Bacterial community in pit mud of Yibin Baijiu in China by high throughput sequencing

Li Chen [a, 1], Yuzhu Li [b, 1], Lei Jin [c], Li He [b], Xiaolin Ao [b], Shuliang Liu [b], Yong Yang [b], Aiping Liu [b], Shujuan Chen [b,*], Likou Zou [c,*]

[a] Yibin Products Quality Supervision and Inspection Institute, Yibin, Sichuan, China
[b] College of Food Science, Sichuan Agricultural University, Ya'an, Sichuan, China
[c] College of Resources, Sichuan Agricultural University, Chengdu, Sichuan, China

Corresponding Author:
Shujuan Chen
No.46 Xinkang Road, Yucheng District, Ya'an City, Sichuan Province, P.R. 625014, China
Email address: chenshujuan1@163.com
Likou Zou
No.211 Huimin Road, Wenjiang District, Chengdu City, Sichuan Province, P.R. 611130, China
Email address: zoulikou@sicau.edu.cn

[1] The authors contributed equally to this work

Abstract

"Yibin Baijiu" (YB) is a special Chinese strong-aroma baijiu (CSAB) origined in Yibin, a
western city of China. YB is fermented in cellars lined with pit mud (PM), of which the
microbiota affect the quality. In the study, 16S rRNA high throughput sequencing was used to
demonstrate the bacterial community structure and diversity in PM of YB. Besides, the
physicochemical characteristics of PM were also measured, including moisture content, pH,
available phosphorous, ammonia nitrogen and humic acid. Results showed that Firmicutes was
the dominant bacteria in all PM samples with the abundance over 70.0%, followed by
Euryarchaeota (11.3%), Bacteroidetes (6.5%), Synergistetes (3.0%), Actinobacteria (1.4%) and
Proteobacteria (1.2%). Moreover, 14 different genera with average relative abundance > 1%
were found. There was no significant difference in Chao1 and Shannon index between the sub-
layer and middle-layer PM (p>0.05). However, LEfSe analysis showed that the relative
abundance of *Lactobacillus* in the sub-layer PM is significantly higher than middle-layer PM. pH
was significantly different (p < 0.01) between the two groups. Canonical correspondence
analysis revealed that bacterial community in PM was correlated significantly with available
phosphorous and pH. Our study provided basic data for further elucidating diversity of
microbiota in pit mud of YB and the potential mechanism of Baijiu production.

Introduction

Baijiu is one of the six most famous distilled liquors worldwide (Fan & Qian, 2006; Jin, Zhu, &
Xu, 2017; McGovern et al., 2004). It included four main aroma types: strong, light, sauce and
rice aromas. Among these, the Chines strong-aroma baijiu (CSAB) accounts for over 70% of
consumption (Liu & Sun, 2018; Liu, Tang, Zhao, et al., 2017). Yibin is one of the birth place of
CSAB culture (Zou, Zhao, & Luo, 2018; Yang et al., 2017; You et al., 2016). It is located at the
confluence of Min River, Jinsha River and Yangtze River with good water quality. Meanwhile, it
has a unique weak acid clay with strong viscosity, good water retention and rich in minerals,
which is an excellent natural material for making cellars (Tang et al., 2012). Moreover, Yibin
enjoys a mid-subtropical humid monsoon climate, which is perfect for the growth of liquor-
making microbes (Zhao et al., 2013). All these provide good resources for the development of
"Yibin Baijiu" with unique flavor.

"Yibin Baijiu" is produced by distilling mixed and fermented grains, containing sorghum,
glutinous rice, rice, wheat and corn, having the characteristics of sweetness, strong aroma and
lasting aftertaste (Zheng et al., 2013; Zheng & Han, 2016). The fermentation of "Yibin Baijiu" is
a traditional solid-state fermentation process that occurs in fermentation pits (Barrios-González,
2012), which serve as solid bioreactors. The pit mud (PM) is a complex ecosystem contains a
variety of microbial communities, like bacteria, archaea and fungi (Liu et al., 2017). Bacterial
communities in PM can break down macromolecules into small peptides and monosaccharides,
which produce aromatic compounds, and thus endow the products flavor (Tao et al., 2014; Zhao,
Zheng, Zhou, & Shi, 2012). Yet, its mechanism still remains to be discovered. To clarify this

61 bacterial effect and optimize the brewing technology, it's essential to study the composition and
62 diversity of microorganisms in PM.
63   Since the 1960s, researchers have used traditional culture method to reveal the microbiota in
64 PM (Wu, Qi, Xue, Guo, & Jie, 1980). *Clostridium, Bacillus, Pseudomonas* and
65 *Sporolactobacillus* were isolated and identified using traditional microbial classification and
66 identification methods (Yue, Hang, Liu, Hu, & Zheng, 2007). However, the method only
67 detected less than 1% of culturable microorganisms, which failed to fully reveal the structure
68 characteristics (Amann, Ludwig, & Schleifer, 1995). Then, molecular methods, such as
69 polymerase chain reaction–denaturing gradient gel electrophoresis (PCR-DGGE) (Deng et al.,
70 2012; Huang, Xiong, Hu, Liang, & Zhao, 2017), phospholipid fatty acid (PLFA) analysis (Zheng
71 et al., 2013) and clone library analysis of the 16s rRNA gene (Ye et al., 2013), have been
72 extensively employed. These methods not only avoid the tedious process of microbial culture,
73 but also contribute to further understanding PM ecosystem. Thus, Firmicutes, Proteobacteria,
74 Bacteroidetes, Actinobacteria, and Synergistetes were found to be predominant in many PM
75 bacterial communities (Ding, Wu, Huang, Li, & Zhou, 2014; Liang, Luo, Zhang, Wu, & Zhang,
76 2016). Although these molecular methods provided a first description of abundant populations of
77 PM, the extent of bacterial diversity remains unexplored. Until recently, 16s rRNA high-
78 throughput sequencing was gradually used (Liu, 2017; Zheng et al., 2015). However, most of
79 these studies just focused on samples from one or few enterprises. Due to few samples used, only
80 limited information can be provided. Moreover, Yibin has a brewing history of nearly a thousand
81 years, suitable ecological factors and relatively special brewing techniques, which make its
82 "Yibin Baijiu" special bacterial flora unique. Therefore, considering the complexity of PM
83 microorganisms, it is necessary to study the PM of "Yibin Baijiu" to enrich the information of
84 PM bacterial communities.
85   In the study,13 baijiu enterprises in Yibin were chosen to examine the microbiota structure via
86 high-throughput sequencing. Meanwhile, the physicochemical characteristics of PM were
87 measured to explore the correlation between bacterial community structure and environmental
88 factors, which help us to understand the microflora composition and main functional
89 microorganism of PM comprehensively.

Materials & Methods

Sampling

92 PM samples were collected in December 2017 from 13 distilleries located in Yibin (28°76′ N,
93 104°63′ E). A total of 68 samples were collected, among which 37 were collected from middle
94 layer of cellar walls and 31 were collected from the substrate layer (100 g PM at each position).
95 PM samples were numbered according to the information of distillery, cellar and layer, and
96 divided into two groups: the middle-layer and the sub-layer (Table S1). The samples were
97 aseptically collected in sterilized bas and kept cold during transport from the distilleries to the
98 laboratory.

