Environmental Microbiome

RESEARCH

Open Access

Structure and metabolic function of spatiotemporal pit mud microbiome

Zhihua Li[1,2,3*], Chi Zhao[1], Zhenyu Mao[2,3], Fengju Zhang[1], Ling Dong[1], Chuan Song[2,3], Yao Chen[2,3], Xin Fu[2,3], Zonghua Ao[2,3], Yanfei Xiong[2,3], Qin Hui[2,3], Weizhi Song[4], Petri Penttinen[5,6] and Suyi Zhang[2,3*]

Abstract

Background Pit mud (PM) hosts diverse microbial communities, which serve as a medium to impart flavor and quality to Baijiu and exhibit long-term tolerance to ethanol and acids, resulting in a unique ecosystem. However, the ecology and metabolic functions of PM remain poorly understood, as many taxa in PM represent largely novel lineages. In this study, we used a combination of metagenomic analysis and chemical derivatization LC–MS analysis to provide a comprehensive overview of microbial community structure, metabolic function, phylogeny, horizontal gene transfer, and the relationship with carboxyl compounds in spatiotemporal PM samples.

Results Our findings revealed three distinct stages in the spatiotemporal changes of prokaryotic communities in PM: an initial phase dominated by *Lactobacillus*, a transitional phase, and a final state of equilibrium. Significant variations in α- and β-diversity were observed across different spatial and temporal PM samples. We identified 178 medium- and high-quality non-redundant metagenome-assembled genomes (MAGs), and constructed their phylogenetic tree, depicting their roles in the carbon, nitrogen, and sulfur cycles. The Wood-Ljungdahl pathway and reverse TCA cycle were identified as the main carbon fixation mechanisms, with both hydrogenotrophic and aceticlastic methanogens playing a major role in methane production, and methylotrophic pathway observed in older PM. Furthermore, we identified relationships between prokaryotes and 29 carboxyl metabolites, including medium- and long-chain fatty acids. Horizontal gene transfer (HGT) was widespread in PM, particularly among clostridia, Bacteroidota, Bacilli, and Euryarchaeota, and was shown to play critical roles in fermentation dynamics, carbon fixation, methane production, and nitrogen and sulfur metabolism.

Conclusion Our study provides new insights into the evolution and function of spatiotemporal PM, as well as its interactions with carboxyl metabolites. *Lactobacillus* dominated in new PM, while methanogens and clostridia were predominant in older or deeper PM layers. The three distinct stages of prokaryotic community development in PM and HGT played critical roles in metabolic function of spatiotemporal PM. Furthermore, this study highlights the importance of α-diversity, β-diversity, methanogens, and *Clostridium* as useful indicators for assessing PM quality in the production of high-quality Baijiu.

Keywords Pit mud, Evolutionary divergence, Metagenome assembled genomes, Horizontal gene transfer, Carboxyl metabolites

*Correspondence:
Zhihua Li
lzh82@hotmail.com
Suyi Zhang
zhangsy@lzlj.com
Full list of author information is available at the end of the article

Introduction

Baijiu, one of the six oldest distilled liquors originating from China, is distinguished by its unique solid-state fermentation process and deep connections to Chinese culture [1]. The distinct flavor profiles of Baijiu are categorized into four main aroma types: strong, light, sauce, and rice aroma taste characteristics [2]. Among these, Chinese strong-flavor Baijiu (CSFB), also known as Luzhou-flavor Baijiu, dominates Chinese Baijiu production, with an impressive annual output of 9.1 million tons, accounting for nearly 70% of total Baijiu production [3]. Typically, CSFB production relies primarily on sorghum as the main raw material, often supplemented with a mixture of other grains such as corn, wheat, sticky rice, and regular rice [4]. The overall production process entails several steps: (1) crushing and mixing the raw materials with a specific ratio of microbial starter Daqu-starter (w/w) to form Jiupei, (2) anaerobically fermenting the mixture in a cellar for a minimum of 60 days, (3) distilling the fermented material to obtain Chinese Baijiu, and (4) aging the resulting strong-flavor baijiu (SFB) in either pottery jars or stainless-steel vessels to enhance its organoleptic qualities [5].

Pit mud (PM) that covers the interior walls of the rectangular fermenting cellars, enclosing the Jiupei (fermented grains), is home to diverse microbial communities [6–8] (Fig. 1). This unique underground fermentation process involves multiple rounds of recycling and occurs in an anaerobic environment, producing ethanol and various organic acids [9]. Continuous batch-to-batch manufacturing activities shape and domesticate the microbial consortia within the PM. Notably, the highest quality CSFB is produced in old cellars that have been in continuous use [10, 11]. One prominent example is the Luzhou Laojiao cellars, located in Luzhou, Sichuan, China. These cellars have been in operation for over 400 years, since the Ming dynasty (1573 CE), and are recognized and protected by the National Cultural Heritage Conservation Board of China. They have also been proposed for inclusion on the Tentative List of World Cultural Heritage sites. The aged PM not only serves as a fermentation vessel, imparting flavor and quality to CSFB, but it is also regarded as a distinct type of fermented soil [10–12], offering valuable insights into long-term microbial tolerance to ethanol and acids under anaerobic conditions.

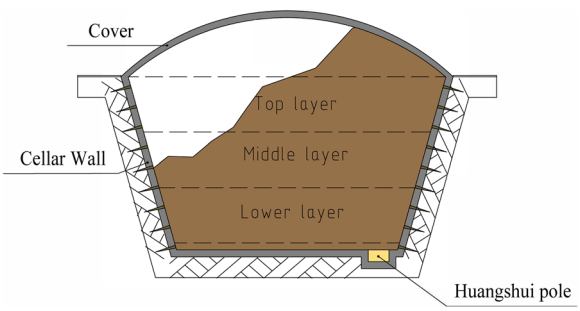


Fig. 1 A diagram of a CSFB fermenting cellar

PM represents a complex ecosystem, housing a wide range of bacteria and archaea, particularly anaerobic microorganisms [12]. The microbial communities in cellars of different ages undergo noticeable shifts, often exhibiting mutual cooperation [13]. PM quality is positively correlated with the presence of *Clostridium* and *Methanobacteria*, whereas lactic acid bacteria (such as *Lactobacillus*, *Pediococcus*, and *Streptococcus*) are negatively associated with PM quality [6]. For instance, the relative abundance of *Clostridia* which contribute to the production of carboxyl metabolites, such as butyric acid and caproic acid, is lower in degraded PM than in both normal and high-quality PM [3–14]. Furthermore, the microbial communities and corresponding carboxyl compounds varied across PM quality grades [14]. However, the ecology and metabolic functions of PM microbial communities remain poorly understood.

Currently, the taxonomic classification of metagenomic DNA sequences from PM is largely incomplete, with many unclassified, uncultured, and yet-to-be-characterized species [7, 8, 15]. These species are often referred to as "microbial dark matter" (MDM) [16, 17]. Microbial metabolic activity plays a crucial role in driving nutrient and energy transformations, as well as maintaining community and functional stability in response to environmental fluctuations [18]. However, knowledge on the MDM species and their metabolic functions in PM remains limited. Recently, scalable software like METABOLIC has enabled metagenome-based interpretation of the microbial community's metabolic and functional potential at the individual genome level, providing insights into biogeochemical cycling, such as carbon, nitrogen, and sulfur cycles [19]. Moreover, chemical derivatization LC–MS is a highly effective technique for identifying and quantifying a wide range of metabolites [20].

We employed metagenomic analysis and chemical derivatization LC–MS analysis to investigate the community structure and metabolic functions of microorganisms in PM. The specific objectives were: (i) to characterize the spatiotemporal microbial community structure and genetic diversity in PM, (ii) to unravel the functions within these complex communities, (iii) to investigate evolutionary processes, including horizontal gene transfer, in PM microbial communities, and (iv) to reveal the relationships between carboxyl metabolites and microbial community structure. The scientific questions are

microbial structure, metabolic function, phylogenetic, horizontal gene transfer as well as the relationship with carboxyl compounds in spatiotemporal PM. The findings of this study will enhance our understanding of the spatiotemporal PM microbiome at a mechanistic level and contribute to optimizing high-quality CSFB production.

Methods

Sample collection, DNA extraction, and sequencing

PM samples were collected on June 21, 2022 and April 22, 2023, at Luzhou Laojiao Group Co., Ltd., in Luzhou (105° 29′ 50″ E, 28° 53′ 47″ N), Sichuan, China. Subsamples (5 g PM per position) were taken from upper, middle, and lower layers on four sides of each 4.6 m long × 3.1 m wide × 2.5 m high cellar (Fig. 1). A total of 144 samples from six old cellars (in operation for more than 100 years) and six new cellars (less than 10 years) were aseptically collected in sterilized plastic bags, immediately cold-transferred to the laboratory, and stored at − 80 °C until further analysis. The four samples from the same layer in each cellar were combined into one composite sample. Total genomic DNA was extracted using the FastDNA SPIN Kit for Soil (MP Biomedicals, USA) according to the manufacturer's protocol. DNA concentration and quality were evaluated using a Nanodrop2000 spectrophotometer (Thermo Scientific, USA) and quantified using a Qubit 3.0 system (Life Invitrogen, USA). Shotgun metagenome sequencing libraries were constructed and sequenced on an Illumina NovaSeq 6000 platform (Illumina Inc, CA, USA), using 2 × 150 bp paired-end (PE150) sequencing. Raw sequence reads were quality-filtered, and adapter sequences and contaminant reads were removed using fastp [21]. After preprocessing, the remaining 2,175,056,900 reads, averaging 60,418,247 reads per sample, were used for further analysis.

