Culture-independent analysis of the bacterial communities in Chinese fermented vegetables and genomic analysis of lactic acid bacteria

Jianming Zhang
Zhejiang Academy of Agricultural Sciences
Hye Seon Song
World Kimchi Institute
Chengcheng Zhang
Zhejiang Academy of Agricultural Sciences
Yeon Bee Kim
World Kimchi Institute
Seong Woon Roh
World Kimchi Institute
Daqun Liu (daqun.liu@hotmail.com)
Zhejiang Academy of Agricultural Sciences https://orcid.org/0000-0002-4232-1916

Research Article

Keywords: culture-independent analysis, bacterial community, Chinese fermented vegetables, genomic analysis

Posted Date: March 15th, 2021

DOI: https://doi.org/10.21203/rs.3.rs-218303/v1

Abstract

Six different fermented vegetables were collected from Zhejiang Province, China, to explore the associated bacterial communities using a high-throughput sequencing platform. A total of 24 phyla, 274 families and 569 genera were identified from six samples. Firmicutes and Proteobacteria were the main phyla in all of samples. Meanwhile, *Brevibacterium* was the major genus in Xiaoshan pickled radish. *Lactobacillus*-related genera and *Vibrio* were the major genera in fermented potherb mustard and its brine. *Enterobacter* and *Cobetia* were the major genera in fermented radish and its brine. *Chromohalobacter* was the major genus in the tuber mustard. These results indicated there were clear differences between the bacterial genera present in Xiaoshan pickled radish, fermented potherb mustard, fermented radish, and tuber mustard. This demonstrated the possible influences of raw materials and manufacturing processes. Furthermore, a large number of lactic acid bacteria were isolated and identified by culture-dependent and 16S rRNA gene sequence analysis, which accounted for more than 68% of all the isolates. In addition, whole genome analysis of *Lactobacillus suantsaii*, *Lactobacillus sakei* subsp. *sakei*, and *Weissella cibaria* showed they had large numbers of genes associated with carbohydrate metabolism. This may explain why these three bacterial strains can grow in fermented vegetable environments.

Introduction

As early as thousands of years ago, the ancestors of the Chinese people started to ferment vegetables to extend shelf life (Guan et al. 2020; Zhang et al. 2021). To date, most Chinese people still make fermented foods using perishable and seasonal vegetables, such as radish, cabbage, mustard, and cucumber (Liang et al. 2018; Zhai et al. 2018; Liu et al. 2019a; Rao et al. 2020). The purpose of fermenting the vegetables is not only to prolong shelf-life, but also to enhance nutrition and improve organoleptic characteristics. There are various fermented vegetables in China, such as paocai in Southwest China, and serofluid dish in Northwest China (Li et al. 2017; Liu et al. 2019b). Zhejiang Province, which is located in Southeast China, has a unique environment with rich topography and abundant rainfall. Due to this unique climate, local people particularly enjoy the fermented vegetable products produced in the region. Xiaoshan pickled radish, fermented potherb mustard, and meigan cai are distinctive traditionally fermented foods that have been consumed for thousands of years in Zhejiang Province (Guan et al. 2020; Zhang et al. 2019). Despite the consistency of the geography and climate, the flavor, taste, and texture of these fermented vegetables differ. The changes in the texture and flavor of the fermented vegetables are dependent on the microorganisms present during the fermentation process. This is because bacteria contribute to the flavor formation and acidification of the foods (Guan et al. 2020). Therefore, various measures have been employed to elucidate the structure and composition of the bacterial communities in the fermented vegetable products.

Over recent decades, culture-independent methods have been successfully employed to analyze the microbial communities in fermented vegetables. For example, denaturing gradient gel electrophoresis has been used to study the microbial ecology of suancai and serofluid dish in China (Wu et al. 2015; Zhang et

al. 2018; Zhou et al. 2018). Real-time quantitative PCR has also been used to detect both bacteria and yeast in Chinese paocai (Xiong et al. 2019). At present, high-throughput sequencing (HTS) is the most common method used to reveal the major bacterial community members in fermented vegetables. HTS can provide more detailed information and deeper insight into the microbial communities compared to conventional molecular method (Chen et al. 2019; Xiao et al. 2020). Nevertheless, HTS is based on community analysis of environmental DNA, and so pure cultures cannot be isolated and their physiological potentiality remains unknown. Therefore, some studies have combined culture-independent and culture-dependent techniques to analyze the bacterial communities and obtain pure cultures (Lu et al. 2020; Wang et al. 2020).

In this study, six fermented vegetable products were collected from Zhejiang. Culture-independent HTS and a culture-dependent method were used in combination to investigate the bacterial communities and isolate specific strains from the six fermented vegetable products. Furthermore, the genomes of *Lactobacillus suantsaii* (*L. suantsaii*) CBA3634, *Lactobacillus sakei* subsp. *sakei* (*L.sakei* subsp*. sakei*) CBA3635, and *Weissella cibaria* (*W. cibaria*) CBA3636, which were isolated from fermented potherb mustard and its brine, were sequenced using PacBio and Illumina platforms. This work aimed to deepen the understanding of the bacterial communities in the fermented vegetable products of Zhejiang and isolate and elucidate key features of the dominant lactic acid bacteria (LAB) responsible for vegetable fermentation. The results of this study will be helpful for stabilizing and enhancing the characteristics of fermented vegetable products of Zhejiang.

Materials And Methods

Sample collection

Samples were collected from six traditional fermented vegetables (C1-C6) from Zhejiang Province in 2018. Xiaoshan pickled radish (C1) was obtained from the factory in Hangzhou and the manufacturing process was as follows. The fresh radish was washed, cut evenly into strips, and air dried for three to five days. The radish was then mixed with a 6–7% (6–7 g/100 g) concentration of sodium chloride and stored at room temperature in the dark for approximately one year. Fermented potherb mustard (*Brassica juncea* var. *multiceps*; C2) and the brine of fermented potherb mustard (C6) were obtained from the factory in Hangzhou. The potherb mustard was pickled in a 16–18% concentration of sodium chloride and had been stored at room temperature in the dark for approximately one year. Fermented radish (C3) and its brine (C5) were obtained from the factory in Hangzhou. The fresh radish was washed and pickled in an 18–20% (18–20 g/100 g) concentration of sodium chloride. Then the radishes were stored at room temperature in the dark for approximately one year. Tuber mustard (*Brassica juncea* var. *tumida*; C4) was obtained from the factory in Hangzhou. Tuber mustard was pickled in a 14–17% concentration of sodium chloride and had been stored at room temperature in the dark for approximately one year.

