

Original Research

Respiratory Microbiome Profiles Associated With Distinct Inflammatory Phenotype and Clinical Indexes in Chronic Obstructive Pulmonary Disease

Tao Yu,¹ Yunru Chen,² Xiaoxia Ren,¹ Ting Yang¹

¹ National Center for Respiratory Medicine, State Key Laboratory of Respiratory Health and Multimorbidity, National Clinical Research Center for Respiratory Diseases, Institute of Respiratory Medicine, Chinese Academy of Medical Sciences and Peking Union Medical College, Department of Pulmonary and Critical Care Medicine, Center of Respiratory Medicine, China-Japan Friendship Hospital, Beijing, China

² Centre for Evidence-based Chinese Medicine, School of Traditional Chinese Medicine, Beijing University of Chinese Medicine, Beijing, China

Address correspondence to:

Xiaoxia Ren

Department of Pulmonary and Critical Care Medicine, Center of Respiratory Medicine
China-Japan Friendship Hospital

No 2, East Yinghua Road

Chaoyang District

Beijing 100029, China

E-mail: xiaoxiaren2011@163.com

Running Head: COPD Respiratory Microbiome Profiles

Keywords: chronic obstructive pulmonary disease; microbiome; cluster analysis\

Abbreviations:

Funding Support: This study was supported by the National Natural Science Foundation of China (grant no. 81800036) and the National High-Level Hospital Clinical Research Funding, Elite Medical Professionals Project of China-Japan Friendship Hospital (grant no. ZRJY2021-TD02).

Date of acceptance: December 12, 2023 | **Published Online Date:** December 19, 2023

Citation: Yu T, Chen Y, Ren X, Yang T. Respiratory microbiome profiles associated with distinct inflammatory phenotype and clinical indexes in chronic obstructive pulmonary disease. *Chronic Obstr Pulm Dis*. 2023; Published online December 19, 2023.

doi: <https://doi.org/10.15326/jcopdf.2023.0445>

This article has an online data supplement.

Abstract

Respiratory microbiome studies have fostered our understanding of the various phenotypes and endotypes of heterogeneous chronic obstructive pulmonary disease (COPD). This study aimed to identify microbiome-driven clusters that reflect the clinical features and dominant microbiota of COPD. This cross-sectional study included 32 patients with stable COPD between December 2019 and December 2020 from the outpatient clinic of China-Japan Friendship Hospital. Sputum samples were tested for 16S rRNA. Patients were classified according to the species level using an unsupervised clustering method to compare the inflammatory phenotypes of two clusters and analyze the correlation between the main bacteria and clinical indicators in each cluster. Patients were further divided into two clusters according to microorganisms. Neutrophils in cluster 1 were significantly increased compared with cluster 2. Cluster 1 was predominantly *Bacteroides*, while cluster 2 was dominated by *Prevotella* and *Fusobacterium* at the genus level. *Fusobacterium* was negatively correlated with the COPD Assessment Test (CAT) score, and *Bacteroides* was positively correlated with the number of acute exacerbations of COPD. This study found that differential flora was negatively associated with CAT scores and the number of acute exacerbations of COPD. This microbiome-driven unbiased clustering method for COPD can help identify new endotype-related COPD phenotypes.

Introduction

Chronic obstructive pulmonary disease (COPD) is a common disease characterized by persistent airway symptoms and restricted airflow, which cause airway and/or alveolar abnormalities.¹ The pathogenesis of COPD is complex; inflammation is a major causative mechanism. In recent years, bacterial colonization in the lungs of patients with COPD has been an important factor contributing to chronic airway inflammation; however, the exact underlying mechanisms are unclear.^{2, 3}

Many studies have found significant bacterial colonization in the lungs of healthy individuals and patients with COPD using 16S rRNA assays.⁴⁻⁷ The main normal colonizing bacteria in the lower respiratory tract of the healthy population are genera from the phyla Bacteroidetes (Bacteroides and Prevotella), Firmicutes (Veillonella and Streptococcus), Proteobacteria, and Fusobacteria (Fusobacterium).⁸ In patients with stable COPD, the types and proportions of airway microorganisms change, with the main colonizing phyla being Aspergillus, Firmicutes, and Actinomycetes, with Veillonella, Haemophilus, Streptococcus, Prevotella, Neisseria, and Moraxella being the dominant genera.^{4, 9} The proportion of Aspergillus increases further when acute exacerbations of slow-onset lung disease occur.¹⁰ Moreover, the alpha and beta diversity of airway microorganisms is significantly lower in patients with COPD and those with acute exacerbations of COPD than in the normal population.¹¹⁻¹³ Additionally, changes in pulmonary microecology correlate with the severity of disease. A study by Sze et al.⁶ found a significant decrease in microbial diversity in patients with COPD with The Global Initiative for Chronic Obstructive Lung Disease (GOLD) stage IV compared with the normal population, mainly associated with Firmicutes and Aspergillus. Dicker et al.¹⁴ also showed that Aspergillus was associated with disease stage, risk of acute exacerbation, blood eosinophil levels, and predicted percentage of forced expiratory volume in 1 s (FEV₁) in

patients with COPD. However, few studies have explored the correlation among airway microorganisms, symptoms, and biological indicators in patients with COPD.

