



## Integrated multi-omics uncover viruses, active fermenting microbes and their metabolic profiles in the *Daqu* microbiome

Xiaoning Huang <sup>a,b</sup>, Rengshu Li <sup>a</sup>, Jinguo Xu <sup>c</sup>, Jiamu Kang <sup>d</sup>, Xiaoxue Chen <sup>a</sup>, Beizhong Han <sup>a</sup>, Yansong Xue <sup>a,\*</sup>

<sup>a</sup> Beijing Laboratory for Food Quality and Safety, College of Food Science and Nutritional Engineering, China Agricultural University, Beijing 10083, China

<sup>b</sup> Department of Bioengineering, University of Illinois Urbana-Champaign, Urbana, IL 61801, USA

<sup>c</sup> College of Food Science and Engineering, Yangzhou University, Yangzhou 225127, China

<sup>d</sup> School of Food Science and Engineering, Hainan University, Haikou 570228, China



### ARTICLE INFO

#### Keywords:

*Daqu*  
Temperature  
Virus  
Microbiome  
Metabolites  
Microbial interactions

### ABSTRACT

The coexistence and coevolution of viruses and fermenting microbes have a significant impact on the structure and function of microbial communities. Although the presence of viruses in *Daqu*, the fermentation starter for Chinese *Baijiu*, has been documented, their specific effects on the community composition and metabolic functions of low, medium, and high-temperature *Daqu* remain unclear. In this study, we employed multi-omics technology to explore the distribution of viruses and active bacteria and fungi in various *Daqu* and their potential metabolic roles. Viral metagenomic sequencing showed a predominance of *Parvoviridae* in High-Temperature *Daqu* (HTQ), while *Genomoviridae* were dominant in Medium-Temperature *Daqu* (MTQ) and Low-Temperature *Daqu* (LTQ). Phages belonging to the *Siphoviridae*, *Podoviridae*, *Herelleviridae*, and *Myoviridae* families showed significantly different abundances across three *Daqu* groups. Metatranscriptomic analysis showed that fungal communities were most active in LTQ, whereas bacterial communities were dominant in MTQ and HTQ. By employing the CRISPR-Cas spacer, a higher predicted number of phage-host linkages was identified in LTQ, particularly with hosts including *Lactobacillus*, *Staphylococcus*, *Acinetobacter*, *Enterobacter*, and *Bacillus*. Correlation analysis showed that bacteria like *Acinetobacter*, *Lactobacillus*, and *Streptococcus* exhibited the strongest associations with metabolites, particularly amino acids and organic acids. The potential phage-induced metabolic differences in the three *Daqu* groups were mainly linked to pathways involved in the metabolism of amino acids, sugars, and organic acids. Overall, our study elucidates the impact of viruses on shaping microbial composition and influencing metabolic functions in *Daqu*. These results improve our comprehension of viruses and microbes in *Daqu* microbial communities and provide valuable insights for enhancing quality control in *Daqu* production.

### 1. Introduction

The symbiotic evolution of microorganisms and phages maintains phenotypic and genotypic diversity, modulates microbial interactions, accelerates evolution of bacteria and phages, influences community structure, and shapes community composition (Chevallereau et al., 2022; Fernández et al., 2018). Bacteriophages (known as phages), the predominant virus type infecting microorganisms, are frequently present in natural microbial ecosystems. For example, phages participate in nutrient cycling in hyperthermophilic composting by recycling bacterial biomass through cell lysis and expressing crucial auxiliary metabolic genes, underscoring their significance in organic matter breakdown

(Liao et al., 2023). Moreover, lactococcal phages are plentiful in cheese starter cultures, produced abundantly and continuously, and although phage replication burdens cells metabolically, it confers a selective advantage to the host lactic acid bacteria (Alexeeva et al., 2018).

*Daqu*, a traditional culture starter for *Baijiu* and vinegar, has a long and ancient history that spans millennia. The development and advancement of *Daqu* manufacturing technology are pivotal achievements in the history of science and technology (Kang, Xue, et al., 2022; Zheng & Han, 2016). In *Daqu*, a complex microbial ecosystem flourishes due to intricate synergistic interactions among various microorganisms, enzymes, and substrates, influenced by multiple fermentation factors (Kang et al., 2024; Liu et al., 2018; Pang et al., 2018). Despite

\* Corresponding author.

E-mail address: [yansong.xue@cau.edu.cn](mailto:yansong.xue@cau.edu.cn) (Y. Xue).

advancements in the fermented food industry, *Daqu* remains essential for saccharification and alcohol fermentation processes. Based on the peak fermentation temperature, *Daqu* can be classified into high-temperature (60–70 °C), medium-temperature (50–60 °C), and low-temperature (40–50 °C) *Daqu* (Wu, Du, & Xu, 2023). These various types of *Daqu*, along with specific manufacturing techniques, result in unique *Baijiu* products. Low-temperature *Daqu* (LTQ) is utilized in the production of light-flavor *Baijiu*, while medium-temperature (MTQ) *Daqu* is employed for strong-flavor *Baijiu*, and high-temperature *Daqu* (HTQ) is used for sauce-flavor *Baijiu* production (Kang, Chen, et al., 2022). Light-flavor, strong-flavor and sauce-flavor *Baijiu* constitute the top three dominant categories of *Baijiu* in China. (Zheng & Han, 2016).

To date, a comprehensive analysis of the microbial structure and composition in *Daqu* production has been conducted, encompassing raw materials, fermentation environments, processes, and fortified fermentation practices (Du et al., 2019; Huang, Fan, et al., 2021; Zhang et al., 2022). Moreover, microbial origins, succession patterns, and interactions involving both biotic and abiotic factors were examined (Fan et al., 2020; Li et al., 2022). Metabolically active microbes are recognized as key drivers of metabolic processes during fermentation. The advancement of metatranscriptomic sequencing, which is RNA-based, offers valuable insights into active gene expression within microbial communities (Jiang et al., 2016). The significance of active microbes in the fermentation processes of light-flavor, strong-flavor, and sauce-flavor *Baijiu* has garnered heightened interest. (Hu et al., 2020; Pan et al., 2022; Song et al., 2017). In MTQ, transcript levels were dominated by Ascomycota, with the synergistic interactions between *Aspergillus* and other microbes driving the metabolism of ester synthesis (Wang, Quan, et al., 2023). The complex microbial interactions within the community have a significant impact on both the microbial composition and overall functionality (Banerjee et al., 2018). Furthermore, fermentation temperature, a key abiotic factor, plays a crucial role in shaping the structure and functionality of the microbial community in *Daqu* (Cai et al., 2021; Fu et al., 2021). The heat produced by microorganisms internally has been recognized as a key factor influencing the microbial community and metabolic profiles in *Daqu* (Xiao et al., 2017).

Viruses have been identified or isolated in different fermented foods, including fermented vegetables (e.g., kimchi), dairy products (cheese, yogurt), soy-based foods (Wu, Li, et al., 2023), and *Daqu* (Du et al., 2023; Kang, Chen, et al., 2022). However, their symbiotic interactions with bacteria and their impact on the structure and function of microbial communities are still not well understood. Given the crucial role of temperature in influencing virus survival and determining their distribution, activity, stability, and evolution (Jansson & Wu, 2023; Jończyk et al., 2011), it is essential to investigate how fermentation temperature affects the diversity and distribution of virus communities in *Daqu*, understand the regulatory function of viruses in the diversity of both host and non-host microorganisms in *Daqu*, and uncover their impact on metabolic profiles.

Due to the ongoing progress in omics technologies, viral omics analysis has become essential for virus research, offering a comprehensive understanding of the genetic information of all virus present (Zhang, Zhang, et al., 2023). Utilizing diverse omics technologies such as metagenomics, transcriptomics, and metabolomics enables the identification of viral community distribution patterns in distinct *Daqu* samples and the investigation of potential metabolic variations due to virus-induced responses in microbial communities. These endeavors are vital for understanding the mechanisms behind solid-state fermentation in *Baijiu* production and offer insights for enhancing the quality control of *Baijiu*.

## 2. Materials and methods

### 2.1. Sample collection

In November 2021, three types of *Daqu* samples were collected from

a prominent manufacturing facility specializing in *Daqu* located in Jinan City, Shandong Province, China. All samples were classified into three groups based on the maximum incubation temperature: LTQ (40 to 50 °C), MTQ (50 to 60 °C), and HTQ (60 to 70 °C). LTQ samples were prepared from barley and pea, while both MTQ and HTQ samples were made using wheat as the raw material according to their specific incubation and storage protocols. The matured *Daqu* were stored in well-ventilated storerooms for sample collection. *Daqu* samples were randomly selected from the top, middle, and bottom layers of stacked bricks in each storeroom and thoroughly mixed into a single sample after crushing. Three samples were collected from three individual storerooms for LTQ, MTQ, and HTQ, respectively. Subsequently, each sample was divided into two subsamples: one stored at –80 °C for viral metagenomics, metatranscriptomics, and metabolomic analysis; the other stored at 4 °C for virus observation and quantification.

### 2.2. Observation and quantification of virus in *Daqu*

To observe viruses in different *Daqu* samples, a confocal laser scanning microscope (CLSM) (Zeiss LSM 980 with Elyra 7, Jena, Germany) was employed to visualize virus particles stained with SYBR Gold fluorescent dye (Huang, Yu, et al., 2021). The visualization of virus particles was obtained using a transmission electron microscope (Jem-2100F, Jeol, Japan) with the phosphotungstic acid counterstaining method (Hara et al., 1991).

