Journal of Inflammation Research

Dovepress
open access to scientific and medical research

Open Access Full Text Article

ORIGINAL RESEARCH

# Shared and Specific Lung Microbiota with Metabolic Profiles in Bronchoalveolar Lavage Fluid Between Infectious and Inflammatory Respiratory Diseases



Yukun He[1]
Wenyi Yu[1]
Pu Ning[1,2]
Qiongzhen Luo[1,3]
Lili Zhao[1]
Yu Xie[1]
Yan Yu[1]
Xinqian Ma[1]
Li Chen[1]
Yali Zheng[1,4]
Zhancheng Gao[1,4]

[1]Department of Respiratory and Critical Care Medicine, Peking University People's Hospital, Beijing, 100044, People's Republic of China; [2]Department of Respiratory and Critical Care Medicine, The Second Affiliated Hospital, Xi'an Jiaotong University, Xi'an, Shaanxi, People's Republic of China; [3]Department of Respiratory & Critical Care Medicine, Beijing Tsinghua Changgung Hospital, School of Clinical Medicine, Tsinghua University, Beijing, People's Republic of China; [4]Department of Respiratory, Critical Care, and Sleep Medicine, Xiang'an Hospital of Xiamen University, School of Medicine, Xiamen University, Xiamen, People's Republic of China

Correspondence: Yali Zheng
Department of Respiratory, Critical Care, and Sleep Medicine, Xiang'an Hospital of Xiamen University, School of Medicine, Xiamen University, Xiamen, People's Republic of China
Email yali_zheng@126.com

Zhancheng Gao
Department of Pulmonary and Critical Care Medicine, Peking University People's Hospital, Beijing, 100044, People's Republic of China
Email zcgao@bjmu.edu.cn

**Background:** Infiltration of the lower respiratory tract (LRT) microenvironment could be significantly associated with respiratory diseases. However, alterations in the LRT microbiome and metabolome in infectious and inflammatory respiratory diseases and their correlation with inflammation still need to be explored.

**Methods:** Bronchoalveolar lavage samples from 44 community-acquired pneumonia (CAP) patients, 29 connective tissue disease-associated interstitial disease (CTD-ILD) patients, and 30 healthy volunteers were used to detect microbiota and metabolites through 16S rRNA gene sequencing and untargeted high-performance liquid chromatography with mass spectrometry.

**Results:** The composition of the LRT microbial communities and metabolites differed in disease states. CAP patients showed a significantly low abundance and both diseases presented a depletion of some genera of the phylum Bacteroidetes, including *Prevotella, Porphyromonas*, and health-associated metabolites, such as sphingosine (d16:1), which were negatively correlated with infectious indicators. In contrast, *Bacillus* and *Mycoplasma* were both enriched in the disease groups. *Streptococcus* was specifically increased in CTD-ILD. In addition, co-elevated metabolites such as FA (22:4) and pyruvic acid represented hypoxia and inflammation in the diseases. Significantly increased levels of amino acids and succinate, as well as decreased itaconic acid levels, were observed in CAP patients, whereas CTD-ILD patients showed only a handful of specific metabolic alterations. Functions related to microbial lipid and amino acid metabolism were significantly altered, indicating the possible contributions of microbial metabolism. Dual omics analysis showed a moderate positive correlation between the microbiome and metabolome. The levels of L-isoleucine and L-arginine were negatively correlated with *Streptococcus*, and itaconic acid positively correlated with *Streptococcus.*

**Conclusion:** In the LRT microenvironment, shared and specific alterations occurred in CAP and CTD-ILD patients, which were associated with inflammatory and immune reactions, which may provide a new direction for future studies aiming to elucidate the mechanism, improve the diagnosis, and develop therapies for different respiratory diseases.

**Keywords:** community-acquired pneumonia, connective tissue disease-associated interstitial disease, microbiome, metabolome, shared and specific alterations

## Introduction

Community-acquired pneumonia (CAP) and connective tissue disease-associated interstitial disease (CTD-ILD) are dynamic, heterogeneous, debilitating lung diseases with multiple comorbidities that affect millions of people worldwide.[1,2] In

Received: 2 October 2021
Accepted: 31 December 2021
Published: 11 January 2022

particular, CAP is acquired outside of a health care setting, and characterized by alveolar infection and intense acute inflammatory responses.[3–6] CTD-ILD is a group of respiratory disorders characterized by dyspnea, and/or impaired pulmonary function,[7] and radiological and histopathological evaluation of the lungs shows patterns of chronic inflammation and fibrosis.[8] These pathophysiological disorders alter lung physiology and could be associated with lung microenvironment dysbiosis.

Human microbial communities are typically complex and involve interactions between immune functions and metabolic pathways to eradicate infection while preventing hyperinflammation.[9–11] In healthy individuals, the composition of the airway microbiota is diverse, well balanced, and may comprise *Prevotella, Streptococcus*, and *Veillonella*, accompanied by other dynamically changing colonizers.[12] Many studies have found significant differences in the composition of the lung microbiome during lung diseases[13] and these diseases are generally associated with a loss of bacterial diversity or the dominance of a small group of taxa.[14]

Although CAP and CTD-ILD are pathogenically and prognostically distinct disorders, they are both characterized by inflammatory destruction of the lung. In upper respiratory tract samples, *Rothia*, and *Streptococcus* were significantly associated with pneumonia and metabolic pathways of sphingolipid, pyruvate, and inositol phosphate were found to be significantly dysregulated in serum of CAP patients.[15] However, studies on the lower respiratory tract (LRT) microenvironment of CTD-ILD are limited. In these two diseases, alterations take place in the LRT microenvironment, and their interactions with the host are worth studying.

Thus, in this study, we explored the composition and functional potential of the microbiota of bronchoalveolar lavage fluid (BALF) samples from 44 CAP patients, 29 CTD-ILD patients, and 30 healthy controls (HC) through 16S ribosomal ribonucleic acid (rRNA) gene sequencing. In addition, we performed metabolomic and lipidomic profiling of BALF samples through untargeted high-performance liquid chromatography with mass spectrometry (HPLC-MS). As a result, we characterized the LRT microenvironment profiles associated with CAP and CTD-ILD and used multi-omics data to identify bacteria–metabolite correlations as a basis for providing new insights into the potential etiologic mechanisms and the tailoring of acute and chronic inflammatory lung disease therapies in the future.

# Materials and Methods

## Patient Screening, Enrollment, and Clinical Examination

Application of the study was submitted to the Ethical Review Committee of Peking University People's Hospital (No.2016PHB202-01) and the CAP cohort was registered at ClinicalTrials.gov (NCT03093220). The study was conducted under the principles of the Helsinki Declaration and patient informed consent was obtained from all participators or their next of kin for incapacitated patients or unconscious subjects who were unable to give informed consents.

Forty-four CAP patients admitted to any of the four participating hospitals between March 2017 and August 2017 as part of a multicenter clinical study were enrolled. Twenty-nine patients with CTD-ILD were admitted to Peking University People's Hospital and thirty healthy volunteers from the medical center were recruited. CAP and CTD-ILD were defined according to the standard published by the America/American Thoracic Society.[2,16]

## BALF Sample Collection, Detection, and Storage

Bronchoscopy was performed within 72 h after hospital admission. Briefly, 100 mL of sterile normal saline was instilled into the diseased region according to CT scan results (for CAP patients), or into the right middle lobe or lingual lobe (for controls). Aliquots were then retrieved by gentle suctioning. Cell counts and protein concentrations of BALF samples were analyzed immediately after collection. The remaining BALF samples underwent centrifugation and then supernatants and precipitates were separated and stored at −80 °C.