Physicochemical properties

100 The basic physicochemical properties of PM, including moisture content, pH, available
101 phosphorous, ammonia nitrogen and humic acid, were also determined as previously described
102 (Shen, 2014).
103 **DNA extraction, PCR amplification, and Illumina HiSeq sequencing**
104 DNA was extracted from the PM by using the FastDNA SPIN kit for Soil (Mpbio, USA)
105 according to the manufacturer's instructions. The crude DNA was detected quantitatively
106 according to absorbance at 260 nm with NanoDrop™ one spectrophotometer (Thermo Fisher
107 Scientific, USA), and the purity was assessed by agarose gel electrophoresis (Beijing Liuyi
108 Biological Technology Co., Ltd. China).
109 To amplify the V3-V4 regions of 16s rRNA genes, a pair of universal bacterial primers was
110 employed in PCR reactions: a forward primer (341F) CCTAYGGGRBGCASCAG and a reverse
111 primer (806R) GGACTACHVGGGTWTCTAAT. The reaction mixtures (25 μl) contained 5 μl
112 of 5×TransStart Fastpfu Buffer (Transgen Biotech), 1 μl of DNA sample, 1 μl of each primer, 2
113 μl 2.5mM dNTPs, and 0.5 μl of TransStart Fastpfu DNA polymerase (Transgen Biotech). The
114 PCR conditions were: 2 min at 95°C; 30 cycles of 20 s at 95°C, 30 s at 55°C, 30 s at 72°C; final
115 72 °C for 5 min and cooling at 4°C. The PCR products were purified with a universal PCR
116 purification kit (Tiangen, Beijing, China) and sent to a commercial sequencing company for high-
117 throughput sequencing using the Illumina Hiseq sequencing platform (MeiYin Health
118 Technology (Beijing) Co., Ltd., China).
119 All the raw HiSeq-generated 16s rRNA gene sequence data were processed in the QIIME
120 pipeline (V1.9.1, http://qiime.org/tutorials/tutorial.html). After initial QC processing, pairs of
121 reads were merged using FLASH (V1.2.7, http://ccb.jhu.edu/software/FLASH/) to acquire raw
122 tags. Then, they were screened by the QIIME unction (Caporaso et al., 2010) to obtain clean
123 tags. Chimeras were detected and removed using the Uchime algorithm to obtain effective tags
124 (Edgar, Haas, Clemente, Quince, & Knight, 2011). The effective tags, at a similarity threshold of
125 97%, were grouped into the same operational taxonomic units (OTUs). A representative
126 sequence for each OTU was extracted and annotated using the SILVA database (V119,
127 https://www.arb-silva.de/).
128 **Data analysis**
129 The community richness and diversity were estimated by the Chao1 (Chao & Bunge, 2002)
130 and Shannon indices (Shannon, 1997), respectively. Bacterial community differences in two
131 groups of PM were evaluated by principal coordinates analysis (PCoA) in Fast UniFrac
132 (http://bmf.colorado.edu/fastunifrac/). Linear discriminant analysis Effect Size (LEfSe) was used
133 to compare the significance of species differences between the two groups (Segata et al., 2011).
134 Canonical correspondence analysis (CCA) between bacterial communities in the PM and
135 physicochemical properties was performed via the vegan package in R (http://vegan.r-forge.r-
136 project.org/). All the statistical analyzes, including spearman's test, were performed with the
137 SPSS software (V22.0, IBM Corp., Armonk, NY, USA). All tests for significance were two
138 sided, and P-values < 0.05 were considered to indicate significance.

139 # Results

Physicochemical properties of the PM

The physicochemical characteristics of PM of the two groups are shown in Table 1. Among all physicochemical indicators analyzed by independent samples t-test, level of pH was significantly different ($p < 0.01$) between the two groups. However, no significant differences in moisture content, ammonia nitrogen, available phosphorous and humic acid were found.

The diversity of bacterial community in PM

Totally, 68 samples were analyzed for amplicon sequencing. On average, 99,639 Clean Tags were generated for each sample with a minimum of 53,888. 53,493–208,410 effective tags for each sample with an average sequence length of 444 bp. Rarefaction curves was used to assess the degree of completion of a taxonomic survey (Good, 1953). The rarefaction curve showed that all curves approached the saturation plateau, proving that the identified sequences of PM samples represent almost all of bacterial sequences (Figure S1).

Eventually, a total of 220,790 OTUs were detected based on 97% similarity in 16s rRNA sequences, of which middle-layer group and sub-layer group had 186,289 and 148,904, respectively. Meanwhile, 114,403 OTUs were shared between the two groups. The mutual and unique OTUs are shown in Venn diagrams for the two groups (Figure S2).

The bacterial richness (Chao1) indices for all samples ranged from 1,214.30 to 42,441.82, and were fairly similar in the two different groups, with $11,296.93 \pm 4,816.73$ in the middle-layer group and $10,549.07 \pm 7,638.48$ in the sub-layer group. Meanwhile, the diversity (Shannon) indices was 7.82 ± 1.68 in the middle-layer group, and 7.57 ± 1.21 in the sub-layer group. There was no significant difference in Chao1 and Shannon between the two groups ($P>0.05$), suggesting similar overall species diversity (Figure 1). Furthermore, the PCoA analysis showed that there was no explicit clustering of samples within the groups, demonstrating significant inter-individual variations in the bacterial communities (Figure 2).

Bacterial community structure in PM

In total, 49 phyla, 332 families and 811 genera were identified from all samples (Table 2). The most abundant phylum was Firmicutes (71.9%), followed by Euryarchaeota (11.3%), Bacteroidetes (6.5%), Synergistetes (3.0%), Actinobacteria (1.4%), Proteobacteria (1.2%), Cloacimonetes (0.6%), Tenericutes (0.4%), Chloroflexi (0.3%) and Cyanobacteria (0.2%) (Figure. 3). Most notably, Firmicutes constituted more than 70.0% of the bacterial microbiota in all the groups, as well as Euryarchaeota and Bacteroidetes were also predominant in all groups with average relative abundances > 5%. The proportion of these three major phyla in the middle-layer and sub-layer group was 71.2% and 72.3%, 12.7% and 10.0%, and 6.9% and 6.3%, respectively.

At the genus level, dominate genera were similar between the two groups. All samples dominated by the 14 major genera (average relative abundance > 1% in all groups) including *Ruminiclostridium* 5 (17.9%), *Gelria* (8.4%), *Methanoculleus* (4.9%), *Syntrophomonas* (3.8%), *Lactobacillus* (3.6%), *Syntrophaceticus* (2.8%), *Proteiniphilum* (2.7%), *Aminobacterium* (2.7%), *Petrimonas* (2.4%), *Sedimentibacter* (2.1%), *Methanosarcina* (1.9%), *Methanobacterium* (1.8%), *Methanobrevibacter* (1.8%) and *Caldicoprobacter* (1.7%), respectively. The most

180   abundant 10 genera are shown in Figure 4. In addition, the relative abundance of *Clostridium*
181   *sensu stricto 14* and *Clostridium sensu stricto 12* also exceeded 1% in the sub-layer group, 1.8%
182   and 1.7% respectively. To determine the classified bacterial taxa with significant abundance
183   differences between two groups, the linear discriminant analysis (LDA) effect size (LEfSe)
184   method was used. The LEfSe analysis result showed that *Lactobacillus* is the key discrepant
185   bacteria, and the relative abundance of *Lactobacillus* in the sub-layer PM is significantly higher
186   than middle-layer PM (Figure 5).

187   **Comparison of dominant bacterial communities in different Baijius' pit mud**
188   Based on previous studies, wer compared the dominant microbiota of strong-aroma baijiu in
189   different regions to ascertain the common dominant bacteria, and the unique dominant bacteria
190   of "Yibin Baijiu" (Table 3). There were four common dominant bacteria in strong-aroma baijiu
191   PM in five different regions (Mianzhu, Luzhou, Jiangsu, Hunan, Yibin), including *Lactobacillus,*
192   *Sedimentibacter, Syntrophomonas* and *Methanobrevibacter*. Besides, it is noteworthy that
193   *Petrimonas, Clostridium, Methanosarcina, Methanoculleus* and *Methanobacterium* were also
194   detected in our study with the relative abundance above 1%, and almost all of them were
195   affiliated with the Euryarchaeota phylum. Three unique bacteria were also found in "Yibin
196   Baijiu" pit mud, namely *Ruminiclostridium 5, Gelria* and *Syntrophaceticus*.

