, et al. QIIME 2 enables comprehensive End-to-End analysis of diverse microbiome data and comparative studies with publicly available data. *Curr Protoc Bioinformatic*. 2020;70(1): e100.

23. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: High-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13(7):581–3.
24. Katoh K, Misawa K, Kuma K, Miyata T. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res*. 2002;30(14):3059–66.
25. Price MN, Dehal PS, Arkin AP. FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS ONE*. 2010;5(3): e9490.
26. McDonald D, Price MN, Goodrich J, Nawrocki EP, DeSantis TZ, Probst A, Andersen GL, Knight R, Hugenholtz P. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. *ISME J*. 2012;6(3):610–8.
27. Bokulich NA, Kaehler BD, Rideout JR, Dillon M, Bolyen E, Knight R, Huttley GA, Caporaso JG. Optimizing taxonomic classification of marker-gene amplicon sequences with QIIME 2’s q2-feature-classifier plugin. *Microbiome*. 2018;6(1):90.
28. Liu C, Cui Y, Li X, Yao M: microeco: an R package for data mining in microbial community ecology. *FEMS Microbiol Ecol* 2021, 97(2).
29. Tian Y, Sun KY, Meng TQ, Ye Z, Guo SM, Li ZM, Xiong CL, Yin Y, Li HG, Zhou LQ. Gut microbiota may not be fully restored in recovered COVID-19 patients after 3-month recovery. *Front Nutr*. 2021;8: 638825.
30. Kim HN, Joo EJ, Lee CW, Ahn KS, Kim HL, Park DI, Park SK. Reversion of Gut Microbiota during the Recovery Phase in Patients with Asymptomatic or Mild COVID-19: Longitudinal Study. *Microorganisms*. 2021;9(6):1237.
31. Gaibani P, D’Amico F, Bartoletti M, Lombardo D, Rampelli S, Fornaro G, Coladonato S, Siniscalchi A, Re MC, Viale P, et al. The Gut Microbiota of Critically Ill Patients With COVID-19. *Front Cell Infect Microbiol*. 2021;11: 670424.
32. Gupta A, Karyakarte R, Joshi S, Das R, Jani K, Shouche Y, Sharma A. Nasopharyngeal microbiome reveals the prevalence of opportunistic pathogens in SARS-CoV-2 infected individuals and their association with host types. *Microbes Infect*. 2022;24(1): 104880.
33. Merenstein C, Liang G, Whiteside SA, Cobian-Guemes AG, Merlino MS, Taylor LJ, Glascock A, Bittinger K, Tanes C, Graham-Wooten J, et al. Signatures of COVID-19 Severity and Immune Response in the Respiratory Tract Microbiome. *mBio*. 2021;12(4):e0177721.
34. Ren Z, Wang H, Cui G, Lu H, Wang L, Luo H, Chen X, Ren H, Sun R, Liu W, et al. Alterations in the human oral and gut microbiomes and lipidomics in COVID-19. *Gut*. 2021;70(7):1253–65.
35. Schnorr SL, Candela M, Rampelli S, Centanni M, Consolandi C, Basaglia G, Turroni S, Biagi E, Peano C, Severgnini M, et al. Gut microbiome of the Hadza hunter-gatherers. *Nat Commun*. 2014;5:3654.
36. Zaura E, Brandt BW, Teixeira de Mattos MJ, Buijs MJ, Caspers MP, Rashid MU, Weintraub A, Nord CE, Savell A, Hu Y, et al. Same Exposure but Two Radically Different Responses to Antibiotics: Resilience of the Salivary Microbiome versus Long-Term Microbial Shifts in Feces. *mBio*. 2015;6(6):e01693–15.
37. Tuddenham SA, Koay WLA, Zhao N, White JR, Ghanem KG, Sears CL. The Impact of Human Immunodeficiency Virus Infection on Gut Microbiota  $\alpha$ -Diversity: An Individual-Level Meta-analysis. *Clin Infect Dis*. 2020;70(4):615–27.
38. Maynard C, Weinkove D. The Gut Microbiota and Ageing. *Subcell Biochem*. 2018;90:351–71.
39. David LA, Maurice CF, Carmody RN, Gootenberg DB, Button JE, Wolfe BE, Ling AV, Devlin AS, Varma Y, Fischbach MA, et al. Diet rapidly and reproducibly alters the human gut microbiome. *Nature*. 2014;505(7484):559–63.
