Current Microbiology (2021) 78:1730–1740
https://doi.org/10.1007/s00284-021-02400-4

Determination of Microbial Diversity and Community Composition in Unfermented and Fermented Washing Rice Water by High-Throughput Sequencing

Youlin Chen[1] · Haiming Chen[1] · Qiuping Zhong[1] · Yong-Huan Yun[1] · Weijun Chen[1,2] · Wenxue Chen[1]

Received: 3 September 2020 / Accepted: 10 February 2021 / Published online: 11 March 2021

Abstract

Washing rice water (WRW) refers to the sewage produced by rice washing in China and other parts of Asia people's daily life. As in the WRW is rich a variety of nutrients, microorganisms are prone to multiply and pollute the environment. In this article, high-throughput sequencing is used to describe the microbial diversity in different fermentation time WRW. The results showed that the sequencing depth effectively covered the microbial species in the samples, and the bacterial community structure in the samples of WRW at different fermentation periods was rich in diversity. Preominant taxa included *Proteobacteria* (62%), *Firmicutes* (28%), approximately *Cyanobacteria* (10%) and *Bacteroidetes* (0.5%). The core WRW microbiome comprises *Trabulsiella, Pseudomonas, Serratia, Lactobacillus, Erwinia, Enterobacter, Clostridium* and *Acinetobacter*, some of which are potential beneficial microbes. The change of microbial community composition with the change of habitat was assessed. It was found that environmental factors had significant influence on the assembly structure of microbial community.

Introduction

Rice is normally washed before cooking. People in China and other parts of Asia have been washing heavy oil dishes with 'milky' water produced by WRW [1, #1]. Recent studies have shown that the particles in WRW have the functions of decontamination, removing dandruff, moistening, brightening and promoting hair to turn black and thick [2, #202]. WRW mainly contains nutritional components of rice bran and aleurone, including starch, protein, oil content, dietary fiber, oryzanol, vitamins, minerals, tocopherols and other bioactive components. Rice bran protein has hypoallergenic, anti-tumor activity and healthcare effect [3, #20], The opioid antagonistic activity of rice bran proteolytic products through the study of rice bran CO_2 supercritical extraction, showed significant hair growth effect, which is equivalent to 3% of hair growth effect of drug minoxidil [4, #28]. The lipids of rice bran can reduce human serum lipid cholesterol [5, #21], and the polysaccharide of rice bran with anti-tumor effect [6, #22]. Consequently, the rich nutrient conditions provide a good environment for the growth of microorganisms.

Understanding the taxonomy, colony characteristics and functional components of WRW microbial community provides valuable information for guiding the isolation and cultivation of microorganisms, and it is likely to improve the ecological environment using the power of microbial community. By exploring the structural and functional genes, as well as the factors that determine the combination of microorganisms, in certain specific environments (gut, soil, ocean) and crop plant species, including rice [7, #51], millet [8, #269], corn [9, #215], citrus [7, #80] and grapevine [10, #161] the rhizosphere microbial community characteristics have made progress. Most studies on related microbial communities have been conducted by means of ribosomal amplicon-based approaches [11, #56]. However, up to now, the 16S rRNA gene has not been reported on the community characteristics of fermentation WRW.

Although it is believed that WRW has strong self-degradation and cleaning ability, when people pour a large amount

✉ Weijun Chen
chenwj@hainanu.edu.cn

✉ Wenxue Chen
hnchwx@163.com

[1] College of Food Sciences & Engineering, Hainan University, 58 People Road, Haikou 570228, China

[2] Chunguang Agro-Product Processing Institute, Wenchang 571333, China

of WRW directly into the sewer, the presence of a large number of microbial colonies will most likely result in great pressure on environmental management. However, some studies have been reported that sewage and other pollutants produced in the production process can be purified and degraded by the action of microorganisms [12, #201]. This degradation ability is most likely due to the large number of microorganisms present in the WRW. Nevertheless, it is still unclear the species composition of these communities, as well as their diversity and the evolutionary relationship between species. To design effective control measures, it is of great importance to better understand the microbial composition of WRW, mainly the community composition and abundance of bacteria, and the results of the evolutionary relationship between their species.

Rice is the staple food for about 50% of the world's population, of which nearly 90% is produced in Asia and consumed in developing countries such as Asia. At present, the world's rice planting area is about 155 million hectares, India's largest is about 44 million hectares, and China's rice planting area is 31 million hectares, accounting for about 20% of the world's planting area. China's total rice production ranks first in the world. Rice is also one of the most important food crops in China. 65% of the population take rice as their staple food. Rice plays an important role in China. Half of the world's rice exports will go to China.

Here, we present the results of this comparative study for the first time, and define the genome and core community composition of WRW from the high-throughput sequencing of community members, laying the foundation for the sustainable production of domestic sewage treatment by using microbial community.

Material and Methods

Preparation of WRW Samples

To represent the difference in fermentation time of organisms in the community of WRW, the samples were collected from rice consumed in people's daily life. Namely, the rice was washed with sterile water in the aseptic environment (rice:water = 1:5, soak for 10 min, stir for 3 s), and then filtered with a 100-mesh sieve to remove residual rice pieces slag. According to the principle of biodegradation technology, the pollutants in the WRW can be digested and metabolized by the microorganism. WRW was divided into equal parts and placed in a sterile fermentation tank for static fermentation at 25 °C. 20 mL of WRW fermentation broth was taken from 0 (T0), 2 (T2), 4 (T4), 6 (T6), and 8 (T8) days, three samples were taken from each fermentation period, that is, each time period was repeated three times, and 15 samples were collected in five different fermentation periods and centrifuged at 12,000 r/min for 10 min to discard the clear liquid, retaining the precipitate, and stored in the refrigerator at − 80 °C for future use.

DNA Extraction

Initially, DNA was extracted from each sample using the Fast DNA SPIN extraction kits (MP Biomedicals, Santa Ana, CA, USA), following the manufacturer's instructions with some modifications. The quantity of DNAs extracted were determined using a NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Waltham, MA, SA) and the quality of DNA extraction was determined by 1.2% agarose gel electrophoresis for microbiome samples of fermented WRW, respectively. DNA samples from three samples collected from the same fermentation period were pooled together and stored at − 80 °C until use. However, when quantifying DNA, this method resulted in low quality DNA ranging from 0 to 8 ng/μL. Therefore, genomic DNA was isolated from 0.5 g fermented WRW samples according to the protocol of Urukawa et al. [13, #282]. At the same time, the extraction reagent blank control was carried out.

