

# Microbiome in Acute Exacerbation and Stable Phase of COPD: A Descriptive and Comparative Study of 16s rRNA Sequencing and Metagenomic Sequencing

**Kang Chang**

China-Japan Friendship Hospital <https://orcid.org/0000-0003-1214-0695>

**Xiaohui Zou**

China-Japan Friendship Hospital

**Zhengtu Li**

Guangzhou Medical College First Affiliated Hospital

**Zifeng Yang**

Guangzhou Medical College First Affiliated Hospital

**Jianpeng Gao**

Beijing Genskey Medical Technology Co.,Ltd

**Guoru Yang**

Weifang No.2 People's Hospital

**Yong Zhang**

Weifang No.2 People's Hospital

**Xue Li**

Weifang No.2 People's Hospital

**Xiaoqi Zhang**

Weifang No.2 People's Hospital

**Yingmei Liu**

China-Japan Friendship Hospital

**Bin Cao** (✉ [caobin\\_ben@163.com](mailto:caobin_ben@163.com))

China-Japan Friendship Hospital

---

## Research article

**Keywords:** Microbiome, 16s rRNA sequencing, metagenomic sequencing, COPD

**Posted Date:** November 11th, 2020

**DOI:** <https://doi.org/10.21203/rs.3.rs-103624/v1>

**License:**  This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

---

# Abstract

**Background:** The role of bacteria in different courses of chronic obstructive pulmonary disease (COPD) has not been very clarified. In this study, sputum samples from patients at different courses of COPD were collected to analyze the differences of the structure and function of respiratory microbiome in different courses of COPD.

**Results:** Our study involved 38 patients with acute exacerbation of COPD (AECOPD). Among them, 42 sputum samples were collected from the acute exacerbation of COPD (Of these 38 patients, 4 patients also collected sputum from the second acute exacerbation phase), and 12 sputum samples were collected from the stable phase of COPD (6 of the above 38 patients). Firmicutes, Proteobacteria, Bacteroidetes and Actinobacteria dominated the phylum in all the COPD samples, while Streptococcus, Neisseria, Haemophilus and Prevotella is the top 4 species in the cohort. Salmonella, Salmonella, Pseudomonas and Proteobacteria achieved higher abundance in AECOPD group. Samples from AECOPD patients showed higher  $\alpha$  diversity. Moreover, genes annotated to K07481(the function is related to transposase) and K16087(the function is related to hemoglobin/transferrin/lactoferrin receptor protein) in stable subgroup showed higher abundance than that in acute exacerbation subgroup. Compared to samples from stable COPD, Proteobacteria contributed the most antibiotic resistance genes.

**Conclusions:** There are differences in the function and metabolism of respiratory microbiome in patients with different course of COPD.

**Word count:** 218/350

## Background

Chronic obstructive pulmonary disease (COPD) is the third leading cause of death in the world, causing about 32 million deaths worldwide in 2017 with the expectation to reach about 44 million by 2040[1, 2]. The incidence of COPD is about 10.1% worldwide, and the number of premature deaths caused by COPD increased by 13.2% from 2007 to 2017[2]. The latest epidemiological study in China shows that the prevalence rate of COPD in China increases significantly with age, with 21.2% between 60 and 69 years old, 35.5% in people over 70 years old[3]. COPD ranks third among the 15 diseases with the highest disease burden in the elderly[4]. In 2017, the case fatality rate of COPD in China was 68/100,000, resulting in 945,000 deaths, which was also the third cause of death[4].

The main feature of acute exacerbations of chronic obstructive pulmonary disease (AECOPD) is that the respiratory symptoms of patients deteriorate and need to clinical treatment [5]. Since bacteria can be cultured in respiratory samples from AECOPD patients, it has always been believed that bacteria maybe an important cause of AECOPD, such as bacterial colonization[6–9]. Studies based on culture techniques usually focus on several known species with respiratory pathogenicity, such as Haemophilus influenzae, Moraxella catarrhalis, Streptococcus pneumoniae, or Pseudomonas aeruginosa (symptoms tend to be more severe). Colonization of these or other types of bacteria can also be found in patients with stable

COPD. These phenomena suggest that the role of bacteria in the deterioration of COPD may be more than infection[10].

The development of Next Generation Sequencing technology provides an opportunity for in-depth study of lung microbiome in healthy people and patients. Some studies based on 16 s rRNA sequencing have found that COPD patients developed unique bacterial community compared with healthy subjects, suggesting that changes in lung microbiome may be related to increased airway inflammation and disease progression[11–13]. Shotgun metagenomic sequencing can be used to sequence microorganisms in samples at the whole genome level, which is more comprehensive and achieve better classification and analysis than 16S rRNA sequencing. Assembling shorter sequencing fragments into larger fragments can increase the ability to distinguish between different microorganisms, improve the resolution of classification, and reach the level of species or strains[14, 15]. In addition, by allocating microbial genes in the metabolic pathway, we have the ability to assess the potential metabolic function of microorganisms and interactions between microorganisms and between microorganisms and humans[16, 17]. This study used both 16 s rRNA sequencing and metagenomics to analyze the sputum samples of AECOPD patients and analyze the role of respiratory microbes in AECOPD.

## Results

### Lung microbiome description, gene prediction and species annotation

Of the 54 sputum samples, 38 were collected from the first acute exacerbation (period A). There were 12 stable period sputum samples (period B) from 6 patients (7 sputum samples were collected in the second month of the stable phase, 4 sputum samples were collected in the 6th month of the stable phase, and 1 sputum was collected in the 12th month of the stable phase). There were 4 sputum samples (period C) from 4 patients during the second acute exacerbation (2 samples were collected on the first day of admission and 2 samples were collected on the third day of admission) (Table 1).

We obtained 310,383 ORFs (Open Reading Frames), with a total length of 173.46 Mbp, GC content of 44.78%, from which 115,179 were assigned to genes. In the ORFs that can be annotated to the NR database (Version: 2018.01) from NCBI, 94.28% were assigned at the kingdom level, 91.89% at the phylum level, 90.35% at the class level, 87.81% at the order level, 85.51% at the family level, 82.71% at the genus level, and 47.85% at the species level. Additional file: Figure S1. Firmicutes, Proteobacteria, Bacteroidetes and Actinobacteria were the top 4 phylums. Streptococcus, Neisseria, Haemophilus and Prevotella were top 4 genus in the cohort, in which many members in these four genus were natural inhabitants in the human up respiratory tract(Figure 1)[18].

### Difference of microbiome between stable and acute exacerbation stage of COPD

Our results show that the sputum microbiome composition of patients with acute exacerbation of COPD is different from that of patients at stable stage. At the phylum level, the relative abundances of Firmicutes and Proteobacteria in stable patients were 15.80% and 23.82%, respectively, while those in

patients with acute exacerbation were 13.50% and 13.05% respectively, with no statistical difference between the two stages. At the genus level, the relative abundance of *Haemophilus* in stable patients was 10.32%, while that in patients with AECOPD was 2.71% (Figure 2, Additional file: Table S1, Additional file: Table S2), with no statistical difference between the two. The results of LEfSe analysis showed that in period A group, *Salmonella*, *Salmonella enterica*, *Pseudomonas* and *Proteobacteria* achieved high abundance. While in period B, *Haemophilus influenzae*, *Pasteurellales* were more abundant (Figure S2). In addition, period A cohort showed higher a diversity than period B group (Figure 3).

### **Analysis of gene function and resistance genes**

In order to further evaluate the function of genes, genes obtained by sequencing were annotated with the following functional database (KEGG, eggnog and CAZy database) by DIAMOND software. There are 201,130 (64.80%) genes that can match the KEGG database, of which 108,824 (35.06%) can match 4,562 KEGG ortholog group (KO) 193014 (62.19%) genes can be assigned to the eggNOG database, 8984 (2.89%) genes can be matched the CAZy database. (Figure S1). We found that the relative abundance of genes annotated to K07481 and K16087 in period B subgroup was higher than that in period A or C subgroup. The relative abundance of the former in period B subgroup was 0.119%, and that in period A and C subgroup was 0.07% and 0.017%, respectively. The relative abundance of the latter in period B subgroup was 0.135%, while that in period A and C subgroups was 0.046% and 0.006%, respectively (Figure S3, Additional file: Table S3). The results of resistance gene annotation showed that 275 genes could be found in the CARD database, including 198 kinds of ARO (the Antibiotic Resistance Ontology) Compared with patients with stable COPD, *Proteobacteria* harbored more ARO genes (44%) in patients with acute exacerbation, suggesting that drug resistance of *Proteobacteria* may play an important role in the process of acute exacerbation of COPD. After assigning ARO to species, in subgroup period A, 44% of ARO belonged to *Proteobacteria*, 13% to *Firmicutes*, 4% to *Actinobacteria*, and 4% to *Bacteroidetes*. In period B subgroup, 44% of ARO belonged to *Proteobacteria*, 13% to *Firmicutes*, 3% to *Actinobacteria*, and 3% to *Bacteroidetes* (Fig. 4).

### **Metabolic pathway analysis**

In this study, there was no difference in the number of genes in different metabolic pathways between period A and period B groups. However, it was found that the relative abundance of genes annotated to the metabolic pathway of purine metabolism (ko00230), ATP binding- cassette (ABC) transporters (ko02010), two-component signal transduction system (ko02020), and ribosome (ko03010) function, the homologous recombination function (ko03440) are higher in period B group than that in period A group (figure 5).

## **Discussion**

The microbiome of respiratory plays an important role in the development of COPD and AECOPD. We found that the relative abundance of *Proteobacteria*, in sputum microbiome of patients with AECOPD decreased significantly compared with patients in stable stage, while the relative abundance of

Firmicutes did not change significantly [19]. In addition, we noted that in stable patients, the relative abundance of *Haemophilus influenzae* was 4.38%, while in patients with AECOPD, the number was only 0.67%.