### 2.3. Viral DNA extraction and shotgun viral metagenome sequencing

A previously established protocols in our laboratory was used to collect the viral particle (Kang et al., 2022a). Briefly, *Daqu* sample (2 g) was suspended or homogenized in precooled sterile SB buffer (0.2 M NaCl, 5 mM CaCl<sub>2</sub>, 50 mM Tris-HCl, 5 mM MgCl<sub>2</sub>, pH 7.5) to concentrate virus-like particles from *Daqu* samples. After at least three cycles of freezing and thawing, each mixture was centrifuged at 12,000 ×g for 5 min to remove precipitates. Microbial cells were filtered through a 0.22 µm filter, and the supernatant was then transferred to an ultracentrifuge tube containing 28 % sucrose and centrifuged at 160, 000 ×g for 2 h (Hitachi, Tokyo, Japan). The precipitate was then resuspended in 200 µL of SB buffer. The obtained virus pellets were treated with DNase I and RNase A (Tiangen Biotech, Beijing, China) to remove free DNA and RNA at 37 °C for 1 h. The suspension was then incubated at 70 °C for 10 min to inactivate the enzyme, followed by centrifugation at 2000 ×g for 5 min. The supernatant was stored at –20 °C for further studies on viral genomes.

The viral DNA were extracted and sequenced at MAGIGENE Biotech Co., Ltd. in Guangzhou, China (<http://www.magigene.com/>). Specifically, total viral nucleic acids were extracted using the MagPure Viral DNA/RNA Mini LQ Kit (R6662-02; Magen, Guangzhou, China) following the manufacturer's protocol. The libraries were pooled and subjected to 150 bp paired-end sequencing using the Novaseq 6000 platform (Illumina, San Diego, CA, USA). For bioinformatics analysis, SOAPnuke (version 1.5.6) was employed to eliminate adapter sequences and eliminate low-quality sequences. The resultant clean reads were subjected to de novo assembly using MEGAHIT version 1.1.2 with default parameters. Subsequently, the assembled contigs underwent mapping against an in-house virus reference database from the GenBank non-redundant nucleotide (NT) database (Zheng et al., 2020), for virus read identification through BLASTx version 2.9.0+. The criteria for identification included alignment similarity ≥80 %, a length of the matched area ≥ 500 bp, and an *E*-value ≤10<sup>–5</sup>. Blast+ with default LCA algorithm parameters was employed for this purpose, and low-quality aligned reads with coverage less than 5 were filtered to enhance the accuracy of taxonomic results. Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) were used to predict the detected viruses to potential host microbes (Huang et al., 2021a).

## 2.4. Total RNA extraction and metatranscriptomic sequencing

The total RNA extraction was performed using a RNeasy Mini Kit (Qiagen, Germany) following the manufacturer's protocol. The RNase-Free DNase Set (Qiagen, Germany) was used to remove DNA. The Ribo-Zero rRNA Removal Kit (Epicentre, an Illumina® company) was then employed to purify the RNA samples by removing ribosomal RNA (rRNA). After RNA quality was checked with NanoDrop 2000 spectrophotometer (Thermo Scientific, Waltham, MA), cDNA synthesis and transcriptome library preparation were carried out using the TruSeqTM RNA sample preparation kit from Illumina (San Diego, CA, USA). The sequencing was conducted on the Illumina HiSeq platform by Majorbio Bio-Pharm Technology Co., Ltd. (Shanghai, China), with bridge PCR using the HiSeq 3000/4000 PE Cluster Kit.

The raw sequencing data underwent assembly following quality control procedures (Li et al., 2009). Subsequently, a non-redundant gene set was derived through the prediction of open reading frames (ORFs) and the elimination of redundancies. Gene abundance in each sample was then calculated and normalized as FPKM (Yang & Kim, 2015). RNA transcripts with significantly different FPKM counts according to edgeR analysis (Robinson et al., 2010). For taxonomic annotation and functional annotation, comparisons were made against the NCBI nr database and KEGG, respectively. Specifically, functional genes, including enzyme-encoding genes linked to *Daqu* fermentation, were annotated. Additionally, the relative abundance of related microorganisms was calculated (Huang et al., 2020).

## 2.5. Metabolite extraction and UHPLC-QE-MS analysis

For metabolite extraction from *Daqu*, the following specific procedures were employed: 50 mg *Daqu* sample was combined with 1.0 mL of extraction solution (methanol: water = 3:1, v/v, containing an isotope-labeled internal standard mixture). The mixture underwent grinding at 35 Hz for 4 min, followed by 5 min of ultrasonication in an ice-water bath three times. After cooling at a low temperature for 1 h, the mixture underwent centrifugation at 12,000 × g for 15 min at 4 °C, and the resulting supernatant was collected for metabolite detection.

Untargeted metabolomic profiling of *Daqu* was conducted using UHPLC-QE-MS. A Vanquish UHPLC system (Thermo Fisher Scientific) equipped with a Waters ACQUITY UPLC HSS T3 column (2.1 mm × 100 mm, 1.8 µm) was employed (Zhang, Shen, et al., 2023). The mobile phase comprised water with 5 mmol/L ammonium acetate and 5 mmol/L acetic acid, along with acetonitrile. The sample tray temperature was maintained at 4 °C, and the injection volume was set at 3 µL. Detection was executed using a Thermo Q Exactive HF-X mass spectrometer controlled by Xcalibur software (version: 4.0.27, Thermo). Both first and second-level mass spectrometry data were collected. The parameters were configured as follows: Sheath gas flow rate: 30 Arb; Aux gas flow rate: 10 Arb; Capillary temperature: 350 °C; Full ms resolution: 60,000; MS/MS resolution: 7500; Collision energy: 10/30/60 in NCE mode; Spray Voltage: 4.0 kV (positive) or -3.8 kV (negative).

The raw data was converted to mzXML format using ProteoWizard. Substances were annotated by matching with a self-built second-level mass spectrometry database. Individual peaks were filtered, retaining only those with missing values in the QC samples <50 % and < 80 % in the actual samples. Data standardization was performed based on the total ion flow of each sample.

## 2.6. Statistical analysis

The identification of biomarkers of viruses and microbes was conducted through Linear Discriminant Analysis Effect Size (LEfSe). A logarithmic LDA score cutoff of ≥3 was applied to detect biomarkers. Multivariate statistical techniques such as Principal Component Analysis (PCA) and Orthogonal Partial Least Square-Discriminant Analysis (OPLS-DA) were used to assess the discriminative ability in classifying

*Daqu* samples. Correlations between active bacteria and phages were determined via Mantel tests using R software. Differential metabolites were subjected to single-variable statistical analysis via Student's t-test. Variable Importance in Projection (VIP) values >1 served as the criterion for selecting distinctive metabolites. Permutation tests, consisting of 200 iterations, were applied to assess the OPLS-DA model parameters R<sup>2</sup> and Q<sup>2</sup>. Metabolites with VIP > 1 and P < 0.05 were considered significantly altered and annotated using KEGG databases through MetaboAnalys (<http://www.metaboanalyst.ca/>). Spearman correlation analysis was conducted using Origin 2023, with correlation coefficients (|r| > 0.7, P < 0.05) considered strong correlation. Network visualization was achieved using Cytoscape v3.7.0.

## 3. Results

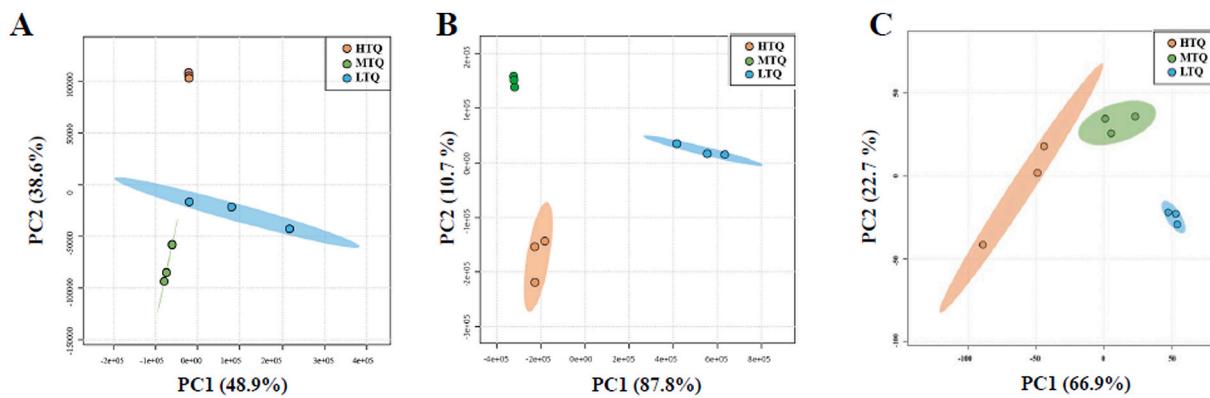
### 3.1. Multi-dimensional differences in the virus, active fermenting microbe, and metabolite profiles of *Daqu*

To explore the intrinsic characteristics of HTQ, MTQ, and LTQ, we initially examined whether significant differences existed among the three groups based on viruses, active fermenting microbes, and metabolites. PCA analysis revealed distinct patterns: HTQ demonstrated minimal variability in viral composition, while LTQ exhibited greater variability (Fig. 1A). In terms of active microbial composition, MTQ exhibited close clustering while for metabolites, LTQ showed close clustering compared to MTQ and HTQ (Fig. 1B and C). Notably, the three *Daqu* groups were distinctly separated based on viral composition, active fermenting microbes, and metabolites. The distinct separation of the three *Daqu* groups highlights the significant impact of incubation temperature on the characteristics of *Daqu*.