Culture-independent HTS analysis of the bacterial communities

Total genomic DNA was extracted from the six samples using a FastDNA SPIN kit for soil (MP Biomedicals, Solon, OH, USA) according to the manufacturer's instructions. The amplicon library was prepared using a two-step PCR approach. In the first step, the hypervariable V3–V4 region of the bacterial 16S rRNA gene was amplified using the primer pair 341F and 805R (Fadrosh et al. 2014). In the second step, the dual index barcodes, i5 and i7, of Illumina Nextera were attached to each sample. The quantity and quality of the constructed library were determined using a Quanti-iT PicoGreen dsDNA Assay kit (Invitrogen, Waltham, MA, USA) and an Agilent 2100 Bioanalyzer System (Agilent Technologies, Santa Clara, CA). The final library products were sequenced using an Illumina MiSeq platform (Illumina, San Diego, CA, USA) according to Illumina's standard protocol. Bioinformatic analyses were performed using the free online platform of Majorbio Cloud Platform (www.majorbio.com) by Shanghai Majorbio Bio-pharm Technology Co. Ltd. according to the method described by Huang et al. (Huang et al. 2020).

Cultivation of LAB

De Man, Rogosa, and Sharpe (MRS) and Rogosa SL media were used to isolate cultivable LAB from all of the six fermented vegetable samples. The samples were serially diluted and plated on MRS and Rogosa SL agar media. Each plate was incubated in aerobic conditions at 30°C for 48 h. After cultivation, forty-four colonies were randomly picked and transferred several times until pure cultures were obtained. The 16S rRNA gene fragments of each isolate were amplified by colony PCR using the bacterial universal primer set 27F and 1492R. The PCR products were analyzed by agarose gel electrophoresis and sequenced using Sanger sequencing. The 16S rRNA gene sequences were identified using the EzBioCloud database (https://www.ezbiocloud.net/identify).

Genomic analysis of the isolated LAB

DNA extraction

Genomic DNA was extracted using an MG genomic DNA Purification Kit (MGmed, Seoul, Korea) according to the manufacturer's protocol. The quality and quantity of extracted genomic DNA were estimated using an Agilent 2100 Bioanalyzer and NanoDrop spectrophotometer (NanoDrop Technologies, Wilmington, DE, UK).

Library construction and sequencing

Whole genome sequencing of the isolates was performed using a combination of PacBio RS II and Illumina HiSeq sequencing platforms. The Illumina data were used to evaluate the complexity of the genome. For the PacBio sequencing, 20 kb SMRTbell libraries were constructed using a standard method and sequenced using the PacBio RS II system with the P6 DNA polymerase and C4 chemistry. For the Illumina HiSeq sequencing, paired-end libraries were constructed for the isolates using a TruSeq Nano DNA kit (350 bp insert size) following the manufacturer's recommendations and were sequenced using the Illumina HiSeq system.

Genome assembly and annotation

The sequencing reads obtained from the PacBio sequencer were *de novo* assembled using the hierarchical genome assembly process. After *de novo* assembly, the Illumina sequencing reads were integrated into the draft assemblies to obtain accurate genome sequences using the Pilon tool. The circular form of the contigs was confirmed using the UGENE software package (Komstantin et al. 2012).

Genome annotation was performed using Pathosystems Resource Integration Center database (PATRIC) and NCBI Prokaryotic Genome Annotation Pipeline (PGAP). The circular genome map was constructed using the CGview Server tool (http://cgview.ca/) based on the CDS annotated by Prokka (Anna et al. 2020). The predicted protein coding sequence (CDS) was assigned a K number using BlastKOALA and classified into a KEGG pathway using the Reconstruct Module tool in KEGG Mapper (Kanehisa et al. 2016). The Prophage Hunter tool was used to predict prophage activity within the genome. The 16S rRNA gene from the genomes were used to construct a phylogenetic tree using MEGA7 according to the neighbor-joining method. The topology of the phylogenetic tree was evaluated with 1,000 bootstrap replications and values >70% were indicated in the tree.

Results

HTS analysis of bacterial community structure

The number of effective sequence reads obtained from the six fermented vegetable samples ranged from 21,388 to 68,573. Based on a similarity threshold of 0.97, a total of 1,131 operational taxonomic units (OTUs) were identified from all the samples. Among them, 67 OTUs were present in all of the samples. In total, 130, 98, 147, 114, 64, and 176 OTUs were unique to the C1, C2, C3, C4, C5, and C6 samples, respectively (Fig. 1A). Species richness estimators (Table 1) were used to reflect the microbial phylotype richness levels. The results revealed that the C3 sample had the highest microbial diversity among the six samples, and rank abundance curves (Fig. 1B) also illustrated this result.

A total of 24 phyla, 274 families, and 569 genera were identified in the six samples. At the phylum level, the major phyla were Firmicutes and Proteobacteria. Firmicutes was the most dominant phylum in the C1 and C2 samples, accounting for 34.48% and 81.97% of the total phyla, respectively. Further, Proteobacteria, the second most dominant phylum, accounted for 33.37% and 15.51% of the total phyla in the C1 and C2 samples, respectively. On the contrary, Proteobacteria was the most dominant phylum in C3, C4, C5, and C6, accounting for 68.56%, 56.73%, 67.92%, and 79.14% of the total phyla, respectively. Firmicutes was the second most dominant phylum in C3, C4, C5, and C6, accounting for 15.82%, 25.27%, 25.81%, and 8.29% of the total phyla, respectively (Fig. 2A). At the family level, Halomonadaceae was most abundant in the C4 (53.83%) and C5 (27.47%) samples. Lactobacillaceae was the most dominant family in C2 (73.73%) and Brevibacteriaceae was the most dominant family in C1. Enterobacteriaceae was the most dominant family in C3, and Beijerinckiaceae was the most dominant family in C6 (Fig. 2B).

At the genus level, *Brevibacterium* (17.39%) was the most abundant in the C1 sample, followed by *Leuconostoc* (13.94%) (Zheng et al., 2020; Kim et al., 2021), *Lactobacillus*-related genera (11.50%), *Pseudomonas* (6.29%), and *Halomonas* (6.19%). In C2, *Lactobacillus*-related genera (73.72%) was the

major genus observed, and this was followed by *Weissella* (7.46%) and *Methylobacterium* (4.72%). *Enterobacter*, *Ralstonia*, *Lactobacillus*-related genera, and *Lactococcus* were the dominant genera in C3, accounting for 6.5%, 5.31%, 5.29%, and 5.15% of the total genera. In C4, *Chromohalobacter*, *Halomonas*, *Halanaerobium*, and *Lactobacillus*-related genera were the dominant genera, accounting for 27.68%, 21.08%, 16.13%, and 13.82% of the total genera. In addition, *Cobetia* (27.10%), *Lactobacillus*-related genera (10.36%), *Enterobacter* (6.84%), *Pseudoalteromonas* (5.55%), and *Lactococcus* (5.42%) were highly abundant in the C5 sample. *Vibrio* (12.37%), *Methylobacterium* (12.33%), *Pseudoalteromonas* (10.78%), *Pseudomonas* (8.82%), *Sphingomonas* (6.52%), and *Enterobacter* (5.95%) were the major genera in the C6 sample (Fig. 2C).