The function of K07481 is mainly related to transposition. A transposon is a mobile element in the genome that can transfer itself from one location of the genome to another, which is the main source of mutation in the genome. The function of K16087 is associated with hemoglobin, transferrin and lactoferrin receptor proteins. It is suggested that the function of K16087 is related to the iron uptake of bacteria in the host. Iron is involved in many important biological activities, including respiration, tricarboxylic acid cycle, gene regulation, DNA biosynthesis and so on. Microbes in the host tend to obtain iron from transferrin, lactoferrin or hemoglobin[20]. The results of this study showed that the relative abundance of genes annotated to K07481 and K16087 increased significantly in the period B subgroup, suggesting that there was a difference in the interaction between airway microorganisms and hosts between patients with stable COPD and patients with acute exacerbation of COPD.

ATP binding cassette transporter is one of the most active intracellular transport systems, which exists widely in bacteria, archaea and eukaryotes. Its main role is to combine ATP hydrolysis with the active transport of a variety of substrates (such as ions, sugars, lipids, proteins and drugs, etc.), and is related to many important biological functions of prokaryotes and eukaryotes[21]. The two-component signal transduction system enables bacteria to sense, respond to, and adapt to changes in the environment or intracellular state[22]. Homologous recombination is a process of gene exchange between homologous DNA sequences and leads to gene conversion or crossover. It is important for the repair of DNA damage, such as single-strand gap, double-strand break (DSB) and so on, and helps to maintain the integrity of the genome[23]. The relative abundance of genes annotated to the above functions decreased in patients with acute exacerbation of COPD, which may reflect the result of complex interaction between airway microbial community and airway environment.

Our research also has some shortcomings. First of all, the sample size of this study is small, including only 38 samples from AECOPD patients; second, this study collects sputum collected by patients through spontaneous expectoration, not directly from the lower respiratory tract; in fact, there has been some debate about the similarities and differences between sputum and other invasive lower respiratory tract samples, such as BAL or bronchial protective brushes. Some studies have found that compared with the samples obtained by bronchial biopsy, there is an enrichment of oral flora in sputum samples. however, further analysis found that these samples are very similar in terms of the abundance of oral flora[24]. Therefore, we believe that although there are differences in microbial composition between samples from the upper respiratory tract and the lower respiratory tract, sputum does reflect the microbial composition of the lower respiratory tract to some extent, especially in terms of abundance[25].

## Conclusions

In conclusion, we showed the taxonomy and functional changes of respiratory microbiome in patients with acute exacerbation and stable stage of COPD, which partly revealed the effects of microbial interaction on host respiratory homeostasis and pathogenicity, and further deepened the understanding of the relationship between respiratory tract microbiome and COPD.

## Methods

### Subjects and samples

A total of 54 sputum samples were collected from 38 patients with AECOPD. Hospitalized patients for AECOPD based on the 2016 GOLD guideline were involved in this study. All patients provide informed consent. The sputum was collected by the method of spontaneous expectoration during the hospitalization. Details for sputum collection and processing, shotgun metagenome sequencing, sequence assembly and gene annotation, see supplementary.

### Bioinformatics analysis

In the aspect of 16s rRNA sequencing, barcodes and sequencing primers were trimmed before merged. Paired-end reads were assembled using FLASH[26], and quality-trimmed using QIIME[27] with default settings. These trimmed sequences were then chimera filtered, singletons discarded, and clustered into operational taxonomic units (OTUs) using a 97% similarity threshold with UPARSE[28]. We used mothur to annotate taxonomic information for the representative sequence of each OTU by comparison to the SILVA\_128[29] reference database.

In terms of metagenomic, all samples were paired-end sequenced on the Illumina platform (insert size 350 bp, read length 150 bp). After quality control, the reads aligned to the human genome with SOAP2[30] were removed. The remaining high-quality reads were used for further analysis. The assembly of reads were executed using SOAPdenovo[31]. Unused reads from each sample were assembled using the same parameters. Genes (minimum length of 100 nucleotides) were predicted on scaffolds (i.e., continuous sequences within scaffolds) longer than 500 bp using MetaGeneMark (prokaryotic GeneMark.hmm version2.10). Then, a non-redundant gene catalogue was constructed with CD-HIT[32] using a sequence identity cut-off of 0.95, with a minimum coverage cut-off of 0.9 for the shorter sequences. To determine the abundance of genes, reads were realigned to the gene catalogue with SOAP2[31] using parameters: -m 200 -x 400 -s 119. Only genes with  $\geq 2$  mapped reads were deemed to be present in a sample. The abundance of genes was calculated by counting the number of reads and normalizing by gene length. The constructed nonredundant gene catalogue was aligned to functional database include KEGG database (<http://www.kegg.jp/kegg/>), eggNOG database (<http://eggnogdb.embl.de/#/app/home>) CAZy database (<http://www.cazy.org/>) by DIAMOND[33]. Each protein was assigned to the database by the highest scoring annotated hit(s) containing at least one HSP scoring over 60 bits. Statistics of the relative abundance of different functional hierarchy calculated by summing relative abundance annotated to the same feature.

# Abbreviations

COPD: chronic obstructive pulmonary disease;

AECOPD: acute exacerbation of chronic obstructive pulmonary disease;

ORFs: Open Reading Frames;

OTU: Operating taxonomic unit;

KEGG: Kyoto Encyclopedia of Genes and Genomes;

eggNOG: Evolutionary genealogy of genes: Non-supervised Orthologous Groups;

CAZy: Carbohydrate-Active enzymes Database;

CARD: the Comprehensive Antibiotic Research Database;

ARO: the Antibiotic Resistance Ontology.

# Declarations

## Competing interests

The authors declare that we have no competing interests.

## Acknowledgements

The authors thank for Xianxia Zhuo, Yeming Wang, Yimin Wang, Haibo Li, Jiankang Zhao, Binghuai Lu for helpful discussions and technical supports.

## Funding

Funding provided by the National Key Research and Development Program of China (2017YFC1309301, 2017YFC1309302), CAMS Innovation Fund for Medical Sciences (CIFMS 2018-I2M-1-003).

## Authors' contributions

Conception and design: XHZ. Acquisition of data: ZTL, ZFY, KC, GUY, XL, XQZ, YZ. Analysis and interpretation of data: KC, LFL, XHZ. Drafting or revising of manuscript: KC, LFL. Final approval of manuscript: XHZ, MEL, BC.

## Availability of data and materials

The data set supporting the results of this article has been deposited in the Sequence Read Archive (SRA) under BioProject accession code PRJNA667797 [<https://www.ncbi.nlm.nih.gov/sra/PRJNA667797>].

## Ethics approval and consent to participate

This study was approved by the Research Ethics Committee of China-Japan Friendship Hospital, and all subjects signed informed consent forms.

## Consent for publication

All authors read and approved the manuscript.

## References

1. Rabe, K.F. and H. Watz, Chronic obstructive pulmonary disease. *LANCET*, 2017. 389(10082): p. 1931-1940.
2. Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet (London, England)*, 2018. 392(10159): p. 1736-1788.
3. Wang, C., et al., Prevalence and risk factors of chronic obstructive pulmonary disease in China (the China Pulmonary Health [CPH] study): a national cross-sectional study. *Lancet*, 2018. 391(10131): p. 1706-1717.
4. M, Z., et al., Mortality, morbidity, and risk factors in China and its provinces, 1990-2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet (London, England)*, 2019. 394(10204): p. 1145-1158.
5. Vogelmeier, C.F., et al., Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. *AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE*, 2017. 195(5): p. 557-582.
6. MONSO, E., et al., BACTERIAL-INFECTION IN CHRONIC OBSTRUCTIVE PULMONARY-DISEASE - A STUDY OF STABLE AND EXACERBATED OUTPATIENTS USING THE PROTECTED SPECIMEN BRUSH. *AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE*, 1995. 152(4): p. 1316-1320.
7. Soler, N., et al., Bronchial microbial patterns in severe exacerbations of chronic obstructive pulmonary disease (COPD) requiring mechanical ventilation. *AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE*, 1998. 157(5): p. 1498-1505.
8. BALL, P., EPIDEMIOLOGY AND TREATMENT OF CHRONIC-BRONCHITIS AND ITS EXACERBATIONS. *CHEST*, 1995. 108S(2): p. S43-S52.
9. Miravittles, M., et al., Relationship between bacterial flora in sputum and functional impairment in patients with acute exacerbations of COPD. *CHEST*, 1999. 116(1): p. 40-46.
10. Huang, Y.J., et al., Understanding the role of the microbiome in chronic obstructive pulmonary disease: principles, challenges, and future directions. *Transl Res*, 2017. 179: p. 71-83.

