

Linking geographic flavor signatures to microbial origin in high-temperature Daqu: An integrated metaproteomics and metabolomics approach

Dandan Song^a, Xian Zhong^a, Guihu Zhang^a, Juan Chen^b, Yansong Xue^c, Liang Yang^{a,*}

^a School of Brewing Engineering, Moutai Institute, Renhuai 564501, China

^b School of Food Engineering, Moutai Institute, Renhuai 564501, China

^c Key Laboratory of Food Bioengineering (China National Light Industry), College of Food Science and Nutritional Engineering, China Agricultural University, Beijing, China

ARTICLE INFO

Keywords:

Enzymatic mechanism
Flavor formation
Functional decoupling
High-temperature Daqu
Metaproteomics
Microbial terroir

ABSTRACT

Elucidating the molecular architecture of microbial terroir is vital for precision fermentation, yet functional decoupling between taxonomic abundance and in situ expression remains a fundamental challenge. To resolve this “abundance-activity paradox,” we integrated metaproteomics, metabolomics, and metagenomics across the Chishui River gradient. We identified distinct chemosensory fingerprints: upstream thermotolerant consortia (*Bacillus* and *Oceanibacillus*) specialize in 2,3,5,6-tetramethylpyrazine biosynthesis mediated by bacterial acetolactate decarboxylase, while downstream microbiota (*Weissella* and *Debaryomyces*) prioritize alcohol and ester formation. Crucially, metaproteomic profiling unmasked the “rare biosphere” as a primary driver of core metabolic fluxes. While *Bacillus* was genomically dominant, keystone functional taxa—specifically low-abundance fungi like *Hyphopichia* and *Paecilomyces*—were the actual executors of rate-limiting starch hydrolysis. Furthermore, functional resilience was uniquely maintained through robust fungal co-occurrence networks despite geographic constraints. This study challenges abundance-centric paradigms, providing an activity-based framework for the rational design of synthetic microbial consortia to standardize flavor while preserving regional identity.

1. Introduction

Within food engineering systems such as fermented foods, microbial communities function as “biological catalysts” that drive critical biochemical transformations governing product quality and safety. Baijiu production, for instance, fundamentally depends on multi-strain solid-state fermentation, wherein the quality and distinctive sensory attributes predominantly rest on the diversity and stability of resident microbiota. As the primary fermentation starter for Jiang-flavor baijiu, high-temperature Daqu serves as a microbial repository governing fermentation kinetics and metabolic outputs (Bresciani et al., 2025; Niu et al., 2025). When uniformly blended with sorghum at a 1:1 ratio, Daqu serves as the fundamental reservoir of microbial consortia, enzyme systems, and critical metabolic precursors for Moutai-flavor Baijiu production. This intricate microbiome executes starch saccharification, proteolysis, and thermophilic fermentation (60–65 °C), yielding

signature flavor compounds dominated by pyrazines and esters (Hou et al., 2024; Luo, Liao, Wu, et al., 2025; Zhu, Chen, et al., 2023). Divergent regional microbiomes uniquely shape local Baijiu styles, positioning high-temperature Daqu as an optimal model for deciphering Baijiu's geographical signatures. Firstly, within the Chishui River Basin—the epicenter of Moutai-flavor Baijiu production—spatial heterogeneity in abiotic factors drives selection-for-function in microbial taxa, establishing the liquor's microbial terroir. Secondly, assembly mechanisms governing microbial community architecture govern metabolic resilience, with endemic strains contributing to region-specific flavor profiles. Thirdly, delineating conserved “core microbiomes” across production systems delivers dual benefits: enabling standardized Daqu production while safeguarding inherent sensory attributes of traditional processing.

Existing studies reveal distinct variations in microbial community structures of high-temperature Daqu across geographical regions. In

* Corresponding author.

E-mail address: moutaiyl0725@163.com (L. Yang).

<https://doi.org/10.1016/j.fochx.2026.103952>

Received 10 January 2026; Received in revised form 27 April 2026; Accepted 5 May 2026

Available online 6 May 2026

2590-1575/© 2026 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

Hubei, *Kroppenstedtia* dominates among bacteria while *Aspergillus* prevails among fungi (Hou et al., 2024). Guizhou samples additionally host *Virgibacillus*, *Oceanibacillus* as core microbiota (Yang, Fan, & Xu, 2024a). Notably, Sichuan-derived *Daqu* exhibits significantly higher abundance of *Saccharopolyspora* (Deng et al., 2020). These findings demonstrate pronounced biogeographical isolation governing high-temperature *Daqu* fermentation, likely driving region-specific divergence in quorum compound synthesis that ultimately shapes baijiu sensory characteristics. While amplicon sequencing (16S/ITS) has characterized dominant taxa including *Bacillus*, *Thermoactinomyces*, *Aspergillus*, and *Thermoascus* (Niu et al., 2025; Y. Wang et al., 2021; Yang, Fan, & Xu, 2024a; Zhang, Tan, et al., 2022), functional predictions based primarily on taxonomic inference overlook gene expression levels, limiting the study of key functional microbes. The abundance of uncultivable or fastidious microorganisms in *Daqu* further impedes mechanistic validation (Zhang, Xu, et al., 2022). While metagenomics offers advantages over amplicon sequencing for functional gene annotation and resolution of core fermentation pathways (Hou et al., 2024; Leech et al., 2020; N. Li et al., 2023; J. Xu, 2024), yet it reveals only the microbiome's genetic potential (i.e., putative function). It is often limited by the assumption that numerically dominant taxa drive community function and flavor formation. In stressful environments like high-temperature *Daqu* fermentation, DNA-based taxonomic abundance can largely decouple from actual metabolic activity, and may sporulate in response to environmental fluctuations, contributing abundant genomic DNA but exhibiting negligible enzymatic activity (Xiaolong Hu et al., 2020). Conversely, low-abundance microorganisms can be highly active and govern rate-limiting metabolic fluxes. Relying solely on metagenomic sequencing only reveals the microbiome's genetic potential (i.e., putative function), obscuring the actual functional executors.

To address this limitation, integrating metaproteomics allows for the direct quantification of expressed enzyme systems (i.e., realized function), overcoming the predictive constraints of metagenomic analyses (Van Den Bossche et al., 2025). This approach shifts the research focus from taxonomic abundance to metabolic activity, redefining keystone species by their real-time catalytic contributions rather than their population size. Consequently, associating regional phenotypic variations in high-temperature *Daqu* with proteomic signatures allows for precise identification of core functional microbiota. This research addresses a significant knowledge gap in brewing microbiology by elucidating the geographical provenance of *Daqu*'s characteristic flavor compounds and their linkage to functional microbial communities. Through an integrated multi-omics strategy combining metaproteomics, metabolomics, and metagenomics, we investigated the phenotypic differentiation of high-temperature *Daqu* across various production areas within the Chishui River Basin, and establishing a ternary framework that links microbial community, function, and flavor. Crucially, we hypothesize that the true drivers of geographic flavor specificity may not be the most abundant taxa, but rather low-abundance microorganisms with high metabolic activity.

2. Materials and methods

2.1. *Daqu* sampling

High-temperature *Daqu* samples were collected from six Maotai-flavor Baijiu factories in the upper, middle, and lower reaches of the Chishui River. Collected samples were snap-frozen in liquid nitrogen, milled into powder, transferred into sterile bags, and stored at -80°C .

2.2. Determination of enzyme activity

Crude enzyme extracts from *Daqu* (HTD) were prepared by homogenizing the sample with a buffer solution ($\text{Na}_2\text{HPO}_4/\text{C}_6\text{H}_8\text{O}_7$, pH 6.0) at a 1:10 (w/v) ratio for enzymatic assays. Glucoamylase and amylase activities were determined using previously described methods (5). One

unit (U) of amylase activity was defined as the amount of enzyme liberating 1 mg of glucose within 5 min, whereas one unit (U) of glucoamylase activity was defined as the amount of enzyme releasing 1 mg of glucose in 30 min. Moisture content of HTD was calculated using the following equation: $\text{Moisture content (\%)} = [(\text{wet weight} - \text{dry weight}) / \text{wet weight}] \times 100$ (Zhang, Luo, et al., 2025).

2.3. Analysis of non-volatile metabolites

2.3.1. Sample preparation and extraction

The *Daqu* sample stored at -80°C refrigerator was thawed on ice. A 400 μL solution (Methanol:Water = 7:3, V/V) containing internal standard was added into 20 mg sample, and vortexed for 3 min. The sample was sonicated in an ice bath for 10 min and vortexed for 1 min, and then placed at -20°C for 30 min. The sample was then centrifuged at 12000 rpm for 10 min (4°C). And the sediment was removed, then centrifuged the supernatant at 12000 rpm for 3 min (4°C). A 200 μL aliquots of supernatant were transferred for LC-MS analysis.

2.3.2. HPLC conditions

All samples were analyzed using two LC-MS/MS methods. One aliquot was analyzed using positive ion conditions and was eluted from T3 column (Waters ACQUITY Premier HSS T3 Column 1.8 μm , 2.1 mm $\times$ 100 mm) using 0.1% formic acid in water as solvent A and 0.1% formic acid in acetonitrile as solvent B in the following gradient: 5 to 20% in 2 min, increased to 60% in the following 3 mins, increased to 99% in 1 min and held for 1.5 min, then come back to 5% mobile phase B within 0.1 min, held for 2.4 min. The analytical conditions were as follows, column temperature, 40°C ; flow rate, 0.4 mL/min; injection volume, 4 μL ; Another aliquot was using negative ion conditions and was the same as the elution gradient of positive mode. All the methods alternated between full scan MS and data dependent MSn scans using dynamic exclusion. MS analyses were carried out using electrospray ionization in the positive ion mode and negative ion mode using full scan analysis over m/z 75–1000 at 35000 resolution. Additional MS settings are: ion spray voltage, 3.5 KV or 3.2 KV in positive or negative modes, respectively; Sheath gas (Arb), 30; Aux gas, 5; Ion transfer tube temperature, 320°C ; Vaporizer temperature, 300°C ; Collision energy, 30,40,50 V; Signal Intensity Threshold, $1.00\text{E}+06$ cps; Top N vs Top speed, 10; Exclusion duration, 3 s. Non-volatile metabolites analyzed via LC-MS/MS were annotated based on accurate mass and MS/MS fragmentation patterns matching with internal databases (MSI Level 2).

2.4. Analysis of volatile compounds

A precisely weighed 2.00 g aliquot of each ground sample was placed into a 20 mL sample vial, followed by the addition of 5 mL of deionized water and 1.8 g of sodium chloride (S. Yang et al., 2022). The volatile compounds were further evaluated by adopting GC-MS (7890B GC System, 5977 A MSD). Each sample (1 μL) was employed for analysis on a DB-FFAP capillary column (60 m $\times$ 250 μm $\times$ 0.25 μm , Agilent Technologies, Santa Clara, CA, USA). Helium (99.999%) was used as a carrier gas at a constant flow rate of 2.0 mL/min, and the inlet temperature was 250°C . The oven temperature was maintained at 40°C initially and was held for 3 min, ramped to 150°C at a rate of $2^{\circ}\text{C}/\text{min}$ and was held for 2 min, and then finally was raised to 230°C at a rate of $7^{\circ}\text{C}/\text{min}$ and held was for 8 min. The temperature of the transfer line was 250°C , and that of the MS ion source was set at 230°C . The ionization energy of the electron impact mass spectra was 70 eV, and the acquisitions were over an m/z scan range of 35–450 amu. The injection was performed in splitless mode. Three parallel experiments were conducted for each sample, and the volatile compounds were identified by comparing their mass spectra in the NIST (2020) baijiu flavor compounds database developed by our group and their RI from the DB-FFAP column with those of the authentic standards. The retention indices of the compounds were calculated using alkanes (C8–C40). The n-alkanes

of C8–C40 were configured in the n-hexane solution, which was selected to take into account the protection of the column. The solution concentration was 20 µg/L. The same heating procedure as the sample was selected. The injection mode was splitless, and the injection volume was 1 µL. After the RI were calculated according to the formula, they were compared with the values in the literature so as to realize the further characterization of the compounds.

2.5. DNA extraction and metagenomic sequencing

0.2 g of *Daqu* was used to extract total genomic DNA with the FastPure Stool DNA Isolation Kit (Magnetic bead) (MJYH, Shanghai, China) according to manufacturer's instructions. Concentration and purity of extracted DNA was determined with SynergyHTX and NanoDrop2000, respectively. DNA quality was checked on 1% agarose gel. DNA extract was fragmented to an average size of about 350 bp using Covaris M220 (Gene Company Limited, China) for paired-end library construction. Paired-end library was constructed using NEXTFLEX Rapid DNA-Seq (Bioo Scientific, Austin, TX, USA). Paired-end sequencing was performed on Illumina NovaSeq™ X Plus (Illumina Inc., San Diego, CA, USA) at Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China) using NovaSeq X Series 25B Reagent Kit according to the manufacturer's instructions (www.illumina.com).

The quality-filtered data were assembled using MEGAHIT (<http://github.com/voutcn/megahit>, version 1.1.2). Contigs with a length ≥ 300 bp were selected as the final assembling result. Open reading frames (ORFs) from each assembled contigs were predicted using Prodigal (Hyatt et al., 2010) (<https://github.com/hyattprodigal/Prodigal>, version 2.6.3) and a length ≥ 100 bp ORFs were retrieved.

A non-redundant gene catalog was constructed using CD-HIT (Fu et al., 2012) (<http://weizhongli-lab.org/cd-hit/>, version 4.7) with 90% sequence identity and 90% coverage. Gene abundance for a certain sample was estimated by SOAPaligner (R. Li et al., 2008) (<https://github.com/ShujiaHuang/SOAPaligner>, version soap2.21 release) with 95% identity.

