The Oral Microbiome and Salivary Proteins Influence Caries in Preschool Children

Wang Chen
Chongqing Medical University https://orcid.org/0000-0001-9709-9657
Qian Jiang
Chongqing Medical University
GuoWei Yan
Chongqing Medical University
DeQin Yang (yangdeqin@hospital.cqmu.edu.cn)
https://orcid.org/0000-0002-8769-0117

Research article

Keywords: dental caries, oral microbiota, salivary proteins, 16S rDNA, iTRAQ

Posted Date: August 28th, 2019

DOI: https://doi.org/10.21203/rs.2.13636/v1

Version of Record: A version of this preprint was published on October 28th, 2020. See the published version at https://doi.org/10.1186/s12903-020-01262-9.

Abstract

Background Oral microbiome and salivary proteins play a critical role in the occurrence and development of caries. In this study, we used metagenomic and metaproteomic analyses to explore the microbiological and proteinic biomarkers and investigate the etiology of caries in preschool children. Our study aims to offer a better comprehension of these factors and the relationship with caries, and these findings might facilitate caries risk assessment and provide a basis for future prevention strategies.

Methods Children six to eight years old living in rural isolated areas with or without caries were recruited. Supragingival plaque and unstimulated saliva were collected for 16S rDNA pyrosequencing and isobaric tags for relative and absolute quantitation (iTRAQ)technique coupled with quantitative nano-flow liquid chromatography-tandem mass spectrometry(LC-MS/MS), respectively.

Results A total of 328486 high-quality 16S rRNA sequences were obtained from 40 samples including 20 plaque and 20 saliva subjects, which was clustered into 14,076 OTUs, representing 18 phyla, 28 classes, 48 orders, 78 families,135 genera, and 410 species. We found the six most abundant phyla (Firmicutes, Bacteroidetes, Fusobacteria, Proteobacteria, Actinobacteria, and Candidate division TM7). The Alpha diversity analysis demonstrated that the richness and diversity of the bacterial communities were similar between caries and caries-free children; differences in the bacterial community composition was analyzed by LEfSe. The core microbiome was defined as 18 predominant genera coming from the saliva and plaque of subjects. For the results of salivary proteome analysis, 9135 unique peptides and 1662 proteins group were identified from 20 salivary samples. 258 proteins were differentially expressed between the caries-free and caries-active group, which might be a potential proteinic biomarker of caries and health status.

Conclusions Dental caries is a multifactorial disease affecting many children around the world. The result of our study revealed the complexities of the oral bacterial community and confirmed “ecological plaque hypothesis”. In addition, as an important host factor of caries the salivary proteins are different in healthy or carious status. Exploration on microbiological or protein biomarkers is of great significance to prevent dental caries.

Background

It is estimated that 2.4 billion people suffer from dental caries, and 621 million of them are children [1]. In preschool children, severe caries can affect their quality of life [2]. Strategies to prevent caries are based on a comprehensive understanding of its etiology and effective control of the risk factors. It is recognized that causes of caries include microorganism in the mouth and host factors. The oral cavity is one of the most diverse and complex microbial environments [3]. Previous studies demonstrated that oral plaque film has high relevance in dental caries. The acid produced from bacteria break the balance of tooth mineralization and demineralization and the host have no rapid response to pH changes, which lead to organic degradation [4]. Saliva is the main microenvironment of oral microorganisms, and to some

extent, saliva microorganism determines the structure of plaque. Salivary protein has a crucial role in monitoring health status or monitoring disease [5]. It was reported that the proteins in saliva could modulate the balance of oral health and homeostasis, maintain a stable ecosystem, and inhibit the growth of cariogenic bacteria [6].

In the past few decades, several investigators have proposed several hypotheses regarding the etiology of caries [7-9], the relationship between bacteria and dental caries, the complexity of the oral bacterial structure, and the difference of bacterial components. Previous studies also mentioned that some salivary proteomic molecules could regulate the oral cavity microbial flora and correlate to caries [6, 10, 11]. Unfortunately, due to differences in samples, technologies, and analytical methods, the results remain controversial and the biomarker information unclear.

Thanks to recent advancement in molecular biology techniques, metagenomic and metaproteomic can be used to obtain a complete analysis of the oral bacteria and proteomic. Next-generation sequencing technologies have been successfully applied in the oral microbial analysis [12-14]. The isobaric tags for relative and absolute quantitation (iTRAQ) is a new technique which uses isotopes to label polypeptides for comparing proteomes quantitatively [15, 16]. To the best of our knowledge, previous studies of caries-related microbiome and proteome were detached. Our present study uses metagenomic and metaproteomic analyses to explore the microbiological and proteinic biomarkers and investigate caries etiology in preschool children.

In this study, we enrolled 6-8 years old preschool children (isolated population) who come from Tujia and Miao minority autonomous county, Pengshui, Chongqing, China. These children have a simple and homogeneous diet; therefore, the impact of different diets and daily living habits is avoided. In the current study, the oral microecological diversity was studied using 16S rDNA pyrosequencing, and the salivary proteins were analyzed using the iTRAQ technique coupled with quantitative nano-flow liquid chromatography-tandem mass spectrometry (LC-MS/MS). Our study aimed to 1) detect the microbiological compositions and to investigate the core microbiome; 2) identify the salivary proteomic and characterize the functional classification in preschool children with or without caries, and 3) attempt to identify microbiological or proteinic biomarkers helpful to prevent dental caries.

Methods

2.1. Subjects Selection and Sample Collection

All the subjects were recruited from Tujia and Miao minority autonomous county, Pengshui, Chongqing, China. This is a remote mountainous area whose population has a simple diet and low mobility. Before enrollment, parent or guardians of the subjects (6-8 years children) were provided with an informed consent explaining the study objectives. According to the criteria from the World Health Organization, 4th-edition publication of “Oral Health Surveys, Basic Methods,” children whose DMFT index was over eight were divided into a caries-active group, and caries-free group (DMFS = 0). Finally, 40 caries-active subjects (20 males and 20 females), and 40 caries-free subjects (20 males and 20 females) were

selected. All the children have 1) no long-term (>3 months) history of living in different places; 2) no antibiotic therapy; 3) no use of fluoride at least three months before the examination; and 4) no other oral diseases or systemic diseases [17]. This experiment was approved by the Ethics Committee of affiliated Hospital of Stomatology of Chongqing Medical University.

Supragingival plaque and unstimulated saliva were collected in the morning before eating, drinking, and tooth brushing. Caries-active plaque samples were collected from each caries site, and a caries-free plaque was collected from healthy molar surfaces. The samples were placed in 1.5 ml sterile Eppendorf tubes. Unstimulated saliva was also collected and transferred to sterile 5 ml microcentrifuge tubes. All the samples were immediately frozen at -20°C and stored at -80°C in the laboratory until further processing.

2.2 Oral Microbiome Analysis

2.2.1. DNA extraction and purification

The genome of all the samples was extracted using Promega Genomic DNA Purification Kit, following the manufacturer's instructions. To detect if the sample was free from contamination, 5 μl DNA samples were taken for agarose gel electrophoresis (110V, 20 min). We selected distinct bands showing no obvious trailing phenomenon, which showed that the genome was relatively complete without significant RNA and protein contamination. Then we evaluate the quality by measuring the absorbance at A260/280 using UV spectrophotometer (DU-800, Beckman Coulter). The samples with the A260: A280 ratios at 1.8-2.0, and the DNA concentrations in 20-100 ng/μl were screened, and the results indicated that the genomic DNA extracted met the requirements for subsequent sequencing [18]. Finally, 40 high-quality samples were selected to perform sequencing analysis: SN (caries-free saliva group, n=10); PN (caries-free plaque group, n=10); SH (caries-active saliva group, n=10); and PH (caries-active plaque group, n=10). The DNA samples were stored at −20°C before use.

2.2.2 PCR Amplification and Pyrosequencing

The general primers for PCR amplification of the bacterial 16S rDNA V1-V3 region were the reverse primer 533R (5'-TTACCGCGGCTGCTGGCAC-3'), and forward primer 8F(5'-AGAGTTTGATCCTGGCTCAG-3'). After adding the tag sequence, the 454 Life Science A or B sequencing adaptor were connected with general primers by linker sequence [19]. PCR amplification was performed using Trans Start Fastpfu DNA Polymerase (TransGen AP221-02), three replicates per sample. The PCR products of the same sample were taken for 2% agarose gel electrophoresis, and the AxyPrep DNA Gel Extraction Kit was used to recover the PCR products. Afterward, the 16S rRNA gene was sequenced on the Roche 454 GS FLX+ Sequencing Method Manual_XLR70 kit.

2.2.3 Bioinformatics Analysis

Ambiguous base, homologous base, and sequences shorter than the original 200 bp sequence were removed or discarded to obtain high-quality sequences [20]. The high-quality sequences (≥ 80% confidence) were compared using the SILVA database [21] (version106) and Mothur software (version

1.31.2) [22] at a 97% similarity level. Based on the results of operational taxonomic units (OTUs) clustering analysis; community richness and diversity indices of ACE, Chao, Shannon, Simpson, and the Good's coverage were calculated. We constructed a circle phylogenetic tree using the ITOL platform to explore the relationships of the general microbial population. The principal coordinates analysis (PCoA) was based on Bray-Curtis distances at an OUT level with 97% identity. PCoA was used to compare the similarities in the bacterial community structures among the four groups. The linear discriminant analysis (LDA) of effect size (LEfSe) was performed to define the biological class features and establish statistical significance [23]. A Venn diagram was made using Mothur software to reveal the core microbiome. The significant differences in microbial community composition were analyzed using one-way ANOVA with SPSS Software (version 25.0), and statistical significance was set at $P< 0.05$.