40. Midani FS, Weil AA, Chowdhury F, Begum YA, Khan AI, Debela MD, Durand HK, Reese AT, Nimmagadda SN, Silverman JD, et al. Human Gut Microbiota Predicts Susceptibility to *Vibrio cholerae* Infection. *J Infect Dis*. 2018;218(4):645–53.
41. Barberio B, Facchin S, Patuzzi I, Ford AC, Massimi D, Valle G, Sattin E, Simionati B, Bertazzo E, Zingone F, et al. A specific microbiota signature is associated to various degrees of ulcerative colitis as assessed by a machine learning approach. *Gut Microbes*. 2022;14(1):2028366.
42. Gou W, Fu Y, Yue L, Chen GD, Cai X, Shuai M, Xu F, Yi X, Chen H, Zhu Y, et al. Gut microbiota, inflammation, and molecular signatures of host response to infection. *J Genet Genomics*. 2021;48(9):792–802.
43. Rashidi A, Ebadi M, Weisdorf DJ, Costalonga M, Staley C. No evidence for colonization of oral bacteria in the distal gut in healthy adults. *P Natl Acad Sci USA*. 2021;118(42):e21114152118.
44. Gupta A, Bhanushali S, Sanap A, Shekatkar M, Kharat A, Raut C, Bhonde R, Shouche Y, Kheur S, Sharma A. Oral dysbiosis and its linkage with SARS-CoV-2 infection. *Microbiol Res*. 2022;261: 127055.
45. Rattanaburi S, Sawaswong V, Chitcharoen S, Sivapornnukul P, Nimsamer P, Suntronwong N, Puenpa J, Poovorawan Y, Payungporn S. Bacterial microbiota in upper respiratory tract of COVID-19 and influenza patients. *Exp Biol Med (Maywood)*. 2022;247(5):409–15.
46. Liu J, Liu S, Zhang Z, Lee X, Wu W, Huang Z, Lei Z, Xu W, Chen D, Wu X, et al. Association between the nasopharyngeal microbiome and metabolome in patients with COVID-19. *Synthetic and systems biotechnology*. 2021;6(3):135–43.
47. Ventero MP, Cuadrat RRC, Vidal I, Andrade BGN, Molina-Pardines C, Haro-Moreno JM, Coutinho FH, Merino E, Regitano LCA, Silveira CB, et al. Nasopharyngeal Microbial Communities of Patients Infected With SARS-CoV-2 That Developed COVID-19. *Frontiers In Microbiology*. 2021;12:637430.
48. Tett A, Pasolli E, Masetti G, Ercolini D, Segata N. Prevotella diversity, niches and interactions with the human host. *Nat Rev Microbiol*. 2021;19(9):585–99.
49. Fujita K, Takata I, Sugiyama H, Suematsu H, Yamagishi Y, Mikamo H. Antimicrobial susceptibilities of clinical isolates of the anaerobic bacteria which can cause aspiration pneumonia. *Anaerobe*. 2019;57:86–9.
50. Zhou H, Shen Y, Shen Q, Zhou J: Thoracic empyema caused by *Prevotella* spp. diagnosed using 16S rDNA sequence analysis. *Clin Respir J* 2015, 9(1):121–124.
51. Tamanai-Shacoori Z, Le Gall-David S, Moussouni F, Sweidan A, Polard E, Bousarghin L, Jolivet-Gougeon A: SARS-CoV-2 and *Prevotella* spp.: friend or foe? A systematic literature review. *J Med Microbiol*. 2022;71(5):001520.
52. Haran JP, Bradley E, Zeamer AL, Cincotta L, Salive MC, Dutta P, Mutaawe S, Anya O, Meza-Segura M, Moormann AM, et al. Inflammation-type dysbiosis of the oral microbiome associates with the duration of COVID-19 symptoms and long COVID. *JCI Insight*. 2021;6(20):e152346.
53. Khan AA, Khan Z. COVID-2019-associated overexpressed *Prevotella* proteins mediated host-pathogen interactions and their role in coronavirus outbreak. *Bioinformatics*. 2020;36(13):4065–9.
54. Larsen JM. The immune response to *Prevotella* bacteria in chronic inflammatory disease. *Immunology*. 2017;151(4):363–74.

## Publisher's Note

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

### Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