Read-based profiling and genomic binning

Microbial community profiling were carried out using Kraken2 with *k-mer* matching to classify reads, followed by estimating relative abundance with Bracken [22, 23]. SPAdes (v3.15.4) was used for metagenomic assembly in meta-mode with default parameters [24]. Draft genomes (bins) were generated using metaBAT2, MaxBin2, and vamb. MaxBin2 results were further refined by GraphBin, which utilized the assembly graph of Maxbin2, and the results from the individual binning approaches were used as inputs for genomic binning consolidated by metaWRAP v1.3.2 (bin_refinement -A metabat2_bins/ -B GraphBin_bins/ -C vamb_bins/ -c 50 -x 10) [25]. The refined bins were combined and dereplicated using dRep (v3.2.2) at a 95% average nucleotide identity (ANI) threshold. Completeness, contamination, and heterogeneity of reconstructed metagenome-assembled genomes (MAGs) were assessed using CheckM (v.1.0.5). A total of 178 high-quality MAGs (completeness > 75% and contamination < 10%) were retained for subsequent analysis [26]. Taxonomic classification was carried out using GTDB-Tk v1.7.0 with the parameters "classify_wf" based on the Genome Taxonomy Database (GTDB) (release 202). Phylogenetic analysis was conducted using IQ-TREE v2.1.2. Detailed MAG taxonomy information is provided in Supplementary Table S1.

MAGs functional annotation

Functional annotation was performed using METABOLIC v4.0, which integrates Kyoto Encyclopedia of Genes and Genomes (KEGG), TIGRfam, Pfam, dbCAN2, and MEROPS databases, along with a protein motif validation step to assess the presence of metabolic pathways based on KEGG modules [19]. Metabolic profiling and functional assessment were conducted using metabolic weight score (MW-score) based on metabolic profiling and gene coverage from metagenomic reads [19].

Horizontal gene transfer identification

Horizontal gene transfers (HGTs) were identified using MetaCHIP across 89 high-quality MAGs (completeness > 90% and contamination < 5%). The inputs for MetaCHIP were MAGs and their taxonomic classifications based on the Genome Taxonomy Database (GTDB), specified at user-defined ranks [27]. The function of HGTs were annotated using eggNOG-Mapper v2.1.12 (http://eggnog-mapper.embl.de/) [28], BlastKOALA and GhostKOALA [29].

Carboxyl metabolites analysis

Carboxyl compound analysis was conducted using chemical derivatization (DmPA bromide) isotope labeling liquid chromatography-mass spectrometry (CIL–LC–MS) following previous protocols [30]. In brief, 100 mg PM in 0.5 mL of LC–MS-grade MeOH/water (4:1, v/v) was homogenized twice in an ice bath using a homogenizer (PRO Scientific, Oxford, CT), incubated at − 20 °C for 10 min, and centrifuged at 12,000 rpm and 4 °C for 10 min, followed by collecting the supernatant. Aliquots of 25 µL and a pooled sample were dried and re-dissolved in 25 µL of LC–MS-grade ACN/water (3:1, v/v) for DmPA bromide derivatization for carboxyl labeling as per the SOP manual. A 10 µL aliquot was directly injected for UHPLC-Q-TOF/MS analysis (Agilent 1290 LC linked to Agilent 6546 Q-TOF Mass Spectrometer). LC–MS analysis, data processing, and metabolite identification were conducted as described previously [30].

Statistical analysis and visualization

Statistical analysis and visualization, including α-diversity (Shannon and Pielou indices), β-diversity (Bray–Curtis dissimilarity), one-way anova and linear discriminant analysis effect size (LEfSe) [31] analyses, Mantel and Wilcoxon rank sum tests, and cladograms of differential features, were conducted using the microeco package (https://chiliubio.github.io/microeco_tutorial/) [32]. Phylogenetic trees were visualized using iTOL (https://itol.embl.de/) through automated annotation file generation [33, 34].

Results

α- and β-diversity

Shannon and Pielou indices were higher in the old PM than in the new PM ($P < 0.05$) (Fig. 2A, B). In both the old and new PM, Shannon and Pielou indices were higher in the lower layer than in the upper layer ($P < 0.05$) (Fig. 2C–F), suggesting that the α-diversity decreases with increasing spatial position within PM. The microbial community compositions in the new and old PM were clearly different (Wilcoxon rank sum test adjusted P value = 7.2e−51) (Fig. 2G, H), with greater within-group dissimilarity in old PM than in new PM (Fig. 2H).

Microbial community structure

Assessing the microbial community composition revealed that the relative abundances of *Methanoculleus*, *Methanosarcina*, *Methanobacterium*, *Clostridium*, *Petrimonas, Paenibacillus*, and *Caproiciproducens* were higher and those of *Lactobacillus* (e.g., bin.4) and *Acetilactobacillus* (e.g., bin.1146) were lower in old PM than in new PM (Fig. 3A) (Supplementary Table S1). Interestingly, the relative abundances of *Lactobacillus* decreased from the upper to lower layer in old PM, while those of *Methanobacterium* (e.g., bin.440) and *Caproiciproducens* increased (Fig. 3B). Similarly, in new PM, the relative abundance of *Lactobacillus* was lower in the lower than in the upper layer, despite it being the dominant genus in new PM (Fig. 3C). On the other hand, *Acetilactobacillus* had higher relative abundances in the lower layer of new PM compared to the upper layer (Fig. 3C).

L. acetotolerans was the dominant species in the *Lactobacillus* genus, especially in new PM (Fig. 4A). In contrast, species such as *Methanoculleus marisnigri*, *Methanoculleus bourgensis*, *Methanosarcina barkeri*, *Petrimonas mucosa*, and *Paenibacillus larvae* were more abundant in old PM. Linear discriminant analysis (LDA) showed that the relative abundances of the classes Clostridia, Methanomicrobia, and Methanobacteria, including *Clostridium* and various Archaea, were higher in old PM, suggesting these taxa increase with PM age (Fig. 4B), likely related to human brewing activities.

Phylogenetic analysis of prokaryotes in PM

For phylogenetic analysis, we identified 178 medium- and high-quality, non-redundant metagenome-assembled genomes (MAGs) (completeness > 75%, contamination < 10%) in PM, including 32 archaeal MAGs and 146 bacterial MAGs (Supplementary Table S1). Among the archaeal genera, methanogens *Methanoculleus* (12 MAGs), *Methanobacterium* (six MAGs), *Methanosarcina* (four MAGs), *Methanothrix*, and *VadinCA11* were predominant, while *Caproicibacter*, *Caproicibacterium*, *Clostridium*, and *Lactobacillus* were the most predominant bacterial genera. Identifying several methanogen MAGs highlighted their important roles in maintaining the structural integrity of PM [7]. At the phylum level, the most abundant bacterial phylum was Firmicutes, followed by Bacteroidota (Fig. 5). In addition, 6 archaeal and 47 bacterial MAGs were unclassified at the species level in the Genome Taxonomy Database (GTDB) (release202) (Fig. 5), suggesting that many microbial dark matter (MDM) strains are present in PM [8, 35]

Microbial functional annotation and pathways present in PM

Functional annotation of the 178 medium- and high-quality MAGs showed that the old PM community could perform various metabolic and biogeochemical transformations related to the carbon, nitrogen, and sulfur cycles (Figs. 6, 7). Regarding the carbon cycle, the MAGs carried genes linked to organic carbon oxidation, carbon fixation, acetate oxidation, hydrogen generation, fermentation, methanogenesis, methanotrophy, and hydrogen oxidation (Fig. 7). Notably, *Methanosarcina*, *Methanoculleus*, *Methanobacterium*, and *Clostridium* were found to utilize the Wood-Ljungdahl pathway, while Methanothrix employed both the Wood-Ljungdahl pathway and the reverse TCA cycle for carbon fixation. Several methanogenic MAGs, such as *Methanoculleus*, *Methanobacterium*, *Methanothrix*, and *Methanosarcina*, contained methane metabolism pathways (Supplementary Table S5). Thirty-two MAGs classified as Archaea contained genes encoding the MCR complex (mcrABC) and methyl-CoM reductase (MCR)-like enzymes, which are essential for methane production. This suggests that these archaea may have the potential for methane or short-chain alkane metabolism. Interestingly, the reverse tricarboxylic acid (TCA) cycle (aclAB) was exclusively