Therefore, this study aimed to detect microbial changes in the sputum of patients with stable COPD using the 16S rRNA method and to explore their relationship with clinical symptoms and biological indicators.

Pre-proof

Materials and Methods

Study population

Thirty-two patients with stable COPD were recruited between December 2019 and December 2020 from the outpatient clinic of China-Japan Friendship Hospital, Beijing, China. The inclusion criteria for patients with COPD were as follows: (i) met the GOLD 2021 diagnostic criteria, (ii) aged > 40 years, and (iii) no acute exacerbation within 1 month. Exclusion criteria were: (i) history of chronic diseases such as severe cardiovascular and cerebrovascular pathology and liver and kidney insufficiency; (ii) history of epilepsy and other neuropsychiatric diseases; (iii) history of anti-tuberculosis drug treatment or active tuberculosis; (iv) history of thoracic, abdominal, or ophthalmic surgery in the last 3 months; (v) pregnant and lactating women; (vi) untreated oncology patients with an expected survival of < 1 year; and (vii) unwillingness to participate and inability to cooperate. All participants completed questionnaires, pulmonary function tests, and computed tomography examinations. All patients signed an informed consent form, and the study protocol was approved by the Ethics Committee of China-Japan Friendship Hospital (2018-148-K105). All experiments were performed in accordance with the Declaration of Helsinki.

16S rRNA gene sequencing

All patients received sputum induction after inhalation of 3% hypertonic saline from an ultrasonic nebulizer, and sputum was stored in a -80 °C refrigerator. Total genomic DNA was extracted from the sputum samples using the cetyltrimethylammonium bromide/sodium dodecyl sulfate method. DNA concentration and purity were monitored on a 1% agarose gel. According to the concentration, DNA was diluted to 1 ng/μL using sterile water. Polymerase chain reaction amplification was performed using fusion primers targeting the V3–V4 region

of the 16S rRNA gene. The fusion primers used were 806R (5'-GGACTACNNGGGTATCTAAT-3') and 341F (5'-CCTACGGGRBGCASCAG-3') with the barcode. Sequencing libraries were generated using TruSeq® DNA PCR-Free Sample Preparation Kit (Illumina, USA) following the manufacturer's recommendations, and index codes were added. Library quality was assessed using a Qubit® 2.0 Fluorometer (Thermo Scientific, USA) and an Agilent Bioanalyzer 2100 system. Finally, the library was sequenced on an Illumina NovaSeq (Illumina, USA) platform, and 250 bp paired-end reads were generated.

Microbiome analyses

Paired-end reads were assigned to samples based on their unique barcodes and truncated by cutting off the barcodes and primer sequences. Fast Length Adjustment of Short Reads was used for sequence assembly. Quality filtering of the raw tags was performed under specific filtering conditions to obtain high-quality clean tags according to QIIME. Finally, the tags were compared with those in the reference database using the UCHIME algorithm. Sequence analysis was performed using Uparse software. Sequences with $\geq 97\%$ similarity were assigned to the same operational taxonomic units.

Microbiome data analysis for the clustered groups

We performed clustering analysis on the samples using species-level abundance data, employing the Partitioning Around Medoid (PAM) clustering algorithm based on the Jensen–Shannon divergence.¹⁵ To determine the optimal number of clusters, we calculated the Calinski–Harabasz (CH) index based on cluster count.¹⁶ The resulting clusters were visualized using the "ade4" package in R for Principal Coordinate Analysis (PCoA) utilizing Jensen–Shannon divergence.¹⁷

Linear Discriminant Effect Size (LEfSe) analysis was conducted to identify the microbiota as biomarkers associated with each cluster. Significant biomarkers were identified using a linear discriminant analysis (LDA) score > 4.0 and a $P < 0.05$ in pairwise comparisons using the Mann–Whitney test and Bonferroni's methods.¹⁸ The resulting biomarkers were visualized using GraPhlAn for cladogram representation and the R statistical package (R version 3.2.5; Institute for Statistics and Mathematics, Vienna, Austria; www.R-project.org) with the "ggplot" package for boxplot visualization, based on the Kruskal–Wallis H test¹⁹.

Statistical analysis

Patients were clustered using PAM clustering and CH index, and the demographic characteristics of different groups of patients were analyzed using the t-test for continuous variables and chi-square test for categorical variables. Continuous data are presented as median \pm standard deviation, and categorical data are presented as numbers and percentages. Spearman's correlation was used to assess the correlation between each group of marker microorganisms and clinical indicators. Statistical significance was set at $P < 0.05$, and the data were analyzed using SPSS (version 26.0; IBM, Armonk, NY, USA).