2.6. Taxonomic and functional annotation

The best-hit taxonomy of non-redundant genes was obtained by aligning them against the NCBI NR database by DIAMOND (Buchfink et al., 2015) (<http://ab.inf.uni-tuebingen.de/software/diamond/>, version 2.0.13) with an e-value cutoff of $1e-5$. Similarly, the functional annotation (GO, KEGG, eggNOG, CAZy, CARD, PHI) of non-redundant genes was obtained. Based on the taxonomic and functional annotation and the abundance profile of non-redundant genes, the differential analysis was carried out at each taxonomic, functional, or gene-wise level by Kruskal-Wallis test.

2.7. Metaproteomics analysis

2.7.1. Elution program

100 µg protein re-suspended with Triethylammonium bicarbonate buffer (TEAB) which with the final concentration of 100 mM. The mixture was reduced with Tris(2-carboxyethyl) phosphine (TCEP) which with the final concentration of 10 mM at 37 °C for 60 min and alkylated with iodoacetamide (IAM) which with the final concentration of 40 mM at room temperature for 40 min in darkness. After centrifugation at 10000 g at 4 °C for 20 min, the pellet was collected, which re-suspended with 100 µL Triethylammonium bicarbonate buffer (TEAB) which with the final concentration of 100 mM. Trypsin was added at 1:50 trypsin to-protein mass ratio and incubated at 37 °C overnight. The uPAC High Throughput column (75 µm $\times$ 5.5 cm, Thermo, USA) was used with solvent A (water with 2% acetonitrile and 0.1% formic acid) and solvent B (water with 80% acetonitrile and 0.1% formic acid). The elution program was used as following gradient: 0–0.1 min, 4%–8% buffer B; 0.1–1 min, 8%–12.5% B; 1–1.1 min, 12.5%–12.6% B; 1.1–3.6

min, 12.6%–22.5% B; 3.6–5.8 min, 22.5%–45% B; 5.8–6.4 min, 45%–99% B; 6.4–8 min, 99%–99% B.

2.7.2. DIA mass detection and protein identification

Data-independent acquisition data were acquired using an Orbitrap Astral mass spectrometer operated in DIA mode. The detection was carried out over a mass range of 70–1050 m/z (MS1), and 150–2000 m/z (MS2). The DIA mass spectrometric raw data were searched and processed with the Spectronaut software (Version 19) software. The parameters are as follows: the peptide length range was set to 7–52; Enzyme cutting site was trypsin/P; The maximum missed cleavage site was 2; Carbamidomethylation of cysteines as fixed modification, and oxidation of methionines and protein N-terminal acetylation as variable modifications; Protein FDR ≤ 0.01 , Peptide FDR ≤ 0.01 , Peptide Confidence $\geq 99\%$, XIC width ≤ 75 ppm. The protein quantification method was MaxLFQ.

2.8. Statistical analysis

The data were analyzed on the free online platform of Majorbio Cloud Platform (www.majorbio.com). *Daqu* sequences were analyzed on the online platform of Majorbio/I-Sanger Cloud Platform (<https://cloud.majorbio.com/>) (Han et al., 2024). Briefly, the raw sequencing reads were trimmed of adapters, and low-quality reads (length < 50 bp or with average quality value < 20) were removed by fastp (<https://github.com/OpenGene/fastp>, version 0.20.0). Reads were aligned to the human genome by BWA (H. Li & Durbin, 2009) (<http://bio-bwa.sourceforge.net>, version 0.7.17) and any hit associated with the reads and their mated reads were removed. The microbe-metabolite co-occurrence network was constructed based on Spearman's rank correlation analysis. Only robust and statistically significant correlations (Spearman correlation coefficient $|r| > 0.6$ and $p < 0.05$) were retained as edges in the network visualization.

3. Results and discussion

3.1. Physicochemical and enzymatic characteristics

Characterization of physicochemical parameters and key enzymatic activities is essential for assessing high-temperature *Daqu* quality and regulating fermentation dynamics (Wang, Wang, Huang, et al., 2020). We analyzed saccharifying power, protease activities, and acidity in *Daqu* samples collected from six production zones spanning the upper, middle, and lower reaches of the Chishui River (Fig. 1A). Downstream *Daqu* exhibited significantly higher amylase activity (*Daqu*₆: 361.49 ± 13.33 U; *Daqu*₅: 343.55 ± 9.67 U) compared to middle- and upstream counterparts (Fig. 1B). Neutral protease activities in downstream *Daqu* (271.63 ± 11.75 U and 244.78 ± 6.40 U) were markedly elevated relative to those from upper and middle reaches (Fig. 1C). Conversely, upstream *Daqu* samples displayed significantly greater acidity (1.40 ± 0.14 mmol/10 g and 1.45 ± 0.094 mmol/10 g) than those from middle and lower reaches (Fig. 1D). These distinctions are likely attributable to divergent raw materials, geographic conditions, microbial consortia, and metabolic functions (Yan et al., 2025; Zhang, Ran, et al., 2025). Overall, substantial variations in saccharification capacities persist among the three high-temperature *Daqu* types despite analogous production protocols, implicating differential microbial composition and metabolic outputs.

3.2. Comparative analysis of non-volatile metabolites in *Daqu* from various geographic origins

Non-targeted metabolomics analysis of *Daqu* revealed 672 non-volatile metabolites detected in anion and cation modes. Cation-mode profiling was dominated by amino acids and derivatives (27.6%), small peptides (27.1%), heterocyclic compounds (16.6%), and organic

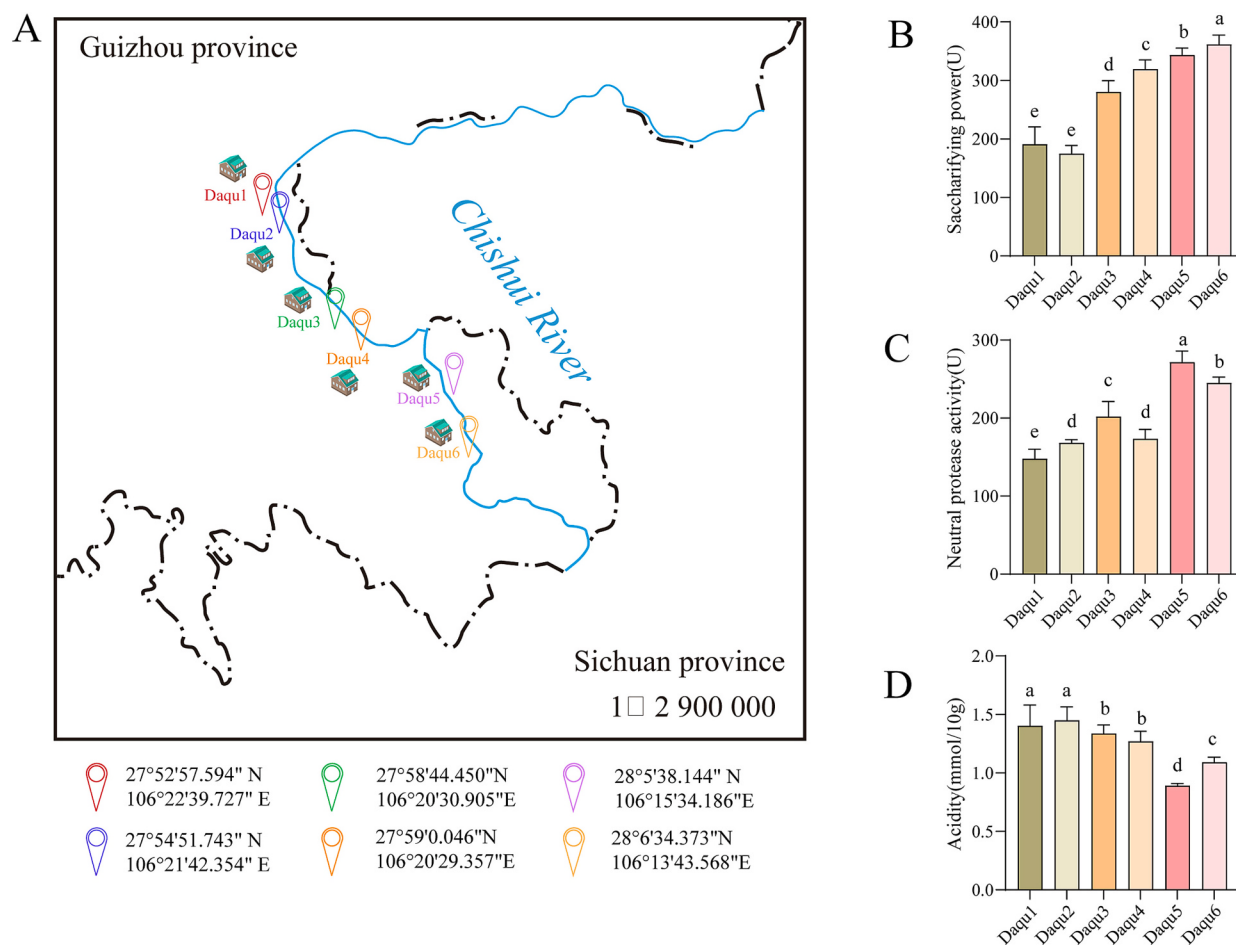


Fig. 1. Enzymatic activities and chemical characteristics of *Daqu* samples. (A) Geographic coordinates and sampling locations of high-temperature *Daqu* along the Chishui River. (B) Saccharification power of *Daqu* across representative regions. (C) Neutral protease activity in regionally collected *Daqu*. (D) Acidity levels in *Daqu* samples from distinct regions.

acids/derivatives (13.1%) (Fig. 2A). Anion mode primarily identified phenolic compounds (35%), amino acids (20.4%), and sugar alcohol derivatives (8.9%) (Fig. 2B). PLS-DA clearly separated upstream (A), midstream (B), and downstream (C) groups along the Chishui River (Fig. 1C, D), with proximal clustering of A and B indicating metabolic similarity between upstream and midstream *Daqu*. Model validity was established through permutation testing ($n = 200$; Fig. 2E, F). Volcano plot analysis (Fig. 2G) identified 25 differential metabolites between A and B, substantially fewer than those between B and C (71 metabolites) or A and C (76 metabolites), highlighting distinct downstream metabolic signatures. Integration of volcano plot thresholds ($p < 0.05$, $|\log_2FC| > 1$; Table S1) and PLS-DA VIP scores (> 1 ; Fig. 2H) defined watershed-specific markers. A-region *Daqu* showed significant enrichment of Ectoine, Proline betaine, 2,3,5,6-Tetramethylpyrazine, and 2-Phenyl-ethanol—microbially derived secondary metabolites. Ectoine and Proline betaine, known as compatible solutes synthesized by halophilic/drought-tolerant bacteria for osmoregulation (Miao et al., 2025; Weinisch et al., 2018), likely support microbial viability in mature *Daqu* (~10% moisture). Elevated levels of organic acids/derivatives (Vanillin, 4-Hydroxybenzaldehyde, d-Camphoric acid, 2-Aminobenzoic acid, Azelaic acid) in A and B may drive higher acidity in A and B. Conversely, C accumulated more amino acids (Glutamine, L-Allothreonine) and small peptides (Glycyl-L-phenylalanine, Ile-Lys), aligning with its enhanced protease activity and suggesting greater protease-producing microbial populations. Core flavor compounds (2,3,5,6-Tetramethylpyrazine and Vanillin in A) exhibited upstream enrichment. Notably, 2,3,5,6-Tetramethylpyrazine contributes nutty, roasted aromas

to high-temperature *Daqu* and Moutai-flavor baijiu (Luo, Liao, Luo, et al., 2025). These watershed-dependent metabolic patterns reflect specialized microbial community structures and enzymatic networks, collectively governing flavor development and microbial adaptation strategies.

3.3. Geographic divergence in volatile compounds of *Daqu* across river reaches

Daqu aroma significantly contributes to the flavor of Jiang-flavor baijiu. Therefore, the aroma of *Daqu* from different production areas influences the style of Jiang-flavor baijiu, making it essential to explore the specificity and core volatile compounds of *Daqu* across different river reaches. Using HS-SPME-GC-MS, we detected 83 flavor compounds in six types of high-temperature *Daqu*. PLS-DA successfully separated the volatile compounds of *Daqu* from three regions (Fig. 3A), indicating distinct flavor profiles among high-temperature *Daqu* from the three river reaches. This further confirms significant aroma differences in high-temperature *Daqu* across regions and suggests notable biogeographic characteristics in its microbial communities.