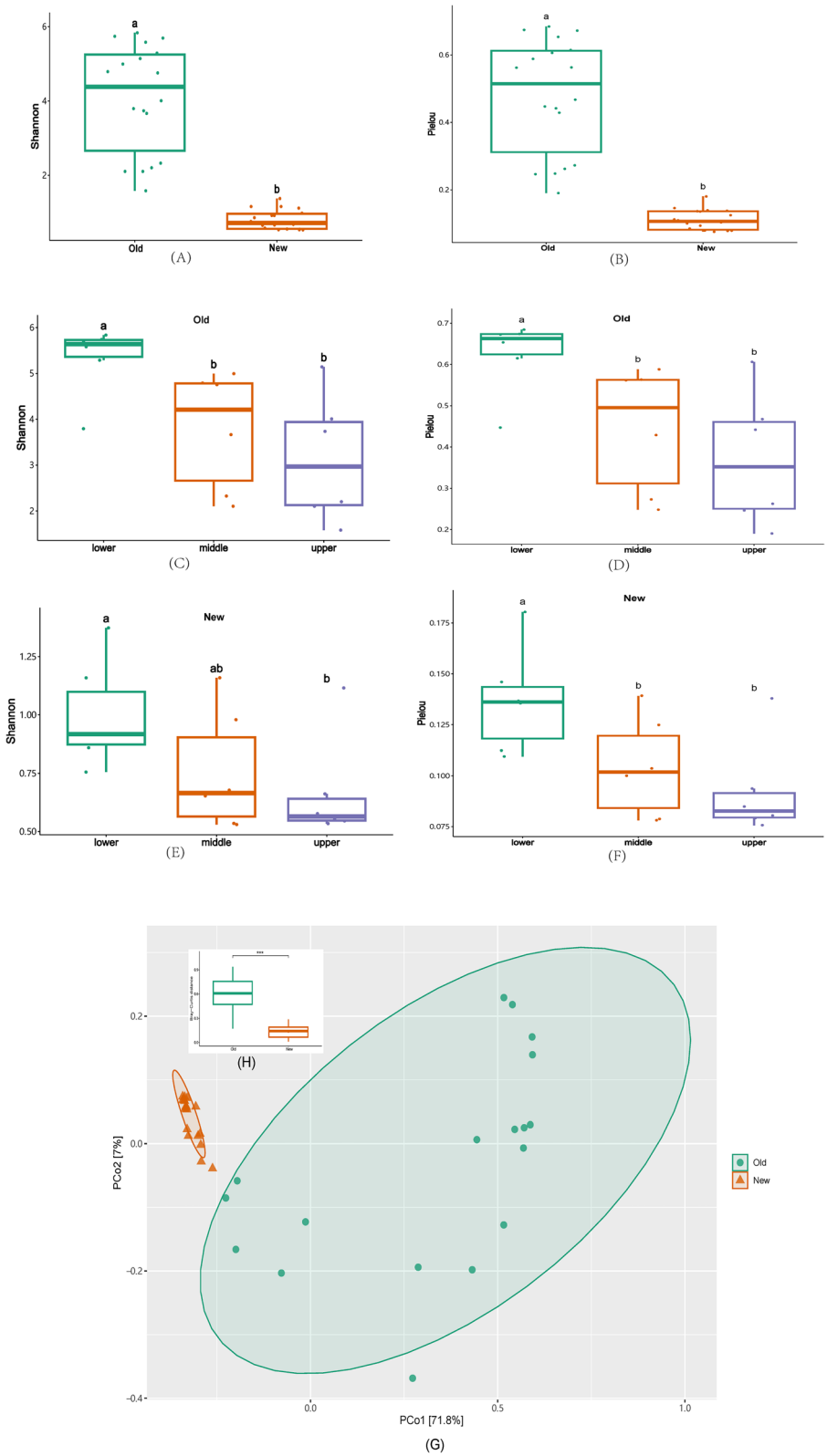


Fig. 2 Analysis of α- and β- diversity in pit mud. **A**, **B** Old pit mud (PM) vs new PM; **C**, **D** old PM at different layers; **E**, **F** new PM at different layers; **G**, **H** PCoA based on Bray–Curtis dissimilarity and Wilcoxon rank sum test analysis between old and new PM

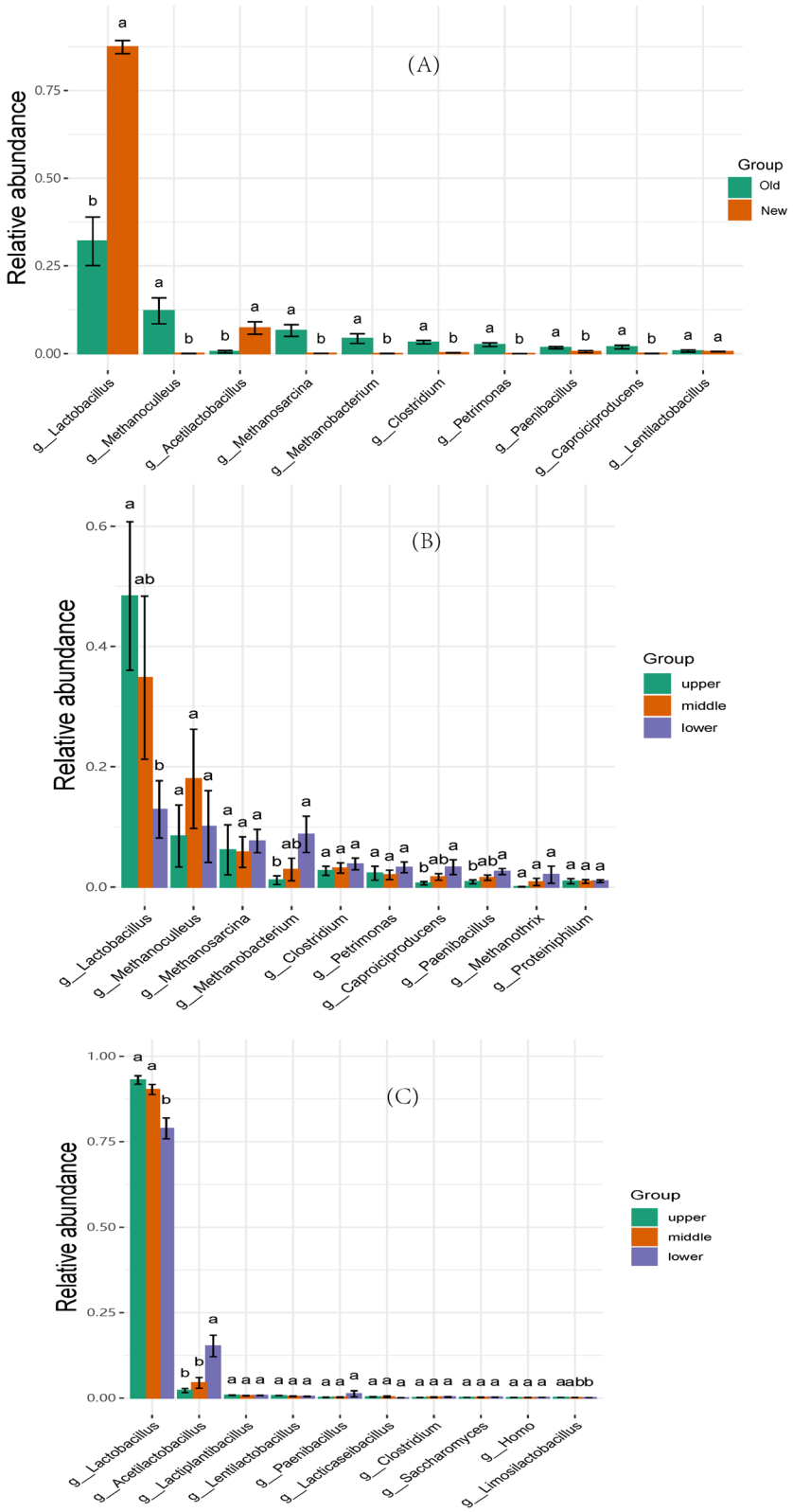
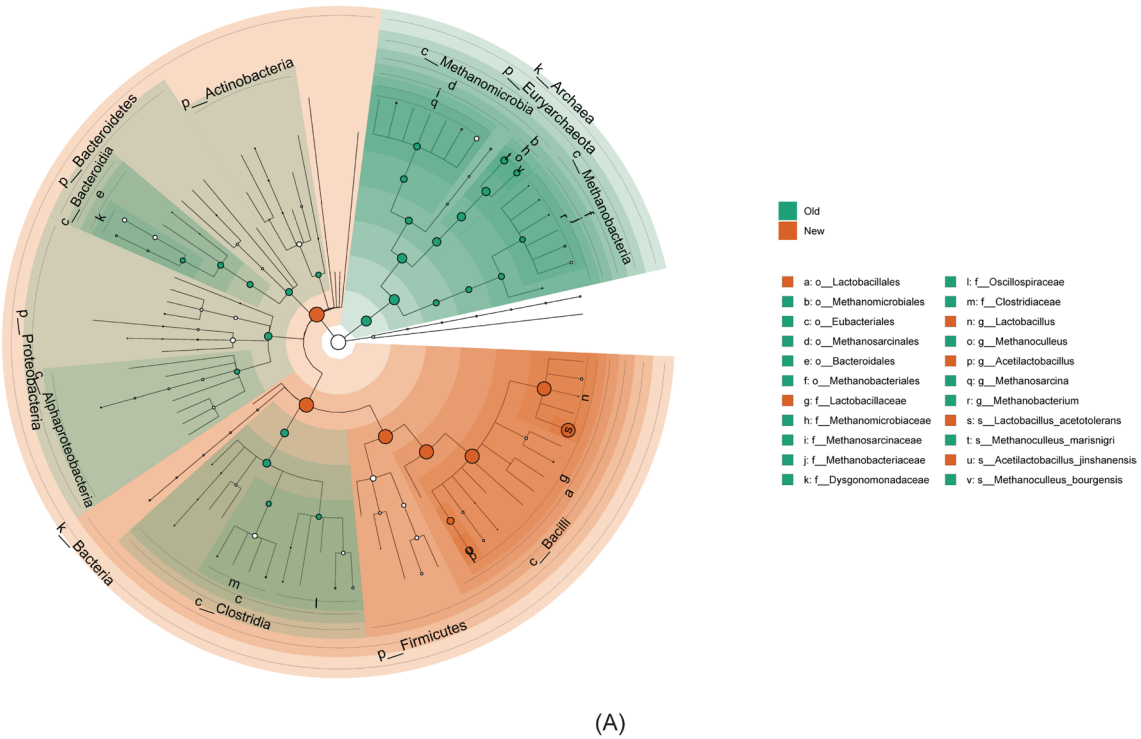
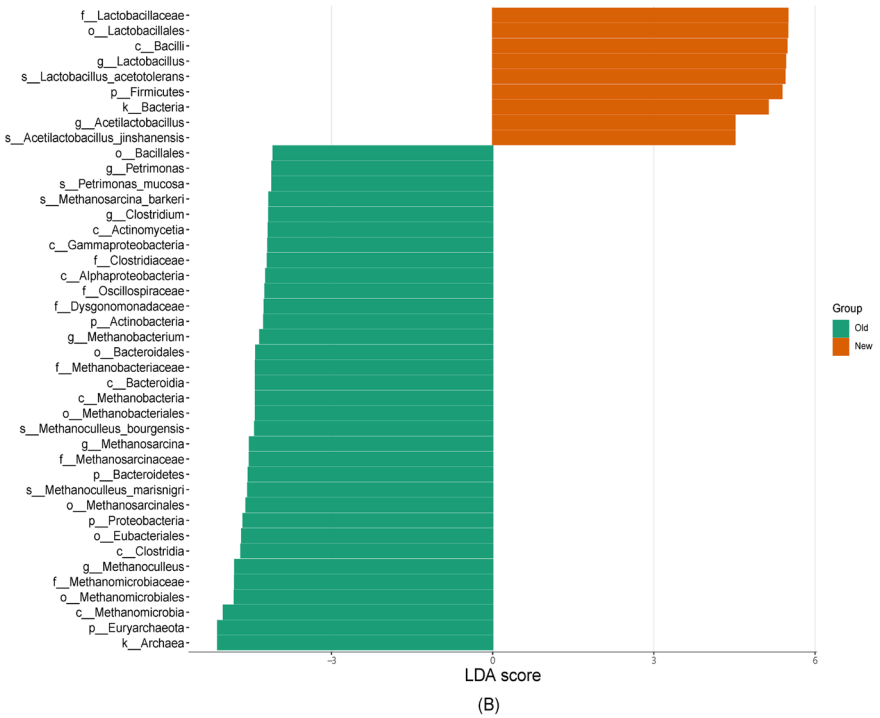