2.3. Salivary Proteomics Analysis

2.3.1 Sample Preparation

Salivary samples from the SN and SH groups used in the metagenomic analysis were selected for further proteomics analyses. A total of 20 saliva proteome samples from each group were pooled (SN=10, SH=10), and the mixture was centrifuged in 5 KDa ultrafiltration tube for concentration until the volume was about 200 μL. Protein quantification was performed using the Bradford assay with bovine serum albumin (BSA) as standard and analyzed with SDS-PAGE. Twenty picograms of protein sample were mixed with SDS-PAGE sample loading buffer (10% SDS, 0.5% BTB, 50% glycerinum , 500 mMDTT, 250 mM Tris HCl pH6.8) in a ratio of 1:5 v/v, incubated in a boiling water bath for 5 min and then centrifuged at 14000g for 20 min. The supernatant was taken for 12.5% SDS-PAGE electrophoresis (14mA, 90min).

2.3.2 Proteins Filter-Aided Sample Preparation (FASP)

The method of filter-aided sample preparation (FASP) was used for protein extraction, digestion, and peptide separation. Samples from the SN and SH groups were mixed with SDT Lysis Buffer (4%SDS,100mM Tris-HCl,1mM DTT pH7.6), incubated in a boiling water bath and then centrifuged in 30KDa ultrafiltration tube to a final volume of 25 μL. To remove large excess of detergent and interfering substances, UA buffer (8M urea,150mM Tris HCl pH8.0) was mixed with protein extract in 30 KDa ultrafiltration tube and centrifuged at 14000 g for 15 min. The filtered liquor was discarded, and the on-filter remaining material was added 100 μL IAA (50nmM IAA in UA) and centrifuged at 14000 g for 10 min. This process of extensive washes and buffer exchange was repeated several times. The peptides were quantified using OD280.

2.3.3 iTRAQ Labeling and SCX Fractionation

Ninety picograms of treated samples from the SN and SH groups were labeled with the iTRAQ Reagent-4plex Multiplex Kit (AB SCIEX) according to the manufacturer's instructions. Peptides from each group were labeled with the following tags: 114 and 116 tags for the SN, 115 and 117 tags for SH, respectively. Each labeled peptide segments were mixed an underwent a strong cation-exchange chromatography

(SCX) fractionation. The SCX gradient information is provided in the **additional Table 1**. According to SCX chromatogram, ten fractions were combined, which then were lyophilized and desalinated using C18 Cartridge (Sigma-Aldrich, St Louis, MO, USA).

2.3.4 Mass Spectrometry Analysis

Peptides were loaded to the Thermo scientific EASY column (2 cm x 100 µm 5µm-C18) and then separated using the same Thermo scientific EASY column (75 µm x 100 mm 3µm-C18) mounted in an EASY-nLC 1000 system with the flow rate of 250 nl/min. Buffer A consisted of 0.1% formic acid, while buffer B consisted of 0.1% formic acid, 84% ACN. The chromatographic column was balanced with 95% buffer A. The flow rate of the gradient started at 0% buffer B, going to 35% buffer B in 100 min, continuing to 100% buffer B in 8 min, and maintaining 100% buffer B in 120 min.

The eluates were injected into a Q-Exactive mass spectrometer (Thermo Fisher Scientific, Waltham, MA, USA), run in positive ion mode with a full MS scan from 300 to 1800 m/z. The MS/MS spectra acquisition parameters were as follows: full scan resolutions set to 70,000 at m/z 200; the AGC target was 3 x 106 with a maximum fill time of 10 ms; dynamic exclusion set to 40 s. We used higher collision energy dissociation (HCD) to collect the mass-charge ratio of peptide fragments. Ten MS2 scans were collected after each full scan. The normalized collision energy (NCE) was 30 eV.

2.3.5 Data Analysis

The raw data were processed using Proteome Discover 1.3 (Thermo Fisher Scientific, Waltham, MA, USA; version 1.3) and searched against the International Protein Index human database (ipi.human.v3.87.fasta) containing 91464 sequences using the Mascot search engine (Matrix Science, Boston, MA, USA; version 2.2). The **additional Table 2** highlights the parameters used for the database search. Proteins were filtered, and the false discovery rate (FDR) of peptide and protein level was less than 1%. The ionic peak strength values of peptides were quantitatively analyzed with the Proteome Discoverer 1.3（Thermo Scientific, San Jose, California ,USA）. The student's t-test was used to evaluate the differences between the two groups, which were considered statistically significant if P＜0.05. Proteins with quantification P-value < 0.05 and fold changes > 1.2 were identified as differentially expressed proteins. Functional classification of differentially expressed proteins was evaluated performing gene ontology (GO) analysis, which includes three-term of biological processes, molecular function, and cellular components.

Results

3.1. Plaque and salivary microbiome

3.1.1. Sequences Information and Bacterial Diversity

All the samples were divided into four groups. After 454 pyrosequencing, a total of 415,203 16S rRNA sequences were obtained from 40 samples, 20 from plaque and 20 from saliva, and 328486 high-quality

sequences (79%) passed the quality-control test. The mean sequence length was 476 bp, with an average sequencing depth of 6,347 reads per sample. All the qualified sequences (≥ 97% similarity level), were compared to the SILVA database (version10.6) using Mothur software (version 1.31.2) and the reads were clustered into 14,076 operational taxonomic units (OTUs) including 6,042 OTUs from caries plaque, 6,757 OTUs from caries-free plaque, 5,406 OTUs from caries saliva, and 5,561 OTUs from caries-free saliva **(Table 1).**

The indices of Shannon, Simpson, Chao, and ACE were calculated to obtain the bacterial richness and diversity; the Good's coverage reflects the sequencing depth. The oral microbial diversity parameters are shown in **Table 2**; there was no significant difference in the richness and diversity of the bacterial communities between caries and caries-free group(P> 0.05). The Good's coverage for each group was over 95%, indicating adequate sequencing depth.

Table 1. Statistics of the Microbial Number in Plaque and Saliva From Several Taxonomic Levels

	PH	PN	SH	SN
phylum	13	15	12	14
class	20	23	20	21
order	33	35	35	36
family	50	52	62	54
genus	82	90	105	97
OTU(0.03)	6042	6757	5406	5661

Table 2. Community Richness Estimator and Diversity Estimator of Every Group

	PH	PN	SH	SN
ACE(0.03)	18587 (18044,19155)	20059 (19502,20641)	14641 (14189,15117)	16707 (16198,17241)
Chao(0.03)	12126 (11549,12764)	12813 (12273,13406)	10159 (9686,10685)	10919 (10410,11482)
Shannon(0.03)	6.74 (6.73,6.75)	6.8 (6.78,6.82)	6.53 ((6.51,6.54)	6.54 6.53,6.56)
Simpson(0.03)	0.0038 (0.0037,0.0039)	0.0034 (0.0033,0.0035)	0.0049 (0.0049,0.005)	0.0045 (0.0044,0.0046)
Coverage(0.03)	0.951656	0.964732	0.959671	0.958081

3.1.2. Bacterial Community Structure and Composition in Different Niches

A total of 18 phyla, 28 classes, 48 orders, 78 families,135 genera, and 410 species were detected in the 40 samples analyzed. Overall, the six most abundant phyla were Firmicutes (33.66%), Bacteroidetes (23.61%), Fusobacteria(19.83%), Proteobacteria(12.89%), Actinobacteria(6.85%), and Candidate division TM7(2%). Together, they represent 98.84% of the total sequences. The reads were dominated by 13 genera including Streptococcus (17.86%), Leptotrichia (14.60%), Prevotella (9.99%), Neisseria (7.51%), Porphyromonas (5.47%), Fusobacterium (5.12%), Capnocytophaga (5.10%), Veillonella (2.86%), Actinomyces (2.79%), Gemella (2.20%), Granulicatella (2.17%), Johnsonella (2.09%), and Derxia (2. 06%). They occupied 79.84% of the whole. **Figure 1a and Figure 1b** shows the taxonomic distributions of the predominant bacteria at the phyla and genera levels. The relative abundance between caries-active and caries-free subjects were compared using the Wilcoxon rank-sum test. The significant difference was detected among four group**(additional Table3 and additional Table4)**.A higher abundance(relative abundance >1%)of Porphyromonas was detected in the SN group(P=0.04396) and Derxia in the SH group(0.04493). A higher abundance (relative abundance >1%)of Capnocytophaga was observed in the PN group than PH group(P=0.01898)**(Fig.2)**

To explore the relationship of the bacterial-community, a circular phylogenetic tree from 133 genera was constructed **(Fig. 3)**. The relative abundance in caries-free and caries group, as well as the community composition at the genus level, could be observed in the Heatmap **(Fig. 4)**, in which we could find that the predominant microbial communities were largely similar, but the variety of individual microorganism is apparent among the four groups.