Isolation and molecular identification of LAB strains

Several colonies on MRS or Rogosa SL plates were isolated and purified from the four samples (C1, C2, C5, and C6). Forty-four isolates were identified via 16S rRNA gene sequence analysis. From C1, eight *Staphylococcus* (six *Staphylococcus piscifermentans* and two *Staphylococcus nepalensis*) and one *Terribacillus* (*Terribacillus goriensis*) were obtained. Five *Lactobacillus* (three *Lactobacillus alimentarius*, one *Lactobacillus suantsaii*, and one *Lactobacillus versmoldensis*) and two *Bacillus* (one *Bacillus altitudinis* and one *Bacillus gossypii*) were obtained from the C2 sample. Eleven *Lactobacillus* (five *Lactobacillus curvatus*, five *Lactobacillus sakei* subsp. *sakei*, and one *Lactobacillus pentosus*) were obtained from the C5 sample. From C6, three *Lactobacillus* (three *Lactobacillus sakei* subsp. *sakei*), eleven *Leuconostoc* (four *Leuconostoc mesenteroides* subsp. *mesenteroides*, five *Leuconostoc mesenteroides* subsp. *jonggajibkimchii*, and two *Leuconostoc carnosum*), two *Weissella* (one *Weissella soli* and one *Weissella cibaria*) and one *Enterococcus* (*Enterococcus casseliflavus*) were obtained. Among them, 30 strains were LAB strains, which accounted for more than 68% of all isolates. Furthermore, the representative LAB strains from the fermented foods, *L. suantsaii* CBA3634, *L. sakei* subsp. *sakei* CBA3635, and *W. cibaria* CBA3636, were selected for whole genome analysis (Table S1).

Whole genome analysis of *L. suantsaii, L. sakei* subsp. *sakei,* and *W.cibaria*

The genomes of the three strains, *L. suantsaii* CBA3634*, L. sakei* subsp. *sakei* CBA3635*,* and *W. cibaria* CBA3636, were sequenced using PacBio sequencer and HiSeq platforms. The *L. suantsaii* CBA3634 genome consisted of a circular chromosome and five plasmids. The length of the genome was 2,661,801 bp with a GC content of 50.49% (Fig. 3A, Table 2). The genome contains 18 rRNA genes, 70 tRNA genes, and 2,362 CDSs. Of the CDSs, 56.0% (1,322 of 2,362) were assigned protein functions and classified into 38 KEGG categories. The top three assigned KEGG categories were “global and overview maps,” “carbohydrate metabolism,” and “amino acid metabolism” (Fig. 3D). One active and two ambiguous prophage regions were predicted in the CBA3634 genome. The active prophage closely matched with the *Lactobacillus* phage LBR48 and the ambiguous prophages closely matched with the *Lactobacillus* phage LBR48 and the *Streptococcus* phage phiJH1301-2.

The complete genome of *L. sakei* subsp. *sakei* CBA3635 was composed of a single circular chromosome of 2,006,110 bp with a GC content of 41.13%, and one circular plasmid of 56,165 bp with a GC content of

41.11%. There were 2,002 CDSs, 21 rRNA operons, and 66 tRNAs in the genome (Fig. 3B and Table 2). The 1,130 CDSs of *L. sakei* subsp. *sakei* CBA3635 were assigned protein functions and classified into 37 KEGG categories as shown in Fig. 3D. With regard to the KEGG categories of strain CBA3635, “global and overview maps” (528 genes), “carbohydrate metabolism” (174 genes), and “translation” (81 genes) were assigned at high proportions. Two active and two ambiguous prophage regions were predicted in the CBA3635 genome. The active prophages closely matched with *Lactobacillus* phage PL-1 and *Staphylococcus* phage IME1361_01. The ambiguous prophages closely matched with *Lactobacillus* prophage Lj965 and *Staphylococcus* phage SPbeta-like.

The complete genome of *W. cibaria* CBA3636 included one circular chromosome and two circular plasmids. The total length was 2,423,709 bp with a GC content of 44.99 % (Fig. 3C and Table 2). The genome contained 2,165 CDSs, 28 rRNA genes, and 89 tRNAs. Of the CDSs, 56.9% (1,232 of 2,165) were assigned protein functions and were divided into 38 KEGG categories. Among them, the highest proportion of genes was assigned to the category of “global and overview maps” (597 genes), followed by “carbohydrate metabolism” (152 genes) and “membrane transport” (87 genes; Fig. 3D). Only three ambiguous prophage regions were predicted in the CBA3636 genome and these were closely related to *Weissella* phage WCP30, *Bacillus* phage Anthony, and *Streptococcus* phage phi-SsUD.1.

The maximum likelihood phylogenetic tree was constructed based on the 16S rRNA sequences of *L. suantsaii* CBA3634*, L. sakei* subsp. *sakei* CBA3635*, W. cibaria* CBA3636, and the reference strains (Fig.4). The phylogenetic tree showed that the three strains are located within two large clades. In addition, *L. suantsaii* CBA3634 was closely related to the reference strain *Lactobacillus suantsaii* L88^{T}, and *L. sakei* subsp. *sakei* CBA3635 was closely related to the reference strains *Lactobacillus sakei* subsp. *sakei* JCM 1157^{T} and *Lactobacillus sakei* subsp. *carnosus* DSM 15831^{T}.

Genome sequence accession numbers

The genomes of *L. suantsaii* CBA3634*, L. sakei* subsp. *sakei* CBA3635*,* and *W. cibaria* CBA3636 were deposited in NCBI with the accession numbers CP059603-CP059608, CP059697-CP059698, and CP059699-CP059701, respectively.

Discussion

Bacteria play a vital role in food fermentation. Bacterial community structure, particularly LAB community structure, is closely associated with the texture, flavor, nutrients and quality of fermented products. However, the understanding of bacterial composition is still limited to several major species responsible for traditional fermented vegetable products (Liu et al. 2017). In this study, the bacterial diversities associated with six fermented vegetables from Zhejiang were investigated. HTS analysis identified 24 phyla, 274 families, and 569 genera based on the OTUs identified in the six samples. The major phyla were Firmicutes and Proteobacteria, the relative abundance of which ranged from 67.85% to 97.48%. This indicated a firm correlation between the two types of bacteria and vegetable fermentation. In fact, most

previous reports have also found that fermented vegetables mainly contain Firmicutes and Proteobacteria at the phylum level (Liang et al. 2018; He et al. 2020).

