

# Symptom-related sputum microbiota in stable chronic obstructive pulmonary disease

Wenqi Diao<sup>1,\*</sup>  
Ning Shen<sup>1,\*</sup>  
Yipeng Du<sup>1,\*</sup>  
John R Erb-Downward<sup>2</sup>  
Xiaoyan Sun<sup>1</sup>  
Chenxia Guo<sup>1</sup>  
Qian Ke<sup>1</sup>  
Gary B Huffnagle<sup>2,3</sup>  
Margaret R Gyetko<sup>2</sup>  
Bei He<sup>1</sup>

<sup>1</sup>Department of Respiratory Medicine, Peking University Third Hospital, Beijing 100191, China;

<sup>2</sup>Division of Pulmonary and Critical Care Medicine, Department of Internal Medicine, University of Michigan Medical School, Ann Arbor, MI, USA; <sup>3</sup>Department of Microbiology and Immunology, University of Michigan Medical School, Ann Arbor, MI, USA

\*These authors contributed equally to this work

Correspondence: Bei He  
Department of Respiratory Medicine,  
Peking University Third Hospital,  
Beijing 100191, China  
Email puh3\_hb@bjmu.edu.cn

**Background:** The role of airway microbiota in COPD is highly debated. Symptomology assessment is vital for the management of clinically stable COPD patients; however, the link between symptoms and the airway microbiome is currently unknown.

**Purpose:** The present study aimed to evaluate the relationship among stable COPD patients.

**Patients and methods:** We conducted pyrosequencing of bacterial 16S rRNA using induced sputum samples in a Han Chinese cohort that included 40 clinically stable COPD patients and 19 healthy controls.

**Results:** Alterations in community composition and core bacterial taxa (*Neisseria subflava*, etc.) were observed in patients with severe symptoms compared with controls. The co-occurrence network indicated that the key microbiota enriched in COPD patients showed higher expression in patients with severe symptoms. The association pattern of symptoms with the sputum microbiome was obviously different from that of lung function in COPD patients.

**Conclusion:** These findings broaden our insights into the relationship between the sputum microbiota and the symptom severity in COPD patients, emphasizing the role of symptoms in the airway microbiome, independent of lung function.

**Keywords:** chronic obstructive pulmonary disease, symptom, sputum, microbiome

## Introduction

COPD is a chronic inflammatory lung disorder that is characterized by irreversible airflow limitation and remains a major public health problem.<sup>1</sup> Cigarette smoking is the principal cause of COPD, but only ~15% of adults with substantial tobacco exposure develop clinical COPD.<sup>2,3</sup> Bacterial colonization or infection is also considered to be an important factor in COPD.<sup>4</sup> Traditional culture-based strategies have demonstrated that bacteria such as *Haemophilus influenzae*, *Streptococcus pneumoniae*, and *Moraxella catarrhalis* may be a key etiological factor exacerbating COPD,<sup>5-7</sup> however, the role of microorganism in clinically stable patients remains unclear, although some studies indicated that their presence may contribute to chronic inflammation in the airway.<sup>8</sup>

In recent years, numerous researchers have revealed the relationship between the lung microbiome and COPD using culture-independent sequencing techniques, such as 16S rRNA pyrosequencing.<sup>9-12</sup> These studies strongly emphasized microbial complexity and the significant heterogeneity among COPD patients.<sup>9-13</sup> This intersubject variation may be caused by the severity of lung function and patient exacerbation status.<sup>14-20</sup> However, the heterogeneity of COPD cannot be fully explained by these factors.<sup>21,22</sup> Symptomology assessment is vital for the management of clinically stable COPD patients; however, the relationship between symptoms and the airway microbiome is currently unknown.<sup>1</sup>

Here, we conducted 16S rRNA-based pyrosequencing to evaluate the relationship between the sputum microbiota and the COPD symptoms in a Han Chinese population.

We reported symptom-related bacterial taxa and their relationships with lung function in COPD patients.

## Patients and methods

### Subjects and study design

This study (NCT03310164) was approved by the Peking University Institutional Review Board (PUIRB, identifier: IRB00001052-11067) and was strictly conducted according to the designed protocols. Written informed consent was obtained from all participants.

Outpatients with COPD ( $n=40$ ) were enrolled in the clinic of Peking University Third Hospital. The inclusion criteria have been previously depicted,<sup>23,24</sup> and the brief description was shown as follows: 1) males 40–80 years of age; 2) a smoking history of  $>10$  pack-years; 3) diagnosis of COPD according to the Global Initiative for Obstructive Lung Disease (GOLD);<sup>1,25</sup> 4) lack of exacerbation events or use of systemic antibiotics within the past 3 months; and 5) lack of severe organ dysfunction or active disease. The control group ( $n=19$ ) consisted of male smokers without COPD.

A spirometry test was performed in all subjects using a ventilator (SensorMedics, Yorba Linda, CA, USA). The predicted percent of forced expiratory volume in 1 second ( $FEV_1\%$ pred) was used to evaluate the severity of lung function, as follows: GOLD I,  $FEV_1\%$ pred  $>80\%$ ; GOLD II,  $80\% > FEV_1\%$ pred  $\geq 50\%$ ; GOLD III,  $50\% > FEV_1\%$ pred  $\geq 30\%$ ; and GOLD IV,  $FEV_1\%$ pred  $<30\%$ . The ratio of residual volume to total lung capacity (RV/TLC) was used to assess the degree of emphysema. The symptoms of COPD patients were quantified with the St George's Respiratory Questionnaire (SGRQ): SGRQ score of  $<25$  was considered to indicate "mild symptoms", and SGRQ score of  $\geq 25$  was considered to indicate "severe symptoms". The cutoff value (SGRQ=25) was equal to the threshold 10 for the COPD Assessment Test (CAT).<sup>1,26</sup> Information regarding age, body mass index (BMI), smoking history, bronchodilators, inhaled corticosteroids (ICSs), exacerbation history in the past 1 year, and comorbidities (hypertension and diabetes) was also collected using questionnaires.

### Sample collection

The subjects were required to stop smoking and stop using corticosteroids and bronchodilators for at least 12 hours prior to sample collection. The sputum induction was performed by an experienced technician under the supervision of a physician. The mouth and posterior oropharynx were rinsed by swishing and gargling with normal saline three times. Then, the subjects received an ultrasonic nebulizer inhalation of 3% hypertonic

saline for up to 20 minutes before coughing up sputum until a volume of at least 1 mL was collected. The collected samples were stored at  $-80^\circ\text{C}$  prior to DNA extraction.