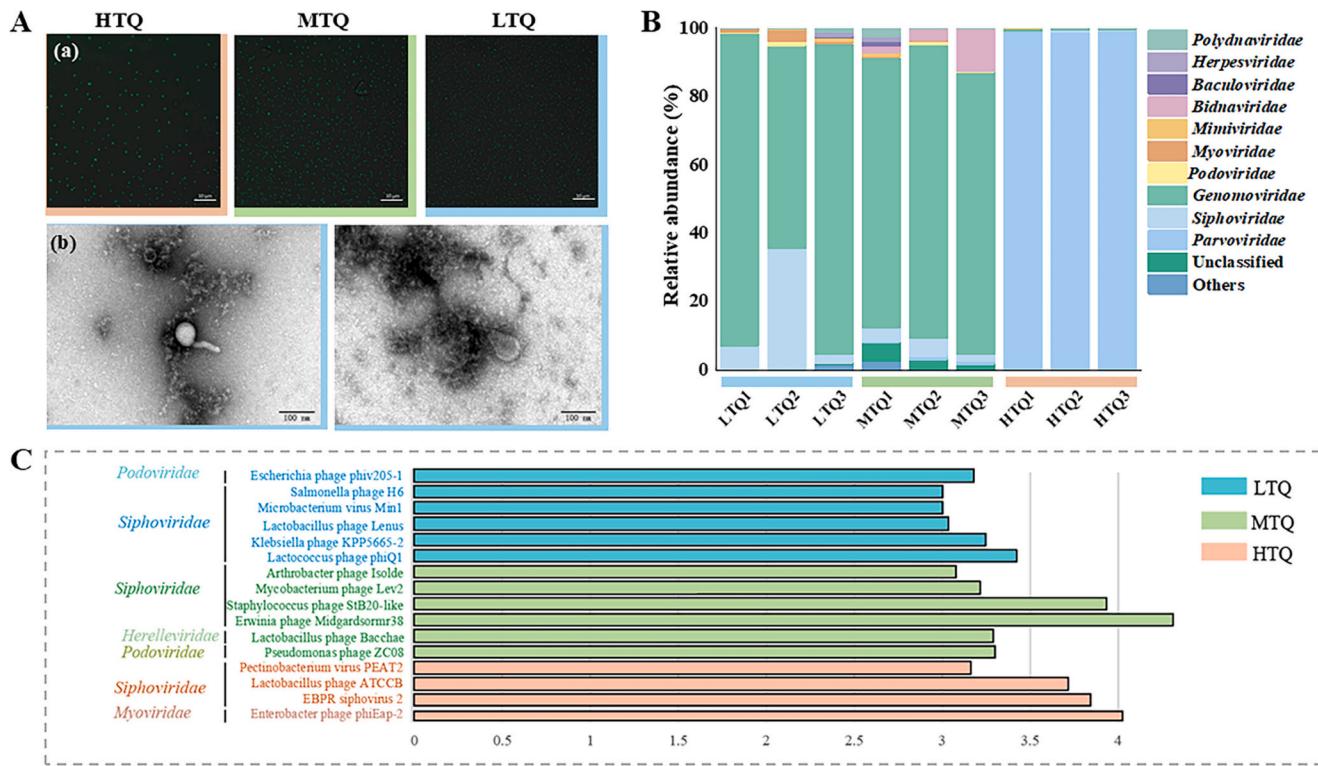
### 3.2. Composition and differentiation of viral community in *Daqu*

To describe the virome composition across different *Daqu* groups, we assessed the abundance of virus particles that existed in three *Daqu* samples using laser confocal microscopy and observed the morphology of viruses in LTQ using transmission electron microscopy. As illustrated in Fig. 2A, the quantity of virus particles was highest in LTQ, followed by MTQ, while HTQ exhibiting the fewest particles, indicating a decrease in virus particle abundance with increasing incubation temperature in different *Daqu*.

Furthermore, we examined the composition and structures of the viral communities in three *Daqu* groups. The statistics of virome sequencing are shown in Table S1. The predominant virus families are depicted in Fig. 2B, with the *Parvoviridae* family comprising the largest proportion in HTQ, averaging 98.98 %. Conversely, *Genomoviridae* was the dominant family in MTQ and LTQ, averaging 84.47 % and 80.55 %, respectively. *Genomoviridae* exhibited considerable variability among different LTQ samples, ranging from 59.38 % to 91.40 %, while *Siphoviridae* accounted for an average of 14.83 % across all LTQ samples. Additionally, *Bidnaviridae*, *Myoviridae*, *Podoviridae*, *Herpesviridae*, *Baculoviridae*, *Mimiviridae*, and *Polydnaviridae* were detected across all *Daqu* samples. Furthermore, viral sequences were annotated using the KEGG database (Fig. S1). At levels 2 and 3, pathways related to cell maintenance and proliferation, such as genetic information processing, signaling, cellular processes, DNA replication and repair exhibited the higher enrichment within the viral community. Additionally, nucleotide, amino acid, lipid, cofactors and vitamins metabolism were also involved in the viral community. The differential viral taxa among different *Daqu* were further compared, LEfSe analysis identified 16 species which exhibited significantly different abundances among the three *Daqu* groups (Fig. 2C). These species were affiliated with four families, with species from *Siphoviridae* being the most abundant in all three *Daqu* groups. Other species belonged to *Podoviridae*, *Herelleviridae*, and *Myoviridae*. *Lactococcus* phage was notably enriched in LTQ, while *Erwinia* phage and *Enterobacter* phage were highly enriched in HTQ. All



**Fig. 1.** Principal Component Analysis of Three Different *Daqu* Samples. (A) Principal component analysis based on virus compositions. PC1 (Principal Component 1) accounted for 48.9 % of the total variability observed, while PC2 (Principal Component 2) explained 38.6 % of the total variability. (B) Principal component analysis based on active microbes composition. PC1 explained a substantial 87.8 % of the total variability, and PC2 contributed 10.7 % to the overall variability. (C) Principal component analysis based on metabolites composition. PC1 elucidated 66.9 % of the total variability, and PC2 explained 22.7 % of the total variability. Each point in the figure represents a sample, with different colors indicating distinct groups.



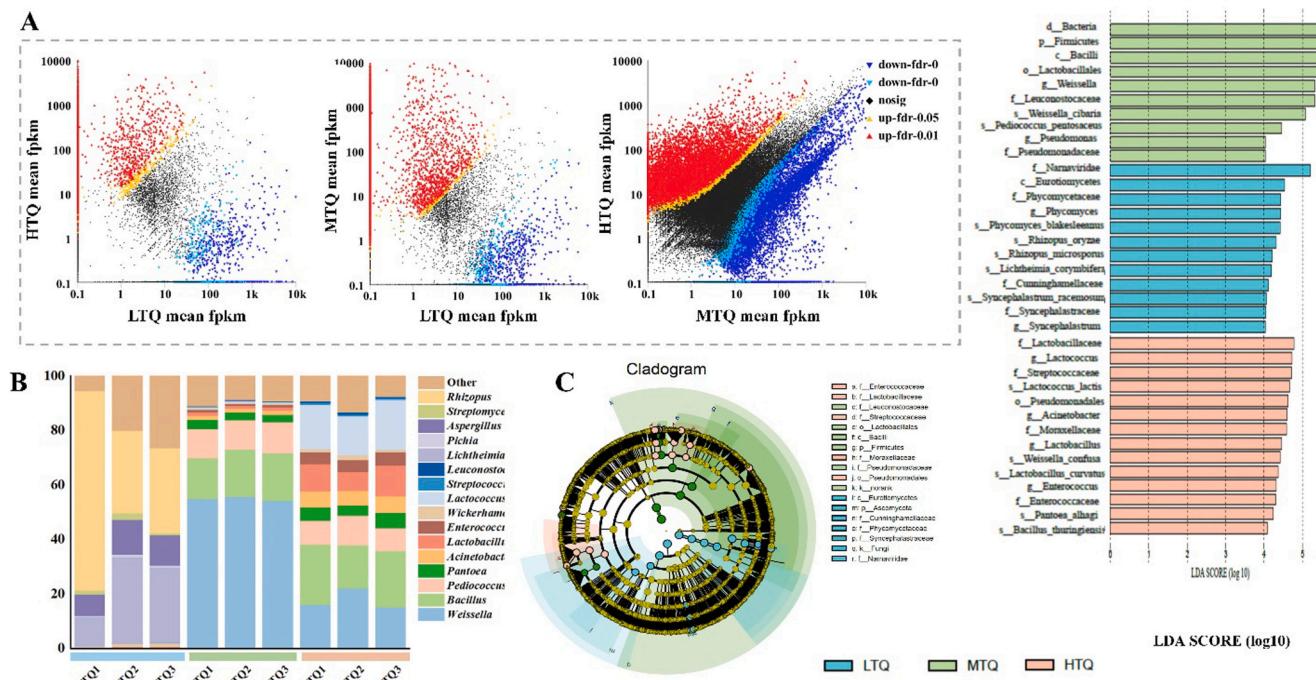
**Fig. 2.** Viral Community Profiles in Different *Daqu* Samples. (A) Confocal Laser Scanning Microscopy (CLSM) (a) and Scanning Electron Microscopy (SEM) (b) images showing viruses in *Daqu*. The number of virus particles was lowest in HTQ, intermediate in MTQ, and highest in LTQ. SEM images illustrate the morphology of virus particles in LTQ. (B) Viral diversity in different *Daqu* samples, depicted at the family level (only the family that occurred at a frequency exceeding 1 % in at least one sample were shown). (C) LEfSe (Linear discriminant analysis Effect Size) analysis highlighting shared phage taxa with statistically significant differences among the samples.

in all, several phage species were identified as lactic acid bacteria-associated phage, including one *Lactococcus* phage, and three *Lactobacillus* phages, highlighting the linkage between lactic acid bacteria and viruses in *Daqu* microbial community.