## 16S Ribosomal RNA Gene Sequencing

Total DNA from BALF samples was extracted using the CTAB/SDS method. 16S rRNA genes of the V3-V4 region were amplified with primers (338F and 806R) and the PCR products were detected by agarose gel electrophoresis (2%) and then mixed in equidensity ratios. The mixture of PCR products was purified with the QIAEX II Gel Extraction Kit (QIAGEN, Germany). Sequencing libraries were generated using the TruSeq® DNA PCR-Free Sample Preparation Kit (Illumina, USA) and then sequenced on the HiSeq2500 platform and 250 bp paired-end raw reads were generated.

## Sequencing of Reagent Controls

To detect contamination introduced by the experimental operations, specimen-free reagent controls for extraction and PCR amplification (no template, ddH2O only) were included in the experiment. Possible contaminants were excluded from the present analysis using the R "Decontam" package.

## 16S rRNA Gene Sequencing Analysis

Paired-end reads were merged and quality filtering, trimming and dereplication of the raw tags were performed under specific conditions to obtain the high-quality clean reads using VSEARCH.[17] Chimeric sequences were detected by aligning clean reads to the Silva database (v.132)[18] with the UCHIME algorithm. The effective reads were denoised to generate amplicon sequence variants (ASVs) using unoise3. Taxonomy assignment was performed on ASVs using USEARCH (v10) against the RDP database (v 11.5)[19] and Greengenes database.[20]

The Abundance-based coverage estimator (ACE) and the Shannon index were calculated to evaluate alpha diversity. Beta diversity was assessed by permutational multivariate analysis of variance (PERMANOVA) and ANOSIM based on Bray-Curtis distances. Differential bacterial taxa among groups were assessed using the "makeContrasts" function within a multifactorial design in the metagenomeSeq v1.35.0 R package. Statistically significant bacterial differences (LDA > 2, $P < 0.05$) associated with different groups were explored using linear discriminant analysis (LDA) effect size (LEfSe). The microbiome phenotypes were predicted by BugBase[21] based on the Greengenes annotation. PICRUSt2 was used to identify predicted associated pathways from the inferred metagenomes of taxa with the "stratified" mode.[22] Comparisons of the predicted pathways were obtained with STAMP. Spearman's rho was calculated using the "corr.test" function within the R package psych v1.8.12105 based on centered log-ratio-transformed genome relative abundance.

## Untargeted Metabolic Profiling Analysis

The detailed method of BALF sample pretreatment is described in the Supplement 3. All compounds in BALF were analyzed using a Cortecs C18 column (2.1 × 100 mm, Waters) on an Ultimate 3000 UHPLC (Dionex) system coupled with a Q Exactive (Orbitrap) mass spectrometer (Thermo Fisher, CA). Detailed parameters for the untargeted metabolic and lipidomic analyses were set following the protocols of our previously reported study.[23] Data-dependent MS/MS acquisition (DDA) of all samples was analyzed using TraceFinderTM (Thermo, CA). Metabolites and lipids were identified based on matching precursor and characteristic fragment masses and then assigned using in-house databases in "screening" mode. Any metabolite feature with more than 20% missing values was removed from the result.[24,25] Missing values were estimated by the Bayesian PCA (BPCA) method. Data were normalized by the QC group and then auto-scaled using MetaboAnalystR 3.0.[26] The normalized data were used for downstream analysis. Principal component analysis (PCA) and orthogonal projections to latent structures discriminant analysis (OPLS-DA) were conducted using SIMCA v14.1 (Umetrics, Sweden).

## Metabolomic and Metagenomic Data Integration

Correlation analysis between metagenomic and metabolomic data was undertaken using ASVs and metabolites. DIABLO function from the mixOmics v6.6.2 R package[27,28] was used to generate integrated metagenomic and metabolomic signatures. The analysis was performed using centered log-ratio-transformed taxa relative abundance and log-transformed auto-scaled metabolite data. The optimum number of components and variables included within the final model was determined using the "tune.block.splsda" function with 50 × 10-fold cross-validation.[29]

## Statistical Analysis

All categorical variables are presented as numbers (percentages), parametric continuous variables are presented as the mean ± SD, and nonparametric continuous variables are presented as median and interquartile ranges (25th and 75th percentiles). Student's *t*-test or analysis of variance (ANOVA) with post-hoc Tukey HSD test were used to analyze continuous parametric data, whereas continuous nonparametric data were analyzed using Mann–Whitney U or a Kruskal–Wallis test. All categorical data were analyzed using the chi-square or Fisher's exact test. All tests were two-sided, the p values were corrected using FDR, and $p < 0.05$ was considered statistically significant.

# Results

## Clinical Characteristics of the Study Population

To identify clinical differences among groups, we first separately described the clinical characteristics of the 44 CAP patients, 29 CTD-ILD patients, and 30 healthy controls (Table 1). Lymphocyte counts and albumin (ALB) levels in peripheral blood (PB) were lower in the CAP patients and CTD-ILD patients than in the HC group, while polymorphonuclear cell (PMN) percentages and albumin (ALB) levels in the BALF, and inflammatory markers in the CAP patients were significantly higher than those in the HC group. Creatinine and total triglyceride (TG) contents were significantly different between the CAP and CTD-ILD groups. No significant differences were observed in age or various comorbidities among the groups. Collectively, these data indicated that CAP patients exhibited substantially different clinical characteristics from those of CTD-ILD or HC subjects.

## The LRT Microbial Community is Altered in CAP and CTD-ILD Patients

To explore the LRT bacterial community composition in different diseases, we conducted high throughput16S rRNA gene sequencing. Rarefaction curves of all samples indicated that the sequencing depth was sufficient to capture the full diversity of taxa associated with each microbial community (Figure S1). In total, 7551 ASVs were identified across all 103 BALF samples (Table S1). After filtering for sequence variants present in at least two samples with a minimum relative abundance of 0.05%,[29] we obtained a total of 1782 sequence variants for community analysis.

In terms of alpha diversity, the ACE index showed a significant decrease in species abundance in the CAP subjects (Figure 1A, all $P < 0.05$), although no significant differences were observed in the Shannon index (Figure 1B). The PERMANOVA index showed statistically significant differences among all groups (Figure 1C and E). NMDS based on the Bray–Curtis distances showed tight clustering of samples from the CTD-ILD group, whereas CAP group samples were distributed diffusely (Figure 1D) and the ANOSIM index presented a larger variability within the CAP group (Figure 1F). We compared the subgroups within the samples and found no significant difference in microbiota composition in terms of smoking status ($P$ = 0.509) or sex ($P$ = 0.085) (Figure S2). Thus, the LRT microbial community was altered in the diseased groups, and the CAP group presented a lower species abundance and higher heterogeneity.

## Shared vs Disease-Specific Microbial Alterations in CAP and CTD-ILD Patients

To identify significantly different genera, we employed the metagenomeSeq R package within a multifactorial design. Moreover, correlation analysis to identify bacteria linked with specific clinical indicators was performed. The four phyla, Proteobacteria, Firmicutes, Bacteroidetes, and Fusobacteria, dominated the community composition across all cohorts (Figure 2A). Specifically, CTD-ILD samples had a higher abundance of Proteobacteria and Firmicutes. *Prevotella* and *Alloprevotella* were decreased in both CAP and CTD-ILD patients and were negatively correlated with white blood cell counts (WBCs) and C-reactive protein (CRP) levels (Figure 2B, Table S2). *Bacillus* was increased in diseased groups. The abundances of the genus *Pseudomonas*, and well-known gut bacteria, such as *Enterococcus*, were all significantly increased in CAP patients. *Pseudomonas* was positively correlated with CRP level. Additionally, *Streptococcus intermedius* which was enriched in the CTD-ILD group was moderately correlated with the erythrocyte sedimentation rate (ESR) level ($r$ = 0.40, $P$ = 0.003). The differentially abundant genera were further confirmed using LEfSe analysis (Figure 2C), revealing that 19 and 2 genera were discriminative for CTD-ILD and CAP patients, respectively. *Streptococcus* and *Haemophilus* were preponderant in CTD-ILD patients, whereas CAP patients presented a higher abundance of *Mycoplasma*.