### Bacterial DNA extraction, polymerase chain reaction (PCR) amplification, and pyrosequencing

The samples were treated with an equal volume of 0.1% dithiothreitol (Calbiochem, La Jolla, CA, USA) for 15 minutes and aliquoted into sterile Eppendorf tubes. After centrifugation at  $13,000\times g$  for 15 minutes, bacterial DNA was extracted from the cell pellets of the aliquoted samples (500  $\mu\text{L}$ ) with a DNA extraction kit (Qiagen, Hilden, Germany) based on a modified protocol.<sup>27</sup> The DNA from sterile water and acute exacerbation (AE) buffer were collected as a potential source of contamination. The total copy number of the 16S rDNA was quantified through real-time PCR (Thermo Fisher Scientific, Waltham, MA, USA).<sup>9</sup> PCR amplification and pyrosequencing were conducted at BGI (Shenzhen, China) according to the protocols supplied by the laboratory of GBH.<sup>9</sup> Briefly, the V3–V5 region of the 16S rRNA gene was amplified by PCR using the universal primers 907R (5'-CCGTCAATTCMTTGGAGTTT-3') and 338F (5'-ACTCCTACGGGAGGCAGCAG-3'), with specific 10-digit barcodes for each sample. The purified amplicons were sequenced with the Roche 454 platform (Roche Applied Science, Basel, Switzerland).

### Microbiome analyses

Mothur software (version 1.38.1, University of Michigan, Ann Arbor, MI, USA) was used to process the raw sequencing data, as described elsewhere.<sup>28</sup> Briefly, we allowed one mismatch to the sample barcode and two mismatches to the adjacent primer. These filtered sequences were aligned through a SILVA-based (version 128) reference database. Chimeric sequences were eliminated with the UCHIME command. The processed sequences were classified with Ribosomal Database Project training set 16.0. Finally, high-quality sequences were annotated to operational taxonomic units (OTUs) at a 3% distance cutoff.

### Statistical analysis

To determine the spatial separation of bacterial communities between groups, principal coordinate analysis was plotted with the Bray–Curtis distance matrix. In addition, permutational multivariate analysis of variance (PERMANOVA) with the Bray–Curtis distance matrix was used to quantify differences in community composition. The mean abundance of OTUs was compared between groups through a Welch's *t*-test or

through ANOVA with Tukey's post hoc test as appropriate. Categorical variables were compared between groups by Fisher's exact test. The associations between OTUs and clinical information were evaluated with Spearman's coefficient. The co-occurrence network was built with Cytoscape (version 3.5.1). Spearman's coefficient of  $>0.5$  was defined as "co-occurrence", and Spearman's coefficient of  $<-0.5$  was defined as "mutual exclusion". A *P*-value of  $<0.05$  was defined as statistically significant. All statistical analyses were performed in R software (version 3.4.0).

## Results

### Study demographics

A total of 59 smokers were enrolled in this study, with 19 healthy controls and 40 COPD patients. The COPD patients were divided into two subgroups: "mild symptoms" (SGRQ  $<25$ ) and "severe symptoms" (SGRQ  $\geq 25$ ). Age, BMI, smoking history, exacerbation ( $\geq 2$ ), medication, and comorbidities were similar between the groups, as summarized in Table 1.

### Shifted community composition in COPD patients with severe symptoms

Figure 1A shows a huge heterogeneity among COPD patients. The patients in the X-zone were obviously different from the controls, but the patients in the Y-zone were similar to

the controls. We observed a separation trend only between COPD patients and the controls, as presented in Figure 1B, but this trend was not statistically significant ( $P=0.101$ , PERMANOVA). We found that the patients in the Z-zone had severe symptoms (SGRQ  $>25$ ) and markedly different from controls, as shown in Figure 1C. Accordingly, patients with severe symptoms were significantly different from controls ( $P=0.034$ , PERMANOVA), and no difference was observed between patients with mild symptoms and controls ( $P=0.286$ , PERMANOVA), as illustrated in Figure 1D. These results suggested that the observed intersubject heterogeneity may be partly explained by the symptom severity of COPD patients.

The PERMANOVA model, which corrected for spirometry severity, the use of ICS or bronchodilators, and comorbidity history (hypertension and diabetes) confirmed the previous univariate PERMANOVA findings, as shown in Table S1. The community composition was not shifted in patients with different severities of lung function and was not affected by the use of ICS, as presented in Figures S1 and S2, respectively.

### Relative abundance of OTUs is altered in COPD patients with severe symptoms

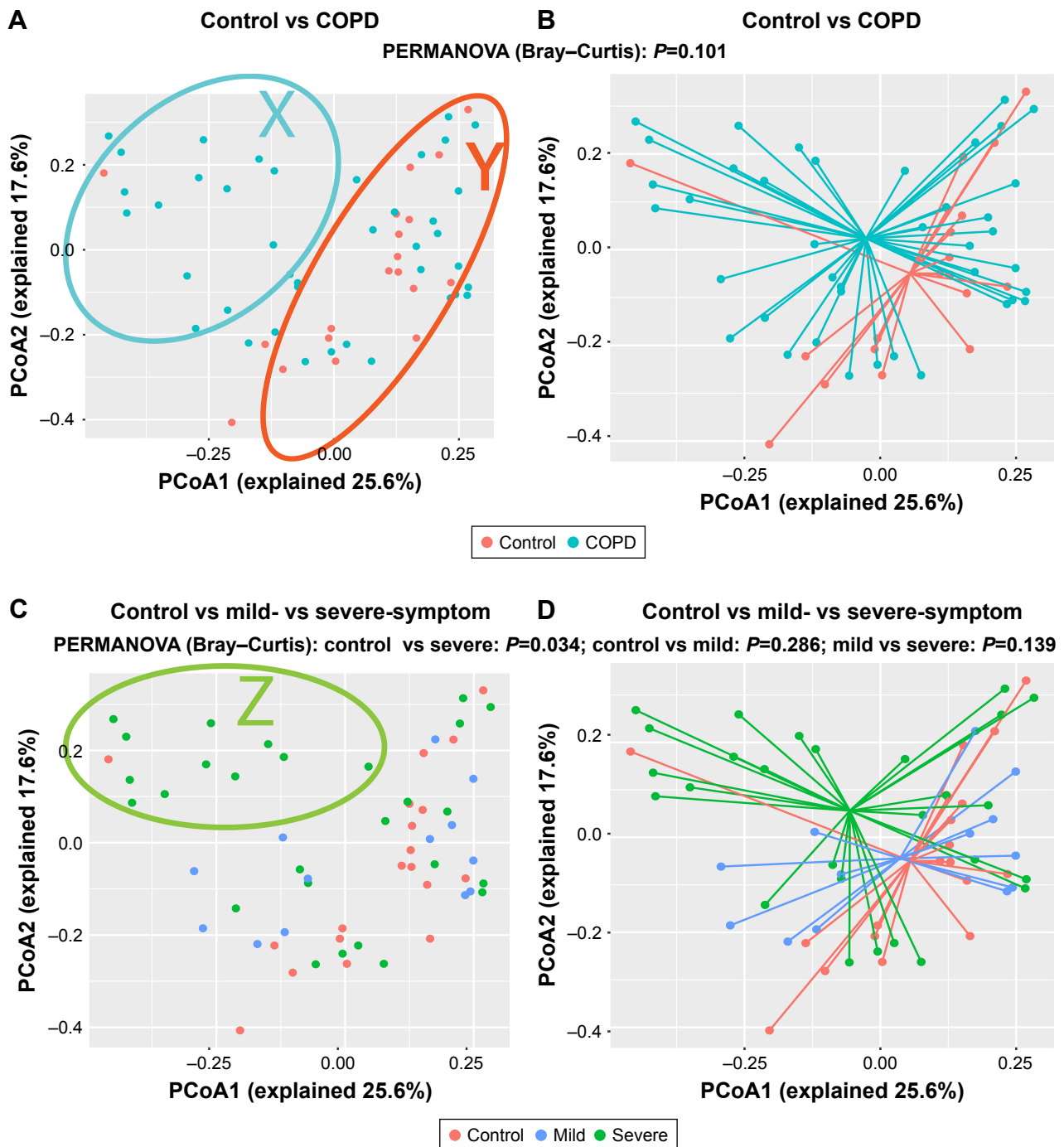
To identify the key microbial drivers, we plotted a rank curve of the top 20 OTUs and analyzed abundance differences between groups, as shown in Figure 2. Compared with the controls,