16S rRNA Gene Amplification and MiSeq Sequencing

As part of related preliminary study, PCR was performed on the V3–V4 hypervariable region of the bacterial 16S rRNA gene, and the community DNA fragments were Paired-end sequenced by the Illumina MiSeq Seq platform. Almost full length 16S rRNA gene was amplified by PCR using forward primer 338F (5′-ACTCCTACGGGAGGCAGCA-3′) and the reverse primer 806R (5′-GGACTACHVGGGTWT CTAAT-3′) [14, #198]. The extracted DNA was amplified by two-step PCR, in the second step, with sample-specific 16-bp barcodes were added into the forward and reverse primers for multiplex sequencing. The PCR amplicons were purified with Agencourt AMPure Beads (Beckman Coulter, Indianapolis, IN) and quantified using the PicoGreen dsDNA Assay Kit (Invitrogen, Carlsbad, CA, USA). After the individual quantification step, the amplicons were collected in equal amounts, and Single Molecule Real Time (SMRT) sequencing technology was performed using the PacBio Sequel platform at Qingdao Personal Biotechnology Co., Ltd (Qingdao, China). To reduce the error rate of sequencing, we derived PacBio circular consensus sequencing (CCS) reads from multiple comparisons of sub-reads. In CCS, DNA polymerase reads a linked circular DNA template many times, which can effectively generate consistent sequence from multiple reads of a single molecule. The original sequence was initially processed by the PacBio SMRT Link portal (version 5.0.1.9585). The minimum prediction accuracy was 99% (minimum prediction accuracy = 3, minimum prediction accuracy = 99) after screening the sequence for at least

three times. 99% of the prediction accuracy is defined as CCS below this threshold is considered noise. Then, PacBio platform was used to fine tune the size of the amplicons to delete sequences larger than 2000 bp and generate files.

Sequence Analysis

As mentioned above, the sequencing data were processed by quantitative analysis of microbial ecology (QIIME2, 2019.4) pipeline [15, #32]. Simply, the original sequenced reads that exactly match the barcode were assigned to the corresponding samples and identified as valid sequences. After detection of chimeras, UCLUST [16, #37] clusters the remaining high-quality sequences into OTUs with 97% sequence consistency. A representative sequence was selected from each OTU using default the parameters. OTU taxonomic classification was conducted by BLAST searching the representative sequences set against the NCBI 16S rRNA Database using the best hit. The OTU table was further generated to record the abundance and classification of each OTU in each sample. In all samples, less than 0.001% of OTUs were discarded. To minimize the difference of sequencing depth among samples, 100 resampled OTU subsets were averaged below 90% of the minimum sequencing depth, and an average circular sparse OTU table was obtained for further analysis.

Amplicon Data Analysis

QIIME (2019.4) and R packages (v3.2.0) were used to analyze the sequence data. The OTU table in QIIME was used to calculate the α diversity index of OTU level, such as Chao1 richness index, Shannon diversity index and Simpson index. Box plot was drawn to compare the abundance and uniformity of OTU among samples. The unweighted UniFrac distance [17, #199] was used for β diversity analysis to study the structural changes of microbial communities in samples. Principal coordinate analysis (PCoA), non metric multidimensional scaling (NMDS) and unweighted arithmetic mean were used to visualize hierarchical clustering of group method (UPGMA) [18, #43]. The difference in unweighted UniFrac distances between paired comparisons was determined by Student's *t*-test and the Monte Carlo permutation test with 1000 permutations. Venn diagram was generated by using R package "VennDiagram" to visualize the shared and unique OTUs among samples or groups, regardless of their relative abundance [19, #48]. Metastats were used to statistically compare the abundance of taxa at the phylum, class, order, family, genus and species levels of samples or groups [20, #46], and the taxonomic flora was visualized as a bar chart. The aim was identified specific biomarkers of WRW at multiple classification levels, the bacterial abundance profiles of WRW samples at different fermentation periods were analyzed using linear discriminant analysis effect size (LEfSe) [21, #33]. In this analysis, we calculated the abundance spectrum of bacteria from phylum to genus at % taxonomic level. LEfSe analysis adopted Kruskal–Wallis test (alpha value 0.05) and LDA score > 2 as the thresholds. PICRUSt (Phylogenetic Investigation of Communities by Reconstruction of Unobserved States) was used to predict microbial functions and was consistent with the Kyoto Encyclopedia of Genes and Genomes (KEGG) database.

Additional Requirements

Nucleotide sequence accession numbers: The sequence data reported in this paper have been deposited in the NCBI database (Accession Numbers: PRJNA657930).

Results

Data Sets

A total of 25.159 million effective sequences were generated from 15 fermented WRW samples, with an average of 167,727 effective sequences per sample (ranging from 157,824 to 182,597). After data quality control, denoising, splicing, and mosaicism screening, 1,743,022 high-quality sequences (accounting for 69% of the total sequences) were observed, with an average of 116,201 sequences per sample (ranging from 100,716 to 129,652; as shown in Table 1). The average sequence length was 411 bp, with the maximum length of 441 bp, and the shortest was 236 bp (Fig. S1). All high-quality sequences were clustered by 97% homology, and 3769 OTUs were obtained. These OTUs were BLAST-searched against in the Greenenes database for classification and allocation. After deleting the OTUs with lower reliability (only 15.5% of all sequences), a modified OTU table is shown, which includes 3769 OTUs, with an average of 754 OTUs per sample (Table 1).

Bacterial Diversity Analysis

The average species diversity index of bacteria in each sample is measured by α diversity index. The α diversity index of Chao1, Observed species, Shannon, Simpson, Faith's PD, Pielou's evenness and Good's coverage are shown in Fig. 1. On the second day, the Shannon value was the highest and the Simpson value was the lowest, indicating that the bacteria diversity of the rice washing water fermentation liquid was the highest in the early stage of fermentation. Observed species index indicated that the number of species contained in the sample on the second day was the highest, and the higher the value was, the higher the species richness of the

Table 1 After removing the primer fragment of the sequence and discarding the sequence of the mismatched primer, Dada2 was called to process the feature sequence, such as quality control, de-noising, splicing, de-chimerism, etc.

Group	Input		Filtered		Denoised		Merged		Non-chimeric		Non-singleton	
	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE	Mean	SE
T0	182,597	14,398	165,673*	12,729	164,839*	127,312	164,178*	12,723	118,706	9652	118,528	9688
T2	157,824	14,539	139,283	12,817	138,159	12,382	136,937	11,659	129,652*	6314	129,460	6256*
T4	162,822	10,378	146,532	9203	145,661	9197	144,647	9090	100,716	7876	100,530	7882
T6	163,561	9680	146,734	10,151	145,959	10,269	145,119	10,267	107,121	20,877	106,970	20,903
T8	171,830	23,102	152,701	17,020	151,984	16,941	151,168	16,855	124,813*	10,599	124,680	10,559*

