ABSTRACT

Milk fan is an acid-curd cheese with strong national characteristics (a traditional dairy product of the Bai nationality with a shape like a piece of paper) and a long history in Yunnan province, China. In our previous study, we characterized the microbial community diversity of milk fan, but the succession of microorganisms associated with flavor formation in milk fan is still unknown. Therefore, we examined the predominant microorganisms and their correlations with the formation of flavor in the fermentation of sour juice and drying of milk fan by gas chromatography mass spectrometry, high-throughput 16S rDNA sequencing, intergenic spacer sequencing and metatranscriptome analysis. We found that the relative abundances of *Lactobacillus* and *Issatchenkia* initially decreased and then increased with time during the fermentation of sour juice. However, the relative abundances of *Acetobacter*, *Leuconostoc*, *Lactococcus*, *Geotrichum*, and *Dipodascus* initially increased and then decreased. During the drying step, the relative abundances of *Lactobacillus* and *Issatchenkia* continuously increased and became the dominant microorganisms in the milk fan. The metatranscriptomes generated from the milk fan showed that “carbohydrate metabolism,” “translation,” and “signal transduction” were the main metabolic functions of the microbial communities. *Rhodotorula* and *Yarrowia* contained more differentially expressed genes than other genera, which indicated they may be associated with the production of the characteristic flavor. Furthermore, a Pearson correlation analysis showed that *Lactococcus*, *Rhodotorula*, *Candida*, *Cutaneotrichosporon*, and *Yarrowia* were significantly positively correlated with more aroma-active compounds, mainly ethyl acetate, 2-heptanone, isovaleraldehyde, butyric acid, nonanal, and hexanal. In conclusion, these findings contribute to a better understanding of the flavor production mechanism during the production of milk fan.

Key words: milk fan, microbial community, microbial succession, high-throughput sequencing, aroma-active compound

INTRODUCTION

Milk fan is a type of acid-curd cheese with a long history and strong national characteristics of the Bai ethnic minority in Yunnan province, China (Tian et al., 2020). Milk fan is a traditional milk product; it is made of fresh milk as the raw material, which is mixed with sour juice fermented from papaya juice after heating, and then subjected to curding, blanching, stretching, forming, and drying (2–3 d of drying; Tian et al., 2019). The unique raw materials and processing method endow milk fan with a good nutritional profile, a delicate taste, a mild aroma and a paper-fan-like shape, which make this product popular (Chen et al., 2021).

Aroma is an important attribute of food, and it can be used as an important indicator of food quality. The varieties of cheeses with different textures and aromas can be attributed to the diversified functional microorganisms inhabiting cheese (Xue et al., 2018). Bertuzzi et al. (2018) found that *Debaryomyces hansenii* was correlated with the production of alcohols and carboxylic acids, *Glutamicibacter arilaitensis* with alcohols, carboxylic acids and ketones, and *Brevibacterium linens* and *Geotrichum candidum* with sulfur compounds in surface-ripened cheese. At present, the research on milk fan has mainly focused on the identification of key aroma compounds, the perceptual interactions among key aroma compounds and the screening and identification of strains (Liu et al., 2009; Tian et al., 2020). In our previous research, key aroma compounds in Chinese milk fan were identified (Tian et al., 2019). We found that *Lactobacillus*, *Rhodotorula*, *Lodderomyces*, and *Debaryomyces* were core functional microbial genera that facilitated formation of the profile of volatile compounds in milk fan (Chen et al., 2021). In a similar study, Jiang et al. (2021) also showed that *Lactobacillus*, *Lactococcus*, and *Streptococcus* were the...
dominant genera in rushan (the Chinese name of milk fan), and that sour juice made a great contribution to shaping the relative abundance of microbes in milk fan.

The metabolic activities of the microbiota in the process of fermenting foods, such as proteolysis, lipolysis, and glycolysis, are important pathways for generating volatile compounds (Sousa et al., 2001; Collins et al., 2003; Marilley and Casey, 2004). These compounds are important for the formation of sensory properties (De Filippis et al., 2016). Therefore, microbial succession during the fermentation process is crucial for the formation of specific flavors and has been studied by high-throughput sequencing with bioinformatics analysis in recent studies. Monnet et al. (2016) applied metagenomic and metatranscriptome sequencing to reblochon-style cheese fermentation processes to determine the microbial diversity. Walsh et al. (2017) used whole-metagenome sequencing combined with metabolomics and flavor analysis to illustrate the dynamic nature of kefir fermentation and microbial succession patterns. Their findings can be applied to optimize the fermentation process, flavors, and other fermented foods. To date, little systematic research has been conducted on the microbial succession of milk fan throughout fermentation. The methods mentioned above have the potential to shed new light on the metabolic and functional capabilities of microbiomes and their succession in milk fan.

The objectives of this study were (1) to investigate the composition and succession of microorganisms in sour juice fermentation and milk fan drying process by high-throughput 16S rDNA and intergenic spacer sequencing (ITS) gene sequencing technology; (2) to determine the metabolic function and gene expression activities of the microbial community using metatranscriptomics; and (3) to analyze the contribution of microorganisms to the formation of characteristic aroma compounds in the different stages of milk fan production. The findings in this study could be important for improving the quality of milk fan and promoting industrial production.

