Alterations of Oropharyngeal Microbiome Reflected Disease Severity and Long-Term Dysbiosis in Patients With COVID-19

Yongzhao Zhou
Sichuan University West China Hospital
Sifen Lu
Sichuan University West China Hospital
Mengjiao Li
Sichuan University West China Hospital
Minjin Wang
Sichuan University West China Hospital
Jinmin Ma
BGI-Shenzhen: BGI Group
Yuanchen Mao
BGI-Shenzhen: BGI Group
Zhongyi Zhu
BGI-Shenzhen: BGI Group
Jiumeng Min
BGI-Shenzhen: BGI Group
Jun Xiao
Sichuan University West China Hospital
Rong Qiu
Suining Central Hospital
Taibing Deng
Guang'an People's Hospital
Li Jiang
Chuanbei Medical College: North Sichuan Medical University
Jing Zhu
Mianyang Central Hospital
Yilan Zeng
public health clinical center of Chengdu
Dan Liu
Sichuan University West China Hospital
weimin li (weimi003@scu.edu.cn)

Sichuan University West China Hospital https://orcid.org/0000-0003-0985-0311

Binwu Ying

Sichuan University West China Hospital

Research

Keywords: oropharynx, microbiome, metatranscriptomics, COVID-19, dysbiosis

Posted Date: May 6th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-490714/v1

Alterations of oropharyngeal microbiome reflected disease severity and long-term dysbiosis in patients with COVID-19

Yongzhao Zhou[1#] Sifen Lu[2#] Mengjiao Li[3#] Minjin Wang[3#] Jinmin Ma[4] Yuanchen Mao[5] Zhongyi Zhu[5,12] Jiumeng Min[5] Jun Xiao[1,6] Rong Qiu[7] Taibing Deng[8] Li Jiang[9] Jing Zhu[10] Yilan Zeng[11] Dan Liu[1*] Binwu Ying[3*] Weimin Li[1,13*]

*Correspondence: **Dan Liu, Binwu Ying, Weimin Li.**

[#] **Yongzhao Zhou, Sifen Lu, Mengjiao Li, Minjin Wang** contributed equally to this work.

Email addresses for the authors：**Yongzhao Zhou**：yongzhaozhou@scu.edu.cn；

Sifen Lu：lusifen@wchscu.cn；**Mengjiao Li**：limengjiao189@163.com；

Minjin Wang: wang.minjin@look.com；**Jinmin Ma**：majinmin@genomics.cn；

Yuanchen Mao：maoyuanchen@bgi.com; **Zhongyi Zhu:** zhuzhongyi@bgi.com;

Jiumeng Min: minjm@bgi.com; **Jun Xiao:** xiaojun5244@163.com;

Rong Qiu: 443182163@qq.com; **Taibing Deng:** 27180619@qq.com;

Li Jiang: 78389789@qq.com; **Jing Zhu:** zhujing_1126@126.com;

Yilan Zeng: zhyilan991827@163.com; **Dan Liu:** liudan10965@wchscu.cn;

Binwu Ying: yingbinwu@scu.edu.cn; **Weimin Li:** weimi003@scu.edu.cn.

1Department of Respiratory and Critical Care Medicine, Frontier Science Center of Disease Molecular Network, West China Hospital, Sichuan University, Chengdu 610041, China

2 Precision Medicine Key Laboratory of Sichuan Province and Precision Medicine Center, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China

3Department of Laboratory Medicine, West China Hospital, Sichuan University, Chengdu, Sichuan 610041, China

4BGI-Shenzhen, Shenzhen, 518083, China

5BGI PathoGenesis Pharmaceutical Technology, BGI-Shenzhen, Shenzhen 518083, China

6Department of Respiratory and Critical Care Medicine, People's Hospital of Ganzi Tibetan Autonomous Prefecture, Ganzi Tibetan Autonomous Prefecture 626000, China

7Department of Respiratory and Critical Care Medicine, Suining Central Hospital, Suining 629000, China

8Department of Respiratory and Critical Care Medicine, Guang'an people's Hospital, Guang'an 638000, China

9Department of Respiratory and Critical Care Medicine, Affiliated Hospital of North Sichuan Medical College, Nanchong 637000, China

10Department of Respiratory and Critical Care Medicine, Mianyang Central Hospital, Mianyang 621000, China

11Department of general medicine, Public Health Clinical Center of Chengdu, Chengdu 610041, China

12 College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

13 Lead contact

Abstract

Background: Recent evidences have shown that gut microbiome of patients with COVID-19 significantly changes and can reflect the severity of the disease. And gut microbiota richness was not restored to normal levels after 6-month recovery. However, *SARS-CoV-2* primarily infects the respiratory tract, few studies investigate whether the alterations of oropharyngeal microbiome is

associated with disease severity in patients with COVID-19, and whether interferences in microbiome composition, if any, eliminate with clearance of the *SARS-CoV-2* virus. We employed metatranscriptomic sequencing to analyse oropharyngeal swabs collected within a week of diagnosis COVID-19 (period of disease group: PDG) and two months after clearance of the *SARS-CoV-2* virus (convalescent group: CG) from 47 patients with COVID-19. Meanwhile, oropharyngeal swabs from 40 healthy subjects were analyzed as healthy control group (HCG).

Results:

Oropharyngeal microbial composition was significantly altered in patients with COVID-19 compared with healthy controls even two months after clearance of the *SARS-CoV-2* virus. Little changes in α-diversity among HCG, PDG and CG ($P > 0.05$), but obviously changes in β-diversity among them. Notably, *Prevotella* increased significantly in PDG than that in HCG (Wilcoxon rank-sum test, $P < 0.001$) and increased gradually along with the severity of patients with COVID-19 aggravated. There was a positive correlation between *Prevotella* and the elevation of Neutrophil percentage ($R = 0.301$, $P = 0.040$). Similarly, *SARS-CoV-2* and *Aspergillus* increased remarkably in critical Patients with COVID-19. There was a negative correlation between *SARS-CoV-2* viral load and platelet counts ($R = -0.330$, $P = 0.022$).

Conclusions: The oropharyngeal microbiome in patients with COVID-19 present persistent dysbiosis even two months after clearance of the *SARS-*

CoV-2 virus. Furthermore, alterations in oropharyngeal microbial composition reflected the severity of disease in patients with COVID-19. Our findings underscore that there is an urgent need to understand the specific roles of oropharyngeal microorganisms in COVID-19 disease progression and rehabilitation.

Keywords: oropharynx, microbiome, metatranscriptomics, COVID-19, dysbiosis

Background

The COVID-19 pandemic is a once-in-a-century public health emergency, the causative pathogen was identified as a virus, named severe acute respiratory syndrome coronavirus 2 (*SARS-CoV-2*)[1]. Although caused only by the *SARS-CoV-2* virus, COVID-19 is likely affected by the rest of the microbiome as evidenced by viral persistence in the gut, and risk factors associated with gut microbial (include bacteria and fungi) dysbiosis[2-5]. Indeed, there is a study suggesting that the disease severity is concordant with composition alteration of the gut microbiome in patients with COVID-19, and the dysbiotic gut microbiota composition in patients with COVID-19 persists after clearance of the virus[6]. Besides, another study found that microbiota richness was not restored to normal levels after 6-month recovery[7]. Recently, Microbiome Centers Consortium has called for researchers try to figure out the connections between the microbiome and COVID-19[2]. *SARS-CoV-2* primarily infects the respiratory tract. Study has already shown that low respiratory tract microbiota in SARS–CoV-2 infected patients were dominated with elevated levels of the upper respiratory tract (URT) commensal microbiota[8], unlike *influenza* virus[9-12], few studies have explored the relationship between *SARS-CoV-2* infection and the oropharyngeal microbiome. Therefore, we investigated whether the alterations of oropharyngeal microbiome are associated with disease severity in patients with COVID-19, and whether interferences in microbiome composition, if any, eliminate with clearance of the *SARS-CoV-2* virus.

Results

Study cohort

We recruited 47 patients with COVID-19 confirmed by positive *SARS-CoV-2* quantitative reverse transcription PCR(RT-qPCR) to establish a longitudinal cohort study. And we collected oropharyngeal swabs at two time points (Within a week of diagnosis COVID-19 (period of disease group: PDG) and two months after clearance of the *SARS-CoV-2* virus (convalescent group: CG)) during February and June 2020. The subjects were classified into mild, moderate, severe, and critical groups based on symptoms as reported before[13, 14]. As healthy controls group (HCG), we collected oropharyngeal swabs from 40 healthy human subjects free of any respiratory infections during the same timeframe. Clinical information and demographic information were presented in **Table 1**.

Table 1 Clinical characteristics of participants

	COVID-19 patients (period of disease group / convalescent group)※	healthy controls
Number of subjects	47	40
Gender (females, males)	27, 20	20, 20
Age, years (mean±SD)	44.7±15.3 years	45.8±16.4 years
Disease severity category	5 mild, 25 moderate, 10 severe, 7 critical	NA
Symptoms at admission, n (%)		
Fever	29(61.7%)	NA
Cough	29(61.7%)	NA
Sputum	19(40.4%)	NA
Sore throat	7(14.9%)	NA
Shortness of breath	17(36.2%)	NA

Received antibiotics during first week hospitalisation by disease severity, n (%)[1]		NA
Mild disease	1(2.1%)	NA
moderate disease	3(6.4%)	NA
Severe disease	5(10.5%)	NA
critical disease	7(14.8%)	NA
Received antivirals during first week hospitalisation by disease severity, n (%)[2]		
Mild disease	4(8.5%)	NA
moderate disease	24(51.0%)	NA
Severe disease	10(21.3%)	NA
critical disease	7(14.8%)	NA

※ The 47 patients with COVID-19 were longitudinal observed in period of disease and two months after clearance of the SARS-CoV-2 virus.

[1] The antibiotics included quinolones and cephalosporins.

[2] The antivirals included Lopinavir/ritonavir, Ribavirin, Oseltamivir and interferon.

The persist dysbiosis of oropharyngeal microbiome in patients with COVID-19 even two months after clearance of the *SARS-CoV-2* virus

We employed metatranscriptomic sequencing in all the 134 oropharyngeal swabs to obtain unbiased microbial profiles. In order to remove possible contamination, blank control samples, including another 6 swabs not sampled were tested. These blank control samples were applied to extract microbial genomic nucleotide and were tested using the same batch of consumables and reagents in the same equipment and laboratory. Then we analyzed the unbiased metatranscriptome sequencing data of 140 samples using a database[15] (built by updating all the complete genomes from NCBI Refseq database in real time) consisting of 4,409 genera (37,933 species) of bacteria, fungi, and viruses for taxonomic classification.

The persist oropharyngeal bacterial dysbiosis: Overall, 534 unique genera of bacteria were detected in oropharyngeal swabs from HCG, PDG and CG.

There were 73 genera with relative abundance greater than 0.01% in 534 genera. The relative abundances of the most top20 abundant bacterial profiles in oropharyngeal swabs from HCG, PDG and CG were shown as **Fig.1a, b, c,** and top20 abundant bacterial profiles of every subjects were shown as **Suppl.Fig.1a, b, c**. Among the 534 genera, the PDG, CG and HCG groups shared with 244 genera and had 63, 36 and 61 unique genera, respectively (**Fig.1e**); Among the 244 shared genera, *Prevotella* were dominant genera of the PDG and CG **(Fig.1f)**. Consistent with previous researches[16, 17], the major bacterial genera in HCG were *Actinomyces* (12.1%), *Acetobacter* (10.9%), *Rothia* (10.4%), *Bacteroides* (9.5%), *Prevotella* (8.8%) (**Fig.1a**). *Actinomyces* and *Rothia* decreased remarkably in group PDG and CG (**Fig.1d**). *Prevotella* (18.3%), *Neisseria* (11.0%), *Alloprevotella* (5.5%), and *Fusobacterium* (4.1%) in PDG were significantly higher than that in HCG (**Fig.1b and Fig.1d**). Two months after *SARS-CoV-2* virus has been cleared, *Prevotella* (19.4%), *Neisseria* (14.2%), *Alloprevotella* (6.9%), and *Fusobacterium* (5.1%) in CG were even higher than that in HCG or in PDG (**Fig.1a, b, c and d**). These results indicated that the bacterial microecology of oropharynx was obviously disturbed by *SARS-CoV-2* infection, and it was hard to get back to normal even two months after the virus has been cleared.

The persist oropharyngeal fungal dysbiosis: Overall, 103 unique genera of fungi were detected in oropharyngeal swabs from HCG, PDG and CG. There were 57 genera with relative abundance greater than 0.01% in 103 genera. The

relative abundances of the most top20 abundant fungal profiles in oropharyngeal swabs from HCG, PDG and CG were shown as **Fig.2a, b, c**. top20 abundant fungal profiles of every subjects were shown as **Suppl.Fig.2a, b, c**. The PDG, CG and HCG shared with 40 genera and had 12, 11 and 6 unique genera, respectively (**Fig.2e**); Among the 40 shared genera, *Aspergillus* were dominant genera of the PDG and CG **(Fig.2f)**. The major genera in HCG were *Malassezia* (16.4%), *Aspergillus* (14.8%), *Clavispora* (14.1%), *Daldinia* (13.6%), *Kwoniella* (7.7%) (**Fig.2a**). Compared with HCG, *Malassezia*, *Clavispora* and *Pseudogymnoascus* decreased remarkably in PDG and CG (**Fig.2d**). *Aspergillus* (20.5%) and *Saccharomyces* (15.7%) in PDG were significantly higher than that in HCG (**Fig.2b, d**). Two months after *SARS-CoV-2* virus has been cleared, *Aspergillus* (20.6%), *Trametes* (10.6%), *Yarrowia* (4.8%), and *Komagataella* (3.6%) in CG were even higher than that in HCG or PDG (**Fig.2a, b, c and d**). As bacterial microecology shown before, fungal microecology rarely returned to normal two months after the *SARS-CoV-2* has been removed.

The persist oropharyngeal viral dysbiosis: *Rhinovirus A, Rhinovirus C* and *Human gamma herpesvirus 4* were the major viruses in oropharyngeal swabs from HCG. In addition to the greatly increase of *SARS-CoV-2*, Human *beta herpesvirus 7* also increased in oropharyngeal swabs from PDG than that in HCG, but *Rhinovirus A* and *Rhinovirus C* decreased significantly. Opportunistic viruses, including Human *gamma herpesvirus 4*, *Human alpha herpesvirus 1*

188 and *Porcine type C oncovirus* were richer in oropharyngeal swabs from CG than
189 that in HCG and PDG(**Fig.3a**).
190 **Little changes in α-diversity among HCG, PDG and CG**: We investigated
191 longitudinal dynamics of oropharynx microbial (include bacteria and fungi)
192 profiles in COVID-19 over time of exacerbation and explored whether recovery
193 from *SARS-CoV-2* infection was associated with restoration of oropharyngeal
194 microbial profiles to a community similar to that of healthy individuals. Results
195 from α-diversity microbial analyses (measured by the Shannon Diversity Index
196 and Observed Diversity Index) have demonstrated little changes in α-diversity
197 among PDG, CG and HCG ($P > 0.05$) (**Fig.4a, b**).
198 **Obviously changes in β-diversity among HCG, PDG and CG:** Principal
199 coordinates analysis (PCoA) based on Bray-Curtis distance was used to
200 visualize the apparent differences in microbial community structure. **Fig. 4c**
201 shows that the PDG and CG are extremely separate from HCG, suggesting that
202 oropharyngeal microbial structures were modulated by *SARS-CoV-2* infection
203 and could not recover even virus was cleared two month later. The analysis of
204 similarities (ANOSIM) plot shows that the inter-group difference was greater
205 than the intra-group difference (**Fig. 4d**, R= 0.246, $P = 0.001$).
206 In conclusion, the *SARS-CoV-2* infection has a strong influence on the
207 composition of oropharyngeal microbiome.
208 **Obviously changes in antibiotic resistance gene expression among HCG,**
209 **PDG and CG:** To characterize antibiotic resistance gene expression profiles,

we analyzed the metatranscriptomic sequence reads based on Comprehensive Antibiotic Resistance Database (CARD 2020; https://card.mcmaster.ca). Across the dataset, we identified genes that confer resistance to 13 classes of antibiotics (**Fig. 5a, b**) with variations across the swabs. The expression of multidrug resistance genes (especially multidrug resistance, *tetracycline* and *glycopeptide*) in the oropharyngeal microbiome of patients with COVID-19 were higher than that in healthy controls, and returned to normal two months after virus clearance.