SE standard error

*Showed significant difference $P < 0.05$, Student's *t*-test

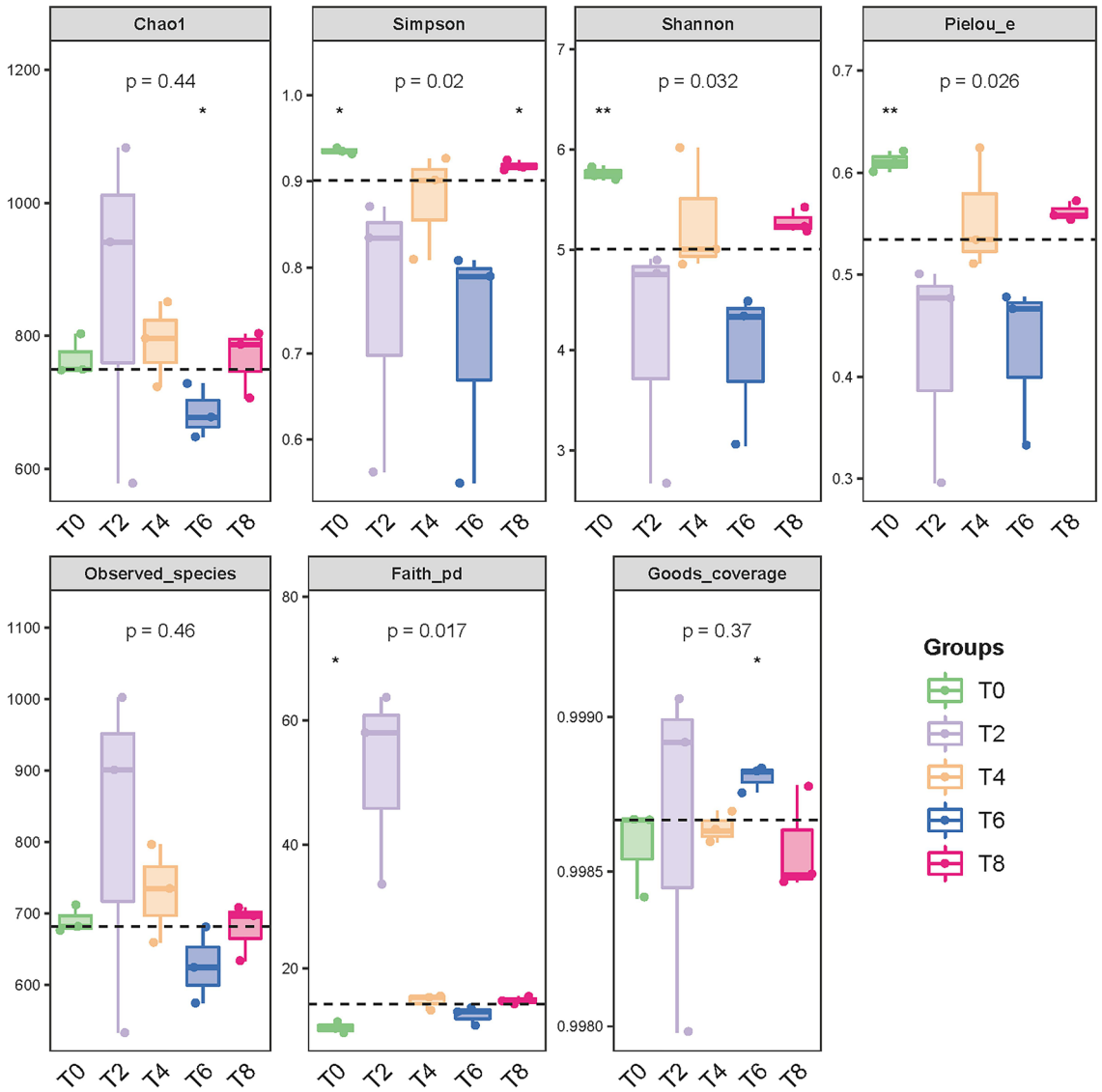

Fig. 1 Taxonomic distribution and diversity comparisons in unfermented and fermented WRW microbiomes. Alpha diversity comparison between the fermentation WRW samples from each different fermentation time based on the Shannon index using the 16S rRNA gene sequence data. Asterisk indicates significant differences

sample was. The differences between the T0 groups were extremely significant. However, the Chao1 richness index in T2 was significantly different in fermented WRW from T0, T4, T6, T8 ($P=0.44$), followed by Faith's PD diversity index was significantly different between them ($P=0.017$). This indicates that the bacterial diversity of pre-fermentation WRW T2 was higher compared with other groups. Good's estimated coverage of over 99% indicates that the current sequencing depth is sufficient to saturate the bacterial diversity of fermentation WRW. In addition, Pielou's evenness index showed that the distribution of the bacterial flora in the WRW samples during the five fermentation periods was very uneven.

Taxonomic Characteristics of the Microbial Community

The distribution of bacteria was characterized based on relative taxonomic abundance. The bacterial sequence mainly included phylum, class, order, family, genus and so on. The taxonomic distribution of dominant bacteria at different levels are provided in Fig. 2 (relative abundance > 0.003%). With respect to the actual abundance, *Proteobacteria* (accounting for 62% of the total sequence) was the most abundant phyla followed by *Firmicutes* (28%), *Cyanobacteria* (about 10%), *Bacteroidetes* (0.5%), *Actinobacteria* (0.1%), accounting for more than 99% of the total sequence. The most common genera were *Clostridium* (25%), *Erwinia* (5%), *Acinetobacter* (3%), *Enterbacter* (2%), *Cronobacter* (2.0%), *unclassified mitochondria* (2%), *Serratia* (1%), representing 92.3% of the total sequences (Fig. S2).

Bacterial Community Structure

The taxonomic characteristics of unfermented and fermented WRW microbial communities were studied. The bacterial diversity of the sample group was significantly different on the second day of fermentation (T2), and the community composition of the sample group in the late fermentation period varied with the change of nutrient composition (as shown in Fig. 3a, $P<0.05$, t test). Obviously, the number of species of WRW bacteria is often very large in T2, and there is a significant difference between the sample community and other sample communities (Fig. 3b). This conclusion laid a theoretical foundation for future researches. The clustering effect was measured by the branch length of the clustering tree, it was found that WRW samples on the second day of fermentation had significant similarity with each sample group, but the similarity between other groups was not significant. (Fig. 3S).