Results

Participant demographics and sputum microbiota composition

A total of 32 patients with COPD were enrolled (Table 1). The mean FEV₁%pred was 42.59%, and the mean COPD Assessment Test (CAT) and Modified British Medical Research Council (mMRC) scores was 24.42 and 2.13, respectively. The predominant phyla observed in the microbiome were Firmicutes, Proteobacteria, and Bacteroidetes, followed by Actinobacteria and Fusobacteria. At the genus level, the most abundant genera were Haemophilus, Streptococcus, Neisseria, and Rothia, although some individual samples also showed a predominance of genera such as Porphyromonas, Acinetobacter, Abiotrophia, Lactobacillus, and unknown members of Prevotellaceae and Corynebacteriaceae (Supplementary Figure 1A-1B).

Participant clustering and characteristics

According to the PAM clustering and CH index, for sputum species-level bacteria, clustering patients into two, three, or four clusters resulted in a higher CH index, whereas the PCoA plots showed better separation (Figure 1A-1B, Supplementary Figure 2A-2B). We compared the clinical indexes of the patients using different clustering methods and found significant differences in blood neutrophil counts, number of exacerbations, and CAT scores when divided into two clusters (Figure 1C, Supplementary Figure 2C-2D). When divided into three or four clusters, there was only one sample in a cluster and the sample size was small; thus, could not be representative to a certain extent. Therefore, we divided the patients into two clusters; the clinical demonstration of each cluster is shown in Table 1.

Dominant microbiotas in the clusters

We used LEfSe analysis to analyze the different microbial distributions in the two patient groups (LDA score: 4.0; Figure 2). At the phylum level, Actinobacteria were the more dominant microorganisms in cluster 1, and Fusobacteria were the predominant microorganisms in cluster 2. At the genus level, cluster 1 was dominated by Bacteroides, while cluster 2 was dominated by Prevotella, Fusobacterium, and unidentified Prevotellaceae. As it is difficult to fully recognize the species level of many microorganisms through 16S rRNA, we compared two clusters at the genus level²⁰ at $P < 0.05$ using the Kruskal–Wallis H test to identify the dominant microbiota related to each cluster (Figure 3). Cluster 2 had more elevated counts of Prevotella, Fusobacterium, Solobacterium, and Gemella than cluster 1, whereas cluster 1 had more elevated counts of Streptomyces, Peptoanaerobacter, and Ottowia than cluster 2. Combining these two analyses, cluster 2 had more Prevotella and Fusobacterium than cluster 1.

Relationship between the sputum flora and clinical phenotypes

To explore the correlation between differentially predominated changes in microbiota tested in the sputum, including Prevotella, Fusobacterium, and Bacteroides, and the clinical index parameters, we performed Spearman correlation analysis (Supplementary Table 1). As shown in Figure 4, there was a positive correlation between Bacteroides and the number of acute exacerbations ($r = 0.417$, $P = 0.018$), a negative correlation between Fusobacterium and CAT scores ($r = -0.357$, $P = 0.045$), and no significant correlation between other bacteria and the number of acute exacerbations or CAT scores.

Discussion

Using the unbiased microbiome profile clustering method for clustering sputum microorganisms in patients with COPD, we divided patients into two clusters. Patients with a lower abundance of *Prevotella* and *Fusobacterium* and a higher abundance of *Bacteroides* had higher CAT scores and a neutrophil inflammatory phenotype. Correlation analysis showed a positive correlation between *Bacteroides* and the number of acute exacerbations and a negative correlation between *Fusobacterium* and CAT scores.

The unsupervised clustering of microorganisms used in this study is a relatively new method. Most studies exploring the correlation between microorganisms and clinical phenotypes are grouped based on clinical phenotypes, with few grouping based on microorganisms. We chose to group patients with COPD according to their microbial characteristics, which is different from the method used in previous studies.²¹⁻²³ Meanwhile, one study grouped patients according to the major genus *Haemophilus* in previous literature but did not use the unsupervised clustering method used in this study,²⁴ which has some bias. In addition, this method has also been used in a study on adult and pediatric asthma, which showed significant correlation between classification and clinical characteristics.^{25, 26} Alongside this cluster method, we also identified the neutrophil inflammatory phenotype, an important phenotype in COPD. Therefore, this cluster analysis method may be more clinically relevant.

In the present study, cluster 2 had a lower acute exacerbation of COPD, neutrophil inflammation, and CAT score compared with cluster 1. Additionally, at the genus level, there was a higher *Prevotella* count; a negative correlation was found between *Prevotella* and disease severity. This is comparable to Mayhew et al.,²⁷ who found that *Haemophilus* and disease severity were positively correlated, while *Prevotella* was negatively correlated. Similarly, another study found that *Prevotella* abundance was significantly higher in healthy

individuals and significantly associated with better lung function and reduced symptoms in patients with COPD.²⁸ Moreover, previous studies have also found that *Fusobacterium* exhibited a significant positive correlation with pre-FEV₁,²⁹ and *Fusobacterium* and *Prevotella jejuni* were significantly more abundant in the infrequent than frequent exacerbator group.²² This confirms that *Fusobacterium* and *Prevotella* are negatively correlated with disease severity.