## Predicted Functional Characterization of the LRT Microbiome

Predicted phenotypes based on taxonomic classification were analyzed with BugBase. The results suggested that gram-positive bacteria were more abundant in the CAP patients than in the HC group, while the abundance of gram-negative bacteria was decreased in the CAP group (Figure S3). No significant difference was observed in the CTD-ILD group.

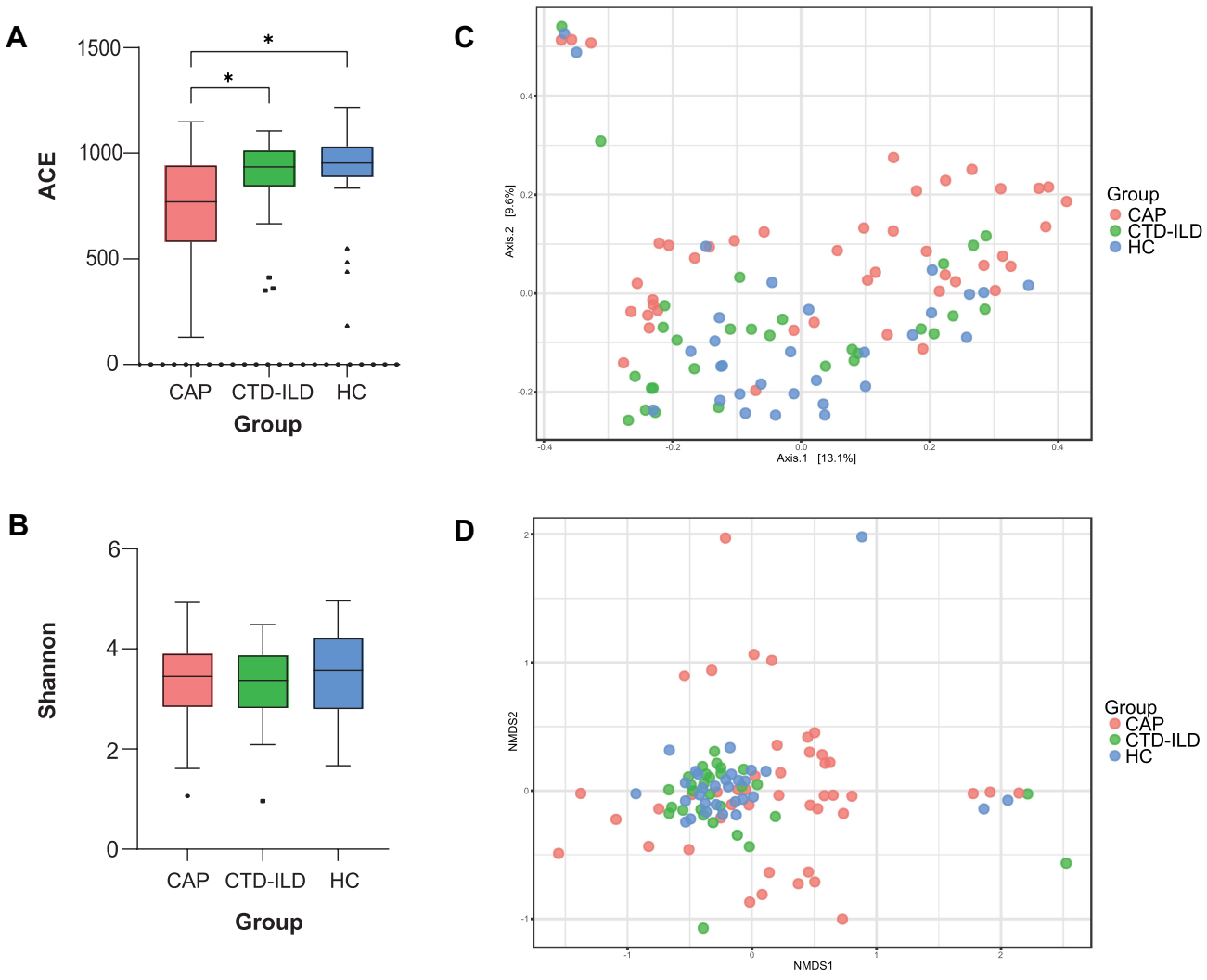
To explore differences in the potential function of microbiota between the groups, we used PICRUSt2 based on the Kyoto Encyclopedia of Genes and Genomes

**Table 1** Demographical and Clinical Features of Included Subjects

| | CAP (n = 44) | CTD-ILD (n = 29) | HC (n = 30) | *p*-value |
|---|---|---|---|---|
| Age | 51.91±3.07 | 51.21±2.46 | 56.10±2.22 | 0.449 |
| Gender, n, (% male) | 28(63.6) | 6(20.7) | 11(36.7) | 0.002 [a,b] |
| Smokers, n (%) | 8(18.2) | 2(6.9) | 0(0) | 0.407 |
| Comorbidities | | | | |
| Diabetes Mellitus, n (%) | 6(13.6) | 6(20.7) | 3(10.0) | 0.087 |
| Hypertension, n (%) | 14(31.8) | 6(20.7) | 11(36.7) | 0.522 |
| Hyperlipidemia, n (%) | 9(20.5) | 5(17.2) | 3(10.0) | 0.119 |
| Coronary Heart Disease, n (%) | 5(11.4) | 2(6.9) | 5(16.7) | 0.166 |
| Laboratory Findings | | | | |
| Peripheral blood | | | | |
| WBC (× $10^{9}$/L) | 6.40(5.13–9.73) | 5.74(4.52–7.80) | 6.87(5.68–8.31) | 0.193 |
| Neutrophils(×$10^{9}$/L) | 4.27(3.55–7.13) | 3.70(3.05–5.35) | 4.03(3.08–4.97) | 0.180 |
| Lymphocytes(×$10^{9}$/L) | 1.03(0.70–1.50) | 1.30(0.80–4.88) | 1.91(1.56–2.66) | 0.000 [a,c] |
| NLR | 3.43(2.31–991) | 0.06(0.05–0.19) | 2.05(1.49–2.63) | 0.000 [a,c] |
| BUN (mmol/L) | 4.18(2.72–5.40) | 4.67(3.3–6.1) | 4.92(4.07–5.6) | 0.243 |
| Cr (μmol/L) | 65.5(55–83.75) | 52(39–67) | 60(56–72) | 0.009 [b] |
| ALT (U/L) | 34.5(12–60.25) | 26(20–35.5) | 17.5(12–29.8) | 0.029 [a] |
| AST (U/L) | 28(20.5–73.75) | 27(20.5–37) | 19(16–22.5) | 0.000 [a,c] |
| CK (U/L) | 77.5(46.5–146) | 64.5(45.5–134) | 58(44–85) | 0.312 |
| ALB (g/L) | 35(29–39.78) | 34.8(31.5–37.9) | 44.1(41–45.4) | 0.000 [a,c] |
| Glucose (mmol/L) | 5.18(4.59–6.48) | 4.7(4.53–6.35) | 4.89(4.62–5.73) | 0.323 |
| TC (mmol/L) | 3.51(2.83–4.36) | 4.78(4.27–5.44) | 4.63(3.82–5.56) | 0.000 [a,b] |
| TG (mmol/L) | 0.92(0.75–1.58) | 1.69(1.22–2.13) | 1.4(0.89–2.37) | 0.01 [b] |
| BALF related | | | | |
| PMN percentages (%) | 18.00(1.75–65.50) | 3.00(1.00–9.75) | 1.00(0.50–2.00) | 0.000 [a] |
| Lymphocyte percentages (%) | 21.00(11.00–43.50) | 35.00(14.75–54.75) | 12.00(8.00–30.75) | 0.059 |
| Eosinophil percentages (%) | 0.00(0.00–1.00) | 0.00(0.00–0.75) | 0.00(0.00–1.00) | 0.537 |
| Albumen concentration (g/L) | 1128.69(312.07–2230.86) | 171.42(75.04–323.68) | 158.88(75.50–220.13) | 0.000 [a,b] |
| Inflammatory markers | | | | |
| PCT (μg/L) | 0.16(0.05–1.02) | 0.06(0.05–0.19) | 0.05(0.05–0.09) | 0.032 |
| CRP (mg/L) | 40.75(12.12–132.75) | 2.52(0.67–7.91) | 1.37(0.78–2.81) | 0.000 [a,b] |
| ESR (mm/h) | 41.00(20.00–63.00) | 16.00(7.00–29.00) | 8.50(6.00–13.50) | 0.000 [a,b] |

**Notes**: [a]Statistical significance exists between CAP and healthy controls; [b]Statistical significance exists between CAP and CTD-ILD [c]Statistical significance exists between CTD-ILD and healthy controls.