Marked Species Analysis

LEfSe and PLS-DA analysis were used to further examine these significant differences. LEfSe analysis was carried out with the collected data to determine the specific taxon, which can be used as a biomarker because of their quantity changing with fermentation time. With an LDA score > 2, a total of 17 genera were identified (Fig. 4a). In Fig. 4b that family and genus level abundance are shown by using a clado-gram. The linear discriminant analysis (LEfSe) of the microbial community dominated by *Clostridium, Erwinia, Acinetobacter, Enterobacter, Serratia, Cronobacter* showed that there are 52 species-specific biomarkers in 5 different fermentation periods of WRW. The sequences generated by the primers of each sample were collected by us into a pool and combined with samples from five different fermentation stages (Fig. 4b). A total of 17 genera were assessed, and the relative abundance of all genera in all samples exceeds 0.1%. 11 of the 17 genus belong to the genus *Proteobacteria*, representing 7 different families. 3 genera and 1 genus were assigned to *Firmicutes* and *Cyanobacteria*, respectively, and

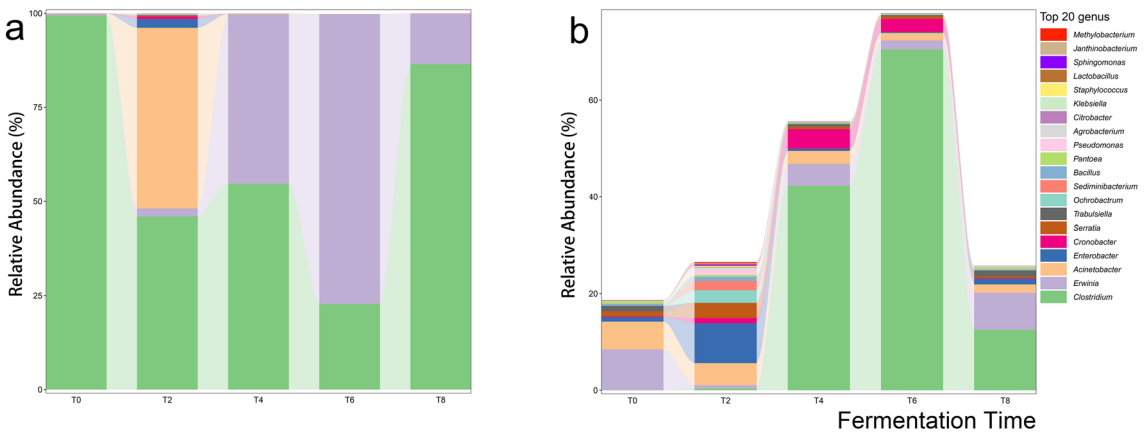

Fig. 2 Distribution of major bacteria at different taxonomic levels (phylum, genus). The predominant taxa; each level ranks in the top 20 (relatively abundant 0.003%)

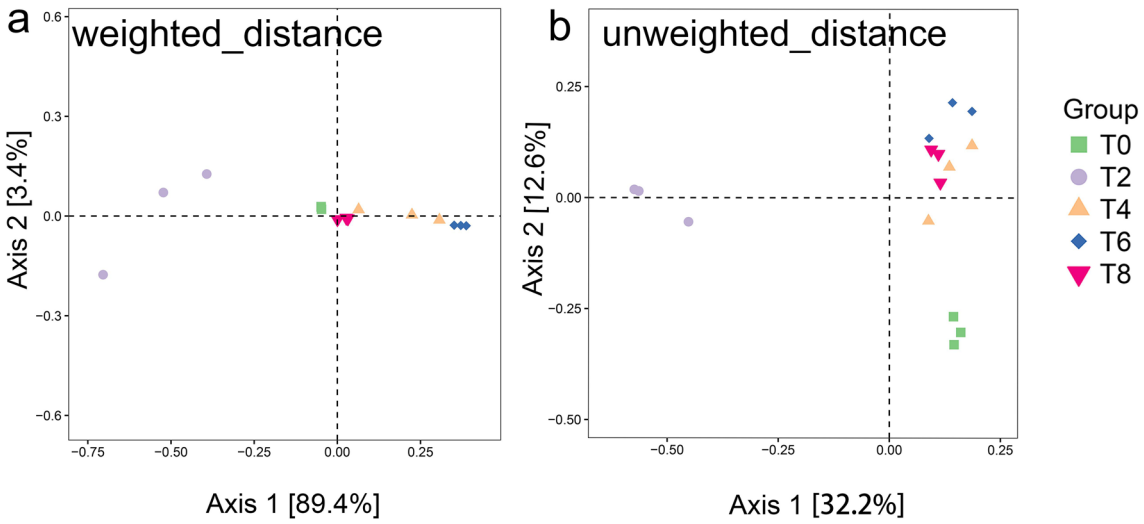

Fig. 3 **a**, **b** PCoA based on the weighted UniFrac distance and unweighted UniFrac distance between unfermented and fermented WRW for each different fermentation time using the 16S rRNA gene sequence data

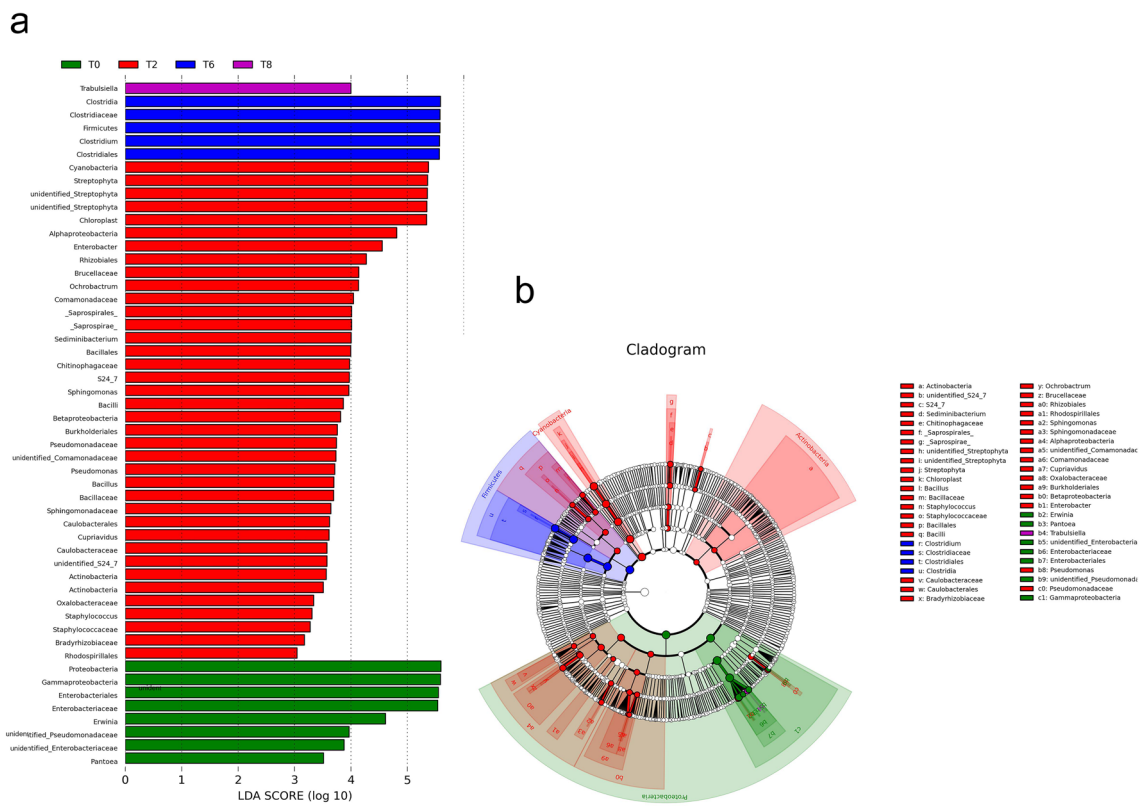

Fig. 4 **a** The ordinate is the taxon with significant difference between groups, and the abscissa is a bar chart to visually show the LDA log score of each taxon. **b** Taxonomic branching maps show the taxonomic hierarchy of the main taxon in the sample community from phylum to genus (from inner to outer). The node size corresponds to the average relative abundance of the taxon. The hollow nodes represent taxa with insignificant inter-group differences

the rest were defined to be other different genera. There were 11 and 4 biomarkers in T2 and T0, respectively. T4 had no specific biomarker, while in T6 and T8 were only 1 different biomarker.