Fig. 3 Differential abundance analyses in pit mud. **A** Top 10 abundant genera in old *vs* new PM; **B** top 10 abundant genera across upper, middle, and lower layers in old PM; **C** top 10 abundant genera across layers in new PM. One-way ANOVA multiple comparison method in "anova" using the microeco package (https://chiliubio.github.io/microeco_tutorial/)

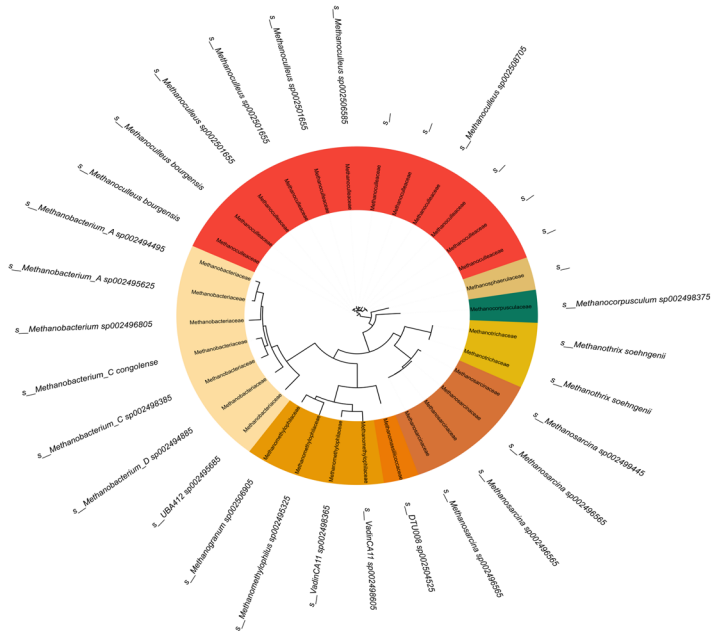


(A)

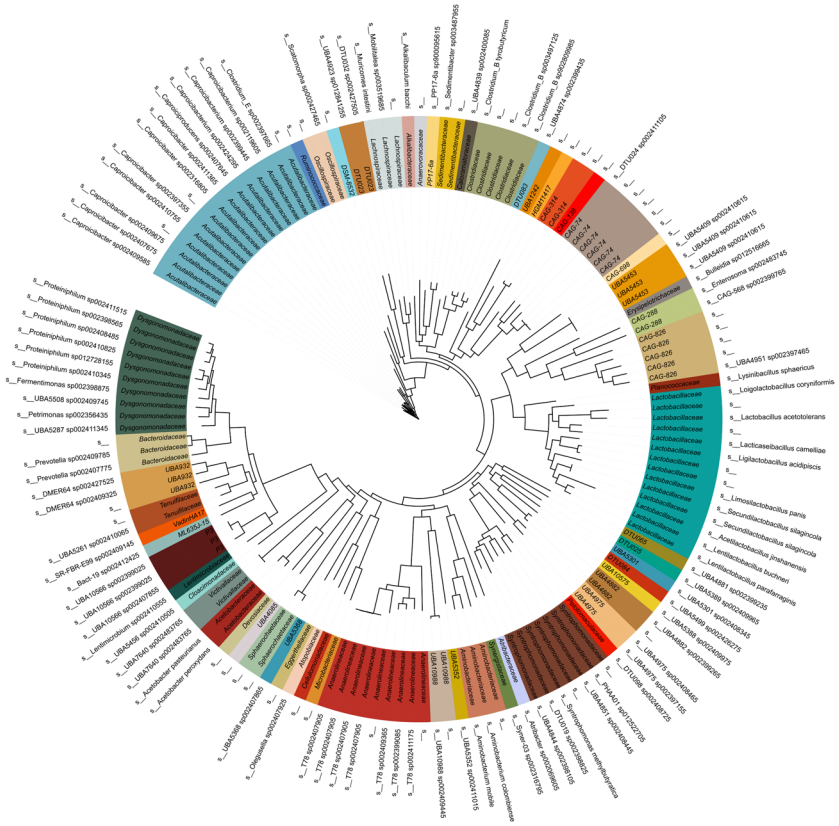


(B)

Fig. 4 Cladogram and linear discriminant analysis between old and new PM



(A)



(B)

Fig. 5 Phylogenetic analysis of Archaea (**A**) and Bacteria (**B**) in old PM based on GTDB-Tk (s_ indicates unclassified at the species level)

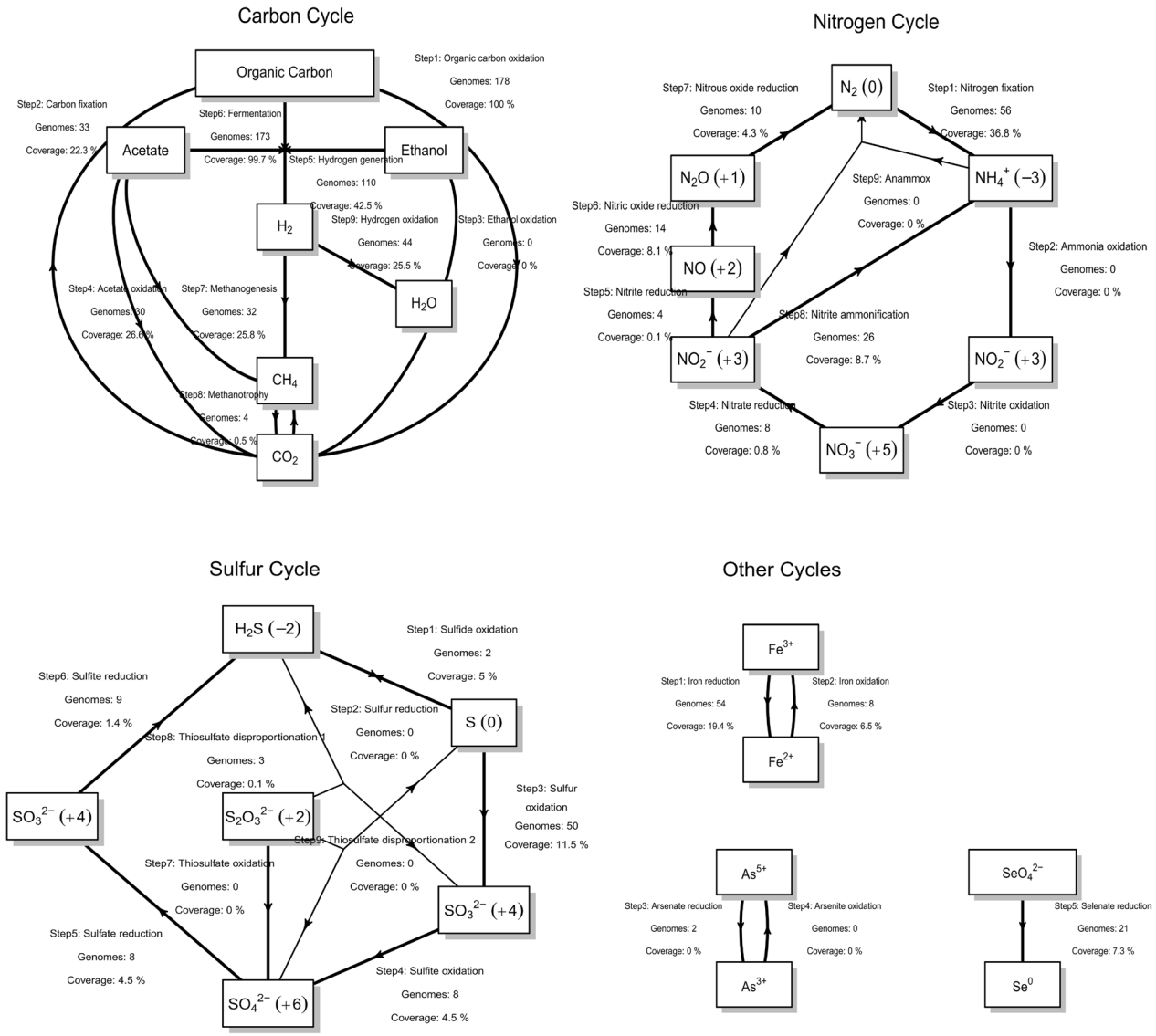


Fig. 6 Schematic diagrams of functional cycling processes present in old PM

detected in *Methanothrix* (Supplementary Table S6), further emphasizing the microbial community's activity in carbon metabolism and carbon fixation in old PM. Similarly, the MAGs carried various genes related to nitrogen and sulfur cycling. Notably, iron reduction emerged as a significant pathway in the microbial community of old PM.

Each transformation/step is represented by an arrow within a cycle. The number of genomes capable of performing the corresponding reaction is indicated, with coverage calculated from the abundance of metabolic genes in these genomes. The chemical states of nitrogen or sulfur atoms are labeled in brackets, representing different steps in the cycle.

Comparison of metabolic profiles between lower and upper microbial communities

Comparing metabolic weight scores revealed that the capacities for aromatics degradation, carbon fixation via the Wood-Ljungdahl pathway, methanogenesis, hydrogen oxidation, and N_2 fixation were higher in the lower than in the upper layer in the old PM. The capacity for complex carbon degradation and iron reduction was higher in the upper layer while that for iron oxidation was higher in the lower layer (Fig. 8), suggesting distinct roles in biogeochemical cycling for the microbial communities in different layers.