Alteration of main oropharyngeal microbial composition reflects disease severity of patients with COVID-19

In order to explore the role of oropharyngeal microbiome in patients with COVID-19, PDG subjects were further categorized as mild, moderate, severe, and critical according to the Diagnosis and Treatment Scheme for COVID-19 released by the National Health Commission of China (Version7)[8, 13, 14], and mild, moderate patients with COVID-19 were divided as mild-moderate group (MMG), severe and critical patients with COVID-19 were divided as severe-critical group (SCG). Associations between alteration of main oropharyngeal microbial composition in patients with COVID-19 and the magnitude of COVID-19 severity were analysed.

***Prevotella* increased gradually along with the symptom aggravated**: Top20 Notably, *Prevotella* increased significantly in PDG than that in HCG (Wilcoxon rank-sum test, $p < 0.001$) (**Fig.1d**). As to the subgroup of PDG, Top20 bacterial

profiles in oropharyngeal swabs from MMG and SCG were shown as **Fig.6a, b**. The abundance of *Prevotella* in SCG was significantly higher (20.1%) than that in MMG (17.3%) (**Fig.6c**). *Prevotella* increased gradually, along with the severity of patients with COVID-19 aggravated gradually. (**Fig.6d**). To investigate the potential effects of the *Prevotella* in oropharynx of patients with COVID-19, we used a spearman analysis strategy for the parallel evaluation of all quantitative clinical traits from COVID-19 patients. Remarkably, there was a positive correlation between *Prevotella* and the elevation of Neutrophil percentage (R = 0.301, p = 0.040) **(Fig.6e)**. These results indicated that it would be necessary to verify the synergic role of *Prevotella* in inducing severity of *SARS-CoV-2* mediated pneumonia.

***Aspergillus* increased remarkably in critical Patients with COVID-19:** It was observed that patients with severe COVID-19 had an increased risk of coexisting *Aspergillus* infection [18]. As to the subgroup of PDG, Top20 fungal profiles in oropharyngeal swabs specimens from MMG and SCG were shown as **Fig.7a, b**. The abundance of *Aspergillus* in SCG was significantly higher (21.7%) than that in MMG (19.8%) (**Fig.7c**). In particularly, *Aspergillus* increased remarkably in critical Patients with COVID-19 (**Fig.7d**). Therefore, monitoring the change in abundance of *Aspergillus* in oropharynx of COVID-19 patients should be considered[18, 19].

Decreased α-diversity appeared to be associated with COVID-19 severity: The α-diversity of oropharyngeal microbial profiles did not differ between

healthy controls and patients with COVID-19 (**Fig. 4a,b,** $P > 0.05$). However, the diversity and richness of oropharyngeal microbial profiles were both significantly lower in SCG group than MMG group. Obviously, the α-diversity decreased gradually along with the symptom aggravated(**Fig. 8a,b**, $P < 0.01$). Besides, Principal coordinates analysis (PCoA) of β-diversity between SCG and MMG demonstrated that the microbial cluster away from each other (**Fig.8c**), there were also significant differences among mild, moderate, severe, and critical groups (**Fig.8d**).

The severity of COVID-19 was positively correlated with higher *SARS-CoV-2* viral load: Fig.3b and **Fig.3c** showed that severe and critical patients with COVID-19 had much higher *SARS-CoV-2* viral load than others. There was a negative correlation between *SARS-CoV-2* viral load and platelet counts ($R = -0.330$, $P = 0.022$) **(Fig.3d)**. These results indicated that *SARS-CoV-2* viral load may be a biomarker to evaluate if they are able to predict lower platelet counts and correlate with the severity of COVID-19 disease. In conclusion, early check of *SARS-CoV-2* viral load by metatranscriptome may be helpful for early detection of critically COVID-19 patients.

Discussion

SARS-COV-2 and other viral infections may alter the microbial composition of the upper respiratory tract, foster enhanced growth of pathogens, and facilitate the subsequent entry of large microbial loads into the lower respiratory tract [8, 12, 20, 21], which are frequently associated with a more severe clinical

course[12, 18, 19, 22-24]. Besides, a specific bacterial formulation has showed a significant ameliorating impact on the clinical conditions of patients positive for *SARS-CoV-2* infection[25]. And given the unequivocal association between gut microbiota composition and COVID-19 disease severity, there is a pressing need to better understand how perturbation of the host oropharyngeal microbiome correlates with *SARS-COV-2* infections[6]. Based on analysis unbiased metatranscriptomic data, our longitudinal study has firstly confirmed that oropharyngeal microbiome composition reflects disease severity and presents long-term dysbiosis in patients with COVID-19.

The bacterial and fungal microecology of oropharynx were significantly perturbed by *SARS-CoV-2* infection, and they were even hard to return back to normal two months after the virus was cleared. Certainly, it is hard to ignore the influence of antibiotic exposure on URT microbiome. However, recent proof demonstrated that antibiotic drugs have short-term influence on the human gut microbiome with the bacterial community largely recovering within 30 days of antibiotic treatment[26, 27]. So, we speculate that *SARS-COV-2* infection is the key factor affecting the microecological dynamics of oropharynx. As other viral infections, antiviral immune responses induced by *SARS-COV-2* is associated with changes in microbial composition and function in the URT, which in turn may alter immune function of URT, thereby enhancing the sustainable proliferation of some bacterial species[12, 21].

Our study conducted in healthy subjects and who underwent *SARS-COV-2*

infection failed to demonstrate obvious changes in α-diversity of the oropharyngeal microbiome. Thus, as other viral infections[12], the effects of *SARS-COV-2* infection on diversity itself are not presently considered a good indicator of risk for complications, including secondary bacterial pneumonias. The effect of *SARS-COV-2* infection on α diversity is variable and is not currently considered a good indicator of the viral infection. Nevertheless, our study in patients with *SARS-COV-2* demonstrated that the α-diversity of oropharyngeal microbiome decreased significantly associated with COVID-19 severity.

In recent studies by next generation sequencing and MinION sequencing, *Prevotella* was found to be the common bacteria in URT of patients with COVID-19 [28, 29]. Notably, rich *Prevotella* was also detected in the intestines of patients with COVID-19[3]. The synergism of *Prevotella* in causing the severity of streptococcus pneumoniae mediated pneumonia was already known[30]. Our data about patients with COVID-19 further demonstrated that *Prevotella* in oropharynx increased gradually consist with the symptom aggravated and could not return back to normal even two months after the *SARS-COV-2* was cleared. Through microaspiration from URT, *Prevotella* could participate in the immune homeostasis of the low respiratory tract [31], and pulmonary inflammation had been associated with accumulation of the pulmonary microbiome with *Prevotella*, this bacterium mainly activated Toll-like receptor 2 and enhanced the expression IL-23 and IL-1 [32]. In vitro, *Prevotella* could

stimulate the production of IL-8, IL-6, and CCL20 in lung epithelial cells, which promoted neutrophil recruitment [33]. In vivo, we also confirmed that there was a positive correlation between *Prevotella* and the elevation of Neutrophil percentage. Latest evidence about host-pathogen interactions in the severity of COVID-19 also confirmed *Prevotella* proteins were over expressed in patients with COVID-19, and *Prevotella* proteins, but not *SARS-COV-2* proteins were involved in multiple interactions with NF-kB, which was involved in increasing clinical severity of COVID-19[34]. In conclusion, *Prevotella* may play an important role in COVID-19 outbreak and should be given attention for understanding disease mechanisms and improving treatment outcomes.

Aspergillus is a common fungal pathogen of secondary infection in patients with critical COVID-19[18, 19, 35, 36]. Our results showed that *Aspergillus* increased remarkably in oropharynx of critical Patients with COVID-19. The respiratory tract is an interconnected system consisting of oropharynx and low respiratory tract, with microaspiration serving as the primary route of *Aspergillus* immigration from the oropharynx to the low respiratory tract. Therefore, for early detection of secondary fungal infections, monitoring the change in abundance of *Aspergillus* in oropharynx of patients with COVID-19 should be considered.

Soon after the onset of the global *SARS-CoV-2* epidemic, the presence of other common respiratory viruses declined and even disappeared[22, 28]. Our results by metatranscriptomics are consistent with the absence of other common respiratory virus coinfection in patients with COVID-19. More

importantly, the severity of COVID-19 was positively correlated with higher *SARS-CoV-2* viral load and opportunistic viruses were rich even fully recovered two months later, and there was a negative correlation between *SARS-CoV-2* viral load and platelet counts. Platelet count showed significantly lower levels in severe patients compared to non-severe patients[37]. Meta-analysis of 1799 patients also revealed those with severe COVID-19 infections had significantly lower platelet counts [37, 38].Therefore, besides keeping an eye on opportunistic viral infection of convalescent patients with COVID-19, early check of *SARS-CoV-2* viral load by metatranscriptome may be helpful for early detection of critically COVID-19 patients.

More importantly, we firstly performed metatranscriptomics to characterize the abundance and diversity of antibiotic resistance genes in the oropharynx microbiome of patients with COVID-19. Higher relative abundance of antibiotic resistance genes in the oropharyngeal microbiome of patients with COVID-19 were decreased to normal two months after virus clearance. In line with previous study[26], Characterization of bacterial antibiotic resistance genes in the URT microbiome of patients with COVID-19 might help better to choose the appropriate antibiotics for bacterial co-infection and secondary infection in patients with COVID-19.

Conclusions

SARS-CoV-2 primarily infects the respiratory tract and significantly perturbs microecology of oropharynx. Alterations in oropharyngeal microbial

composition, particularly of *Prevotella* and *Aspergillus*, reflect the severity of disease in patients with COVID-19. Furthermore, the oropharyngeal microbiome in patients with COVID-19 present persistent dysbiosis two months after clearance of the *SARS-CoV-2* virus. These findings underscore an urgent need to understand the specific roles of oropharyngeal microorganisms in COVID-19 disease progression and rehabilitation as well as co-infection and secondary infection.

Abbreviations

SARS-CoV-2: severe acute respiratory syndrome coronavirus 2

COVID-19: coronavirus disease 2019

URT: upper respiratory tract

PDG: period of disease group

CG: convalescent group

HCG: healthy controls group

RT-qPCR: quantitative reverse transcription polymerase chain reaction

Methods

Study design and specimens collect

We recruited 47 patients with COVID-19 confirmed by positive *SARS-CoV-2* quantitative reverse transcription PCR(RT-qPCR) to establish a longitudinal cohort study. The severity status of confirmed COVID-19 cases was categorized as mild, moderate, severe, and critical according to the Diagnosis and Treatment Scheme for COVID-19 released by the National Health

Commission of China (Version 7). As healthy controls group, we recruited 40 healthy human subjects free of any respiratory infections during the same timeframe. Written consents were obtained from patients or their guardian(s). Clinical information and demographic information were collected. Oropharyngeal samples were obtained for routine diagnostic purposes, and the residual clinical diagnostic samples were provided for research and surveillance purposes. We collected oropharyngeal swabs at two time points (Within a week of diagnosis COVID-19 and two months after clearance of the *SARS-CoV-2* virus. During the same timeframe, we collected oropharyngeal swabs from 40 healthy human subjects free of any respiratory infections. In order to remove possible contamination, blank control samples, including another 6 swabs not sampled, were collected. After collection, the samples were placed into a tube containing 2mL viral RNA preservation solution.

Samples processing and quantitative real-time PCR test for COVID-19

Total RNA was extracted using QIAamp viral RNA mini kit (QIAGEN 52904, Germany) following the manufacturer's protocol.

Real-time RT-PCR was performed by amplifying two target genes, including open reading frame 1ab (ORF1ab) and nucleocapsid protein (N) using RT-PCR kit (Sansure Biotech Inc, China) on a Real-time PCR thermal cycler (ABI 7500 system, Applied Biosystems instruments, USA).

Metatranscriptomic sequencing

After extraction, the first strand cDNA was prepared by priming with random

hexamers followed. Standard RNAseq library was constructed using NGS library construction kit (Genskey 1906). The libraries were sequenced on Illumina NextSeq 500-sequencer using a 75-cycle single-end sequencing strategy and each sample will get nearly 20 million reads.

Data processing

First, software fastp (parameters: -q 15 –u 40 -l 50, version: 0.19.5) [39] was used to filter low-quality reads, remove adapter and Komplexity (parameters: -t 0.4, version: Nov-2019)[40]was used to remove low complexity reads for raw data. Then, filtered reads were mapped to the ensembl 84 (GRCh38) human reference genome to remove host sequences using HISAT2 (version 2.1.0) with default parameters [41]. Next, unmapped reads were annotated taxonomic classifications by Kraken2 (version 2.0.9, parameters: --threads 24 --confidence 0.1)[42]with a self-build database (built by downloading all the complete genomes from NCBI Refseq database, including the *SARS-COV-2* reference NC_045512.2. And only genomes from archaea, bacteria, fungi, protozoa and viral were selected for building a classification database for Kraken2 (k=35, ℓ=31)). Then, abundance of taxonomy was estimated using Bracken (version 2.5, parameters: -r 75 -l G, S -t 0) [43]. Finally, microorganisms satisfying the following criteria were retained for the subsequent analysis: (1) archaea, bacteria, fungi, or virus; (2) 10-fold higher filtered RPM than that in the negative control (NC); (3) no batch effect; (4) no known contamination;(5) only human virus.

In order to obtain information of antibiotic resistance gene. First, unmapped reads were assemblied using SPAdes (version: 3.13.0) with default parameters [44] and scaffolds less than150 bp were removed. Then, MetaGeneMark [45] was used to predict protein coding genes from the above filtered scaffolds. Next, the abundance of antibiotic resistance genes were annotated by aligning the sequence of the above predicted genes to Comprehensive Antibiotic Resistance Database (CARD 2020; https://card. mcmaster.ca) [46] using the HISAT2 alignment method with default parameters. Each gene was filtered to have at least two matching reads. Finally, only antibiotic resistance genes occurred in >25% of samples were used for further statistical analyses.

Statistical analysis

Statistical analyses were done using R (version 3.5.1). The α-diversity was analyzed by vegan function with R packages. Principal co-ordinates analysis (PCoA) was based on Bray-Curtis dissimilarity (BCD) distance matrices[47]. Differences between groups in the α-diversity were evaluated using the Mann-Whitney test or the Kruskal-Wallis test. False discovery rate (FDR) values were estimated using the Benjamini–Yekutieli method to control for multiple testing [48]. Principal component analysis (PCA) was analyzed by prcomp function with R packages [49]. Analysis of similarities (ANOSIM) was used to statistically test whether there is a significant difference between two or more groups of sampling units[50]. The log10 value was used to show the abundance of antibiotic resistance gene. Finally, the spearman correlation method [51] was

used to analyze the correlation analysis between microbe and clinical index, and a correlation coefficient of more than 0.300 is considered [52]. A p-value of < 0.05 was considered statistically significant. Concerning symbols used to represent higher orders of significance, $P < 0.05$ was indicated by *, $P < 0.01$ by ** and $P < 0.001$ by ***.