**Table 1** Baseline data of participants

Variables	Control (n=19)	COPD (n=40)		
		All (n=40)	Mild symptom (n=13)	Severe symptom (n=27)
Age (years)	58.7±6.5	61.5±6.5	63±5.2	60.8±7.1
BMI (kg/m <sup>2</sup> )	25±3.0	24.1±3.5	23.9±2.3	24.2±4.1
Smoking history (pack-years)	33.8±14.8	37.1±21.1	30.2±13.6	40.4±23.4
FEV <sub>1</sub> %pred	98.4±11	57.7±17.4**	57.7±12.0**	57.7±19.7**
RV/TLC	35.4±9.1	52.8±12.4**	51.5±12.7**	53.4±12.4**
SGRQ score	NA	37.1±17.7	18.5±4.2	46±14.4
ICS	NA	11 (27.5%)	5 (38.4%)	6 (22.2%)
Bronchodilator	NA	38 (95.0%)	12 (92.3%)	26 (96.3%)
Exacerbation ( $\geq 2$ ) <sup>a</sup>	NA	1 (2.5%)	0 (0%)	1 (3.7%)
Hypertension	3 (15.8%)	7 (17.5%)	2 (15.4%)	5 (18.5%)
Diabetes	2 (10.5%)	4 (10.0%)	1 (7.7%)	3 (11.1%)
GOLD classification				
I	NA	5 (12.5%)	0 (0%)	5 (18.5%)
II	NA	19 (47.5%)	8 (61.5%)	11 (40.7%)
III	NA	14 (35.0%)	5 (38.5%)	9 (33.3%)
IV	NA	2 (5.0%)	0 (0%)	2 (7.4%)

**Notes:** Continuous variables are presented as "mean±SD"; categorical variables are shown as "number (%)". \*\* $P<0.01$ : the group is significantly different from control. "Mild symptom" represents COPD patients of SGRQ  $<25$ ; and "severe symptom" denotes COPD patients of SGRQ  $\geq 25$ . <sup>a</sup>Number of 2 or greater than exacerbation events in the past 1 year.

**Abbreviations:** BMI, body mass index; FEV<sub>1</sub>%pred, the predicted percent of forced expiratory volume in 1 second; RV/TLC, ratio of residual volume to total lung capacity; SGRQ, St George's respiratory questionnaire; ICS, inhaled corticosteroid; GOLD, Global Initiative for Obstructive Lung Disease; NA, not applicable.



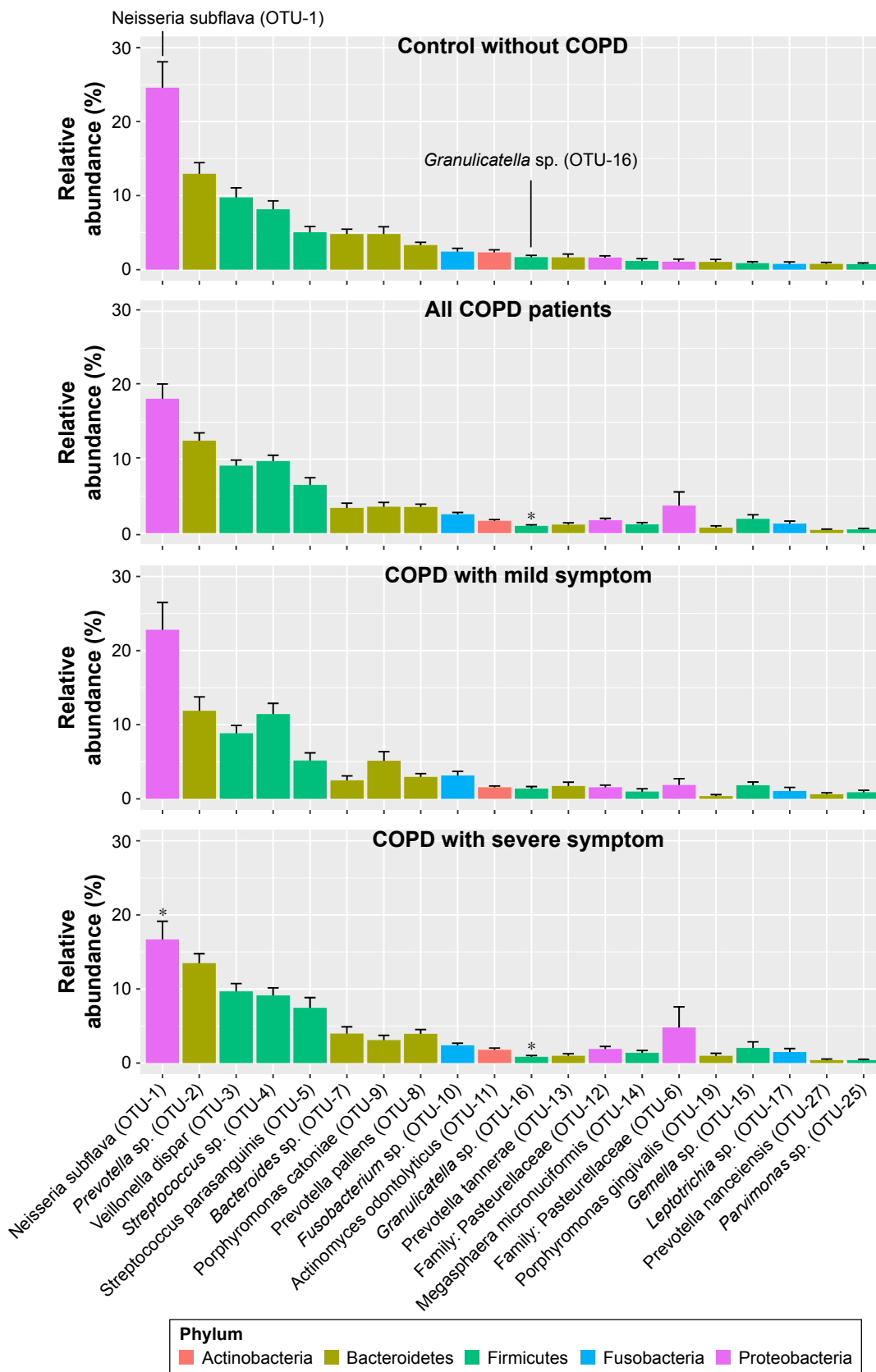
**Figure 1** Community composition is shifted in COPD patients with severe symptoms.