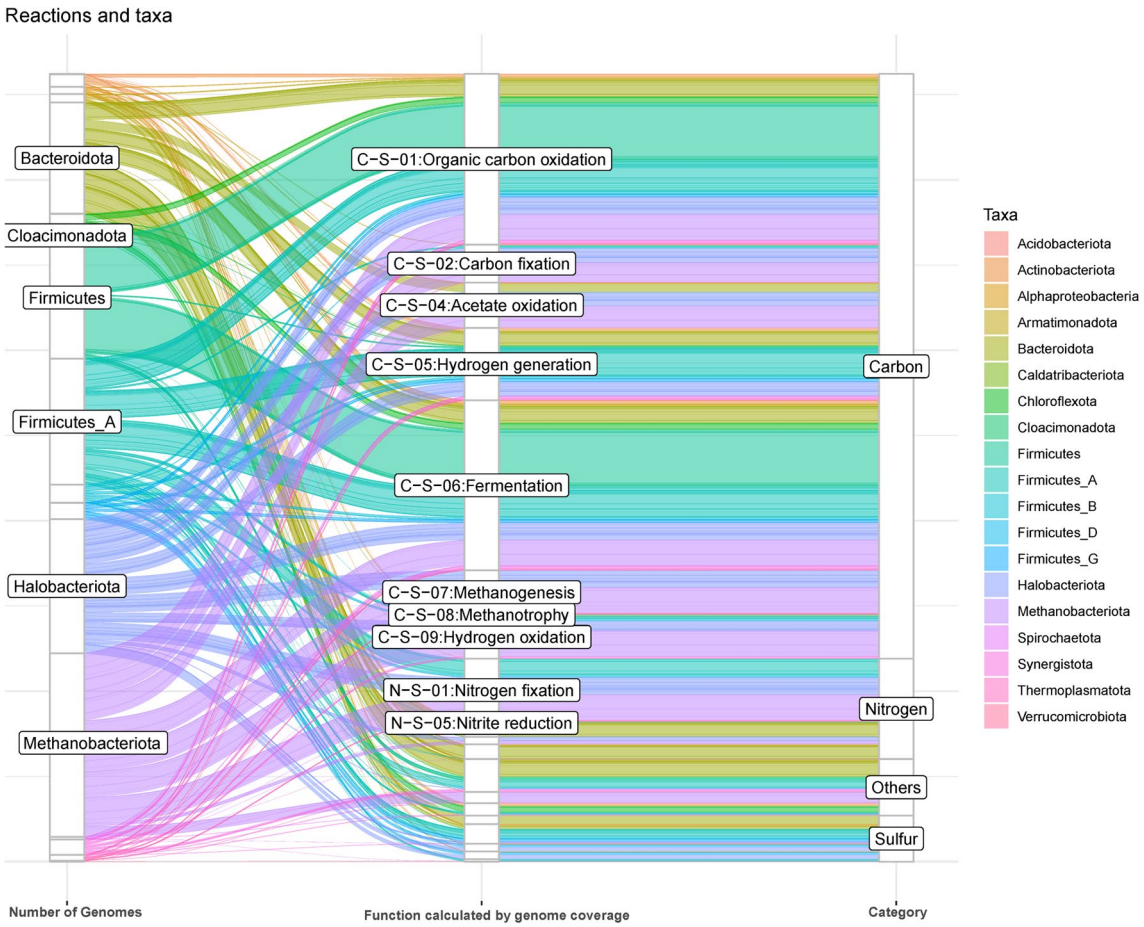


Fig. 7 Sankey diagram representing metabolic and biogeochemical processes, as well as entire elemental cycles. The taxonomic groups shown in the diagram include primarily Firmicutes (c__Bacilli), Firmicutes_A (c__Clostridia), Halobacteriota (c__Methanomicrobia, c__Methanosarcinia), and Methanobacteriota (c__Methanobacteria). The first column represents taxonomic groups based on the number of genomes; the second column shows the contribution of each microbial group to metabolic functions, calculated based on genome coverage; the third column illustrates each group's contribution to functional categories and biogeochemical cycles

Extensive horizontal gene transfers in PM

Assessing horizontal gene transfer (HGT) across 89 MAGs (Supplementary Table S2) revealed 1,067 HGT candidates, out of which 110 were removed as false positives. Of the remaining 957 HGT candidates, 251 were further validated and the direction of gene flow was determined using the phylogenetic (PG) approach. These HGTs mainly occurred within classes such as Bacteroidia, Clostridia, Methanomicrobia, and Bacilli (Fig. 9A). HGT events were also observed between genera, such as *Caproicibacter* and *Clostridium*, *Caproicibacter* and *Caproicibacterium*, and *Proteiniphilum* and *Petrimonas* (Fig. 9B). The number of genes transferred from Clostridia to other classes was higher than that from other classes to Clostridia. Functional annotation of the HGT genes showed their involvement in a range of biological processes, including genetic information processing, carbohydrate metabolism, and signaling and cellular processes (Fig. 9C). Furthermore, KEGG mapper reconstruction revealed that the HGT genes included carbon fixation, methane metabolism, nitrogen metabolism, and sulfur metabolism in prokaryotes related genes (data not shown). The HGT genes were assigned to bacterial phyla Clostridia, Bacteroidota, Bacilli, and archaeal phyla Euryarchaeota and Candidatus Thermoplasmatota (Fig. 9D).

Relationship between taxa and carboxyl metabolites

We identified 29 organic compounds, including amino acids, fatty acids, and phenylpropanoic acids, in PM (Supplementary Table S3). Assessing the relationships between taxa and carboxyl metabolites showed that six genera correlated with a range of carboxyl compounds (Fig. 10). For instance, *Lactobacillus* correlated with 2-phenylbutric acid, hendecanoic acid, and tridecylic acid ($P < 0.05$). Regarding medium-chain/

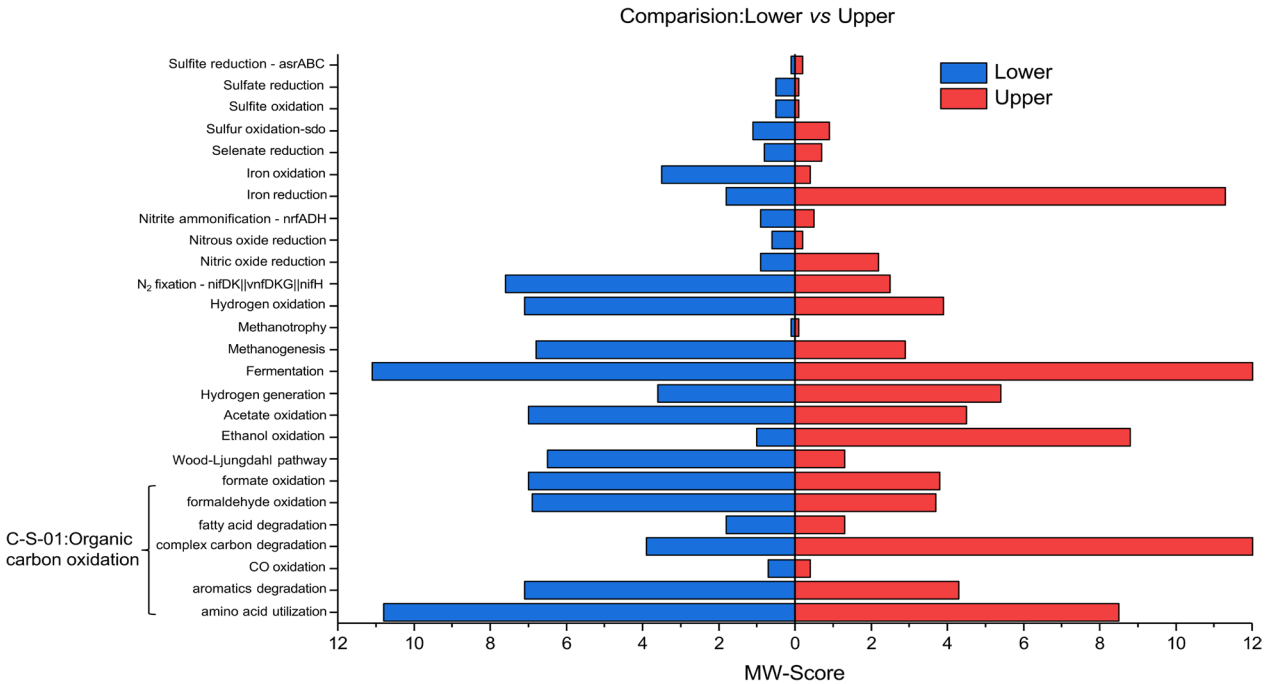


Fig. 8 Comparison of metabolic weight scores (MW-score) between the lower and upper layer of old PM

long-chain fatty acids, *Lactobacillus*, *Methanoculleus*, and *Petrimonas* correlated with tetradecanedioic acid; *Methanoculleus*, *Methanosarcina*, *Petrimonas*, and *Caproiciproducens* correlated with octadecanedioic acid; *Lactobacillus*, *Petrimonas*, and *Caproiciproducens* correlated with pentadecylic acid; *Methanoculleus*, *Methanosarcina*, *Petrimonas*, and *Caproiciproducens* correlated with 2-hydroxydecanoic acid; and *Methanoculleus*, *Methanosarcina*, *Petrimonas*, and *Paenibacillus* correlated with hydroxyoctanoic acid ($P < 0.05$).

Discussion

Using continuously utilized old pit mud (PM) or lower-layer PM results in higher-quality Chinese strong-flavor Baijiu (CSFB) compared to using new PM or upper-layer PM [10, 11, 36]. The prokaryotic communities in PM change with the age of the cellar, which impacts the quality of CSFB and contributes to a unique ecosystem [37]. However, the ecological and metabolic functions of these communities remain poorly understood.