Sixteen characteristic compounds (VIP ≥ 1 , concentration $\geq 10 \mu\text{g/g}$) contributing most to regional differences were identified based on content and differential metrics (Fig. 3C). Comparative analysis of these volatile compounds (Fig. 3D) revealed that pyrazines, such as tetramethyl-pyrazine, trimethyl-pyrazine, and 2,3-dimethyl-pyrazine, were more abundant in upstream and downstream *Daqu*. Specifically, tetramethyl-pyrazine levels in upstream *Daqu* ($169.35 \pm 34.26 \mu\text{g/g}$)

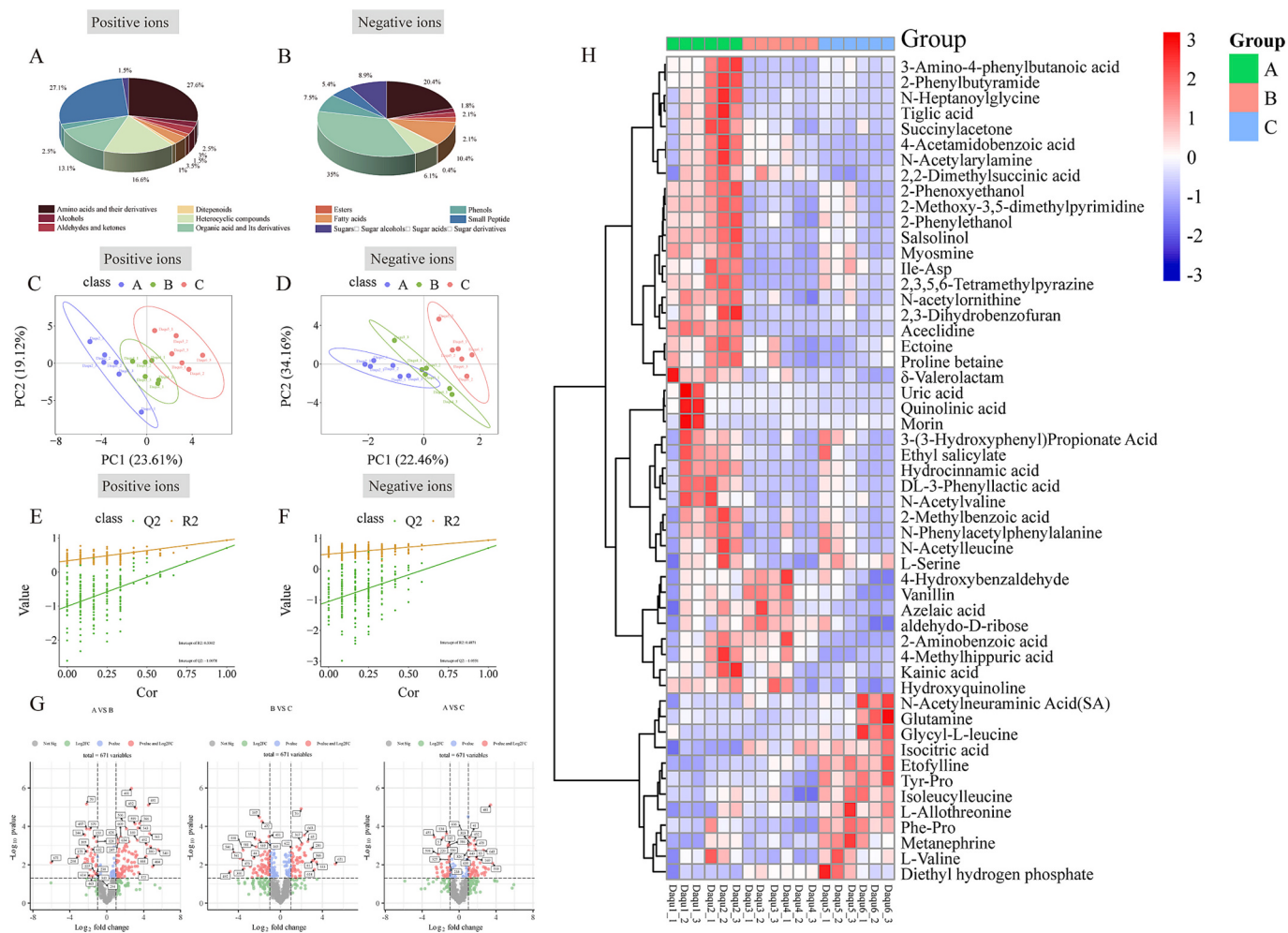


Fig. 2. Differences in non-volatile metabolites of *Daqu* from different river basins. (A, B) Differences in compound categories and proportions under two detection modes. (C, D) PLS-DA analysis of *Daqu* compounds differences. (E, F) 200 permutation tests for PLS-DA. (G) Fold change comparisons of *Daqu* compounds across different basins. (H) Analysis of compound concentrations in *Daqu* from different basins.

were significantly higher than in midstream ($7.72 \pm 2.64 \mu\text{g/g}$) and downstream *Daqu* ($143.19 \pm 49.5 \mu\text{g/g}$), aligning with Non-targeted metabolomics' compound trends. Additionally, alcohols like benzyl alcohol, phenylethyl alcohol, and 3-methyl-1-butanol were elevated in downstream *Daqu*. The downstream 3-methyl-1-butanol content ($2.71 \pm 0.55 \mu\text{g/g}$) significantly exceeded midstream ($1.63 \pm 0.78 \mu\text{g/g}$) and upstream levels ($1.15 \pm 0.32 \mu\text{g/g}$); phenylethyl alcohol reached $31.65 \pm 6.21 \mu\text{g/g}$, notably higher than upstream ($19.48 \pm 4.9 \mu\text{g/g}$) and midstream *Daqu* ($22.64 \pm 5.67 \mu\text{g/g}$).

Midstream *Daqu* exhibited higher levels of compounds such as 2-methyl-hexanoic acid and 1,2,4-trimethoxybenzene (Fig. 3D). These volatile disparities further demonstrate significant geographic differentiation in volatile metabolite composition among *Daqu* along the Chishui River. Pyrazines—known for nutty and roasted aromas—are typically produced by thermotolerant microorganisms (e.g., *Bacillus*) during high-temperature fermentation (He et al., 2025; Luo, Liao, Wu, et al., 2025). Alcohols like 3-methyl-1-butanol and phenylethyl alcohol derive from yeast-mediated Ehrlich pathway conversion of amino acids (Siripong et al., 2020), suggesting downstream conditions (e.g., lower temperature, higher humidity, or nutrient availability) favor yeast growth, enhancing alcohol fermentation and aroma synthesis.

Integrated enrichment analysis of volatile and non-volatile compounds (Fig. 3E) identified piperidinones, phenylpropanoic acids, and benzenes with high enrichment ratios, though amino acids, peptides, analogues, and methoxyphenols showed the most significant

enrichment ($P < 0.00001$). As illustrated in Fig. 3F, amino acids, peptides, and analogues encompassed the most differentially abundant metabolites ($n = 10$), followed by methoxyphenols ($n = 3$) and benzoic acids/derivatives ($n = 3$), indicating pronounced disparities in nitrogen metabolism and aromatic compound synthesis among microbial communities across river reaches. Studies reveal that divergent amino acids (arginine, proline, glutamate, glutamine, tryptophan) correlate strongly with *Bacillus* (Xing et al., 2025), likely linked to its protease-producing capacity (Y. Liu et al., 2022; Zeng et al., 2022). This underscores *Bacillus*' critical role in nitrogen metabolism during high-temperature *Daqu* fermentation. Notably, methoxyphenols—particularly 2-methoxy-4-vinylphenol and vanillin—were markedly enriched in our study (Fig. 3F). These key aroma compounds (G. Shi et al., 2024) may arise from microbial ferulate decarboxylase activity (Langos & Granvogl, 2016; Witthuhn et al., 2012). Collectively, these results demonstrate marked geographical divergence in non-volatile and volatile metabolites in Chishui River *Daqu*, with compounds such as tetramethylpyrazine, 4-vinylguaiacol, phenethyl alcohol, and 3-methyl-1-butanol being most prominent. These differences likely stem from microbial community variations. Subsequent research should focus on enzymes and functional microorganisms linked to these differential compounds.

3.4. Geographical heterogeneity of microbial communities

The microbiota involved in HTD fermentation primarily originates

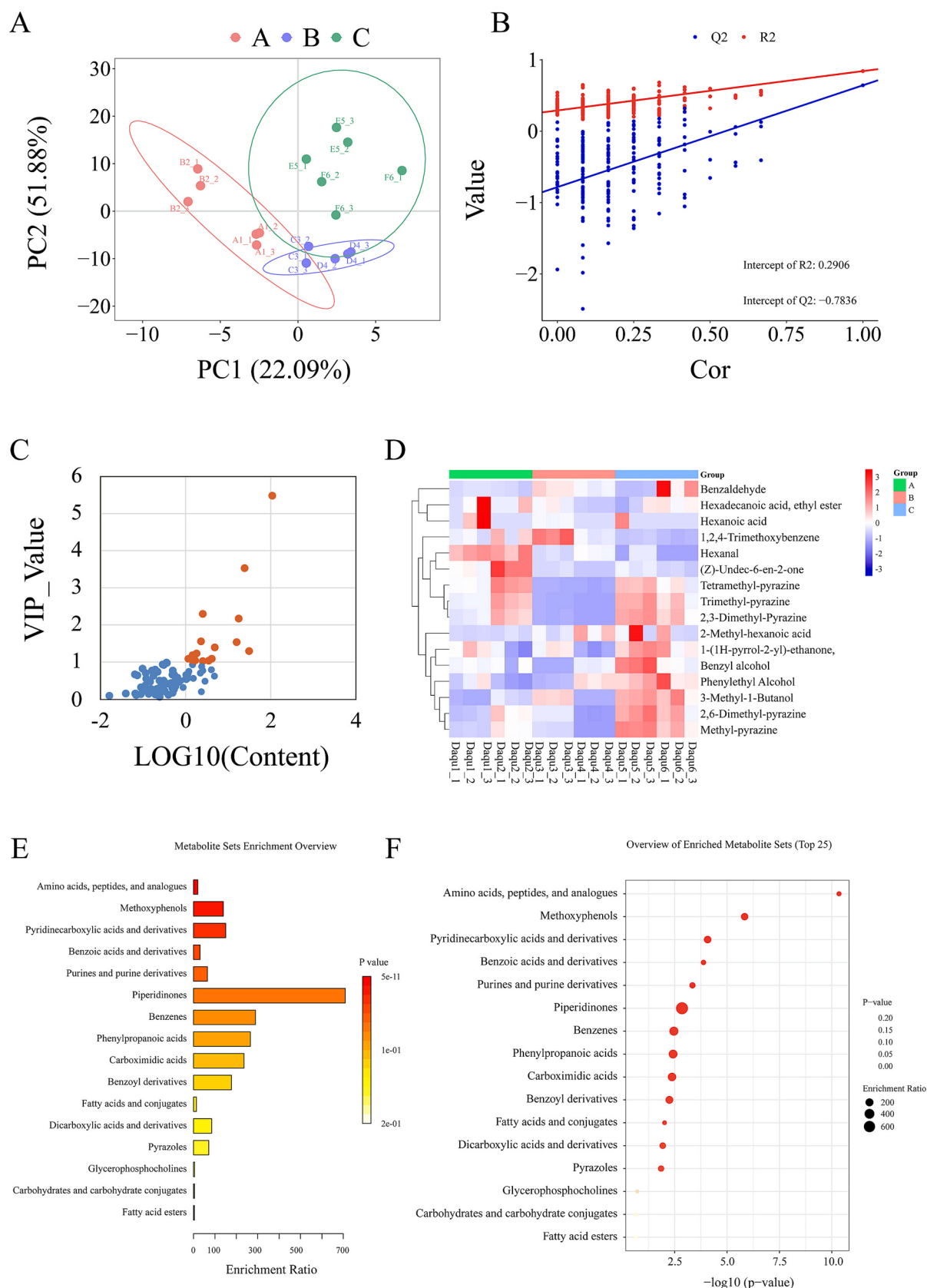


Fig. 3. (A) PLS-DA of volatile compounds across *Daqu* variants. (B) Permutation testing. (C) VIP-Content score plot. (D) Abundance of differential metabolites among *Daqu* types. (E) KEGG enrichment analysis for non-volatile/volatile compounds. (F) Significance of pathway-based compound enrichment.

from natural enrichment of local environmental microorganisms and directional introduction of exogenous microbes, leading to significant differences in microbial composition and metabolic activity among different *Daqu* samples. Based on metagenomic analysis, we explored the microbial community structural variations in high-temperature *Daqu* across different river basins. We displayed the microbial composition of *Daqu* from distinct basins (Fig. 4A, B). Using RPKM values, we calculated the relative abundance of microorganisms in various *Daqu*. At the phylum level, predominant bacteria in high-temperature *Daqu* were Bacillota and Actinomycetota, with minor Pseudomonadota present in downstream *Daqu*; predominant fungi were Ascomycota and Mucoromycota, with minor Basidiomycota in downstream samples. At the

genus level, high-temperature *Daqu* across all regions consisted mainly of bacteria such as *Kroppenstedtia*, *Bacillus*, *Lentibacillus*, *Oceanobacillus*, *Desmospora*, and *Scopulibacillus*; fungi included *Aspergillus*, *Paecilomyces*, *Rasamsonia*, *Talaromyces*, *Lichtheimia*, *Monascus*, and *Penicillium*. Unlike upstream *Daqu*, midstream *Daqu* contains more *Pallidibacillus* and *Saccharopolyspora*; downstream *Daqu* contains more *Weissella*. Significant differences are also observed in fungal communities. The mean relative abundance of *Monascus* in upstream regions is 7.3%, which sharply declines in midstream and downstream areas. *Talaromyces* shows lower abundance in downstream *Daqu*, while *Meyerozyma*, *Clavispora*, and *Debaryomyces* exhibit higher relative abundance. These results suggest distinct microbial community structures in *Daqu* across different river