**Notes:** Unsupervised PCoA was plotted based on the Bray–Curtis distance. PERMANOVA was used to determine differences in community composition between groups. (A) Control and COPD are labeled red and blue, respectively. In the X-zone (blue circle), the samples mostly belong to the COPD group; in the Y-zone (red circle), the control group and the COPD group cannot be clearly separated. (B) The control group and the COPD group with centroids added. (C) The COPD patients included two subgroups: “mild symptoms” (SGRQ <25) and “severe symptoms” (SGRQ  $\geq$ 25). Control, mild symptoms, and severe symptoms are labeled red, blue, and green, respectively. In the Z-zone (green circle), most of the samples belong to the severe symptoms group. (D) The control, mild symptoms, and severe symptoms groups with centroids added.

**Abbreviations:** PCoA, principal coordinate analysis; PERMANOVA, permutational multivariate analysis of variance; SGRQ, St George’s Respiratory Questionnaire.

*Granulicatella* sp. (OTU-16) was reduced in all patients and in the group of patients with severe symptoms. In addition, *Neisseria subflava* (OTU-1) was reduced only in patients with severe symptoms but was not altered in all COPD patients. These findings suggested that decreased *Granulicatella* sp.

and *N. subflava* may be the main contributors to the community discrepancy between patients with severe symptoms and controls. The two bacteria were not detected in sterile water or AE buffer, as shown in Figure S3, demonstrating that they did not originate from contamination.



**Figure 2** Relative abundance of OTUs is altered in COPD patients with severe symptoms.

**Notes:** The top 20 most abundant OTUs identified in the sputum are presented. Asterisks imply that OTUs are significantly altered in this group compared to the controls. Welch's t-test was used to determine the statistical difference between the controls and COPD patients, and ANOVA with Tukey's multiple comparisons test was used to detect the statistical discrepancy between the controls and COPD patients with mild (SGRQ<25) and severe symptom (SGRQ≥25). Values are shown as the mean ± SEM.

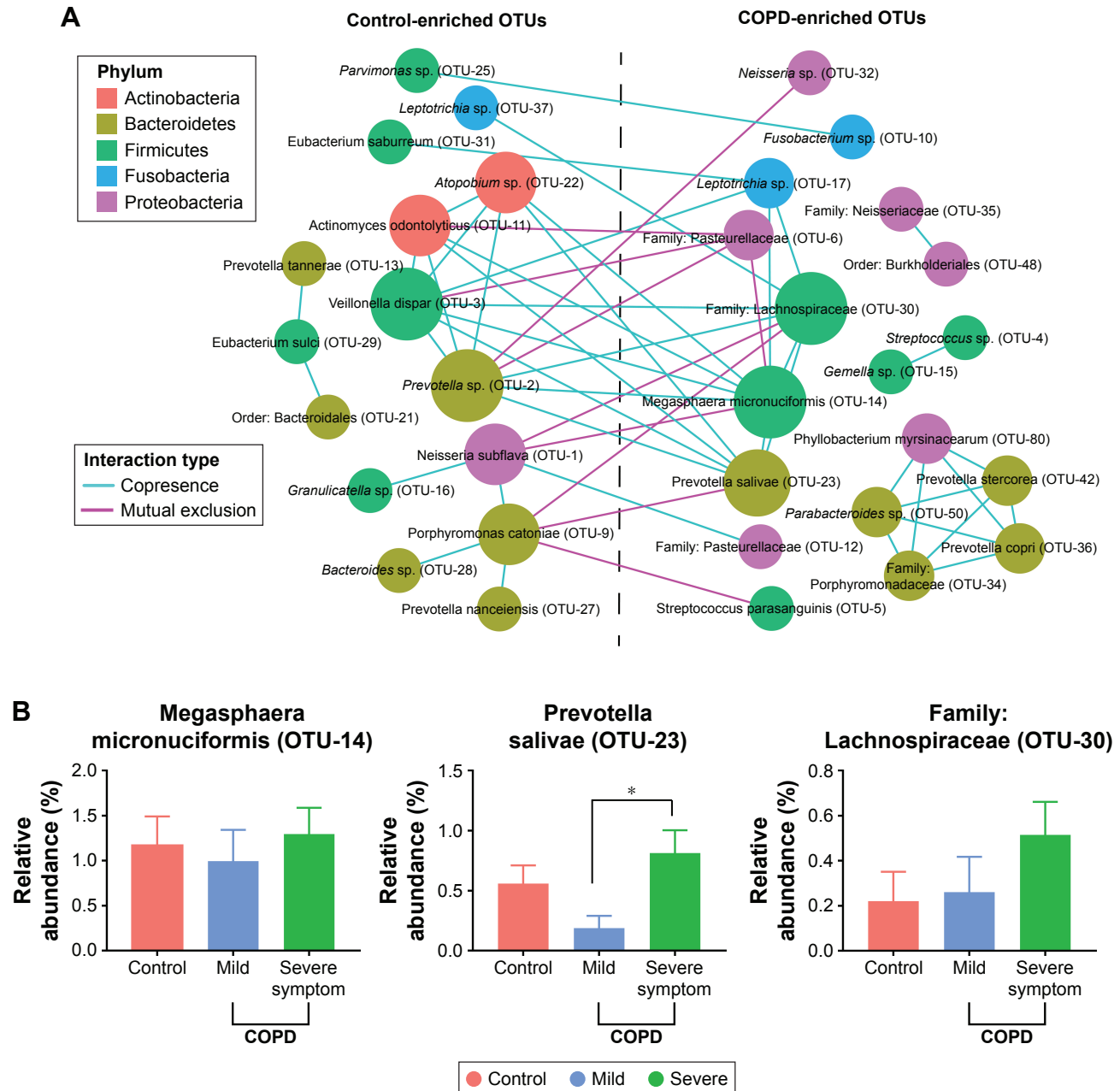
**Abbreviations:** ANOVA, analysis of variance; OTU, operational taxonomic unit; SEM, standard error of the mean; SGRQ, St George's Respiratory Questionnaire.