We investigated the spatiotemporal changes in prokaryotic communities in PM. The results indicated that the evolution of these communities includes three distinct phases: an initial phase dominated by *Lactobacillus*, a transitional phase, and a state of equilibrium. The *Lactobacillus*-dominated phase was characterized by abundant *Lactobacillus*, especially *L. acetotolerans*, leading to increased acidity and decreased alpha diversity and species richness in anaerobic conditions. The lactic acid produced by *Lactobacillus* during fermentation, along with pH, significantly influences the structure of the prokaryotic community in PM [38]. The transitional phase was marked by increased diversity, particularly among archaea, accompanied by a decline in *Lactobacillus* abundance. Methanogens, *Clostridium*, *Petrimonas*, *Paenibacillus*, and *Caproiciproducens* were more prevalent in old PM or the lower layer, compared to new PM or the upper layer. Notably, many of these microorganisms were associated with various carboxyl compounds, including medium-chain/long-chain fatty acids. The equilibrium phase was characterized by intricate interactions among different taxa, especially between archaea and bacteria, which is crucial for sustaining the degradation process and preventing hydrogen buildup or excessive acidification in PM [39]. Maintaining microbial diversity and balance between archaea and bacteria are essential for adapting to environmental changes and facilitating diverse functional capabilities. During CSFB fermentation, polymers in the starch-rich raw materials are broken down and metabolized by microorganisms in PM into various small flavor compounds, such as alcohols, acids, and esters. Therefore, preserving microbial diversity and balance in PM is critical for ensuring efficient responses to environmental variations and sustaining diverse functional abilities, which serve as quality benchmarks for PM.

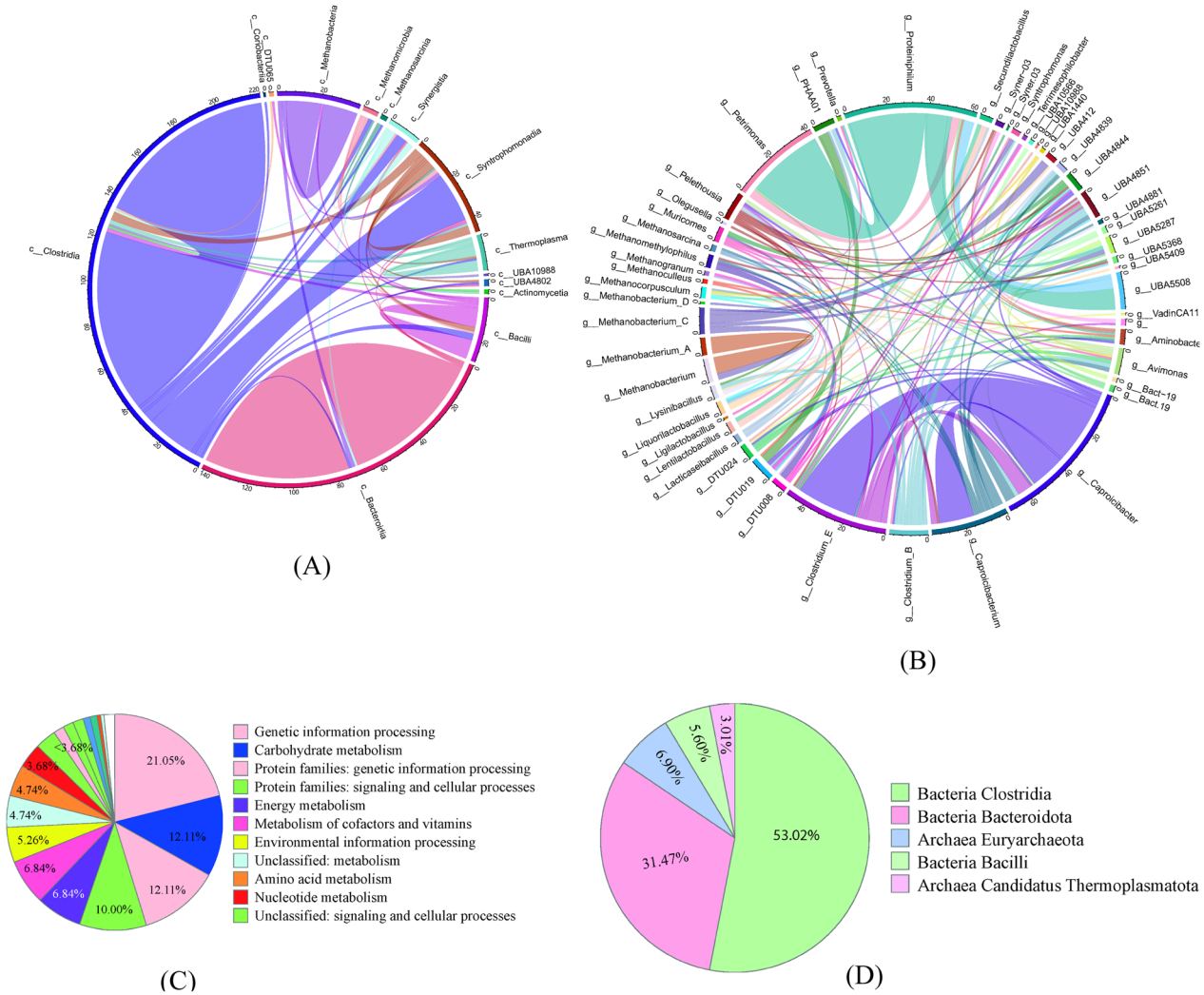


Fig. 9 Analysis of horizontal gene transfer (HGT) in PM (**A** HGTs at the class level; **B** HGTs at the genus level; **C** Functional categorization of HGT genes based on KEGG; **D** Taxonomic analysis of HGT genes)

Lactobacillus dominates not only in new PM but also in Jiupei samples, where it produces large amounts of lactic acid. High levels of lactic acid can lead to the formation of ethyl lactate, which negatively affects CSFB quality [8]. Moreover, lactic acid can cause significant acidification, thereby degrading PM quality [11]. The shift from *Lactobacillus* dominated state to a more diverse state can be influenced by selection and drift, driven by factors such as the materials utilized in Jiupei (fermented grains), the fermentation processes (long-term tolerance to ethanol and acids), and environmental conditions like temperature in anaerobic settings [40]. The activities of microbial communities in PM are sensitive to environmental factors like ethanol concentration, pH, nutrients, and the brewing process. Exposure to high concentrations of ethanol and acid stress can inhibit mRNA synthesis, cause irregular transcription termination, and alter membrane-lipid composition, all of which are detrimental to cell viability [41]. However, periodic batch-to-batch fermentation allows functional microbial taxa in PM to adapt and display metabolic versatility in anaerobic conditions, supporting processes such as hydrolysis, acidogenesis, acetogenesis, and methanogenesis [42]. Consequently, PM can be classified as "single-stage" anaerobic digestion system that produces organic acids like acetate, H_2, and CO_2 from sugar fermentation and fatty acid degradation, and supports the growth of methanogens through hydrolysis and acidogenesis [43].

When new Jiupei is introduced into PM, significant shifts occur in fermentation conditions, including changes in ethanol concentration, pH, dissolved oxygen, and nutrient levels. Methanogens, which collaborate with bacteria in decomposing organic matter and converting it into organic acids and amino acids