**Abbreviations**: BALF, bronchoalveolar lavage; PMN, polymorphonuclear leukocyte; TC, serum total cholesterol; TG, serum total triglycerides; AST, aspartate aminotransferase; ALT, alanine aminotransferase; CK, creatine kinase; Cr, creatinine; WBC, white blood cell; N, neutrophil; CRP, C-reactive protein; ESR, erythrocyte sedimentation rate; PCT, procalcitonin; ALB, albumin.




**E**

| Group | Df | SumsOfSqs | R2 | Pr(>F) |
|---|---|---|---|---|
| CAP vs. HC | 1(73) | 1.1365 | 0.04298 | 0.001 |
| CAP vs. CTD-ILD | 1(72) | 1.0518 | 0.03895 | 0.001 |
| CTD-ILD vs. HC | 1(58) | 0.5356 | 0.02711 | 0.021 |

**F**

| Group | R-value | P-value |
|---|---|---|
| CAP vs. HC | 0.03423 | 0.151 |
| CAP vs. CTD-ILD | 0.07261 | 0.038 |
| CTD-ILD vs. HC | 0.04236 | 0.019 |

**Figure 1** Lower respiratory tract microbiome alpha and beta diversity of taxonomic analysis. (**A**) Comparison of the ACE index based on the genus profile in different groups for assessment of microbiome alpha diversity of three groups. * $P \leq 0.05$. (**B**) Comparison of the Shannon index based on the genus profile in different groups. (**C**) Beta diversity was assessed by PERMANOVA based on Bray-Curtis distances using principal coordinate analysis (PCoA). (**D**) Beta diversity was assessed by ANOSIM based on Bray-Curtis distances using non-metric multidimensional scaling (NMDS). (**E**) PERMANOVA index of groups. (**F**) ANOISM index of groups.

(KEGG)[30] and Metacyc[31] databases. No significant differences were observed in the overall predicted pathways among all samples (Figure S4). However, several annotated functions were differentially expressed among the groups at the individual pathway level (Tables S3–6). The functions "glycerolipid metabolism" and "androstenedione degradation" were significantly enriched in the CAP group whereas "D-arginine and D-ornithine metabolism" was significantly enriched in the CTD-ILD group. These results suggested that the microbial metabolic capability

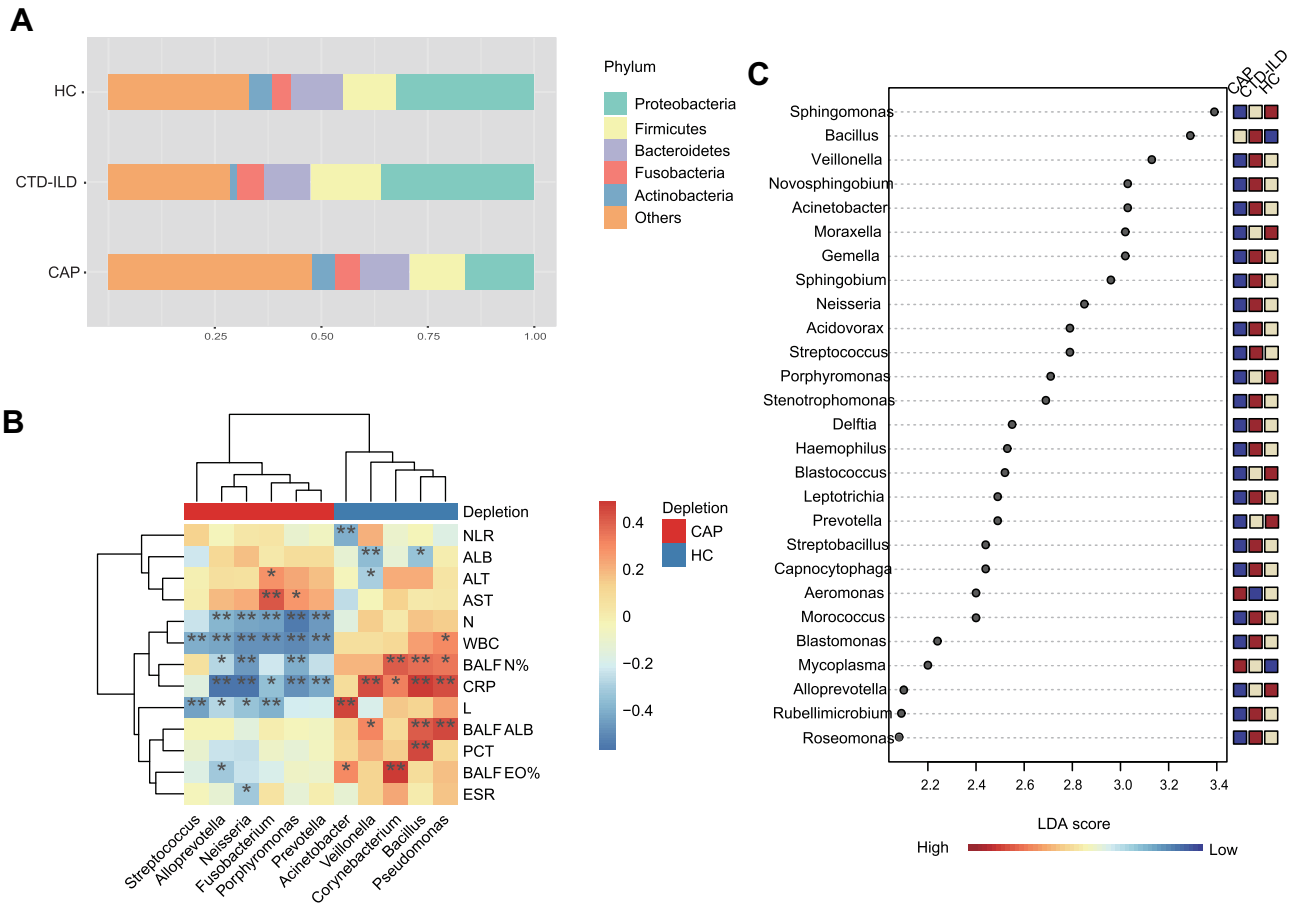



**Figure 2** Taxonomic analysis of the lower respiratory tract microbiome. (**A**) The relative abundance of microbial communities at the phylum level among groups. (**B**) Spearman's rho calculated between ASVs and clinical indicators. (**C**) Linear discriminant analysis of effect size (LEfSe) of ASVs enriched in each group. Enrichment in different groups indicated by the colored bar along the left of the heatmap. Black stars within heatmap boxes indicate significant results (* $P \leq 0.05$, ** $P \leq 0.01$), Benjamini–Hochberg adjustment for multiple comparisons. ASV abundances were centered with log-ratio transformation prior to analysis.