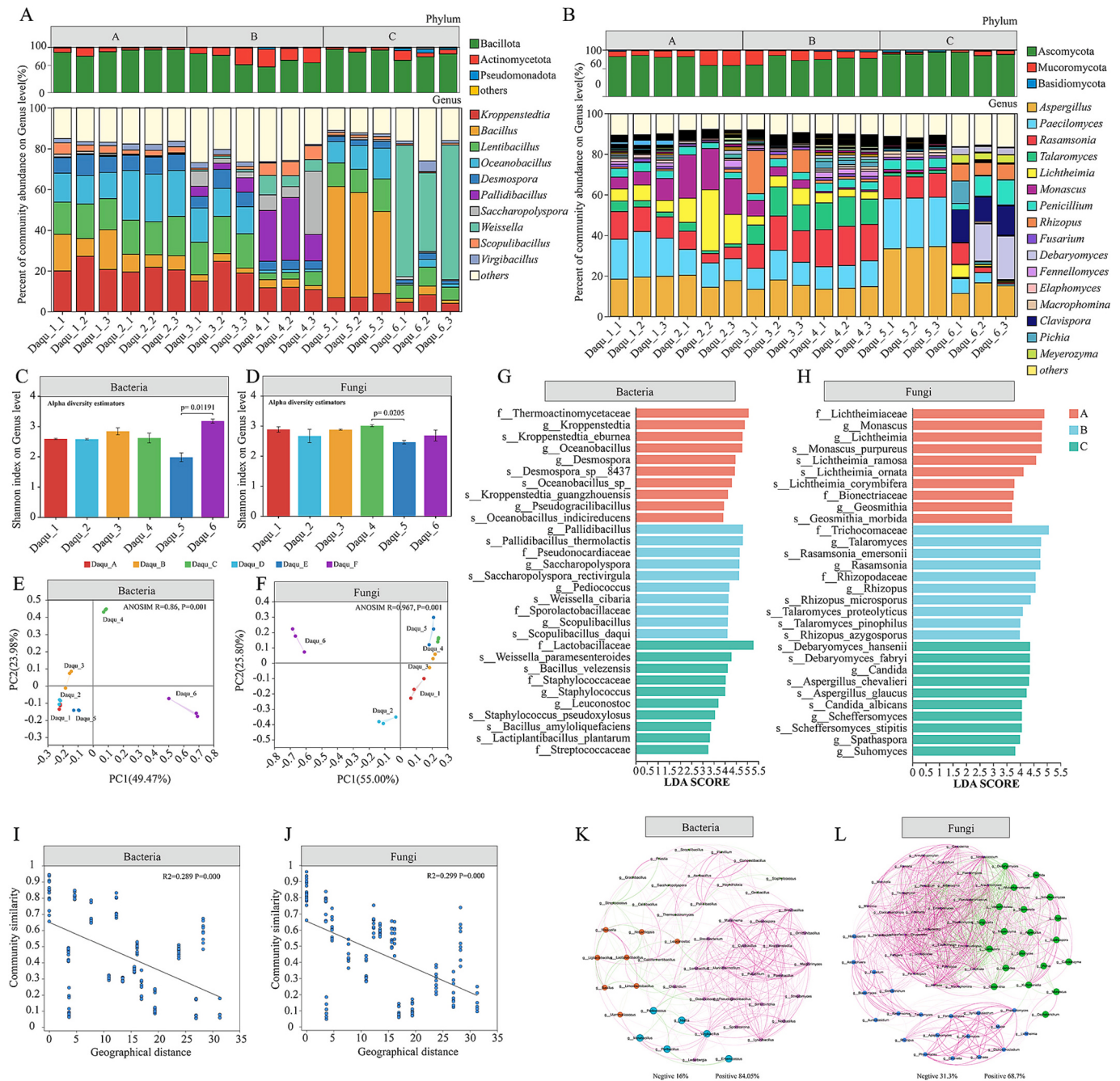


Fig. 4. Microbial community profiling of *Daqu* across distinct fermentation stages. (A, B) Taxonomic composition of *Daqu* microbiota from sampled basins; (C, D) Alpha-diversity indices; (E, F) Non-metric multidimensional scaling (NMDS) ordination based on Bray-Curtis dissimilarity; (G, H) Linear discriminant analysis (LDA) effect sizes for fungal and bacterial taxa across basins; (I, J) Distance-decay relationships characterizing microbial community similarity; (K, L) Fungal-bacterial co-occurrence network patterns.

basins. Changes in microbial α -diversity of high-temperature *Daqu* were studied (Figs. 4C, D). The Shannon index shows no significant overall variation among six regions (Fig. 4C), indicating that species richness remains relatively stable across basins despite significant compositional differences. This stability likely results from selective pressures in the high-temperature fermentation environment maintaining the core microbiome. Non-metric multidimensional scaling (NMDS) revealed bacterial community clustering: *Daqu1*, *Daqu2*, *Daqu3*, and *Daqu5* grouped together but differed significantly from *Daqu4* and *Daqu6* (ANOSIM $R = 0.86$, $P = 0.001$) (Fig. 4E). Fungal communities showed clear separation: *Daqu1* and *Daqu2* clustered distinctly from *Daqu3* and *Daqu4*, while *Daqu6* exhibited pronounced isolation (ANOSIM $R = 0.967$, $P = 0.001$).

To identify basin-specific microbial biomarkers, LEfSe analysis revealed: In watershed A, the LDA value of Thermoactinomycetaceae was 5.056, marking it as a significant biomarker. Key taxa also included *Kroppenstedtia* (4.91) and *Desmospora* (4.49), accompanied by *Oceanobacillus* (4.42) and *Pseudogracilibacillus* (3.92); Basin B was characterized by *Pallidibacillus* (4.53), *Saccharopolyspora* (4.51), and *Scopulibacillus* (4.13); downstream regions were dominated by *Lactobacillaceae* (5.25), primarily *Weissella* (4.25). *Bacillus* and *Staphylococcus* (4.07) were differential genera downstream, with *Bacillus* represented mainly by *B. velezensis* (4.15) and *B. amyloliquefaciens* (3.42). For fungi, upstream *Daqu* was distinguished by *Lichtheimiaceae* (mainly *Monascus* (4.73) and *Lichtheimia* (4.68)); Basin B was associated with *Trichocomaceae* (*Rasamsonia* (4.72), *Talaromyces* (4.68), *Rhizopus* (4.52)); Basin C was typified by *Weissella*, *Debaryomyces*, and *Aspergillus*. Studies reveal that microorganisms including *Saccharopolyspora*, *Pallidibacillus*, and *Scopulibacillus* exhibit thermotolerance up to 60 °C (Wu et al., 2016; Yao et al., 2016); In contrast, the upstream signature genus *Kroppenstedtia* survives temperatures ≤ 50 °C, while *Oceanobacillus* tolerates ≤ 45 °C (Nam et al., 2008); Downstream high-temperature *Daqu* is dominated by *Lactobacillaceae* with lower thermal resilience (Y. Sun et al., 2022), suggesting that microbial thermal adaptation drives community divergence across river basins—a variation attributable to distinct *Daqu*-production processes and environmental conditions (Yuan et al., 2025). Given that dispersal shapes community assembly (Wang, Wang, Shen, et al., 2020), we investigated spatial constraints on basin-specific microbiomes through distance-decay analysis (Fig. 4I, J). Similarity in bacterial communities of high-temperature *Daqu* declined significantly across geographical distances ($R^2 = 0.289$; $p < 0.001$), paralleled by fungal communities ($R^2 = 0.299$; $p < 0.001$). This distance-decay pattern confirms dispersal limitation in shaping *Daqu* microbiota composition, consistent with observations in freshwater bacterial assemblages, phyllosphere bacteria, and diverse ecosystems (Finkel et al., 2012; L. Liu et al., 2015). Co-occurrence network analysis further revealed that despite the distance-decay effect, specific microorganisms contribute to maintaining the stability of the high-temperature *Daqu* microbiome. Positive associations dominated both bacterial (84.05% of edges) and fungal networks, with fungal interactions being notably more extensive (1064 edges vs. 370 bacterial edges). Key bacterial genera exhibiting high positive connectivity—*Lentibacillus* (22 edges), *Oceanobacillus* (24 edges), and *Kroppenstedtia* (23 edges)—are likely ecological keystone taxa underpinning bacterial community stability. The fungal co-occurrence network largely comprised rare genera, suggesting their persistence may be facilitated by synergistic interactions (M. Zhu, Yin, et al., 2022). Crucially, *Paecilomyces* (43 edges) and *Penicillium* (42 edges), displaying the highest network connectivity, emerged as principal stabilizers of the fungal assemblage.

The present study demonstrates distinct microbial structures in high-temperature *Daqu* across different watersheds, suggesting these microbial biomarkers may correlate with variations in *Daqu*'s functional phenotypes. Furthermore, metagenomic analyses from multiple studies have characterized the microbial communities in high-temperature *Daqu*, identifying dominant bacterial genera such as *Bacillus*, *Kroppenstedtia*, *Oceanobacillus*, *Saccharopolyspora*, and *Staphylococcus*, along

with key fungal genera including *Paecilomyces*, *Aspergillus*, *Rasamsonia*, *Thermomyces*, and *Thermoascus* (Luo, Liao, Wu, et al., 2025; Yang, Fan, & Xu, 2024b; Zhu, Cheng, et al., 2023), consistent with our findings. Moreover, the decay effect of geographic distance influences the similarity of microbial community composition in *Daqu* from different basins, which further verifies that the microbial community composition of high-temperature *Daqu* is constrained by dispersal processes. The divergence in microbial functionality—where upstream habitats select for pyrazine biosynthesis and downstream habitats favor alcohol/ester formation—reflects intense selection by long-term environmental pressures (G. Zhang et al., 2025). Upstream environments, with potentially lower ambient humidity and specific local micro-climates, create selective pressures that favor robust, thermotolerant r-strategists like *Bacillus*. Conversely, downstream facilities, characterized by higher environmental water activity and unique long-term domestication pools within the fermentation facilities, provide a niche highly conducive to the proliferation of fermentative yeasts and molds. However, significant microbial interactions sustain the basic microbial community of high-temperature *Daqu*, wherein its dominant microbial genera also function as crucial ecological keystones.

3.5. Region-specific heterogeneity in microbial community functionality

Metagenomic analysis revealed the compositional structure and functional potential of high-temperature *Daqu* microbial communities, identified key microbial taxa in samples from different watersheds, and predicted their potential metabolic pathways. These findings provide crucial insights for understanding the microbial basis of *Daqu* fermentation. However, metagenomics can only reflect “who is there” and “what they can do” but cannot directly confirm whether these genes are expressed and functionally active during actual fermentation. To elucidate the functional dynamics of the microbial community in *Daqu*, we employed proteomics to profile the expression of key enzymes and further investigated how microbiome functional expression contributes to the formation of *Daqu* with distinct functional attributes. We identified 35,610 polypeptides (Fig. S1A), with protein abundance peaking in the 11–20 kDa molecular weight range, followed by the 21–30, 31–40, and 41–50 kDa ranges (Fig. S1B). Functional annotation (GO/KEGG/COG) revealed that Molecular Function comprised the largest category, dominated by “catalytic activity” (20,918 proteins), signifying enzymatic predominance (Fig. S1C). COG annotation identified “amino acid transport and metabolism” as the most abundant functional category, succeeded by “carbohydrate transport and metabolism,” “energy production and conversion,” “lipid transport and metabolism,” and “general function prediction” (Fig. S1D), underscoring the central role of enzymes in defining *Daqu* functionality and flavor. Concordant KEGG analysis showed “carbohydrate metabolism” as the top pathway (5692 proteins), followed by “amino acid metabolism” (4176), “energy metabolism” (2896), and “translation” (2825) (Fig. S1E). These annotations consistently prioritized proteins integral to carbohydrate/amino acid metabolism and energy conversion, thereby pinpointing core functional microbiota governing *Daqu*'s metabolic phenotype. Research indicates that amino acid differences in *Daqu* serve as a crucial marker for evaluating its quality. Amino acids, particularly arginine, exhibit significant disparities between the two. Environments characterized by high temperature, high humidity, and low acidity contribute to the abundance of amino acids in *Daqu*, while mechanical pressing affects the stability of the core microbiota. The presence of amino acids is considered a novel potential biomarker influencing *Daqu* quality. (Xing et al., 2025).

Using metagenomic sequencing data, NR-based taxonomic annotation of detected proteins was performed. Species abundances were derived by aggregating proteins mapped to each taxon, generating taxonomically resolved abundance profiles (Fig. 5A). Fungal proteins dominated the proteome across *Daqu* from three geographically distinct regions, with the genus *Trichosporon* exhibiting the highest mean relative abundance (24%), followed by *Bacillus* (9.3%), *Paecilomyces* (6.9%),

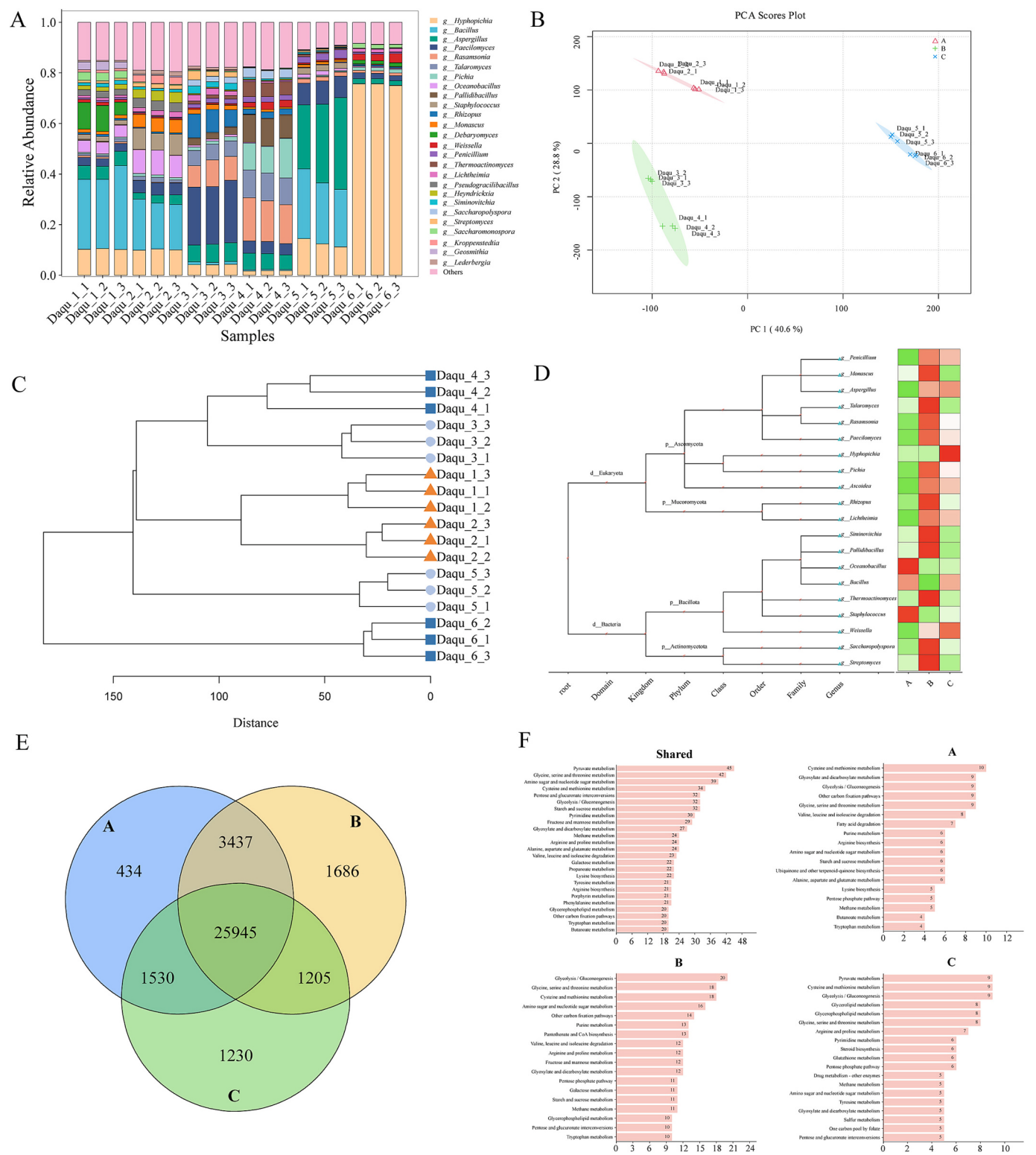


Fig. 5. Microbial community analysis across sample groups. (A) Stacked bar plot showing relative abundances of microbial taxa, with colors indicating genera. (B) Principal coordinates analysis (PCoA) ordination based on Bray-Curtis dissimilarity, depicting structural divergence among groups (color-coded). (C) Hierarchical clustering analysis of samples using protein profiles. (D) Heatmap displaying microbial abundances annotated at the protein level, with rows representing taxonomic units and columns denoting samples. (E) Unique and shared protein sets identified in *Daqu* samples across watersheds. (F) Distribution frequency of unique and shared proteins in KEGG pathways.