## Interaction of OTUs enriched in controls and COPD patients

Figure 3A depicts the interaction between control- and COPD-enriched OTUs. The blue and red edges represent co-occurrence and mutual exclusion, respectively. The size of the nodes (OTUs) is proportional to the degree of connectivity. Larger nodes are highly connected with other

nodes and may play a more important role in the whole network. Overall, microbial co-occurrence is present within individual groups, and microbial mutual exclusion is present between control- and COPD-enriched groups.

Most of the key microbiota enriched in controls were highly abundant bacteria, such as *N. subflava* (OTU-1), *Prevotella* sp. (OTU-2), and *Veillonella dispar* (OTU-3),



**Figure 3** Interaction of OTUs enriched in controls and COPD patients.

**Notes:** (A) The color of the nodes represents different phyla, and the size of the nodes is proportional to their degree of connectivity. The bigger node is highly connected with others and may play a more important role in the whole network. The blue and red edges represent co-occurrence (Spearman's coefficient  $>0.5$  and  $*P<0.05$ ) and co-exclusion (Spearman's coefficient  $<-0.5$  and  $*P<0.05$ ), respectively. (B) Relative abundance of the key OTUs enriched in COPD patients between COPD patients with mild symptoms (SGRQ  $<25$ ) and COPD patients with severe symptoms (SGRQ  $\geq 25$ ). Statistical significance ( $*P<0.05$ ) was determined through ANOVA with Tukey's post hoc test.

**Abbreviations:** ANOVA, analysis of variance; OTU, operational taxonomic unit; SGRQ, St George's Respiratory Questionnaire.

which accounted for 42.2% of the total abundance. Additionally, *Granulicatella* sp. (OTU-16) was not a core bacterium but a commensal with OTU-1.

In contrast to the control-enriched OTUs, the pivotal bacteria enriched in COPD patients were low-abundance bacteria, including *Megasphaera micronuciformis* (OTU-14), *Prevotella salivae* (OTU-23), and the family Lachnospiraceae (OTU-30). Noticeably, the three OTUs had a higher abundance in patients with severe symptoms than in patients with mild symptoms, although a significant difference was identified only for *P. salivae* (OTU-23), as shown in Figure 3B. These results suggested that the core bacteria enriched in COPD patients may influence symptoms.

## Relationship between the sputum flora and the clinical phenotypes

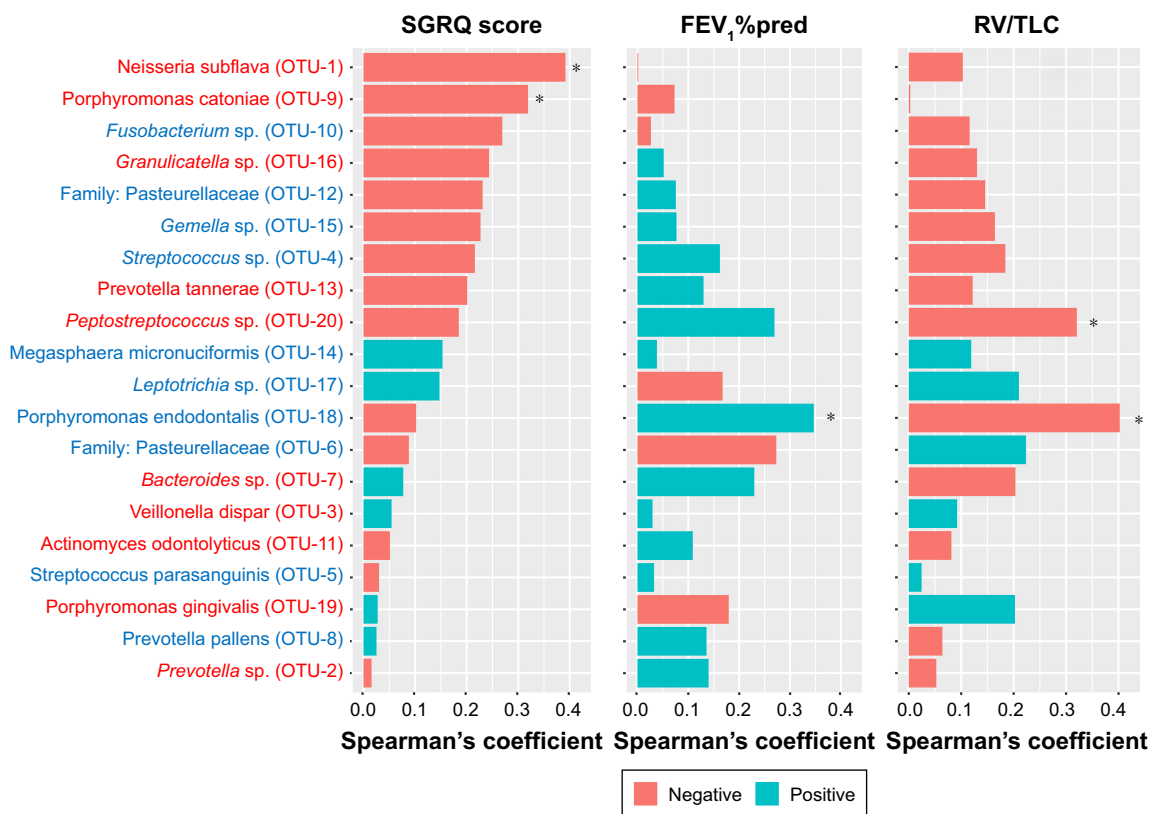
To investigate the association between specific bacteria with lung function and symptoms in COPD patients, we calculated Spearman's coefficient for the 20 most abundant OTUs, as illustrated in Figure 4. A higher SGRQ score, a

lower FEV<sub>1</sub>%pred, and a higher RV/TLV denoted worse symptoms, airflow limitation, and emphysema, respectively. The association pattern of symptoms with the microbiota was obviously different from the association patterns of FEV<sub>1</sub>%pred and RV/TLC, and the association patterns of the latter two parameters were very similar.

## Discussion

To the best of our knowledge, this investigation is the first study to explore the relationship between symptoms and the sputum microbiome in clinically stable COPD patients. The key findings are the identification of alterations in community composition and core bacterial taxa such as *N. subflava* in stable COPD patients with more severe symptoms.