**Abbreviations**: NLR, Neutrophil to Lymphocyte Ratio; AST, aspartate aminotransferase; ALT, alanine aminotransferase; ALB, albumin; WBC, white blood cell; N, neutrophil; CRP, C-reactive protein; ESR, Erythrocyte Sedimentation Rate; PCT, procalcitonin; BALF_N%, the percentage of polymorphonuclear leukocyte in bronchoalveolar lavage; BALF_L%, the percentage of lymphocyte in bronchoalveolar lavage; BALF_EO%, the percentage of eosinophil in bronchoalveolar lavage; BALF, ALB albumin in bronchoalveolar lavage.

could contribute to a disease-associated disruption of the LRT metabolic profiles.

## Shared vs Disease-Specific Metabolic Alterations in CAP and CTD-ILD Patients

To explore the metabolic and lipidomic profiles of the LRT among CAP and CTD-ILD patients, we conducted an untargeted HPLC-MS analysis of 103 paired BALF samples. This assay identified a total of 505 compounds that likely originated from either the microbiota or the host. According to the PCA result, we found that, in agreement with NMDS analysis of microbiota, CAP samples were more diffusely grouped (Figure 3A). OPLS-DA score plots showed a clear separation among all groups (Figure S5).

To identify metabolic alterations, we calculated the fold changes (FCs), *P* values, and OPLS-DA variable importance in the projection (VIP) scores for all metabolic features. Using a linear model with adjustment for covariates (age, sex, smoking, DM2, and hyperlipidemia), we found that 47 compounds showed the same trend in disease groups, and 181 and 36 compounds were identified as specific to CAP and CTD-ILD patients, respectively ($P \leq 0.05$, and VIP > 1) (Figure 3B, Tables S7–9).

Eight lipid species were increased in both diseases. Notably, PI (18:0/20:4) was moderately associated with elevated CRP (r = 0.451) and neutrophils in BALF (r= 0.417), and PC (18:0p/20:4) was moderately associated with elevated ESR (r = 0.564) (Figure 3C, Table S10). These two lipids have one of the acyl chains constituted by C20:4, that is, the

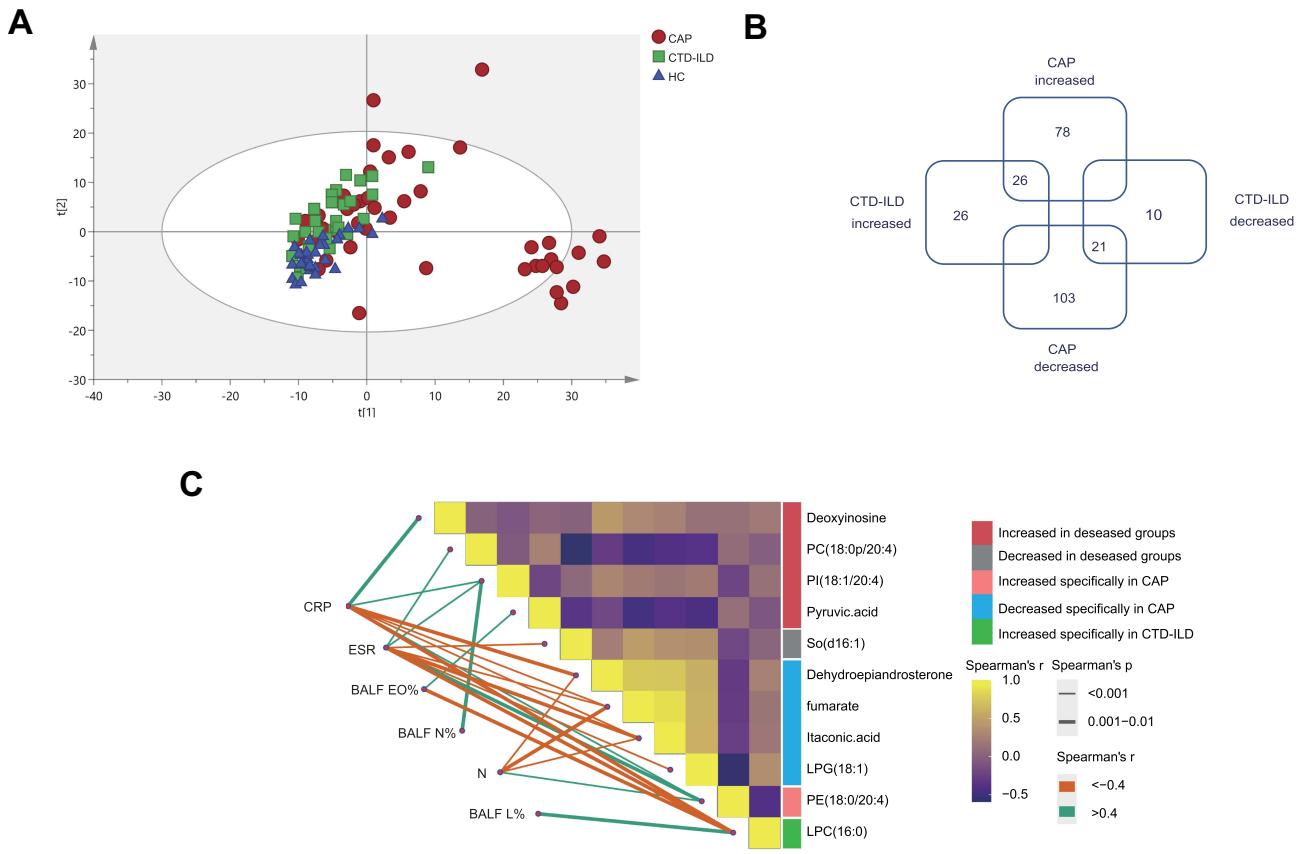



**Figure 3** Metabolic and lipidomic profiles of CAP patients and controls. (**A**) A principal component analysis (PCA) score plot of metabolic and lipidomic profiles in BALF samples. PCA score plot colored according to sample group: red circles, severe CAP (SCAP); blue triangles, healthy control (HC); green squares, CTD-ILD. (**B**) Venn plot to identify the differential metabolites associated with CAP and CTD-ILD patients ($P \leq 0.05$, and VIP > 1). (**C**) Correlation analysis among metabolites and relationships with inflammatory mediators. The green lines represent positive correlations and the Orange lines represent negative correlations. WBC white blood cell, N neutrophil, CRP C-reactive protein, ESR Erythrocyte Sedimentation Rate, BALF_N% the percentage of polymorphonuclear leukocyte in bronchoalveolar lavage, BALF_L% the percentage of lymphocyte in bronchoalveolar lavage, BALF_EO%, the percentage of eosinophil in bronchoalveolar lavage.

structure of arachidonic acid, which can be oxidized to prostaglandins, well-known players in the inflammatory response. Pyruvic acid was also elevated in both diseases due to active glycolysis. In contrast, sphingosine [So] (d16:1) was decreased in the disease groups and inversely related to ESR level (r value = −0.562).

For CAP-specific metabolic alterations, the levels of 12 kinds of amino acid and nucleotide-related metabolites in BALF were significantly increased, suggesting proteolytic activity and DNA damage induced by oxidative stress. Besides, we found significantly lower levels of fumarate and itaconic acid and increased succinate levels in the CAP groups, indicating remodeling of the tricarboxylic acid (TCA) cycle accompanying inflammatory macrophage activation. However, only a handful of metabolites were uniquely altered in CTD-ILD.

Collectively, the co-elevated metabolites represented general disorders in the disease states, such as hypoxia and inflammation. In acute infection, metabolic disorders are more evident in the LRT.

## Dual-Omics Analysis Reveals the Relationship Between the LRT Microbiota and Metabolites in CAP and CTD-ILD

We next sought to examine the associations between the microbiome and metabolome. A moderate positive correlation was observed in the dual omics analysis (Figure 4A). Ultimately, two major microbiome/metabolite clusters were defined (Figure 4B). The first cluster showed associations between a group of five species and nine metabolites. All five species were depleted in CAP patients. *Porphyromonas*, and *Streptococcus* were negatively correlated with the amino acids, Phe, Arg, and Ile, and positively correlated with fumarate and itaconic acid. The second network did not contain any differential metabolites and therefore may represent interactions unrelated to disease states.