This study had some limitations. First, the sample size was not sufficiently large, and a larger sample size is needed to confirm the results. Second, this was a cross-sectional study, which can only show a correlation between microorganisms and clinical indicators but does not confirm the existence of a causal relationship. Finally, this was a clinical study lacking relevant basic experiments for validation. Therefore, additional data are required to support the results of this study.

Conclusion

This microbiome-driven unbiased clustering method for COPD can help identify new endotype-related COPD phenotypes. This study divided patients into two clusters for microbial clustering and found that differential flora was negatively associated with CAT scores and the number of acute exacerbations of COPD. Thus, understanding the mechanisms that drive bacterial composition and disease will help to prevent and discover new phenotypes of COPD and treat this condition.

Acknowledgments

The authors thank Dr. Lichao Sun (Department of Emergency, China-Japan Friendship Hospital, Beijing, China) and Professor Hongbo Zhang (Department of Emergency, China-Japan Friendship Hospital, Beijing, China) for their help with data acquisition in this study.

Author contributions: X.X.R. and T.Y.: Concept and design. R.R.C.: Data collection. T.Y.: Data analyses. T.Y.: Manuscript writing. All authors: Manuscript proofing and Data interpretation. X.X.R.: Guarantor.

Disclosure statement: The authors declare that there are no competing interests to declare.

Data availability statement: The datasets used and/or analyzed in the current study are available from the corresponding author upon reasonable request.

References

1. Sheikh K, Coxson HO, Parraga G. This is what COPD looks like. *Respirology*. 2016;21:224–136. doi: 10.1111/resp.12611
2. Dima E, Kyriakoudi A, Kaponi M, et al. The lung microbiome dynamics between stability and exacerbation in chronic obstructive pulmonary disease (COPD): current perspectives. *Respir Med*. 2019;157:1–6. doi: 10.1016/j.rmed.2019.08.012.
3. Wang Z, Locantore N, Haldar K, et al. Inflammatory endotype-associated airway microbiome in chronic obstructive pulmonary disease clinical stability and exacerbations: A multicohort longitudinal analysis. *Am J Respir Crit Care Med*. 2021;203:1488–1502. doi: 10.1164/rccm.202009-3448OC.
4. 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:e16384. doi: 10.1371/journal.pone.0016384.
5. Sze MA, Dimitriu PA, Hayashi S, et al. The lung tissue microbiome in chronic obstructive pulmonary disease. *Am J Respir Crit Care Med*. 2012;185:1073–1080. doi: 10.1164/rccm.201111-2075OC.
6. 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;192:438–445. doi: 10.1164/rccm.201502-0223OC.
7. 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:e0149095. doi: 10.1371/journal.pone.0149095.
8. Yatera K, Noguchi S, Mukae H. The microbiome in the lower respiratory tract. *Respir Investig*. 2018;56:432–439. doi: 10.1016/j.resinv.2018.08.003.

9. Garcia-Nuñez M, Millares L, Pomares X, et al. Severity-related changes of bronchial microbiome in chronic obstructive pulmonary disease. *J Clin Microbiol.* 2014;52:4217–4223. doi: 10.1128/jcm.01967-14.
10. López Caro JC, Santibáñez M, García Rivero JL, et al. Sputum microbiome dynamics in chronic obstructive pulmonary disease patients during an exacerbation event and post-stabilization. *Respiration.* 2019;98:447–454. doi: 10.1159/000501988.
11. Wang Z, Bafadhel M, Haldar K, et al. Lung microbiome dynamics in COPD exacerbations. *Eur Respir J.* 2016;47:1082–1092. doi: 10.1183/13993003.01406-2015.
12. Wang Z, Singh R, Miller BE, et al. Sputum microbiome temporal variability and dysbiosis in chronic obstructive pulmonary disease exacerbations: an analysis of the COPDMAP study. *Thorax.* 2018;73:331–338. doi: 10.1136/thoraxjnl-2017-210741.
13. Wang Z, Maschera B, Lea S, et al. Airway host-microbiome interactions in chronic obstructive pulmonary disease. *Respir Res.* 2019;20:113. doi: 10.1186/s12931-019-1085-z.
14. Dicker AJ, Huang JTJ, Lonergan M, et al. The sputum microbiome, airway inflammation, and mortality in chronic obstructive pulmonary disease. *J Allergy Clin Immunol.* 2021;147:158–167. doi: 10.1016/j.jaci.2020.02.040.
15. Arumugam M, Raes J, Pelletier E, et al. Enterotypes of the human gut microbiome. *Nature.* 2011;473:174–180. doi: 10.1038/nature10187.
16. Caliński T, Harabasz J. A dendrite method for cluster analysis. *Commun Stat Theor Methods.* 1974;1:1–27. doi: 10.1080/03610927408827101.
17. Chen J, Bittinger K, Charlson ES, et al. Associating microbiome composition with environmental covariates using generalized UniFrac distances. *Bioinformatics.* 2012;28:2106–2113. doi: 10.1093/bioinformatics/bts342.