FEV<sub>1</sub>%pred, RV/TLC, history of frequent exacerbation, and comorbidities (diabetes and hypertension) were very similar between symptomatic ( $\geq 25$ ) and asymptomatic ( $< 25$ ) patients in our study. The inconsistency of symptoms may be related to disease heterogeneity. Previous studies on large cohorts have revealed that symptom reports are highly



**Figure 4** Relationship between the sputum flora and the clinical phenotypes in COPD patients.

**Notes:** The associations between OTUs and clinical indexes (SGRQ score, FEV<sub>1</sub>%pred, RV/TLC, etc.) were determined by Spearman's coefficient. Red and blue bars denote negative and positive associations, respectively. Asterisks indicate that the association is significant ( $P < 0.05$ ). Y-axis text labeled in red and blue denotes control-enriched OTUs and COPD-enriched OTUs, respectively.

**Abbreviations:** OTU, operational taxonomic unit; SGRQ, St George's Respiratory Questionnaire; FEV<sub>1</sub>%pred, predicted percent of forced expiratory volume in 1 second; RV/TLV, ratio of residual volume to total lung capacity.

variable, even in patients with the same GOLD stage.<sup>29,30</sup> Another possible explanation is that only 22.2% of symptomatic patients in the present study previously used ICS and the proportion seems to be lower than that of asymptomatic patients (38.4%), although statistical difference is not significant. The low usage rate of ICS caused an increase in current symptoms in these patients.

A common view is that the airway microbiota composition is altered in patients with very severe lung function and with AE, although considerable heterogeneity of the airway microbiota exists between patients.<sup>10,20</sup> A huge microbial variation between patients was also observed in the present study. Our key finding is the identification of symptom-related sputum microbial alterations, which are not associated with lung function. Compared with previous cohorts, the patients in this study had relatively mild pulmonary function deficits and included only two patients staged GOLD IV. Moreover, the use of ICS did not differ significantly between patients with mild and severe symptoms. It was reported that patients with more severe symptoms may have higher inflammation levels.<sup>31</sup> Therefore, the shift in the sputum microbiome in symptomatic patients may represent an adaptation to the inflammatory airway environment.

Previous studies have reported that *Neisseria* is a highly abundant genus in the sputum and is not significantly altered in COPD patients.<sup>12,18</sup> In this study, *N. subflava* (OTU-1) was also not shifted in COPD patients, but it was decreased in patients with severe symptoms and negatively correlated with SGRQ scores, which implied that *N. subflava* may be associated with COPD symptoms. In addition, *N. subflava* played an important role in the co-occurrence network, and the reduction of this species may drive the observed shift at the community level. The genus *Granulicatella* was reported to decrease in COPD patients,<sup>32</sup> which was verified in our study. We also found that *Granulicatella* sp. was commensal with *N. subflava* (OTU-1) and was reduced more significantly in patients with severe symptoms. Accordingly, several low-abundance COPD-enriched core microbiota, such as *M. micronuciformis* (OTU-14), *P. salivae* (OTU-23), and the family Lachnospiraceae (OTU-30), showed an increasing trend in patients with severe symptoms, and these taxa were mutually exclusive from the key control-enriched taxa. Thus, the three OTUs may represent potentially pathogenic microorganisms (PPM) in COPD. In addition, these OTUs also present higher expression in patients with severe symptoms, suggesting that PPMs may be related to symptoms. These results imply that common sputum bacteria may be replaced by other low-abundance taxa, causing increased COPD symptoms.

In this study, we also found that the association pattern of the sputum microbiota with symptoms was obviously different from that of lung function in COPD patients. Therefore, symptomology should also be addressed independent of lung function in future microbiome studies.

## Limitations

This study has several limitations. First, nonsmokers and females were not included because COPD patients living in cities in China are almost male smokers. Second, this investigation was a small cross-sectional study, which restricts the dynamic observations for these results. Third, we used only the more comprehensive SGRQ questionnaire to evaluate symptomatology instead of the CAT, which was not collected during enrollment. Fourth, information regarding the environmental and occupational exposure of the subjects was not collected in this study. Finally, the causality between key bacteria and symptoms has not been determined using animal models.

## Conclusion

This preliminary study highlighted the potential link between the microbiota and symptomology in COPD patients, emphasizing the role of symptoms in the airway microbiome independent of lung function. Because both the airway microbial environment and the patient symptoms undergo dynamic changes and are treatable, understanding the interplay between these factors could facilitate individualized treatment.

## Acknowledgments

The authors thank Jie Li (BGI, Shenzhen, China) for advice on bioinformatics and statistical analyses in this study. The present study was funded by the National Natural Science Foundation of China (grant nos: 81270097, 81470235, and 81670034) and Beijing Medical University (grant nos: 20110176 and 20160529).

## Author contributions

All authors declared their personal contributions to this manuscript: BH conducted the study design and critically revised the article; WD conducted bioinformatics analyses and interpretation and wrote the manuscript; NS and YD were involved in statistical analyses and enrolled subjects; XS and CG collected throat samples and performed lung function tests for the subjects; QK gathered clinical information from the subjects; and JRE-D, GBH, and MRG were involved in the study design. All authors contributed toward data analysis, drafting and revising the paper and agree to be accountable for all aspects of the work.



## Disclosure

The authors report no conflicts of interest in this work.