Core Microorganisms and Function Predictions

Our ultimate goal was to test the existence of an identifiable common core microbial community [22, #151], a core was defined as the group of members shared by the microbial community in the Venn diagram, and the core was represented with overlapping withs 97% consistency. 1004, 1791, 1206, 998 and 1003 OTUs in T0, T2, T4, T6 and T8 groups were generated, respectively. As shown in Fig. 5, 201 OTUs in 5 groups were identified, accounting for 5.33% of all OTUs (3769 OTUs), and the total OTUs abundances approximately 5.34%.

In the Venn diagram, the unique OTUs in each group was also observed. 355, 1451, 374, 326 and 309 unique OTUs were found in the T0, T2, T4, T6 and T8 groups, respectively. The top 10 phyla and genera with the highest mean OTU frequency were selected for display (Fig. 4S). Among the five groups, there were four phyla with the same OTU frequency, which were *Proteobacteria, Firmicutes, Bacteroidetes* and *Cyanobacteria*, with eight genera, which belonged to *Acetobacter, Clostridium, Enterobacter, Erwinia, Latobacillus, Serratia, Pseudomonas, Trabulsiella*. In the five groups, the OTU sequences with the same percentage of abundance contained four phyla (*Proteobacteria, Firmicutes, Bacteroidetes* and *Cyanobacteria*) and five genera (*Clostridium, Acinetobacter, Enterobacter, Erwinia* and *Semiminiature*). According to the 16S rRNA composition data of each sample, PICRUSt analysis was performed. The KOs were mainly involved in 33 KEGG level 2 pathways (Fig. 5S). A total of 33 metabolic functions were predicted,

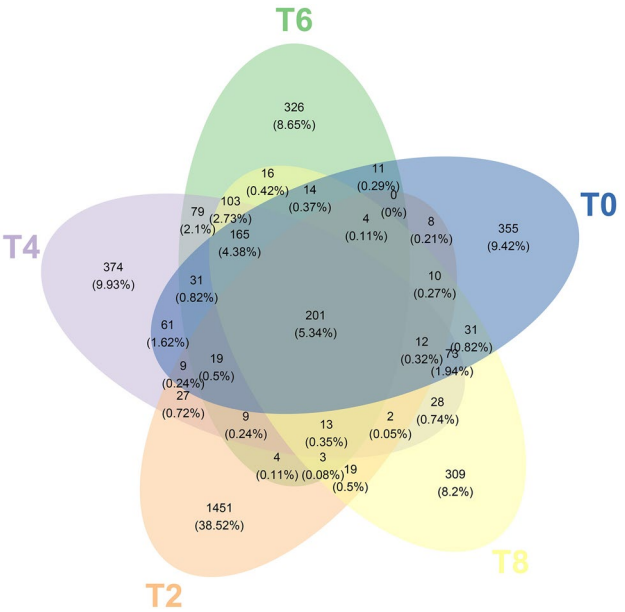

Fig. 5 Venn diagram analysis. Venn diagram shows the shared and unique OTUs at 97% identity in the five groups

carbohydrate metabolism (15.4%), Metabolism of cofactors and vitamins (12.1%), amino acid metabolism (10.9%), Metabolism of terpenoids and polyketides (9.7%), Metabolism of other amino acids (8.1%) were the most abundant. We also studied bacterial functions related to human diseases, such as cancer, infectious diseases, cardiovascular diseases, neurodegenerative diseases, and immune system diseases.

Discussion

In this study, the diversity, taxonomic composition and community abundance of microbial communities in WRW were studied for better and more comprehensive realizing of microbial communities characteristics. In this paper, the microorganisms in the fermented WRW were detected and analyzed, and the characteristics of biodiversity and bacterial abundance in different environments were determined.

16S rRNA amplicon sequencing has been widely used in the classification of various microbial communities. Previously researches on microbial communities by amplification sequencing [23, #62; 24, #164] were mainly focused on soil, rhizosphere, intestine, mouth and ocean. New insights into the composition and structure of microbial communities have been provided through the introduction of high-throughput sequencing technology. The bacterial 16S rRNA gene consists of 9 hypervariable regions, and different microbial diversity was usually represented by sequences generated using different combinations of these regions. Here, the bacterial diversity and community structure were explorated, which is to use high-throughput sequencing to sequence the WRW at five different fermentation periods from the highly variable V3–V4 region of 16S rRNA. Compared with previous studies used basic molecular techniques such as DGGE, qPCR and microarray chips [25, #165; 26, #166] is different in that we not only identified the major taxa, but also identified low abundance or rare taxa, such as *Methylobacterium, Janthinobacterium, Sphinomonas, Lactobacillus, Staphylococcus*. The V3–V4 region has been used in some studies [27, #176; 28, #144], as it can provide higher phylogenetic resolution and better diversity and richness. In addition, more sequences of this region were stored in the database, which greatly facilitated diversity analysis [29, #167].

1,743,022 high-quality sequences were obtained with an average of 116,201 per sample sequence (Table 1), which was much higher than previous studies on oral microbes [30, #169; 31, #118] for 16S rRNA sequencing studies. Approximately 1000 sequences was covered the depth of each sample, which provided a wonderful balance between sample size and sampling depth and was able to detect species with reasonable accuracy at 1% abundance. Therefore,

the sequencing depth of our study (116,201 sequences per sample) was reasonable and large enough to detect the vast majority of bacterial species in fermented WRW samples. Furthermore, Good's coverage also indicated that our sequencing depth was sufficient to reflect the total bacterial diversity of the WRW flora, with an estimated value of > 99% (Fig. 1), including some extremely low rare species richness [32, #171].

In this study, the sequences of *Proteobacteria* and *Firmicutes* were the main microbial communities in the fermented WRW samples. Together, these phyla accounted for more than 90% of the total number of microorganisms detected. Similar proportions have been highlighted in other early studies of water [33, #283].

According to PCoA and hierarchical cluster analysis, there is a similar community structure among these five groups of samples (Figs. 3 and 5S). Samples from the T4, T6, and T8 groups were gathered together, while the T0 and T2 groups appeared to be more variable. Although there are similar bacterial members, the abundance of some bacteria differs significantly between groups. In the early stage (T0), the sample itself carries a certain amount of microorganisms, which were enriched in nutrients and moderate pH content. In the middle stage (T2), a large number of microorganisms grow, resulting in increased species richness. In the late fermentation period (T8), due to poor nutrients and low pH, the microbial flora biosynthesizes some nutrients, and the reproduction of most microorganisms is suppressed. It can be inferred that the bacterial composition in the late stage of fermentation is obviously affected by environmental factors. This is consistent with the results of other studies on fermentation materials [34, #177].