Acknowledgements

We thank Yannan Lini for assistance with processing language editing of the manuscript. We thank Genskey for their assistance in the sequencing work.

Authors' contributions

WML, BWY, DL, YZZ, MJL and MJW contributed substantially to the conception and design of the study and interpretation and wrote the manuscript. RQ, TBD, LJ, JZ and YLL undertook clinical sample collection and processing and analyses of clinical records. YZZ, SFL, MJL, MJW, JMM, YCM, ZYZ, JMM and JX contributed substantially to the acquisition of data and the analysis. YZZ and SFL contributed substantially to data interpretation. The authors read and approved the final manuscript.

Funding

This work was funded by the Sichuan Provincial Program for Diagnostic and Treatment of COVID-19 (2020YFS0002, 2020YFS0001, 2020YFS0572).

Availability of data and materials

The sequencing data from this study have been deposited in the CNSA (https://db.cngb.org/cnsa/) of CNGBdb with accession number CNP0001393.

Ethics approval and consent to participate

This study was approved by The Biomedical Research Ethics Committee of West China Hospital and the document IDs are NO.2020(100), NO.2020 (193) and NO.2020 (267). Written consents were obtained from patients or their guardian(s). All work was performed in accordance with the Helsinki Declaration.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have not competing interests.

References

1. Zhu N, Zhang D, Wang W, Li X, Yang B, Song J, Zhao X, Huang B, Shi W, Lu R *et al*: **A Novel Coronavirus from Patients with Pneumonia in China, 2019**. *N Engl J Med* 2020, **382**(8):727-733.
2. Microbiome Centers Consortium CC: **Coordinating and Assisting Research at the SARS-CoV-2/Microbiome Nexus**. *mSystems* 2020, **5**(6).
3. Tao W, Wang X, Zhang G, Guo M, Ma H, Zhao D, Sun Y, He J, Liu L, Zhang K *et al*: **Re-detectable positive SARS-CoV-2 RNA tests in patients who recovered from COVID-19 with intestinal infection**. *Protein Cell* 2020.
4. Zuo T, Zhang F, Lui GCY, Yeoh YK, Li AYL, Zhan H, Wan Y, Chung ACK, Cheung CP, Chen N *et al*: **Alterations in Gut Microbiota of Patients With COVID-19 During Time of Hospitalization**. *Gastroenterology* 2020, **159**(3):944-955.e948.

5. Zuo T, Zhan H, Zhang F, Liu Q, Tso EYK, Lui GCY, Chen N, Li A, Lu W, Chan FKL *et al*: **Alterations in Fecal Fungal Microbiome of Patients With COVID-19 During Time of Hospitalization until Discharge**. *Gastroenterology* 2020, **159**(4):1302-1310 e1305.

6. Yeoh YK, Zuo T, Lui CY, Zhang F, Ng SC: **Gut microbiota composition reflects disease severity and dysfunctional immune responses in patients with COVID-19**. *Gut* 2021:gutjnl-2020-323020.

7. 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* 2021.

8. Shen Z, Xiao Y, Kang L, Ma W, Shi L, Zhang L, Zhou Z, Yang J, Zhong J, Yang D *et al*: **Genomic Diversity of Severe Acute Respiratory Syndrome-Coronavirus 2 in Patients With Coronavirus Disease 2019**. *Clin Infect Dis* 2020, **71**(15):713-720.

9. Qin T, Geng T, Zhou H, Han Y, Ren H, Qiu Z, Nie X, Du T, Liang J, Du P *et al*: **Super-dominant pathobiontic bacteria in the nasopharyngeal microbiota as causative agents of secondary bacterial infection in influenza patients**. *Emerg Microbes Infect* 2020, **9**(1):605-615.

10. Martin-Loeches I, M JS, Vincent JL, Alvarez-Lerma F, Bos LD, Sole-Violan J, Torres A, Rodriguez A: **Increased incidence of co-infection in critically ill patients with influenza**. *Intensive Care Med* 2017, **43**(1):48-58.

11. Kaul D, Rathnasinghe R, Ferres M, Tan GS, Barrera A, Pickett BE, Methe BA, Das SR, Budnik I, Halpin RA *et al*: **Microbiome disturbance and resilience dynamics of the upper respiratory tract during influenza A virus infection**. *Nat Commun* 2020, **11**(1):2537.

12. Hanada S, Pirzadeh M, Carver KY, Deng JC: **Respiratory Viral Infection-Induced Microbiome Alterations and Secondary Bacterial Pneumonia**. *Front Immunol* 2018, **9**:2640.

13. Jian W, Jun L, Xinguo Z, Chengyuan L, Wei W, Dawei W, Wei X, Chunyu Z, Jiong Y, Bin J: **Clinical Characteristics of Imported Cases of Coronavirus Disease 2019 (COVID-19) in Jiangsu Province: A Multicenter Descriptive Study**. *Clinical Infectious Diseases*.

14. Lu J, du Plessis L, Liu Z, Hill V, Kang M, Lin H, Sun J, Francois S, Kraemer MUG, Faria NR *et al*: **Genomic Epidemiology of SARS-CoV-2 in Guangdong Province, China**. *Cell* 2020, **181**(5):997-1003 e1009.

15. Zhang H, Ai JW, Yang W, Zhou X, Zhang W: **Metatranscriptomic Characterization of COVID-19 Identified A Host Transcriptional Classifier Associated With Immune Signaling**. *Clinical Infectious Diseases* 2020.

16. **The nasal cavity microbiota of healthy adults**. *Microbiome* 2014, **2**(1):1-5.

17. Lemon KP, Klepac-Ceraj V, Schiffer HK, Brodie EL, Lynch SV, Kolter R: **Comparative analyses of the bacterial microbiota of the human nostril and oropharynx**. *mBio* 2010, **1**(3).

18. Koehler P, Cornely OA, Bottiger BW, Dusse F, Eichenauer DA, Fuchs F, Hallek M, Jung N, Klein F, Persigehl T *et al*: **COVID-19 associated pulmonary aspergillosis**. *Mycoses* 2020, **63**(6):528-534.

19. Gangneux JP, Bougnoux ME, Dannaoui E, Cornet M, Zahar JR: **Invasive fungal diseases during COVID-19: We should be prepared**. *J Mycol Med* 2020, **30**(2):100971.

20. Wolter N, Tempia S, Cohen C, Madhi SA, Venter M, Moyes J, Walaza S, Malope-Kgokong B, Groome M, du Plessis M *et al*: **High nasopharyngeal pneumococcal density, increased by viral coinfection, is associated with invasive pneumococcal pneumonia**. *J Infect Dis* 2014, **210**(10):1649-1657.

21. Haocheng Z, Jing-Wen A, Wenjiao Y, Xian Z, Fusheng H, Shumei X, Weiqi Z, Yang L, Yiqi Y, Xuejing G: **Metatranscriptomic Characterization of COVID-19 Identified A Host Transcriptional**

Classifier Associated With Immune Signaling. *Clinical Infectious Diseases*.

22. Calcagno A, Ghisetti V, Burdino E, Trunfio M, Allice T, Boglione L, Bonora S, Di Perri G: **Coinfection with other respiratory pathogens in COVID-19 patients**. *Clin Microbiol Infect* 2020.

23. Langford BJ, So M, Raybardhan S, Leung V, Westwood D, MacFadden DR, Soucy JR, Daneman N: **Bacterial co-infection and secondary infection in patients with COVID-19: a living rapid review and meta-analysis**. *Clin Microbiol Infect* 2020.

24. Kim D, Quinn J, Pinsky B, Shah NH, Brown I: **Rates of Co-infection Between SARS-CoV-2 and Other Respiratory Pathogens**. *JAMA* 2020, **323**(20):2085-2086.

25. d'Ettorre G, Ceccarelli G, Marazzato M, Campagna G, Pinacchio C, Alessandri F, Ruberto F, Rossi G, Celani L, Scagnolari C *et al*: **Challenges in the Management of SARS-CoV2 Infection: The Role of Oral Bacteriotherapy as Complementary Therapeutic Strategy to Avoid the Progression of COVID-19**. *Front Med (Lausanne)* 2020, **7**:389.

26. Seelbinder B, Chen J, Brunke S, Vazquez-Uribe R, Santhaman R, Meyer AC, de Oliveira Lino FS, Chan KF, Loos D, Imamovic L *et al*: **Antibiotics create a shift from mutualism to competition in human gut communities with a longer-lasting impact on fungi than bacteria**. *Microbiome* 2020, **8**(1):133.

27. MacPherson CW, Mathieu O, Tremblay J, Champagne J, Nantel A, Girard SA, Tompkins TA: **Gut Bacterial Microbiota and its Resistome Rapidly Recover to Basal State Levels after Short-term Amoxicillin-Clavulanic Acid Treatment in Healthy Adults**. *Sci Rep* 2018, **8**(1):11192.

28. Wang Z, Hu X, Li Z, Tu C, Wang Y, Pang P, Zhang H, Zheng X, Liang Y, Shan H *et al*: **Effect of SARS-CoV-2 Infection on the Microbial Composition of Upper Airway**. *Infect Drug Resist* 2020, **13**:2637-2640.

29. Moore SC, Penrice-Randal R, Alruwaili M, Randle N, Armstrong S, Hartley C, Haldenby S, Dong X, Alrezaihi A, Almsaud M *et al*: **Amplicon-Based Detection and Sequencing of SARS-CoV-2 in Nasopharyngeal Swabs from Patients With COVID-19 and Identification of Deletions in the Viral Genome That Encode Proteins Involved in Interferon Antagonism**. *Viruses* 2020, **12**(10).

30. Nagaoka K, Yanagihara K, Morinaga Y, Nakamura S, Harada T, Hasegawa H, Izumikawa K, Ishimatsu Y, Kakeya H, Nishimura M: **Prevotella intermedia Induces Severe Bacteremic Pneumococcal Pneumonia in Mice with Upregulated Platelet-Activating Factor Receptor Expression**. *Infection and Immunity* 2014, **82**(2):587-593.

31. Huffnagle GB, Dickson RP, Lukacs NW: **The respiratory tract microbiome and lung inflammation: a two-way street**. *Mucosal Immunol* 2017, **10**(2):299-306.

32. Segal LN, Alekseyenko AV, Clemente JC, Kulkarni R, Wu B, Chen H, Berger KI, Goldring RM, Rom WN, Blaser MJ: **Enrichment of lung microbiome with supraglottic taxa is associated with increased pulmonary inflammation**. *Microbiome* 2013, **1**.

33. Larsen, Madura J: **The immune response to Prevotella bacteria in chronic inflammatory disease**. *Immunology* 2017, **151**(4).

34. Khan AA, Khan Z: **COVID-2019-associated overexpressed Prevotella proteins mediated host-pathogen interactions and their role in coronavirus outbreak**. *Bioinformatics* 2020, **36**(13):4065-4069.

35. Vaillancourt M, Jorth P: **The Unrecognized Threat of Secondary Bacterial Infections with COVID-19**. *mBio* 2020, **11**(4).

36. Biesen SV, Kwa D, Bosman RJ, Juffermans NP: **Detection of Invasive Pulmonary Aspergillosis in COVID-19 with Non-directed Bronchoalveolar Lavage**. *American Journal of Respiratory and*

584 *Critical Care Medicine* 2020.

585 37. Kermali M, Khalsa RK, Pillai K, Ismail Z, Harky A: **The role of biomarkers in diagnosis of COVID-**
586 **19 - A systematic review**. *Life Sci* 2020, **254**:117788.

587 38. Lippi G, Plebani M, Henry BM: **Thrombocytopenia is associated with severe coronavirus**
588 **disease 2019 (COVID-19) infections: A meta-analysis**. *Clinica Chimica Acta* 2020.

589 39. Chen S, Zhou Y, Chen Y, Gu J: **fastp: an ultra-fast all-in-one FASTQ preprocessor**. *Bioinformatics*
590 2018, **34**(17):i884-i890.

591 40. Zijie S, Yan X, Lu K, Wentai M, Leisheng S, Li Z, Zhuo Z, Jing Y, Jiaxin Z, Donghong Y: **Genomic**
592 **Diversity of Severe Acute Respiratory Syndrome–Coronavirus 2 in Patients With Coronavirus**
593 **Disease 2019**. *Clinical Infectious Diseases*.

594 41. Kim D, Langmead B, Salzberg SL: **HISAT: A fast spliced aligner with low memory requirements**.
595 *Nature Methods* 2015, **12**(4).

596 42. Wood DE, Lu J, Langmead B: **Improved metagenomic analysis with Kraken 2**. *Genome biology*
597 2019, **20**(1).

598 43. Lu J, Breitwieser FP, Thielen P, Salzberg SL: **Bracken: estimating species abundance in**
599 **metagenomics data**. *Peerj Computer Science* 2017, **3**(1):e104.

600 44. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI,
601 Pham S, Prjibelski AD: **SPAdes: a new genome assembly algorithm and its applications to**
602 **single-cell sequencing**. *Journal of Computational Biology* 2012, **19**(5):455-477.

603 45. Wenhan Z, Alexandre L, Mark B: **Ab initio gene identification in metagenomic sequences**.
604 *Nucleic Acids Research* 2010, **38**(12):e132.

605 46. Alcock BP, Raphenya AR, Lau TTY, Tsang KK, Mégane B, Arman E, William H, Nguyen A-LV, Cheng

606 AA, Sihan L: **CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic**
607 **resistance database**. *Nucleic Acids Research*.
608 47. Kristi B, Brett WM, Sharon WT, Martin M, Mia J, Melissa Z, Taylor MW, Douglas RG:
609 **Differentially Regulated Host Proteins Associated with Chronic Rhinosinusitis Are Correlated**
610 **with the Sinonasal Microbiome**. *Frontiers in Cellular and Infection Microbiology* 2017, **7**:504-.
611 48. Benjamini Y, Yekutieli D: **The control of the false discovery rate in multiple testing under**
612 **dependency**. *Annals of Statistics* 2001, **29**(4).
613 49. Dubey G, Kollah B, Gour VK, Shukla AK, Mohanty SR: **Diversity of bacteria and archaea in the**
614 **rhizosphere of bioenergy crop Jatropha curcas**. *Biotech* 2016, **6**(2):257.
615 50. Marti, J., Anderson, Daniel, C., I., Walsh: **PERMANOVA, ANOSIM, and the Mantel test in the**
616 **face of heterogeneous dispersions: What null hypothesis are you testing?** *Ecological*
617 *Monographs* 2013.
618 51. De Winter JCF, Gosling SD, Potter J: **Comparing the Pearson and Spearman Correlation**
619 **Coefficients Across Distributions and Sample Sizes: A Tutorial Using Simulations and**
620 **Empirical Data**. *Psychological Methods* 2016, **21**(3):273-290.
621 52. Mukaka M: **Statistics corner: A guide to appropriate use of correlation coefficient in medical**
622 **research**. *Malawi Medical Journal* 2012, **24**(3):69-71.