At the genus level, however, the abundance of different genera varied greatly among the different fermented vegetable products. This result was in accordance with the results of previous studies. For example, *Lactobacillus*-related genera and *Pediococcus* were found to be the major genera in the suancai samples (Liu et al. 2019a), whereas *Bacillus* and *Bacteroides* were the main genera in shuidouchi samples (Chen et al. 2019). The results of the present study showed that *Brevibacterium* was the most abundant genus in Xiaoshan pickled radish. *Lactobacillus*-related genera and *Vibrio* were the most abundant genera in fermented potherb mustard and its brine. *Enterobacter* and *Cobetia* were the most abundant in fermented radish and its brine. *Chromohalobacter* was the most abundant in tuber mustard. These differences in bacterial community may result from the different types of raw materials used during fermentation. Additionally, besides the effects of the raw vegetable used, manufacturing processes or environmental conditions such as temperature, humidity and salinity can also significantly affect bacterial communities and the final quality of the fermented vegetable product (Xu et al. 2020).

Comparing the results obtained from the culture-dependent and culture-independent methods, many microorganisms such as *Lactobacillus*-related genera, *Leuconostoc,* and *Weissella* were encountered using both methods. In particular, *Lactobacillus*-related genera was found to be abundant using both approaches, which effectively confirmed the existence and composition of these microbes and communities in the fermented vegetable products. However, for the C4 samples, no colonies were obtained on the MRS medium. Some halophilic bacteria, such as *Chromohalobacter*, *Halomonas,* and *Halanaerobium*, were dominant in C4 samples and these halophilic bacteria need to be grown in a salty medium. Thus, the development of a more specific medium for target strain growth and reproduction will be critical for future study.

Whole genome analysis of strains CBA3634, CBA3635, and CBA3636 showed that these three LAB had a large number of genes associated with carbohydrate metabolism. Carbohydrate metabolism includes subcategories related to glucose, fructose, mannose, galactose, ascorbate, aldarate, starch, sucrose, amino sugar, nucleotide sugar, pyruvate, glyoxylate, and dicarboxylate, among others. These results indicate that *L. suantsaii* CBA3634*, L. sakei* subsp. *sakei* CBA3635*,* and *W. cibaria* CBA3636 are capable of metabolizing various carbohydrates. In addition, a tertiary level analysis of the carbohydrate metabolism in the genomes of the three strains indicated that their genes are involved in utilizing a variety of carbon sources such as starch, sucrose, pentose, galactose, fructose, and mannose. According to previous reports (Tashakor et al. 2017), *L. sakei* can indeed use different carbon sources, such as sucrose, maltose, fructose, and arabinose. Furthermore, it has been shown that *Weissella* can use glucose, D-maltose, D-ribose, mannitose, and trehalose. This ability to use multiple carbon sources could explain why *L. sakei* subsp. *sakei* CBA3635 and *W. cibaria* CBA3636 are able to grow in the fermented vegetable environment, as it may enhance their survival, competitiveness, and persistence.

Active or ambiguous prophages were detected in the *L. suantsaii* CBA3634*, L. sakei* subsp. *sakei* CBA3635*,* and *W. cibaria* CBA3636 genomes. In general, the presence of prophages in the genome has been considered to have a negative effect on the host, but recently it has been shown that prophages can provide several benefits to the host. Prophages can act in bacterial cellular processes such as antibiotic resistance, stress response, and virulence, and also mediate defense against superinfection (Song et al. 2019). The prophages were found in the genomes of LAB including *Lactobacillus, Weissella,* and *Leuconostoc*. Prophages present in LAB genomes can help the host to survive in their habitats (Durmaz et al. 2008; Wang et al. 2010; Liu et al. 2009; Joseph et al. 2016; Panthee et al. 2019).

Some strains of *L. sakei* are known to produce bacteriocins, such as sakacin A, sakacin G, sakacin P, sakacin Q, and sakacin X (Kim et al. 2020). The bioinformatic analysis performed in the present study showed that the *L. sakei* subsp. *sakei* CBA3635 genome contained a sakacin P response regulator gene. However, the sakacin P structural gene *sppA* was not found in the genome. Mretr *et al*. also found some of *L. sakei* strains that did not produce sakacin P, but contained sakacin P-related synthetic genes (Mretr et al. 2005). However, the role of the sakacin P-related genes in non-producers is not known. These results indicate that *L. sakei* subsp. *sakei* CBA3635 cannot produce sakacin. In addition, *Weissella* species have been known to synthesize extracellular exopolysaccharides (EPS) (Ye et al. 2018). EPS-producing LAB have been known to have an ability to tolerate gastrointestinal conditions, reduce pathogenic biofilm formation, and help in adhesion to epithelial cells than non-EPS-producing LAB. The bioinformatic analysis performed in the present study showed that *W. cibaria* CBA3636 genome contained the exopolysaccharide biosynthesis protein, which is an EPS-related gene. Nucleotide sugar biosynthesis is essential for EPS. Through further exploration of the KEGG database, it was found that *W. cibaria* CBA3636 had a high number of genes (29 genes) related to amino sugar and nucleotide sugar biosynthesis. This shows that *W. cibaria* CBA3636 is a potential EPS-producing bacterial strain.

Conclusion

Using HTS, the bacterial communities of six fermented vegetable products from Zhejiang Province in China, were found to be highly diverse. Firmicutes and Proteobacteria were the main phyla found in all of the fermented vegetable products. However, *Brevibacterium* was the major genus in Xiaoshan pickled radish; *Lactobacillus*-related genera and *Vibrio* were dominant in fermented potherb mustard and its brine; *Enterobacter* and *Cobetia* were the major genera in fermented radish and its brine; and *Chromohalobacter* was the major genus in tuber mustard. These results indicated that different types of raw materials and manufacturing processes led to differences in the bacterial genera present. A culture-dependent method was used to isolate the cultivable LAB strains. In total, 44 isolates were successfully obtained and identified by 16S rRNA gene sequencing. Among them, 30 strains were LAB strains, which accounted for more than 68% of all of the isolates. Whole genome analysis of *L. suantsaii* CBA3634*, L. sakei* subsp. *sakei* CBA3635, and *W. cibaria* CBA3636 showed that these three bacteria had a large number of genes associated with carbohydrate metabolism. This illustrated that *L. suantsaii* CBA3634*, L. sakei* subsp. *sakei* CBA3635, and *W. cibaria* CBA3636 can grow in fermented vegetable environments.