197   In addition, the common dominant microorganism of strong-aroma baijiu and those of other
198   aroma types (miscellaneous-aroma and sauce-aroma) were compared. The dominant microbiota
199   composition in the Baijiu pit mud of different aroma types were quite different. Only
200   *Lactobacillus* is the dominant microorganism shared by three aroma types of baijiu pit mud. In
201   addition, *Enterobacter* is another dominant microorganism shared by miscellaneous-aroma and
202   sauce-aroma, which was not found in strong-aroma baijiu pit mud. The detailed dominant
203   microflora information of 3 aroma types of Baijiu pit mud are shown in Table 4.

204   **Correlation between environmental factors and bacterial community**
205   CCA was performed to identify the major environmental factors which influence the variation in
206   bacterial communities. The contributions of environmental factors were as followed: available
207   phosphorous (R^2=0.237, P=0.0005) > pH (R^2=0.210, P=0.0015) > humic acid (R^2=0.056,
208   P=0.1559) > ammonia nitrogen (R^2=0.053, P=0.1724) > moisture content (R^2=0.049, P=0.1959).
209   The CCA results indicated that available phosphorous and pH had a greater impact on PM
210   bacterial community structure among these five factors. Furthermore, correlation tests between
211   environment factors and the most abundant 35 genera were performed using spearman's
212   correlation coefficient. The result showed that pH was negatively correlated with the relative
213   abundance of *Clostridium sensu stricto 12* (R= -0.502, P=0.002), *Lachnoclostridium* (R= -0.473,
214   P=0.008) and *Lactobacillus* (R= -0.451, P=0.020), at the genus level (Figure 6).

215   **Discussion**
216   In the study, Illumina Hiseq sequencing platform was used to explore the population diversity of
217   bacteria in middle-layer and sub-layer pit mud of "Yibin Baijiu". According to previous studies,
218   the total number of effective tags of 16s rRNA of microorganisms in pit mud obtained from the
219   Roche 454 or Illumina Miseq sequencing platforms were less than 10,000 (Liu et al., 2017; Tao

220   et al., 2014). In our study, more than 200,000 effective tags were detected in a single sample with
221   an average of 94,023 tags for all samples. Thus, the result of sequencing can adequately reflect
222   almost all of the sample diversity noted.
223   A total of 49 phyla, 332 families and 811 genera were detected, which may be helpful to
224   further understand the phylogenetic status of a massive number of uncultured microbes in the pit
225   mud ecosystem. Moreover, although the relative abundance of Firmicutes, Euryarchaeota,
226   Bacteroidetes and Synergistetes differed slightly between middle-layer and sub-layer samples,
227   they were still the key components of the microbiota in the PM of "Yibin Baijiu". Previous
228   studies also have found that the three most abundant phyla in PM were Firmicutes,
229   Euryarchaeota and Bacteroidetes (Liang et al., 2015; Liu et al., 2017; Tao et al., 2014).
230   Firmicutes is a syntropyhic bacteria that can degrade volatile fatty acids such as butyrate and its
231   analogs, and represented mainly by the Clostridia and Bacilli (Garcia-Peña et al., 2011).
232   Euryarchaeota includes diverse groups of methanogens that are interspersed with non-
233   methanogenic lineages. Methanogens are often symbiotic with hexanoic acid bacteria and can
234   effectively increase the content of ethyl hexanoate in strong-aroma baijiu through the "hydrogen
235   transfer" effect between species (Zhao et al., 2017). Bacteroidetes includes three major types of
236   bacteria: Bacteroidia, Flavobacteriia, and Sphingobacteriia. Most of them are capable of
237   degrading macromolecular carbohydrates such as cellulose to produce acid (Guo, Cheng, Sun,
238   Zhu & Wu, 2014). These phyla in PM all play the vital role in the fermentation of baijiu.
239   In our study, the dominant genera of aroma PM were also very different compared with
240   previous studies. Using PCR-DGGE and high-throughput sequencing technology, Huang et al.
241   (2017) concluded that *Corynebacterium, Myroides, Sphingobacterium, Lactobacillus,*
242   *Clostridium, Acetobacter, Alcaligenes, Enterobacter* and *Acinetobacter* were the dominant
243   bacterial genera in the PM of miscellaneous-aroma baijiu. Wang et al. (2016) detected that
244   *Weissella, Lactobacillus, Bacillus, Enterobacter* and *Paenibacillus* were the dominant bacteria in
245   sauce-aroma baijiu pit mud using nested PCR-DGGE. Wang et al. (2015) proved that
246   *Ruminococcaceae, Thermoanaerobacteriaceae* and *Clostridiaceae* were the dominant bacteria in
247   sauce-aroma baijiu pit mud by Illumina high-throughput sequencing platform. Bian et al. (2012)
248   studied the archaea community structure in sauce-aroma baijiu PM through a culture-
249   independent method named ARDRA, and found that they were mainly affiliated with
250   Euryarchaeota, including *Methanoculleus, Methanosarcina, Methanosaeta,* and
251   *Methanobacterium*.
252   Obviously, dominant bacteria have great contribution to the formation of different aroma types
253   of baijiu. Among baijiu PM of several aroma types, *Lactobacillus* was a dominant genus, thus, it
254   is speculated that *Lactobacillus* was an important bacteria of baijiu brewing. *Gelria* was found to
255   be the dominant genus in our study, which was rarely reported in previous studies. This
256   illustrates the diversity and complexity of the bacterial composition of PM. Furthermore, Fan et
257   al. (2019) reported that *Gelria* in Daqu was associated with volatile compounds such as alcohols,
258   esters, aldehydes, alkanes, benzene and nitrogen compounds in Baijiu. These volatile compounds
259   are recognized as the key odorants in Baijiu, which effect both the flavor and mouth-feel