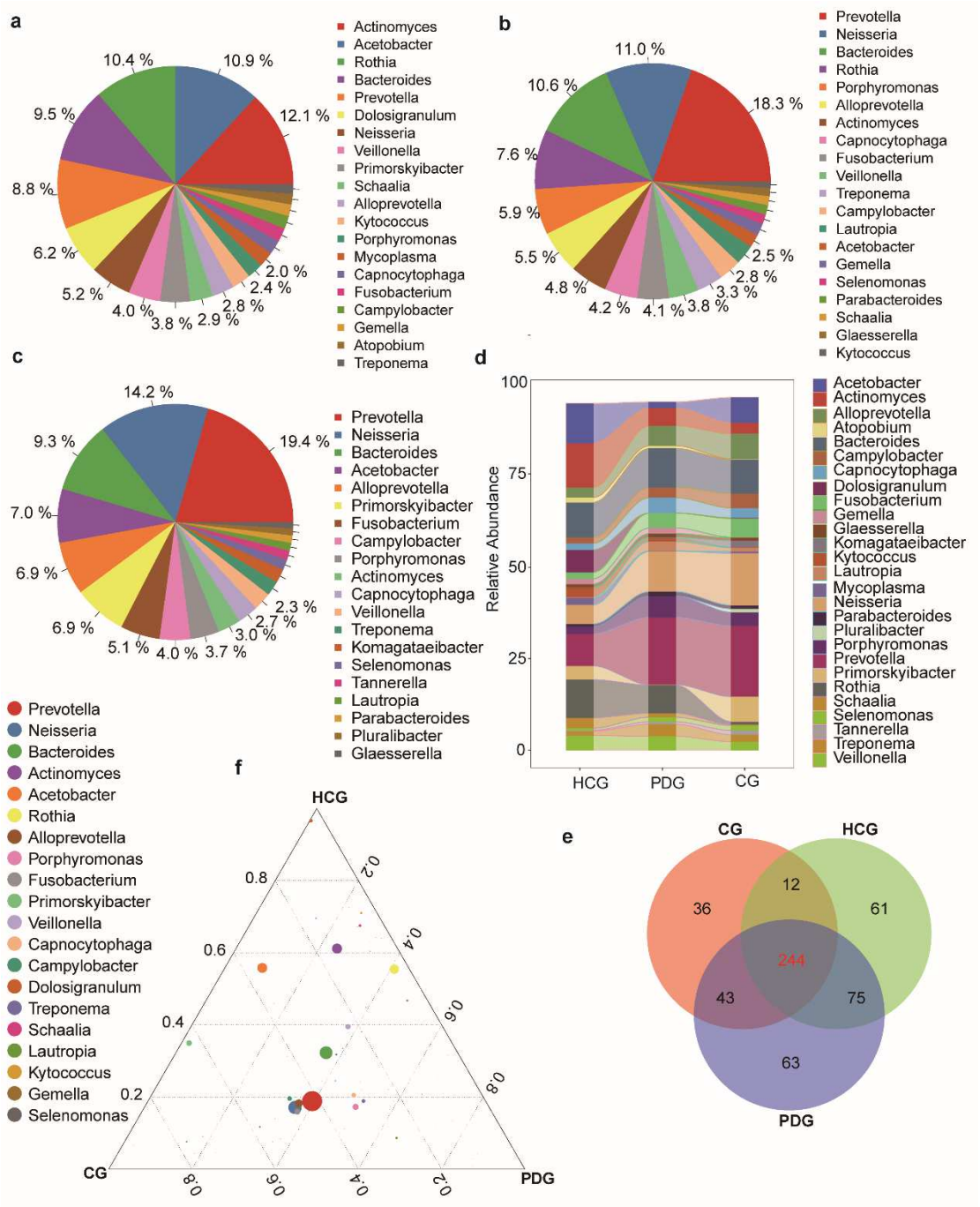
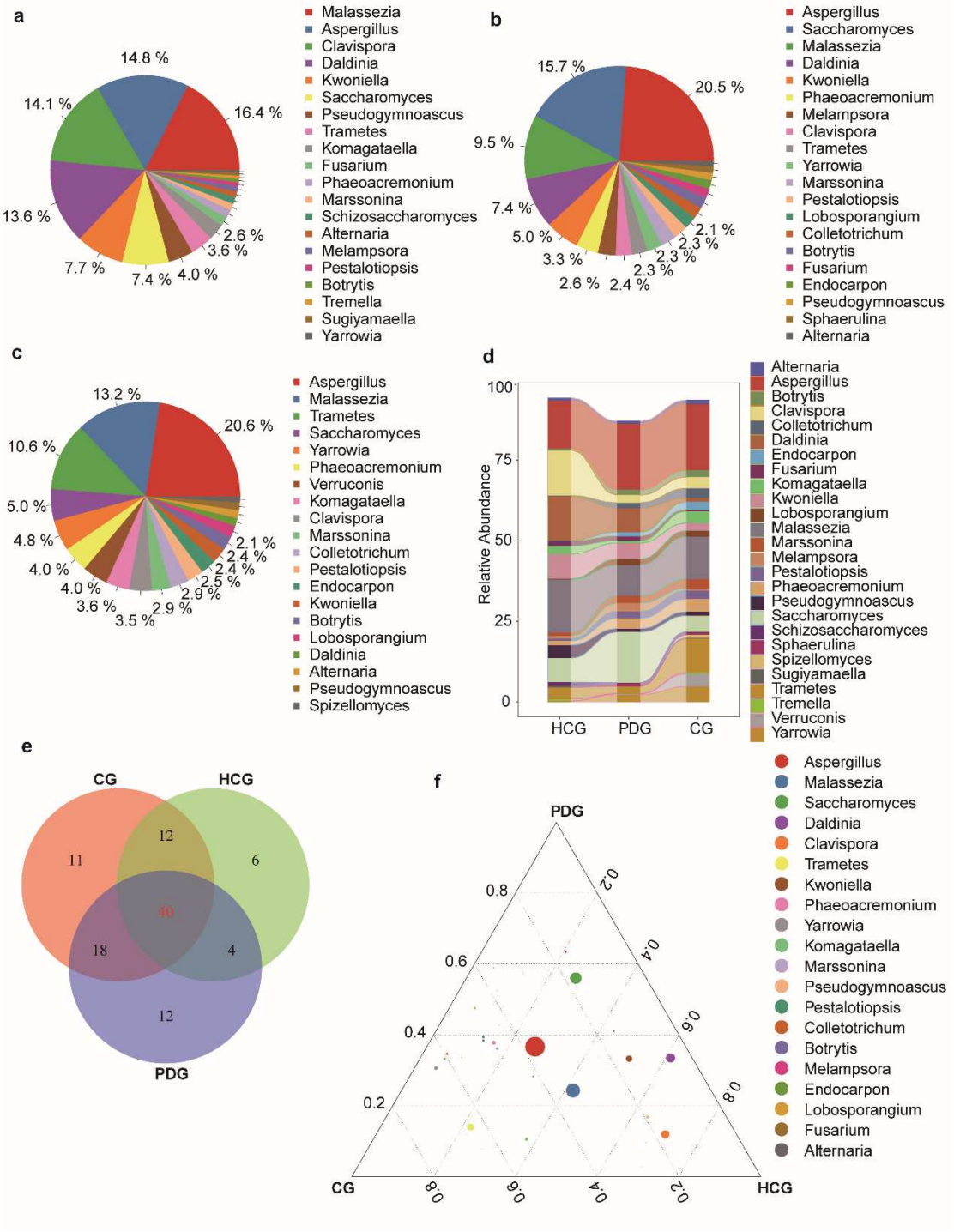


Fig.1 Oropharyngeal bacterial characterization and alterations among HCG, PDG and CG. Average compositions of relative abundance of the top20 bacterial communities for each group at the genus level, **a** for HCG, **b** for PDG and **c** for CG. **d** Alterations of top20 bacterial communities among the three groups at the genus level, Especially, *Prevotella* is significantly higher in PDG

than that in HCG (Wilcoxon rank-sum test, $p < 0.001$). **e** Venn diagram showing shared and unique genus of the three groups. **f** Ternary plots depicting the 244 shared bacterial landscape on genus level among HCG, PDG and CG group. The sum of the proportion for one specific bacteria in the three groups was set as 1, and the proportion of one specific bacteria in each group is equal to its corresponding relative abundance divided by the relative abundance sum of this bacteria in the three groups. The size of the circles represents the relative abundance of the genus.



637

638 **Fig.2 Oropharyngeal fungal characterization and alterations among HCG, PDG**

639 **and CG.** Average compositions of relative abundance of the top20 fungal

640 communities for each group at the genus level, **a** for HCG, **b** for PDG and **c** for

641 CG. **d** Alterations of top20 fungal communities among the three groups at the

642 genus level. **e** Venn diagram showing shared and unique genus of the three

groups. **f** Ternary plots depicting the 244 shared fungal landscape on genus level among HCG, PDG and CG group. The sum of the proportion for one specific bacteria in the three groups was set as 1, and the proportion of one specific fungi in each group is equal to its corresponding relative abundance divided by the relative abundance sum of this fungi in the three groups. The size of the circles represents the relative abundance of the genus.

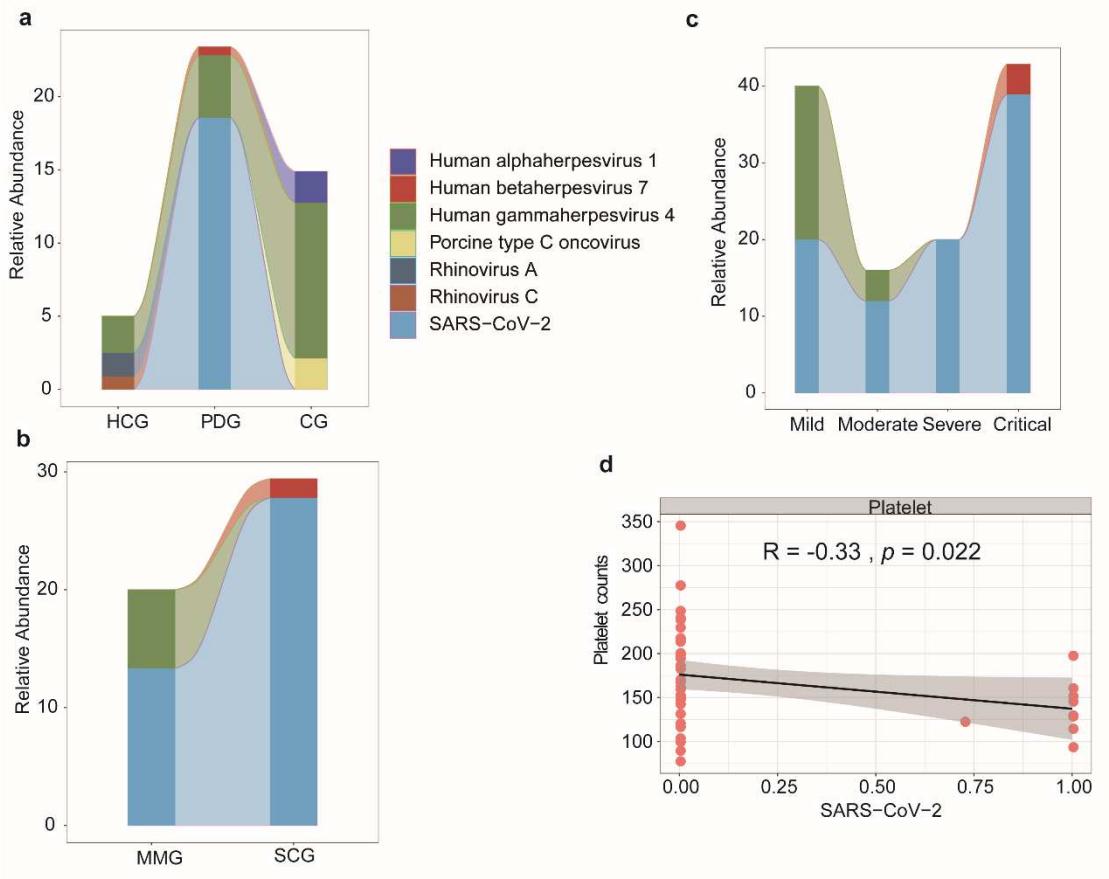


Fig.3 Oropharyngeal *SARS-CoV-2* viral load was positively correlated with the severity of COVID-19. a The relative abundance changes of 7 oropharyngeal viral species (including SARS-COR-2) among HCG, PDG and CG group, **b** among MMG and SCG group, **c** among Mild, Moderate, Severe and Critical group. **d** Negative correlation between *SARS-CoV-2* viral load (x-axis) and

platelet counts (y-axis) (Spearman correlation analysis: R = −0.330, *P* = 0.022), Linear regression lines are shown in each scatter plot in black, and shaded regions represent 95% confidence intervals.

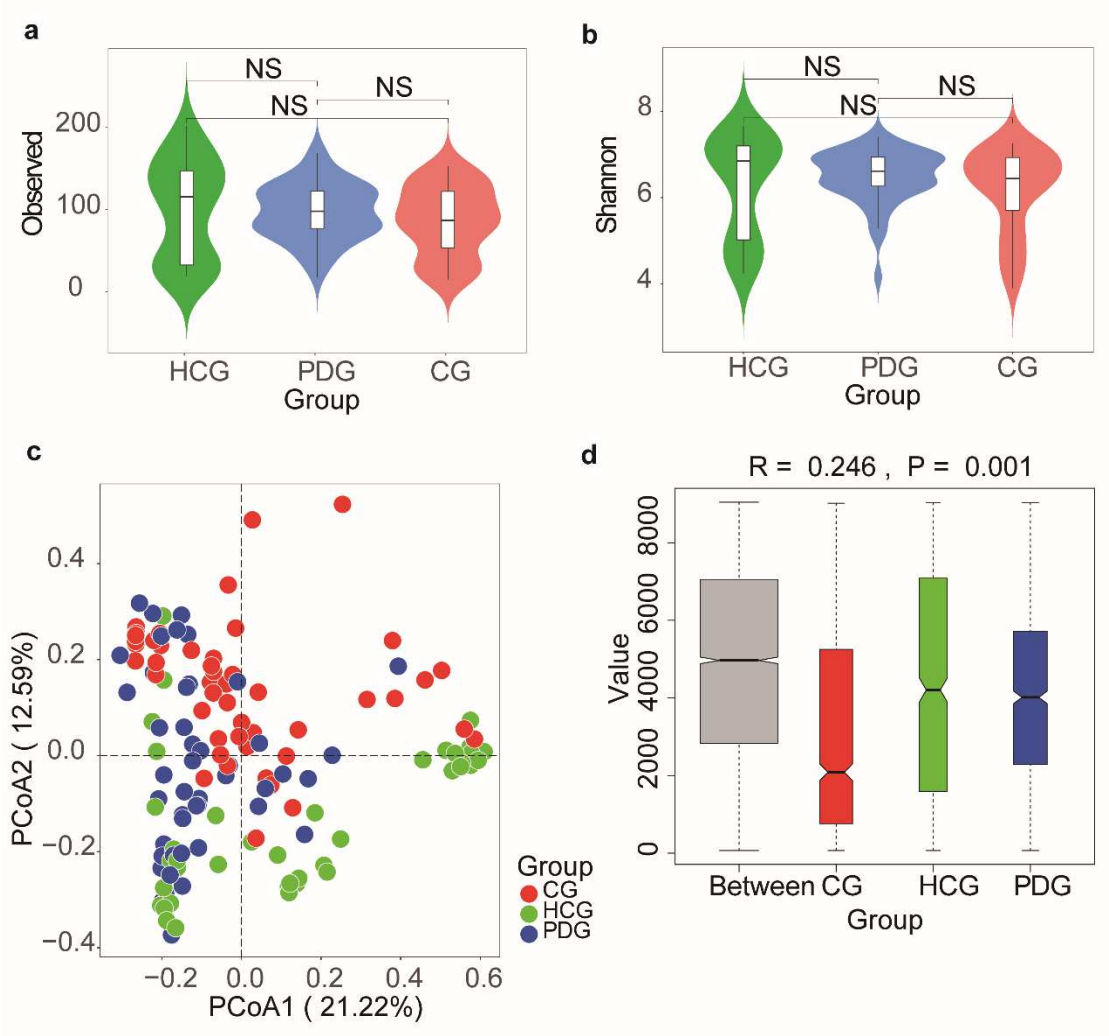


Fig.4 Oropharynx microbial alterations in diversity among HCG, PDG and CG. **a** Little changes in Observed index among HCG, PDG and CG (NS: *P* > 0.05). **b** Little changes in Shannon index among HCG, PDG and CG (NS: *P* > 0.05). **c** Principal coordinate analysis (PCoA) of Bray–Curtis distance showed obviously changes in β-diversity among HCG, PDG and CG. The colors represent three different groups. PCoA1 and PCoA2 represent the top two principal coordinates that captured most of the diversity. The fraction of diversity captured by the

coordinate is valued as a percentage. **d**. The analysis of similarities (ANOSIM) based on Bray–Curtis distance shows that R > 0 and $P < 0.05$, indicating that the inter-group difference among HCG, PDG and CG was greater than the intra-group difference in each group. The R means the statistical value of ANOSIM, y-axis value means Bray–Curtis rank, and 999 was used for the permutations. NS means no significant.