The analysis of similarities (ANOSIM) was performed to compare similarities and dissimilarities in the bacterial community structures among the four groups. The principal coordinates analysis (PCoA), based on the Bray-Curtis distances, demonstrated segregations between samples from dental plaque and saliva. As shown in **Figure 5**, the microbiota from the SN group overlapped with SH, so as PN and PH groups. A clear difference can be observed within the saliva and the plaque samples. The PCoA result indicated that the oral microbial community compositions are similar between caries-active patients and healthy controls; nevertheless, there were some dissimilarities in two different niches. These differences were also observed using the nonmetric multidimensional scaling (NMDS) analysis **(Additional Fig.1).**

The LEfSe analysis was performed to represent differences in the bacterial community composition among the four groups, and their effect sizes were represented in a taxonomic tree. **Figure 6a** shows cladograms representing the microbial community with significant differences at different levels. Significant differences detected in the PH group were Dialister and Selenomonas at the genus level. Capnocytophaga (belonging to class Flavobacteria, order Flavobacteriales, family Flavobacteriaceae), Fusobacterium (belonging to family Fusobacteriaceae), and Desulfuromonadales_g_unclassified (belonging to order Desulfuromonadales) exhibited relatively higher abundance in the PN group. As for the saliva group, Actinomyces and Mogibacterium were significantly enriched in the SH group, while the relative abundance of Haemophilus and Porphyromonas was higher in the SN group ($LDA > 2$, $P<0.05$). **Figure 6b** shows the LDA score representing the impact of differential features among groups.

3.1.3. The Core Microbiome

A Venn diagram was used to display the core microbiome, the overlapping areas in the circles stand for the members shared among the four groups in each taxonomical level. The oral microbiome analysis revealed an overlap of shared OUTs and genera. As the Venn diagram shows **(Fig.7a,7b)**; 14076 OTUs were identified including 5406, 6042, 5661, and 6757 OTUs in the SH, PH, SN and PN groups, respectively. A total of 1328 OTUs and 71 genera were common among the four groups, occupying 9.4% of all the OTUs (14076 OTUs), and 52.6% of all the genera (135 genera) detected. We detected 18 predominant genera uniform in the samples from saliva and plaque subjects including Abiotrophia, Actinomyces, Bergeyella, Campylobacter, Capnocytophaga, Corynebacterium, Derxia, Fusobacterium, Gemella, Granulicatella, Johnsonella, Neisseria, Porphyromonas, Prevotella, Propionibacterium, Streptococcus, Veillonella, and Ruminococcaceae uncultured. This shared microbiome supports the existence of an “oral core microbiome” and this shared genera may be part of the oral core microbiome in the dental plaque and saliva of caries from patient and healthy controls.

3.2. Differentially Expressed Proteins Related to Dental Caries and Its Functional Classification

Proteomic analysis of saliva samples was performed using the iTRAQ-coupled LC-MS/MS method to detect protein biomarkers of caries risk in children. Two saliva samples from the SN (caries-free saliva group, n=10) and SH (caries-active saliva group, n=10) were used for this study. The protein bands of the saliva samples from subjects with and without caries were not entirely consistent with SDS-PAGE electrophoresis **(Additional Fig.2),** indicating the existence of differentially expressed salivary protein between healthy and cariogenic children. The salivary protein samples from the SN and SH groups were also used (90 μg for each group) for the iTRAQ analysis. After querying the database, a total of 9135 unique peptides and 1662 proteins group (unique peptides ≥ 1) were identified, including 1626 proteins with quantitative information **(Additional file1)**. Pearson correlation between each experimental group and its replicate showed good reproducibility **(Appendix Fig.3a,3b).** We found 258 proteins to be differentially expressed according to the criteria of P-value < 0.05 and ratio- fold change > 1.2. Some differential expressed proteins between caries and healthy saliva were listed in the **additional Table5.**

Gene ontology analysis was performed to explore the biological function of the differentially expressed proteins base on their biological processes, molecular function, and cellular components. The proteins involved in the metabolic process (16.91%), regulation of biological process (12.99%), and response to the stimulus (12.54%) were enriched in the SH group compared with healthy controls **(Figure. 8a)**. The majority of differentially expressed proteins were found in the cytoplasm (17.05%), extracellular (14.26%), and membrane (13.77%) **(Figure. 8b)**. GO analysis showed that proteins involved in the protein binding (35.76%) and catalytic activity (16.94%) were enriched in SH group **(Figure.8c).**

Discussion

The etiological concepts of oral infectious diseases, including caries and periodontal disease, has gradually changed from a single pathogen theory to a microecological imbalance theory [4, 24, 25].

Therefore, a systems biology approach is required to explain the complex interactions between the microbiome and the host. As far as we know, an approximate of 1000 bacterial species have been found in the oral cavity [26] due to the advent of molecular analysis methods such as the polymerase chain reaction denaturing gradient gel electrophoresis(PCR-DGGE), human oral Microbe identification microarray (HOMIM), next-generation sequence technologies (NGS), and others [24, 25, 27]. Recently, 16S rRNA sequence analysis has been applied in the study of uncultured oral microbial communities; this method is an advantageous molecular analysis technology for investigating the oral bacteria diversity and microbial community composition in oral diseases. Meanwhile, host factors should not be neglected. Salivary proteins play an essential role in the occurrence and development of caries. Proteomics has advanced significantly over the past decades, and it has been applied for the study of caries and other oral diseases [28, 29]. In this study, we preliminarily explored microbiome and host factors in childhood caries to find biomarkers using the high-throughput technique of 16S rDNA pyrosequencing and iTRAQ-coupled LC-MS/MS.

We used 16S rRNA pyrosequencing to analyze plaque and salivary bacteria from caries-active and caries-free samples. After 454 pyrosequencing, a total of 415,203 16S rRNA sequences were obtained from 40 samples, including 20 from plaque and 20 from saliva subjects and the sequences were clustered into 14,076 OTUs. Good's coverage for each group was over 95%, indicating adequate sequencing depth. Eighteen phyla, 28 classes, 48 orders, 78 families,135 genera, and 410 species were detected with 454 pyrosequencing, and these results exceeded the data of the previous HOMIM analysis of our group [28, 29]. Some rare and not-yet-cultivated bacteria species that might be related to caries were detected in the present study.

According to the results of alpha diversity indices, the richness and diversity of the bacterial communities in caries groups were similar to the caries-free group, as previously found in other studies [17, 30, 31]. This result might show that the diversity of the microbial community has little effect on caries. Xiao et al., [18] studied the supragingival plaque of dental caries and demonstrated a higher bacterial diversity of healthy dental plaques compared to dental caries. The dubious results may be due to the difference between individuals, the selection process of subjects, sequencing technology methods, and other factors. These discrepant data further confirmed the complexity of oral microbiota. Moreover, we found the six most abundant phyla including *Firmicutes, Bacteroidetes, Fusobacteria, Proteobacteria, Actinobacteria, and Candidate division TM7,* which were in agreement with the results of previous studies [13, 17, 18]. The comparison of the dominant bacterial community on phylum and genus level were similar among the four groups, but the relative abundance was different. At the genus level, 135 genera were detected, including 13 prevalent genera, roughly similar to previous studies [30, 32, 33]. *Porphyromonas* were detected at high levels in the SN group; higher abundances of *Capnocytophaga* were observed in the PN group, and higher abundances of Actinomyces were observed in the PH group($P<0.05$). *Porphyromonas* are recognized as a significant pathogen of chronic adult periodontitis [34], but in caries, it seems to be related to dental health [17, 32]. At the beginning of caries, several microorganisms gather on the tooth surface in an ordered way, and then the oral ecosystem is broken. Acidogenic and acid-tolerating species such as *Streptococcus* and *Lactobacillus* shift

toward community dominance [35, 36]. In this study, the detection rate of well-recognized cariogenic (*Streptococcus*) was high in both the caries-free and caries group and had no significant difference. Also, Lactobacillus was detected in only four caries-active saliva samples and one caries-active plaque sample with low detection rate (about 0.01%). These outcomes may derive from different categories of severity, sampling location, phase of bacterial colonization, and other reasons. To understand bacterial ecology in oral diseases, further exploration of the dynamic alterations in oral microbiota during the caries process is needed. *Dialister Selenomonas, Actinomyces*, and *Mogibacterium* were identified at significantly higher levels in the caries-active sample with the LEfSe analysis, which might be recognized as a potential bacterial biomarker. These results indicated that the cause of dental caries is not the activity of specific microorganisms, some cariogenic bacteria are also part of the normal oral flora and they are a constant variable in the occurrence and development of dental caries [37].

Oral health and disease are correlated with the interplay inside the oral microbial community. Saliva, as the main microenvironment of oral bacteria, is considered a significant influence on the colonization of microorganisms [38]. The result of PCoA analysis revealed clear segregation between samples from dental plaque and saliva, meaning the distribution of microorganism structures in plaque were different from those of saliva. Ren et al.,[13] suggested that dental plaque had significant phylogenetic differences compared with saliva and tongue coating. The reasons for this situation are probably related to the physicochemical features at different sites, such as pH, oxygen concentration, and bacterial adherence [39].

Human microbiological studies support the concept of a “core microbiome,” which is referred to the microbiome shared by most individuals in a specific environment of the body such as the skin, nasal cavity, intestinal tract, and oral cavity [40-43]. In our study, the Venn diagram shows that 52.6% of all the genera were shared and 18 predominant genera uniform was identified in saliva and plaque subjects, indicating the existence of “oral core microbiome,” as suggested by a previous study [18]. However, Jiang et al. evaluated the microbiota conservation level and found that OTU sharing decreases with each additional individual sample. They supported the notion that there was no species-level “oral core microbiome” of saliva in children dental caries [17]. In fact, the core microbiome contributes to the functional stability and microecological balance of healthy oral cavity.

For the result of salivary proteome analysis, we detected differentially expressed proteins and its functional classification between the SN and SH groups. A total of 9135 unique peptides and 1662 proteins group were identified. Compared with the method of electrospray ionization ion-trap tandem mass spectrometry (ESI-MS/MS) used in our previous study, the number of proteins and peptides identified in our present study was higher [44]. The iTRAQ technique uses stable isotopes to label samples in each group and then performs tandem mass spectrometry, which is more convenient and sensitive for quantitative analysis. Considering that physiological processes influence the component of saliva, all the saliva specimens were collected using a standard method to guarantee the quality of samples. Pearson correlation between each experimental group and its replicate had good reproducibility indicating the reliability of the data.