MATERIALS AND METHODS

Milk Fan Production and Sample Collection

The production of the milk fan used in the present study was adapted from Xue et al. (2018) with some minor modifications. Sour juice samples were prepared by boiling papaya juice and water (1:3, wt/wt) for 30 min. It was naturally fermented for about 6 d. Pasteurization at 70°C for 30 s was used in milk sterilization. When the milk was cooled down to 40 to 50°C, sour juice was mixed with the milk at the volume ratio 1:2. Then the mixture was stirred until solidification occurred under heat and acidity. The solidified curd was picked up by bamboo chopsticks and kneaded in a pot by hand. The homogenized curd was then stretched into an oval slice and wrapped counter-clockwise on a bamboo chopstick. The curd on the bamboo stick was naturally dried for 24 h (room temperature: 18–20°C, relative humidity: 80–85%). The bamboo chopstick was removed to obtain milk fan.

Sour juice samples were collected at 0 (S0), 3 (S3), and 6 (S6) d since fermentation. Milk fan samples were collected at 0 (R0) and 24 (R24) h during the drying stage (Figure 1). Five biological replicates were collected at each sampling point. To reduce the sampling bias, sour juice samples were collected from the upper, middle and lower parts of the fermentation tank after the sour juice was fully stirred. The whole milk fan samples were cut into evenly sized cubes and then mashed with a mortar. Each sample was divided into 3 parts. The samples for DNA extraction were stored at 4°C; samples for metatranscriptomic analysis were snap-frozen in liquid nitrogen and stored at –80°C; and samples for the analysis of volatile compounds were stored at –20°C.

The contents of moisture, water activity (Aw), fat, and protein, salt content, and the pH value of milk fan were determined according to standard methods no. 926.08, 978.18, 933.05, and 991.22, 983.14, respectively, as described by AOAC International (2000).

DNA Extraction, PCR Amplification, and Illumina MiSeq Sequencing

Fifty milliliters of a sour juice sample was placed into a sterile centrifuge tube. After centrifugation at 2,650 × g for 10 min at 4°C, the precipitated pellet was collected and placed in a refrigerator (–80°C). A mashed milk fan sample of 3 g was weighed and placed in 27 mL of sterile normal saline (0.85 g/100 mL). The thoroughly mixed sample suspension was centrifuged at 4°C for 2 min at 2,650 × g. The precipitates were collected and placed in a sterile centrifuge tube in a refrigerator (–80°C). DNA in the collected samples (including sour juice samples and milk fan samples) was extracted using a TIANamp Stool DNA Kit (fecal genome DNA extraction kit; Beijing, China). The extracted DNA was stored at –20°C.

To examine bacteria, the V3 to V4 domains of 16S rDNA were amplified by PCR (95°C for 5 min; followed by 25 cycles of 9°C for 30 s, 55°C for 30 s, and 72°C for 40 s with a final extension of 72°C for 10 min).

Journal of Dairy Science Vol. 106 No. 11, 2023
using the primer pairs 338-F (5′-ACTCCTACGGAGGCAGCAG-3′) and 806-R (5′-GGACTACHVGGGTWTCTAAT-3′) with specific barcodes. To examine fungi, the ITS1 rDNA regions were amplified by PCR (95°C for 2 min, followed by 30 cycles at 95°C for 30 s, 61°C for 30 s, and 72°C for 45 s with a final extension at 72°C for 10 min) with the primers ITS1F (5′-CTTGTTGATYATGGGTAAGTG-3′) and ITS2R (5′-GCTGCGTTCTTCATCAGTG-3′) with specific barcodes. The purified amplicons were pooled at equimolar concentrations, and further paired-end sequencing was performed using a MiSeq platform (Illumina Inc., San Diego, CA).

**Bioinformatics Analysis**

Sequences were analyzed using the QIIME-1.9.1 pipeline (Caporaso et al., 2010). Low-quality sequences were filtered, and the remaining sequences were assigned using effective tags to operational taxonomic units (OTU) on the basis of 97% similarity of sequences determined by Uparse software-7.0.1090. Representative OTU with a high frequency of occurrence were selected and annotated with taxonomic information (e.g., phylum, family, and genus levels; Xue et al., 2018).

Alpha diversity indices (e.g., abundance-based coverage estimator [ACE], Chao1, and observed species) were calculated to estimate the species richness and relative diversity level in sour juice and milk fan samples (Wang and Shao, 2018). Differences of bacterial community structures among sour juice and milk fan samples were assessed using a phylogeny-based metric, weighted UniFrac distance. If the calculated weighted UniFrac distance between samples is relatively small, the samples are more similar and share more microbial lineages of common evolutionary history.