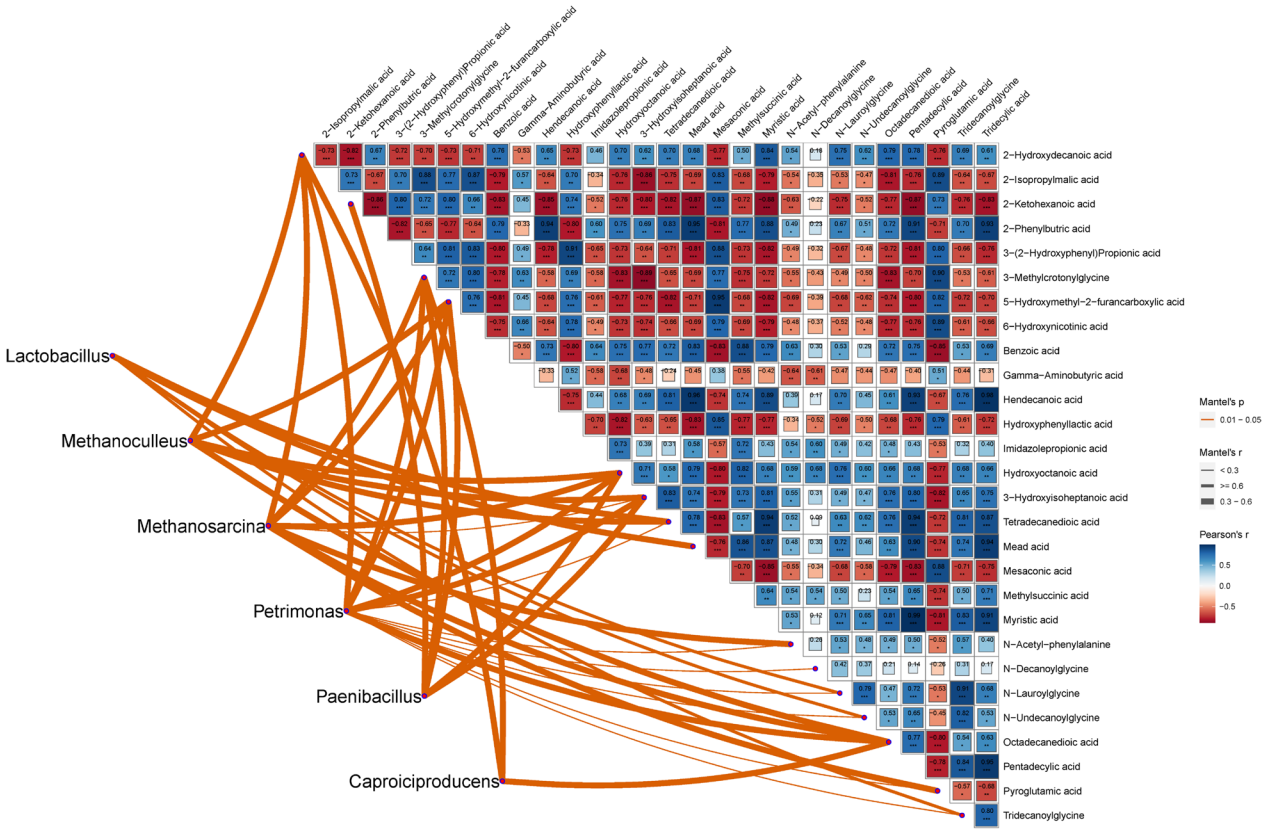


Fig. 10 Relationships between taxa and carboxyl metabolites

through various metabolic pathways, play a crucial role in this process [19, 37]. The composition and activities of methanogens in PM are influenced by pH; for instance, *Methanocorpusculum* thrives at pH 7.0, while *Methanosarcina* grows well at both pH 6.0 and 8.0 [44]. Additionally, pH influences the types of fermentation that occur, with higher pH favoring butyric acid-type and mixed acid-type fermentations, while lower pH promotes ethanol fermentation [45]. Temperature also plays a significant role in methanogenesis. Moderate temperatures around 35 °C typically lead to a combination of hydrogenotrophic and acetoclastic methanogenesis. At lower temperatures around 5–10 °C, methanogens primarily metabolize acetate to methane. Conversely, at temperatures exceeding 40 °C, hydrogenotrophic methanogenesis dominates, leading to marked changes in the structure and function of the methanogen community [46]. This dynamic interplay of factors is crucial for carbon cycling in PM. Notably, hydrogen produced during fermentation must be released, as a high partial pressure of hydrogen ($\geq 10^{-4}$ atm) inhibits acetogenesis, which produces volatile fatty acids (VFAs) such as propionic acid, butyric acid, and acetic acid [43]. The adaptation of niche-selected microbial populations in PM to anaerobic conditions is essential, and measures should be taken to enhance methanogenic activity by adjusting pH, temperature, and substrate to improve CSFL aroma production.

Our results also provide a foundation for understanding the roles of methanogens and *Clostridium* in carbon cycling under anaerobic conditions in PM. Thirty-three metagenome-assembled genomes (MAGs), mostly assigned to methanogens and *Clostridium*, demonstrated the capacity for carbon fixation through the Wood-Ljungdahl pathway (WLP) and/or the reverse TCA cycle. Furthermore, substrates such as acetate, lactate, butyric acid, and/or ethanol can be converted into caproate and hexanoic acid through carboxylic acid chain elongation processes between bacteria and methanogens [10]. However, methanogenic activity can be inhibited by acid accumulation [47].

CH_4 production is carried out by mainly three methanogenic pathways, i.e., hydrogenotrophic, acetoclastic, and methylotrophic pathways [48]. While many methanogenic species remain uncultivable, genomic-centered metagenomic approaches offer insights into their phylogeny and functions. In this study, both hydrogenotrophic (e.g., *Methanoculleus* and *Methanobacterium*) and acetoclastic (e.g., *Methanosarcina*) methanogens dominated

the methanogens in old PM. In particular, *Methanobacterium* showed higher abundance in the lower layer of old PM, indicating that CO_2/H_2 and acetate were the main substrates for methanogenesis in old PM [14, 37]. Detecting methanogens such as *Methanocorpusculum*, *Methanothrix*, *Methanogranum*, and *Methanomethylophilus* suggested the presence of methylotrophic pathway in old PM [49]. Many archaeal and bacterial MAGs remained unclassified at the species level, highlighting the complex ecology of PM and expanding our understanding of the phylogenetic diversity of prokaryotes in PM. However, since the activity of the PM microbiome cannot be concluded based on DNA sequencing, future studies should focus on metatranscriptomics to uncover the mechanisms by which these microorganisms respond to environmental fluctuations.

Interestingly, we observed extensive horizontal gene transfer (HGT) in PM. HGT is an important force in microbial evolution and adaptation to environmental and genetic perturbations [50]. Most of the HGTs were involved in pathways related to genetic information processing, carbon fixation, methane metabolism, nitrogen metabolism, and sulfur metabolism in prokaryotes, suggesting that HGT plays a critical role in substrate utilization and competition, enhancing microbial adaptation to long-term ethanol and acid fermentation fluctuations in PM. Contrary to previous studies [11], the number of genes transferred from Clostridia to other classes was higher than from other classes to Clostridia. Since HGT can maintain stable biodiversity even when competing species co-exist [51], the results indicated that HGT may contribute to the evolutionary complexity in PM.

The genera *Lactobacillus*, *Methanoculleus*, *Methanosarcina*, *Petrimonas*, *Paenibacillus*, and *Caproiciproducens* contributed to the carboxyl compounds that, along with ethanol, are crucial for the flavor of Chinese Baijiu [52]. *Lactobacillus*, *Caproiciproducens*, and *Petrimonas* produce organic acids and short-chain fatty acids under anaerobic fermentation [53]. Additionally, methanogens, in collaboration with fermentative bacteria, enhance the production of organic acids through syntrophic interactions in PM, contributing to carbon cycling and generating protons and hydroxyl ions [54]. Therefore, the results show that methanogens and *Clostridium* play critical roles in maintaining the structure and function of PM and in producing carboxyl compounds that affect the quality of CSFB, and explain why high-quality Baijiu is produced exclusively in old cellars.

Conclusions

In conclusion, our study provides a comprehensive understanding of the microbial community structure, metabolic functions, phylogenetics, horizontal gene transfer, and their relationship with carboxyl compounds in spatiotemporal PM. *Lactobacillus* dominated in new PM, while methanogens and clostridia were predominant in older or deeper PM layers. We demonstrated that the evolutionary divergence of spatiotemporal prokaryotic communities in PM can be categorized into three distinct stages. HGT played a critical role in metabolic function of spatiotemporal PM. Additionally, α-diversity, β-diversity, methanogens, and *Clostridium* can serve as valuable indicators to assess the quality of PM for producing high-quality Baijiu in the industry. These findings enhance our understanding of microbial ecology and metabolic processes in PM, with important implications for the production of Chinese Baijiu.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s40793-025-00668-8.

Supplementary material 1

Supplementary material 2

Supplementary material 3

Author contributions

The research was designed by ZL and SZ. Sampling was done by CZ, ZM, CS, YC, XF, YX, HQ and ZA. Laboratory experimentation was done by FZ, and LD. Sequence processing, data curation and data analyses were done by ZL, CZ, and SZ. The original draft was written by ZL and SZ, then reviewed and edited by all authors. ZS performed the HGT analysis, PP, edited the language, ZL and SZ supervised entire research. All authors approved the final manuscript.

Funding

This work was supported by the Youth Fund of Postdoctor (2022HX04), Yibin City Agricultural Science and Technology Innovation Project (2024NYHZ007).

Availability of data and materials

The datasets in this study are publicly available in the NCBI BioProject with accession No. PRJNA1076102.

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Author details

[1]Present Address: Institute of Agro-Products Processing Science and Technology, Sichuan Academy of Agricultural Sciences, Chengdu 610066, Sichuan, China. [2]Luzhou Laojiao Co., Ltd., Luzhou 646000, China. [3]Present Address: National Engineering Research Center of Solid-State Brewing, Luzhou 646000, China. [4]Department of Ocean Science, The Hong Kong University of Science and Technology, Clear Water Bay, Hong Kong, China. [5]College of Resources, Sichuan Agricultural University, Chengdu 611130, China. [6]Faculty of Agriculture and Forestry, University of Helsinki, 00014 Helsinki, Finland.

Received: 31 July 2024 Accepted: 7 January 2025
Published online: 20 January 2025