Two hundred and fifty-eight proteins were found to be differentially expressed, which might play a part in the process of childhood dental caries. Some important proteins were included in differentially expressed proteins, such as *lactoferrin, matrix metalloproteinase-9, cystatin-B, protein S100-A9, mucin-7, and proline-rich protein*, which have demonstrated a potential relationship with caries in previous studies [45-47]. Lactoferrin is an antibacterial protein with the iron-chelating property directly binding to bacteria and agglutinate *S. mutans*. The combined bacteria is easy to be removed with the mechanical saliva action [48, 49]. Also, it was reported that there was a high correlation between matrix metalloproteinase-9 and caries lesion depth [47]. MMPs and cysteine cathepsins could affect the caries process in the early phases of demineralization [50]. The result of the GO analysis shows that differentially expressed proteins were associated with metabolic process and regulation of the biological process, mainly in the protein binding. As common salivary proteins, mucin-7 binding to proline-rich protein could be adsorbed onto the tooth surface to form a pellicle that regulates the bacteria adhesion and modulate the demineralization/remineralization process [51, 52]. The molecular sequencing techniques make precise identification of proteins. However, because of the complexity of saliva and immature technologies, proteinic information in our present research is not complete, and some low abundance proteins from the microorganism and its metabolite were not explored. The specific mechanism and more detailed information about the proteins in the saliva need to be further investigated. There is still a long way to devise strategies that modulate interactions of microbiota and salivary proteins for the treatment of oral diseases.

Conclusion

In summary, 18 phyla, 28 classes, 48 orders, 78 families,135 genera, and 410 species were detected using 454 pyrosequencing technology in 40 samples. Alpha diversity analysis demonstrated that the richness and diversity of the bacterial communities were similar between caries and caries-free children. Meanwhile, a series of bacteria related to caries was detected, which could be taken as potential microbial biomarkers for caries in children six to eight years old. The diverse oral bacteria maintain the balance of the oral environment, and this study confirms “ecological plaque hypothesis.” A portion of the detected microorganisms was shared in all the samples, including plaque and saliva, supporting the existence of an oral core microbiome. In salivary protein, we explored the relationship between human saliva protein and oral disease, as well as the functional classification of differentially expressed proteins. Using the iTRAQ technique, a total of 9135 unique peptides and 1662 proteins group were identified, and 258 proteins were found to be differentially expressed. These differentially expressed proteins might be a potential proteinic biomarker of caries or health status. Our study provides an in-depth comprehension of the etiology of children caries from the perspective of microorganisms and host factors. These microbiological and proteinic biomarkers could help caries risk assessment and provide a basis for further prevention strategies.

List Of Abbreviations

DMFT: missing (due to caries), or filled tooth surfaces in primary teeth

iTRAQ: isobaric tags for relative and absolute quantitation

LC-MS/MS: quantitative nano-flow liquid chromatography-tandem mass spectrometry

OUTs: operational taxonomic units

PCoA: principal coordinates analysis

FASP: Filter aided proteome preparation

GO: gene ontology

LEfSe: linear discriminant analysis of effect size

SN: caries-free saliva group

PN: caries-free plaque group

SH: caries-active saliva group

PH: caries-active plaque group

Declarations

Ethics approval and consent to participate

This experiment was approved by the Ethics Committee of Affiliated Hospital of Stomatology of Chongqing Medical University and the approval number is CQHS-IRB-2016-05. Informed and written consent was obtained from the guardians of all children.

Consent for publication

Not applicable

Availability of data and materials

The datasets used during the current study are available from the corresponding author on reasonable request.

Competing interests

The authors declare that they have no competing interests

Funding

This work was financially supported by the National Natural Science Foundation of China (No.31571508 and No.81170950) to Dq.Y .The National Natural Science Foundation had no role in design of the study, the collection, analysis, and interpretation of data and writing the manuscript. It supported the funding of the study only.

Authors' contributions

WC and GWY designed and performed the experiment. WC analyzed the data and drafted the initial manuscript. QJ revised the manuscript .DQY reviewed the manuscript, and approved the final manuscript as submitted .All authors gave final approval and agreed to be accountable for this work.

Acknowledgements

We would like to thanks all the children, parents and staff and the Affiliated Hospital of Stomatology of Chongqing Medical University for their help and participation in this study. We also acknowledge Shanghai Majorbio Bio-pharm Technology Co., Ltd for their kind help in 454 pyrosequencing and iTRAQ-coupled LC-MS/MS technology.

References

1. **Global, regional, and national incidence, prevalence, and years lived with disability for 310 diseases and injuries, 1990-2015: a systematic analysis for the Global Burden of Disease Study 2015**. *Lancet (London, England)* 2016, **388**(10053):1545-1602.
2. Listl S, Galloway J, Mossey PA, Marcenes W: **Global Economic Impact of Dental Diseases**. *Journal of dental research* 2015, **94**(10):1355-1361.
3. Duran-Pinedo AE, Frias-Lopez J: **Beyond microbial community composition: functional activities of the oral microbiome in health and disease**. *Microbes and infection* 2015, **17**(7):505-516.
4. Takahashi N, Nyvad B: **Ecological Hypothesis of Dentin and Root Caries**. *Caries research* 2016, **50**(4):422-431.
5. de Almeida Pdel V, Gregio AM, Machado MA, de Lima AA, Azevedo LR: **Saliva composition and functions: a comprehensive review**. *The journal of contemporary dental practice* 2008, **9**(3):72-80.
6. Van Nieuw Amerongen A, Bolscher JG, Veerman EC: **Salivary proteins: protective and diagnostic value in cariology?** *Caries research* 2004, **38**(3):247-253.
7. Marsh PD: **Microbial ecology of dental plaque and its significance in health and disease**. *Advances in dental research* 1994, **8**(2):263-271.
8. Kleinberg I: **A mixed-bacteria ecological approach to understanding the role of the oral bacteria in dental caries causation: an alternative to Streptococcus mutans and the specific-plaque hypothesis**. *Critical reviews in oral biology and medicine : an official publication of the American Association of Oral Biologists* 2002, **13**(2):108-125.

9. Takahashi N, Nyvad B: **Caries ecology revisited: microbial dynamics and the caries process**. *Caries research* 2008, **42**(6):409-418.
10. Hemadi AS, Huang R, Zhou Y, Zou J: **Salivary proteins and microbiota as biomarkers for early childhood caries risk assessment**. *Int J Oral Sci* 2017, **9**(11):e1.
11. Koscielniak D, Jurczak A, Zygmunt A, Krzysciak W: **Salivary proteins in health and disease**. *Acta biochimica Polonica* 2012, **59**(4):451-457.
12. Camelo-Castillo AJ, Mira A, Pico A, Nibali L, Henderson B, Donos N, Tomas I: **Subgingival microbiota in health compared to periodontitis and the influence of smoking**. *Frontiers in microbiology* 2015, **6**:119.
13. Ren W, Zhang Q, Liu X, Zheng S, Ma L, Chen F, Xu T, Xu B: **Exploring the oral microflora of preschool children**. *Journal of microbiology (Seoul, Korea)* 2017, **55**(7):531-537.
14. Zhang M, Chen Y, Xie L, Li Y, Jiang H, Du M: **Pyrosequencing of Plaque Microflora In Twin Children with Discordant Caries Phenotypes**. *PloS one* 2015, **10**(11):e0141310.
15. Treumann A, Thiede B: **Isobaric protein and peptide quantification: perspectives and issues**. *Expert review of proteomics* 2010, **7**(5):647-653.
16. Wang K, Wang X, Zheng S, Niu Y, Zheng W, Qin X, Li Z, Luo J, Jiang W, Zhou X *et al*: **iTRAQ-based quantitative analysis of age-specific variations in salivary proteome of caries-susceptible individuals**. *Journal of translational medicine* 2018, **16**(1):293.
17. Jiang S, Gao X, Jin L, Lo EC: **Salivary Microbiome Diversity in Caries-Free and Caries-Affected Children**. *International journal of molecular sciences* 2016, **17**(12).
18. Xiao C, Ran S, Huang Z, Liang J: **Bacterial Diversity and Community Structure of Supragingival Plaques in Adults with Dental Health or Caries Revealed by 16S Pyrosequencing**. *Frontiers in microbiology* 2016, **7**:1145.
19. Morris RM, Nunn BL, Frazar C, Goodlett DR, Ting YS, Rocap G: **Comparative metaproteomics reveals ocean-scale shifts in microbial nutrient utilization and energy transduction**. *The ISME journal* 2010, **4**(5):673-685.
20. Chen H, Jiang W: **Application of high-throughput sequencing in understanding human oral microbiome related with health and disease**. *Frontiers in microbiology* 2014, **5**:508.
21. Pruesse E, Quast C, Knittel K, Fuchs BM, Ludwig W, Peplies J, Glockner FO: **SILVA: a comprehensive online resource for quality checked and aligned ribosomal RNA sequence data compatible with ARB**. *Nucleic acids research* 2007, **35**(21):7188-7196.
22. Schloss PD, Westcott SL, Ryabin T, Hall JR, Hartmann M, Hollister EB, Lesniewski RA, Oakley BB, Parks DH, Robinson CJ *et al*: **Introducing mothur: open-source, platform-independent, community-supported software for describing and comparing microbial communities**. *Applied and environmental microbiology* 2009, **75**(23):7537-7541.
23. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C: **Metagenomic biomarker discovery and explanation**. *Genome biology* 2011, **12**(6):R60.