18. Segata N, Izard J, Waldron L, et al. Metagenomic biomarker discovery and explanation. *Genome Biol.* 2011;12:R60. doi: 10.1186/gb-2011-12-6-r60.
19. Asnicar F, Weingart G, Tickle TL, Huttenhower C, Segata N. Compact graphical representation of phylogenetic data and metadata with GraPhlAn. *PeerJ.* 2015;3:e1029. doi: 10.7717/peerj.1029.
20. Shima K, Coopmeiners J, Graspeuntner S, Dalhoff K, Rupp J. Impact of micro-environmental changes on respiratory tract infections with intracellular bacteria. *FEBS Lett.* 2016;590:3887–3904. doi: 10.1002/1873-3468.12353.
21. Yang CY, Li SW, Chin CY, et al. Association of exacerbation phenotype with the sputum microbiome in chronic obstructive pulmonary disease patients during the clinically stable state. *J Transl Med.* 2021;19:121. doi: 10.1186/s12967-021-02788-4.
22. Li W, Wang B, Tan M, Song X, Xie S, Wang C. Analysis of sputum microbial metagenome in COPD based on exacerbation frequency and lung function: a case control study. *Respir Res.* 2022;23:321. doi: 10.1186/s12931-022-02246-9.
23. Halder K, George L, Wang Z, et al. The sputum microbiome is distinct between COPD and health, independent of smoking history. *Respir Res.* 2020;21:183. doi: 10.1186/s12931-020-01448-3.
24. Diver S, Richardson M, Halder K, et al. Sputum microbiomic clustering in asthma and chronic obstructive pulmonary disease reveals a *Haemophilus*-predominant subgroup. *Allergy.* 2020;75:808–817. doi: 10.1111/all.14058.
25. Kim YH, Park MR, Kim SY, Kim MY, Kim KW, Sohn MH. Respiratory microbiome profiles are associated with distinct inflammatory phenotype and lung function in children with asthma *J Investig Allergol Clin Immunol.* 2023;0 doi: 10.18176/jiaci.0918.

26. Abdel-Aziz MI, Brinkman P, Vijverberg SJH, et al. Sputum microbiome profiles identify severe asthma phenotypes of relative stability at 12 to 18 months. *J Allergy Clin Immunol*. 2021;147:123–134. doi: 10.1016/j.jaci.2020.04.018.
27. Mayhew D, Devos N, Lambert C, et al. Longitudinal profiling of the lung microbiome in the AERIS study demonstrates repeatability of bacterial and eosinophilic COPD exacerbations. *Thorax*. 2018;73:422–430. doi: 10.1136/thoraxjnl-2017-210408.
28. Ramsheh MY, Haldar K, Esteve-Codina A, et al. Lung microbiome composition and bronchial epithelial gene expression in patients with COPD versus healthy individuals: a bacterial 16S rRNA gene sequencing and host transcriptomic analysis. *Lancet Microbe*. 2021;2:e300–e310. doi: 10.1016/S2666-5247(21)00035-5.
29. Wang Z, Liu H, Wang F, et al. A refined view of airway microbiome in chronic obstructive pulmonary disease at species and strain-levels. *Front Microbiol*. 2020;11:1758. doi: 10.3389/fmicb.2020.01758.

Table 1 Characteristics of patients with COPD

Variables	Total (n = 32)	Cluster 1 (n = 15)	Cluster 2 (n = 17)	P
Age (years)	67.12 ± 8.74	69.47 ± 9.90	65.06 ± 7.26	0.156
Sex (male, %)	29 (90.62%)	13 (86.67%)	16 (94.12%)	0.471
BMI (kg/m ²)	22.66 ± 3.74	21.43±3.69	23.74 ± 3.54	0.080
Smoking history				
Current smoker	8	6	2	0.075
Ex-smoker	20	9	11	
Non-smoker	3	0	3	
Smoking index	39.92 ± 29.32	46.20 ± 17.07	33.63 ± 37.48	0.247
Number of exacerbations	2.75 ± 2.214	3.60 ± 2.82	2.00 ± 1.12	0.039
ICS (%)	15 (46.88%)	8 (53.33%)	7 (41.18%)	0.492
FEV ₁ %pred (%)	42.59 ± 24.99	40.25 ± 23.38	44.44 ± 26.92	0.686
CAT score	24.42 ± 8.61	28.00 ± 7.85	21.06 ± 8.13	0.022
mMRC score	2.13 ± 1.20	2.20 ± 1.08	2.06 ± 1.34	0.757
Blood neutrophils (x10 ⁹ /L)	5.58 ± 2.80	6.63 ± 3.29	4.66 ± 1.95	0.045
Blood eosinophils (x10 ⁹ /L)	0.22 ± 0.146	0.26 ± 0.17	0.19 ± 0.11	0.206
Complications				
Diabetes (n, %)	7 (21.88%)	4 (26.67%)	3 (17.65%)	0.552
Hypertension (n, %)	8 (25.00%)	2 (13.33%)	6 (35.29%)	0.194