Declarations

Acknowledgments and Funds

This work was supported by the National Natural Science Foundation of China (NO. 31671865); the Key Research and Development Program of Zhejiang Province (NO. 2020c04008), China; and the World Institute of Kimchi funded by the Ministry of Science and ICT (KE2001-2), Republic of Korea.

Author contributions

Daqun Liu and Seong Woon Roh designed experiments. Jianming Zhang and Hye Seon Song analyzed experimental results and wrote the manuscript. Chengcheng Zhang, Yeon Bee Kim, Seong Woon Roh carried out experiments.

Compliance with ethical standards

Conflict of interest The authors declared that they had no competing interests.

Ethical approval Not applicable.

Consent for publication All authors approved the manuscript.

References

Anna S, Torsten S, Simon G (2020) Genome annotation with Prokka (Galaxy TrainingMaterials)./training-material/topics/genome-annotation/tutorials/annotation-with-prokka/tutorial.html Online; accessed Tue Sep 01 2020.

Chen YH, Liu XW, Huang JL, Baloch S, Xu X, Pei XF (2019) Microbial diversity and chemical analysis of Shuidouchi, traditional Chinese fermented soybean. Food Res Int. 116, 1289-1297.

Durmaz E, Miller MJ, AzcaratePeril MA, Toon SP, Klaenhammer TR (2008) Genome sequence and characteristics of lrm1, a prophage from industrial *Lactobacillus rhamnosus* strain m1. Appl Environ Microbiol. 74(15), 4601-4609.

Fadrosh DW, Ma B, Gajer P, Sengamalay N, Ott S, Brotman RM, Ravel J (2014) An improved dual-indexing approach for multiplexed 16S rRNA gene sequencing on the Illumina MiSeq platform. Microbiome. 2(1), 6.

Guan QQ, Zheng W, Huang T, Xiao YS, Liu Z.G, Peng Z, Gong DM, Xie MY, Xiong T (2020) Comparison of microbial communities and physiochemical characteristics of two traditionally fermented vegetables. Food Res Int. 128, 1-12.

He Z, Chen HY, Wang XY, Lin XP, Ji CF, Li SJ, Liang HP (2020) Effects of different temperatures on bacterial diversity and volatile flavor compounds during the fermentation of Suancai, a traditional fermented vegetable food from northeastern China. Lwt-Food Sci. Technol. 118, 108773.

Huang TT, Wu ZY, Zhang WX (2020) Effects of garlic addition on bacterial communities and the conversions of nitrate and nitrite in a simulated pickle fermentation system. Food Control. 113, 107215.

Joseph BD, Jason Q, Edze RW, Angus B, David SG, Alan RD, Karen LM (2016) Prophages mediate defense against phage infection through diverse mechanisms. International Society for Microbial Ecology. 6, 1-13.

Kanehisa M, Sato Y, Morishima K (2016) BlastKOALA and GhostKOALA: KEGG Tools for Functional Characterization of Genome and Metagenome Sequences. J Mol Biol. 428(4), 726-731.

Kim E, Cho EJ, Yang SM, Kim,MJ, Kim HY. Novel approaches for the identification of microbial communities in kimchi: MALDI-TOF MS analysis and high-throughput sequencing. Food Microbiology. 94: 103641: (2021).

Kim KH, Chun BH, Baek JH, Roh SW, Lee SH, Jeon CO (2020) Genomic and metabolic features of *Lactobacillus sakei* as revealed by its pan-genome and the metatranscriptome of kimchi fermentation. Food Microbiol. 86, 103341.

Konstantin O, Olga G, Mikhail F, Team TU (2012) Unipro ugene: a unified bioinformatics toolkit. Bioinformatics. 28(8), 1166-7.

Li QQ, Kang JM, Ma Z, Li XQ, Liu LL, Hu XZ (2017) Microbial succession and metabolite changes during traditional serofluid dish fermentation. Lwt-Food Sci. Technol. 84, 771-779.

Liang HP, Chen HY, Zhang WX, Yu CX, Ji CF, Lin XP (2018) Investigation on microbial diversity of industrial, Zhacai Paocai during fermentation using high-throughput sequencing and their functional characterization. Lwt-Food Sci. Technol. 91, 460-466.

Liu DQ, Tong C (2017) Bacterial community diversity of traditional fermented vegetables in China. Lwt-Food Sci.Technol. 40-48.

Liu JY, Li AH, Ji C, Yang WM (2009) First description of a novel *Weissella* species as an opportunistic pathogen for rainbow trout Oncorhynchus mykiss (Walbaum) in China. Vet Microbiol. 136(3-4), 314-320.

Liu ZG, Li JY, Wei BL, Huang T, Xiao YS, Peng Z, Xie MY, Xiong T (2019a) Bacterial community and composition in Jiang-shui and Suan-cai revealed by high-throughput sequencing of 16S rRNA. Int J Food Microbiol. 306, 1-9.

Liu ZG, Peng Z, Huang T, Xiao YS, Li JY, Xie MY, Xiong T (2019b) Comparison of bacterial diversity in traditionally homemade Paocai and Chinese spicy cabbage. Food Microbiol. 83, 141-149.

Lu Y.H, Tan XY, Lv YP, Yang GH, Chi YL, He Q (2020) Physicochemical properties and microbial community dynamics during Chinese horse bean-chili-paste fermentation, revealed by culture-dependent and culture-independent approaches. Food Microbiol. 85, 103309.

Mretr T, Naterstad K, Wang E, Aasen I, Chaillou S, Zagorec M, Axelsson L (2005) Sakacin P non-producing *Lactobacillus sakei* strains contain homologues of the sakacin P gene cluster. Res Microbiol. 156(9), 0-96.

Panthee S, Paudel A, Blom J, Hamamoto H, Sekimizu K (2019) Complete Genome Sequence of *Weissella hellenica* 0916-4-2 and Its Comparative Genomic Analysis. Front Microbiol. 10, 1-13.

Rao Y, Tao YF, Chen X, She X, Qian Y, Li YL, Du YW, Xiang WL, Li HH, Liu L (2020) The characteristics and correlation of the microbial communities and flavors in traditionally pickled radishes. Lwt-Food Sci. Technol. 118, 1-8.

Song WC, Sun, HX, Zhang C, Cheng, L, Peng, Y, Deng, ZQ, Wang D, Wang Y, Hu M, Liu WN, Yang HM, Shen Y, Li JH, You LC, Xiao MF (2019) Prophage hunter: an integrative hunting tool for active prophages. Nucleic Acids Res. 47, 74-80.