24. Xu P, Gunsolley J: **Application of metagenomics in understanding oral health and disease**. *Virulence* 2014, **5**(3):424-432.
25. McLean JS: **Advancements toward a systems level understanding of the human oral microbiome**. *Frontiers in cellular and infection microbiology* 2014, **4**:98.
26. Wade WG: **The oral microbiome in health and disease**. *Pharmacological research* 2013, **69**(1):137-143.
27. Ye DD, Fan MM, Guan Q, Chen HJ, Ma ZS: **[A review on the bioinformatics pipelines for metagenomic research]**. *Dong wu xue yan jiu = Zoological research* 2012, **33**(6):574-585.
28. Xin BC, Luo AH, Qin J, Paster BJ, Xu YL, Li YL, Yang DQ: **Microbial diversity in the oral cavity of healthy Chinese Han children**. *Oral diseases* 2013, **19**(4):401-405.
29. Luo AH, Yang DQ, Xin BC, Paster BJ, Qin J: **Microbial profiles in saliva from children with and without caries in mixed dentition**. *Oral diseases* 2012, **18**(6):595-601.
30. Xu H, Hao W, Zhou Q, Wang W, Xia Z, Liu C, Chen X, Qin M, Chen F: **Plaque bacterial microbiome diversity in children younger than 30 months with or without caries prior to eruption of second primary molars**. *PloS one* 2014, **9**(2):e89269.
31. Zhou J, Jiang N, Wang S, Hu X, Jiao K, He X, Li Z, Wang J: **Exploration of Human Salivary Microbiomes–Insights into the Novel Characteristics of Microbial Community Structure in Caries and Caries-Free Subjects**. *PloS one* 2016, **11**(1):e0147039.
32. Xu Y, Jia YH, Chen L, Huang WM, Yang DQ: **Metagenomic analysis of oral microbiome in young children aged 6-8 years living in a rural isolated Chinese province**. *Oral diseases* 2018, **24**(6):1115-1125.
33. Ling Z, Kong J, Jia P, Wei C, Wang Y, Pan Z, Huang W, Li L, Chen H, Xiang C: **Analysis of oral microbiota in children with dental caries by PCR-DGGE and barcoded pyrosequencing**. *Microbial ecology* 2010, **60**(3):677-690.
34. Nakayama K: **Molecular genetics of Porphyromonas gingivalis: gingipains and other virulence factors**. *Current protein & peptide science* 2003, **4**(6):389-395.
35. Marsh PD: **Dental plaque as a biofilm and a microbial community - implications for health and disease**. *BMC oral health* 2006, **6 Suppl 1**:S14.
36. Selwitz RH, Ismail AI, Pitts NB: **Dental caries**. *Lancet (London, England)* 2007, **369**(9555):51-59.
37. Takahashi N, Nyvad B: **The role of bacteria in the caries process: ecological perspectives**. *Journal of dental research* 2011, **90**(3):294-303.
38. Peyyala R, Ebersole JL: **Multispecies biofilms and host responses: "discriminating the trees from the forest"**. *Cytokine* 2013, **61**(1):15-25.
39. Mark Welch JL, Rossetti BJ, Rieken CW, Dewhirst FE, Borisy GG: **Biogeography of a human oral microbiome at the micron scale**. *Proceedings of the National Academy of Sciences of the United States of America* 2016, **113**(6):E791-800.

40. Zaura E, Keijser BJ, Huse SM, Crielaard W: **Defining the healthy "core microbiome" of oral microbial communities**. *BMC microbiology* 2009, **9**:259.
41. Turnbaugh PJ, Gordon JI: **The core gut microbiome, energy balance and obesity**. *The Journal of physiology* 2009, **587**(Pt 17):4153-4158.
42. Huse SM, Ye Y, Zhou Y, Fodor AA: **A core human microbiome as viewed through 16S rRNA sequence clusters**. *PloS one* 2012, **7**(6):e34242.
43. Li K, Bihan M, Methe BA: **Analyses of the stability and core taxonomic memberships of the human microbiome**. *PloS one* 2013, **8**(5):e63139.
44. Yan G, Huang W, Xue H, Jia Y, Yang D: **[Relationship between dental caries and salivary proteome by electrospray ionization ion-trap tandem mass spectrometry in children aged 6 to 8 years]**. *Hua xi kou qiang yi xue za zhi = Huaxi kouqiang yixue zazhi = West China journal of stomatology* 2014, **32**(3):297-302.
45. Fine DH, Toruner GA, Velliyagounder K, Sampathkumar V, Godboley D, Furgang D: **A lactotransferrin single nucleotide polymorphism demonstrates biological activity that can reduce susceptibility to caries**. *Infection and immunity* 2013, **81**(5):1596-1605.
46. Hong SW, Seo DG, Baik JE, Cho K, Yun CH, Han SH: **Differential profiles of salivary proteins with affinity to Streptococcus mutans lipoteichoic acid in caries-free and caries-positive human subjects**. *Molecular oral microbiology* 2014, **29**(5):208-218.
47. Ballal V, Rao S, Bagheri A, Bhat V, Attin T, Zehnder M: **MMP-9 in Dentinal Fluid Correlates with Caries Lesion Depth**. *Caries research* 2017, **51**(5):460-465.
48. Moslemi M, Sattari M, Kooshki F, Fotuhi F, Modarresi N, Khalili Sadrabad Z, Shadkar MS: **Relationship of Salivary Lactoferrin and Lysozyme Concentrations with Early Childhood Caries**. *Journal of dental research, dental clinics, dental prospects* 2015, **9**(2):109-114.
49. Jenssen H, Hancock RE: **Antimicrobial properties of lactoferrin**. *Biochimie* 2009, **91**(1):19-29.
50. Mazzoni A, Tjaderhane L, Checchi V, Di Lenarda R, Salo T, Tay FR, Pashley DH, Breschi L: **Role of dentin MMPs in caries progression and bond stability**. *Journal of dental research* 2015, **94**(2):241-251.
51. Frenkel ES, Ribbeck K: **Salivary mucins protect surfaces from colonization by cariogenic bacteria**. *Applied and environmental microbiology* 2015, **81**(1):332-338.
52. Siqueira WL, Zhang W, Helmerhorst EJ, Gygi SP, Oppenheim FG: **Identification of protein components in in vivo human acquired enamel pellicle using LC-ESI-MS/MS**. *Journal of proteome research* 2007, **6**(6):2152-2160.