At the genus level, *Enterobacterium* and *Clostridium* were the main species in the early stage of fermentation, which grew in the medium with low carbohydrate and high protein. At the beginning of WRW fermentation, it is characterized by the increase of short chain fatty acid production and butyric acid proportion, which corresponds to the increase of donor dependent proportion of *Bacteroides* and *Firmicutes*, and it has been intensively enriched in the community of fermentation (T2) [27, #176]. With the increase of the number of bacteria and the concentration of signal molecules, the expression of virulence factors, mucopolysaccharides and other related genes can be activated. Due to the antagonistic effect of dominant bacteria, the number of some bacteria decreased gradually. Therefore, the ecological balance of bacterial flora may be destroyed during the whole fermentation process, leading to the formation of some disease substances [35, #180].

This observation supports that the difference in the species composition between samples (groups) does not only indicates the difference in the composition of all species, but also the difference in the distribution of some components. They may have already manifested significant differences at phylum level.

The existence of "core microbiome" was first proposed [36, #184], referring to the organisms, genes or functions shared by all or most individuals in specific human habitat, such as the mouth, nose, skin and intestine. One of the objectives of the Human Microbiome Project is to determine whether there is an identifiable core microbiome between individuals. Most studies have focused on the human gut and investigate the relationship between diseases and the core gut microbiome such as obesity, metabolic syndrome and diabetes [37, #185; 38, #182; 39, #186]. 16S rRNA sequencing technology was used to conduct an in-depth study of the citrus rhizosphere core microbiome, and 1620 overlapping OTUs were found from 23 sites in 8 citrus producing countries [39, #186]. This study has provided 201 overlapping OTUs in the five groups of Venn diagram (Fig. 5), which are defined as the core microbial groups of WRW samples, regardless of their abundance. These widely distributed core microbial communities may play an important role in the stability and function of fermentation WRW. According to the PICRUSt analysis, we predicted that the functions of bacteria mainly focused on carbohydrate metabolism, amino acid metabolism, metabolism of cofactors and vitamins, metabolism of other amino acids, metabolism of terpenoids polyketides, membrane transport, replication and repair, cell mobility and energy metabolism, etc. As suggested by [40, #181], starting from the Venn diagram representing members, the core microbial community is defined by combining composition, phylogeny, persistence and connectivity, which increased the complexity layer. The proper definition of the core microbiota depends on the ecological issues being addressed. Our research is mainly to determine the existence of the core microbial community, so we usually define the core as a group of members shared in the microbial community, and the overlapping circle area was expressed in the Venn diagram. In the future, we wouled be apply a variety of definitions to the core microbiota enabled us to better enhance the ecological significance.

There were significant differences in the microbial classification of unfermented WRW samples in different fermentation periods. However, a given bacterial genotype obviously selects a specific core microbial community [41, #82], which may play an important role in microbial assembly. Up to now, most of the core microbiome of plants have been identified by taxonomic markers [11, #268; 42, #60]. However, some have stressed that more attention should be paid to the identification of microorganisms with common functions in specific environmental factors, such microbial communities based on functional definitions should be conducive to manipulate communities for beneficial purposes. Focusing only on diversity may hinder further study of the research system. Therefore, the diversity of WRW was determined

by high-throughput sequencing technology to analyze in different fermentation periods. Based on the diversity information of bacterial community, such studies are still at a very early stage to understand the internal mechanism of bacterial community pattern. Some of the core microorganisms were found in this study overlapped with those previously found in Arabidopsis [43, #153], millet [44, #52], sugarcane [11, #268] and coolola [42, #60], suggesting that many environmental factors driving community aggregation may be common in ecosystem species. Furthermore, some core microorganisms associated with the WRW are beneficial. For example, P*seudomonas, Bradyrhizobium, Agrobacterium* and *Rhizobium* were found to inhibit plant diseases in different environments [45, #191; 46, #189]. The identification of microorganisms in WRW are provided a useful starting point for future research, and the characteristics of microbial colonies can be determined by using synthetic flora. Although we have defined core microbial communities according to the above criteria, some of them may be common in other habitats and may not necessarily be specific to WRW. Therefore, further experiments are needed to determine the specific microbial community of fermented WRW.

In most reports, although a great diversity of bacterial communities have been revealed, the dominant bacterial community in any given environment is relatively small. For example, in soil samples, only 2% of the bacterial groups were reported that account for nearly half of the soil bacteria in different locations around the world [23, #284]. The bulk soil and citrus rhizosphere samples were collected from different continents that only a few bacterial groups, such as *Proteobacteria, Acidobacteria, Actinobacteria* and *Bacteroidetes* are abundant, whether they are bulk soil or citrus rhizosphere [39, #186]. These abundant bacterial groups were also dominant members of the rhizosphere of other plants [7, #193; 47, #158]. However, we also discovered that the bacteria in the fermented WRW samples were statistically more abundant *Proteobacteria, Actinobacteria, Firmicutes, Bacteroidetes* and so on. Some groups of *Proteobacteria* and *Bacteroides* were enriched from the early fermentation stage to the post fermentation stage, while some autotrophic micro organs, such as *Cyanobacteria*, were gradually consumed in the microbial community of the samples. The enrichment of microorganisms in fermented WRW can be attributed to their lifestyles (for example, the rapid growth of symbiotic bacteria and the ability to utilize various C sources existing in fermented WRW) [48, #192]. Our study revealed that although there are different bacterial group components in different fermentation time, these components are key similarity at genus level, but they are different in quantity, which indicates host driven selection for specific traits. In fact, in this study, the taxonomy, abundance and diversity characteristics of the microbial community were determined in WRW, enabled us to better determine the microbial taxon and its nature in this habitat. The experimental basis is provided for the development of traditional fermented WRW in the daily chemical industry. Meanwhile, our results provide valuable information for guiding the isolation and cultivation of microorganisms, which can be useful to improve the ecological environment. At present, the diversity and functional characteristics of the bacteria in fermented WRW is rare reported.

Conclusion

In conclusion, 16S rRNA sequencing technology has greatly expanded our understanding of bacterial diversity and community structure of fermented WRW. We have observed a large number of colonies with 1,743,022 high-quality sequences and 3769 OTUs, with 97% homology, belonging to 41 phyla, 108 classes, 197 orders, 318 families, 501 genera and 595 species. The bacterial diversity of fermentation WRW was not affected by fermentation time, and the diversity richness decreased with the extension of fermentation time. Our results showed that the community structure was similar in different fermentation time. In five groups of 201 overlapping OTUs, an identifiable core microflora was identified. In addition, the findings of this study are novel, because the detection of unfermented WRW provides a powerful basis for the analysis of whether bacteria come from the water source of WRW or from the resident bacteria of rice itself. In this study, a more comprehensive description of bacteria can give a more comprehensive understanding of the whole community composition. The analysis of α diversity, β diversity and PCoA provided a theoretical basis for further study. So far, the phylogeny and interaction of Chinese WRW have not been thoroughly elucidated. This study provides valuable information for guiding the isolation and cultivation of microorganisms, and also provides the possibility for improving the ecological environment by using the power of microorganisms.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00284-021-02400-4.