11. Huang, Y.J., et al., Airway Microbiome Dynamics in Exacerbations of Chronic Obstructive Pulmonary Disease. *JOURNAL OF CLINICAL MICROBIOLOGY*, 2014. 52(8): p. 2813-2823.
12. Zakharkina, T., et al., Analysis of the Airway Microbiota of Healthy Individuals and Patients with Chronic Obstructive Pulmonary Disease by T-RFLP and Clone Sequencing. *PLOS ONE*, 2013. 8(e683027).
13. Molyneaux, P.L., et al., Outgrowth of the Bacterial Airway Microbiome after Rhinovirus Exacerbation of Chronic Obstructive Pulmonary Disease. *AMERICAN JOURNAL OF RESPIRATORY AND CRITICAL CARE MEDICINE*, 2013. 188(10): p. 1224-1231.
14. Jovel, J., et al., Characterization of the Gut Microbiome Using 16S or Shotgun Metagenomics. *Front Microbiol*, 2016. 7: p. 459.
15. Scholz, M., et al., Strain-level microbial epidemiology and population genomics from shotgun metagenomics. *Nat Methods*, 2016. 13(5): p. 435-8.
16. Joice, R., et al., Determining microbial products and identifying molecular targets in the human microbiome. *Cell Metab*, 2014. 20(5): p. 731-741.
17. Quince, C., et al., Shotgun metagenomics, from sampling to analysis. *Nat Biotechnol*, 2017. 35(9): p. 833-844.
18. Bathorn, E., et al., Airways inflammation and treatment during acute exacerbations of COPD. *International journal of chronic obstructive pulmonary disease*, 2008. 3(2): p. 217-29.
19. Z, W., et al., Lung microbiome dynamics in COPD exacerbations. *The European respiratory journal*, 2016. 47(4): p. 1082-92.
20. KD, K. and V. HJ, Structural biology of bacterial iron uptake. *Biochimica et biophysica acta*, 2008. 1778(9): p. 1781-804.
21. CF, H., ABC transporters: from microorganisms to man. *Annual review of cell biology*, 1992. 8: p. 67-113.
22. JA, H., Two-component and phosphorelay signal transduction. *Current opinion in microbiology*, 2000. 3(2): p. 165-70.
23. L, M., F. F and G. S, DNA polymerase delta is preferentially recruited during homologous recombination to promote heteroduplex DNA extension. *Molecular and cellular biology*, 2008. 28(4): p. 1373-82.
24. J, D., et al., Bacterial biogeography of adult airways in atopic asthma. *Microbiome*, 2018. 6(1): p. 104.
25. Ditz, B., et al., Sputum microbiome profiling in COPD: beyond singular pathogen detection. *Thorax*, 2020. 75(4): p. 338-344.
26. T, M. and S. SL, FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinformatics (Oxford, England)*, 2011. 27(21): p. 2957-63.
27. J, K., et al., Using QIIME to analyze 16S rRNA gene sequences from microbial communities. *Current protocols in bioinformatics*, 2011: p. Unit 10.7.

28. RC, E., UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nature methods*, 2013. 10(10): p. 996-8.
29. P, Y., et al., The SILVA and "All-species Living Tree Project (LTP)" taxonomic frameworks. *Nucleic acids research*, 2014. 42: p. D643-8.
30. R, L., et al., SOAP2: an improved ultrafast tool for short read alignment. *Bioinformatics (Oxford, England)*, 2009. 25(15): p. 1966-7.
31. R, L., et al., SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *GigaScience*, 2012. 1(1): p. 18.
32. W, L. and G. A, Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics (Oxford, England)*, 2006. 22(13): p. 1658-9.
33. B, B., X. C and H. DH, Fast and sensitive protein alignment using DIAMOND. *Nature methods*, 2015. 12(1): p. 59-60.
34. Singh, D., et al., Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease: the GOLD science committee report 2019. *Eur Respir J*, 2019. 53(5).

## Tables

Table 1. Patients and samples information table

Sample number	Patient number	Sex	Age	COPD severity classification	
				GOLD classification	ABCD Assessment system
S3002AD1	S002	M	72	NA	NA
S3003AD3	S003	M	58	NA	D
S3004AD1	S004	M	80	NA	D
S3005AD1	S005	M	61	NA	NA
S3009AD1	S009	M	63	NA	D
S3009BM2	S009	M	63	NA	C
S3009BM6	S009	M	63	NA	C
S3010AD1	S010	M	82	GOLD3	D
S3010BM2	S010	M	82	GOLD3	D
S3010BM6	S010	M	82	NA	NA
S3011AD1	S011	M	74	NA	NA
S3012AD1	S012	M	67	NA	D
S3013AD1	S013	M	76	NA	B
S3013BM12	S013	M	76	NA	B
S3013BM2a	S013	M	76	GOLD3	B
S3013BM2b	S013	M	76	GOLD3	B
S3013BM6	S013	M	76	NA	A
S3016AD1	S016	M	68	GOLD3	D
S3016BM2	S016	M	68	GOLD3	D
S3016BM6	S016	M	68	GOLD3	D
S3016CD3	S016(Second hospitalization)	M	68	NA	D
S3018AD1	S018	M	66	GOLD3	D
S3021BM2	S021	NA	NA	NA	NA
S3021AD1	S021	M	81	GOLD3	D
S3022AD1	S022	M	82	NA	NA
S3023AD1	S023	M	80	NA	D

S3024AD1	S024	M	79	GOLD3	D
S3024BM2	S024	M	79	GOLD3	D
S3026AD5	S026	M	69	NA	NA
S3027AD5	S027	F	79	GOLD3	D
S3027CD3	S027(Second hospitalization)	F	79	NA	D
S3030AD3	S030	M	67	NA	D
S3030CD1	S030(Second hospitalization)	M	67	NA	D
S3031AD1	S031	M	67	NA	C
S3032AD1	S032	M	67	GOLD4	D
S3033AD1	S033	M	84	NA	D
S3033CD1	S033(Second hospitalization)	M	84	NA	D
S3034AD1	S034	M	79	NA	NA
S3039AD1	S039	M	77	NA	D
S3042AD1	S042	M	73	NA	D
S3044AD1	S044	M	80	GOLD3	D
S3045AD1	S045	M	68	GOLD3	D
S3046AD1	S046	M	84	NA	D
S3047AD1	S047	M	60	NA	D
S3048AD1	S048	M	61	NA	B
S3049AD1	S049	M	73	NA	D
S3050AD5	S050	M	70	GOLD3	B
S3052AD3	S052	M	82	GOLD3	D
S3053AD1	S053	M	62	NA	D
S3055AD1	S055	M	66	NA	C
S3056AD1	S056	M	83	NA	D
S3058AD1	S058	M	64	GOLD3	B
S3060AD1	S060	F	77	NA	D
S3061AD1	S061	M	75	NA	D

Description: NA means data missing.

GOLD classification: GOLD1:FEV<sub>1</sub>(% predicted)≥80%, GOLD2: 50%≤FEV<sub>1</sub>(% predicted)≤79%, GOLD3: 30%≤FEV<sub>1</sub>(% predicted)≤49%, GOLD4: FEV<sub>1</sub>(% predicted) <30%. FEV<sub>1</sub>: Forced expiratory volume in one second[34].

ABCD Assessment system: A: mMRC 0-1 CAT<10 and moderate or severe exacerbation history is 0 or 1(not leading to hospital admission), B: mMRC ≥2, CAT≥10 and moderate or severe exacerbation history is 0 or 1(not leading to hospital admission), C: mMRC 0-1 CAT<10 and the number of history of moderate or severe exacerbation≥2 or one with hospital admission, D: mMRC ≥2 CAT≥10 and the number of history of moderate or severe exacerbation≥2 or one with hospital admission [34].

## Figures

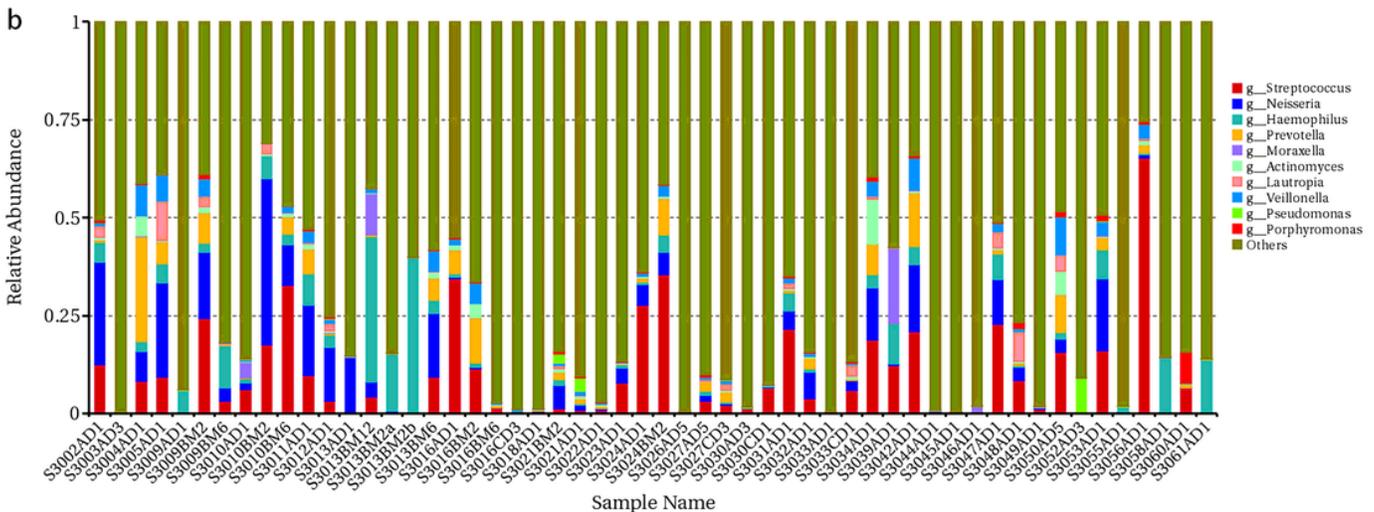
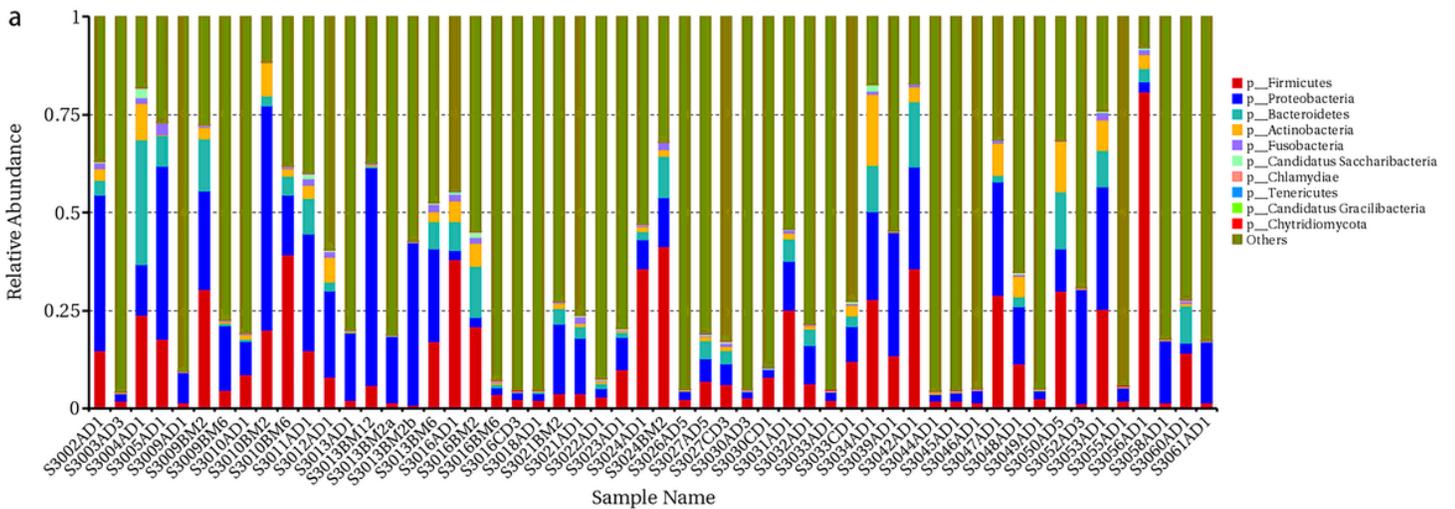