260   properties. Therefore, as a unique dominant genus of "Yibin Baijiu" PM, we speculated that
261   *Gelria* may contribute to the formation of unique flavour of "Yibin Baijiu".
262   At the genus level, the dominant bacteria in PM at different regions producing strong-aroma
263   baijiu were compared. A total of 11 dominant genera were detected by Liu et al. (2017) in
264   CSAB's PM collected from Luzhou, Sichuan province (about 110 kilometers from Yibin city),
265   among which 8 genera were also the dominant bacteria in "Yibin Baijiu" pit mud. Tao et al.
266   (2014) reported that 14 dominant genera in CSAB's PM which collected from Mianzhu, Sichuan
267   province (about 300 kilometers from Yibin city), 6 of which were different with those of "Yibin
268   Baijiu" pit mud. By comparison, we found that the composition of dominant genera in PM in
269   Mianzhu area was more different from that in Yibin area due to their distant geographical
270   location. Compared with those of Jiangsu area strong-aroma baijiu PM (about 1,700 kilometers
271   away from Yibin city), 7 of the 17 dominant genera in Jiangsu area baijiu PM were shared by
272   "Yibin Baijiu" PM. The dominant genera in Jiangsu area baijiu PM and "Yibin Baijiu" PM were
273   greatly different, which further suggested that geographical location has a great influence on
274   baijiu pit mud bacterial community. Previous studies also have demonstrated that composition of
275   PM bacterial community varied by geographical location (Huang, Wang, Wei, Liu, & Li, 2015;
276   Liang et al., 2016). Interestingly, the dominant genera in the baijiu PM of Hunan region, which is
277   470 kilometers away from Yibin City, was also similar to the dominant genera in "Yibin Baijiu"
278   pit mud. It was speculated that the identical latitude of Hunan and Yibin, where both have the
279   same climate, especially temperature, has a significant impact on the diversity and structure of
280   microbiota. Although the far distance between the two places, dominant genera of these two
281   baijiu pit muds have were similar.
282   The difference between the physicochemical indicators of the middle-layer and sub-layer pit
283   mud of "Yibin Baijiu" was analyzed by t-test. Results showed that the pH of the sub-layer was
284   significantly lower than that of the middle-layer. Meanwhile, CCA analysis showed that the pH
285   value of PM has a great correlation with the bacterial community structure in PM. Spearman
286   analysis revealed that the *Clostridium sensu stricto 12, Lachnoclostridium* and *Lactobacillus*
287   were negatively significantly correlated with pH value. LEfSe also showed that *Lactobacillus*
288   and *Clostridiaceae_1* were significantly less abundant in the middle-layer of PM with higher pH
289   level. Specifically, the above three genera that were significantly related to the pH value all
290   belong to the most abundant phylum Firmicutes in PM. Among them, *Clostridium sensu stricto*
291   *12* and *Lachnoclostridium* were affiliated with order Clostridiales, while *Lactobacillus* was
292   affiliated with order Lactobacillales. Order Clostridiales bacteria have metabolic ability of
293   converting organic substances into organic acids (such as butyric, caproic acids, etc.), alcohols,
294   CO_2 /H_2, and minerals (Kenealy, Cao, & Weimer, 1995; Zheng et al., 2013). Moreover,
295   Clostridiales are important contributors to form ethyl butyrate and ethyl caproate when butyric
296   and caproic acids reacted with alcohols by enzymatic and non-enzymatic catalysis (Ding, Wu,
297   Zhang, Zheng, & Zhou, 2014). Lactobacillales also play an important role during baijiu brewing
298   because the lactic acid they produce can form ethyl lactate. Ethyl lactate, ethyl butyrate and ethyl
299   caproate are vital flavour compounds in CSAB (Fan & Qian, 2005). Thus, pH level largely

300 determines the quality of "Yibin Baijiu". Interestingly, a strong correlation of the bacterial
301 community with the available phosphorous level was also observed by CCA analysis. Available
302 phosphorus provides nutrients for the growth and reproduction of microorganisms. It is a
303 component constituting the cell nuclear membrane and a source of energy substance ATP. Zheng
304 et al. (2013) reported that the low pH and high content of available phosphorus of PM may result
305 in the orthogenesis of the microorganism and Clostridiales, Lactobacillaceae and Bacillales-
306 dominated microbial community structure. However, in the spearman analysis, no bacteria were
307 found to have a significant correlation with available phosphorous level. The reason may be that
308 the spearman analysis only performed a pairwise correlation between environmental factors and
309 the most abundant 35 genera and lacked a reflection of the overall situation. In addition, the
310 available phosphorous level may have a significant correlation with some genera with low
311 abundance in the PM, which were not shown in the spearman's heat map.

Conclusions

313 This study demonstrated the population diversity of bacteria in pit mud of "Yibin Baijiu" from
314 13 distilleries using HiSeq sequencing. Our work suggested the presence of a complex
315 community structure and abundant species in the pit mud of "Yibin Baijiu". No significant
316 difference was observed in bacterial diversity between middle-layer and sub-layer pit mud. The
317 LEfSe analysis showed that the relative abundance of *Lactobacillus* in the sub-layer PM is
318 significantly higher than middle-layer PM. pH was significantly different (p < 0.01) between the
319 two groups. Available Phosphorous and pH had stronger influence on bacterial community
320 structure in PM. Our study provided basic data for further elucidating diversity of microbiota in
321 pit mud of YB and the potential mechanism of Baijiu production.

Acknowledgements

323 This study was supported by Yibin Science and Technology Bureau (2017CNY001-4).

Conflicts of Interest: The authors declare no conflict of interest

References

326 Amann, R. I., Ludwig, W., & Schleifer, K. H. 1995. Phylogenetic identification and in situ
327   detection of individual microbial cells without cultivation. *Microbiological Reviews, 59* (1),
328   143-169. https://doi.org/10.1016/S0882-4010(95)90076-4
329 Barrios-González, J. 2012. Solid-state fermentation: Physiology of solid medium, its molecular
330   basis and applications. *Process Biochemistry, 47* (2), 175-185.
331   https://doi.org/10.1016/j.procbio.2011.11.016
332 Bian, M. H., Ye, G. B., Yang, X. D., Ni, B., & Li, D. Y. 2012. Research on Archaea Community
333   Structure in the Pit Mud of Maotai-flavor Liquor. *Liquor Making Science & Technology* (8),
334   51-53. https://doi.org/10.13746/j.njkj.2012.08.033
335 Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K.,
336   Fierer, N., Pena, A. G., Goodrich, J. K., Gordon, J. I., Huttley, G. A., Kelley, S. T., Knights,

337   D., Koenig, J. E., Ley, R. E., Lozupone, C. A., McDonald, D., Muegge, B. D., Pirrung, M.,
338   Reeder, J., Sevinsky, J. R., Turnbaugh, P. J., Walters, W. A., Widmann, J., Yatsunenko, T.,
339   Zaneveld, J., & Knight, R. 2010. QIIME allows analysis of high-throughput community
340   sequencing data. *Nature Methods, 7* (5), 335-336. https://doi.org/10.1038/nmeth.f.303
341 Chao, A., & Bunge, J. 2002. Estimating the number of species in a stochastic abundance model.
342   *Biometrics, 58* (3), 531-539. https://doi.org/10.1111/j.0006-341X.2002.00531.x
343 Deng, B., Shen, C. H., Shan, X. H., Ao, Z. H., Zhao, J. S., Shen, X. J., & Huang, Z. G. 2012.
344   PCR-DGGE analysis on microbial communities in pit mud of cellars used for different
345   periods of time. *Journal of the Institute of Brewing, 118* (1), 120-126.
346   https://doi.org/10.1002/jib.2
347 Ding, X. F., Wu, C. D., Huang, J., Li, H., & Zhou, R. Q. 2014. Eubacterial and archaeal
348   community characteristics in the man-made pit mud revealed by combined PCR-DGGE and
349   FISH analyses. *Food Research International, 62*, 1047-1053.
350   https://doi.org/10.1016/j.foodres.2014.05.045
351 Ding, X. F., Wu, C. D., Zhang, L. Q., Zheng, J., & Zhou, R. Q. 2014. Characterization of
352   eubacterial and archaeal community diversity in the pit mud of Chinese Luzhou - flavor
353   liquor by nested PCR–DGGE. *World Journal of Microbiology Biotechnology, 30* (2), 605.
354   https://doi.org/10.1007/s11274-013-1472-4
355 Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., & Knight, R. 2011. UCHIME improves
356   sensitivity and speed of chimera detection. *Bioinformatics, 27* (16), 2194-2200.
357   https://doi.org/10.1093/bioinformatics/btr381
358 Fan, G., Fu, Z., Sun, B., Zhang, Y., Wang, X., Xia, Y., Huang M., & Li, X. 2019. Roles of aging
359   in the production of light-flavored Daqu. *Journal of bioscience and bioengineering*, 127 (3),
360   309-317. https://doi.org/10.1016/j.jbiosc.2018.08.005
361 Fan, W., & Qian, M. C. 2005. Headspace Solid Phase Microextraction and Gas
362   Chromatography−Olfactometry Dilution Analysis of Young and Aged Chinese "Yanghe
363   Daqu" Liquors. *Journal of Agricultural and Food Chemistry, 53* (20), 7931-7938.
364   https://doi.org/10.1021/jf051011k
365 Fan, W., & Qian, M. C. 2006. Characterization of aroma compounds of chinese "Wuliangye"
366   and "Jiannanchun" liquors by aroma extract dilution analysis. *Journal of Agricultural and
367   Food Chemistry, 54* (7), 2695-2704. https://doi.org/10.1021/jf052635t
368 Garcia-Peña, E. I., Parameswaran, P., Kang, D. W., Canul-Chan, M., and Krajmalnik-Brown, R.
369   2011. Anaerobic digestion and co-digestion processes of vegetable and fruit residues:
370   Process and microbial ecology, *Bioresour. Technol.* 102, 9447–9455.
371   https://doi.org/10.1016/j.biortech.2011.07.068
372 Good, I. J. 1953. The Population Frequency of Species and the Estimation of Population
373   Parameters. *Biometrika, 40* (3/4), 237-264. https://doi.org/10.1093/biomet/40.3-4.237
374 Guo, X. H., Cheng, W., Sun F. O., Zhu, W. J. & Wu, W. X. 2014. A comparison of microbial
375   characteristics between the thermophilic and mesophilic anaerobic digesters exposed to