## References

- Vogelmeier CF, Criner GJ, Martinez FJ, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD Executive Summary. *Am J Respir Crit Care Med*. 2017;195(5):557–582.
- Fletcher C, Peto R. The natural history of chronic airflow obstruction. *Br Med J*. 1977;1(6077):1645–1648.
- Franklin W, Lowell FC, Michelson AL, Schiller IW. Chronic obstructive pulmonary emphysema; a disease of smokers. *Ann Intern Med*. 1956;45(2):268–274.
- Sethi S. Infection as a comorbidity of COPD. *Eur Respir J*. 2010;35(6):1209–1215.
- Papi A, Bellettato CM, Braccioni F, et al. Infections and airway inflammation in chronic obstructive pulmonary disease severe exacerbations. *Am J Respir Crit Care Med*. 2006;173(10):1114–1121.
- Rosell A, Monsó E, Soler N, et al. Microbiologic determinants of exacerbation in chronic obstructive pulmonary disease. *Arch Intern Med*. 2005;165(8):891–897.
- Lieberman D, Lieberman D, Ben-Yaakov M, et al. Infectious etiologies in acute exacerbation of COPD. *Diagn Microbiol Infect Dis*. 2001;40(3):95–102.
- Matkovic Z, Miravittles M. Chronic bronchial infection in COPD. Is there an infective phenotype? *Respir Med*. 2013;107(1):10–22.
- Erb-Downward JR, Thompson DL, Han MK, et al. Analysis of the lung microbiome in the “healthy” smoker and in COPD. *PLoS One*. 2011;6(2):e16384.
- Sze MA, Dimitriu PA, Hayashi S, et al. The lung tissue microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2012;185(10):1073–1080.
- Sze MA, Dimitriu PA, Suzuki M, et al. Host response to the lung microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2015;192(4):438–445.
- Cameron SJ, Lewis KE, Huws SA, et al. Metagenomic sequencing of the chronic obstructive pulmonary disease upper bronchial tract microbiome reveals functional changes associated with disease severity. *PLoS One*. 2016;11(2):e0149095.
- Pragman AA, Kim HB, Reilly CS, Wendt C, Isaacson RE. The lung microbiome in moderate and severe chronic obstructive pulmonary disease. *PLoS One*. 2012;7(10):e47305.
- Galiana A, Aguirre E, Rodriguez JC, et al. Sputum microbiota in moderate versus severe patients with COPD. *Eur Respir J*. 2014;43(6):1787–1790.
- García-Nuñez M, Millares L, Pomares X, et al. Severity-related changes of bronchial microbiome in chronic obstructive pulmonary disease. *J Clin Microbiol*. 2014;52(12):4217–4223.
- Huang YJ, Sethi S, Murphy T, Nariya S, Boushey HA, Lynch SV. Airway microbiome dynamics in exacerbations of chronic obstructive pulmonary disease. *J Clin Microbiol*. 2014;52(8):2813–2823.
- Millares L, Ferrari R, Gallego M, et al. Bronchial microbiome of severe COPD patients colonised by *Pseudomonas aeruginosa*. *Eur J Clin Microbiol Infect Dis*. 2014;33(7):1101–1111.
- Su J, Liu HY, Tan XL, et al. Sputum bacterial and fungal dynamics during exacerbations of severe COPD. *PLoS One*. 2015;10(7):e0130736.
- Wang Z, Bafadhel M, Haldar K, et al. Lung microbiome dynamics in COPD exacerbations. *Eur Respir J*. 2016;47(4):1082–1092.
- Huang YJ, Erb-Downward JR, Dickson RP, Curtis JL, Huffnagle GB, Han MK. Understanding the role of the microbiome in chronic obstructive pulmonary disease: principles, challenges, and future directions. *Transl Res*. 2017;179:71–83.
- Einarsson GG, Comer DM, McIlreavey L, et al. Community dynamics and the lower airway microbiota in stable chronic obstructive pulmonary disease, smokers and healthy non-smokers. *Thorax*. 2016;71(9):795–803.
- Huffnagle GB. Another piece in the ‘research mosaic’ that describes the role of the lung microbiome in COPD. *Thorax*. 2016;71(9):777–778.
- Diao W, Shen N, Du Y, et al. Identification of thyroxine-binding globulin as a candidate plasma marker of chronic obstructive pulmonary disease. *Int J Chron Obstruct Pulmon Dis*. 2017;12:1549–1564.
- Diao WQ, Shen N, Du YP, Yp D, et al. Fetuin-B (FETUB): a plasma biomarker candidate related to the severity of lung function in COPD. *Sci Rep*. 2016;6:30045.
- Vestbo J, Hurd SS, Agustí AG, et al. Global strategy for the diagnosis, management, and prevention of chronic obstructive pulmonary disease: GOLD executive summary. *Am J Respir Crit Care Med*. 2013;187(4):347–365.
- Jones PW, Tabberer M, Chen WH. Creating scenarios of the impact of COPD and their relationship to COPD Assessment Test (CAT™) scores. *BMC Pulm Med*. 2011;11:42.
- Mason KL, Erb Downward JR, Mason KD, et al. *Candida albicans* and bacterial microbiota interactions in the cecum during recolonization following broad-spectrum antibiotic therapy. *Infect Immun*. 2012;80(10):3371–3380.
- Schloss PD, Westcott SL, Ryabin T, et al. Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities. *Appl Environ Microbiol*. 2009;75(23):7537–7541.
- Agusti A, Calverley PM, Celli B, et al. Characterisation of COPD heterogeneity in the ECLIPSE cohort. *Respir Res*. 2010;11:122.
- Weatherall M, Marsh S, Shirtcliffe P, Williams M, Travers J, Beasley R. Quality of life measured by the St George’s Respiratory Questionnaire and spirometry. *Eur Respir J*. 2009;33(5):1025–1030.
- Gupta N, Pinto L, Benedetti A, et al. The COPD assessment test: can it discriminate across COPD subpopulations? *Chest*. 2016;150(5):1069–1079.
- Ogorodova LM, Fedosenko SV, Popenko AS, et al. [Comparison study of oropharyngeal microbiota in case of bronchial asthma and chronic obstructive pulmonary disease in different severity levels]. *Vestnik Rossiiskoi Akademii Meditsinskikh Nauk*. 2015;70(6):669–678. Russian.

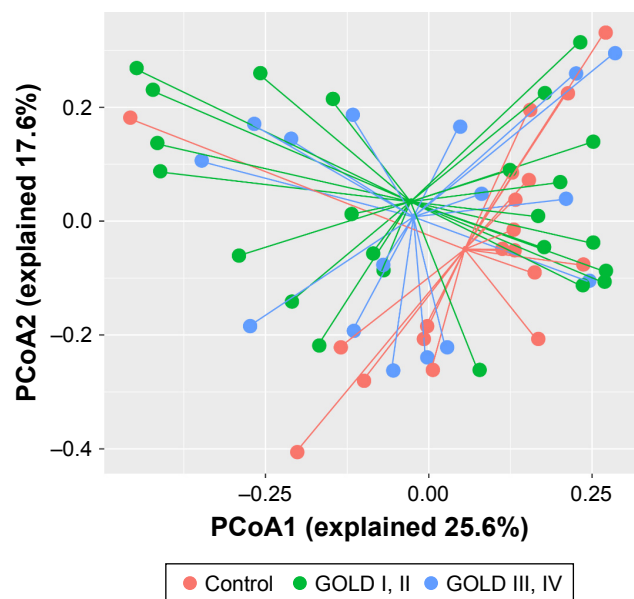
## Supplementary materials

**Table S1** Relationship between clinical index and sputum microbiome by PERMANOVA

Variables	Control vs mild	Control vs severe	Mild vs severe
Symptom severity	0.258	0.030	0.173
Inhaled corticosteroid	0.07	0.237	0.985
Spirometry severity	0.323	0.549	0.643
Hypertension	0.523	0.925	0.536
Exacerbation (>2) <sup>#</sup>	NA	0.141	0.200
Diabetes	0.739	0.927	0.689
Bronchodilator	0.852	NA	0.868

**Notes:** PERMANOVA, correcting for spirometry severity, use of inhaled corticosteroid and bronchodilator, and history of hypertension and diabetes, suggested sputum microbiome is different between control and COPD with severe symptom ( $P=0.033$ ). <sup>#</sup>Number of 2 or greater than exacerbation events in the past 1 year.