### 3.3. Composition and differentiation of active microbial community in *Daqu*

A total of 145.13 Gbp of raw data and 128.77 Gbp of clean data were generated through metatranscriptomic sequencing. The Q20 value, exceeding 98 % for all samples, indicates high sequence accuracy

(Table S2) (Pan et al., 2022). To compare the expression and distribution of differentially expressed genes between two *Daqu* groups, all genes expression levels are visualized (Fig. 3A). In the comparison between HTQ and LTQ, 4380 gene expressions were downregulated and 70,668 were upregulated. Similarly, in the comparison between MTQ and LTQ, 4380 gene expressions were downregulated and 66,037 were upregulated. For the differentiation between HTQ and MTQ, 34573 gene expressions were downregulated and 49,417 were upregulated. These findings illustrate significant variations in gene expression among different *Daqu* groups. To compare the active microbial compositions of different *Daqu* groups, taxonomic composition was annotated for



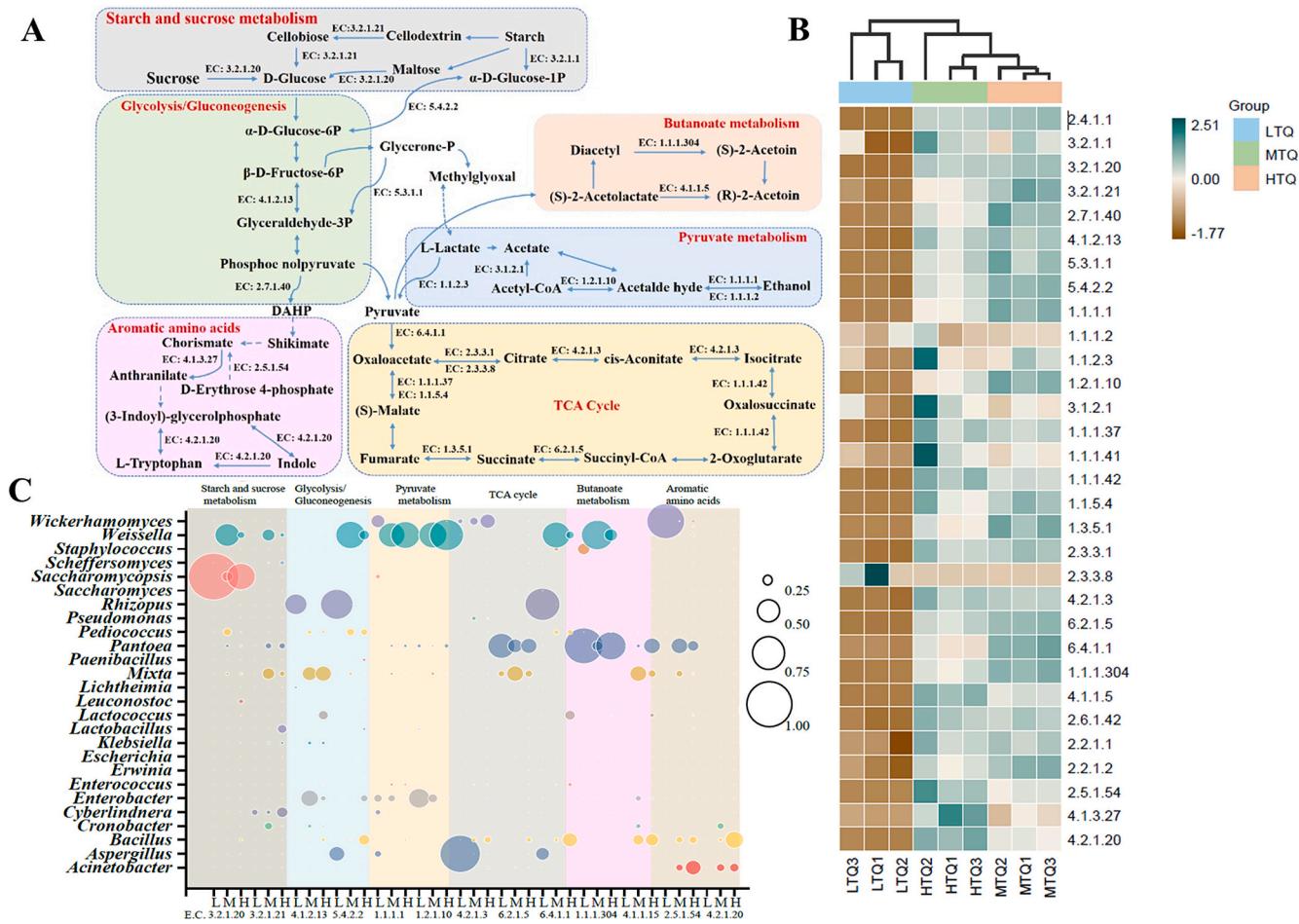
**Fig. 3.** Active Microbial Community Profiles in Different *Daqu* Samples (A) Scatter Plot of gene expression levels ( $\log_2$  FPKM) detected in different *Daqu* samples. The horizontal and vertical axes respectively indicate the average gene expression levels (FPKM values) of the two *Daqu* samples. Each dot represents a specific gene, where “down-fdr-0.01” indicates significantly downregulated genes with a  $P < 0.01$  (dark blue), “down-fdr-0.05” represents significantly downregulated genes with  $P$ -values between 0.01 and 0.05 (light blue), “nosig” indicates non-significant differentially expressed genes (black), “up-fdr-0.05” denotes significantly upregulated genes with  $P$ -values between 0.01 and 0.05 (yellow), and “up-fdr-0.01” signifies significantly upregulated genes with a  $P < 0.01$  (red) (B) Microbial composition and structure estimated through meta-transcriptome analysis (only the genera that occurred at a frequency exceeding 1 % in at least one sample were shown). (C) LEfSe analysis revealing shared microbial taxa with significant differences among the samples. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

expressed transcripts. In total, 58 phyla and 787 genera were annotated. The dominant microbes at the genus level are shown in Fig. 3B. In LTQ, *Rhizopus*, *Lichtheimia*, and *Aspergillus* were predominant with abundances of 44.81 %, 23.38 %, and 17.42 %, respectively. *Streptomyces* and *Lactobacillus* were the dominant bacteria in LTQ with abundances of 1.58 % and 0.26 %, respectively. In both MTQ and HTQ, bacterial communities were more dominant than fungal communities. In MTQ, *Weissella*, *Bacillus*, and *Pediococcus* were the top three microbes with abundances of 54.88 %, 16.55 %, and 10.94 %, respectively. In HTQ, *Bacillus* (19.57 %), *Weissella* (17.70 %), *Lactococcus* (19.25 %), and *Lactobacillus* (9.5 %) were prevalent. These results indicate that fungal communities were most influential in LTQ, while bacterial communities were more active in MTQ and HTQ. To further analyze the differences in microbial composition among the three *Daqu* groups, we conducted LEfSe analysis to identify significant differences in microbial composition. LEfSe bars displayed the discrimination of different microbes in the various *Daqu* groups ( $LDA > 4$ ). A total of 36 differential taxa were specifically identified. HTQ exhibited the highest number of distinct taxa, primarily comprising active bacteria, including *Bacillus*, *Lactococcus*, *Lactobacillus*, and *Streptococcus*. Fungi such as *Phycomyces*, *Rhizopus*, and *Lichtheimia* were significantly enriched in LTQ. Meanwhile, *Weissella* and *Pediococcus* were significantly abundant in MTQ (Fig. 3C).

After elucidating the microbial composition among distinct *Daqu* groups, our subsequent objective was to identify functional distinctions among them. To achieve this, COG and KEGG annotations were conducted to delineate the active genes and associated pathways expressed across different *Daqu* groups. The composition and abundance of the top 50 COG and KEGG function classifications were depicted. Notably, MTQ and HTQ exhibited higher abundances based on annotations compared to LTQ. Specifically, COG annotations revealed that translation,

ribosomal structure and biogenesis, along with carbohydrate transport and metabolism, were the most prevalent categories (Fig. S2A). Similarly, KEGG analysis indicated enrichments in pathways related to glycolysis/gluconeogenesis, biosynthesis of amino acids, ribosome, and carbon metabolism across the various *Daqu* groups (Fig. S2B).

Metabolic pathways, characterized by a sequence of enzyme-catalyzed reactions, play a crucial role in *Daqu* production and *Baijiu* fermentation. Through COG and KEGG annotations, along with our previous work, key pathways including starch and sucrose metabolism, glycolysis/gluconeogenesis, aromatic amino acid metabolism, pyruvate metabolism, butanoate metabolism, and the tricarboxylic acid (TCA) cycle were identified as central process (Huang et al., 2020; Zhang, Shen, et al., 2023). Within each pathway, pivotal enzymes and their respective abundances were illustrated (Fig. 4A and B). Notably, the majority of enzymes exhibited higher abundances in MTQ and HTQ compared to LTQ, with the exception of E.C. 2.3.3.8 (ATP citrate synthase). Furthermore, the top abundant enzymes in each metabolic category, along with the microbial contributors to these enzymes, were depicted in Fig. 4C. For instance, *Saccharomyces* was found to be the predominant contributor, accounting for 99.68 % for E.C. 3.2.1.20 (alpha-glucosidase) in LTQ, whereas *Weissella* contributed 48.37 % and 14.74 % for the enzyme in MTQ and HTQ, respectively. In the glycolysis/gluconeogenesis category, *Rhizopus* was the principal contributor to enzymes E.C. 4.1.2.13 (fructose-bisphosphate aldolase) and E.C. 5.4.2.2 (phosphoglucomutase) in LTQ, while *Bacillus*, *Pediococcus*, and *Weissella* were associated with E.C. 5.4.2.2 in HTQ. Regarding pyruvate metabolism, *Aspergillus* emerged as the predominant producer of alcohol dehydrogenase (EC 1.1.1.1) in LTQ, whereas *Weissella* dominated in MTQ and HTQ. In amino acid metabolism, *Wickerhamomyces* predominantly contributed to E.C. 2.5.1.54 in LTQ, while *Acinetobacter* and *Bacillus* were primary contributors in MTQ and HTQ. These findings



**Fig. 4.** Metabolic Pathway Profile Based on Meta-Transcriptome Analysis. (A) Schematic overview of the metabolic processes in *Daqu*, encompassing six main sections: starch and sucrose metabolism, glycolysis/gluconeogenesis, butanoate metabolism, pyruvate metabolism, aromatic amino acids, and the TCA cycle. (B) Abundance of key enzyme-encoding genes annotated using KEGG. The heatmap is scaled by relative abundances for each row, ranging from low to high relative abundance. (C) Contribution rate of microorganisms to the enzyme-encoding genes annotated by KEGG in different *Daqu* samples.

underscore the varied microbial contributions to specific functions across different *Daqu* samples, highlighting the complexity of microbial roles in the fermentation processes.