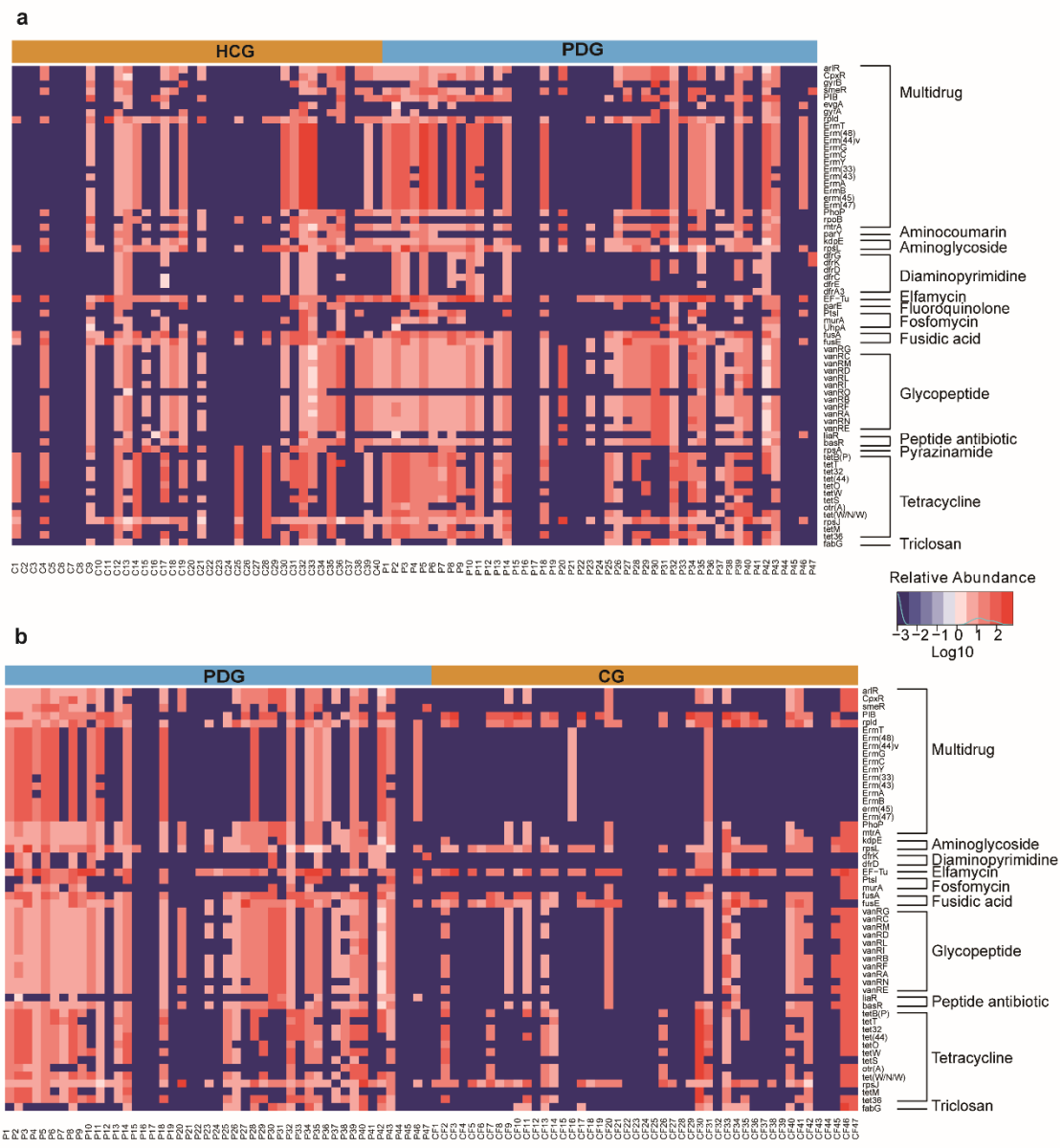


Fig.5 Obviously changes of oropharynx microbial antibiotic resistance genes

expression among HCG, PDG and CG. a Heatmap for the relative abundances of microbial antibiotic resistance genes expression between HCG and PDG. **b** Heatmap for the relative abundances of microbial antibiotic resistance genes expression between PDG and CG. The relative abundance changes from low to high is showed by color changes from blue to red. The log10 value was used to show the abundance of antibiotic resistance gene.

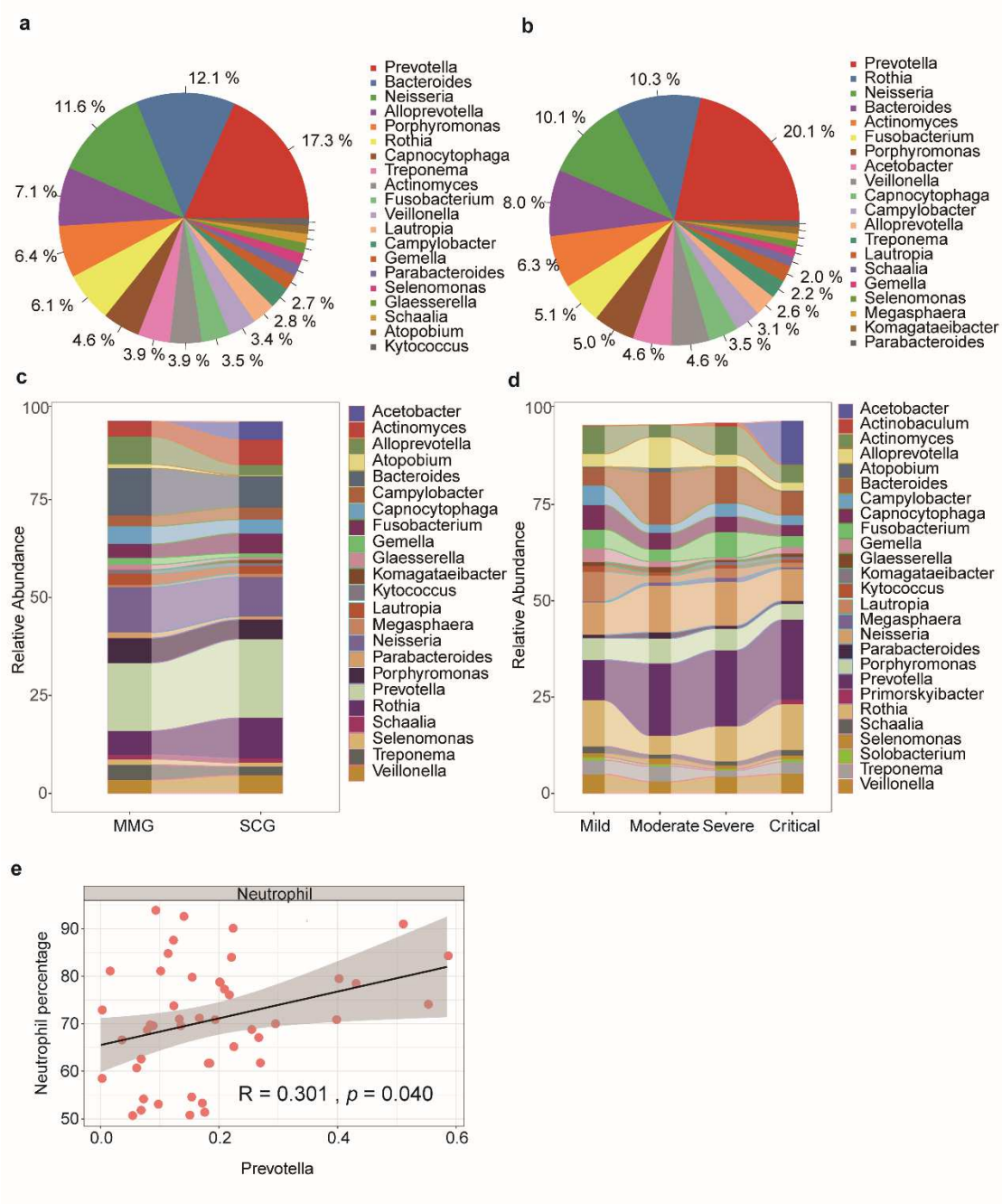


Fig.6 Alterations of oropharyngeal bacteria reflected disease severity in patients with COVID-19. a, b Average compositions of relative abundance of the top20 bacterial communities of MMG and SCG groups at the genus level. **c** Alterations of top20 bacterial communities among the MMG and SCG groups at the genus level. **d** Alterations of top20 bacterial communities among the Mild, Moderate, Severe and Critical groups at the genus level. Especially, the alteration of *Prevotella* gradually increased along with the symptom aggravated in patients with COVID-19. **e** Positive correlation between *Prevotella* relative abundance (x-axis) and the elevation of Neutrophil percentage (y-axis) (Spearman correlation analysis: R = 0.301, $p = 0.040$), Linear regression lines are shown in each scatter plot in black, and shaded regions represent 95% confidence intervals.

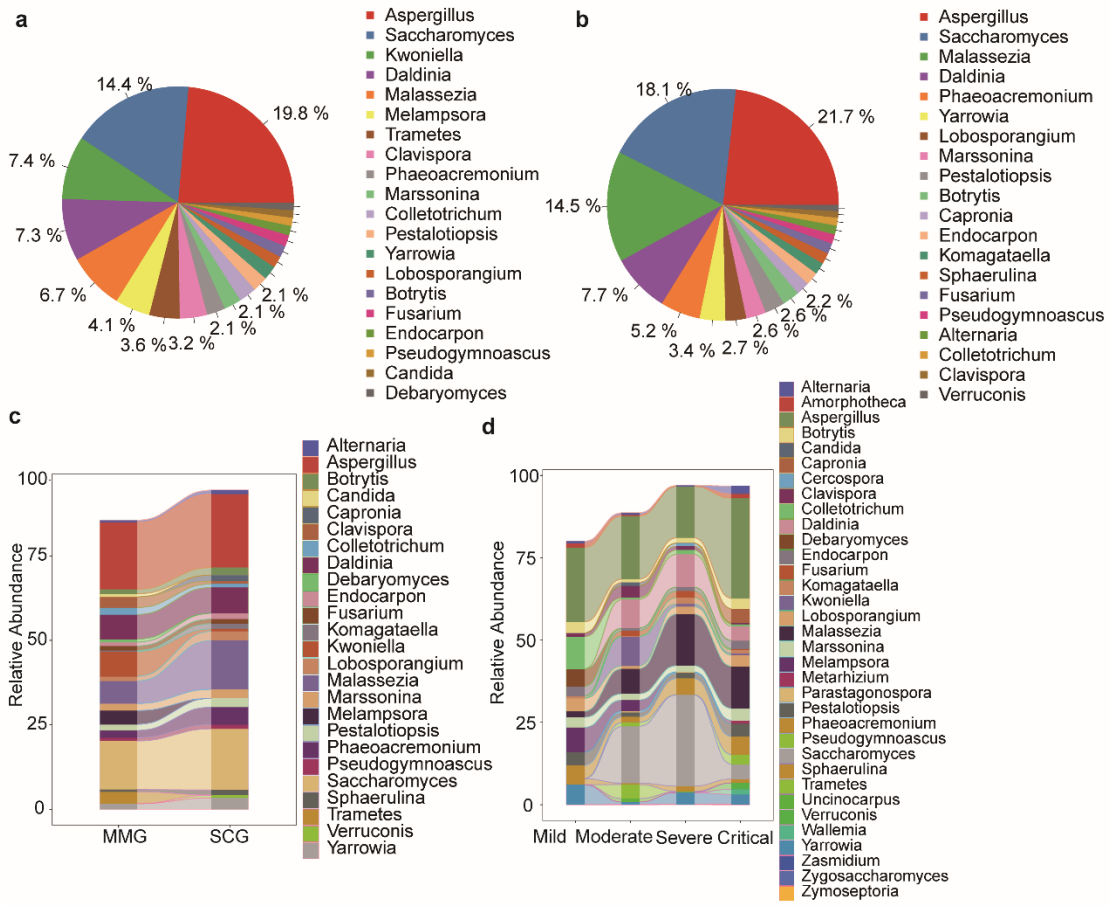


Fig.7 Alterations of oropharyngeal fungi reflected disease severity in patients with COVID-19. a, b Average compositions and relative abundance of the top20 fungal communities of MMG and SCG groups at the genus level. **c** Alterations of top20 fungal communities among the MMG and SCG groups at the genus level. **d** Alterations of top20 fungal communities among the Mild, Moderate, Severe and Critical groups at the genus level. Especially, *Aspergillus* increased remarkably in critical Patients with COVID-19.