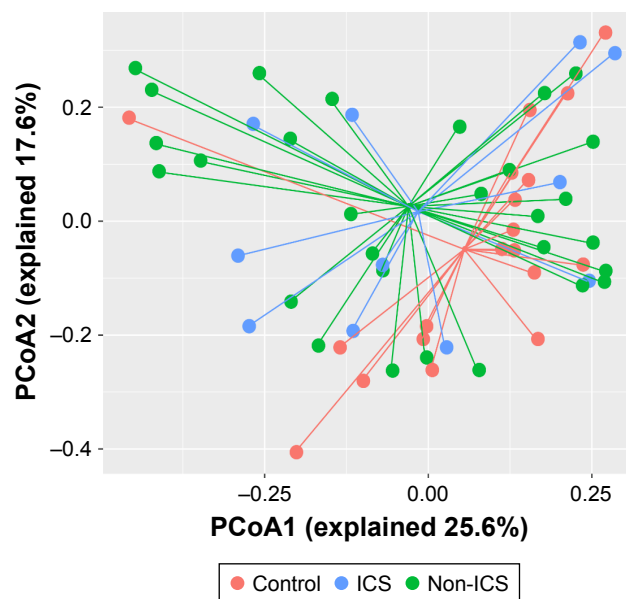
**Abbreviations:** NA, not applicable; PERMANOVA, permutational multivariate analysis of variance.



**Figure S1** Community composition was not altered between controls and patients with different severity of lung function.

**Notes:** Unsupervised PCoA and PERMANOVA were used to determine the discrepancy of community level between controls, patients with different severity of lung function using Bray–Curtis distance. No statistical difference was identified between groups (control vs GOLD I, II,  $P=0.162$ ; control vs GOLD III, IV,  $P=0.105$ ; GOLD I, II vs GOLD III, IV,  $P=0.723$ ).

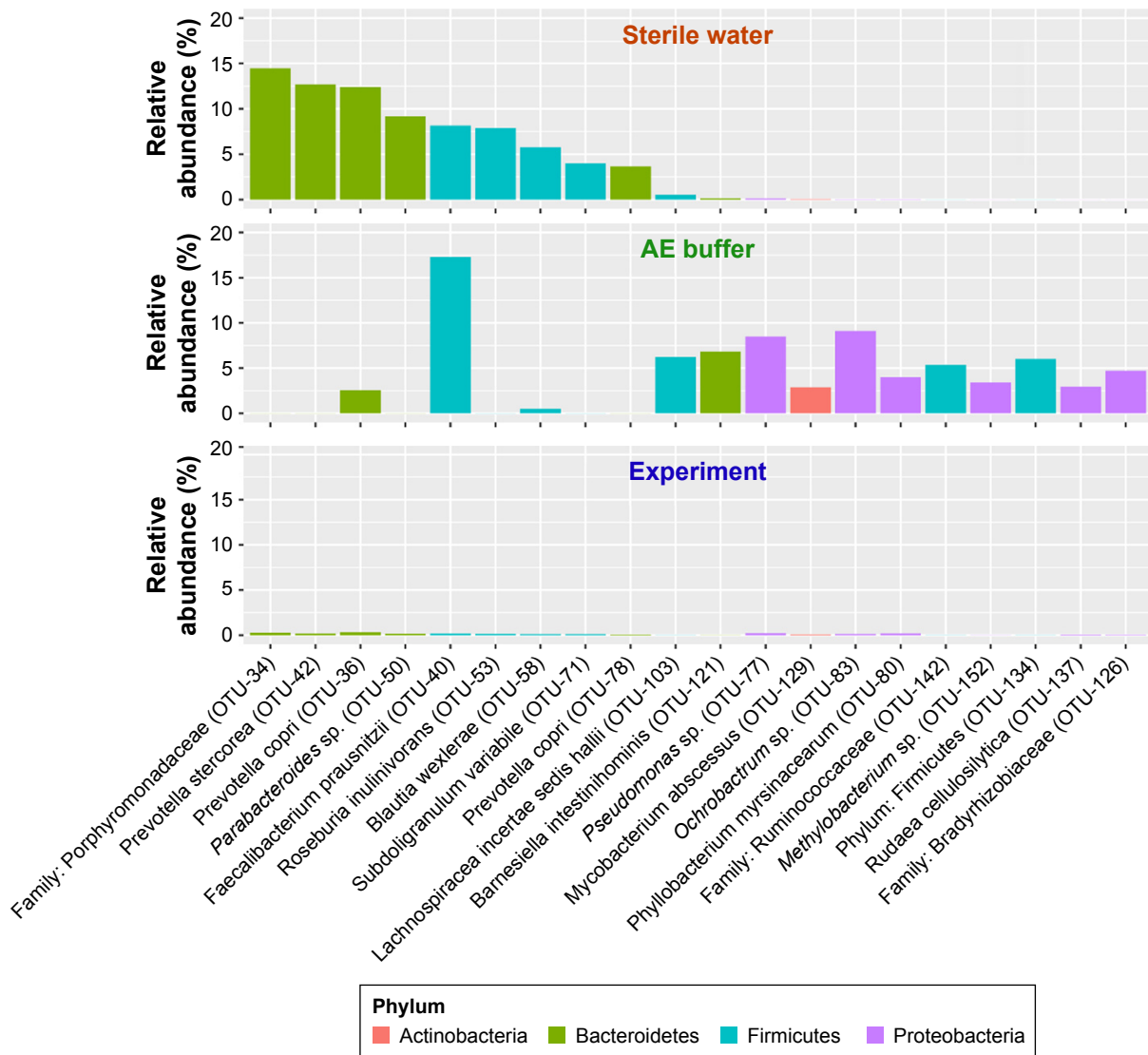
**Abbreviations:** GOLD, Global Initiative for Obstructive Lung Disease; PCoA, principal coordinate analysis; PERMANOVA, permutational multivariate analysis of variance.



**Figure S2** Community composition was not affected by the use of inhaled corticosteroid.

**Notes:** Unsupervised PCoA and PERMANOVA were used to determine the discrepancy of community level between controls and patients with or without the use of ICS using Bray–Curtis distance. No statistical difference was identified between groups (control vs ICS,  $P=0.221$ ; control vs non-ICS,  $P=0.145$ ; ICS vs non-ICS,  $P=0.957$ ).

**Abbreviations:** ICS, inhaled corticosteroid; PCoA, principal coordinate analysis; PERMANOVA, permutational multivariate analysis of variance.



**Figure S3** Potential contaminated OTUs in negative control.

**Note:** The potentially contaminated OTUs in sterile water and AE buffer were present, which showed extremely low abundance in experimental samples and did not include the previously identified core microbiota (*Neisseria subflava*, etc.).

**Abbreviations:** AE, acute exacerbation; OTU, operational taxonomic unit.