Figure 1

Histogram of relative abundance of the top ten species of phylum(a) and genus(b).

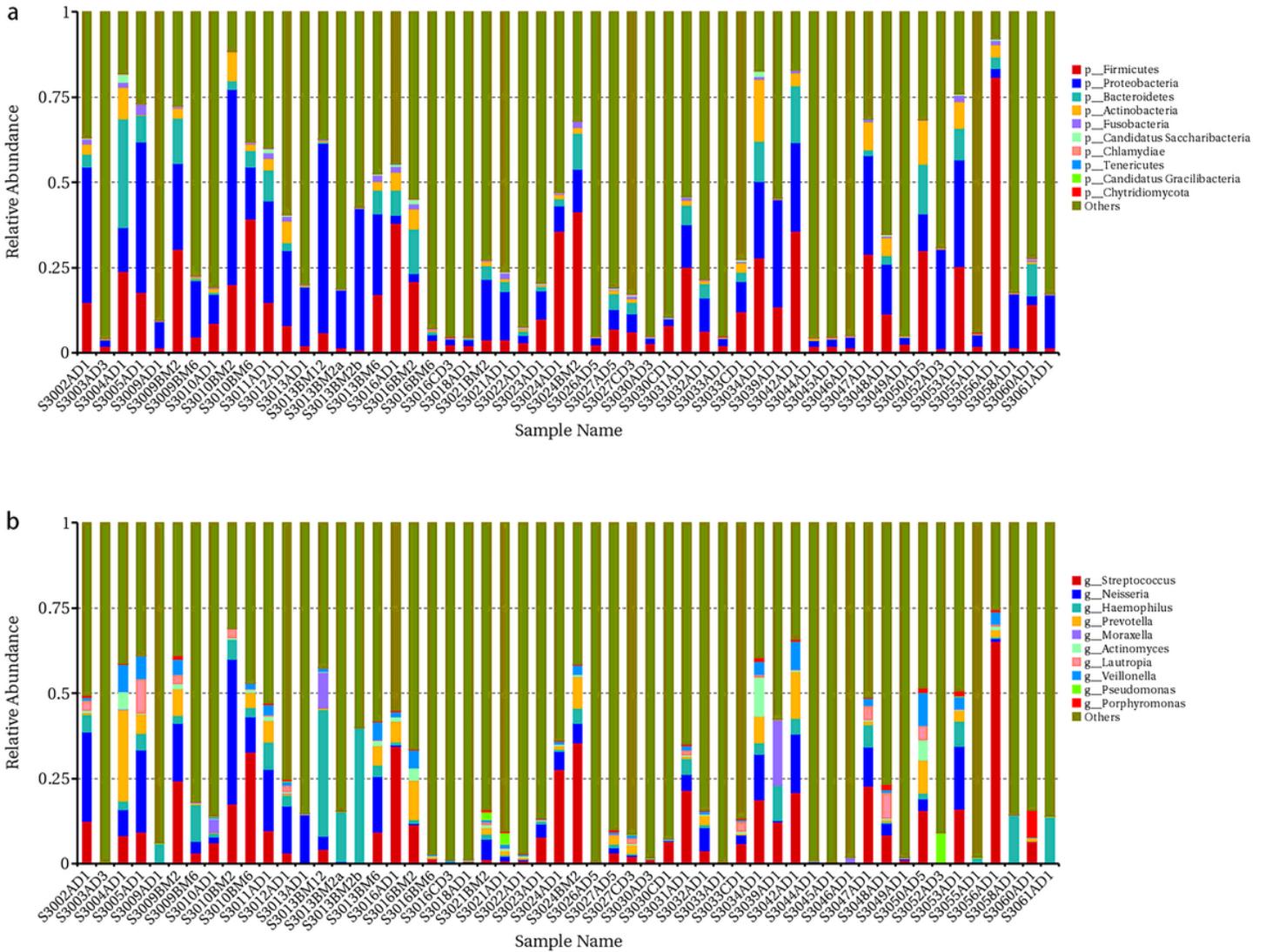
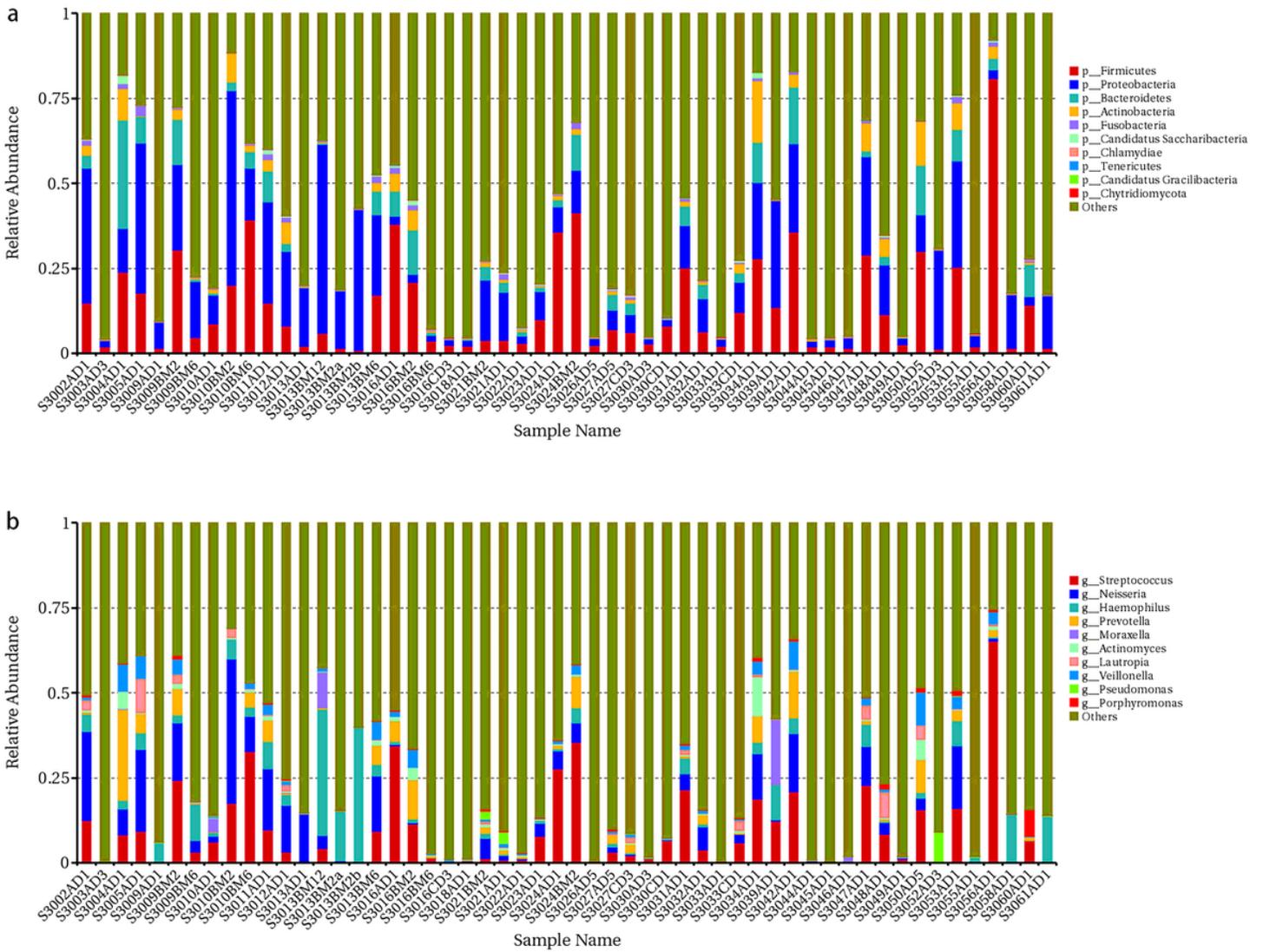


Figure 1

Histogram of relative abundance of the top ten species of phylum(a) and genus(b).