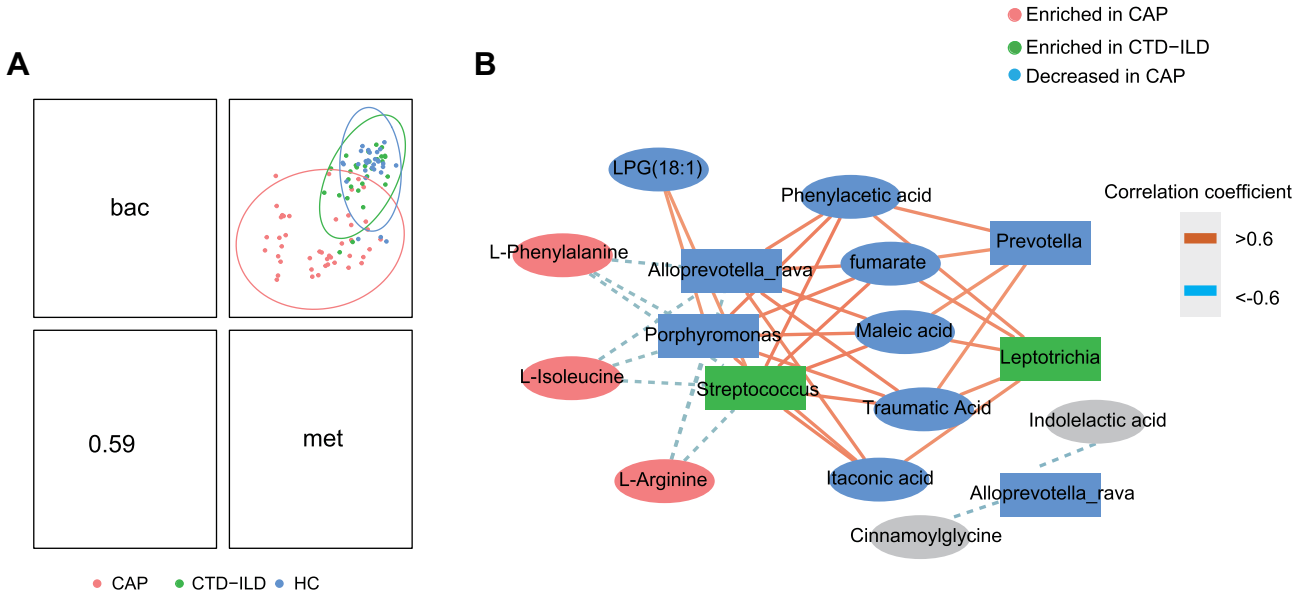



**Figure 4** Integration of respiratory microbiomes and metabolomes. (**A**) Integration of microbiome and metabolome datasets using DIABLO software. DIABLO sample plot demonstrating the overall correlation between microbiome and metabolome data. (**B**) DIABLO sample plot demonstrating the microbiome-metabolome network. A positive correlation between nodes is indicated by Orange connecting lines, and a negative correlation is indicated by blue. Species and metabolites enriched in CAP and CTD-ILD are denoted by red and green, respectively. Microbiome data filtered for genomes with a minimum of 0.05% relative abundance in ≥2 samples. Microbiome data are centered log-ratio-transformed relative abundance. Metabolomics data are log-transformed auto-scaled values.

## Discussion

Several studies have suggested that infiltration of the LRT microenvironment can be related to the pathogenesis of respiratory diseases.[32,33] However, due to obvious limitations in obtaining lower airway samples and heterogeneity of lung diseases, the LRT microbiome and associated metabolites have remained poorly characterized. In this study, we present the first analysis to our knowledge of the microbiota and metabolome of the human LRT in both CAP and CTD-ILD patients. We revealed that these two diseases had shared and specific alterations in the lung microenvironment, which could be informative for future research exploring disease mechanisms, or could suggest potential therapeutic targets through validation of their effects in model organisms.

Importantly, both diseases showed a depletion of *Prevotella, Alloprevotella*, and *Porphyromonas* compared to the HC group, and all three genera belong to the phylum Bacteroidetes. Duncan reported that most Bacteroidetes are pH sensitive and fail to prosper in acidic environments,[34] so this finding may be explained by the decreased local pH of the LRT microenvironment due to inflammation.[35] In the CAP subjects, loss of overall species abundance was observed, which may diminish colonization resistance against pathogens. We found that some gut-associated bacteria were specifically more abundant in CAP patients, and previous studies have indicated that increased lung and gut permeability may induce bacterial migration in various respiratory diseases, such as acute respiratory disease syndrome (ARDS).[33,36] Besides, *Pseudomonas*, which is a frequent pathogenic colonizer in COPD and asthma patients[37] and predispose the host to more severe respiratory viral pneumonia,[38] was increased in the CAP group and correlated with CRP level. Exoproteins from *Pseudomonas* were reported to disrupt the mucosal barrier and induce IL- 6 production potentially contributing to mucosal inflammation.[39] Thus, in CAP patients even without ARDS, the overgrowth of certain pathogens under acute infection may lead to the breakdown of the lung barrier and the induction of gut-lung bacterial translocation, which further correlates with more severe inflammation.[33,37,38,40]

In the CTD-ILD group, many putative commensal microbes, such as *Veillonella* and *Streptococcus*, were enriched compared with HC. *Streptococcus intermedius* enrichment was associated with increased ESR. Liu reported that *Streptococcus intermedius* secreted a histone-like DNA binding protein, which induced proinflammatory cytokine production in a macrophage-derived cell line.[41] Therefore, the increased abundance of *Streptococcus intermedius* and its interaction with the

host immune system could be related to the inflammatory status of CTD-ILD.

Predicted phenotypes based on BugBase suggested that gram-positive bacteria were more abundant in the CAP patients, while the abundance of gram-negative bacteria was decreased in the CAP group. According to guidelines for the diagnosis and treatment of adult community-acquired pneumonia in China,[42] common pathogens of CAP include gram-positive bacteria such as *Streptococcus pneumoniae* and *Staphylococcus aureus*, as well as mycoplasma pneumoniae/chlamydia. But for the hospital acquired pneumonia (HAP), common pathogens are gram-negative bacteria. This high level of the analysis may indicate a higher overall level of gram-positive bacteria in the lower respiratory flora in CAP patients, which may be related to the etiology of CAP.

PICRUSt2 analysis also suggested that microbial metabolic function may contribute to the LRT metabolome, and hence respiratory health. The predicted function associated with "D-arginine and D-ornithine metabolism" was significantly enriched in the CTD-ILD group, which plays a vital role in pulmonary fibrosis pathogenesis. Arginine can be converted into ornithine through arginase and proline, as a metabolite from ornithine, participates in fibrosis.[43] Thus, arginase inhibitors might be potential therapeutic agents for CTD-ILD by rectifying immune imbalance and arginine metabolism disorder.[44]

Dual-omics analysis revealed that *Streptococcus* was negatively correlated with three amino acids but positively correlated with itaconic acid. *Streptococcus* is one of the most abundant genera inhabiting the respiratory tract of the healthy population.[45] LEfSe analysis showed that *Streptococcus* was increased in CTD-ILD patients but decreased in CAP patients. Besides, we found that the CAP group was highly enriched for various amino acids, which is consistent with the finding in the inflamed CF airway[46] and had a significantly lower level of itaconic acid. Itaconic acid is derived from activated macrophages and can inhibit inflammation at low doses but promote inflammatory apoptosis at high doses.[47] Thus, in the HC group, the balanced composition of commensal flora may lead to relative amino acid starvation in the airway that could restrict viral invasion[11] and may also inhibit excessive inflammation. Once dysbiosis occurs in the LRT, depletion or abnormal enrichment of commensal flora will cause the host's original balanced immune state to be disrupted, and metabolic reprogramming of immune cells may also occur, causing invasion of pathogens and then an excessive inflammatory response.