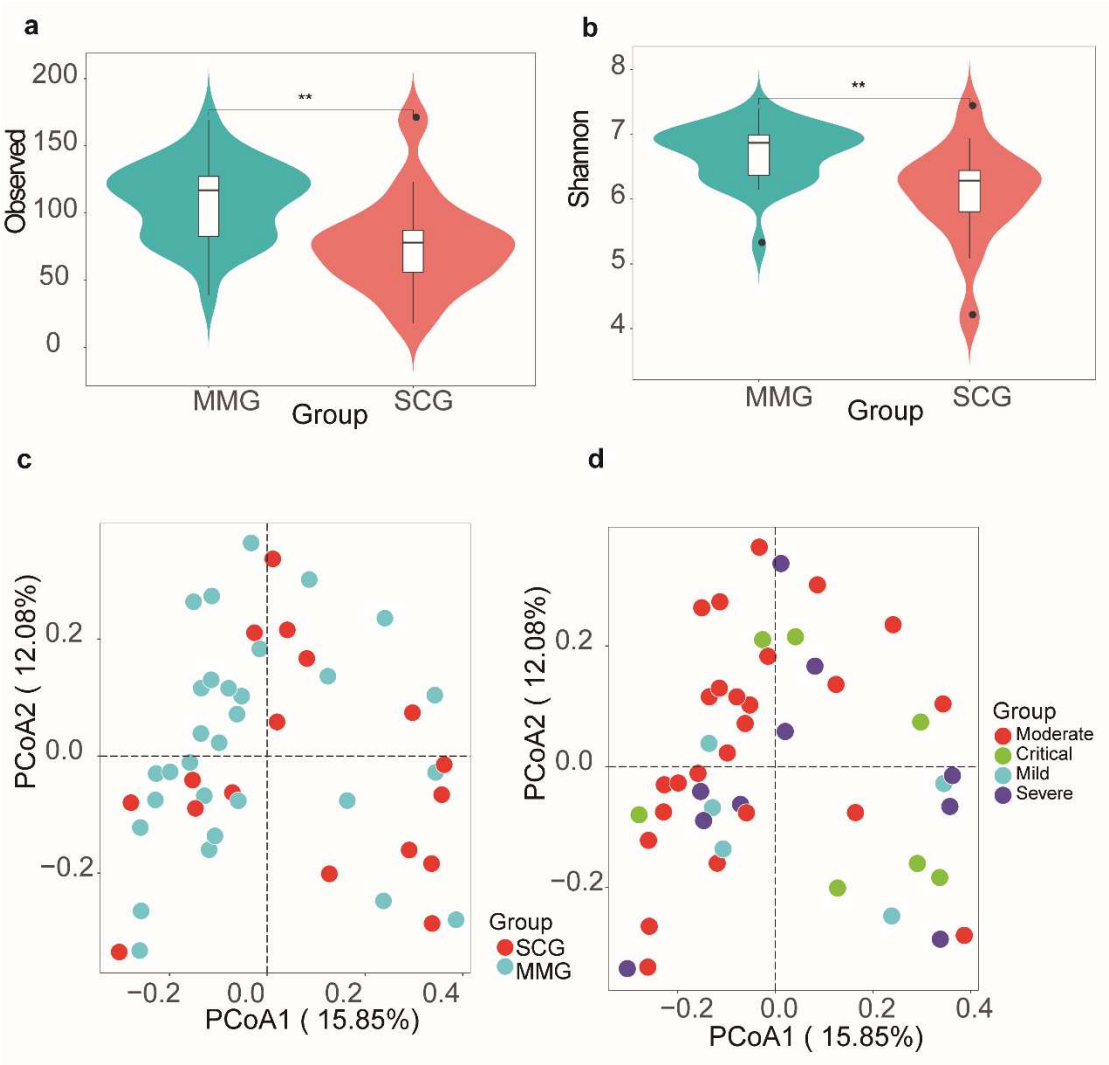
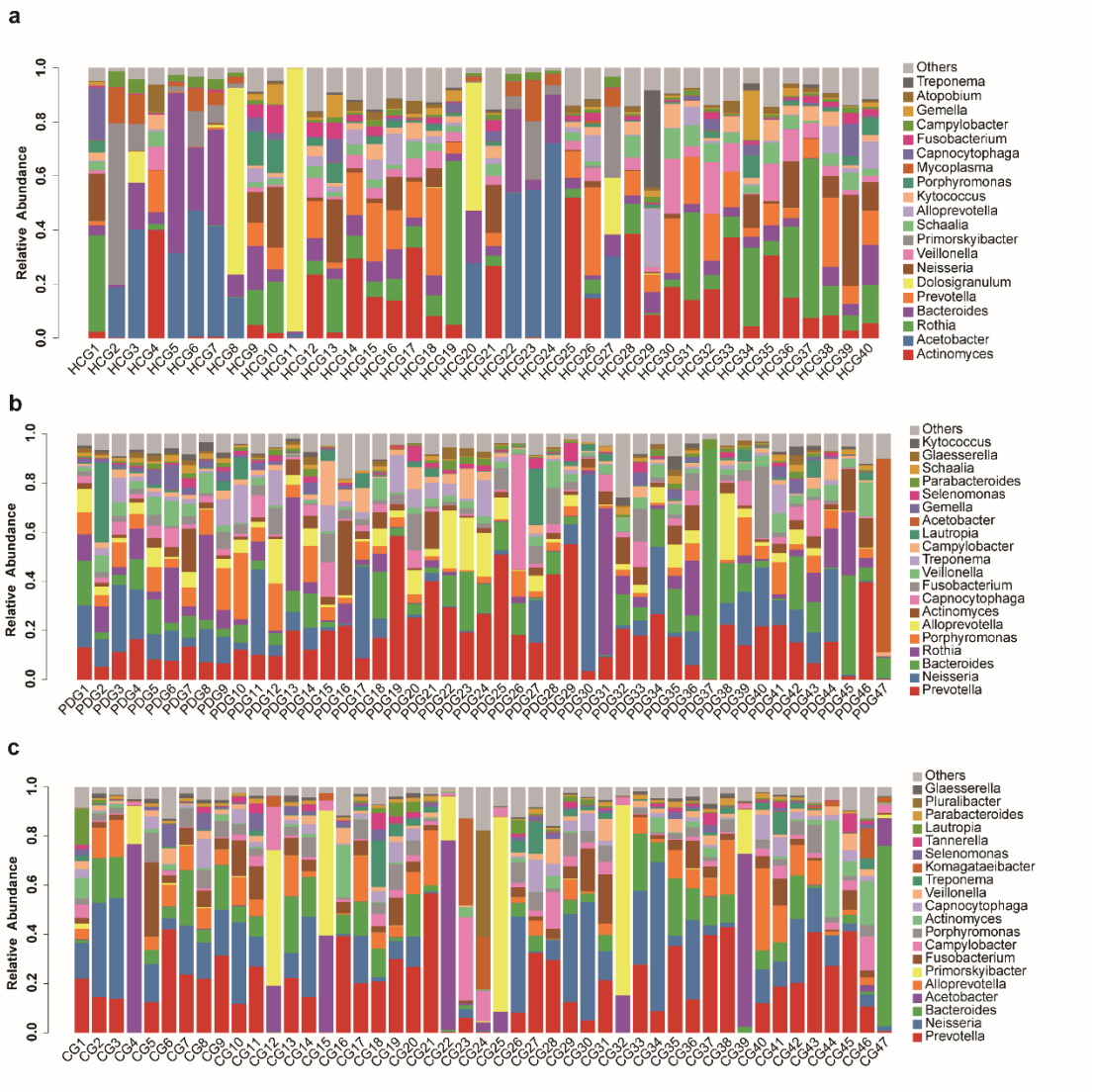
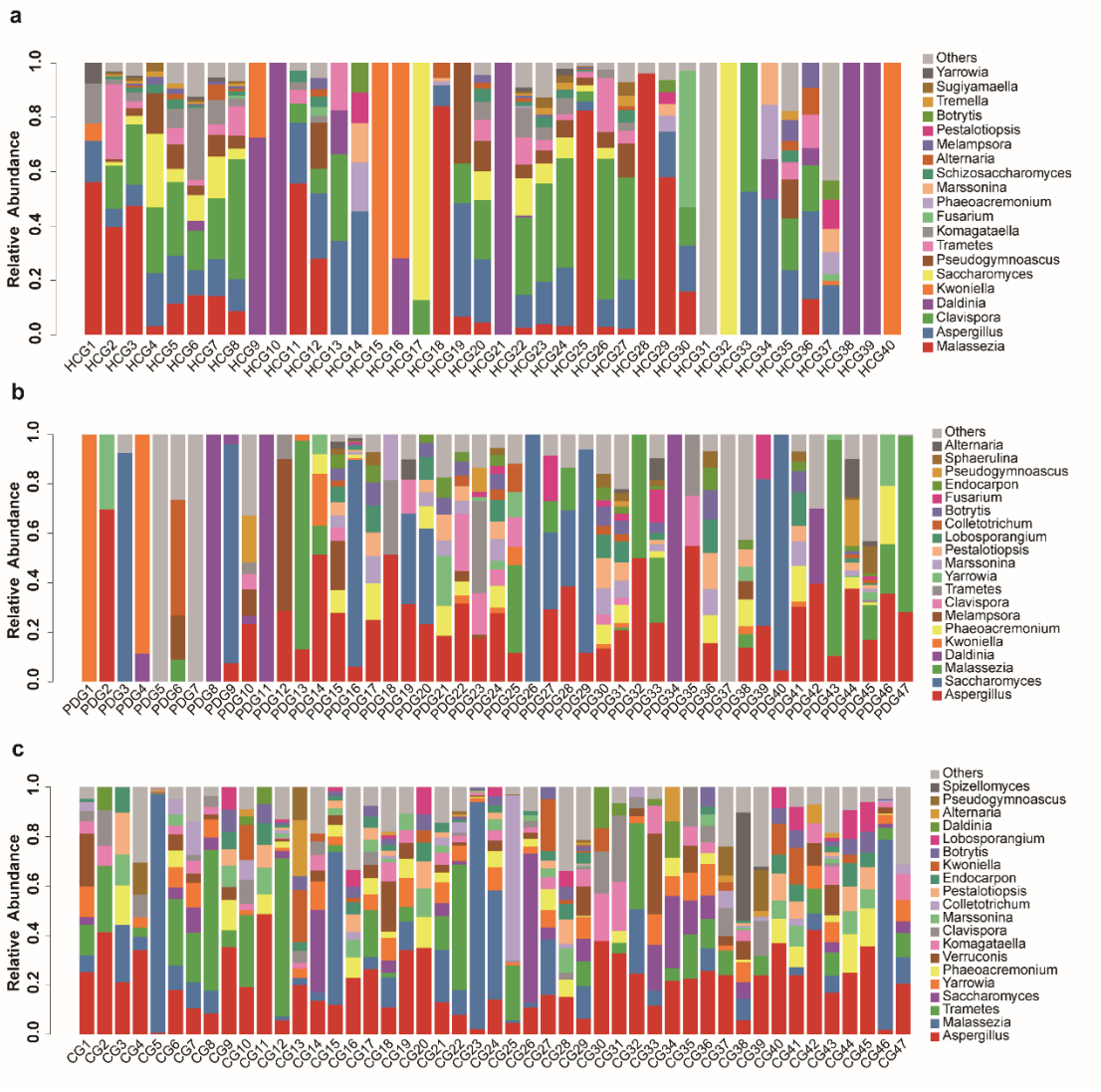


Fig.8 Decreased oropharyngeal microbiome in diversity appeared to be associated with COVID-19 severity. a Obviously decreased oropharyngeal microbiome in Observed index between MMG and SCG (**: $P < 0.001$). **b** Obviously decreased oropharyngeal microbiome in Shannon index between MMG and SCG (**: $P < 0.001$). **c** Principal coordinate analysis (PCoA) of Bray–Curtis distance showed obviously changes in β-diversity among MMG and SCG groups. **d** Principal coordinate analysis (PCoA) of Bray–Curtis distance showed obviously changes in β-diversity among Mild, Moderate, Severe and Critical groups. The colors represent different groups. PCoA1 and PCoA2 represent

the top two principal coordinates that captured most of the diversity. The fraction of diversity captured by the coordinate is valued as a percentage.



Suppl.Fig.1 Oropharyngeal bacterial characterization of every subjects among HCG, PDG and CG. Average compositions of relative abundance of the top20 bacterial communities for each group at the genus level, **a** for HCG, **b** for PDG and **c** for CG. shared bacterial landscape on genus level among HCG, PDG and CG group.



Suppl.Fig.1 Oropharyngeal fungal characterization of every subjects among HCG, PDG and CG. Average compositions of relative abundance of the top20 fungal communities for each group at the genus level, **a** for HCG, **b** for PDG and **c** for CG. shared bacterial landscape on genus level among HCG, PDG and CG group.

Figures

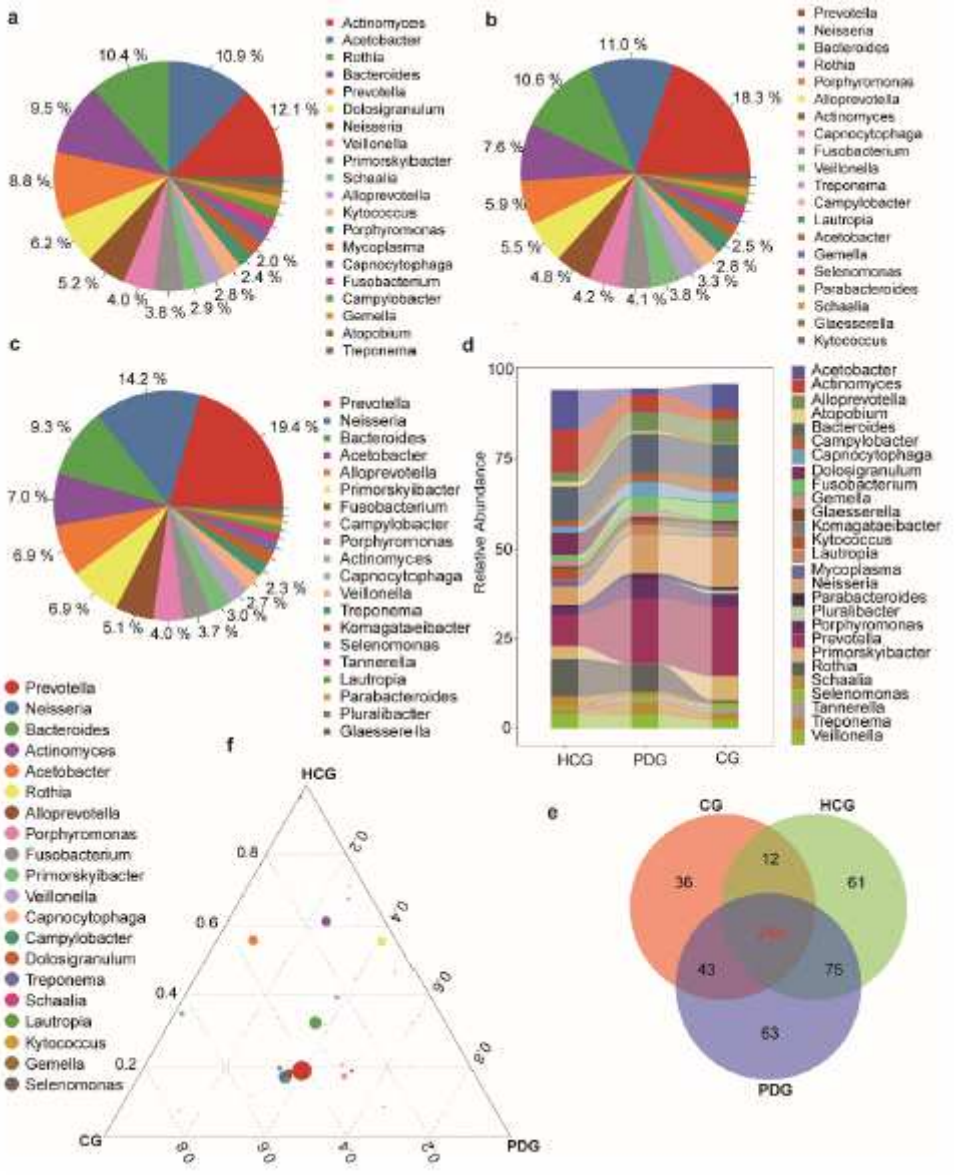


Figure 1

Oropharyngeal bacterial characterization and alterations among HCG, PDG and CG. Average compositions of relative abundance of the top20 bacterial communities for each group at the genus level, a for HCG, b for PDG and c for CG. d Alterations of top20 bacterial communities among the three groups at the genus level, Especially, Prevotella is significantly high er in PDG than that in HCG (Wilcoxon rank than that in HCG (Wilcoxon rank–sum test, sum test, pp□□0.001)0.001). . ee Venn diagram showing Venn diagram showing shared and unique genus of the three groups. shared and unique genus of the three groups. ff Ternary plots depicting the 244 Ternary plots depicting the 244 shared bacterial landscape on genus level among HCG, PDG and CG group. shared bacterial landscape on genus level among HCG, PDG and CG group. The sum of the proporThe sum of the proportion for one specific bacteria in the three groups was set tion for one specific bacteria in the three groups was set as 1, and the proportion of one

specific bacteria in each group is equal to its as 1, and the proportion of one specific bacteria in each group is equal to its corresponding relative abundance divided by the relative abundance sum of corresponding relative abundance divided by the relative abundance sum of this bacteria in the three groups. The sizethis bacteria in the three groups. The size of the circles represents the relative of the circles represents the relative abundance of the genus.abundance of the genus.