Tashakor A, Hosseinzadehdehkordi M, Emruzi Z, Gholami D (2017) Isolation and identification of a novel bacterium, *Lactobacillus sakei* subsp. *dgh* strain 5, and optimization of growth condition for highest antagonistic activity. Microb Pathogenesis. 106,78-84.

Wang DD, Chen G, Tang Y, Li H, Shen WX, Wang M, Liu SL, Qin W, Zhang QS (2020) Effects of temperature on Paocai bacterial succession revealed by culture-dependent and culture-independent methods. Int J Food Microbiol. 317, 108463.

Wang XX, Kim YH, Ma Q, Hong SH, Pokusaeva K, Sturino JM,Wood TK (2010) Cryptic prophages help bacteria cope with adverse environments. Nat Commun. 1(9), 147.

Wu RN, Yu ML, Liu XY, Meng LS, Wang QQ, Xue YT, Wu JR, Yue XQ (2015) Changes in flavour and microbial diversity during natural fermentation of Suan-cai, a traditional food made in Northeast China. Int J Food Microbiol. 211, 23-31.

Xiao YS, Huang T, Huang CL, Hardie J, Peng Z, Xie MY (2020) The microbial communities and flavour compounds of Jiangxi yancai, Sichuan paocai and Dongbei suancai: Three major types of traditional Chinese fermented vegetables. Lwt-Food Sci. Technol. 121, 1-9.

Xiong T, Chen JK, Huang T, Xie MY, Xiao YS, Liu CG, Peng Z (2019) Fast evaluation by quantitative PCR of microbial diversity and safety of Chinese Paocai inoculated with *Lactobacillus plantarum* NCU116 as the culture starter. Lwt-Food Sci.Technol. 101, 201-206.

Xu XX, Wu BB, Zhao WT, Lao F, Chen F, Liao XJ, Wu JH (2020) Shifts in autochthonous microbial diversity and volatile metabolites during the fermentation of chili pepper (*capsicum frutescens* L.). Food Chem.

127512.

Ye GB, Chen YH, Wang CL, Yang RR, Bin XY (2018) Purification and characterization of exopolysaccharide produced by *Weissella cibaria* YB-1 from pickle Chinese cabbage. Int J Biol Macromol. 120, 1315-1321.

Zhai YK, Ilenys PD, Diaz JT (2018) Viability of commercial cucumber fermentation without nitrogen or air purging. Trends Food Sci Tech. 81, 185-192.

Zhang CC, Chen JB, Liu XQ, Liu DQ (2019) Bacterial community and quality characteristics of the fermented potherb mustard (brassica juncea var. multiceps) under modified atmospheres. Food Res Int. 116, 266-275.

Zhang J, Wu SS, Zhao LH, Ma QL, Li XB, Ni MY, Zhou T, Zhu HL (2018) Culture-dependent and -independent analysis of bacterial community structure in Jiangshui, a traditional Chinese fermented vegetable food. Lwt-Food Sci. Technol. 96, 244-250.

Zhang JM, Zhang CC, Wu WC, Lv XC, Xin XT, Liu DQ, Hu HJ, Guo ST (2021). Correlation of the bacterial communities with umami components, and chemical characteristics in zhejiang xuecai and fermented brine. Food Res Int, 140, 109986.

Zheng J, Wittouck S, Salvetti E, Franz CMAP, Harris HMB, Mattarelli P, O'Toole PW, Pot B, Vandamme P, Walter J, Watanabe K, Wuyts S, Felis GE, Ganzle MG, Lebeer S. A taxonomic note on the genus *Lactobacillus*: description of 23 novel genera, emended description of the genus *Lactobacillus* Beijerinck 1901, and union of *Lactobacillaceae* and *Leuconostocaceae*. International Journal of Systematic and Evolutionary Microbiology. 70: 2782–2858 (2020).

Zhou Q, Zang SZ, Zhao ZN, Li XL (2018) Dynamic changes of bacterial communities and nitrite character during northeastern Chinese sauerkraut fermentation. Food sci biotechnol. 27, 79-85.

Tables

Table 1. The diversity indices of bacterial communities in the six samples

Sample	ACE	Chao1	Coverage (%)
C1	567.79	549.75	0.99
C2	449.97	436.89	0.99
C3	653.07	643.01	0.99
C4	349.38	341.33	0.99
C5	475.71	467.44	0.99
C6	590.61	573.40	0.99

Table 2. General genome features of *Lactobacillus suantsaii* CBA3634, *Lactobacillus sakei* subsp. *sakei* CBA3635, and *Weissella cibaria* CBA3636

Feature	CBA3634	CBA3635	CBA3636
Genome size	2,661,801	2,006,110	2,423,709
GC content (%)	50.49	41.13	44.99
rRNAs	18	21	28
tRNAs	70	66	89
Predicted genes	2,362	2,002	2,165

Figures

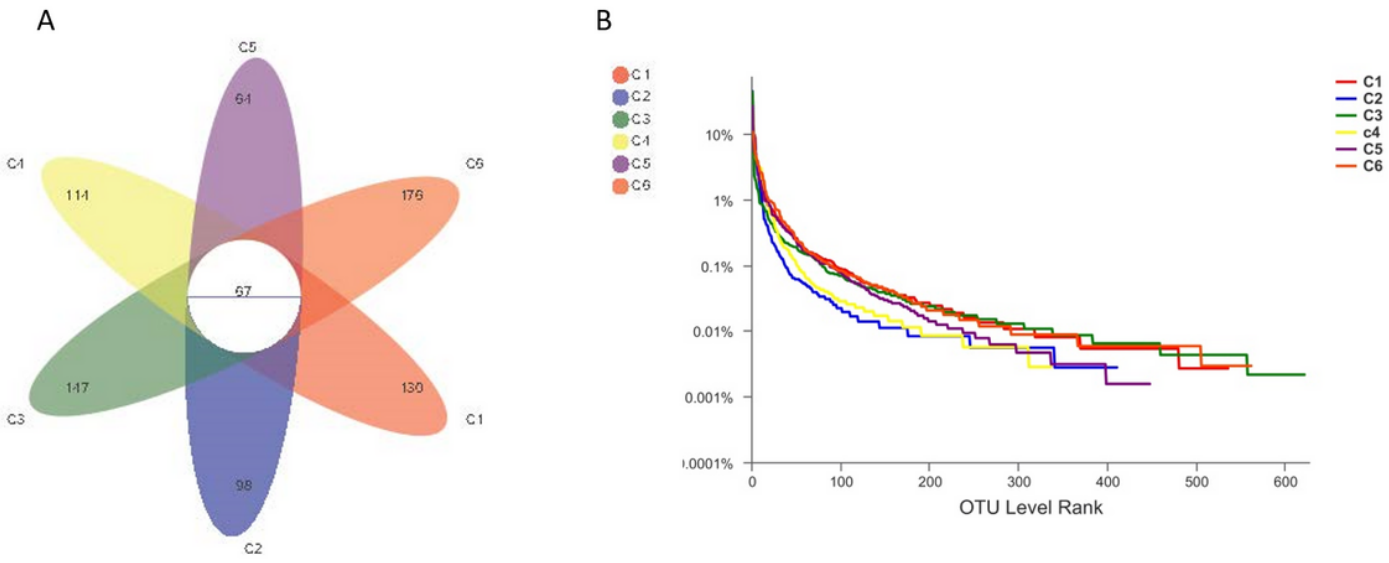

Figure 1

Venn diagram of bacteria (A) and rank-abundance curves of bacteria (B) in the six samples.