376   elevated food waste loadings, Bioresour Technol, 152: 420-428.
377   https://doi.org/10.1016/j.biortech.2013.11.012
378   Hu, X., Du, H., Ren, C., & Xu, Y. 2016. Illuminating Anaerobic Microbial Community and
379   Cooccurrence Patterns across a Quality Gradient in Chinese Liquor Fermentation Pit Muds.
380   *Applied and environmental microbiology, 82* (8), 2506-2515.
381   https://doi.org/10.1128/AEM.03409-15
382   Huang, Y. N., Xiong, X. M., Hu, Y. L., Liang, Y. X., & Zhao, S. M. 2017. Bacterial community
383   and diversity in pit mud of Baiyunbian liquor analyzed by PCR-DGGE and high-throughput
384   sequencing. *Microbiology China, 44* (2), 376-383.
385   https://doi.org/10.13344/j.microbiol.china.160080
386   Huang, Z. G., Wang, Y. L., Wei, C. H., Liu, Y. M., & Li, M. 2015. Diversity of Bacteria
387   Community Structure in Pit Mud from Different Regions and Analysis of Its Correlations
388   with Physicochemical Factors. *Liquor Making Science & Technology* (11), 18-25.
389   https://doi.org/10.13746/j.njkj.2015097
390   Jin, G., Zhu, Y., & Xu, Y. 2017. Mystery behind Chinese liquor fermentation. *Trends in Food*
391   *Science & Technology, 63*, 18-28. https://doi.org/10.1016/j.tifs.2017.02.016
392   Kenealy, W. R., Cao, Y., & Weimer, P. J. 1995. Production of caproic acid by cocultures of
393   ruminal cellulolytic bacteria and Clostridium kluyveri grown on cellulose and ethanol.
394   *Applied Microbiology Biotechnology, 44* (3-4), 507-513.
395   https://doi.org/10.1007/BF00169952
396   Liang, H., Li, W., Luo, Q., Liu, C., Wu, Z., & Zhang, W. 2015. Analysis of the bacterial
397   community in aged and aging pit mud of Chinese Luzhou-flavour liquor by combined
398   PCR-DGGE and quantitative PCR assay. *Journal of the Science of Food and Agriculture,*
399   *95* (13), 2729-2735. https://doi.org/10.1002/jsfa.7013
400   Liang, H. P., Luo, Q. C., Zhang, A., Wu, Z. Y., & Zhang, W. X. 2016. Comparison of bacterial
401   community in matured and degenerated pit mud from Chinese Luzhou-flavour liquor
402   distillery in different regions. *Journal of the Institute of Brewing, 122* (1), 48-54.
403   https://doi.org/10.1002/jib.296
404   Liu, H., & Sun, B. 2018. Effect of Fermentation Processing on the Flavor of Baijiu. *Journal of*
405   *Agricultural and Food Chemistry, 66* (22), 5425-5432.
406   https://doi.org/10.1021/acs.jafc.8b00692
407   Liu, M., Tang, Y., Zhao, K., Liu, Y., Guo, X., Ren, D., Yao, W., Tian, X., Gu, Y., Yi, B., &
408   Zhang, X. 2017. Determination of the fungal community of pit mud in fermentation cellars
409   for Chinese strong-flavor liquor, using DGGE and Illumina MiSeq sequencing. *Food*
410   *Research International, 91*, 80-87. https://doi.org/10.1016/j.foodres.2016.11.037
411   Liu, M. K., Tang, Y. M., Guo, X. J., Zhao, K., Tian, X. H., Liu, Y., Yao, W. C., Deng, B., Ren,
412   D. Q., & Zhang, X. P. 2017. Deep sequencing reveals high bacterial diversity and
413   phylogenetic novelty in pit mud from Luzhou Laojiao cellars for Chinese strong-flavor
414   Baijiu. *Food Research International, 102*, 68-76.
415   https://doi.org/10.1016/j.foodres.2017.09.075

416 Liu, M. K., Tang, Y. M., Zhao, K., Gu Y. F., Ren, D. Q., Yao, W. C., Tian, X. H., Liu, Y., &
417   Zhang, X. P. 2017. Recent advances in research on the community, isolation, and
418   application of microbes in the pit mud used in manufacture of Chinese Strong-flavor Baijiu.
419   *Microbiology China, 44* (5), 1222-1229. https://doi.org/10.13344/j.microbiol.china.160559
420 McGovern, P. E., Zhang, J., Tang, J., Zhang, Z., Hall, G. R., Moreau, R. A., Nunez, A., Butrym,
421   E. D., Richards, M. P., Wang, C. S., Cheng, G., Zhao, Z., & Wang, C. 2004. Fermented
422   beverages of pre- and proto-historic China. *Proceedings of the National Academy of*
423   *Sciences of the United States of America, 101* (51), 17593-17598.
424   https://doi.org/10.1073/pnas.0407921102
425 Segata, N., Izard, J., Waldron, L., Gevers, D., Miropolsky, L., Garrett, W. S., & Huttenhower, C.
426   2011. Metagenomic biomarker discovery and explanation. *Genome Biology, 12* (6), R60.
427   https://doi.org/10.1186/gb-2011-12-6-r60
428 Shannon, C. E. 1997. The mathematical theory of communication. 1963. *MD Computers, 14* (4),
429   306-317. https://doi.org/10.1063/1.3067010
430 Shen, Y. F. 2014. *Handbook of Chinese liquor making technology (2nd ed)*. Beijing: China Light
431   Industry Press.
432 Tang S. Y., Wang R., Chen X. W., Long P., Wang F. & Chen L. 2012. Review of Wuliangye old
433   pit mud Research. *Food and Fermentation Technology*, *48* (6): 7-11.
434   https://doi.org/10.3969/j.issn.1674-506X.2012.06-00
435 Tao, Y., Li, J. B., Rui, J. P., Xu, Z. C., Zhou, Y., Hu, X. H., Wang, X., Liu, M. H., Li, D. P., &
436   Li, X. Z. 2014. Prokaryotic Communities in Pit Mud from Different-Aged Cellars Used for
437   the Production of Chinese Strong-Flavored Liquor. *Applied and environmental*
438   *microbiology, 80* (7), 2254-2260. https://doi.org/10.1128/AEM.04070-13
439 Tao, Y., Rui, J. P., Li, J. B., Xu, Z. C., Li, D. P., Tang, Q. L., Wei, Y., & Li, X. Z. 2014.
440   Microbial community compositions and diversity in pit mud of Chinese Luzhou-flavor
441   liquor. *CIESC Journal, 65* (5), 1801-1807. https://doi.org/10.3969/j.issn.0438-
442   1157.2014.05.034
443 Wang, L., Wang, Y. Y., Wang, H. Y., Liu, G., Yang, F., Jiang, H. G., Jin, X. U., Wang, D. Q., &
444   Jin, T. 2015. Microbial Composition of Bottom Pit Mud in Jiangxiang Baijiu(Liquor) Pit.
445   *Liquor Making Science & Technology* (1), 12-15. https://doi.org/10.13746/j.njkj.2014513
446 Wang, Q., Zhang, H., & Liu, X. 2016. Microbial Community Composition Associated with
447   Maotai Liquor Fermentation. *Journal of Food Science, 81* (6), M1485-1494.
448   https://doi.org/10.1111/1750-3841.13319
449 Wang, X., Du, H., & Xu, Y. 2017. Source tracking of prokaryotic communities in fermented
450   grain of Chinese strong-flavor liquor. *International Journal of Food Microbiology, 244*, 27-
451   35. https://doi.org/10.1016/j.ijfoodmicro.2016.12.018
452 Wu, Y. Y., Qi, Y. P., Xue, C. J., Guo, S. Z., & Jie, L. X. 1980. Microbial ecology distribution
453   and anaerobic fermentation characteristics in the pit mud of Luzhou-flavor liquor.
454   *Microbiology China* (3), 108-112. https://doi.org/10.13344/j.microbiol.china.2007.02.013