Data were represented as mean±SD for continuous variables where appropriate and n (%) for categorical variables. Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; ICS, inhaled corticosteroid; FEV₁%pred, predicted percentage of forced

expiratory volume in 1 s; CAT score, COPD Assessment Test; mMRC score, Modified British Medical Research Council

Pre-proof

Figure captions

Figure 1 Participant clustering and characteristics. (A) Calinski–Harabasz (CH) index according to cluster number using partitioning around the medoid clustering method based on Jensen–Shannon divergence at the species level. (B) Two-dimensional (2D) principal coordinate analysis (PCoA) plot for cluster 2. (C) Comparison of blood inflammatory phenotypes and CAT scores in clusters 1 and 2. CAT: COPD Assessment Test; COPD, chronic obstructive pulmonary disease

Figure 2 Dominant microbiotas in the clusters

(A) Linear discriminant analysis (LDA) effect size analysis across the two clusters with $P < 0.05$ and an LDA score > 4.0 . (B) Cladogram showing differentially abundant taxa according to each cluster.

Figure 3 Comparison of microbiota at the genus level among the clusters

Kruskal–Wallis H test with $P < 0.05$.

Figure 4 Correlation between CAT scores and the number of acute exacerbations vs. *Fusobacterium* and *Bacteroides*

CAT: COPD Assessment Test; COPD, chronic obstructive pulmonary disease

Figure 1

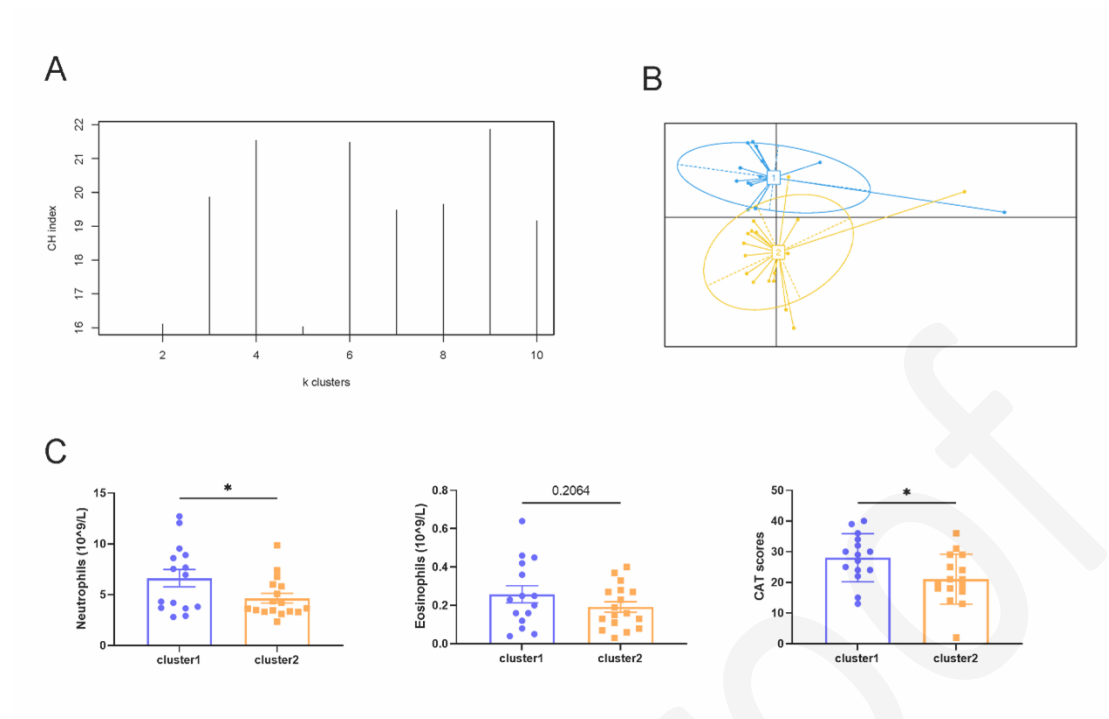


Figure 2

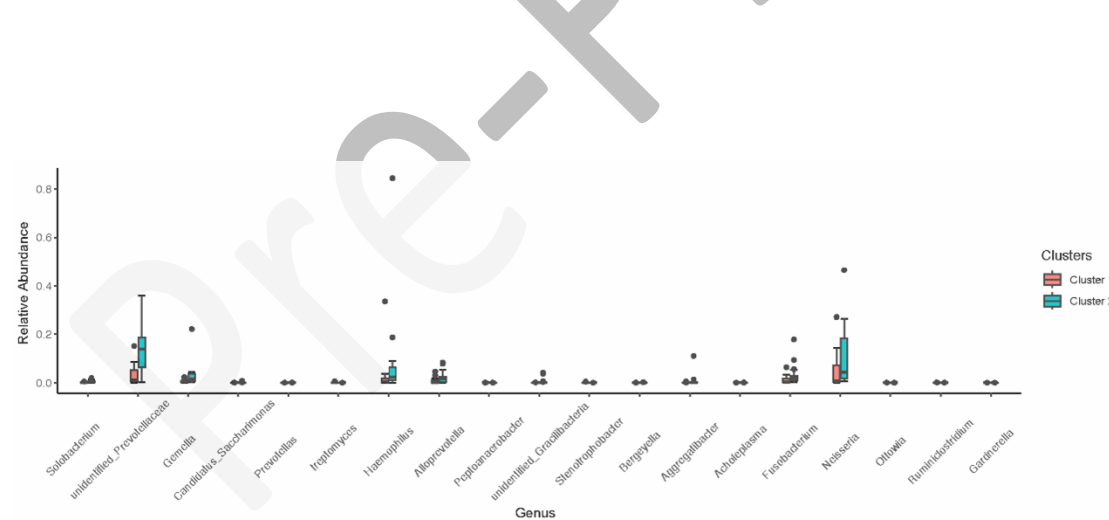


Figure 3

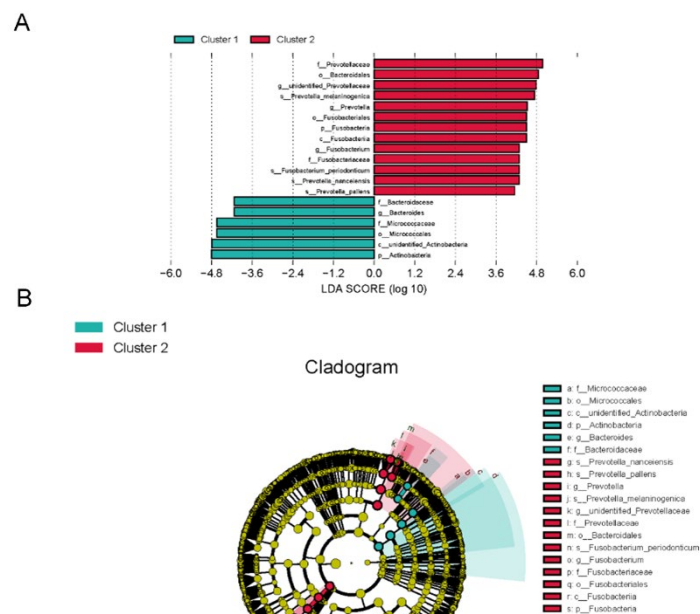
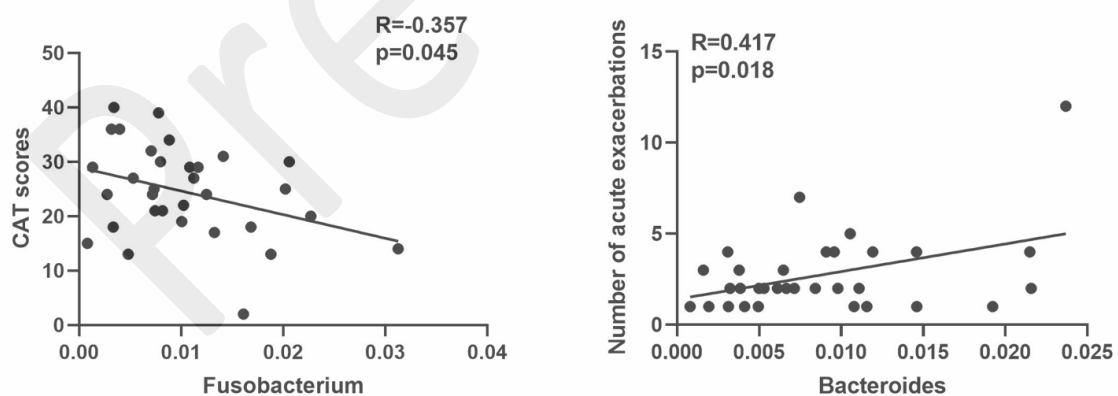
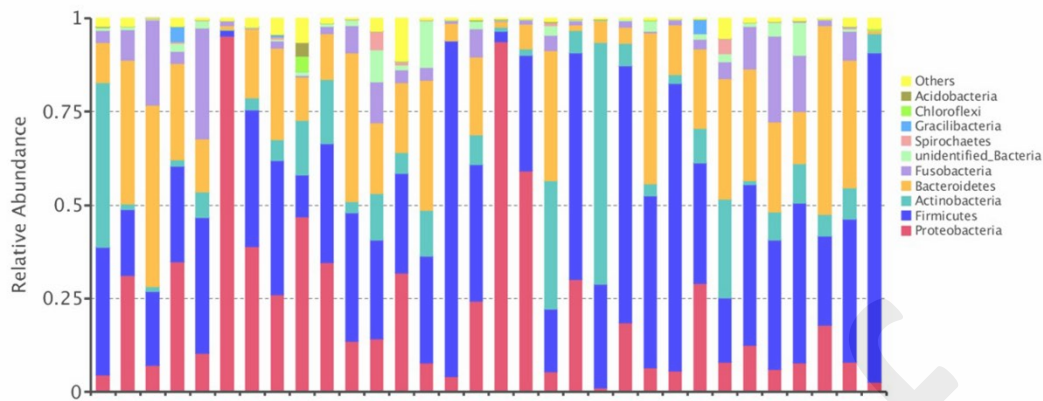