Figures

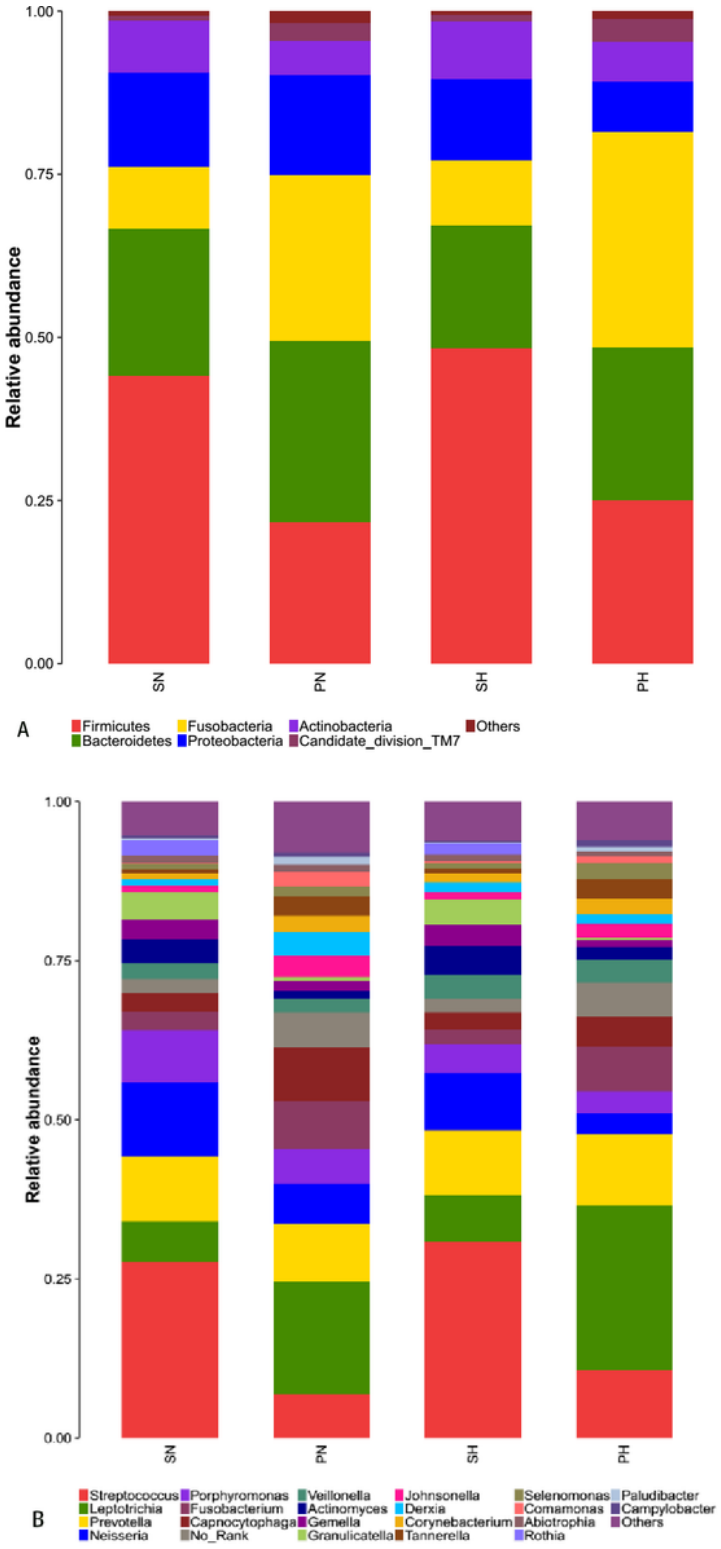
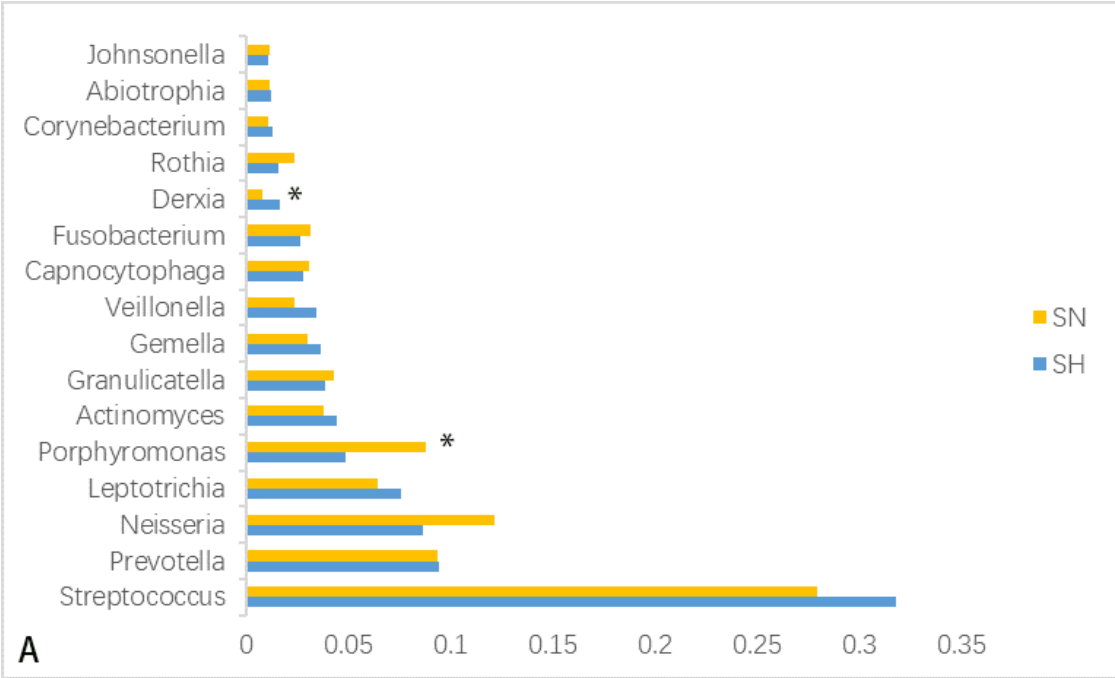


Figure 1

The Distributions of Predominant Taxa at the Phylum and Genus Level in Each Group. The distributions of predominant taxa (relative abundance >2% on average) on (A)phylum and (B)genus level.



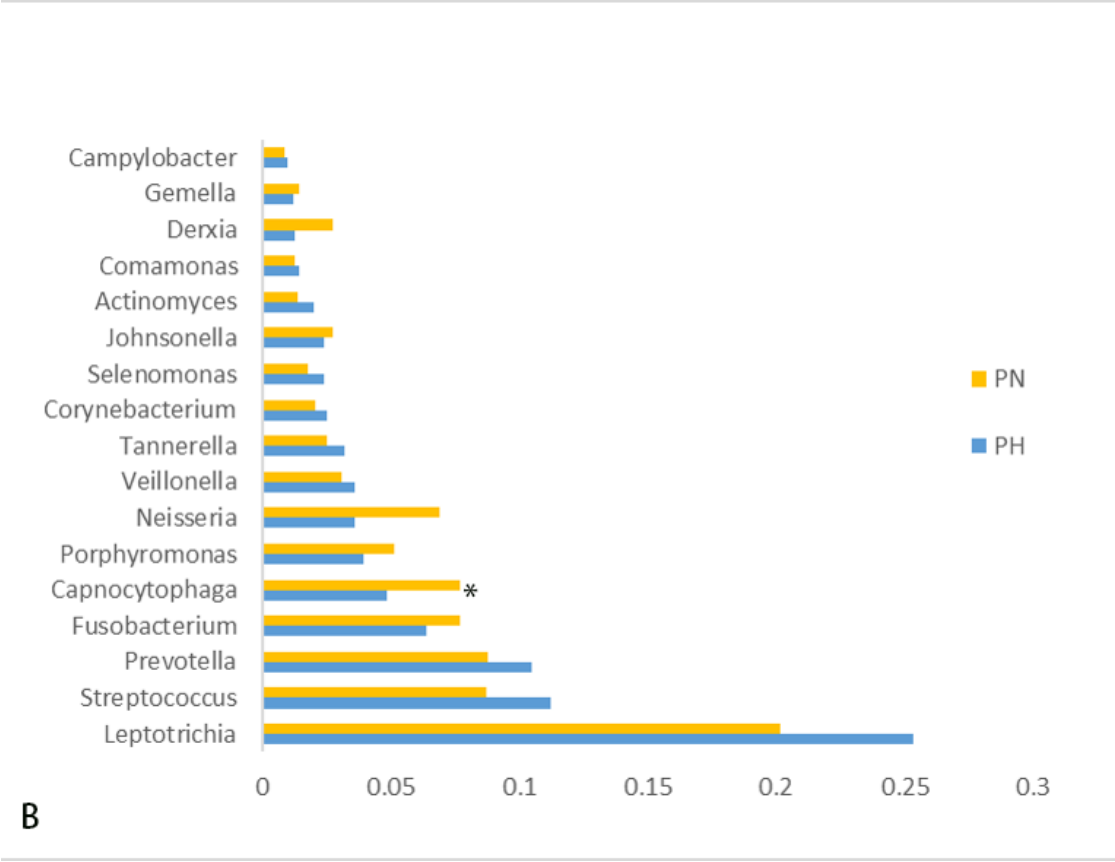


Figure 2

The Relative Abundance Comparison of the Predominant Bacteria at the Genus Level. Wilcoxon rank-sum test analyzes the difference. * represents a significant difference ($P<0.05$)

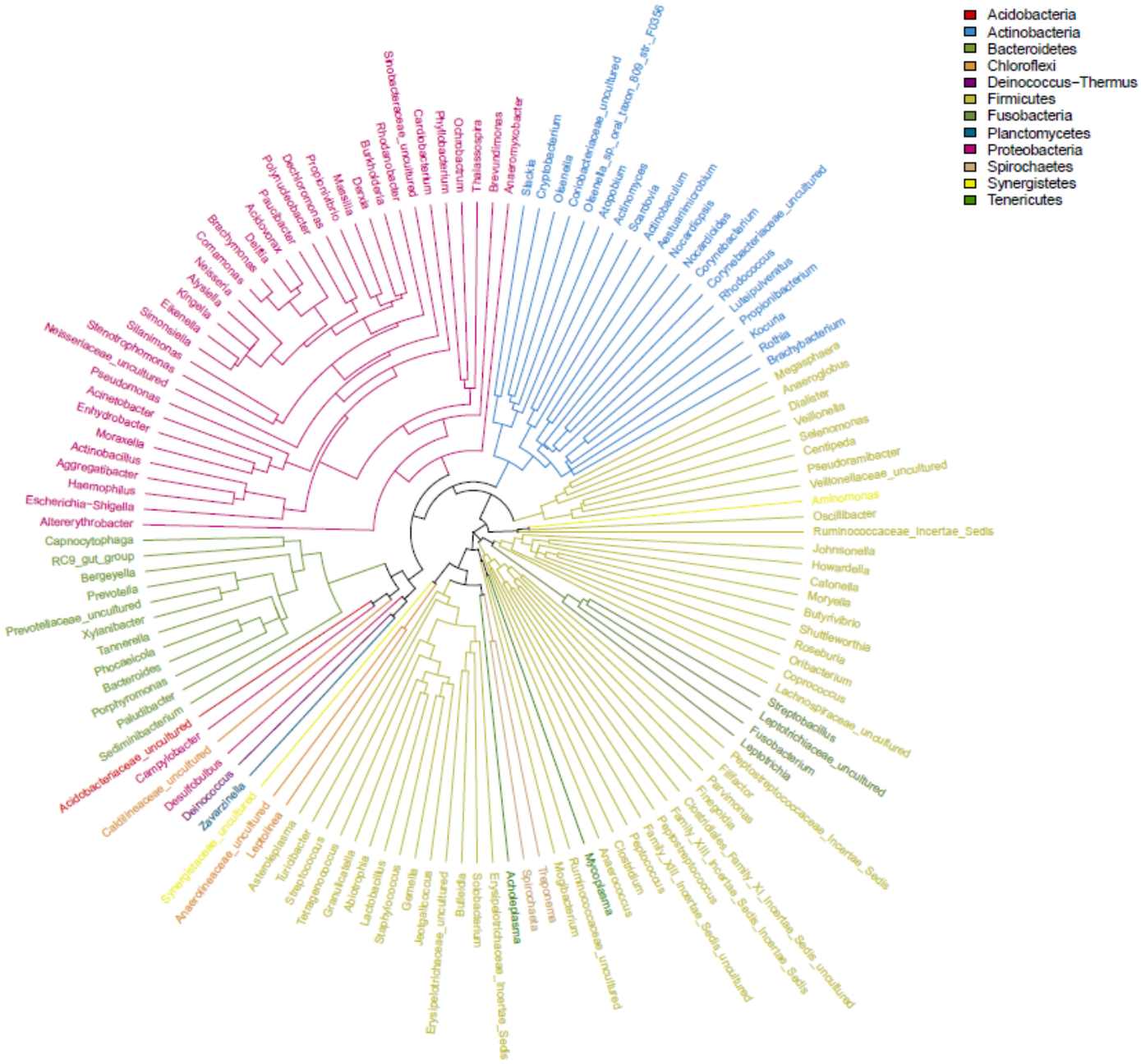


Figure 3

Circular Phylogenetic Tree at the Level of the Total 133 Genera. The tree was generated with Fasttree .The color-coded branches from the inner to the outer circles corresponds to the taxa from the phylum to genus level based 12 abundant phylum.