#### 3.4. Predicted phage-host linkages in different *Daqu* groups

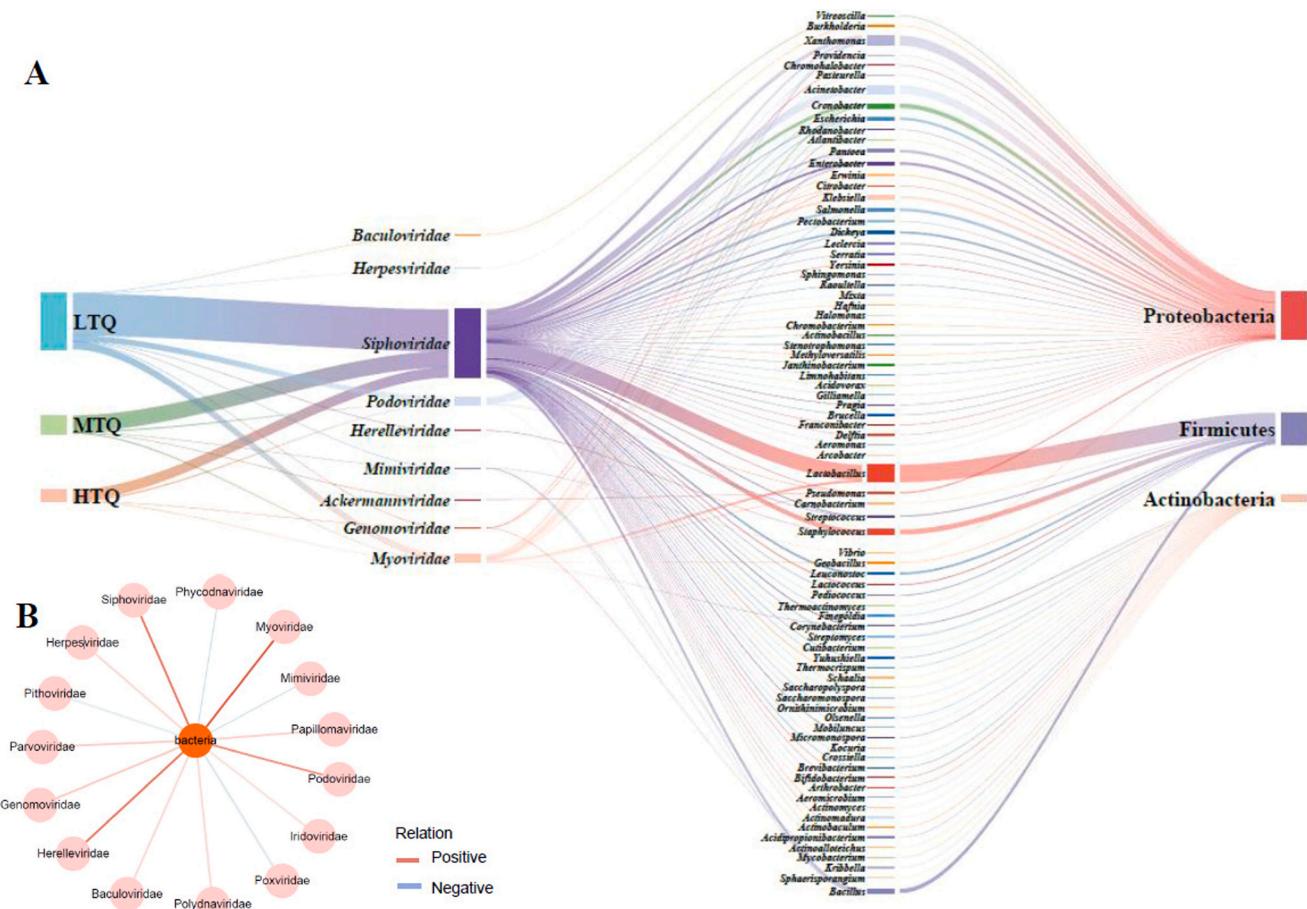
Mantel tests were conducted to identify correlation between phage and bacterial community. *Siphoviridae*, *Herelleviridae*, *Podoviridae* and *Myoviridae* showed a relatively strong positive correlation with bacterial community of *Daqu* (Fig. 5B). Furthermore, the host spectra of phages were estimated through matching the CRISPR-Cas spacer database which is currently the main bioinformatic method for predicting virus-host linkages (Fig. 5A). A total of 1613 matches were identified between phages and their potential host bacteria. Across the three *Daqu* groups, 573 contigs assigned to 9 phage families were matched to 3 host phyla. There were 350 contigs from LTQ, 113 contigs from MTQ and 74 contigs from HTQ. Notably, the predicted number of phages and their hosts in the LTQ was over three times that of the MTQ and HTQ. *Siphoviridae* was the most prevalent phage originated from different *Daqu* groups. In LTQ, besides *Siphoviridae*, *Myoviridae* and *Podoviridae* were also predominant. Specifically, 300 contigs were linked to Proteobacteria, 198 contigs were linked to Firmicutes, and 42 contigs were linked to Actinobacteria. At the genus level, 80 bacterial genera were matched as potential hosts of the phages existed in different *Daqu*. Among them, *Siphoviridae*, *Myoviridae*, and *Podoviridae* demonstrated broad host ranges, while *Baculoviridae* and *Ackermannviridae* exhibited

specificity towards contigs associated with a single bacterial genus. Within *Daqu* samples, *Siphoviridae* (427 contigs) exhibited the highest number of predicted hosts, predominantly infecting bacterial genera such as *Lactobacillus*, *Staphylococcus*, *Acinetobacter*, *Enterobacter*, and *Bacillus*. Additionally, *Lactobacillus* was identified as a potential host for *Myoviridae*. *Myoviridae* also displayed infectivity towards *Xanthomonas*, *Cronobacter*, *Pantoea*, *Escherichia*, and *Acinetobacter*. Meanwhile, *Podoviridae* was prevalent annotated across three different *Daqu* groups, and the predicted hosts of *Podoviridae* included *Xanthomonas*, *Escherichia*, *Enterobacter*, *Streptococcus*, and *Bacillus*.

#### 3.5. Composition and differentiation of metabolites in *Daqu*

UHPLC-QE-MS coupled with multivariate statistical analysis was used to investigate differences in metabolites among three types of *Daqu*. A total of 1347 metabolites were detected across all samples, and Table S3 provides a list of those with relative content greater than 0.1. The Venn diagram shows that 451 metabolites were common to all *Daqu* samples, while LTQ had the highest number of unique metabolites, totaling 161 (Fig. S3A). Hierarchical cluster analysis (HCA) identified three sub-clusters, with each representing data from the three *Daqu* samples, indicating significant differences in metabolite contents among the groups (Fig. S3B). Next, differential metabolites among the three *Daqu* groups were identified based on VIP > 1, P < 0.05, and absolute log<sub>2</sub> fold changes (log<sub>2</sub> FC) > 1. Specifically, we found 84 differential metabolites between HTQ and LTQ, 57 differential metabolites between

A



**Fig. 5.** Correlation between Phage and Bacteria in Different *Daqu* Samples. (A) Predicted Phage–Host Linkages in Different *Daqu* Samples. From left to right, the first columns display three different *Daqu* samples. The second column illustrates viral contigs linking to hosts identified in *Daqu* samples. The third column depicts the distribution of bacterial hosts assigned using CRISPR approaches at the genus level. The last column displays the distribution of bacterial hosts at the phylum level. The height of each column section represents the abundance of viruses and hosts. (B) Association between phage and bacterial community by Mantel tests.

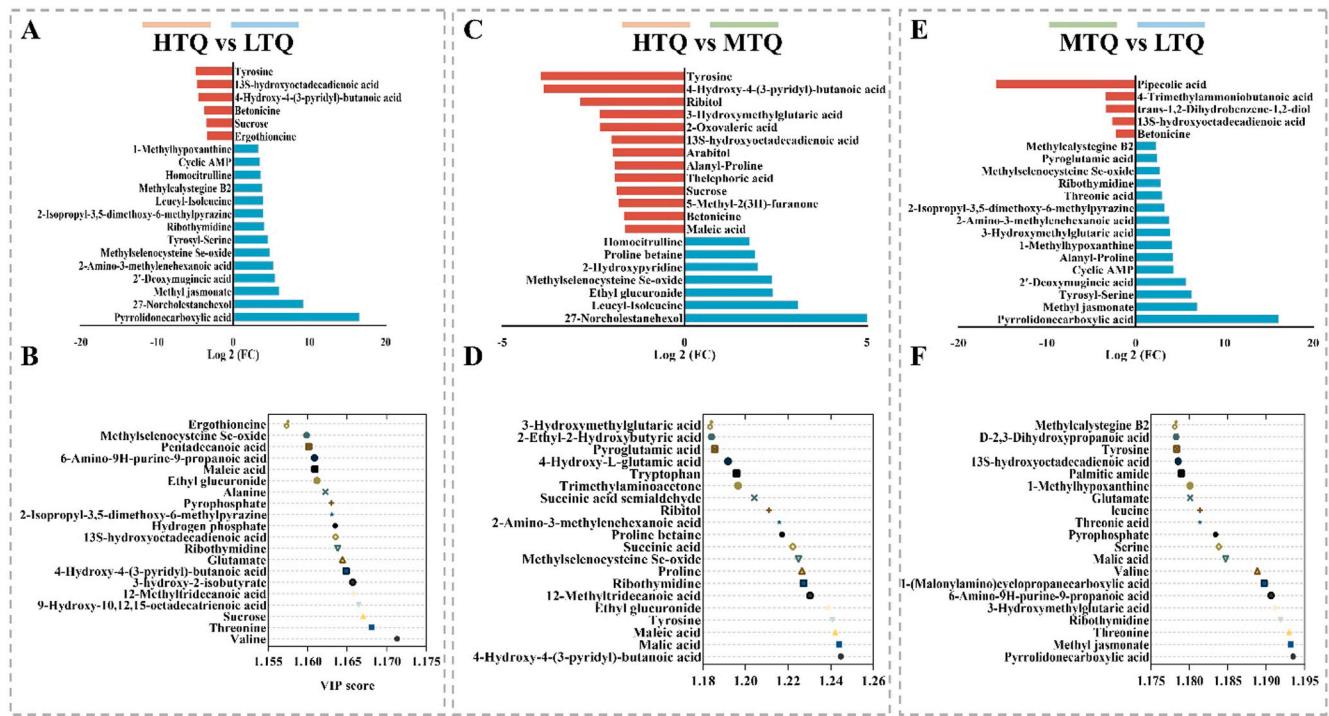
HTQ and MTQ, and 67 differential metabolites between MTQ and LTQ (Table S4). The top 20 differential metabolites with  $\log_2 \text{FC} > 1$  and  $P < 0.05$  are shown in Fig. 6A, C, and E, while the top 20 differential metabolites with  $\text{VIP} > 1$  are shown in Fig. 6B, D, and F. A variety of organic acids and amino acids were key differentiators among the different *Daqu* groups. For instance, tyrosine valine, alanine, and threonine (Fig. 6A and B); 4-hydroxy-4-(3-pyridyl)-butanoic acid, malic acid, succinic acid and proline (Fig. 6C and D); pipecolic acid, pyrrolidonecarboxylic acid, leucine, and L-threonine (Fig. 6E and F) were representative differential metabolites in HTQ vs. LTQ, HTQ vs. MTQ, and MTQ vs. LTQ, respectively. 13S-hydroxyoctadecadienoic acid, 3-aminoacproic acid, alanyl-proline, sucrose, tyrosyl-serine were identified as differential metabolites across the three *Daqu* groups in pairwise comparison. Subsequently, according to KEGG enrichment analysis based on the differential metabolite, we found that amino acid metabolism and sugar metabolism were significantly enriched in all groups (Fig. S4).