References

1. Hong J, Zhao D, Sun B. Research progress on the profile of trace components in Baijiu. Food Rev Int. 2021;39:1666–93.
2. Hong J, Tian W, Zhao D. Research progress of trace components in sesame-aroma type of Baijiu. Food Res Int. 2020;137:109695. https://doi.org/10.1016/j.foodres.2020.109695.
3. Liu MK, Liu CY, Tian XH, Feng J, Guo XJ, Liu Y, et al. Bioremediation of degraded pit mud by indigenous microbes for Baijiu production. Food Microbiol. 2022;108:104096. https://doi.org/10.1016/j.fm.2022.104096.
4. Xing-Lin H, De-Liang W, Wu-Jiu Z, Shi-Ru J. The production of the Chinese Baijiu from sorghum and other cereals. J Inst Brew. 2017;123:600–4.
5. Huang Z, Zeng Y, Sun Q, Zhang W, Wang S, Shen C, et al. Insights into the mechanism of flavor compound changes in strong flavor baijiu during storage by using the density functional theory and molecular dynamics simulation. Food Chem. 2022;373:131522. https://doi.org/10.1016/j.foodchem.2021.131522.
6. Hu X, Du H, Ren C, Xu Y. Illuminating anaerobic microbial community and cooccurrence. Appl Environ Microbiol. 2016;82:2506–15.
7. Lu M, Zhou W, Ji F, Wu J, Nie Y, Ren C, et al. Profiling prokaryotic community in pit mud of Chinese strong-aroma type liquor by using oligotrophic culturing. Int J Food Microbiol. 2021;337:108951. https://doi.org/10.1016/j.ijfoodmicro.2020.108951.
8. Fu J, Chen L, Yang S, Li Y, Jin L, He X, et al. Metagenome and analysis of metabolic potential of the microbial community in pit mud used for Chinese strong-flavor liquor production. Food Res Int. 2021;143:110294.
9. Qian W, Lu Z, Chai L, Zhang X, Li Q, Wang S. Cooperation within the microbial consortia of fermented grains and pit mud drives organic acid synthesis in strong-flavor Baijiu production. Food Res Int. 2021;147:110449.
10. Chai LJ, Qian W, Zhong XZ, Zhang XJ, Lu ZM, Zhang SY, et al. Mining the factors driving the evolution of the pit mud microbiome under the impact of long-term production of strong-flavor Baijiu. Appl Environ Microbiol. 2021;87:1–16.
11. Mei J, Chai L, Zhong X, Lu Z, Zhang X, Wang S, et al. Microbial biogeography of pit mud from an artificial brewing ecosystem on a large time scale: all roads lead to Rome. Msystems. 2023;8:e00564-23.
12. Gao J, Liu G, Li A, Liang C, Ren C, Xu Y. Domination of pit mud microbes in the formation of diverse flavour compounds during Chinese strong aroma-type Baijiu fermentation. Lwt. 2021;137:110442. https://doi.org/10.1016/j.lwt.2020.110442.
13. Deng B, Shen C, Shan X, Ao Z, Zhao J, Shen X, et al. PCR-DGGE analysis on microbial communities in pit mud of cellars used for different periods of time. J Inst Brew. 2012;118:120–6.
14. Hu X, Du H, Ren C, Xu Y. Illuminating anaerobic microbial community and cooccurrence patterns across a quality gradient in Chinese liquor fermentation pit muds. Appl Environ Microbiol. 2016;82:2506–15.
15. Vollmers J, Wiegand S, Lenk F, Kaster AK. How clear is our current view on microbial dark matter? (Re-)assessing public MAG & SAG datasets with MDMcleaner. Nucleic Acids Res. 2022;50:e76.
16. Oh J, Byrd AL, Deming C, Conlan S, Kong HH, Segre JA, et al. Biogeography and individuality shape function in the human skin metagenome. Nature. 2014;514:59–64. https://doi.org/10.1038/nature13786.
17. Rodríguez del Río Á, Giner-Lamia J, Cantalapiedra CP, Botas J, Deng Z, Hernández-Plaza A, et al. Functional and evolutionary significance of unknown genes from uncultivated taxa. Nature. 2023;626:377–84.
18. Baker BJ, De Anda V, Seitz KW, Dombrowski N, Santoro AE, Lloyd KG. Diversity, ecology and evolution of Archaea. Nat Microbiol. 2020;5:887–900. https://doi.org/10.1038/s41564-020-0715-z.
19. Zhou Z, Tran PQ, Breister AM, Liu Y, Kieft K, Cowley ES, et al. METABOLIC: high-throughput profiling of microbial genomes for functional traits, metabolism, biogeochemistry, and community-scale functional networks. Microbiome. 2022;10:1–22. https://doi.org/10.1186/s40168-021-01213-8.
20. Zhao S, Li L. Chemical derivatization in LC-MS-based metabolomics study. TrAC Trends Anal Chem. 2020;131:115988. https://doi.org/10.1016/j.trac.2020.115988.
21. Chen S, Zhou Y, Chen Y, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34:i884–90.
22. Lu J, Rincon N, Wood DE, Breitwieser FP, Pockrandt C, Langmead B, et al. Metagenome analysis using the Kraken software suite. Nat Protoc. 2022;17:2815–39.
23. Lu J, Breitwieser FP, Thielen P, Salzberg SL. Bracken: estimating species abundance in metagenomics data. PeerJ Comput Sci. 2017;3:e104.
24. Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, et al. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. J Comput Biol. 2012;19:455–77.
25. Uritskiy GV, DiRuggiero J, Taylor J. MetaWRAP—a flexible pipeline for genome-resolved metagenomic data analysis. Microbiome. 2018;6:1–13.
26. Zhang RY, Wang YR, Liu RL, Rhee SK, Zhao GP, Quan ZX. Metagenomic characterization of a novel non-ammonia-oxidizing Thaumarchaeota from hadal sediment. Microbiome. 2024;12:1–16. https://doi.org/10.1186/s40168-023-01728-2.
27. Song W, Wemheuer B, Zhang S, Steensen K, Thomas T. MetaCHIP: community-level horizontal gene transfer identification through the combination of best-match and phylogenetic approaches. Microbiome. 2019;7:1–14.
28. Huerta-Cepas J, Forslund K, Coelho LP, Szklarczyk D, Jensen LJ, Von Mering C, et al. Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. Mol Biol Evol. 2017;34:2115–22.
29. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol. 2016;428:726–31. https://doi.org/10.1016/j.jmb.2015.11.006.
30. Li Z, Dong L, Zhao C, Zhang F, Zhao S, Zhan J, et al. Development of a high-coverage quantitative metabolome analysis method using four-channel chemical isotope labeling LC–MS for analyzing high-salt fermented food. J Agric Food Chem. 2022;70:8827–37.
31. Segata N, Izard J, Waldron L, Gevers D, Miropolsky L, Garrett WS, et al. Metagenomic biomarker discovery and explanation. Genome Biol. 2011;12:R60.
32. Liu C, Cui Y, Li X, Yao M. Microeco: an R package for data mining in microbial community ecology. FEMS Microbiol Ecol. 2021;97:1–9.
33. Hua Z, Zhou X. Data and text mining itol. toolkit accelerates working with iTOL (Interactive Tree of Life) by an automated generation of annotation files. 2023;39:0–3.
34. Letunic I, Bork P. Interactive tree of life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. 2021;49:W293–6.
35. Parks DH, Chuvochina M, Rinke C, Mussig AJ, Chaumeil P, Hugenholtz P. GTDB: an ongoing census of bacterial and archaeal diversity through a phylogenetically consistent, rank normalized and complete genome-based taxonomy. Nuclear Acids Res. 2022;50:785–94.
36. Hu X, Tian R, Wang K, Cao Z, Yan P, Li F, et al. The prokaryotic community, physicochemical properties and flavors dynamics and their correlations in fermented grains for Chinese strong-flavor Baijiu production. Food Res Int. 2021;148:110626. https://doi.org/10.1016/j.foodres.2021.110626.
37. Zhang H, Meng Y, Wang Y, Zhou Q, Li A, Liu G, et al. Prokaryotic communities in multidimensional bottom-pit-mud from old and young pits used for the production of Chinese Strong-Flavor Baijiu. Food Chem. 2020;312:126084. https://doi.org/10.1016/j.foodchem.2019.126084.
38. Guo MY, Hou CJ, Bian MH, Shen CH, Zhang SY, Huo DQ, et al. Characterization of microbial community profiles associated with quality of Chinese strong-aromatic liquor through metagenomics. J Appl Microbiol. 2019;127:750–62.
39. Lin Q, De Vrieze J, Li J, Li X. Temperature affects microbial abundance, activity and interactions in anaerobic digestion. Bioresour Technol. 2016;209:228–36.
40. Van Rossum T, Ferretti P, Maistrenko OM, Bork P. Diversity within species: interpreting strains in microbiomes. Nat Rev Microbiol. 2020. https://doi.org/10.1038/s41579-020-0368-1.
41. Haft RJF, Keating DH, Schwaegler T, Schwalbach MS, Vinokur J. Correcting direct effects of ethanol on translation and transcription machinery confers ethanol tolerance in bacteria. PNAS. 2014;111:E2576–85.
42. De Bernardini N, Basile A, Zampieri G, Kovalovszki A, De Diego Diaz B, Offer E, et al. Integrating metagenomic binning with flux balance analysis

to unravel syntrophies in anaerobic CO_2 methanation. Microbiome. 2022;10:1–18. https://doi.org/10.1186/s40168-022-01311-1.
43. Van DP. A review of anaerobic digestion systems for biodegradable waste: configurations, operating parameters, and current trends. Environ Eng Res. 2020;25:1–17.
44. Zhang L, Loh KC, Lim JW, Zhang J. Bioinformatics analysis of metagenomics data of biogas-producing microbial communities in anaerobic digesters: a review. Renew Sustain Energy Rev. 2019;100:110–26. https://doi.org/10.1016/j.rser.2018.10.021.
45. Li Z, Gu J, Ding J, Ren N, Xing D. Molecular mechanism of ethanol-H_2 co-production fermentation in anaerobic acidogenesis: challenges and perspectives. Biotechnol Adv. 2021;46:107679. https://doi.org/10.1016/j.biotechadv.2020.107679.
46. Soil SR, Chen Y, Wu N, Liu C, Mi T, Li J, et al. Community characteristics of methanogens and methanogenic pathways in science of the total environment methanogenesis pathways of methanogens and their responses to substrates and temperature in sediments from the South Yellow Sea. Sci Total Environ. 2022;815:152645.
47. Tao Y, Wang X, Li X, Wei N, Jin H, Xu Z, et al. The functional potential and active populations of the pit mud microbiome for the production of Chinese strong-flavour liquor. Microb Biotechnol. 2017;10:1603–15.
48. Yang Y, Chen J, Tong T, Xie S, Liu Y. Influences of eutrophication on methanogenesis pathways and methanogenic microbial community structures in freshwater lakes. Environ Pollut. 2020;260:114106.
49. Söllinger A. Methylotrophic methanogens everywhere—physiology and ecology of novel players in global methane cycling. Biochem Soc Trans. 2019;47:1895–907.
50. Wang R, Wu J, Jiang N, Lin H, An F, Wu C, et al. Recent developments in horizontal gene transfer with the adaptive innovation of fermented foods. Crit Rev Food Sci Nutr. 2023;63:569–84. https://doi.org/10.1080/10408398.2022.2081127.
51. Zhu S, Hong J, Wang T. Horizontal gene transfer is predicted to overcome the diversity limit of competing microbial species. Nat Commun. 2024;15:1–9.
52. Xie X, Lu X, Zhang X, Zheng F, Yu D, Li C, et al. In-depth profiling of carboxyl compounds in Chinese Baijiu based on chemical derivatization and ultrahigh-performance liquid chromatography coupled to high-resolution mass spectrometry. Food Chem X. 2022;15:100440.
53. Tao Y, Li J, Rui J, Xu Z, Zhou Y, Hu X, et al. Prokaryotic communities in pit mud from different-aged cellars used for the production of Chinese strong-flavored liquor. Appl Environ Microbiol. 2014;80:2254–60.
54. Liu MK, Tang YM, Guo XJ, Zhao K, Tian XH, Liu Y, et al. Deep sequencing reveals high bacterial diversity and phylogenetic novelty in pit mud from Luzhou Laojiao cellars for Chinese strong-flavor Baijiu. Food Res Int. 2017;102:68–76. https://doi.org/10.1016/j.foodres.2017.09.075.

Publisher's Note