Acknowledgements This study was supported by the Natural Science Foundation of Hainan Province of China (2019RC128; 2019RC120) and the Hainan University Start-up Scientific Research Projects of China (kyqd1630, kyqd1551).

Declarations

Conflict of interest All authors declare that they have no conflicts of interest.

References

1. Chen B, Hu Z, Li H, Li Z, Li Q, Chen Y (2019) Intact starch granules for pickering emulsion: exploring mechanism of cleaning with washing rice water and floury soup. Colloid Surf A 561:155–164. https://doi.org/10.1016/j.colsurfa.2018.10.045
2. Lam S, Velikov KP, Velev OD (2014) Pickering stabilization of foams and emulsions with particles of biological origin. Curr Opin Colloid Interface Sci. https://doi.org/10.1016/j.cocis.2014.07.003
3. Han S-W, Chee K-M, Cho S-J (2015) Nutritional quality of rice bran protein in comparison to animal and vegetable protein. Food Chem. https://doi.org/10.1016/j.foodchem.2014.09.127
4. Choi JS, Jeon MH, Moon WS, Moon JN, Cheon EJ, Kim JW, Jung SK, Ji YH, Son SW, Kim MR (2014) In vivo hair growth-promoting effect of rice bran extract prepared by supercritical carbon dioxide fluid. Biol Pharm Bull 37(1):44–53. https://doi.org/10.1248/bpb.b13-00528
5. Mendel F (2013) Rice brans, rice bran oils, and rice hulls: composition, food and industrial uses, and bioactivities in humans, animals, and cells. J Agric Food Chem. https://doi.org/10.1021/jf403635v
6. Li W, Xiaoxuan L, Zhengxing C (2009) Sulfated modification of the polysaccharides obtained from defatted rice bran and their antitumor activities. Int J Biol Macromol. https://doi.org/10.1016/j.ijbiomac.2008.12.006
7. Edwards J, Johnson C, Santosmedellin C, Lurie E, Podishetty NK, Bhatnagar S, Eisen JA, Sundaresan V (2015) Structure, variation, and assembly of the root-associated microbiomes of rice. Proc Natl Acad Sci USA 112(8):201414592. https://doi.org/10.1073/pnas.1414592112
8. Zhang Y, Xu J, Riera N, Jin T, Li J, Wang N (2017) Huanglongbing impairs the rhizosphere-to-rhizoplane enrichment process of the citrus root-associated microbiome. Microbiome. https://doi.org/10.1186/s40168-017-0304-4
9. Walters WA, Jin Z, Youngblut ND, Wallace JG, Sutter JL, Zhang W, Gonzalezpena A, Peiffer JA, Koren O, Shi Q (2018) Large-scale replicated field study of maize rhizosphere identifies heritable microbes. Proc Natl Acad Sci USA 115(28):7368–7373. https://doi.org/10.1073/pnas.1800918115
10. Marasco R, Rolli E, Fusi M, Michoud G, Daffonchio D (2018) Grapevine rootstocks shape underground bacterial microbiome and networking but not potential functionality. Microbiome 6(1):3. https://doi.org/10.1186/s40168-017-0391-2
11. Hamonts K, Trivedi P, Garg A, Janitz C, Grinyer J, Holford P, Botha FC, Anderson IC, Singh BK (2018) Field study reveals core plant microbiota and relative importance of their drivers. Environ Microbiol 20(1):124–140. https://doi.org/10.1111/1462-2920.14031
12. Atashgahi S, Sanchezandrea I, Heipieper HJ, Der Meer JRV, Stams AJM, Smidt H (2018) Prospects for harnessing biocide resistance for bioremediation and detoxification. Science 360(6390):743–746. https://doi.org/10.1126/science.aar3778
13. Urakawa H, Martenshabbena W, Stahl DA (2010) High abundance of ammonia-oxidizing archaea in coastal waters, determined using a modified DNA extraction method. Appl Environ Microbiol 76(7):2129–2135. https://doi.org/10.1128/AEM.02692-09
14. Caporaso JG, Lauber CL, Walters WA, Berglyons D, Lozupone CA, Turnbaugh PJ, Fierer N, Knight R (2011) Global patterns of 16S rRNA diversity at a depth of millions of sequences per sample. Proc Natl Acad Sci USA 108:4516–4522. https://doi.org/10.1073/pnas.1000080107
15. Caporaso JG, Kuczynski J, Stombaugh J, Bittinger K, Bushman FD, Costello EK, Fierer N, Pena AG, Goodrich JK, Gordon JI (2010) QIIME allows analysis of high-throughput community sequencing data. Nat Methods 7(5):335–336. https://doi.org/10.1038/nmeth.f.303
16. Edgar RC (2010) Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26(19):2460–2461. https://doi.org/10.1093/bioinformatics/btq461
17. Lozupone CA, Lladser ME, Knights D, Stombaugh J, Knight R (2011) UniFrac: an effective distance metric for microbial community comparison. ISME J 5(2):169–172. https://doi.org/10.1038/ismej.2010.133
18. Ramette A (2007) Multivariate analyses in microbial ecology. FEMS Microbiol ECOL 62(2):142–160. https://doi.org/10.1111/j.1574-6941.2007.00375.x
19. Zaura E, Keijser BJF, Huse SM, Crielaard W (2009) Defining the healthy "core microbiome" of oral microbial communities. BMC Microbiol 9(1):259–259. https://doi.org/10.1186/1471-2180-9-259
20. White JR, Nagarajan N, Pop M (2009) Statistical methods for detecting differentially abundant features in clinical metagenomic samples. PLOS Comput Biol. https://doi.org/10.1371/journal.pcbi.1000352
21. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, Huttenhower C (2011) Metagenomic biomarker discovery and explanation. Genome Biol 12(6):R60. https://doi.org/10.1186/gb-2011-12-6-r60
22. Lozupone CA, Hamady M, Kelley ST, Knight R (2007) Quantitative and qualitative beta diversity measures lead to different insights into factors that structure microbial communities. Appl Environ Microbiol 73(5):1576–1585. https://doi.org/10.1128/AEM.01996-06
23. Delgadobaquerizo M, Oliverio AM, Brewer TE, Benaventgonzalez A, Eldridge DJ, Bardgett RD, Maestre FT, Singh BK, Fierer N (2018) A global atlas of the dominant bacteria found in soil. Science 359(6373):320–325. https://doi.org/10.1126/science.aap9516
24. Perez-Jaramillo JE, Carrion VJ, Bosse M, Ferrao LFV, de Hollander M, Garcia AAF, Ramirez CA, Mendes R, Raaijmakers JM (2017) Linking rhizosphere microbiome composition of wild and domesticated *Phaseolus vulgaris* to genotypic and root phenotypic traits. ISME J 11(10):2244–2257. https://doi.org/10.1038/ismej.2017.85
25. Muyzer G, De Waal EC, Uitterlinden GA (1993) Profiling of complex microbial populations by denaturing gradient gel electrophoresis analysis of polymerase chain reaction-amplified genes coding for 16S rRNA. Appl Environ Microbiol 59(3):695–700. https://doi.org/10.1128/AEM.59.3.695-700.1993
26. Preza D, Olsen I, Willumsen T, Boches SK, Cotton SL, Grinde B, Paster BJ (2009) Microarray analysis of the microflora of root caries in elderly. Eur J Clin Microbiol 28:509–517. https://doi.org/10.1007/s10096-008-0662-8)
27. De Paepe K, Verspreet J, Courtin CM, De Wiele TV (2020) Microbial succession during wheat bran fermentation and colonisation by human faecal microbiota as a result of niche diversification. ISME J 14(2):584–596. https://doi.org/10.1038/s41396-019-0550-5
28. Song C, Wang B, Tan J, Zhu L, Lou D, Cen X (2016) Comparative analysis of the gut microbiota of black bears in China using high-throughput sequencing. Mol Genet Genomics 292(2):407–414. https://doi.org/10.1007/s00438-016-1282-0
29. Kim M, Morrison M, Yu Z (2011) Evaluation of different partial 16S rRNA gene sequence regions for phylogenetic analysis of microbiomes. J Microbiol Methods 84(1):81–87. https://doi.org/10.1016/j.mimet.2010.10.020
30. Ling Z, Kong J, Jia P, Wei C, Wang Y, Pan Z, Huang W, Li L, Chen H, Xiang C (2010) Analysis of oral microbiota in children with dental caries by PCR-DGGE and barcoded pyrosequencing.