### 3.6. Correlations between core active microbes and metabolites

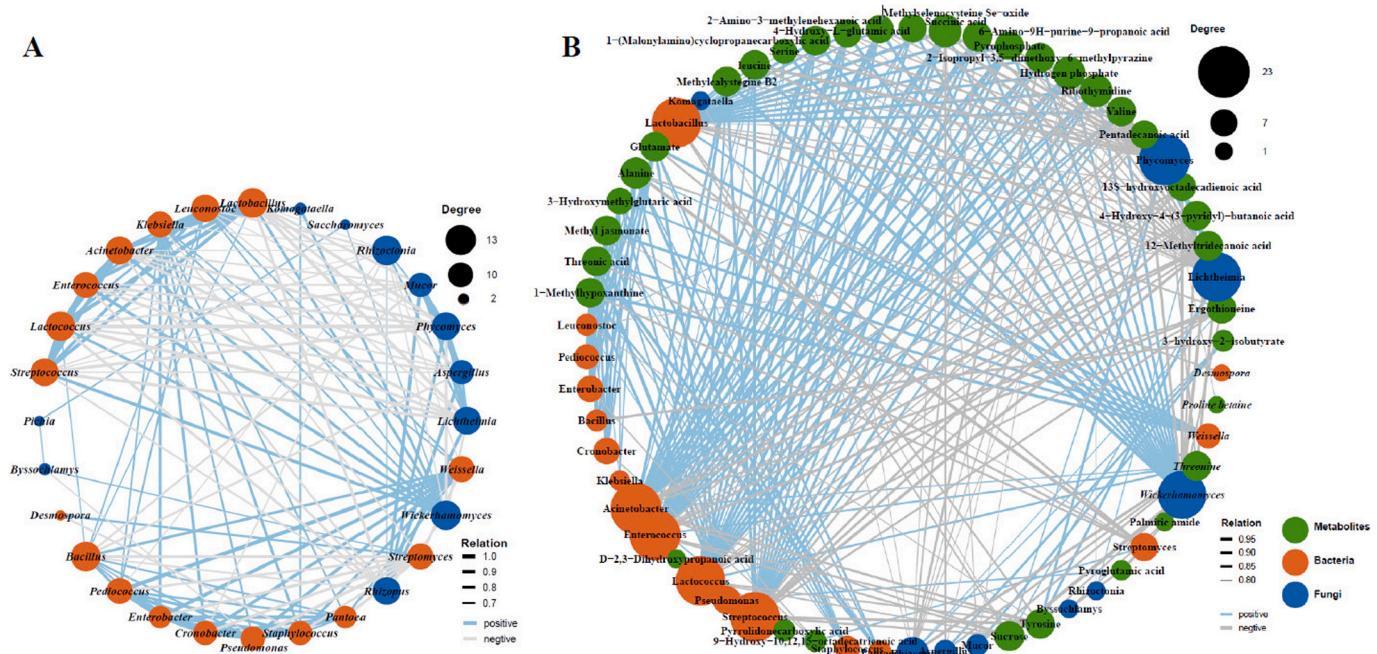
Spearman correlation analysis was conducted to evaluate microbial interaction and relationships between microbes and metabolites. For microbial interaction, a total of 266 pairwise correlations were observed ( $|r| > 0.7$ ,  $P < 0.05$ ), comprising 158 positive and 108 negative correlations (Fig. 7A). Bacteria showed predominantly positive correlations with each other, but exhibited negative correlations with fungi, especially molds. Conversely, fungi displayed positive correlations among themselves but negative correlations with bacteria. Notably,

*Lactobacillus* exhibited positive associations with several bacteria including *Bacillus*, *Streptococcus*, *Acinetobacter*, and *Enterococcus*, while showing negative associations with fungi such as *Rhizopus*, *Lichtheimia*, and *Aspergillus*, except for *Wickerhamomyces*, which showed a positive interaction with *Lactobacillus*. This synergistic interactions between *Lactobacillus* and *Wickerhamomyces* contributed to the synthesis of ester compounds (Wang, Fan, et al., 2020). *Wickerhamomyces*, uniquely among fungi, showed numerous positive correlations with bacteria including *Leuconostoc*, *Acinetobacter*, *Enterococcus*, *Lactococcus*, and *Streptococcus*. Additionally, *Wickerhamomyces* showed positive correlations with *Pichia* and *Saccharomyces*. *Staphylococcus* was positively correlated with *Weissella*, *Pediococcus*, *Enterobacter*, *Pantoea* and *Cronobacter*, but negatively correlated with *Streptomyces*. *Acinetobacter* exhibited negative associations with several fungi including *Rhizopus*, *Lichtheimia*, *Aspergillus*, and *Phycomyces*, while showing positive relationships with bacteria such as *Enterococcus*, *Lactococcus*, and *Streptococcus*. *Bacillus* displayed positive correlations with *Weissella*, *Leuconostoc*, *Pediococcus*, and *Enterobacter*, but exhibited negative correlations with fungi such as *Rhizopus*, *Aspergillus*, and *Mucor*. Additionally, *Lactobacillus*, *Streptococcus*, *Bacillus*, *Acinetobacter*, *Enterobacter*, *Staphylococcus*, *Pantoea*, and *Cronobacter* were identified as host bacteria for phages detected in *Daqu* samples (Fig. 5). Phages directly impacted their host bacteria, which indirectly influenced microbial interactions with other bacteria and fungi.

For relationships between microbes and metabolites, a total of 409 strong pairwise correlations were identified ( $|r| > 0.8$ ,  $P < 0.05$ ), consisting of 236 positive and 173 negative correlations (Fig. 7B). Bacteria



**Fig. 6.** Differential Metabolites Analysis among Different *Daqu* Samples. (A) Fold change analysis of significant differential metabolites and (B) Important metabolites identified by variable importance in projection (VIP) score obtained from OPLS-DA model between HTQ and LTQ. (C) Fold change analysis of significant differential metabolites and (D) Important metabolites identified by variable importance in projection (VIP) score obtained from OPLS-DA model between HTQ and MTQ. (E) Fold change analysis of significant differential metabolites and (F) Important metabolites identified by variable importance in projection (VIP) score obtained from OPLS-DA model between MTQ and LTQ.



**Fig. 7.** Correlation Network between Differential Metabolites and Microbial Communities in *Daqu* Samples. (A) Microbial interactions based on Spearman correlation analysis. (B) Relationships between microorganisms and metabolites. Orange circles represent bacteria, navy circles represent fungi, green circles represent metabolites. Blue lines indicate positive correlations, while the grey lines represent negative correlations. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

such as *Acinetobacter*, *Lactobacillus*, and *Streptococcus* exhibited the highest number of correlations with metabolites, primarily showing positive associations with amino acids and organic acids. For example,

*Lactobacillus* demonstrated positive correlations with valine, threonine, serine, leucine, pentadecanoic acid, succinic acid, 2-amino-3-methylenehexanoic acid, and pyrrolidonecarboxylic acid. Among fungi,

*Lichtheimia*, *Phycomyces*, and *Wickerhamomyces* displayed the most correlations with metabolites. *Lichtheimia* and *Phycomyces* exhibited positive correlations with sucrose and organic acids such as 12-methyltridecanoic acid, 3-hydroxy-2-isobutyrate, and 4-hydroxy-4-(3-pyridyl)-butanoic acid, but negative correlations with amino acids like serine and leucine. *Wickerhamomyces* showed positive correlations with valine and threonine but negative correlations with organic acids including 12-methyltridecanoic acid and succinic acid. The Maillard reactions between amino acid and reducing sugars has been shown to significantly contribute to the different colour. Additionally, the concentration of amino acids influences the formation of flavor compounds, including esters and alcohols (Wang, Wu, et al., 2020; Zhang, Shen, et al., 2023). Phages were detected in various types of *Daqu*, indicating their potential influence on host bacteria (Fig. 5), which can lead to subsequent changes in microbial interactions and metabolite profiles (Fig. 7). Together, these findings highlight the potential impact of phage-mediated differences on microbial populations and metabolic functions across different types of *Daqu*.