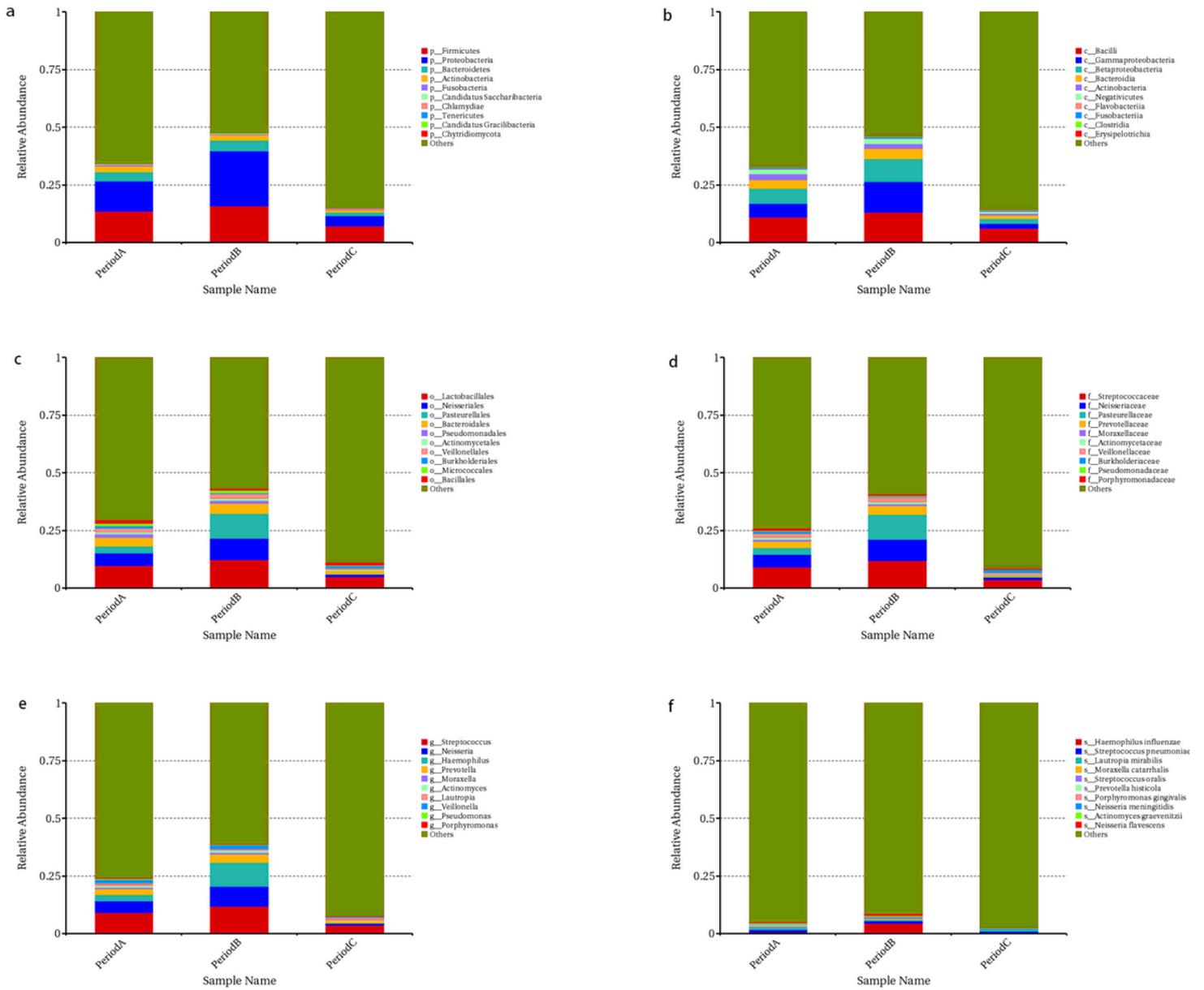
**Figure 1**

Histogram of relative abundance of the top ten species of phylum(a) and genus(b).



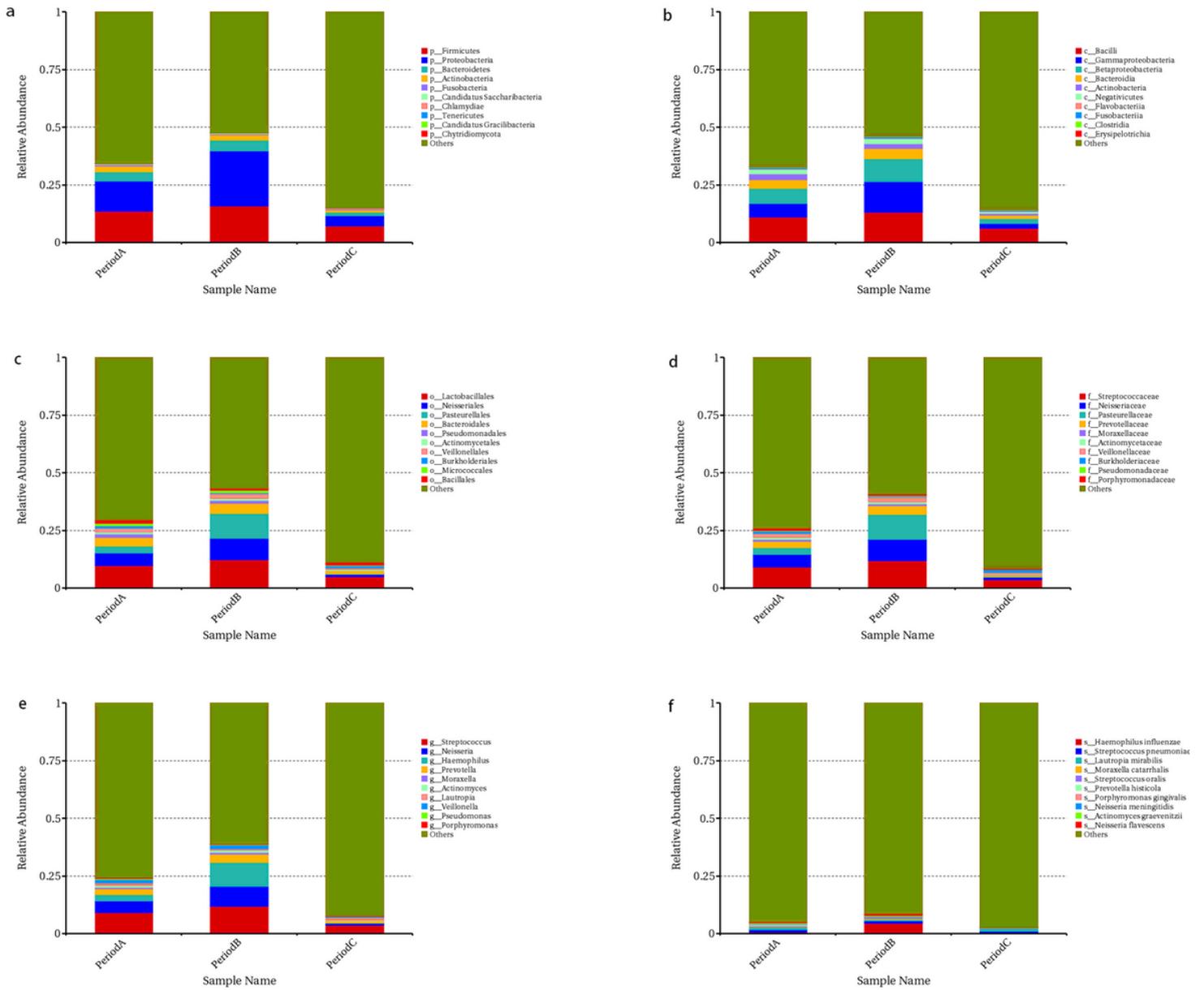
**Figure 2**

Histogram of relative abundance of species with different classification levels (a: phylum, b: class, c: order, d: family, e: genus, f: species).



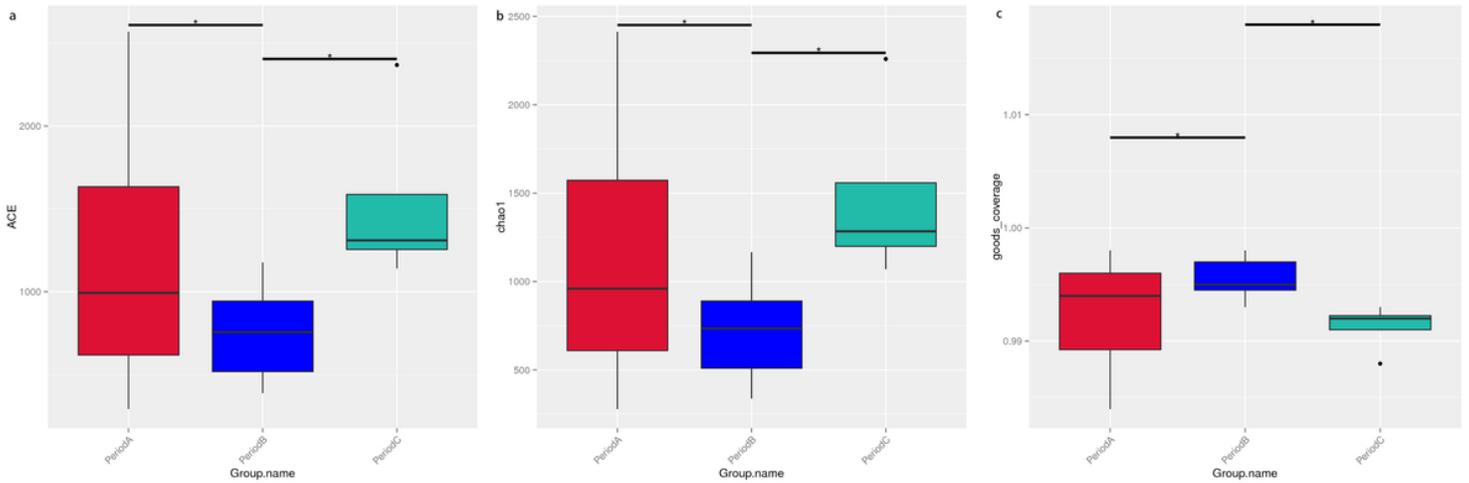
**Figure 2**

Histogram of relative abundance of species with different classification levels (a: phylum, b: class, c: order, d: family, e: genus, f: species).