Oceanobacillus (5.2%), *Rasamsonia* (5.04%), and *Aspergillus* (4.5%). Significant spatial heterogeneity in microbial protein abundance was observed across production regions. Studies demonstrate that fungal diversity and network interactions in solid-state fermentation systems enhance the functional efficacy of fermentation processes. Such microbial complexity elevates fermented product quality through facilitation of carbon metabolic pathways and biosynthesis of flavor compounds. These findings strongly suggest that fungal communities play pivotal roles in shaping the functional enzyme repertoire and flavor characteristics of *Daqu* (Zhang, Xu, et al., 2022).

A critical question arises regarding how these low-abundance fungi maintain dominant enzymatic activities in high-temperature *Daqu*, where core temperatures typically exceed 60 °C—a threshold generally inhibitory to fungal growth. This phenomenon can be explained by a combination of biological, spatial, and temporal mechanisms. First, from a biological standpoint, certain endemic fungal populations in *Daqu* exhibit remarkable intrinsic thermotolerance, capable of surviving temperatures up to 55–60 °C (Zhou et al., 2022). These robust strains likely secrete thermotolerant isozymes that maintain catalytic stability even under high thermal stress. Second, the solid-state architecture of *Daqu* bricks creates highly heterogeneous spatial micro-niches (Yang, Fu, et al., 2024). While the core undergoes intense thermogenesis, the outer crust and peripheral layers experience significant heat dissipation, providing a cooler thermal refuge that permits fungal survival and active enzyme secretion. Third, and crucially from a temporal perspective, *Daqu* production involves a prolonged storage (maturation) phase following peak fermentation (M. Zhu, Zheng, et al., 2022). As the *Daqu* cools, it creates an environmental window that permits the resuscitation of heat-stressed fungi and the secondary colonization of ambient fungi from the storage environment. These newly introduced or recovered fungi can rapidly proliferate and secrete enzymes during storage, thereby contributing significantly to the accumulated functional proteome observed in mature *Daqu*.

In the upstream region, the abundances of *Oceanobacillus*, *Debaromyces*, *Monascus*, *Saccharomonospora*, *Heyndrickxia*, and *Staphylococcus* were significantly higher. The midstream region exhibited higher abundances of *Rhizopus*, *Penicillium*, *Thermoactinomyces*, *Pallidibacillus*, *Rasamsonia*, *Saccharopolyspora*, and *Paecilomyces*. Downstream regions were dominated by *Hyphopichia*, *Pichia*, and *Weissella*. The spatial distribution of microbial protein abundances largely aligned with metagenomic data but showed notable differences. Specifically, proteomic data revealed lower protein abundances for *Kroppenstedtia*, *Lentibacillus*, and *Pallidibacillus*—dominant in metagenomes—whereas fungi like *Monascus*, *Penicillium*, and *Lichtheimia*, though prevalent in metagenomes, exhibited variable and non-dominant abundances in the proteome. Notably, *Hyphopichia*, nearly undetected in metagenomes, consistently dominated the proteomes of all *Daqu* samples, suggesting that low-abundance microorganisms may critically contribute to enzymatic functions. While this “abundance-activity paradox” underscores the potential of rare biosphere members acting as keystone functional executors, it is imperative to interpret this discrepancy with methodological caveats. The divergence between genomic potential and expressed protein profiles may not be exclusively biological; several technical biases inherent to multi-omics approaches could also contribute to this pattern. First, bias in DNA extraction efficiency is a well-documented challenge in microbiome research; thick-walled fungal spores or specific cellular structures of *Hyphopichia* might resist complete lysis during total DNA extraction, leading to a significant underestimation of its true genomic abundance (Luise R et al., 2025). Second, inherent biological variations, such as differences in rRNA gene copy numbers among diverse fungal species, can further skew taxonomic profiling and relative abundance calculations in sequencing data (Matchado et al., 2024). Third, the metagenomic assembly process inherently struggles with low-abundance taxa (Anwar et al., 2025). It is highly probable that the fragmented reads of *Hyphopichia* experienced “assembly dropouts,” failing to assemble into recognizable contigs due

to insufficient sequencing depth for the rare biosphere. Conversely, from a metaproteomic perspective, the high sequence homology among related fungal proteins might lead to incorrect protein annotations or false-positive taxonomic assignments during database searching, artificially inflating the apparent functional dominance of *Hyphopichia* (Conti et al., 2023). Therefore, while our integrated omics approach strongly suggests that low-abundance taxa drive critical metabolic fluxes, acknowledging these technical limitations is crucial. Future in vitro validation—such as absolute quantification via qPCR or targeted parallel reaction monitoring proteomics—is required to definitively decouple actual biological activity from methodological artifacts. Principal component analysis (PCA) of *Daqu* proteins from upstream, midstream, and downstream sections of the Chishui River (Fig. 5B) revealed distinct separation, indicating functionally unique microbial communities in different river segments that may influence *Daqu* functionality. Hierarchical clustering analysis using Euclidean distance matrices and average linkage (Fig. 5C) grouped samples primarily by watershed, with downstream samples showing lower dispersion (*Daqu*_6 formed a separate cluster). We further analyzed microbial relative abundances across taxa (Fig. 5D). Results showed higher bacterial abundances (*Oceanobacillus*, *Staphylococcus*, *Bacillus*) upstream; increased fungal abundances (*Penicillium*, *Monascus*, *Aspergillus*, *Talaromyces*, *Rasamsonia*) alongside elevated bacteria (*Pallidibacillus*, *Thermoactinomyces*, *Saccharopolyspora*, *Streptomyces*) midstream; and higher abundances of *Hyphopichia*, *Weissella*, and *Aspergillus* downstream.

Furthermore, we quantified core and basin-specific microbial proteins in *Daqu* samples from distinct river watersheds. The three basins shared 25,945 proteins, with Group A exhibiting 434 differentially abundant proteins, while Groups B and C contained 1686 and 1230 unique proteins, respectively. Enzyme enrichment analysis based on KEGG annotations revealed *Daqu*'s functional signatures linked to volatile compound profiles. Within the shared enzymatic repertoire, pyruvate metabolism represented the most dominant pathway (45 enzymes), succeeded by glycine, serine and threonine metabolism (42 enzymes), amino sugar and nucleotide sugar metabolism (39 enzymes), cysteine and methionine metabolism (34 enzymes), pentose and glucuronate interconversions (32 enzymes), glycolysis/gluconeogenesis (32 enzymes), and starch and sucrose metabolism (32 enzymes). Critically, carbohydrate- and amino acid-related pathways demonstrated cross-basin conservation. Differential enzyme analysis indicated that group-specific enzymatic markers consistently correlated with cysteine and methionine metabolism, glyoxylate and dicarboxylate metabolism, glycolysis/gluconeogenesis, and degradation pathways for valine, leucine, and isoleucine. In addition to the unique enzymes, we further analyzed the differences in shared proteins among different *Daqu* samples (Fig. 6A). Based on fold change (FC) and *p*-values, we compared the proteomic differences between *Daqu* from different drainage basins. Compared to Group B, Group A exhibited 7406 up-regulated proteins and 12,735 down-regulated proteins. Group A showed 10,567 up-regulated proteins and 12,846 down-regulated proteins versus Group C. Group B demonstrated 12,777 up-regulated and 9474 down-regulated proteins compared to Group C, further indicating functional expression disparities among *Daqu* from three distinct basins.

Using predefined thresholds ($p < 0.01$, $FC \geq 2$ or ≤ 0.5), we screened KEGG pathways related to *Daqu* functions based on prior research (Luo, Liao, Wu, et al., 2025), and identified relevant enzymes and associated microbes. In carbohydrate metabolism pathways (ko00500), Group A showed significantly higher abundance of enzymes linked to *Thermoascus*, *Monascus*, *Desmospora*, *Oceanobacillus*, *Lentibacillus*, and *Bacillus* compared to Group B. Group B exhibited greater abundance of *Thermomyces*, *Talaromyces*, *Rasamsonia*, *Paecilomyces*, and *Aspergillus* (Fig. 6B). For Group C, *Lichtheimia*, *Hyphopichia*, *Weissella*, and *Bacillus* also participated in carbohydrate degradation, with abundances significantly higher than Group A. These differences may reflect selective pressure from differential substrate availability on microbiota across basins. Bacteria typically adapt better to rapidly degrade simple

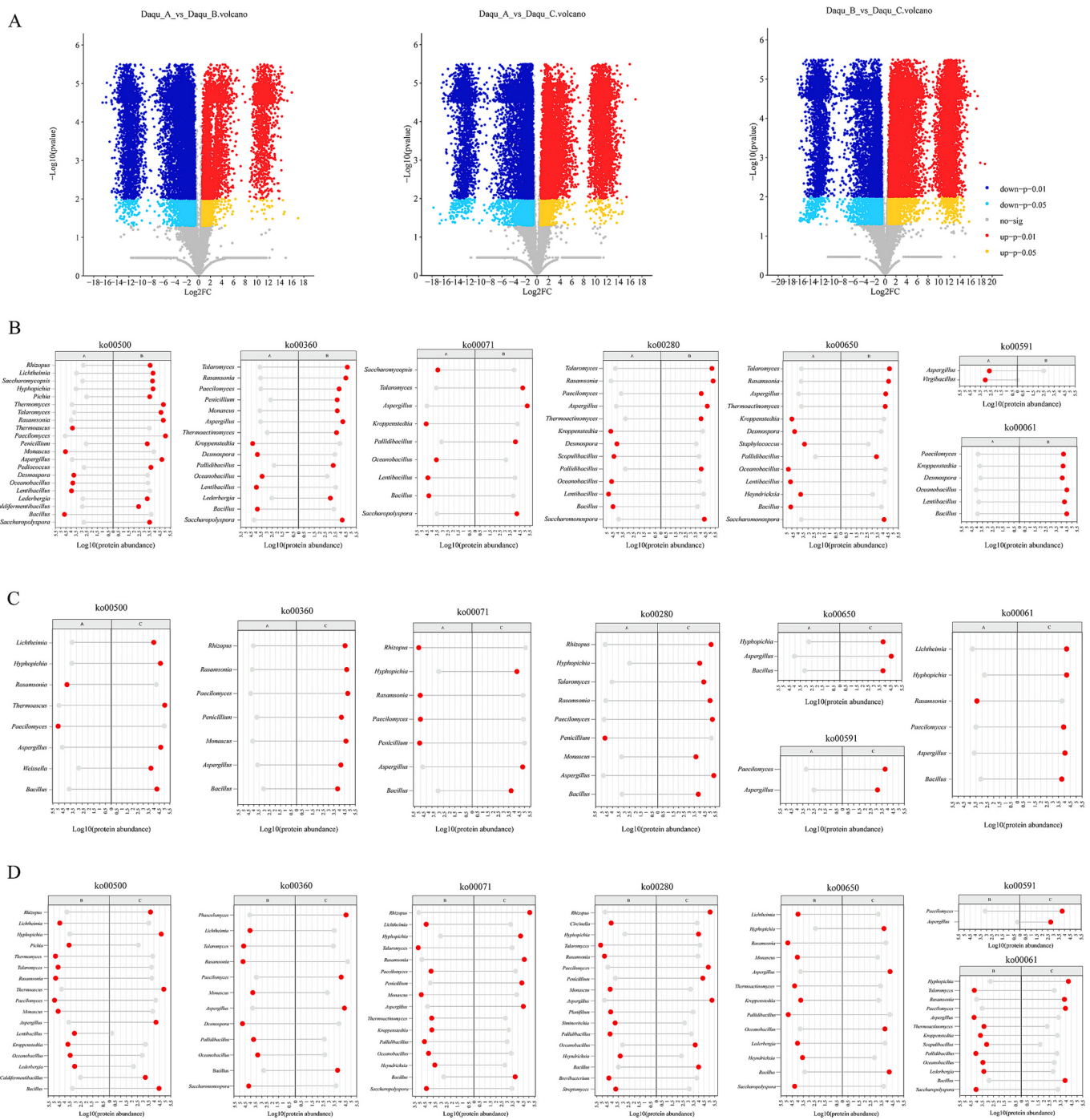


Fig. 6. (A) Differentially expressed proteins in *Daqu* across distinct drainage basins. (B) Differentially abundant microorganisms within functional KEGG pathways for Group A versus Group B. (C) Differentially abundant microorganisms within functional KEGG pathways for Group B versus Group C. (D) Differentially abundant microorganisms within functional KEGG pathways for Group C versus Group D.