455 Yang, M. M., Chen, L., Liu, Q., Li, H., Hu, Y. J., Liu, Y. M., & Yang, Z. J. 2017. Correlation
456   analysis between physicochemical indexes and sensory evaluation of the geographical
457   indication product "Yibin Jiu" base liquor. *China Brewing, 36* (01), 142-145.
458   https://doi.org/10.11882/j.issn.0254-5071.2017.01.030

459 Ye, G. B., Luo, H. B., Yang, X. D., Li, D. Y., Wang, Y., & Ni, B. 2013. Community Stricture of
460   Prokaryotes in Pit Mud of Lu-Flavor Liquor from Luzhou Prefecture Based on Culture-
461   independent Approach. *Journal of Food Science, 34* (17). https://doi.org/10.7506/spkx1002-
462   6630-201317038

463 You, L., Wang, S., Zhou, R., Hu, X., Chu, Y., & Wang, T. 2016. Characteristics of yeast flora in
464   Chinese strong-flavoured liquor fermentation in the Yibin region of China. *Journal of the
465   Institute of Brewing, 122* (3), 517-523. https://doi.org/10.1002/jib.352

466 Yue, Y. Y., Hang, W. X., Liu, X., Hu, C., & Zheng, S. Y. 2007. Isolation and Identification of
467   Facultative Anaerobes in the Pit Mud of Chinese Luzhou-flavor Liquor. *Microbiology
468   China, 34* (2), 251-255. https://doi.org/10.3969/j.issn.0253-2654.2007.02.013

469 Zhao, D., Zheng, J., Peng, Z. Y., Lv, X. L., Yang, K. Z. & Zhang, J. M. 2017. Prokaryotic
470   microbial community in wuliangye pit mud using nextseq high-throughput sequencing
471   technology. *Food and Fermentation Industries,* (09), 5-12.
472   https://doi.org/10.13995/j.cnki.11－1802/ts.014156

473 Zhao, J. S., Zheng, J., Zhou, R. Q., & Shi, B. 2012. Microbial community structure of pit mud in
474   a Chinese strong aromatic liquor fermentation pit. *Journal of the Institute of Brewing, 118*
475   (4), 356-360. https://doi.org/10.1002/jib.52

476 Zhao X. L., Zhou Y. K., Qiao Z. W., Li H., Li W. H., Liu B. H., Yang Q. S., Zhang L. J., Du S.
477   T., Xu Z. H. & Yang J. Z. 2013. A Study on Unique Climate Characteristics of Yibin of
478   Sichuan in Liquor Production. *Liquor Making*, (5):5-10.

479 Zheng, J., Liang, R., Zhang, L., Wu, C., Zhou, R., & Liao, X. 2013. Characterization of
480   microbial communities in strong aromatic liquor fermentation pit muds of different ages
481   assessed by combined DGGE and PLFA analyses. *Food Research International, 54* (1),
482   660-666. https://doi.org/10.1016/j.foodres.2013.07.058

483 Zheng, Q., Lin, B., Wang, Y., Zhang, Q., He, X., Yang, P., Zhou, J., Guan, X., & Huang, X.
484   2015. Proteomic and high-throughput analysis of protein expression and microbial diversity
485   of microbes from 30- and 300-year pit muds of Chinese Luzhou-flavor liquor. *Food
486   Research International, 75*, 305-314. https://doi.org/10.1016/j.foodres.2015.06.029

487 Zheng, X. W., & Han, B. Z. 2016. Baijiu (白酒), Chinese liquor: History, classification and
488   manufacture. *Journal of Ethnic Foods, 3* (1), 19-25.
489   https://doi.org/10.1016/j.jef.2016.03.001

490 Zou, W., Zhao, C., & Luo, H. 2018. Diversity and Function of Microbial Community in Chinese
491   Strong-Flavor Baijiu Ecosystem: A Review. *Front Microbiol, 9*, 671.
492   https://doi.org/10.3389/fmicb.2018.00671

Result of physicochemical indicators of middle and sub layer groups in PMs

**: difference is significant at the 0.01 level.

1 **Table 1** Result of physicochemical indicators of middle and sub layer groups in PMs

Sampling	Value for indicated PM sample (mean ± SD)		P-value
	Middle	**Sub**	
moisture content (%)	38.03 ± 0.07	39.32 ± 0.08	0.509
pH**	6.00 ± 1.49	4.69 ± 1.06	0.000
ammonia nitrogen (mg/100g)	58.51 ± 23.80	54.16 ± 22.85	0.448
available phosphorous (mg/100g)	106.55 ± 118.52	90.31 ± 107.68	0.560
humic acid (‰)	3.10 ± 1.67	3.10 ±1.35	0.999

2 **: difference is significant at the 0.01 level.

Number of classified species at different taxonomic taxa

Table 2 Number of classified species at different taxonomic taxa

Group	Phylum	Class	Order	Family	Genus
Middle	45	88	151	291	680
Sub	46	88	160	304	713
All	49	94	172	332	811

Comparison of dominant microorganism in PM in different regions of strong-aroma baijiu

*Ratio: The ratio of the number of dominant genera which shared with "Yibin Baijiu" pit mud to the total number of dominant genera.

1 **Table 3** Comparison of dominant microorganism in PM in different regions of strong-aroma baijiu

Locations	Dominant microorganism	Ratio*	Reference
Mianzhu (31°43′N, 104°12′E) 300 km	*Lactobacillus, Petrimonas, Clostridium IV, Sedimentibacter, Syntrophomonas, Spirochaetes SHA-4, unclassified Porphyromonadaceae, Anaerobrancaceae, Clostridiaceae 1, Ruminococcaceae, Methanoculleus, Methanosarcina, Methanobacterium, Methanobrevibacter*	8/14	(Tao et al., 2014)
Luzhou (28°53′N, 105°27′E) 70 km	*Methanobrevibacter, Caproiciproducens, Petrimonas, Lactobacillus, Sedimentibacter, Proteiniphilum, Syntrophomonas, Aminobacterium, Christensenellaceae R-7, Caldicoprobacter, Olsenella*	8/11	(Liu, Tang, Guo, et al., 2017)
Jiangsu (35°20′N, 121°57′E) 1700 km	*Lactobacillus, Ruminococcus, Caloramator, Clostridium, Sedimentibacter, Syntrophomonas, Sporanaerobacter, Pelotomaculum, T78, Prevotella, Blvii28, Methanobacterium, Methanobrevibacter, Methanosaeta, Methanoculleus, Methanosarcina, Nitrososphaera*	7/17	(Hu, Du, Ren, & Xu, 2016)
Hunan (28°21′N, 109°45′E) 470 km	*Petrimonas, Lactobacillus, Sedimentibacter, Clostridium, Ruminococcus, Syntrophomonas, Symbiobacterium, Methanobacterium, Methanobrevibacter, Methanoculleus Methanosarcina*	8/11	(Wang, Du, & Xu, 2017)
Yibin (28°76′N, 104°63′E) 0 km	*Ruminiclostridium 5, Lactobacillus, Gelria, Methanoculleus, Syntrophomonas, Syntrophaceticus, Aminobacterium, Proteiniphilum, Petrimonas, Sedimentibacter, Methanobrevibacter, Caldicoprobacter, Methanosarcina and Methanobacterium*	14/14	This work
Unique	*Ruminiclostridium 5, Gelria, Syntrophaceticus*	——	This work
Common	*Lactobacillus, Sedimentibacter, Syntrophomonas, Methanobrevibacter*	——	All

2 *Ratio: The ratio of the number of dominant genera which shared with "Yibin Baijiu" pit mud to the total

3 number of dominant genera.