Figure 4

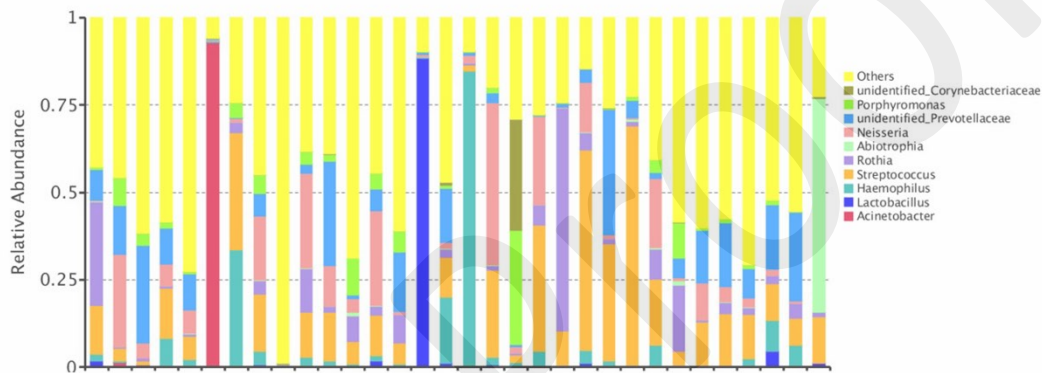


Online Supplement

A



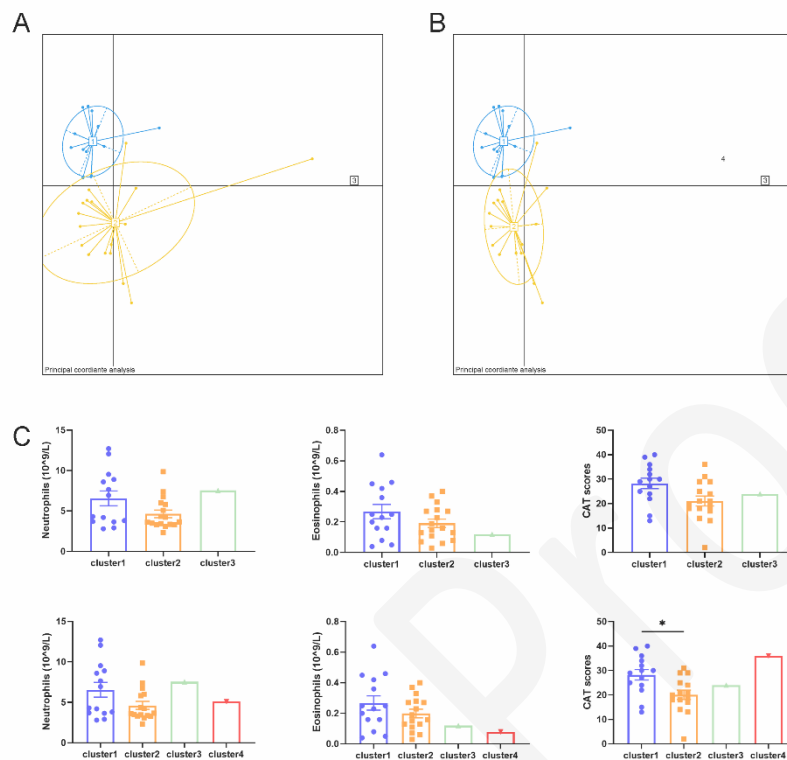
B



Supplementary Figure 1 Composition of the respiratory bacterial community in 32 patients

(A) Microbiome profiles of all stable samples at the phylum level. (B) Microbiome profiles of patients with COPD at the genus level.

COPD, chronic obstructive pulmonary disease



Supplementary Figure 2 Participant clustering and characteristics. (A-B) Two-dimensional (2D) principal coordinate analysis (PCoA) plot for 3 and 4 cluster. (C) Comparison of blood inflammatory phenotypes and CAT scores in clusters. CAT: COPD Assessment Test; COPD, chronic obstructive pulmonary disease

	<i>Prevotella</i>	<i>Fusobacterium</i>	<i>Bacteroides</i>
BMI	0.610	0.461	0.075
Smoking index	0.995	0.925	0.425
Number of exacerbations	0.144	0.309	0.018
FEV ₁ %pred (%)	0.917	0.845	0.910
CAT score	0.555	0.045	0.143
mMRC score	0.780	0.617	0.916
Blood neutrophils	0.901	0.799	0.776
Blood eosinophils	0.686	0.598	0.492

Supplementary Table 1 Correlation analysis between the dominant genera in each cluster and clinical index

Abbreviations: BMI, body mass index; COPD, chronic obstructive pulmonary disease; FEV₁%pred, predicted percentage of forced expiratory volume in 1 s; CAT, COPD Assessment Test; mMRC score, Modified British Medical Research Council