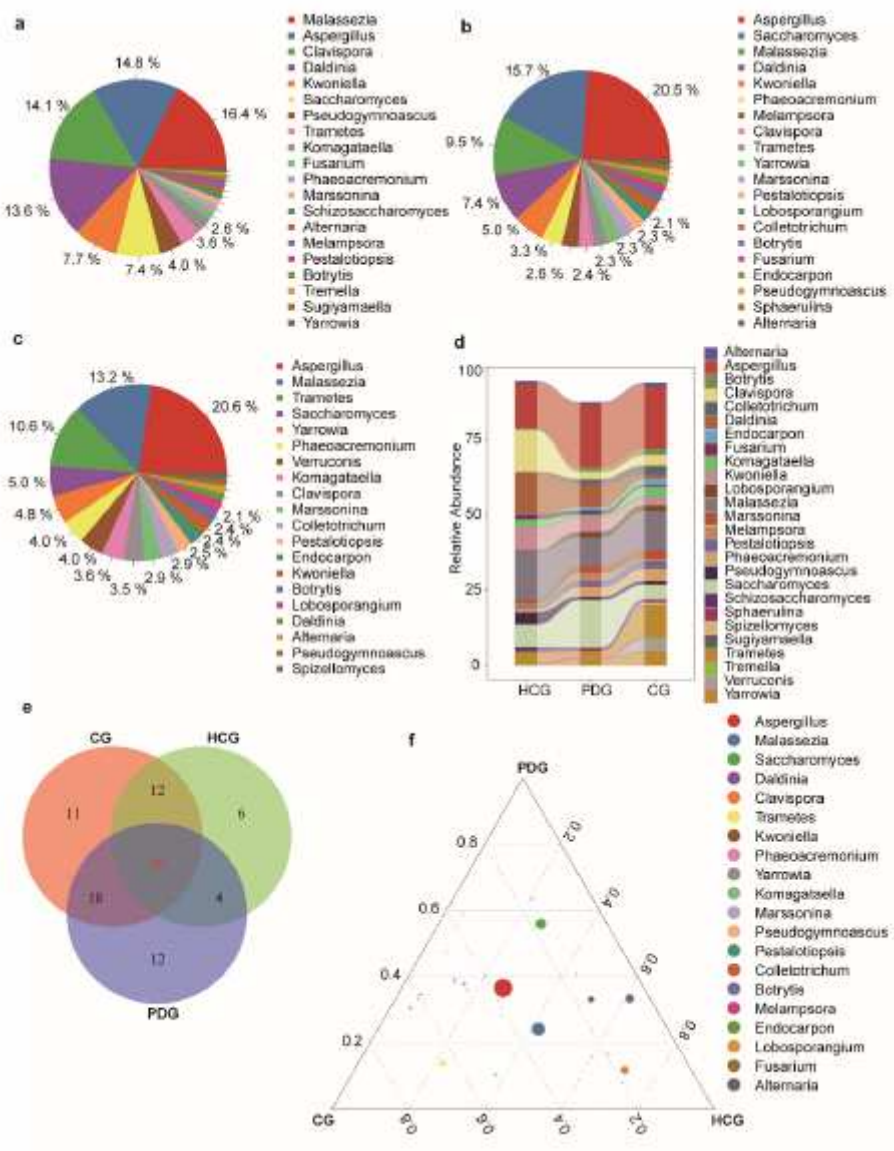


Figure 2

Oropharyngeal fungal characterization and alterations among HCG, PDG and CG. Average compositions of relative abundance of the top20 fungal communities for each group at the genus level, a for HCG, b for PDG and c for CG. d Alterations of top20 fungal communities among the three groups at the genus level. e Venn diagram showing shared and unique genus of the three groups. groups. ff Ternary plots depiTernary plots depicting the 244 shared cting the 244 shared fungalfungal landscape on genus landscape on genus level among HCG, PDG and CG group. The sum of the proportion for one level among HCG, PDG and CG group. The sum of the proportion for one specific bacteria in the three groups was set as 1, and the proportion of one specific bacteria in the three groups was set as 1, and the

proportion of one specific fungi in each group is equal to its corresponding relspecific fungi in each group is equal to its corresponding relative abundance ative abundance divided by the relative abundance sum of this fungi in the three groups. The divided by the relative abundance sum of this fungi in the three groups. The size of the circles represents the relative abundance of the genus.size of the circles represents the relative abundance of the genus.

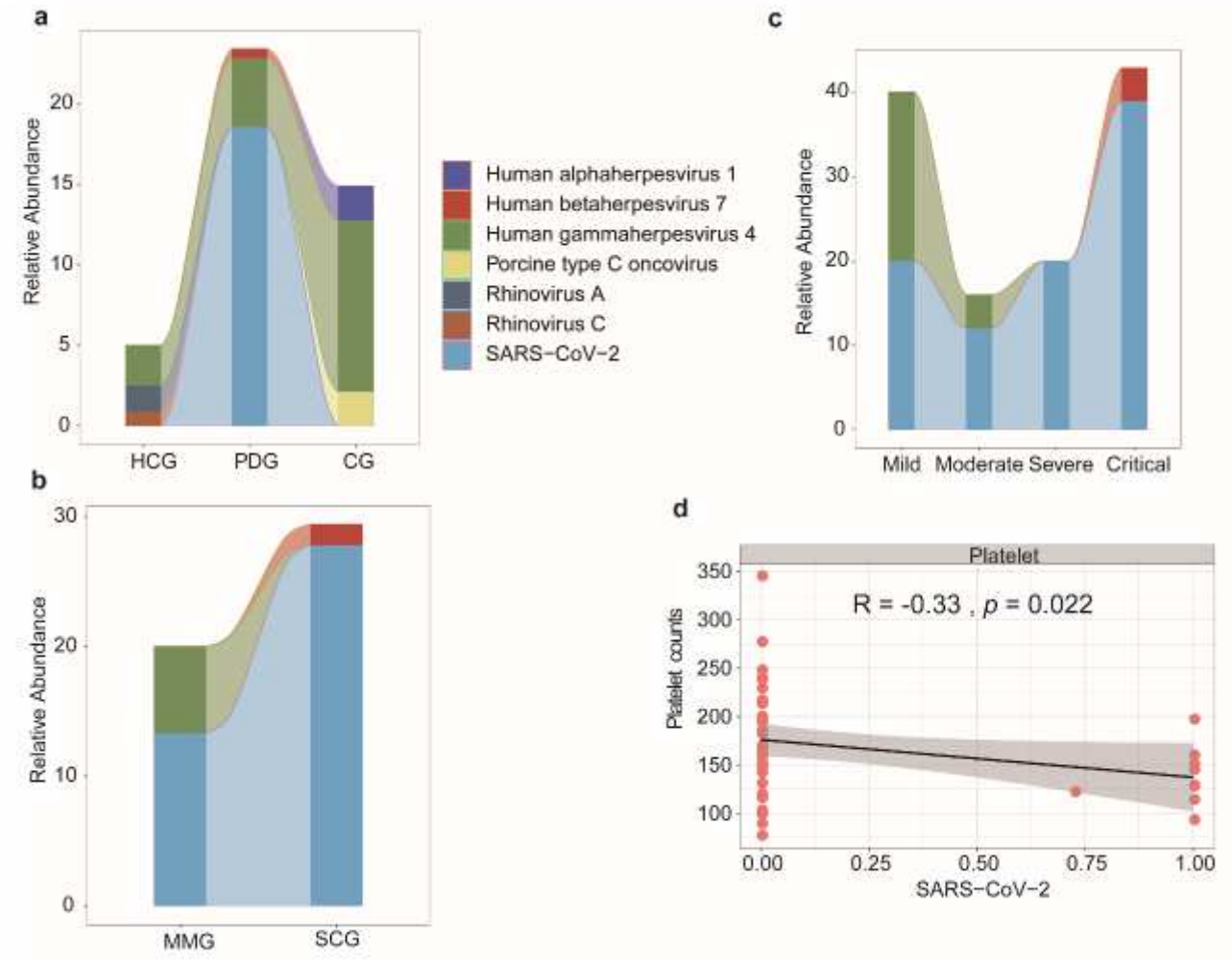


Figure 3

Oropharyngeal SARS CoV 2 viral load was positively correlated with the severity of COVI D 19. a The relative abundance changes of 7 oropharyngeal viral species (including SARS-COR-2) among HCG, PDG and CG group, b among MMG and SCG group, c among Mild, Moderate, Severe and Critical group. d Negative correlation between SARS-CoV-2 viral load (x-axis) and platelet counts (y-axis) (Spearman correlation analysis: R = −0.330, P = 0.022), Linear regression lines are shown in each scatter plot in black, and shaded regions represent 95% confidence intervals.

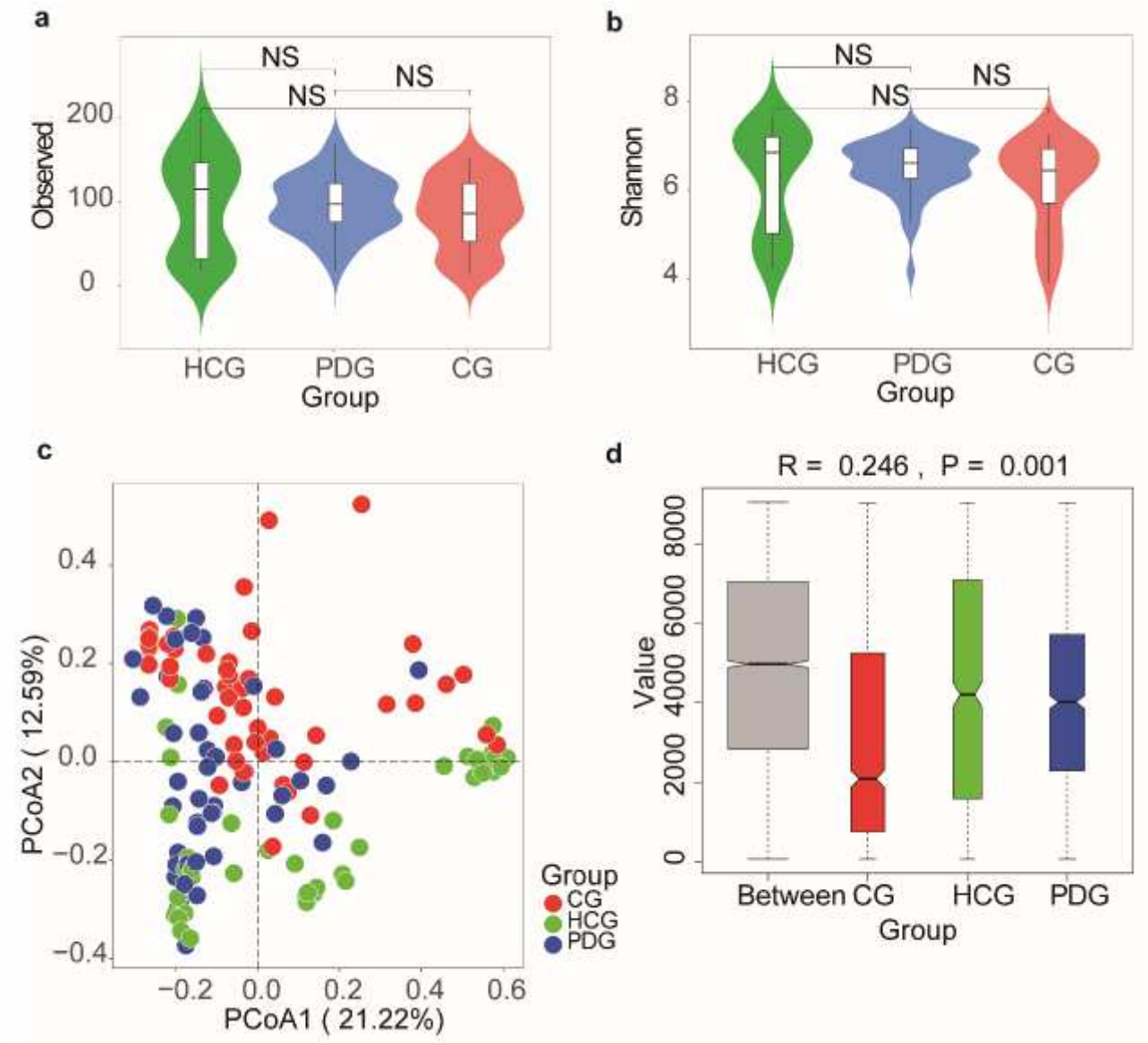


Figure 4

Oropharynx microbial alterations in diversity among HCG, PDG and CG. a Little changes in Observed index among HCG, PDG and CG (NS: P 0.05). b Little changes in Shannon index among HCG, PDG and CG (NS: P 0.05). c Principal coordinate analy sis (PCoA) of Bray Curtis distance showed obviously changes in β diversity among HCG, PDG and CG. The colors represent three different groups. PCoA1 and PCoA2 represent the top two principal coordinates that captured most of the diversity. The fraction of diversity captured by the coordinate is valued as a percentage. coordinate is valued as a percentage. dd. The analysis of similarities (ANOSIM) . The analysis of similarities (ANOSIM) based on Braybased on Bray––Curtis distance shows that R > 0 and Curtis distance shows that R > 0 and PP < 0.05, indicating that < 0.05, indicating that the interthe inter–group difference among HCG, PDG and CG was greater than the intragroup difference among HCG, PDG and CG was greater than the intra–group difference in each group. The R means the statistical value of ANOSIM, group difference in each group. The R means the statistical value of ANOSIM, yy–axis value means Brayaxis value means Bray––Curtis rank, and 999 was

used for the permutations.Curtis rank, and 999 was used for the permutations. NSNS means no significant.means no significant.