carbohydrates (e.g., glucose, maltose, sucrose) (Jiang et al., 2020), while fungi likely play key roles in decomposing complex polysaccharides like starch and cellulose (S.-B. Yang et al., 2025). This emphasizes the importance of microbial functional complementarity in maintaining ecosystem stability. From an ecological and evolutionary perspective, the 'abundance-activity paradox' observed in *Bacillus* highlights a classic temporal decoupling between biomass accumulation and metabolic activity (Bressuire-Isaard et al., 2018). During the initial stages of *Daqu* fermentation, when temperatures are moderate, *Bacillus* likely adopts an r-selection strategy, rapidly proliferating to occupy the ecological niche. However, as the core temperature peaks (>60 °C), these bacteria

undergo sporulation to endure the severe environmental stress. Consequently, while metagenomic sequencing captures the massive genomic DNA from these resilient endospores, their actual protein synthesis is virtually halted, explaining their low in situ protein expression in the mature *Daqu*. Different microorganisms may optimize resource utilization and avoid competitive exclusion through division of labor. In the phenylalanine-derived aromatic volatile compound pathway (ko00360), the abundances of bacteria such as *Kroppenstedtia*, *Desmospora*, *Pallidibacillus*, *Oceanobacillus*, *Lentibacillus*, and *Bacillus* in Group A were significantly higher than those in Group B. In contrast, fungi including *Talaromyces*, *Ramsonia*, *Paecilomyces*, and *Penicillium* in

Group B exhibited greater enzyme expression. Furthermore, in Group C, except for *Phascolomyces*, *Paecilomyces*, *Aspergillus*, and *Bacillus*, the abundances of related enzymes in other microorganisms were lower than in Group B. This indicates that Group B has distinctive characteristics in aromatic amino acid metabolism, followed by Group C. Studies show that phenylalanine, as an aromatic amino acid, produces metabolites such as 2-phenylethanol (2-PE), which is highly valued for its rose-like aroma. Prior research revealed that phenylalanine supplementation during soy (tofu) whey fermentation significantly altered flavor profiles by promoting sugar utilization and enhancing the formation of higher alcohols and their corresponding esters. This further underscores the critical role of fungi in aromatic amino acid metabolism (Chua & Liu, 2020), consistent with our earlier detection of notably higher alcohol content in *Daqu* starter from the C watershed. In the long-chain volatile compound pathway (ko00071), the abundances of enzymes expressed by microorganisms like *Saccharomycopsis*, *Kroppenstedtia*, *Oceanobacillus*, and *Lentibacillus* in Group A were significantly higher than in Group B, while fungi such as *Rhizopus*, *Rasamsonia*, *Paecilomyces*, and *Penicillium* showed higher abundances than in Group C. Compared to Group B, Group C displayed elevated fungal enzyme expression, whereas Group B primarily exhibited bacterial enzyme expression. In the volatile compound pathway of branched-chain amino acids, Group A mainly shows high expression of bacteria-related enzymes, including *Kroppenstedtia*, *Desmospora*, *Oceanobacillus*, and *Lentibacillus*. Compared with Group A, Group B exhibits high expression of fungi-related enzymes such as *Rasamsonia*, *Paecilomyces*, and *Aspergillus*. Additionally, Group C also displays high expression of fungi-related enzymes relative to Group A, including *Rhizopus*, *Talaromyces*, and *Rasamsonia* (Fig. 6C). In the pyrazine synthesis pathway, acetoin is a critical precursor compound. Its metabolism involves the ko00650 pathway, where Group A expresses predominantly bacterial enzymes. The abundance of bacteria-related enzymes in Group B is significantly higher than in Group C (Fig. 6D). Unlike other groups, the expression of *Bacillus* enzymes is substantially higher in Groups A and C compared to Group B. Studies demonstrate that *Bacillus* can efficiently produce acetoin and further convert it to tetramethylpyrazine (TMP) (Peng et al., 2020). Our previous research revealed significantly higher 4-methylpyrazine content in Groups A and C than in Group B, indicating that enhanced *Bacillus*-related enzyme expression likely drives this outcome. In the fatty acid synthesis pathway (ko00061), Group B exhibited higher abundances of bacteria-associated enzymes such as *Kroppenstedtia*, *Thermoactinomyces*, *Scopulibacillus*, and *Oceanobacillus*, while Group C showed higher abundances of fungi-associated enzymes such as *Hyphopichia*, *Rasamsonia*, and *Paecilomyces*. For volatile compounds derived from long-chain fatty acids (ko00591), Group C demonstrated significantly higher enzyme abundances of microbes including *Aspergillus* and *Paecilomyces* than other groups, indicating these compounds are primarily mediated by fungi. Previous studies employing high-throughput amplicon sequencing combined with metagenomic and metatranscriptomic analyses identified *Kroppenstedtia* as an important contributor to fatty acid synthesis during fermentation of Chinese soy-sauce aroma-type baijiu. This finding highlights the significance of *Kroppenstedtia* in fatty acid synthesis (J. Zhang et al., 2024). Furthermore, members of *Actinobacteria* possess diverse fatty acid synthesis capabilities, exemplified by the FAS II system in *Streptomyces* (Gago et al., 2011), though the fatty acid synthesis system in *Thermoactinomyces* warrants further investigation. These patterns strongly suggest functional niche partitioning of microbial communities under varying environmental or treatment conditions.

3.6. Metabolic network underlying core HTD metabolism

To determine the expression levels of enzyme-coding genes critical for substrate degradation and characteristic flavor formation, and to assess the contributions of core microorganisms, we reconstructed a metabolic network encompassing seven key pathways based on KEGG database annotations (Fig. S2). We identified 42 relevant enzymes, and

their protein abundance profiles are presented in the heatmap (Fig. 7A). Given the complexity and diversity of starch saccharification routes, we evaluated enzymes across multiple pathways to pinpoint the principal routes for starch decomposition in *Daqu*. Key enzyme abundances within the starch decomposition module were substantially higher in B and C samples compared to the A. This finding aligns with the lower saccharification power observed in A. The starch breakdown pathways primarily involved conversion to glucose via intermediate steps like dextrin and maltose. Phenylalanine degradation largely involved EC 2.6.1.5 and EC 2.6.1.9, while branched-chain amino acid degradation was chiefly mediated by EC 1.4.1.9. Starch degradation capacity and flavor compound production are paramount for evaluating HTD quality. These processes integrate substrate degradation with precursor biosynthesis, driven by synergistic microbial metabolism. Liquefaction and saccharification powers measure microbial starch degradation efficiency and function by swelling and rupturing starch granules, creating a homogeneous paste to enhance subsequent saccharification. Starch-to-glucose conversion primarily occurs via maltose alpha-D-glucosyltransferase/alpha-amylase (EC 3.2.1.1), producing dextrin and maltose intermediates, or directly via EC 3.2.1.3. Notably, EC 3.2.1.3 was predominantly produced by fungi (*Paecilomyces*, *Lichtheimia*, *Monascus*, *Rhizomucor*, *Thermomyces*, *Saccharomycopsis*) (Fig. 7C), although some bacteria (*Bacillus*, *Kroppenstedtia*, *Lederbergia*) also contributed. Bacteria like *Weissella* and *Bacillus* possessed enzymes targeting maltose and dextrin (e.g., EC 3.2.1.10, EC 2.4.1.8). Consequently, *Bacillus* and *Kroppenstedtia* likely obtained carbon directly from starch depolymerization, whereas bacteria like *Weissella* appeared dependent on fermentable sugars (maltose, glucose) supplied by others, suggesting microbial community assembly in *Daqu* is influenced by substrate selectivity. Crucially, within B and C, the significantly higher abundance of EC 3.2.1.3 attributed to *Paecilomyces* likely underpinned their enhanced saccharification power compared to A. Tetramethylpyrazine is generated through a highly efficient bio-chemo synergistic mechanism rather than a strictly enzymatic pathway. Upstream *Daqu* exhibited elevated abundances of bacterial acetolactate decarboxylase (predominantly from *Bacillus*), which is crucial for the massive biological accumulation of the precursor, acetoin (P. wang et al., 2017). However, from a food chemistry perspective, the final conversion of acetoin to TMP is primarily driven by the abiotic, thermally-induced Maillard reaction, naturally facilitated by the >60 °C peak temperature (H. Xu et al., 2024). Thus, TMP enrichment is the cumulative result of biological precursor supply coupled with optimal abiotic thermal conditions. This correlates with the higher pyrazine content found in A. Long-chain fatty aldehyde synthesis involved enzymes like EC 2.1.3.199, predominantly expressed by fungi (*Paecilomyces*, *Hyphopichia*). For ester synthesis, short-chain esters depended mainly on carboxylesterase (EC 3.1.1.1) from bacteria (*Pallidibacillus*, *Bacillus*), while long-chain esters relied on triacylglycerol lipase (EC 3.1.1.3) from fungi (*Paecilomyces*, *Monascus*). Research has shown that *Monascus purpureus* can synthesize short-chain fatty acid esters in aqueous phases, and our findings are consistent with this (Y. Xu et al., 2021). Aromatic amino acid degradation involved EC 2.6.1.5 (expressed by fungi like *Thermoascus*, *Pichia*) and EC 2.6.1.9 (expressed by bacteria like *Bacillus*, *Pallidibacillus*, *Oceanobacillus*). Aryl-alcohol dehydrogenase (EC 1.1.1.90) was uniquely expressed by *Desmospora*. However, given the high abundance and overlapping function of enzymes like EC 1.1.1.- or EC 1.1.1.2 in HTD, these likely contribute to the synthesis of aromatic alcohols (e.g., phenethyl alcohol) and were expressed mainly by fungi (*Hyphopichia*, *Monascus*, *Lichtheimia*). BCAAs were degraded primarily via leucine dehydrogenase (EC 1.4.1.9) expressed by bacteria (*Pallidibacillus*, *Lentibacillus*, *Bacillus*), with subsequent conversion of aldehydes to acids (e.g., isovaleric acid) by aldehyde dehydrogenases (EC 1.2.1.3, EC 1.2.1.5, EC 1.2.1.10), the importance of branched-chain amino acids in microbial metabolism can be understood through their role in energy metabolism. Branched-chain amino acids serve as crucial energy sources, particularly in the absence of glucose. Research indicates that branched-chain amino acids can be

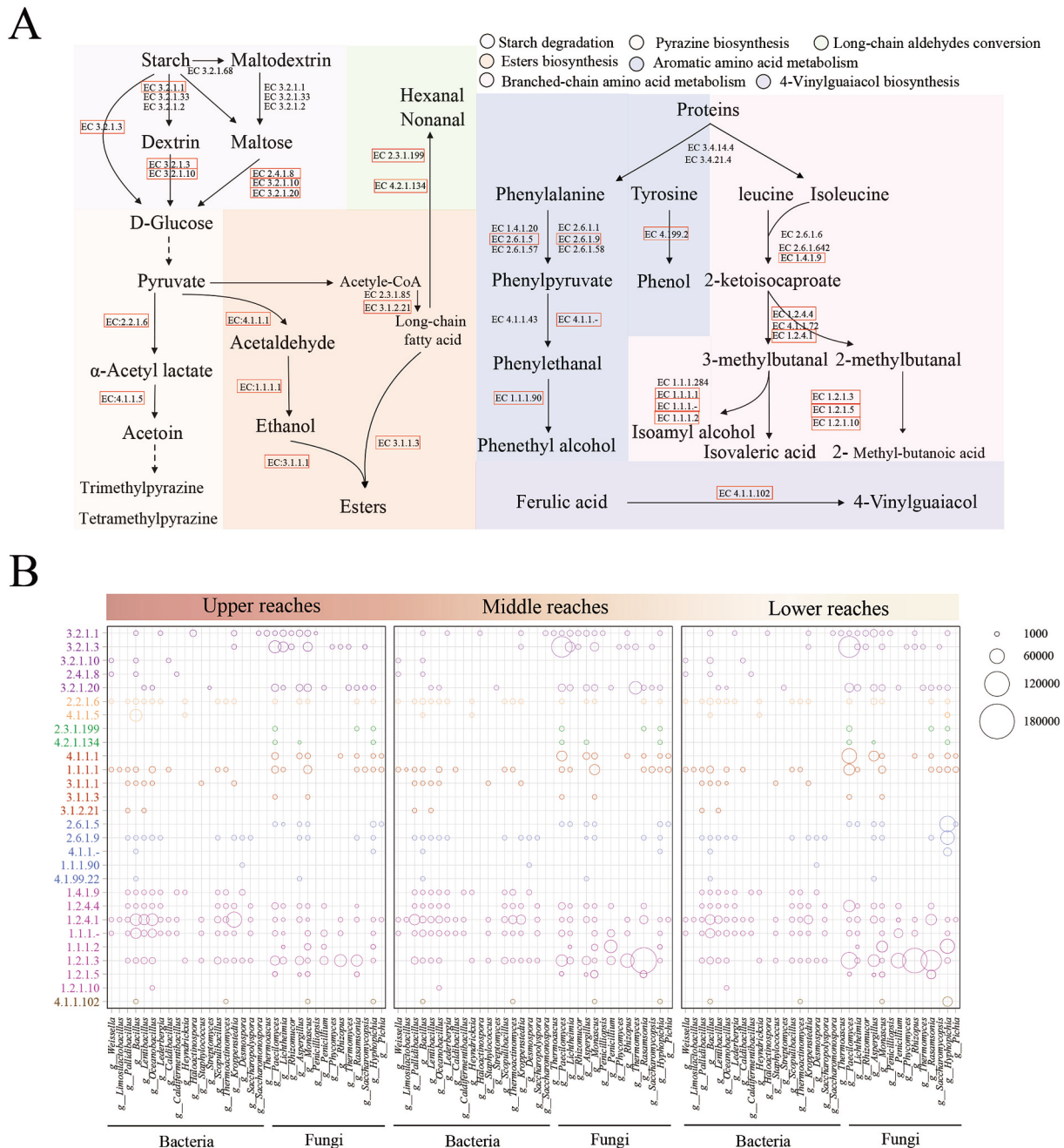


Fig. 7. (A) The organization of the high-temperature *Daqu* metabolic network comprising seven core modules; (B) Spatial distribution of microbiota and enzymatic profiles associated with substrate catabolism and flavor-biosynthetic pathways within the *Daqu* ecosystem.