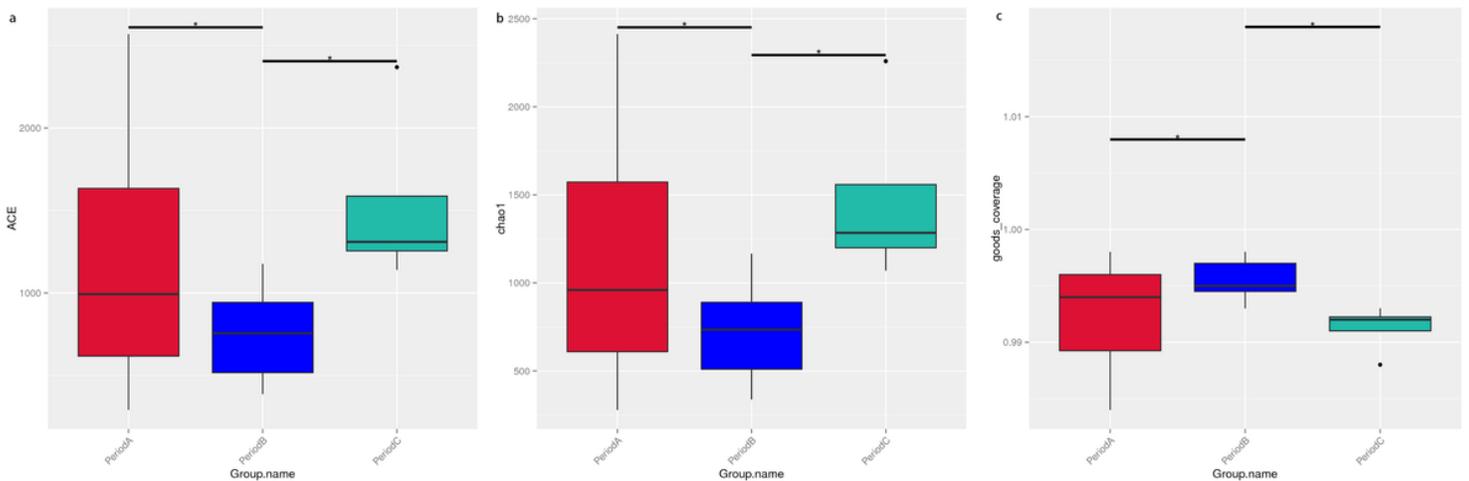
**Figure 2**

Histogram of relative abundance of species with different classification levels (a: phylum, b: class, c: order, d: family, e: genus, f: species).



**Figure 3**

Box chart for analysis of differences between groups of a diversity index. Description: The ACE estimator (figure 3a) and the Chao1 estimator (figure 3b) are used to evaluate community richness of microbiome, the goods coverage estimator is used to evaluate sequencing depth (figure 3c).



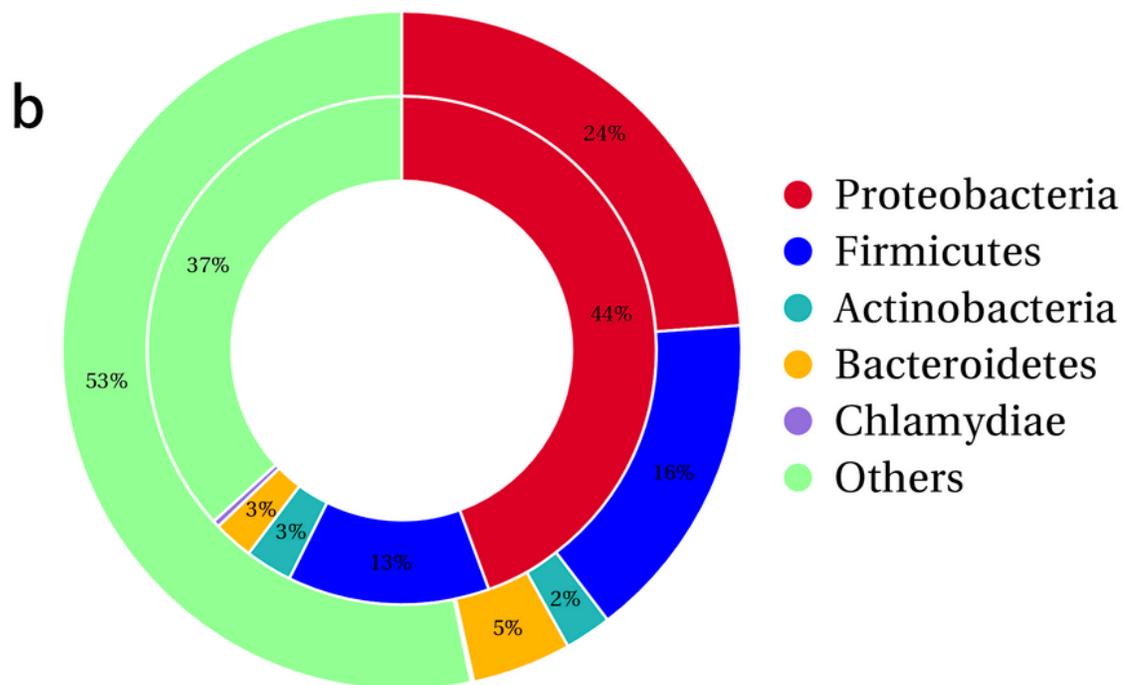
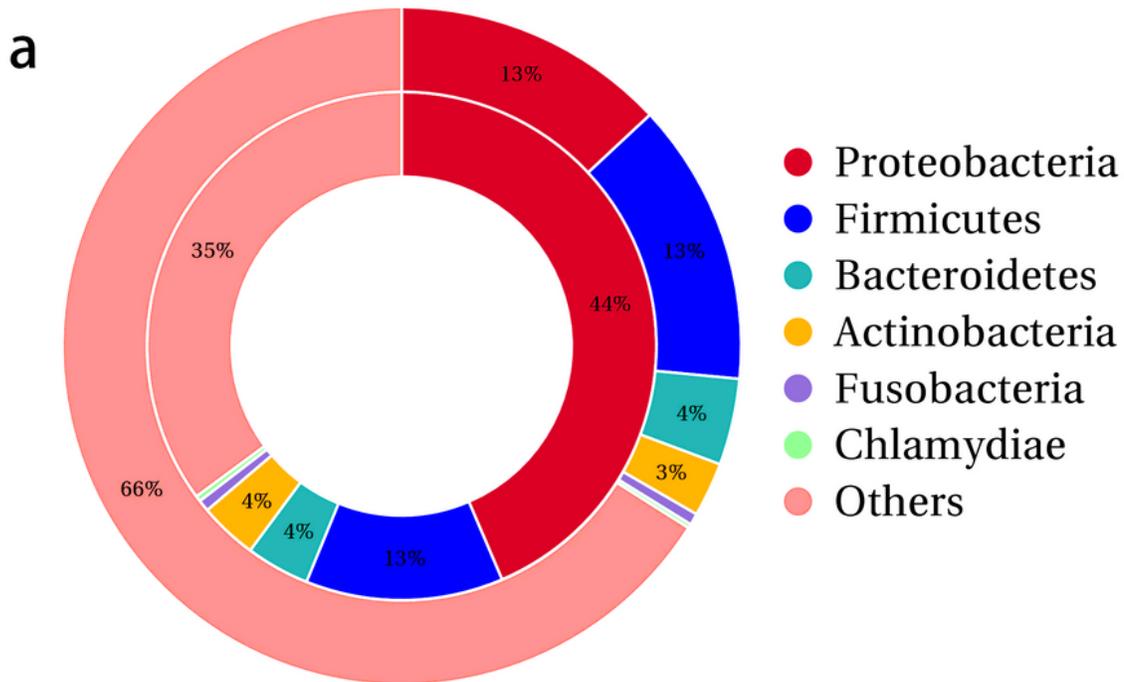
**Figure 3**

Box chart for analysis of differences between groups of a diversity index. Description: The ACE estimator (figure 3a) and the Chao1 estimator (figure 3b) are used to evaluate community richness of microbiome, the goods coverage estimator is used to evaluate sequencing depth (figure 3c).