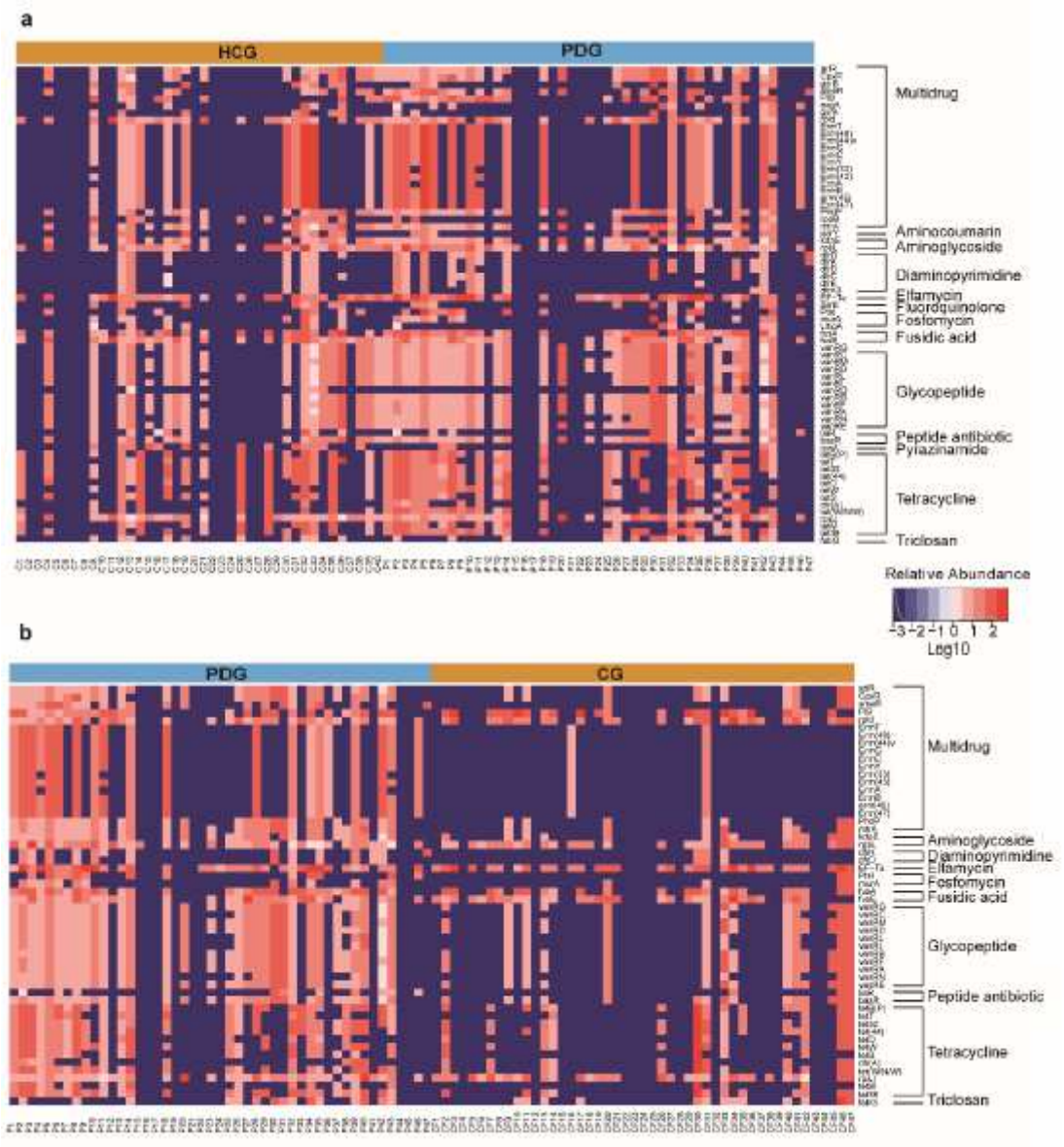


Figure 5

Obviously changes of oropharynx microbial antibiotic resistance genes expression among HCG, PDG and CG. a expression among HCG, PDG and CG. a Heatmap for the relative abundances Heatmap for the relative abundances of microbial antibiotic resistance genes expression between HCG and PDG. of microbial antibiotic resistance genes expression between HCG and PDG. bb Heatmap Heatmap for the relative abundances of microbial antibiotic resistance genes for the relative abundances of microbial antibiotic resistance genes expression between PDG and CG. The relative abundance changes from low expression between PDG and CG. The relative abundance changes from low to high is showed by color changes from blue to red. The log10 value was used to high is showed by color changes from blue to red. The log10 value was used to show the abundance ofto show the abundance of antibiotic reantibiotic resistance genesistance gene..

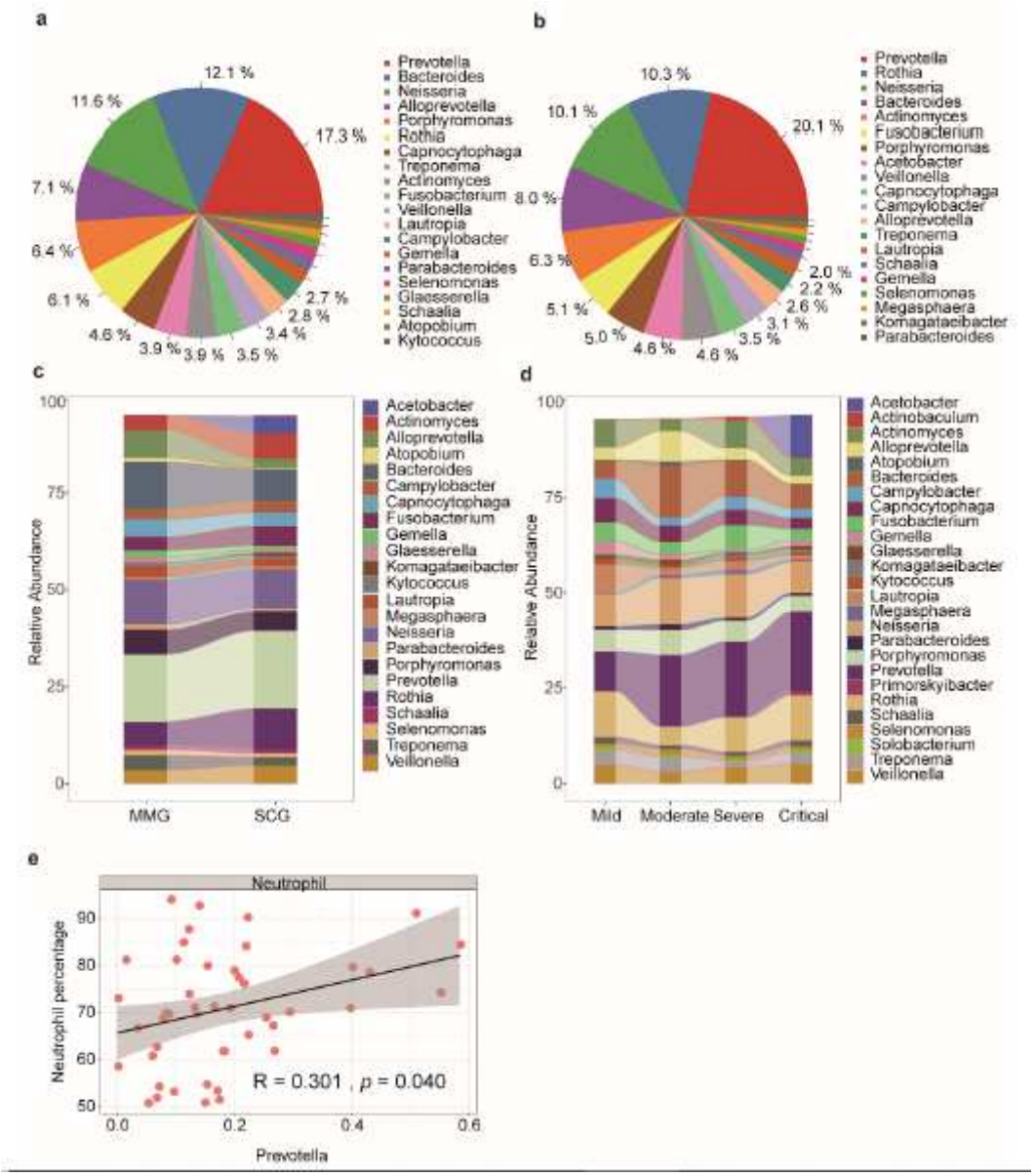


Figure 6

Alterations of oropharyngeal bacteria reflected disease severity in patients with COVID 19. a, b Average compositions of relative abundance of the top20 bacterial communities of MMG and SCG groups at the genus level. c Alterations of top20 bacterial communities among the MMG and SCG groups at the genus level. d Alterations of top20 bacterial communities among the Mild, Moderate, Severe and Critical groups at the genus level. Especially, the alteration of Prevotella gradually increased along with the symptom aggravated in patients with COVID 19. e Positive correlation between Prevotella relative abundance (x-axis) and the elevation of Neutrophil percentage (y-axis) (Spearman correlation analysis: R = 0.301, p = 0.040), Linear regression lines are shown in each scatter plot in black, and shaded regions represent 95% confidence intervals

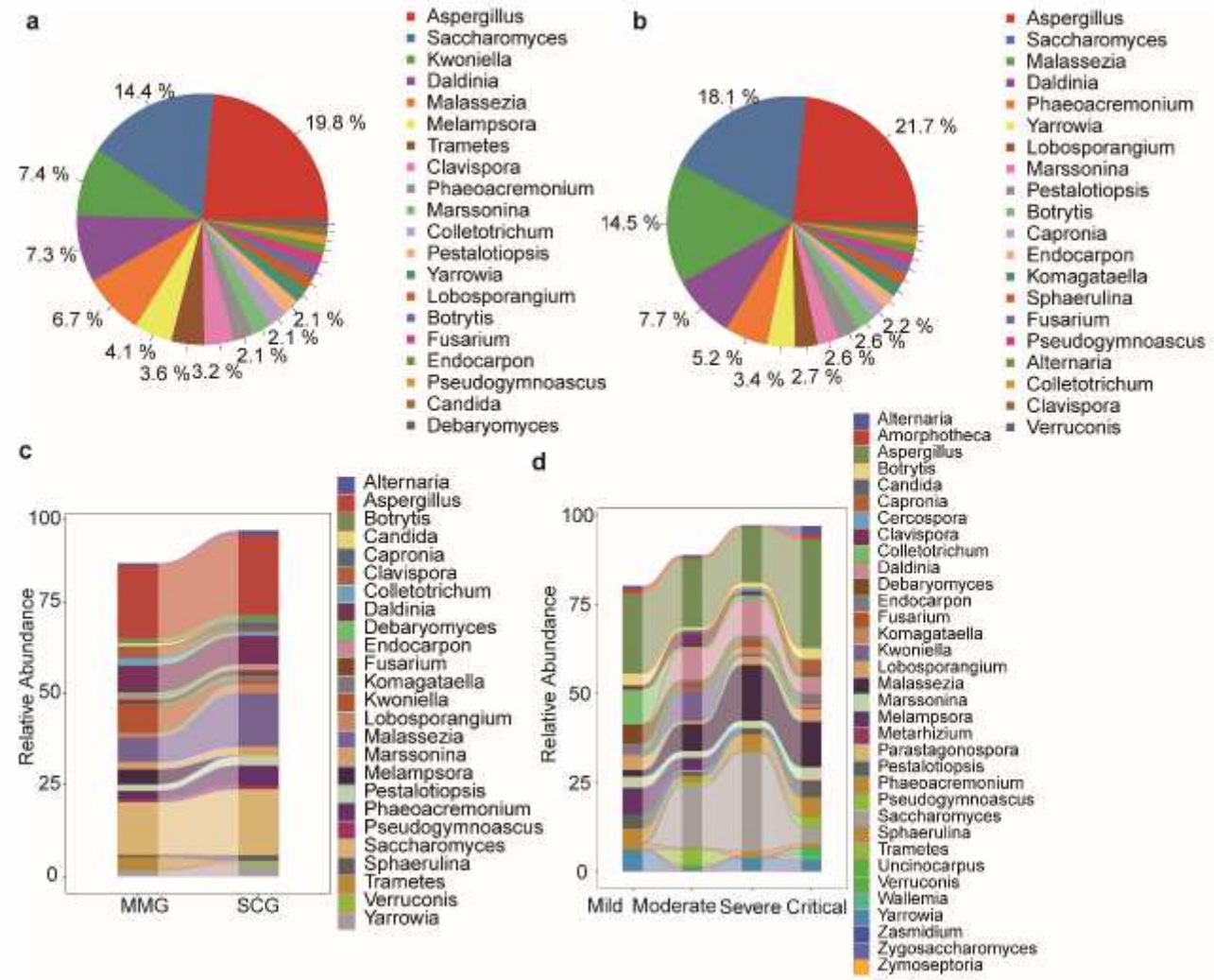


Figure 7

Alterations of oropharyngeal fungi reflected disease severity in patients with COVID 19. a, b Average compositions and relativ e abundance of the top20 fungal communities of MMG and SCG groups at the genus level. c Alterations of top20 fungal communities among the MMG and SCG groups at the genus level. d Alterations of top20 fungal communities among the Mild, Moderate, Severe and Critical groups at the genus level. Especially Aspergillus increased remarkably in critical Patients with COVID 19.

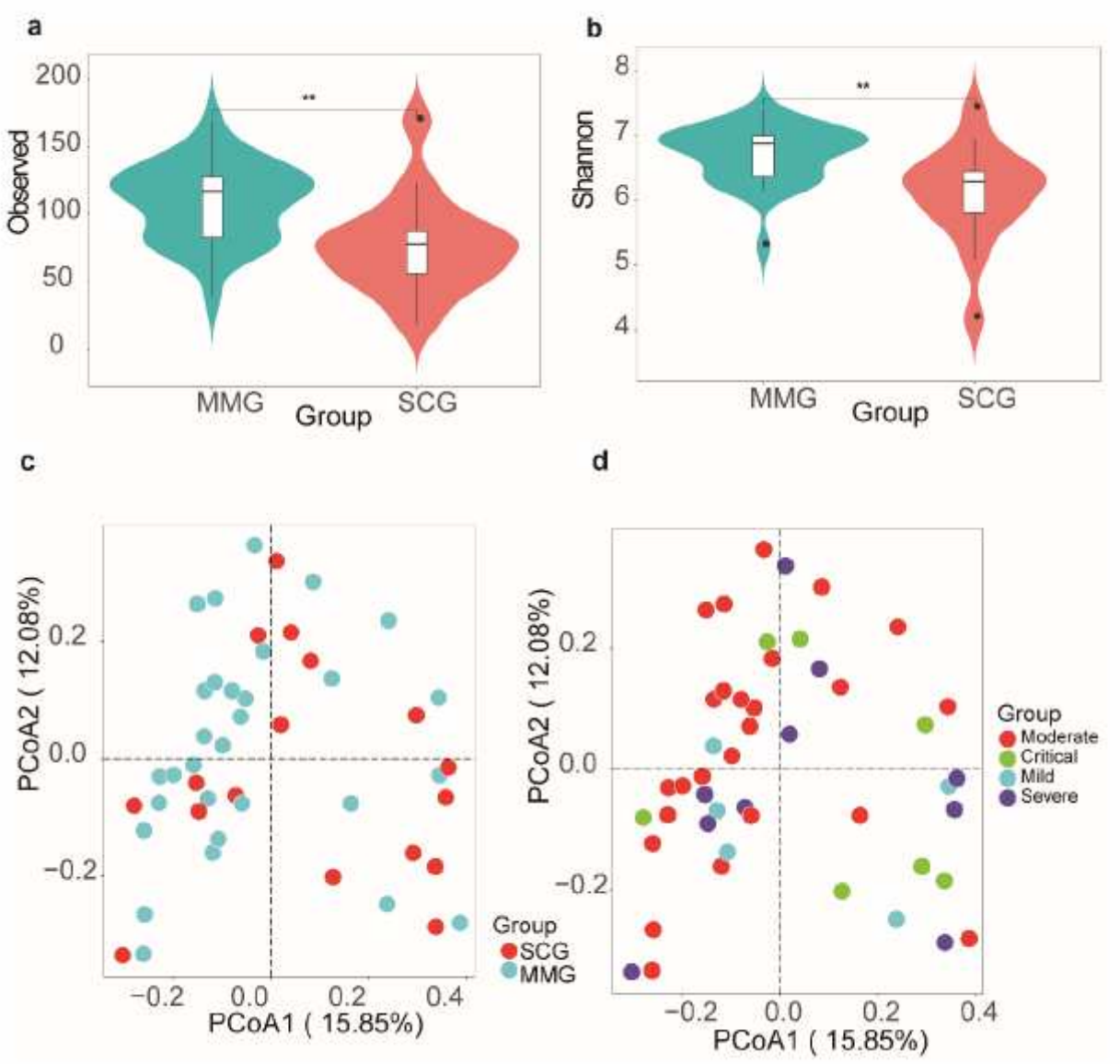


Figure 8

Decreased oropharyngeal microbiome in diversity appeared to be associated with COVID 19 severity. a Obviously decreased oropharyngeal microbiome in Observed index between MMG and SCG (**: P 0.001). b Obviously decreased oropharyngeal microbiome in Shannon index between MMG and SCG (**: P 0.001). c Principal coordinate analysis (PCoA) of Bray Curtis distance showed obviously changes in β diversity among MMG and SCG groups. d Principal coordinate analysis (PCoA) of Bray Curtis distance showed obviously changes in β diversity among Mild, Moderate, Severe and Critical groups. The colors represent different groups. PCoA1 and PCoA 2 represent the top two principal coordinates that captured most of the diversity. The the top two principal coordinates that captured most of the diversity. The fraction of diversity captured by the coordinate is valued as a percentage.fraction of diversity captured by the coordinate is valued as a percentage.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- FigS1.png