converted into acetyl-CoA and enter the tricarboxylic acid cycle, thereby supplying energy to cells (Chen et al., 2024). Although bacteria such as *Pallidibacillus*, *Lentibacillus*, and *Bacillus* can also encode these aldehyde dehydrogenases, they exhibit high expression in fungi (*Rasamsonia*, *Thermomyces*, *Rhizopus*) and even higher expression in B and C, indicating fungi as key producers of characteristic branched-chain acids like isovaleric acid. Similarly, the biosynthesis of 4-vinylguaiacol is associated with *Bacillus*, *Thermoactinomyces*, *Hyphopichia*, and *Monascus*. Among these, *Hyphopichia* shows significantly higher abundance of ferulic acid decarboxylase (EC 4.1.1.102), particularly enriched in C. This finding aligns with the peak levels of 4-vinyl guaiacol detected in downstream *Daqu*. Numerous studies have demonstrated that microorganisms including *Bacillus*, *Enterobacter*, *Streptomyces*, and *Lactobacillus* convert ferulic acid to 4-vinylguaiacol via ferulic acid decarboxylase (EC

4.1.1.102) (Hadiza et al., 2012; Hunter et al., 2012; L.-H. Sun et al., 2018). Nevertheless, the microbial origins of guaiacol derivatives synthesis in (HTD) remain elusive. We plan to experimentally verify the key microorganisms contributing to guaiacol production in subsequent studies. Furthermore, our co-occurrence networks reveal a robust ‘cross-feeding’ mechanism. The high-abundance bacteria (such as *Bacillus*) function as pioneer degraders during early fermentation stages. By secreting robust extracellular proteases and amylases, they depolymerize recalcitrant macromolecules into bioavailable small peptides and fermentable sugars. This early-stage mass degradation paves the way for the later-stage fungi, supplying them with the essential small-molecule precursors required to drive the ultimate synthesis of complex flavor compounds (S. Liu et al., 2024). It is important to note that *Daqu* fermentation is a highly dynamic process lasting over a month. The

current study relies on a single-time-point multi-omics snapshot of the mature *Daqu* products. While this effectively captures the cumulative outcomes and final active microbiota, it may introduce bias regarding the temporal succession of the community. Future studies employing time-series-based multi-omics sampling are warranted to fully map the dynamic cross-feeding networks and temporal shifts in gene expression.

4. Conclusion

This integrated multi-omics study systematically analyzed high-temperature *Daqu* starters across the Chishui River, elucidating the mechanistic basis for geographic heterogeneity in Maotai-flavor baijiu terroir. Geographic segregation drove distinct microbial community assembly: Upstream sites were dominated by thermotolerant *Kroppenstedtia* and *Saccharopolyspora*, concomitant with elevated pyrazine biosynthesis; whereas downstream proliferation of *Weissella* and *Hyphopichia* promoted alcohol generation. Crucially, metaproteomics confirmed that metabolic functional divergence was governed by in situ enzyme activities—starch saccharification was primarily mediated by *Paecilomyces* and *Hyphopichia* derived α -amylase, while upstream pyrazine accumulation was regulated by *Bacillus*-originated acetolactate decarboxylase. Although spatial distance directly diminished microbial taxonomic similarity, fungal co-occurrence networks sustained ecosystem stability. The disconnect between phylogenetic biomarkers and functional drivers underscores the necessity of activity-based microbial profiling for terroir prediction. These results establish robust microbiome-function-flavor linkages, offering industrial standardization biomarkers while preserving regional distinctiveness. Future work should prioritize in vitro validation of enzymatic pathways and practical implementation of synthetic microbial consortia. Collectively, our findings advocate for a paradigm shift in the assessment and engineering of fermentation starters: from a taxonomy-centric to a function- and activity-centric perspective. The identified functional niches and active microbial consortia provide a actionable roadmap for the rational design of synthetic communities, aiming to achieve both flavor standardization and the preservation of coveted geographical identities in fermented food production.

CRediT authorship contribution statement

Dandan Song: Writing – review & editing, Writing – original draft, Validation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Xian Zhong:** Writing – review & editing, Validation, Formal analysis. **Guihu Zhang:** Writing – review & editing. **Juan Chen:** Writing – review & editing. **Yansong Xue:** Writing – review & editing, Conceptualization. **Liang Yang:** Writing – review & editing, Project administration, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China (NSFC 32460269); Guizhou Provincial Basic Research Program (Natural Science) (No. QianKeHeJiChu-ZD [2025] 018); Guizhou Key Laboratory of Microbial Resources Exploration in Fermentation industry (Qiankehe Platform Talent-ZDSYS [2023] 007); The Science and Technology Innovation Team of Moutai Institute (MTXYTD202501); Moutai Institute & Guizhou Guiding distillery (XYNJ20240374, XYNJ20250032); Moutai Institute & Guangdong Li'er'an Chemical Industry Group Co., Ltd. (XYNJ20240104).

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.fochx.2026.103952>.

Data availability

Data will be made available on request.

References

- Anwar, S., Lamaudiere, M., Hassall, J., Dehinsilu, J., Bhuller, R. K., Hold, G. L., ... Sergaki, C. (2025). DNA reference reagents isolate biases in microbiome profiling: A global multi-lab study. *mSystems*, 10(11), e00466–e00525. <https://doi.org/10.1128/mSystems.00466-25>
- Bresciani, L., Custer, G. F., Koslicki, D., & Dini-Andreote, F. (2025). Interplay of ecological processes modulates microbial community reassembly following coalescence. *The ISME Journal*, 19(1), Article wraf041. <https://doi.org/10.1093/ismej/wraf041>
- Bressuire-Isoard, C., Broussolle, V., & Carlin, F. (2018). Sporulation environment influences spore properties in *Bacillus*: Evidence and insights on underlying molecular and physiological mechanisms. *FEMS Microbiology Reviews*, 42(5), 614–626. <https://doi.org/10.1093/femsre/fuy021>
- Buchfink, B., Xie, C., & Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nature Methods*, 12(1), 59–60. <https://doi.org/10.1038/nmeth.3176>
- Chen, H., Gao, Q., Liu, B., Zhang, Y., Fang, J., Wang, S., Chen, Y., & Chen, C. (2024). Identification of the global regulatory roles of RraA via the integrative transcriptome and proteome in *Vibrio alginolyticus*. *mSphere*, 9(7), Article e00020–24. <https://doi.org/10.1128/msphere.00020-24>
- Chua, J.-Y., & Liu, S.-Q. (2020). Effect of single amino acid addition on growth kinetics and flavor modulation by *Torulaspora delbrueckii* in soy (tofu) whey alcoholic beverage fermentation. *Food Research International*, 135, Article 109283. <https://doi.org/10.1016/j.foodres.2020.109283>
- Conti, A., Casagrande Pierantoni, D., Robert, V., Corte, L., & Cardinali, G. (2023). MinION sequencing of yeast mock communities to assess the effect of databases and ITS-LSU markers on the reliability of metabarcoding analysis. *Microbiology Spectrum*, 11(1), e01052–e01122. <https://doi.org/10.1128/spectrum.01052-22>
- Deng, L., Mao, X., Liu, D., Ning, X.-Q., Shen, Y., Chen, B., Nie, H.-F., Huang, D., & Luo, H.-B. (2020). Comparative analysis of physicochemical properties and microbial composition in high-temperature *Daqu* with different Colors. *Frontiers in Microbiology*, 11, Article 588117. <https://doi.org/10.3389/fmicb.2020.588117>
- Finkel, O. M., Burch, A. Y., Elad, T., Huse, S. M., Lindow, S. E., Post, A. F., & Belkin, S. (2012). Distance-decay relationships partially determine diversity patterns of Phyllosphere Bacteria on Tamrix trees across the Sonoran Desert. *Applied and Environmental Microbiology*, 78(17), 6187–6193. <https://doi.org/10.1128/AEM.00888-12>
- Fu, L., Niu, B., Zhu, Z., Wu, S., & Li, W. (2012). CD-HIT: Accelerated for clustering the next-generation sequencing data. *Bioinformatics*, 28(23), 3150–3152. <https://doi.org/10.1093/bioinformatics/bts565>
- Gago, G., Diacovich, L., Arabolaza, A., Tsai, S.-C., & Gramajo, H. (2011). Fatty acid biosynthesis in actinomycetes. *FEMS Microbiology Reviews*, 35(3), 475–497. <https://doi.org/10.1111/j.1574-6976.2010.00259.x>
- Hadiza, A. A., Shahid, I., Kim, W. C., & Maznah, I. (2012). Biotransformation of ferulic acid to 4-vinyl guaiacol by *Lactobacillus farcinis*. *African Journal of Biotechnology*, 11(5), 1177–1184. <https://doi.org/10.5897/AJB11.520>
- Han, C., Shi, C., Liu, L., Han, J., Yang, Q., Wang, Y., Li, X., Fu, W., Gao, H., Huang, H., Zhang, X., & Yu, K. (2024). Majorbio cloud 2024: Update single-cell and multiomics workflows. *iMeta*, 3(4), Article e217. <https://doi.org/10.1002/imt.2.217>
- He, Y., Liu, Y., Sun, J., Qiao, M., Wang, L., Zhao, Q., Xiao, D., & Guo, X. (2025). Characterization of the key aroma-active compounds causing the Douchi aroma in high-temperature *Daqu* via integrated molecular sensory science, metabolomics, and molecular docking. *Food Chemistry*, 493, Article 146125. <https://doi.org/10.1016/j.foodchem.2025.146125>
- Hou, Q., Wang, Y., Qu, D., Zhao, H., Tian, L., Zhou, J., Liu, J., & Guo, Z. (2024). Microbial communities, functional, and flavor differences among three different-colored high-temperature *Daqu*: A comprehensive metagenomic, physicochemical, and electronic sensory analysis. *Food Research International*, 184, Article 114257. <https://doi.org/10.1016/j.foodres.2024.114257>
- Hunter, W. J., Manter, D. K., & Van Der Lelie, D. (2012). Biotransformation of ferulic acid to 4-Vinylguaiacol by *Enterobacter soli* and *E. Aerogenes*. *Current Microbiology*, 65(6), 752–757. <https://doi.org/10.1007/s00284-012-0222-4>
- Hyatt, D., Chen, G.-L., LoCascio, P. F., Land, M. L., Larimer, F. W., & Hauser, L. J. (2010). Prodigal: Prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics*, 11(1), 119. <https://doi.org/10.1186/1471-2105-11-119>
- Jiang, N., Ma, X.-D., Fu, L.-H., Li, C.-X., Feng, J.-X., & Duan, C.-J. (2020). Identification of a unique 1,4- β -D-glucan glucosylhydrolase of glycoside hydrolase family 9 from *Cytophaga hutchinsonii*. *Applied Microbiology and Biotechnology*, 104(16), 7051–7066. <https://doi.org/10.1007/s00253-020-10731-8>
- Langos, D., & Granvogl, M. (2016). Studies on the simultaneous formation of aroma-active and toxicologically relevant vinyl aromatics from free phenolic acids during wheat beer brewing. *Journal of Agricultural and Food Chemistry*, 64(11), 2325–2332. <https://doi.org/10.1021/acs.jafc.5b05606>