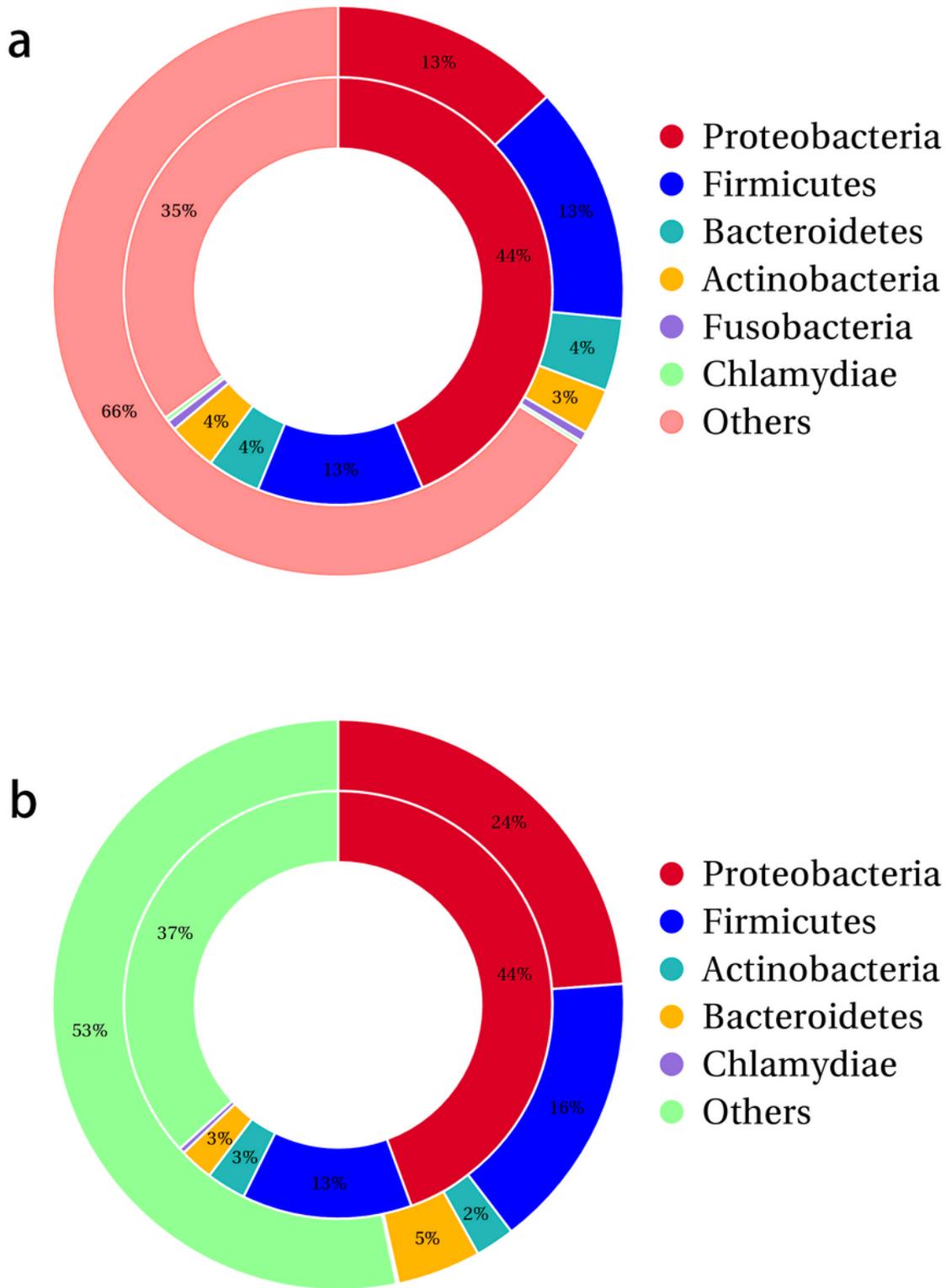
**Figure 3**

Box chart for analysis of differences between groups of a diversity index. Description: The ACE estimator (figure 3a) and the Chao1 estimator (figure 3b) are used to evaluate community richness of microbiome, the goods coverage estimator is used to evaluate sequencing depth (figure 3c).



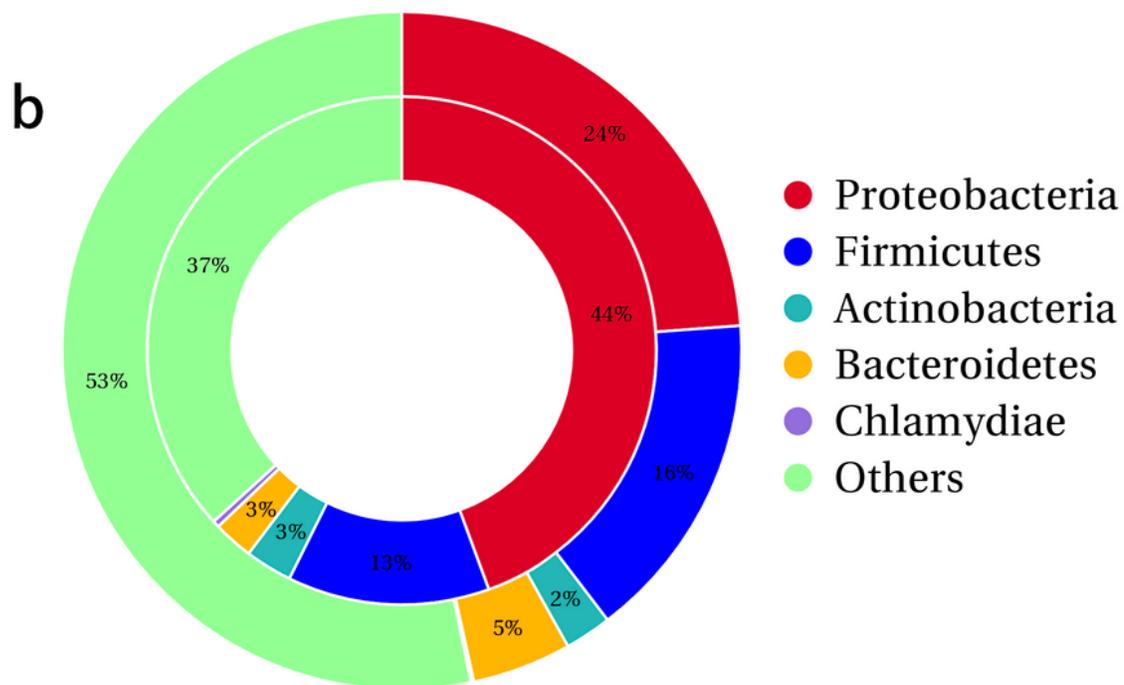
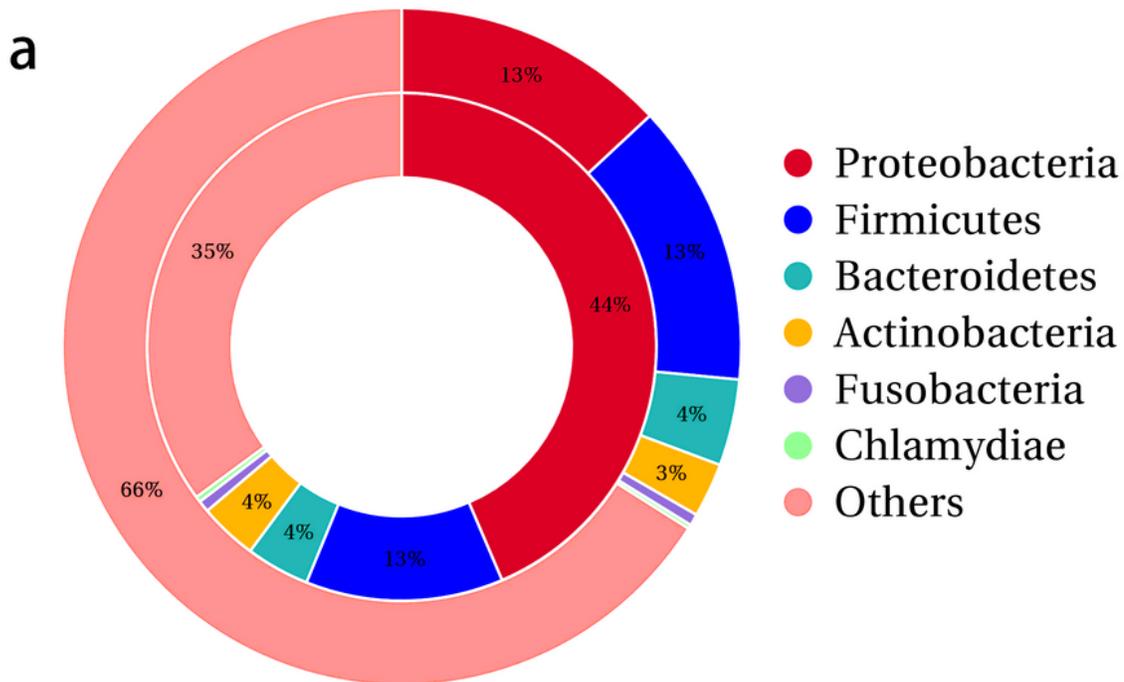
**Figure 4**

Circle diagram of species attribution analysis of resistance genes. Description: The inner circle is the species distribution of ARO(Antibiotic Resistance Ontology), and the outer circle is the species distribution of all sample genes in the group.



**Figure 4**

Circle diagram of species attribution analysis of resistance genes. Description: The inner circle is the species distribution of ARO(Antibiotic Resistance Ontology), and the outer circle is the species distribution of all sample genes in the group.