Comparison of dominant microorganism in PM of different aroma types of baijiu

1   **Table 4** Comparison of dominant microorganism in PM of different aroma types of baijiu

Representative aroma	Dominant microorganism	Reference
miscellaneous-aroma	*Corynebacterium, Myroides, Sphingobacterium, Lactobacillus, Clostridium, Acetobacter, Alcaligenes, Enterobacter and Acinetobacter*	(Huang et al., 2017)
sauce-aroma	Clostridiaceae, Ruminococcaceae and Thermoanaerobacteriaceae	(Wang et al., 2015)
	Methanoculleus, Methanosarcina, Methanosaeta and Methanobacterium	(Bian, Ye, Yang, Ni, & Li, 2012)
	Weissella, Lactobacillus, Bacillus, Enterobacter and Paenibacillus	(Wang, Zhang, & Liu, 2016)
strong-aroma (common)	*Lactobacillus, Sedimentibacter, Syntrophomonas and Methanobrevibacter*	——

Figure 1

The Chao1 richness indices (A) and the Shannon diversity indices (B).

The Box-plots exhibit the first (25%) and third (75%) quartiles, the median and the maximum and minimum observed values within each data set.
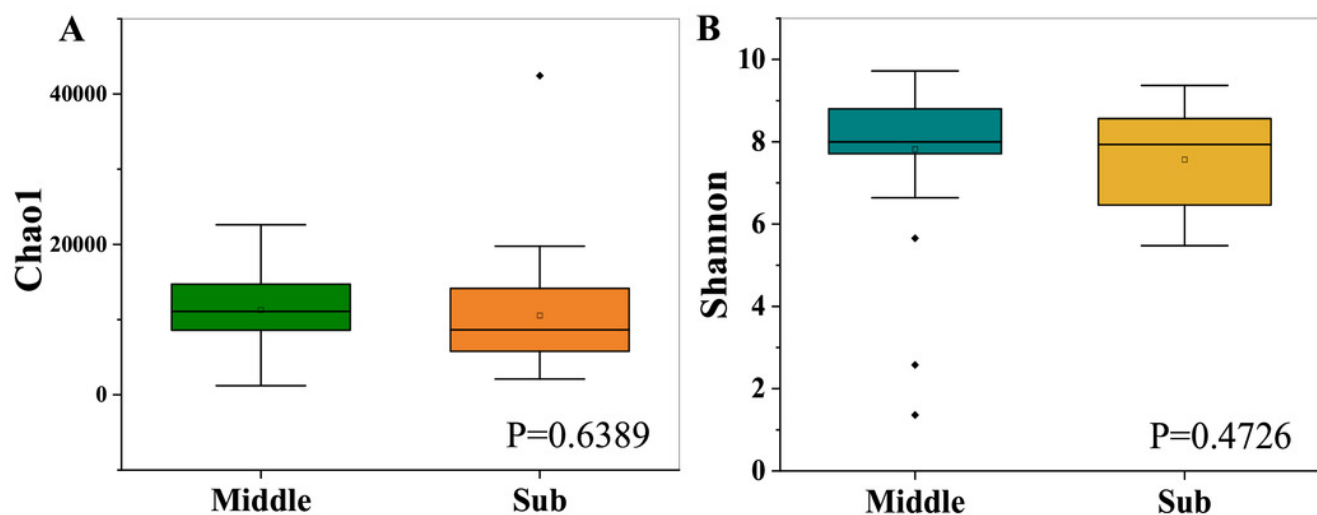
Figure 2

PCoA of Weighted Unifrac distance (A) and Unweighted Unifrac distance (B) for the microbiota at PM of middle-layer group and sub-layer group.

The principal coordinate combination with the highest contribution rate is selected for graph display.
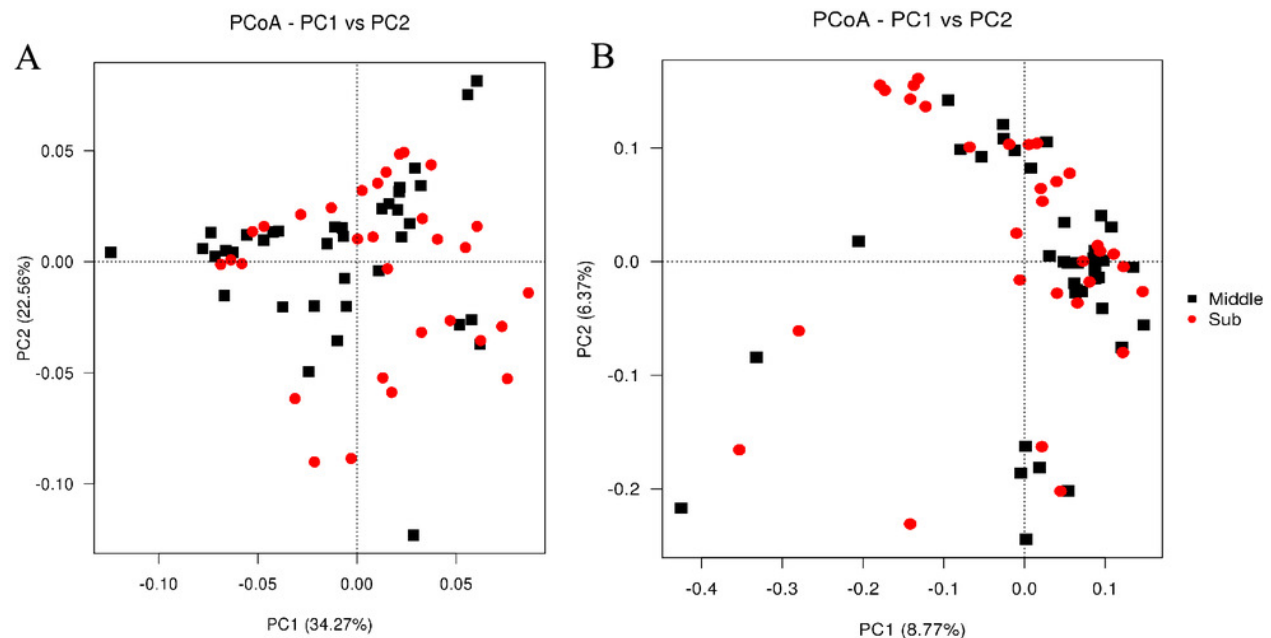
Figure 3

Taxonomic classifications of the bacterial communities in PM samples from sub-layer and middle-layer at the phylum level.