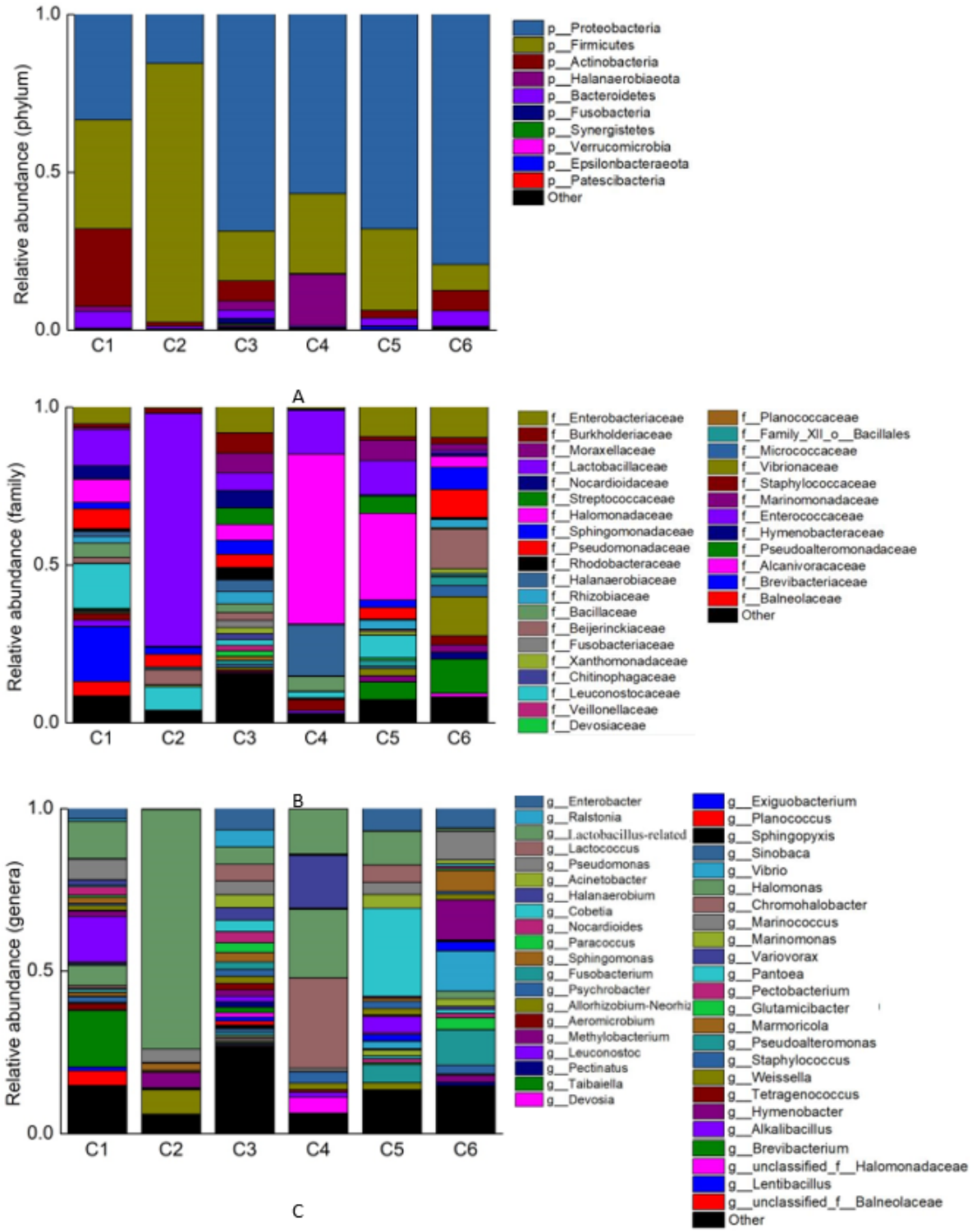

Figure 2

Bacterial communities in the six samples at the phylum (A), family (B), and genus (C) level.