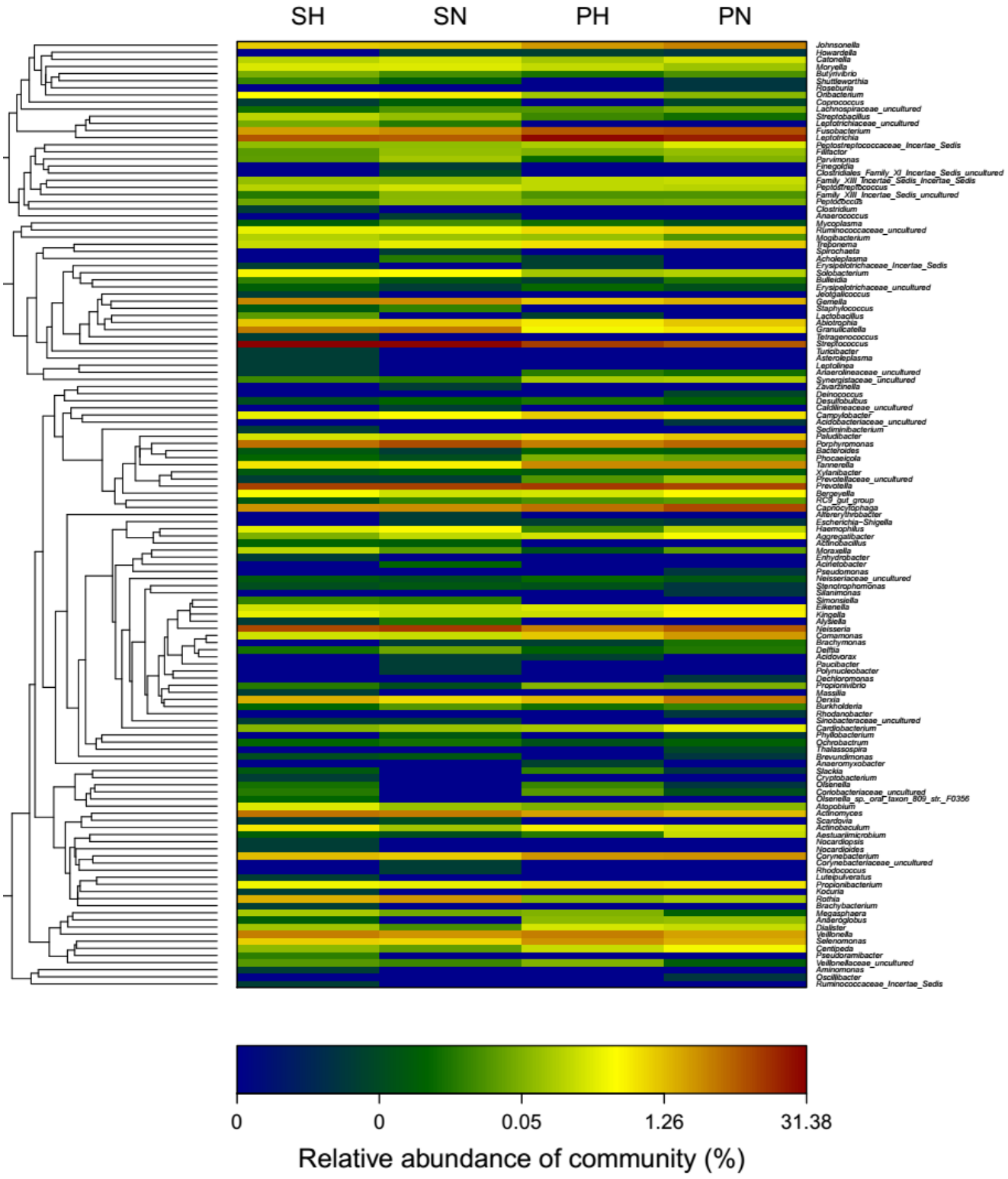


Figure 4

Relative Abundance of the 133 bacterial Genera in Each Group. Each column represents one genus of the groups. The relative abundance (%) is indicated according to the color scale at the bottom of the plot.

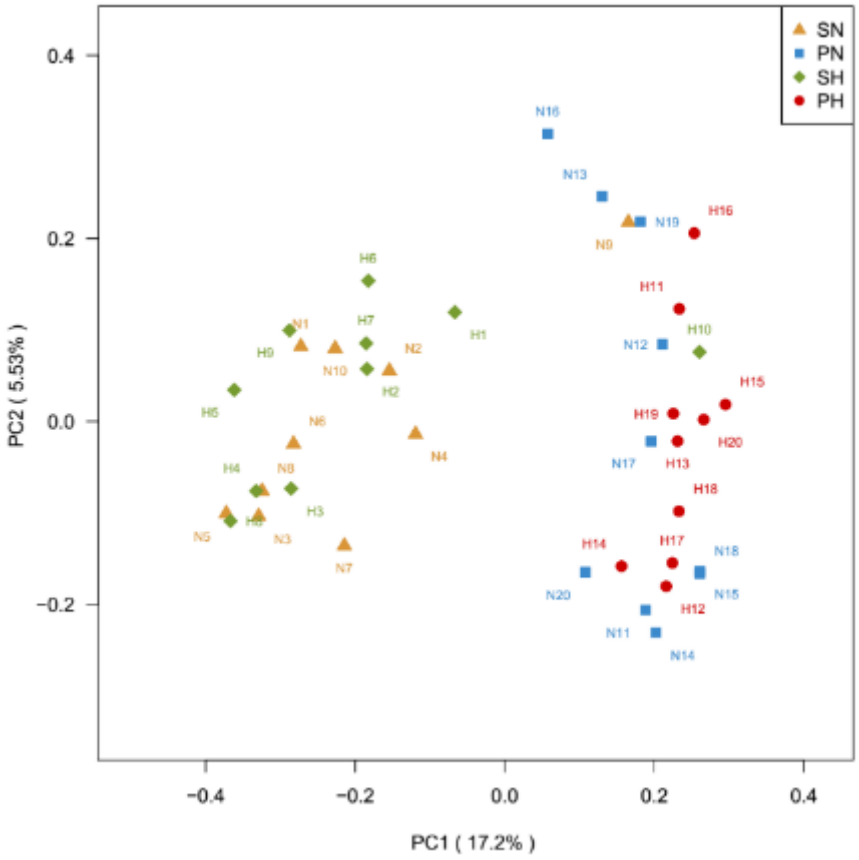


Figure 5

PCoA based on the Bray-Curtis Distances at OUT level with 97%.A dot represents each sample. Orange triangles represent the SN samples. Green rhombus represents the PN samples. Blue squares represent the PN samples. Red balls represent the PH samples. PC1 explained 17.2% of the variation observed, while PC2 explained 5.53% of the variation.