In this context, it is possible that targeting specific metabolites and/or pathways in combination with narrow-spectrum antibiotics for specific microbes, can represent a potential therapeutic strategy to control acute or chronic inflammation in respiratory diseases.

The present study had some limitations. First, in all identified ASVs, the proportion of "unassigned" in the BALF samples reached 35% and varied from 10% to 40% in the previous studies,[37,48,49] which may be caused by the imperfect recall (not all sequences or taxa are detected) or imperfect precision (additional false sequences or taxa are detected) of the 16S V3-V4 rRNA sequencing.[50] Meanwhile, the different phenomena of species within genera highlight the interspecies variability that complicates microbiome data interpretation, so it is important to identify organisms at the species level. However, resolution of 16S method at the species level is especially low, so bacterial community analysis using full-length 16S rRNA gene or metagenomics is needed. Second, the discovered statistical correlations between metabolites and clinical indicators were generally moderate, which could likely be improved with a further larger study. Third, the relationship between bacteria and metabolites was limited and further verification using metatranscriptome analysis and animal models is necessary.

Despite these limitations, our study is the first to our knowledge to compare the LRT environment between acute infections and auto-immune disorders, providing a benchmark for future studies evaluating associations between the LRT microenvironment and the immune homeostasis. Our results provide more nuanced insight into dysbiosis of the LRT microenvironment, revealing shared and specific alterations in CAP and CTD-ILD patients. These findings may provide new insights into the etiologies and pathophysiological mechanisms of different respiratory diseases. However, the causal interactions between the observed metabolic and microbiome changes and the host in disease development remain to be elucidated. Moreover, the alterations in the LRT microenvironment should be interpreted with caution, as many identified alterations may be indicative of general disease states. Thus, microbes or metabolites that are non-specifically associated with multiple diseases would not be useful as diagnostics. On the other hand, bacteria and metabolites that are associated with specific diseases could be developed into diagnosis and therapy targets.

## Data Sharing Statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://www.ncbi.nlm.nih.gov/, PRJNA751994.

## Ethics Approval and Informed Consent

Application of the study was submitted and approved by the Ethical Review Committee of Peking University People's Hospital (No.2016PHB202-01) and the CAP cohort was registered at ClinicalTrials.gov (NCT03093220). The study was conducted under the principles of the Helsinki Declaration and patient informed consent was obtained from patients or their next of kin.

## Consent for Publication

All authors agreed to the publication of the article.

## Acknowledgments

The authors wish to thank staff members of the cooperating hospitals (Department of Pulmonary and Critical Care Medicine, Shanghai Pulmonary Hospital, Shanghai, China; Department of Pulmonary and Critical Care Medicine, West China Hospital, Sichuan, China; and Department of Pulmonary and Critical Care Medicine, Second Hospital of Jilin University, Jilin, China) for assistance with samples and clinical data collection. They are also grateful to the Metabolomics Facility at the Technology Center for Protein Sciences, Tsinghua University for technical support.

## Funding

This work is funded by the National Natural Science Foundation of China (No.82000019, No. 81870010), National Key Research and Development Programme of China (No. 2016YFC0903800).

## Disclosure

The authors declare that they have no competing interests.

## References

1. Schwarzkopf L, Witt S, Waelscher J, Polke M, Kreuter M. Associations between comorbidities, their treatment and survival in patients with interstitial lung diseases - a claims data analysis. *Respir Res*. 2018;19(1):73. doi:10.1186/s12931-018-0769-0
2. Mandell LA, Wunderink RG, Anzueto A, et al. Infectious Diseases Society of America/American Thoracic Society consensus guidelines on the management of community-acquired pneumonia in adults. *Clin Infect Dis*. 2007;44(Suppl 2):S27–72. doi:10.1086/511159
3. Welte T, Torres A, Nathwani D. Clinical and economic burden of community-acquired pneumonia among adults in Europe. *Thorax*. 2012;67(1):71–79. doi:10.1136/thx.2009.129502
4. Cruz CS, Ricci MF, Vieira AT. Gut Microbiota Modulation as a Potential Target for the Treatment of Lung Infections. *Front Pharmacol*. 2021;12:724033. doi:10.3389/fphar.2021.724033
5. Wunderink RG, Waterer G. Advances in the causes and management of community acquired pneumonia in adults. *BMJ*. 2017;358:j2471. doi:10.1136/bmj.j2471
6. Magill SS, Edwards JR, Bamberg W, et al. Multistate point-prevalence survey of health care-associated infections. *N Engl J Med*. 2014;370(13):1198–1208. doi:10.1056/NEJMoa1306801
7. Salisbury ML, Han MK, Dickson RP, Molyneaux PL. Microbiome in interstitial lung disease: from pathogenesis to treatment target. *Curr Opin Pulm Med*. 2017;23(5):404–410. doi:10.1097/MCP.0000000000000399
8. Travis WD, Costabel U, Hansell DM, et al. An official American Thoracic Society/European Respiratory Society statement: update of the international multidisciplinary classification of the idiopathic interstitial pneumonias. *Am J Respir Crit Care Med*. 2013;188(6):733–748. doi:10.1164/rccm.201308-1483ST
9. Dethlefsen L, McFall-Ngai M, Relman DA. An ecological and evolutionary perspective on human-microbe mutualism and disease. *Nature*. 2007;449(7164):811–818. doi:10.1038/nature06245
10. Mendez R, Banerjee S, Bhattacharya SK, Banerjee S. Lung inflammation and disease: a perspective on microbial homeostasis and metabolism. *IUBMB Life*. 2019;71(2):152–165. doi:10.1002/iub.1969
11. Tome D. Amino acid metabolism and signalling pathways: potential targets in the control of infection and immunity. *Eur J Clin Nutr*. 2021;75(9):1319–1327. doi:10.1038/s41430-021-00943-0
12. Woods DF, Flynn S, Caparros-Martin JA, Stick SM, Reen FJ, O'Gara F. Systems Biology and Bile Acid Signalling in Microbiome-Host Interactions in the Cystic Fibrosis Lung. *Antibiotics*. 2021;10(7):254. doi:10.3390/antibiotics10070766
13. Moffatt MF, Cookson WO. The lung microbiome in health and disease. *Clin Med (Lond)*. 2017;17(6):525–529. doi:10.7861/clinmedicine.17-6-525
14. Tunney MM, Einarsson GG, Wei L, et al. Lung microbiota and bacterial abundance in patients with bronchiectasis when clinically stable and during exacerbation. *Am J Respir Crit Care Med*. 2013;187(10):1118–1126. doi:10.1164/rccm.201210-1937OC
15. Pettigrew MM, Tanner W, Harris AD. The Lung Microbiome and Pneumonia. *J Infect Dis*. 2021;223(12 Suppl 2):S241–S245. doi:10.1093/infdis/jiaa702
16. Fischer A, Antoniou KM, Brown KK, et al. An official European Respiratory Society/American Thoracic Society research statement: interstitial pneumonia with autoimmune features. *Eur Respir J*. 2015;46(4):976–987. doi:10.1183/13993003.00150-2015
17. Rognes T, Flouri T, Nichols B, Quince C, Mahe F. VSEARCH: a versatile open source tool for metagenomics. *PeerJ*. 2016;4:e2584. doi:10.7717/peerj.2584
18. Quast C, Pruesse E, Yilmaz P, et al. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Res*. 2013;41(Database issue):D590–6. doi:10.1093/nar/gks1219
19. Cole JR, Wang Q, Fish JA, et al. Ribosomal Database Project: data and tools for high throughput rRNA analysis. *Nucleic Acids Res*. 2014;42(Databaseissue):D633. doi:10.1093/nar/gkt1244
20. DeSantis TZ, Hugenholtz P, Larsen N, et al. Greengenes, a chimera-checked 16S rRNA gene database and workbench compatible with ARB. *Appl Environ Microbiol*. 2006;72(7):5069–5072. doi:10.1128/AEM.03006-05
21. BugBase [homepage on the Internet]. BugBase predicts organism-level microbiome phenotypes. Available from: https://bugbase.cs.umn.edu/documentation.html. Accessed December 31, 2021.