The "Others" refers to the group which could not be accurately assigned to any known bacterial taxonomic group at the phylum level, as well as the group which ranked after 10 phyla with most relative abundance.
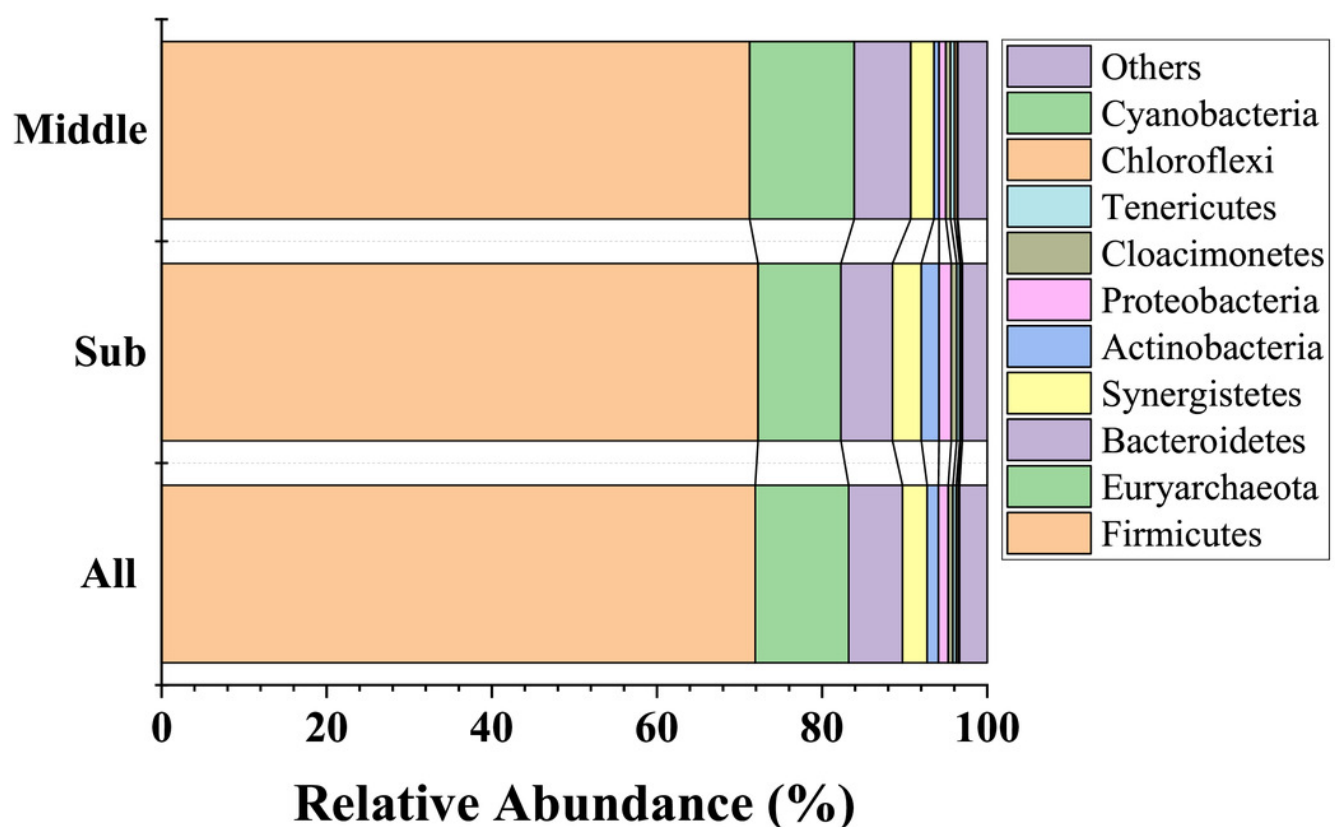
Figure 4

The most abundant 10 genera in Bacterial community at genus level.
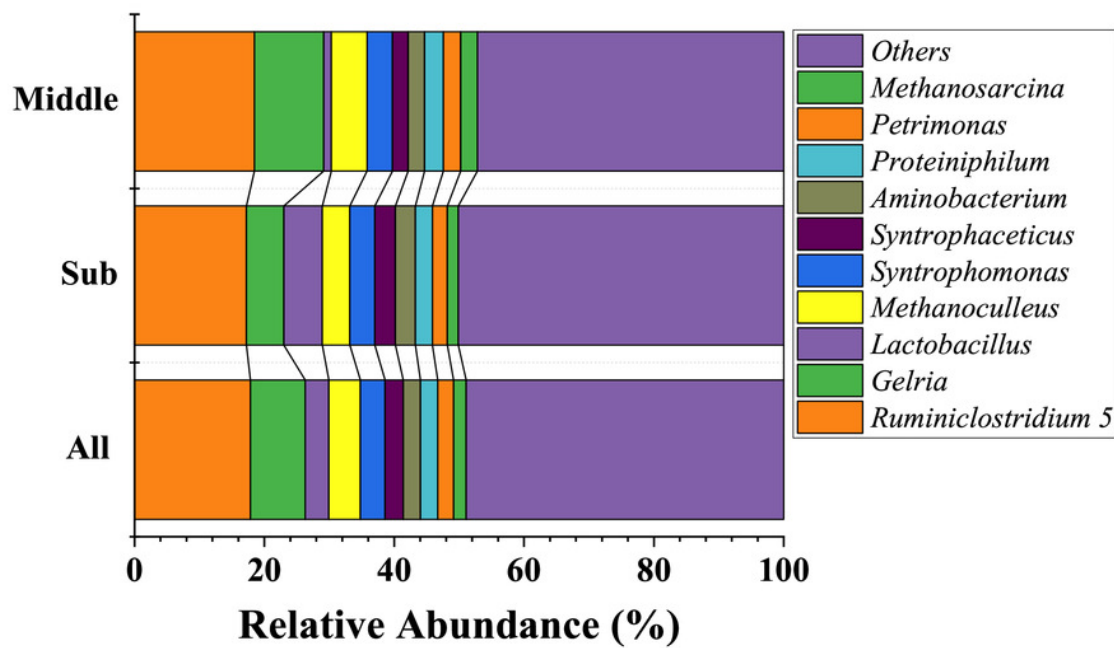
Figure 5

LDA score map with a threshold value of 4.0.

The length of the histogram represents the effect size of the species with significant difference in abundance in different groups.
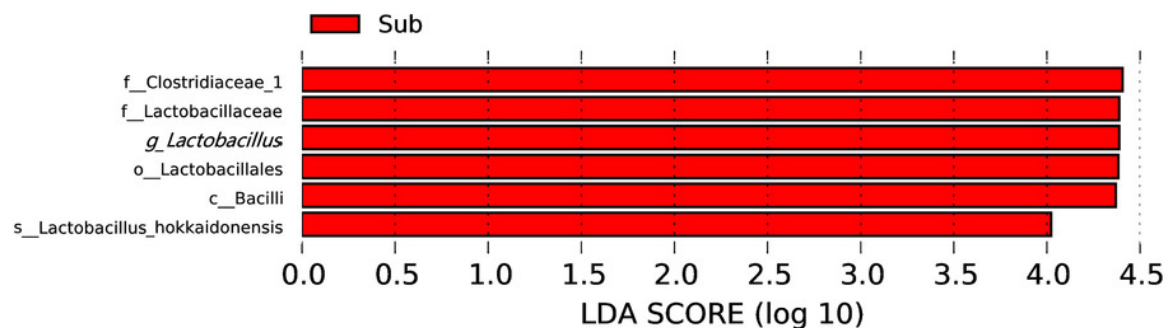
Figure 6

Heat map showing the spearman correlation coefficient of environmental factors and bacterial community in PM.

The horizontal direction is the species information, the longitudinal direction is the environmental factor information, the heat map corresponds to the Spearman correlation coefficient R, and the R value ranges from -1 to 1, R<0 is a negative correlation, and R>0 is a positive correlation.