- Leech, J., Cabrera-Rubio, R., Walsh, A. M., Macori, G., Walsh, C. J., Barton, W., Finnegan, L., Crispie, F., O'Sullivan, O., Claesson, M. J., & Cotter, P. D. (2020). Fermented-food metagenomics reveals substrate-associated differences in taxonomy and health-associated and antibiotic resistance determinants. *mSystems*, 5(6), e00522–20. doi: <https://doi.org/10.1128/mSystems.00522-20>.
- Li, H., & Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics (Oxford, England)*, 25(14), 1754–1760. <https://doi.org/10.1093/bioinformatics/btp324>
- Li, N., Fu, J., Zhang, G., Liu, J., Li, Z., Luo, R., & Li, L. (2023). Investigating the mechanism of the flavor formation in Sichuan sun vinegar based on flavor-orientation and metagenomics. *Current Research in Food Science*, 6, Article 100460. <https://doi.org/10.1016/j.crfis.2023.100460>
- Li, R., Li, Y., Kristiansen, K., & Wang, J. (2008). SOAP: Short oligonucleotide alignment program. *Bioinformatics*, 24(5), 713–714. <https://doi.org/10.1093/bioinformatics/btn025>
- Liu, L., Yang, J., Yu, Z., & Wilkinson, D. M. (2015). The biogeography of abundant and rare bacterioplankton in the lakes and reservoirs of China. *The ISME Journal*, 9(9), 2068–2077. <https://doi.org/10.1038/ismej.2015.29>
- Liu, S., Zhao, L., Li, M., Zhu, Y., Liang, D., Ma, Y., Sun, L., Zhao, G., & Tu, Q. (2024). Probiotic bacillus as fermentation agents: Status, potential insights, and future perspectives. *Food Chemistry: X*, 22, Article 101465. <https://doi.org/10.1016/j.fochx.2024.101465>
- Liu, Y., Fu, J., Wang, L., Zhao, Z., Wang, H., Han, S., Sun, X., & Pan, C. (2022). Isolation, identification, and whole-genome sequencing of high-yield protease bacteria from Daqu of ZhangGong Laojiu. *PLoS One*, 17(4), Article e0264677. <https://doi.org/10.1371/journal.pone.0264677>
- Luo, Y., Liao, H., Luo, Y., Gao, L., & Xia, X. (2025). Temperature-driven microbial assembly and flavor compound dynamics in high-temperature Daqu: Adaptive co-evolution for efficient Tetramethylpyrazine production. *International Journal of Food Microbiology*, 441, Article 111325. <https://doi.org/10.1016/j.ijfoodmicro.2025.111325>
- Luo, Y., Liao, H., Wu, L., Wu, M., Luo, Y., Yao, Y., Ji, W., Gao, L., & Xia, X. (2025). Temperature adaptability drives functional diversity and horizontal gene transfer within microbial communities in Daqu solid-state fermentation. *Bioresource Technology*, 433, Article 132770. <https://doi.org/10.1016/j.biortech.2025.132770>
- Matchado, M. S., Rühlemann, M., Reitmeier, S., Kacprowski, T., Frost, F., Haller, D., ... List, M. (2024). On the limits of 16S rRNA gene-based metagenome prediction and functional profiling. *Microbial Genomics*, 10(2). <https://doi.org/10.1099/mgen.0.001203>
- Miao, Q., Wang, Z.-B., Xie, Y., Hu, R., Huang, T., & Ni, S.-Q. (2025). Multi-omics evidence supporting the role of Halomonas enrichment contribution to high-salinity denitrification. *Water Research*, 287, Article 124413. <https://doi.org/10.1016/j.watres.2025.124413>
- Nam, J.-H., Bae, W., & Lee, D.-H. (2008). Oceanobacillus caeni sp. Nov., isolated from a Bacillus-dominated wastewater treatment system in Korea. *International Journal of Systematic and Evolutionary Microbiology*, 58(5), 1109–1113. <https://doi.org/10.1099/ijs.0.65335-0>
- Niu, J., Li, W., Du, B., Wu, Y., Lang, Y., Sun, B., Sun, W., & Li, X. (2025). Temporal heterogeneity of microbial communities and flavor metabolism during storage of high-temperature Daqu. *Food Chemistry*, 464, Article 141577. <https://doi.org/10.1016/j.foodchem.2024.141577>
- Peng, Q., Zhao, X., Wen, J., Huang, M., Zhang, J., & Song, F. (2020). Transcription in the acetoin catabolic pathway is regulated by AcoR and CcpA in Bacillus thuringiensis. *Microbiological Research*, 235, Article 126438. <https://doi.org/10.1016/j.micres.2020.126438>
- Shi, G., Fang, C., Xing, S., Guo, Y., Li, X., Han, X., Lin, L., & Zhang, C. (2024). Heterogeneous mechanism in high-temperature Daqu fermentation by traditional craft and mechanical craft: From microbial assembly patterns to metabolism phenotypes. *Food Research International*, 187, Article 114327. <https://doi.org/10.1016/j.foodres.2024.114327>
- Siripong, W., Angela, C., Tanapongpipat, S., & Runguphan, W. (2020). Metabolic engineering of Pichia pastoris for production of isopentanol (3-Methyl-1-butanol). *Enzyme and Microbial Technology*, 138, Article 109557. <https://doi.org/10.1016/j.enzmictec.2020.109557>
- Sun, L.-H., Lv, S.-W., Yu, F., Li, S.-N., & He, L.-Y. (2018). Biosynthesis of 4-vinylguaiacol from crude ferulic acid by Bacillus licheniformis DLF-17056. *Journal of Biotechnology*, 281, 144–149. <https://doi.org/10.1016/j.jbiotec.2018.07.021>
- Sun, Y., Peng, C., Wang, J., Guo, S., Sun, Z., & Zhang, H. (2022). Mesopic fermentation contributes more to the formation of important flavor compounds and increased growth of Lactobacillus casei Zhang than does high temperature during milk fermentation and storage. *Journal of Dairy Science*, 105(6), 4857–4867. <https://doi.org/10.3168/jds.2021-20949>
- Van Den Bossche, T., Armengaud, J., Benndorf, D., Blakeley-Ruiz, J. A., Brauer, M., Cheng, K., ... Li, L. (2025). The microbiologist's guide to metaproteomics. *American Chemical Society (ACS)*. <https://doi.org/10.26434/chemrxiv-2025-rh2q9>
- Wang, P., Wu, Q., Jiang, X., Wang, Z., Tang, J., & Xu, Y. (2017). Bacillus licheniformis affects the microbial community and metabolic profile in the spontaneous fermentation of daqu starter for Chinese liquor making. *International Journal of Food Microbiology*, 250, 59–67. <https://doi.org/10.1016/j.ijfoodmicro.2017.03.010>
- Wang, Y., Cai, W., Wang, W., Shu, N., Zhang, Z., Hou, Q., Shan, C., & Guo, Z. (2021). Analysis of microbial diversity and functional differences in different types of high-temperature Daqu. *Food Science & Nutrition*, 9(2), 1003–1016. <https://doi.org/10.1002/fsn3.2068>
- Wang, Y., Wang, K., Huang, L., Dong, P., Wang, S., Chen, H., Lu, Z., Hou, D., & Zhang, D. (2020). Fine-scale succession patterns and assembly mechanisms of bacterial community of Litopenaeus vannamei larvae across the developmental cycle. *Microbiome*, 8(1), 106. <https://doi.org/10.1186/s40168-020-00879-w>
- Wang, Z., Wang, C., Shen, C., Wang, S., Mao, J., Li, Z., Gänzle, M., & Mao, J. (2020). Microbiota stratification and succession of amylase-producing Bacillus in traditional Chinese Jiuqu (fermentation starters). *Journal of the Science of Food and Agriculture*, 100(8), 3544–3553. <https://doi.org/10.1002/jsfa.10405>
- Weinisch, L., Kühner, S., Roth, R., Grimm, M., Roth, T., Netz, D. J. A., ... Filker, S. (2018). Identification of osmoadaptive strategies in the halophile, heterotrophic ciliate Schmidingeriothrix salinarum. *PLoS Biology*, 16(1), Article e2003892. <https://doi.org/10.1371/journal.pbio.2003892>
- Witthuhn, R. C., Van Der Merwe, E., Venter, P., & Cameron, M. (2012). Guaiacol production from ferulic acid, vanillin and vanillic acid by alicyclobacillus acidoterrestris. *International Journal of Food Microbiology*, 157(1), 113–117. <https://doi.org/10.1016/j.ijfoodmicro.2012.04.022>
- Wu, H., Liu, B., & Pan, S. (2016). Saccharopolyspora subtropica sp. Nov., a thermophilic actinomycete isolated from soil of a sugar cane field. *International Journal of Systematic and Evolutionary Microbiology*, 66(5), 1990–1995. <https://doi.org/10.1099/ijsem.0.000976>
- Xing, S., Shi, G., Lu, J., Fang, C., Li, C., Yuan, S., Shi, F., Lin, L., & Zhang, C. (2025). The discrepancy in amino acids within high-temperature Daqu: A novel metabolic marker for the quality evaluation of Daqu. *Food Chemistry*, 470, Article 142645. <https://doi.org/10.1016/j.foodchem.2024.142645>
- Xu, H., Ye, J., Lin, L., Lian, M., & Chen, J. (2024). Regulation of tetramethylpyrazine formation by the phenolics-Fenton coupled redox cycling system. *Journal of Agricultural and Food Chemistry*, 72(21), 12251–12259. <https://doi.org/10.1021/acs.jafc.4c02576>
- Xu, J. (2024). Waste to resource: Mining antimicrobial peptides in sludge from metagenomes using machine learning. *Environment, International*, 186, 108574. <https://doi.org/10.1016/j.envint.2024.108574>
- Xu, Y., Wang, X., Liu, X., Li, X., Zhang, C., Li, W., Sun, X., Wang, W., & Sun, B. (2021). Discovery and development of a novel short-chain fatty acid ester synthetic biocatalyst under aqueous phase from Monascus purpureus isolated from baijiu. *Food Chemistry*, 338, Article 128025. <https://doi.org/10.1016/j.foodchem.2020.128025>
- Yan, C., Huang, Z., Tu, R., Zhang, L., Wu, C., Wang, S., Huang, P., Zeng, Y., & Shi, B. (2025). Revealing the differences in microbial community and quality of high-temperature daqu in the Southern Sichuan–Northern Guizhou Region. *Foods*, 14(4), Article 570. <https://doi.org/10.3390/foods14040570>
- Yang, L., Fan, W., & Xu, Y. (2024a). Effects of storage period and season on the microecological characteristics of Jiangxiang high-temperature Daqu. *Food Research International*, 196, Article 115034. <https://doi.org/10.1016/j.foodres.2024.115034>
- Yang, L., Fan, W., & Xu, Y. (2024b). Qu-omics elucidates the formation and spatio-temporal differentiation mechanism underlying the microecology of high temperature Daqu. *Food Chemistry*, 438, Article 137988. <https://doi.org/10.1016/j.foodchem.2023.137988>
- Yang, S., Duan, J., Lv, S., Xu, L., & Li, H. (2022). Revealing the changes in compounds when producing strong-flavor Daqu by statistical and instrumental analysis. *Fermentation*, 8(12), 720. <https://doi.org/10.3390/fermentation8120720>
- Yang, S.-B., Fu, J.-J., He, J.-H., Zhang, X.-J., Chai, L.-J., Shi, J.-S., Wang, S.-T., Zhang, S.-Y., Shen, C.-H., Lu, Z.-M., & Xu, Z.-H. (2024). Decoding the Qu-aroma of medium-temperature Daqu starter by volatiles, aroma recombination, omission studies and sensory analysis. *Food Chemistry*, 457, Article 140186. <https://doi.org/10.1016/j.foodchem.2024.140186>
- Yang, S.-B., Lu, Z.-M., Shen, H.-J., Fu, J.-J., Yang, Y., Zhang, X.-J., Chai, L.-J., Wang, S.-T., Zhang, S.-Y., Shen, C.-H., Shi, J.-S., & Xu, Z.-H. (2025). Unraveling Qu-aroma variation between inner and outer layers of medium-temperature Daqu: A multi-omics and sensory approach. *International Journal of Food Microbiology*, 442, Article 111392. <https://doi.org/10.1016/j.ijfoodmicro.2025.111392>
- Yao, S., Zhai, L., Xin, C., Liu, Y., Xu, L., Zhang, X., Zhao, T., Zhang, L., & Cheng, C. (2016). Scopulibacillus daqui sp. Nov., a thermophilic bacterium isolated from high temperature daqu. *International Journal of Systematic and Evolutionary Microbiology*, 66(11), 4723–4728. <https://doi.org/10.1099/ijsem.0.001417>
- Yuan, H., Zheng, J., Ding, L., Wang, H., Jiang, Q., Zhang, C., Xie, T., Nan, G., Li, L., & Lou, K. (2025). Temperature-driven divergence in microbial consortia and physicochemical functionality: A comparative study of high- and medium-temperature Daqu. *Microorganisms*, 13(6), 1312. <https://doi.org/10.3390/microorganisms13061312>
- Zeng, Y., Wang, Y., Chen, Q., Xia, X., Liu, Q., Chen, X., Wang, D., & Zhu, B. (2022). Dynamics of microbial community structure and enzyme activities during the solid-state fermentation of Forgoed Daqu: A starter of Chinese strong flavour baijiu. *Archives of Microbiology*, 204(9), 577. <https://doi.org/10.1007/s00203-022-03198-w>
- Zhang, G., Luo, M., Xu, C., Pan, X., Yi, G., Xiao, W., Wang, C., Zhao, Z., & Li, R. (2025). Microbial and functional gene dynamics in long-term fermented mariculture sediment. *Journal of Marine Science and Engineering*, 13(5), 975. <https://doi.org/10.3390/jmse13050975>
- Zhang, H., Tan, Y., Wei, J., Du, H., & Xu, Y. (2022). Fungal interactions strengthen the diversity-functioning relationship of solid-state fermentation systems. *mSystems*, 7 (4), Article e00401–22. <https://doi.org/10.1128/msystems.00401-22>
- Zhang, J., Du, R., Niu, J., Ban, S., Zhang, Y., Xu, L., Nie, H., Wu, Q., & Xu, Y. (2024). Daqu and environmental microbiota regulate fatty acid biosynthesis via driving the core microbiota in soy sauce aroma type liquor fermentation. *International Journal of Food Microbiology*, 408, Article 110423. <https://doi.org/10.1016/j.ijfoodmicro.2023.110423>
- Zhang, Y., Xu, J., Ding, F., Deng, W., Wang, X., Xue, Y., Chen, X., & Han, B.-Z. (2022). Multidimensional profiling indicates the shifts and functionality of wheat-origin

- microbiota during high-temperature Daqu incubation. *Food Research International*, 156, Article 111191. <https://doi.org/10.1016/j.foodres.2022.111191>
- Zhang, Z., Ran, X., Guo, Z., Hou, Q., Qu, D., Wang, C., Xu, Y., & Wang, Y. (2025). Microbial diversity, functional properties, and flavor characteristics of high-temperature Daqu with different colors. *Food Research International*, 212, Article 116406. <https://doi.org/10.1016/j.foodres.2025.116406>
- Zhou, Q., Ma, K., Song, Y., Wang, Z., Fu, Z., Wang, Y., Zhang, X., Cui, M., Tang, N., & Xing, X. (2022). Exploring the diversity of the fungal community in Chinese traditional baijiu daqu starters made at low-, medium- and high-temperatures. *LWT*, 162, Article 113408. <https://doi.org/10.1016/j.lwt.2022.113408>
- Zhu, C., Cheng, Y., Shi, Q., Ge, X., Yang, Y., & Huang, Y. (2023). Metagenomic analyses reveal microbial communities and functional differences between Daqu from seven provinces. *Food Research International*, 172, Article 113076. <https://doi.org/10.1016/j.foodres.2023.113076>
- Zhu, M., Yin, H., Yuan, Y., Liu, H., Qi, X., Ren, Y., & Dang, Z. (2022). Discrepancy strategies of sediment abundant and rare microbial communities in response to floating microplastic disturbances: Study using a microcosmic experiment. *Science of the Total Environment*, 835, Article 155346. <https://doi.org/10.1016/j.scitotenv.2022.155346>
- Zhu, M., Zheng, J., Xie, J., Zhao, D., Qiao, Z.-W., Huang, D., & Luo, H.-B. (2022). Effects of environmental factors on the microbial community changes during medium-high temperature daqu manufacturing. *Food Research International*, 153, Article 110955. <https://doi.org/10.1016/j.foodres.2022.110955>
- Zhu, Q., Chen, L., Peng, Z., Zhang, Q., Huang, W., Yang, F., Du, G., Zhang, J., & Wang, L. (2023). The differences in carbohydrate utilization ability between six rounds of sauce-flavor daqu. *Food Research International*, 163, Article 112184. <https://doi.org/10.1016/j.foodres.2022.112184>