22. Douglas GM, Maffei VJ, Zaneveld JR, et al. PICRUSt2 for prediction of metagenome functions. *Nat Biotechnol*. 2020;38(6):685–688. doi:10.1038/s41587-020-0548-6
23. Tang H, Wang X, Xu L, et al. Establishment of local searching methods for orbitrap-based high throughput metabolomics analysis. *Talanta*. 2016;156–157:163–171. doi:10.1016/j.talanta.2016.04.051
24. Bijlsma S, Bobeldijk I, Verheij ER, et al. Large-scale human metabolomics studies: a strategy for data (pre-) processing and validation. *Anal Chem*. 2006;78(2):567–574. doi:10.1021/ac051495j
25. Yang J, Zhao X, Lu X, Lin X, Xu G. A data preprocessing strategy for metabolomics to reduce the mask effect in data analysis. *Front Mol Biosci*. 2015;2:4. doi:10.3389/fmolb.2015.00004
26. Pang Z, Chong J, Li S, Xia J. MetaboAnalystR 3.0: toward an Optimized Workflow for Global Metabolomics. *Metabolites*. 2020;10(5):186. doi:10.3390/metabo10050186
27. Rohart F, Gautier B, Singh A, Le Cao KA. mixOmics: an R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol*. 2017;13(11):e1005752. doi:10.1371/journal.pcbi.1005752
28. Singh A, Shannon CP, Gautier B, et al. DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays. *Bioinformatics*. 2019;35(17):3055–3062. doi:10.1093/bioinformatics/bty1054
29. Bowerman KL, Rehman SF, Vaughan A, et al. Disease-associated gut microbiome and metabolome changes in patients with chronic obstructive pulmonary disease. *Nat Commun*. 2020;11(1):5886. doi:10.1038/s41467-020-19701-0
30. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: integrating viruses and cellular organisms. *Nucleic Acids Res*. 2021;49(D1):D545–D551. doi:10.1093/nar/gkaa970
31. Caspi R, Billington R, Keseler IM, et al. The MetaCyc database of metabolic pathways and enzymes - a 2019 update. *Nucleic Acids Res*. 2020;48(D1):D445–D453. doi:10.1093/nar/gkz862
32. O'Dwyer DN, Ashley SL, Gurczynski SJ, et al. Lung Microbiota Contribute to Pulmonary Inflammation and Disease Progression in Pulmonary Fibrosis. *Am J Respir Crit Care Med*. 2019;199(9):1127–1138. doi:10.1164/rccm.201809-1650OC
33. Dickson RP, Singer BH, Newstead MW, et al. Enrichment of the lung microbiome with gut bacteria in sepsis and the acute respiratory distress syndrome. *Nat Microbiol*. 2016;1(10):16113. doi:10.1038/nmicrobiol.2016.113
34. Duncan SH, Louis P, Thomson JM, Flint HJ. The role of pH in determining the species composition of the human colonic microbiota. *Environ Microbiol*. 2009;11(8):2112–2122. doi:10.1111/j.1462-2920.2009.01931.x
35. Lardner A. The effects of extracellular pH on immune function. *J Leukoc Biol*. 2001;69(4):522–530.
36. Bingula R, Filaire M, Radosevic-Robin N, et al. Desired Turbulence? Gut-Lung Axis, Immunity, and Lung Cancer. *J Oncol*. 2017;2017:5035371. doi:10.1155/2017/5035371
37. Zheng J, Wu Q, Zou Y, Wang M, He L, Guo S. Respiratory Microbiota Profiles Associated With the Progression From Airway Inflammation to Remodeling in Mice With OVA-Induced Asthma. *Front Microbiol*. 2021;12:723152. doi:10.3389/fmicb.2021.723152
38. Lehtinen MJ, Hibberd AA, Mannikko S, et al. Nasal microbiota clusters associate with inflammatory response, viral load, and symptom severity in experimental rhinovirus challenge. *Sci Rep*. 2018;8(1):11411. doi:10.1038/s41598-018-29793-w
39. Tuli JF, Ramezanpour M, Cooksley C, Psaltis AJ, Wormald PJ, Vreugde S. Association between mucosal barrier disruption by Pseudomonas aeruginosa exoproteins and asthma in patients with chronic rhinosinusitis. *Allergy*. 2021;76(11):3459–3469. doi:10.1111/all.14959
40. Huang Y, Mao K, Chen X, et al. S1P-dependent interorgan trafficking of group 2 innate lymphoid cells supports host defense. *Science*. 2018;359(6371):114–119. doi:10.1126/science.aam5809
41. Liu D, Yumoto H, Hirota K, et al. Histone-like DNA binding protein of Streptococcus intermedius induces the expression of pro-inflammatory cytokines in human monocytes via activation of ERK1/2 and JNK pathways. *Cell Microbiol*. 2008;10(1):262–276. doi:10.1111/j.1462-5822.2007.01040.x
42. Qu JM, Cao B. [Guidelines for the diagnosis and treatment of adult community acquired pneumonia in China (2016 Edition)]. *Zhonghua Jie He He Hu Xi Za Zhi*. 2016;39(4):241–242. doi:10.3760/cma.j.issn.1001-0939.2016.04.001. Chinese.
43. Olson JW, Hacker AD, Atkinson JE, Altiere RJ, Gillespie MN. Polyamine content in rat lung during development of hypoxia-induced pulmonary hypertension. *Biochem Pharmacol*. 1986;35(4):714–716. doi:10.1016/0006-2952(86)90372-2
44. Gao L, Zhang JH, Chen XX, et al. Combination of L-Arginine and L-Norvaline protects against pulmonary fibrosis progression induced by bleomycin in mice. *Biomed Pharmacother*. 2019;113:108768. doi:10.1016/j.biopha.2019.108768
45. Ren L, Wang Y, Zhong J, et al. Dynamics of the Upper Respiratory Tract Microbiota and its Association with Mortality in COVID-19. *Am J Respir Crit Care Med*. 2021;204(12):1379–1390. doi:10.1164/rccm.202103-0814OC
46. Wolak JE, Esther CR Jr, O'Connell TM. Metabolomic analysis of bronchoalveolar lavage fluid from cystic fibrosis patients. *Biomarkers*. 2009;14(1):55–60. doi:10.1080/13547500802688194
47. Muri J, Wolleb H, Broz P, Carreira EM, Kopf M. Electrophilic Nrf2 activators and itaconate inhibit inflammation at low dose and promote IL-1beta production and inflammatory apoptosis at high dose. *Redox Biol Sep*. 2020;36:101647. doi:10.1016/j.redox.2020.101647
48. Wang K, Huang Y, Zhang Z, et al. A Preliminary Study of Microbiota Diversity in Saliva and Bronchoalveolar Lavage Fluid from Patients with Primary Bronchogenic Carcinoma. *Med Sci Monit*. 2019;25:2819–2834. doi:10.12659/MSM.915332
49. Cho SY, Choi JH, Lee SH, Choi YS, Hwang SW, Kim YJ. Metataxonomic investigation of the microbial community in the trachea and oropharynx of healthy controls and diabetic patients using endotracheal tubes. *PLoS One*. 2021;16(11):e0259596. doi:10.1371/journal.pone.0259596
50. Callahan BJ, McMurdie PJ, Rosen MJ, Han AW, Johnson AJ, Holmes SP. DADA2: high-resolution sample inference from Illumina amplicon data. *Nat Methods*. 2016;13(7):581–583. doi:10.1038/nmeth.3869