Microb Ecol 60(3):677–690. https://doi.org/10.1007/s00248-010-9712-8
31. Xiao C, Ran S, Huang Z, Liang J (2016) Bacterial diversity and community structure of supragingival plaques in adults with dental health or caries revealed by 16S pyrosequencing. Front Microbiol 7:1145. https://doi.org/10.3389/fmicb.2016.01145
32. Schloss PD, Gevers D, Westcott SL (2011) Reducing the effects of PCR amplification and sequencing artifacts on 16S rRNA-based studies. PLoS ONE. https://doi.org/10.1371/journal.pone.0027310
33. Dai Z, Sevillano-Rivera MC, Calus ST et al (2020) Disinfection exhibits systematic impacts on the drinking water microbiome. Microbiome 8:42. https://doi.org/10.1186/s40168-020-00813-0
34. Fang R, Dong Y, Chen F, Chen Q (2015) Bacterial diversity analysis during the fermentation processing of traditional Chinese yellow rice wine revealed by 16S rDNA 454 pyrosequencing. J Food Sci. https://doi.org/10.1111/1750-3841.13018
35. Reading NC, Sperandio V (2006) Quorum sensing: the many languages of bacteria. FEMS Microbiol Lett 254(1):1–11. https://doi.org/10.1111/j.1574-6968.2005.00001.x
36. Turnbaugh PJ, Gordon JI (2009) The core gut microbiome, energy balance and obesity. J Physiol 587(17):4153–4158. https://doi.org/10.1113/jphysiol.2009.174136
37. Neu J, Lorca GL, Kingma SDK, Triplett EW (2010) The intestinal microbiome: relationship to type 1 diabetes. Endocrinol Metab Clin N Am 39(3):563–571. https://doi.org/10.1016/j.ecl.2010.05.008
38. Tilg H (2010) Obesity, metabolic syndrome, and microbiota: multiple interactions. J Clin Gastroenterol. https://doi.org/10.1097/MCG.0b013e3181dd8b64
39. Xu J, Zhang Y, Zhang P, Trivedi P, Riera N, Wang Y, Liu X, Fan G, Tang J, Colettafilho HD (2018) The structure and function of the global citrus rhizosphere microbiome. Nat Commun 9(1):4894. https://doi.org/10.1038/s41467-018-07343-2
40. Shade A, Handelsman J (2012) Beyond the Venn diagram: the hunt for a core microbiome. Environ Microbiol 14(1):4–12. https://doi.org/10.1111/j.1462-2920.2011.02585.x
41. Lemanceau P, Blouin M, Muller D, Moenneloccoz Y (2017) Let the core microbiota be functional. Trends Plant Sci 22(7):583–595. https://doi.org/10.1016/j.tplants.2017.04.008
42. Yeoh YK, Dennis PG, Paungfoolonhienne C, Weber LC, Brackin R, Ragan MA, Schmidt S, Hugenholtz P (2017) Evolutionary conservation of a core root microbiome across plant phyla along a tropical soil chronosequence. Nat Commun 8(1):215. https://doi.org/10.1038/s41467-017-00262-8
43. Bulgarelli D, Rott M, Schlaeppi K, Van Themaat EVL, Ahmadinejad N, Assenza F, Rauf P, Huettel B, Reinhardt R, Schmelzer E (2012) Revealing structure and assembly cues for Arabidopsis root-inhabiting bacterial microbiota. Nature 488(7409):91–95. https://doi.org/10.1038/nature11336
44. Jin T, Wang Y, Huang Y, Xu J, Zhang P, Wang N, Liu X, Chu H, Liu G, Jiang H (2017) Taxonomic structure and functional association of foxtail millet root microbiome. GigaScience 6(10):1–12. https://doi.org/10.1093/gigascience/gix089
45. Riera N, Handique U, Zhang Y, Dewdney MM, Wang N (2017) Characterization of antimicrobial-producing beneficial bacteria isolated from Huanglongbing escape citrus trees. Front Microbiol 8:2415–2415. https://doi.org/10.3389/fmicb.2017.02415
46. Trivedi P, He Z, Van Nostrand JD, Albrigo G, Zhou J, Wang N (2012) Huanglongbing alters the structure and functional diversity of microbial communities associated with citrus rhizosphere. ISME J 6(2):363–383. https://doi.org/10.1038/ismej.2011.100
47. Donn S, Kirkegaard JA, Perera G, Richardson AE, Watt M (2015) Evolution of bacterial communities in the wheat crop rhizosphere. Environ Microbiol 17(3):610–621. https://doi.org/10.1111/1462-2920.12452
48. Leff JW, Jones SE, Prober SM, Barberan A, Borer ET, Firn J, Harpole WS, Hobbie SE, Hofmockel KS, Knops JMH (2015) Consistent responses of soil microbial communities to elevated nutrient inputs in grasslands across the globe. Proc Natl Acad Sci USA 112(35):10967–10972. https://doi.org/10.1073/pnas.1508382112