#### 4. Discussion

The interaction, co-evolution, and metabolic characteristics of viruses and active fermenting microbes play a crucial role in shaping microbiome functions (Chevallereau et al., 2022; Huang et al., 2021a). *Daqu*, a traditional starter culture, aligns with these aspects, but the composition and mechanisms remain unclear. In this study, a multi-omics method was employed to investigate the viruses, active fermenting microbes, and metabolic patterns of *Daqu* at different temperatures. The analysis revealed that phages like *Parvoviridae*, *Genomoviridae*, and *Siphoviridae* prevail in distinct *Daqu* groups. Moreover, our study demonstrated that phages interacting with their host bacteria could modulate microbial relationships, consequently affecting the composition and metabolic activity of *Daqu* microbial populations. These results enhance comprehension of the phage role in *Daqu*, aiding in the development of customized microbial communities to improve quality control in *Daqu* production processes with precision.

Our findings emphasized the substantial role of fungal communities in LTQ, while bacteria showed higher activity in MTQ and HTQ. Previous studies employing culture-dependent or genomic sequencing methods have characterized the microbial composition of various *Daqu* samples. Particularly in HTQ, *Thermomyces*, *Mucor*, *Saccharomyces*, *Pichia*, *Bacillus*, *Weissella*, *Lactobacillus*, and *Lactococcus* were dominant (Sakandar et al., 2020), consistent with our findings where *Bacillus*, *Weissella*, *Lactococcus*, and *Lactobacillus* prevailed. Previous investigations showed prevalent presence of *Thermoascus*, *Rhizomucor*, *Aspergillus*, *Saccharomycopsis*, *Pichia*, *Lichtheimia*, *Bacillus*, *Weissella*, and *Lactobacillus* in MTQ (Sakandar et al., 2020), consistent with our outcomes except for the prevalence of *Pediococcus*. In LTQ, *Rhizopus*, *Lichtheimia*, *Saccharomycopsis*, *Pichia*, *Wickerhamomyces*, *Weissella*, *Lactobacillus*, and *Bacillus* were dominant in previous studies (Fan et al., 2020; Sakandar et al., 2020), supporting our findings with predominance of *Rhizopus*, *Lichtheimia*, and *Aspergillus*. Traditional culture methods provide isolated colonies but miss up to 90 % of non-culturable microorganisms, leading to bias from microbial preferences for culture conditions. Despite amplicon sequencing being more sensitive, it remains susceptible to primer biases and selectivity. Shotgun metagenomic sequencing eliminates primer-related errors but cannot distinguish between viable and non-viable bacteria within the population. The application of metatranscriptomic sequencing allows for detailed taxonomic and functional characterization of microbial communities, offering insights into their composition and active metabolic functions (Jiang et al., 2016). Regarding metabolic function annotation, most enzymes involved in the primary fermentation process showed higher levels in MTQ and HTQ than in LTQ, indicating a greater contribution from bacteria than fungi.

Phages, viruses that infect bacteria, are widely acknowledged as the

most prevalent biological entities on Earth, coexisting in all bacterial-friendly environments. Phages are thought to have a significant impact on the taxonomic and functional composition of microbial communities, playing a role in shaping their stability. Phage predation is associated with bacterial population evolution, potentially crucial in molding the flavor profiles of fermented foods (Waller et al., 2014; White et al., 2022; Zhang et al., 2023a). In our study, *Parvoviridae* was most prevalent in HTQ, whereas *Genomoviridae* dominated MTQ and LTQ. Our analysis of host-phage connections showed that *Siphoviridae*, *Podoviridae*, and *Myoviridae* from various *Daqu* groups were closely linked with bacteria. *Siphoviridae* exhibited associations with bacteria such as *Lactobacillus*, *Staphylococcus*, *Acinetobacter*, *Enterobacter*, and *Bacillus*. *Podoviridae* was also prevalent predicted across three *Daqu* groups, with predicted hosts including *Xanthomonas*, *Escherichia*, *Enterobacter*, *Streptococcus*, and *Bacillus*. *Myoviridae* displayed infectivity towards *Xanthomonas*, *Cronobacter*, *Pantoea*, *Escherichia*, and *Acinetobacter*. Although *Siphoviridae*, *Podoviridae*, and *Myoviridae* were not the most abundant phages in the *Daqu* microbial community, they contributed significantly to microbe-phage interactions within the *Daqu* microbiome. These phages can directly regulate bacterial populations through a lytic cycle and establish prey-predator relationships with their host bacteria. Moreover, phages can influence host fitness when entering a lysogenic cycle. Additionally, phages encoding bacterial functional genes mediate their transduction between hosts, improving host adaptability through horizontal gene transfer (Wang, Wu, et al., 2023). Apart from controlling host population sizes, phages can trigger significant changes in bacterial transcriptomes and proteomes, affecting metabolism (Fernández et al., 2018). Furthermore, phages can have indirect effects on non-host populations. Phage-induced bacterial lysis offers essential nutrients that can facilitate the growth of other microorganisms. This phenomenon has been observed in Kimchi fermentation, where a phage lysate of *Weissella cibaria* promoted the growth of *Leuconostoc citreum* (Kong & Park, 2019). Alterations in host bacteria population or traits caused by phage infection would modify microbial interactions among phage hosts and other microbes.

Microbial interactions are essential in shaping the structure, stability, and function of microbial communities (Banerjee et al., 2018; Liu et al., 2020). In the *Baijiu* fermentation process, these interactions are considered a key factor that impacts the quality and flavor (Chen et al., 2021; Fan et al., 2020; Sakandar et al., 2020). Our study found that bacteria showed mainly positive correlations with other bacteria but negative correlations with fungi, particularly molds. Conversely, molds tended to exhibit positive correlations among themselves but negative correlations with bacteria. Correlation analysis of microbes and metabolites showed that unique microbial populations influenced variations in metabolite profiles among various *Daqu* samples. Furthermore, KEGG annotation emphasized the substantial enrichment of amino acid and sugar metabolism in diverse *Daqu* groups. Future research should focus on isolating phages from *Daqu* microbiome through culture-dependent methods such as the double-layer agar plate technique (Kropinski et al., 2009). The isolated phages will facilitate the exploration of the metabolism and interactions among core phages and the dominant bacteria and fungi in *Daqu* within a simulated fermentation system.

#### 5. Conclusion

In conclusion, viral communities, active fermenting microbial communities, and metabolic profiles varied significantly across low, medium, and high-temperature *Daqu* groups. Phages of *Siphoviridae*, *Podoviridae*, *Herelleviridae*, and *Myoviridae* showed varying levels of abundance in these *Daqu* groups. Furthermore, microbes like *Bacillus*, *Lactobacillus*, *Streptococcus*, *Phycomyces*, *Rhizopus*, and *Lichtheimia* were notably enriched in various *Daqu* groups. The existence of phage communities impacted the population dynamics of their host bacteria. Correlation networks were constructed to clarify the complex

relationships among phages, microbes, and metabolites. Our findings highlight the role of phage communities in influencing the metabolic functions observed in different *Daqu* groups.

## CRediT authorship contribution statement

**Xiaoning Huang:** Writing – original draft, Visualization, Methodology, Formal analysis. **Rengshu Li:** Writing – review & editing, Validation, Methodology. **Jinguo Xu:** Methodology. **Jiamu Kang:** Writing – review & editing, Visualization. **Xiaoxue Chen:** Writing – review & editing. **Beizhong Han:** Conceptualization. **Yansong Xue:** Writing – review & editing, Supervision, Project administration, 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.

## Acknowledgements

This work was supported by The National Natural Science Foundation of China (Grant No. 32272275), Natural Science Foundation of Sichuan Province (Grant No. 2025ZNSFSC0228) and China Alcoholic Drinks Association (Deciphering the phage diversity and distribution characteristics in *Daqu*).

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.foodres.2025.116061>.

## Data availability

All raw sequence data generated in this research has been deposited in the NCBI Sequence Read Archive under accession number: PRJNA1132293 and PRJNA1132256.