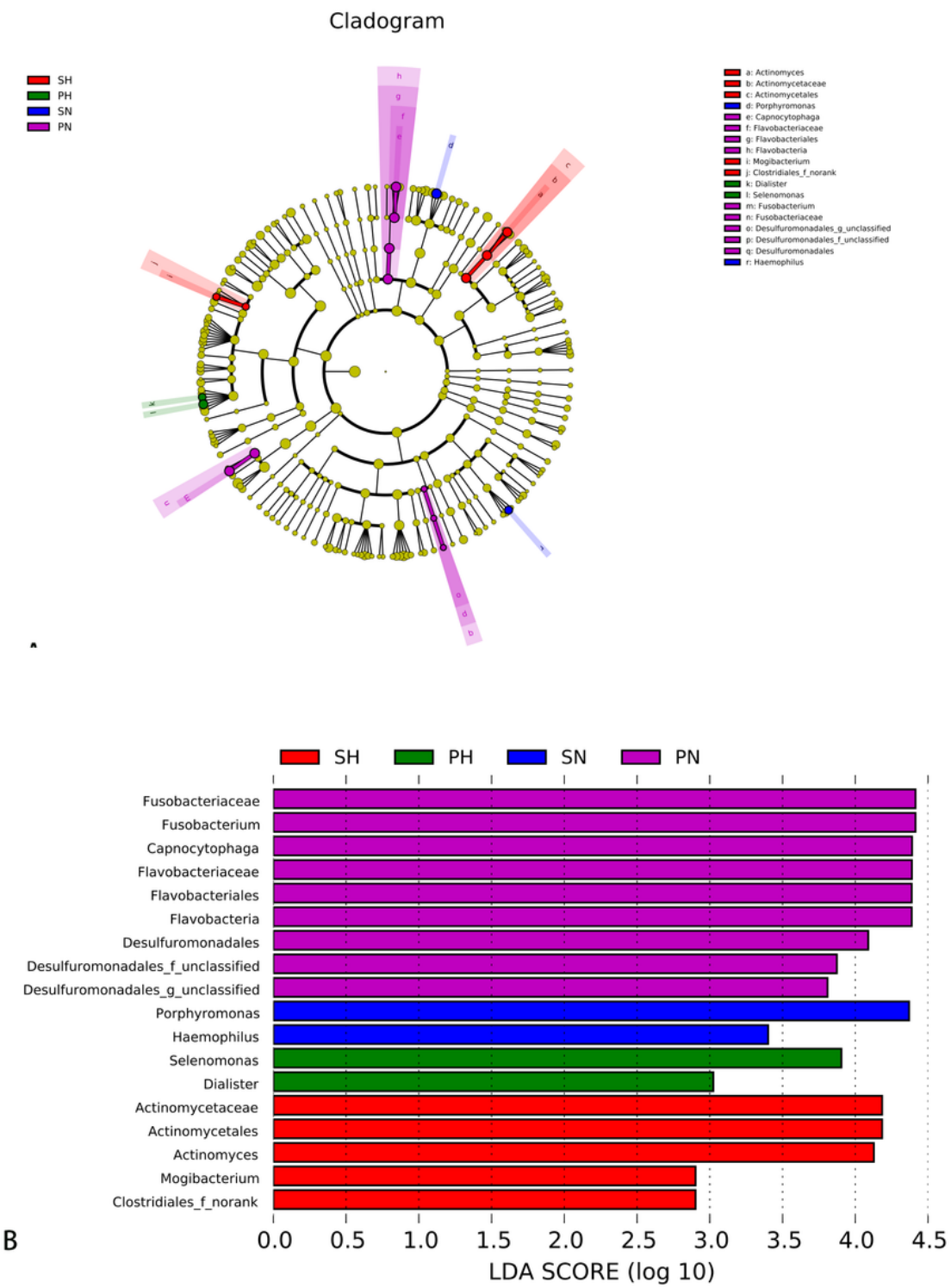


Figure 6

LEfse Analysis Based on LDA Shows the Microbial Variations (A) Cladogram for the taxonomic representation of the significant differences among the four groups. The colored nodes from the inner to the outer circles represent taxa from the phylum to genus level. The colorized point represents the different taxa with more a significant role. (B) Histogram of the LDA scores: the higher the LDA score, the

greater the species impact on the different effects. The threshold on the logarithmic LDA score for discriminative features was set at 2.0.

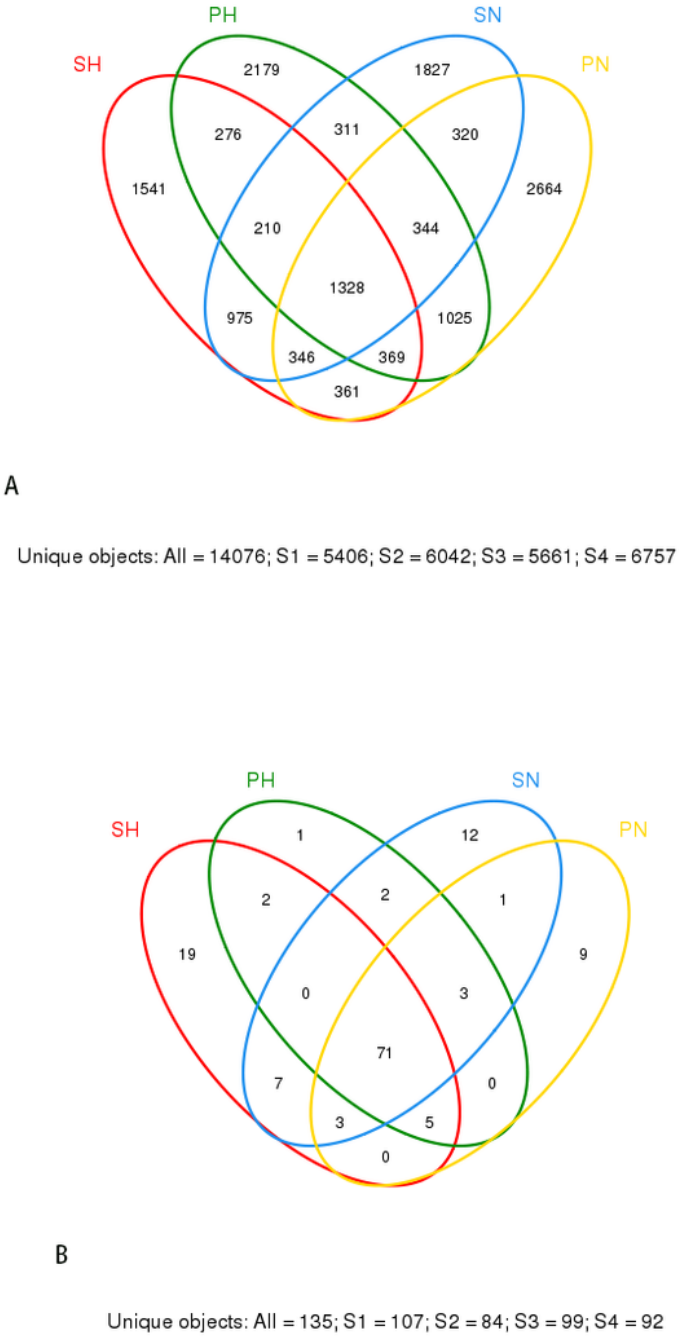
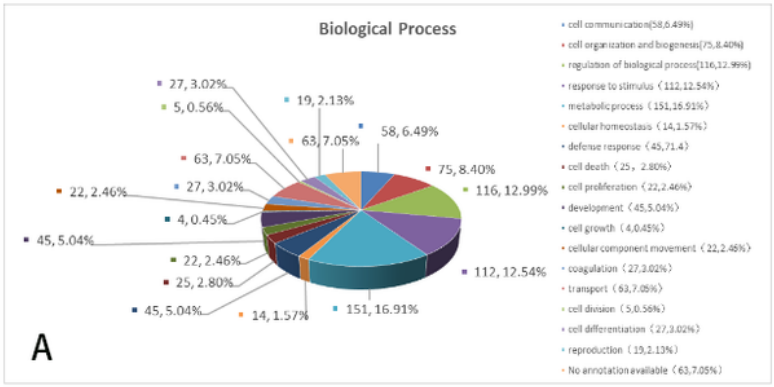
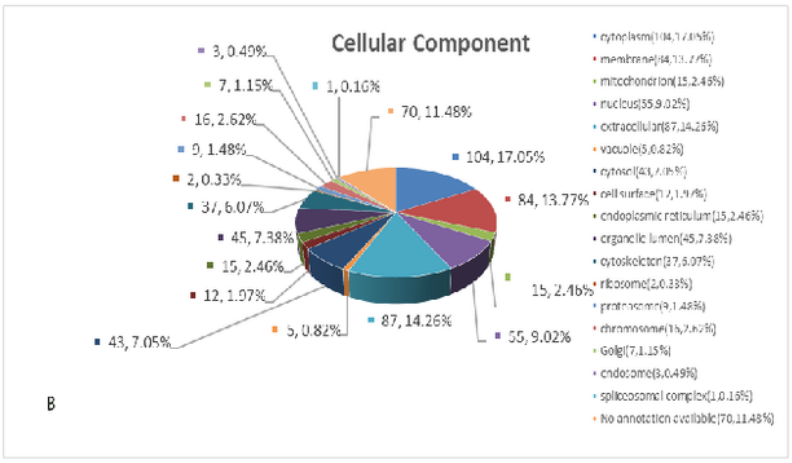


Figure 7

A Venn Diagram Showing Shared (A) OTUs and (B) genera with 97% Identity





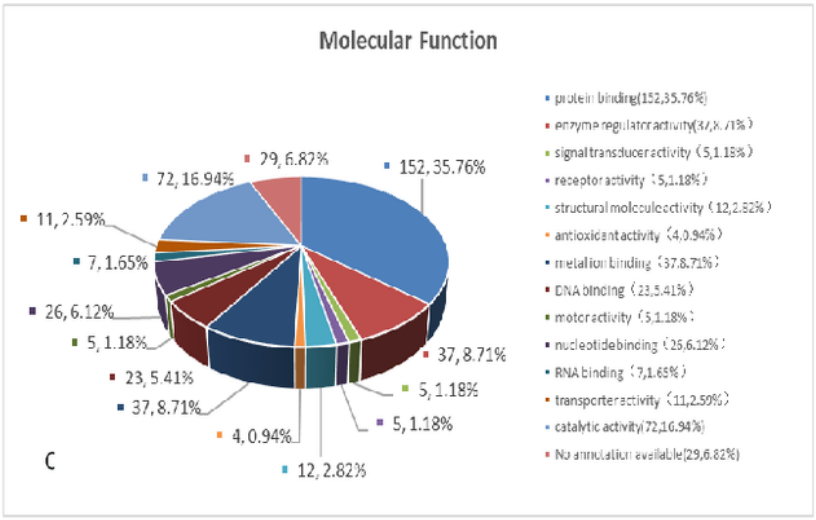


Figure 8

Gene Ontology Analysis of the Differentially Expressed Proteins.(A)biological processes (B) cellular components (C) molecular function

Supplementary Files

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

- supplement1.docx
- supplement1.png
- supplement3.png
- supplement4.png
- supplement4.docx
- supplement4.docx
- supplement7.docx
- supplement8.docx
- supplement9.xlsx
- supplement10.png