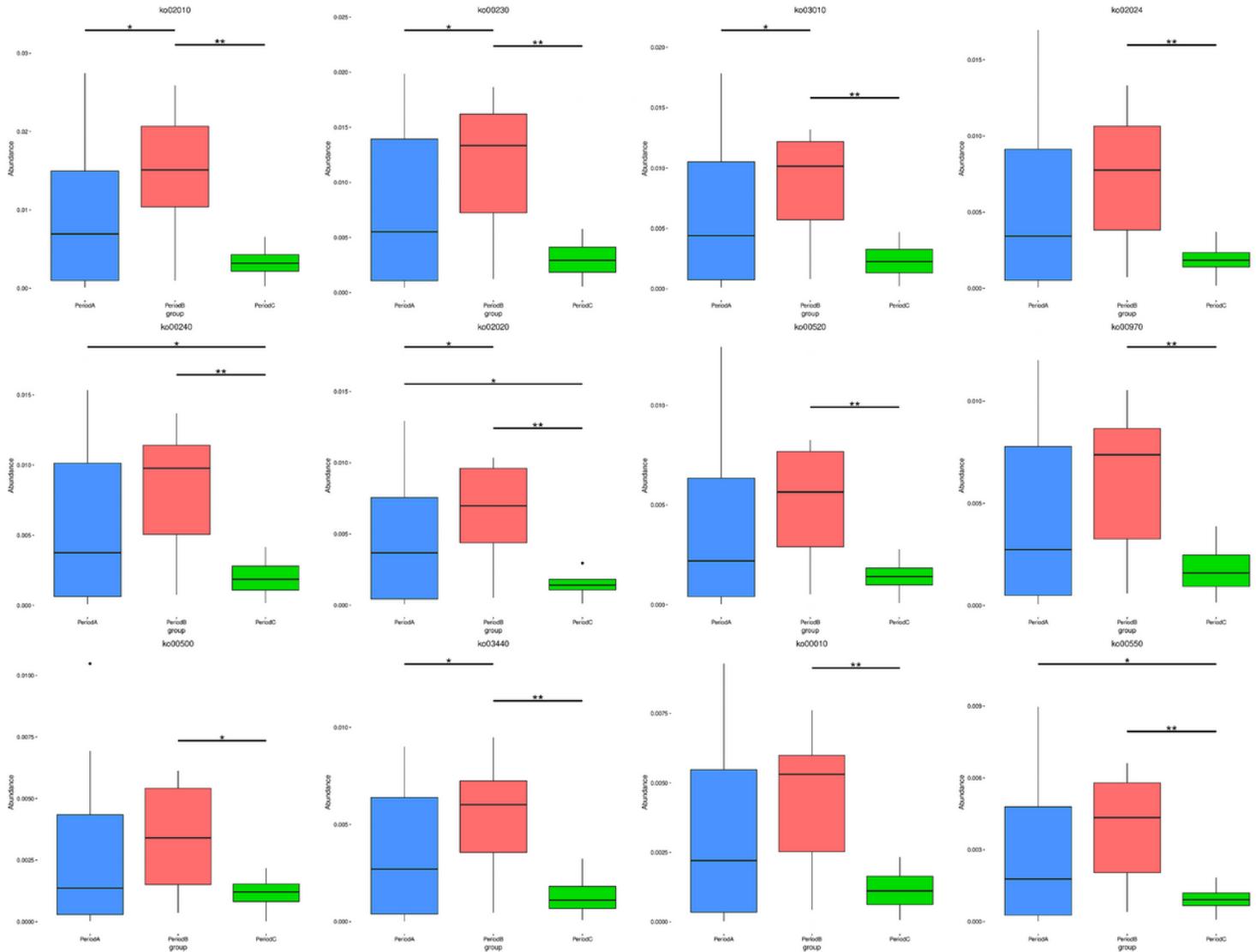
**Figure 4**

Circle diagram of species attribution analysis of resistance genes. Description: The inner circle is the species distribution of ARO(Antibiotic Resistance Ontology), and the outer circle is the species distribution of all sample genes in the group.



**Figure 5**

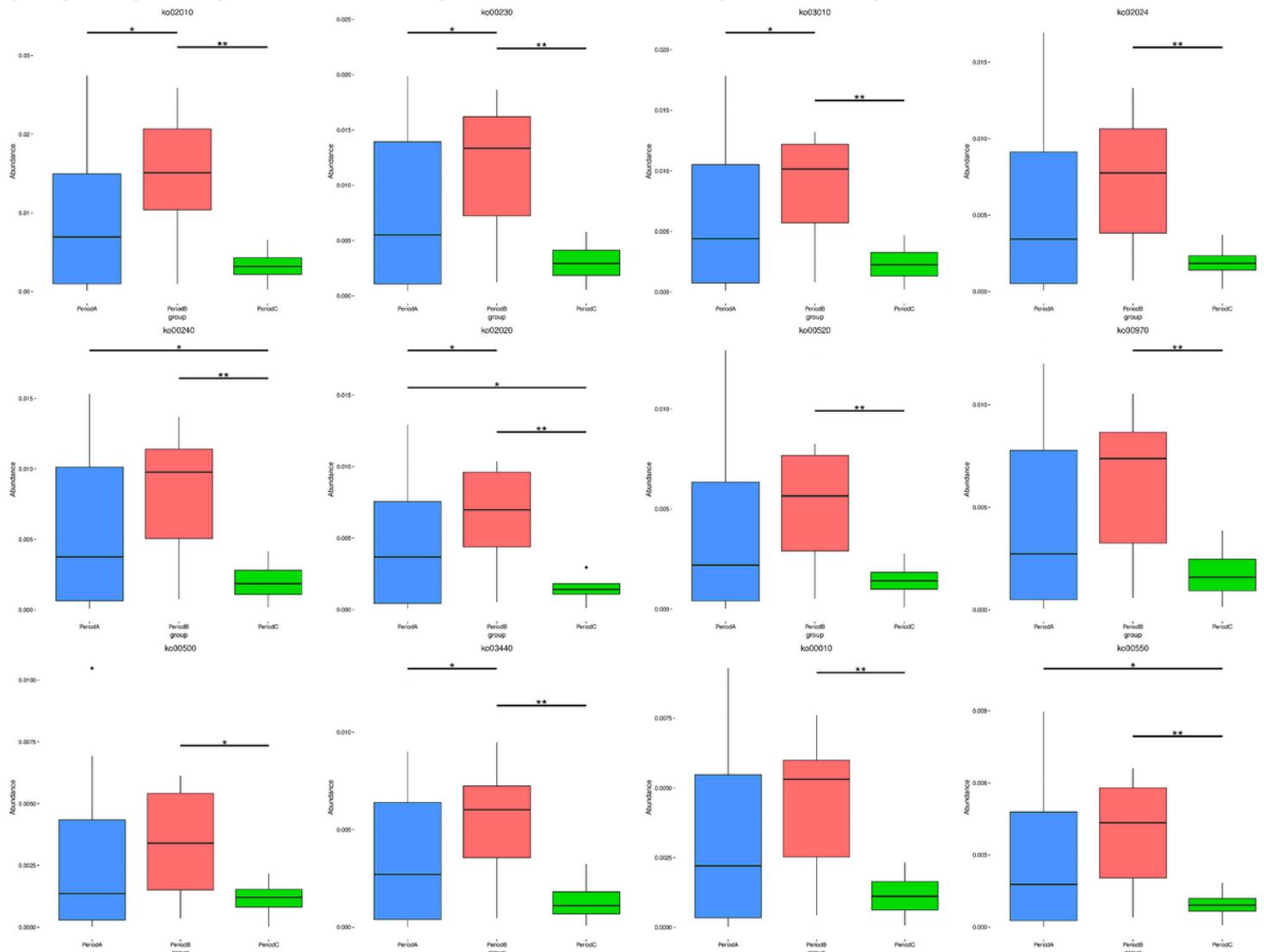
Functional Box Diagram for Metastat Analysis of functional differences between groups(KEGG). Description: The horizontal axis in the picture is the sample grouping, and the longitudinal axis is the relative abundance of the corresponding function. The horizontal line represents two groups with significant differences, and the absence of them means that there is no difference in this function between the two groups. "\*" means that there is a significant difference between the two groups ( $q$  value  $< 0.05$ ), and "\*\*" means that there is a very significant difference between the two groups ( $q$  value  $< 0.01$ ). The Metastats method is used to take hypothetical test of the functional abundance data between groups to get the p value, and through the correction of the p value, the q value is obtained.



**Figure 5**

Functional Box Diagram for Metastat Analysis of functional differences between groups(KEGG). Description: The horizontal axis in the picture is the sample grouping, and the longitudinal axis is the relative abundance of the corresponding function. The horizontal line represents two groups with significant differences, and the absence of them means that there is no difference in this function

between the two groups. "\*" means that there is a significant difference between the two groups (q value < 0.05), and "\*\*" means that there is a very significant difference between the two groups ((q value < 0.01). The Metastats method is used to take hypothetical test of the functional abundance data between groups to get the p value, and through the correction of the p value, the q value is obtained.



**Figure 5**

Functional Box Diagram for Metastat Analysis of functional differences between groups(KEGG). Description: The horizontal axis in the picture is the sample grouping, and the longitudinal axis is the relative abundance of the corresponding function. The horizontal line represents two groups with significant differences, and the absence of them means that there is no difference in this function between the two groups. "\*" means that there is a significant difference between the two groups (q value < 0.05), and "\*\*" means that there is a very significant difference between the two groups ((q value < 0.01). The Metastats method is used to take hypothetical test of the functional abundance data between groups to get the p value, and through the correction of the p value, the q value is obtained.

## Supplementary Files

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

- [COPDadditionalfile.docx](#)
- [COPDadditionalfile.docx](#)
- [COPDadditionalfile.docx](#)
- [OnlinefigureS1.Png](#)
- [OnlinefigureS1.Png](#)
- [OnlinefigureS1.Png](#)
- [OnlinefigureS2.Png](#)
- [OnlinefigureS2.Png](#)
- [OnlinefigureS2.Png](#)
- [OnlinefigureS3.Png](#)
- [OnlinefigureS3.Png](#)
- [OnlinefigureS3.Png](#)