## References

- Alexeeva, S., Guerra Martínez, J. A., Spus, M., & Smid, E. J. (2018). Spontaneously induced prophages are abundant in a naturally evolved bacterial starter culture and deliver competitive advantage to the host. *BMC Microbiology*, 18, 1–16.
- Banerjee, S., Schlaepi, K., & van der Heijden, M. G. (2018). Keystone taxa as drivers of microbiome structure and functioning. *Nature Reviews Microbiology*, 16(9), 567–576.
- Cai, W., Wang, Y., Ni, H., Liu, Z., Liu, J., Hou, Q., Shan, C., Yang, X., & Guo, Z. (2021). Diversity of microbiota, microbial functions, and flavor in different types of low-temperature *Daqu*. *Food Research International*, 150, Article 110734.
- Chen, S., Huang, J., Qin, H., Zhou, R., Yang, Y., Qiu, C., & Zhang, S. (2021). Characterizing the interaction relationship of the microbial communities between *Zaopei* and pit mud disturbing by *Daqu*. *Food Science and Biotechnology*, 30(10), 1357–1367.
- Chevallereau, A., Pons, B. J., van Houte, S., & Westra, E. R. (2022). Interactions between bacterial and phage communities in natural environments. *Nature Reviews Microbiology*, 20(1), 49–62.
- Du, H., Chen, B., Fu, W., Yang, F., Lv, X., Tan, Y., Xi, X., Wang, L., & Xu, Y. (2023). Composition and function of viruses in sauce-flavor *baijiu* fermentation. *International Journal of Food Microbiology*, 387, Article 110055.
- Du, H., Wang, X., Zhang, Y., & Xu, Y. (2019). Exploring the impacts of raw materials and environments on the microbiota in Chinese *Daqu* starter. *International Journal of Food Microbiology*, 297, 32–40.
- Fan, Y., Huang, X., Chen, J., & Han, B. (2020). Formation of a mixed-species biofilm is a survival strategy for unculturable lactic acid bacteria and *Saccharomyces cerevisiae* in *Daqu*, a Chinese traditional fermentation starter. *Frontiers in Microbiology*, 11, Article 510900.
- Fernández, L., Rodríguez, A., & García, P. (2018). Phage or foe: An insight into the impact of viral predation on microbial communities. *The ISME Journal*, 12(5), 1171–1179.
- Fu, G., Deng, M., Chen, K., Chen, Y., Cai, W., Wu, C., ... Wan, Y. (2021). Peak-temperature effects of starter culture (*Daqu*) on microbial community succession and volatile substances in solid-state fermentation (*Jiutai*) during traditional Chinese special-flavour *baijiu* production. *Lwt*, 152, Article 112132.
- Hara, S., Terauchi, K., & Koike, I. (1991). Abundance of viruses in marine waters: Assessment by epifluorescence and transmission electron microscopy. *Applied and Environmental Microbiology*, 57(9), 2731–2734.
- Hu, X., Wang, K., Chen, M., Fan, J., Han, S., Hou, J., ... Wei, T. (2020). Profiling the composition and metabolic activities of microbial community in fermented grain for the Chinese strong-flavor *baijiu* production by using the metatranscriptome, high-throughput 16S rRNA and ITS gene sequencings. *Food Research International*, 138, Article 109765.
- Huang, D., Yu, P., Ye, M., Schwarz, C., Jiang, X., & Alvarez, P. J. (2021). Enhanced mutualistic symbiosis between soil phages and bacteria with elevated chromium-induced environmental stress. *Microbiome*, 9, 1–15.
- Huang, X., Fan, Y., Lu, T., Kang, J., Pang, X., Han, B., & Chen, J. (2020). Composition and metabolic functions of the microbiome in fermented grain during light-flavor *baijiu* fermentation. *Microorganisms*, 8(9), 1281.
- Huang, X., Fan, Y., Meng, J., Sun, S., Wang, X., Chen, J., & Han, B.-Z. (2021). Laboratory-scale fermentation and multidimensional screening of lactic acid bacteria from *Daqu*. *Food Bioscience*, 40, Article 100853.
- Jansson, J. K., & Wu, R. (2023). Soil viral diversity, ecology and climate change. *Nature Reviews Microbiology*, 21(5), 296–311.
- Jiang, Y., Xiong, X., Danksa, J., & Parkinson, J. (2016). Metatranscriptomic analysis of diverse microbial communities reveals core metabolic pathways and microbiome-specific functionality. *Microbiome*, 4, 1–18.
- Jończyk, E., Kłak, M., Miedzybrodzki, R., & Górska, A. (2011). The influence of external factors on bacteriophages. *Folia Microbiologica*, 56, 191–200.
- Kang, J., Chen, X., Han, B. Z., & Xue, Y. (2022). Insights into the bacterial, fungal, and phage communities and volatile profiles in different types of *Daqu*. *Food Research International*, 158, Article 111488.
- Kang, J., Huang, X., Li, R., Zhang, Y., Chen, X. X., & Han, B. Z. (2024). Deciphering the core microbes and their interactions in spontaneous *baijiu* fermentation: A comprehensive review. *Food Research International*, 188, Article 114497.
- Kang, J., Xue, Y., Chen, X., & Han, B. Z. (2022). Integrated multi-omics approaches to understand microbiome assembly in *Jiutai*, a mixed-culture starter. *Comprehensive Reviews in Food Science and Food Safety*, 21(5), 4076–4107.
- Kong, S. J., & Park, J. H. (2019). Effect of bacteriophages on viability and growth of co-cultivated *Weissella* and *Leuconostoc* in kimchi fermentation. *Journal of Microbiology and Biotechnology*, 29(4), 558–561.
- Kropinski, A. M., Mazzocco, A., Waddell, T. E., Lingohr, E., & Johnson, R. P. (2009). Enumeration of bacteriophages by double agar overlay plaque assay. *Bacteriophages: Methods and Protocols, Volume 1: Isolation, characterization, and interactions*, 69–76.
- Li, R., Yu, C., Li, Y., Lam, T.-W., Liu, S.-M., Kristiansen, K., & Wang, J. (2009). SOAP2: An improved ultrafast tool for short read alignment. *Bioinformatics*, 25(15), 1966–1967.
- Li, Y., Liu, S., Zhang, S., Liu, T., Qin, H., Shen, C., ... Yin, Q. (2022). Spatiotemporal distribution of environmental microbiota in spontaneous fermentation workshop: The case of Chinese *baijiu*. *Food Research International*, 156, Article 111126.
- Liao, H., Liu, C., Ai, C., Gao, T., Yang, Q. E., Yu, Z., ... Friman, V. P. (2023). Mesophilic and thermophilic viruses are associated with nutrient cycling during hyperthermophilic composting. *The ISME Journal*, 17(6), 916–930.
- Liu, F., Mao, J., Kong, W., Hua, Q., Feng, Y., Bashir, R., & Lu, T. (2020). Interaction variability shapes succession of synthetic microbial ecosystems. *Nature Communications*, 11(1), 309.
- Liu, J., Chen, J., Fan, Y., Huang, X., & Han, B. (2018). Biochemical characterisation and dominance of different hydrolases in different types of *Daqu*—a Chinese industrial fermentation starter. *Journal of the Science of Food and Agriculture*, 98(1), 113–121.
- Pan, Y., Wang, Y., Hao, W., Duan, C., Wang, S., Wei, J., & Liu, G. (2022). Metatranscriptomics unravel composition, drivers, and functions of the active microorganisms in light-flavor liquor fermentation. *Microbiology Spectrum*, 10(3), e02151–e02121.
- Pang, X. N., Han, B. Z., Huang, X. N., Zhang, X., Hou, L. F., Cao, M., ... Chen, J. Y. (2018). Effect of the environment microbiota on the flavour of light-flavour *baijiu* during spontaneous fermentation. *Scientific Reports*, 8(1), 3396.
- Robinson, M. D., McCarthy, D. J., & Smyth, G. K. (2010). Edger: A Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics*, 26 (1), 139–140.
- Sakandar, H. A., Hussain, R., Khan, Q. F., & Zhang, H. (2020). Functional microbiota in Chinese traditional *baijiu* and *Mijiu Qu* (starters): A review. *Food Research International*, 138, Article 109830.
- Song, Z., Du, H., Zhang, Y., & Xu, Y. (2017). Unraveling core functional microbiota in traditional solid-state fermentation by high-throughput amplicons and metatranscriptomics sequencing. *Frontiers in Microbiology*, 8, 1294.
- Waller, A. S., Yamada, T., Kristensen, D. M., Kultima, J. R., Sunagawa, S., Koonin, E. V., & Bork, P. (2014). Classification and quantification of bacteriophage taxa in human gut metagenomes. *The ISME Journal*, 8(7), 1391–1402.
- Wang, B., Wu, Q., Xu, Y., & Sun, B. (2020). Synergistic effect of multiple saccharifying enzymes on alcoholic fermentation for Chinese *baijiu* production. *Applied and Environmental Microbiology*, 86(8), e00013–e00020.
- Wang, R., Wu, J., Jiang, N., Lin, H., An, F., Wu, C., Yue, X., Shi, H., & Wu, R. (2023). Recent developments in horizontal gene transfer with the adaptive innovation of fermented foods. *Critical Reviews in Food Science and Nutrition*, 63(5), 569–584.
- Wang, W., Fan, G., Li, X., Fu, Z., Liang, X., & Sun, B. (2020). Application of *Wickerhamomyces anomalus* in simulated solid-state fermentation for *baijiu* production: Changes of microbial community structure and flavor metabolism. *Frontiers in Microbiology*, 11, Article 598758.
- Wang, Y., Quan, S., Zhao, Y., Xia, Y., Zhang, R., Ran, M., ... Zhang, W. (2023). The active synergistic microbiota with *Aspergillus* as the core dominates the metabolic network of ester synthesis in medium-high temperature *Daqu*. *Food Microbiology*, 115, Article 104336.

- White, K., Yu, J.-H., Eraclio, G., Dal Bello, F., Nauta, A., Mahony, J., & van Sinderen, D. (2022). Bacteriophage-host interactions as a platform to establish the role of phages in modulating the microbial composition of fermented foods. *Microbiome Research Reports*, 1(1).
- Wu, Q., Li, L., Xiang, P., Zhang, T., Peng, L., Zou, L., & Li, Q. (2023). Phages in fermented foods: Interactions and applications. *Fermentation*, 9(3), 201.
- Wu, S., Du, H., & Xu, Y. (2023). *Daqu* microbiota adaptability to altered temperature determines the formation of characteristic compounds. *International Journal of Food Microbiology*, 385, Article 109995.
- Xiao, C., Lu, Z. M., Zhang, X. J., Wang, S. T., Ao, L., Shen, C. H., ... Xu, Z. H. (2017). Bioheat is a key environmental driver shaping the microbial community of medium-temperature *Daqu*. *Applied and Environmental Microbiology*, 83(23). e01550–01517.
- Yang, I. S., & Kim, S. (2015). Analysis of whole transcriptome sequencing data: Workflow and software. *Genomics & informatics*, 13(4), 119.
- Zhang, H., Zhang, H., Du, H., Yu, X., & Xu, Y. (2023). The insights into the phage communities of fermented foods in the age of viral metagenomics. *Critical Reviews in Food Science and Nutrition*, 12, 1–13.
- Zhang, Y., Shen, Y., Niu, J., Ding, F., Ren, Y., Chen, X., & Han, B. Z. (2023). Bacteria-induced amino acid metabolism involved in appearance characteristics of high-temperature *Daqu*. *Journal of the Science of Food and Agriculture*, 103(1), 243–254.
- Zhang, Y., Xu, J., Ding, F., Deng, W., Wang, X., Xue, Y., ... 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.
- Zheng, L., Liang, X., Shi, R., Li, P., Zhao, J., Li, G., Wang, S., Han, S., Radosevich, M., & Zhang, Y. (2020). Viral abundance and diversity of production fluids in oil reservoirs. *Microorganisms*, 8(9), 1429.
- Zheng, X. W., & Han, B. Z. (2016). *Baijiu* (白酒), Chinese liquor: History, classification and manufacture. *Journal of Ethnic Foods*, 3(1), 19–25.