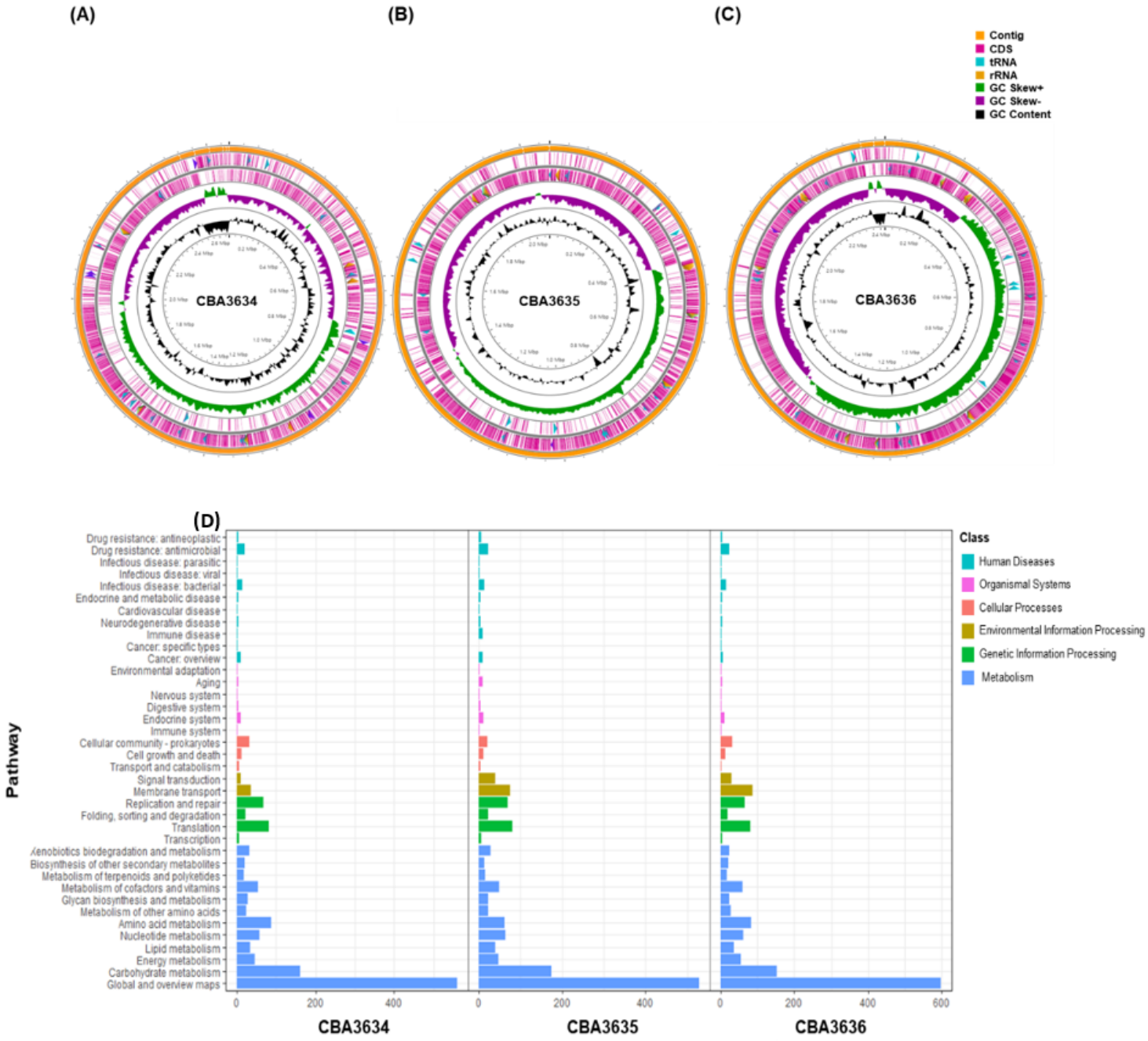

Figure 3

The genome map and KEGG pathways of Lactobacillus suantsaii CBA3634, Lactobacillus sakei subsp. sakei CBA3635, and Weissella cibaria CBA3636. (A) is genome map of CBA3634; (B) is genome map of CBA3635; (C) is genome map of CBA3636. The meaning of each circle is indicated by the legend in the figure. (D) is KEGG pathways of CBA3634, CBA3635, and CBA3636.

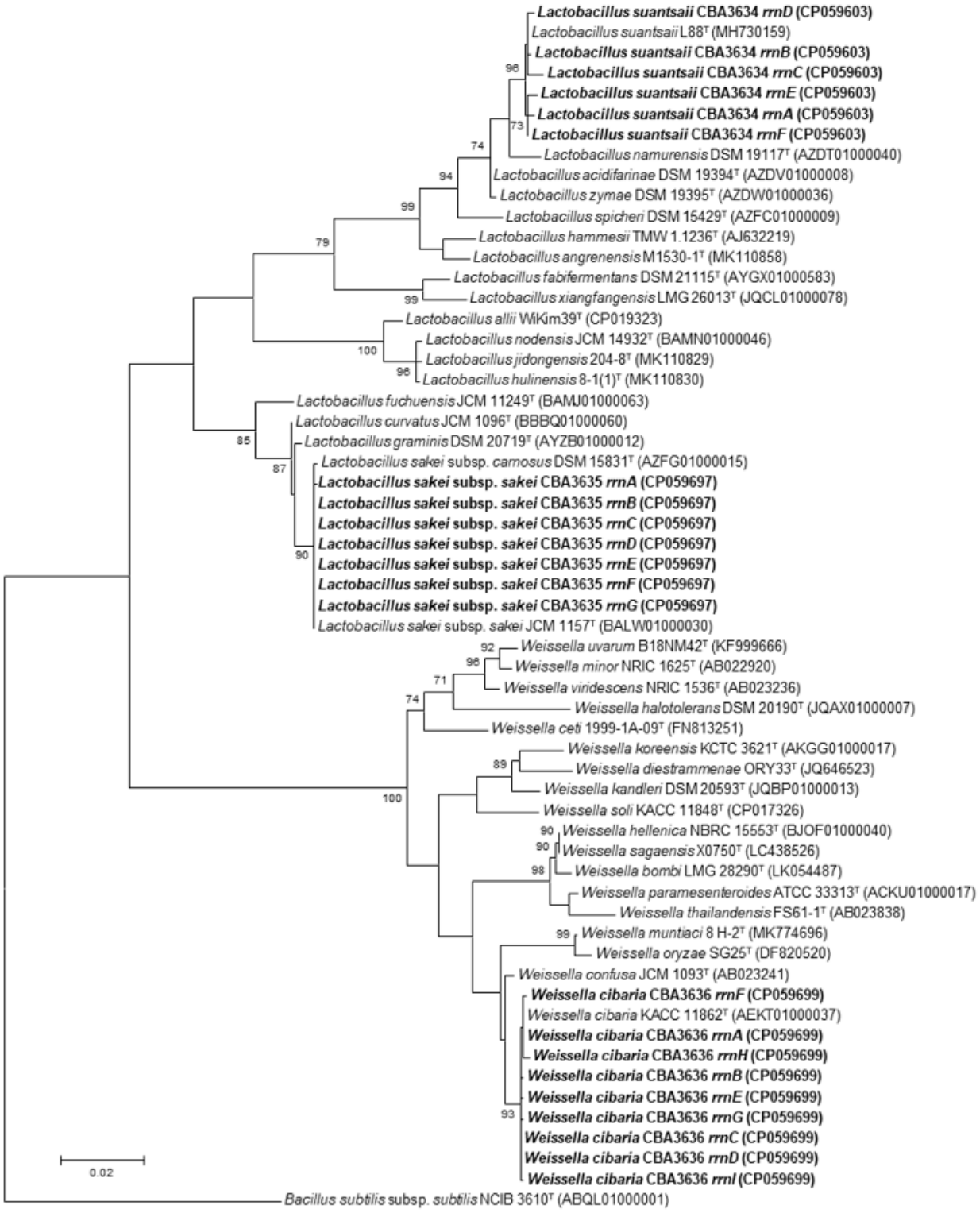

Figure 4

Phylogenetic tree of Lactobacillus suantsaii CBA3634, Lactobacillus sakei subsp. sakei CBA3635, Weissella cibaria CBA3636, and related strains based on 16S rRNA gene sequences.

Supplementary Files

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

- Supplement.docx