Putative mRNA reads were subject to BLASTX comparisons against a nonredundant protein database obtained from the National Center for Biotechnology Information (E value, <10⁻⁵). The least common ancestor-based algorithm implemented in MEGAN was used to determine the taxonomic level of each gene. MetaGene Annotator was applied to assemble contigs, and reads longer than 100 base pairs were identified. BLASTP was used to query the predicted proteins by matching sequence reads against the integrated nonredundant protein database (E value, <10⁻⁵). Functional annotation was performed using Kyoto Encyclopedia of Genes and Genomes (KEGG; version April 2011; http://www.genome.jp/kegg/) and Clusters of Orthologous

---

**RNA Extraction, Preparation, and High-Throughput Sequencing**

Three independent RNA extracts were extracted from 0.5-g milk fan samples using an EZNA Soil RNA Kit (Omega BioTek Inc., Norcross, GA) in accordance with the manufacturer’s instructions (Xue et al., 2018). The extracts were then added to DNase I (TaKara, Beijing, China) and incubated at 37°C for 1 h to eliminate traces of DNA. The RNA quality was determined using an Agilent 2100 bioanalyzer and quantified using a NanoDrop spectrophotometer (Thermo Scientific, Wilmington, DE). Ribosomal RNA was removed from the total RNA samples using an Epicenter Ribo-Zero rRNA Removal Kit (Epicenter, Madison, WI). Before the construction of libraries using the TruSeq RNA sample preparation kit (Illumina), 3 independent RNA extracts of each sample were combined. The pooled libraries were sequenced on a MiSeq platform (Illumina) in accordance with standard protocols to generate more than 33 million reads, averaging 356 nucleotides per sample.
Groups (COG; http://www.ncbi.nlm.nih.gov/COG/). Transcript expression was normalized using fragments per kilobase of transcript per million fragments mapped (Trapnell et al., 2010).

The amplicon databases were submitted to the National Center for Biotechnology Information Sequence Read Archive. These databases are available under the accession numbers PRJNA884308 and PRJNA884265 for the high-throughput DNA sequencing data (16S rDNA and ITS sequences) and metatranscriptome data, respectively.

Volatile Compound Analysis

The method used for volatile compound analysis was similar to that used in our previous study with some minor modifications (Chen et al., 2021). Samples (sour juice or milk fan, 3 g) were mixed with 100 μL of internal standard (2-octanoyl, 13 μg/L), and the mixture was transferred to a 20-mL vial. After equilibrating at 60°C for 5 min, a solid-phase microextraction fiber (Supelco Inc., Bellefonte, PA) coated with divinylbenzene-carboxen-polydimethylsiloxane (100 μm in thickness) was exposed to the headspace of a 20-mL vial for 35 min, and the vial was stirred continuously at 250 rpm. The GC temperature protocol was as follows. The initial temperature was 40°C for 4 min, and it was then increased to 100°C at 3°C/min for 2 min, to 150°C at 4°C/min, and finally to 230°C at 10°C/min for 5 min. The aroma components were identified by comparison of their mass spectrometry spectra with NIST17 libraries and comparison of their Kovats retention indices with those reported in the literature (http://webbook.nist.gov; Cao et al., 2021).

Statistical Analysis

The data of volatile compounds were analyzed by ANOVA followed by Duncan’s multiple range tests, using IBM SPSS Statistics 21 (IBM Corp., Armonk, NY). P-values <0.05 were considered statistically significant. R software (version 3.2.5; https://www.R-project.org/) was used to calculate the Pearson correlation coefficients between aroma-active compounds and dominant microorganisms in milk fan samples, and Cytoscape 3.4.0 was used to draw a correlation network.

RESULTS

Sequencing Data of Sour Juice and Milk Fan Samples

The total numbers of effective sequences in the samples ranged from 41,592 to 50,246 for bacteria, and from 66,917 to 72,560 for fungi. All of the sequences were clustered with a 97% sequence identify cut-off. The number of OTU for bacteria ranged from 26 to 104, and that for fungi ranged from 16 to 51.

Physical and Chemical Analysis

In the process of sour juice fermentation, the pH value of sour juice gradually decreased owing to acid production (S0: 6.8 ± 0.3, S3: 4.4 ± 0.2, S6: 2.1 ± 0.1). The mean protein, fat, pH value, salt content, moisture and Aw of the fresh milk fan (R0) and milk fan (R24) are shown in Supplemental Table S1 (https://doi.org/10.6084/m9.figshare.23893725.v1, Tian, 2023). The pH value of milk fan increased from 4.9 (R0) to 5.3 (R24), and similar changes have also been reported in milk fan samples (Jiang et al., 2021). Furthermore, the drying procedure dramatically reduced the moisture content and nearly doubled the protein and fat content.

Alpha and β Diversity Analysis of Microbial Communities in Milk Fan Samples at Different Production Stages

The α diversity indices (e.g., ACE, Chao1, Shannon and Simpson) of sour juice and milk fan samples at different production stages are shown in Figure 2. The first 2 indices, ACE (see Figure 2A, E) and Chao1 (see Figure 2B, F), represent the richness of the community, while the Shannon and Simpson indices reflect the species diversity of the community (He and Chung, 2020). We observed some features of the variation in microbial community richness during milk fan production. In the sour juice fermentation stages (S0–S6), the ACE and Chao1 indices of the samples decreased, the Shannon (see Figure 2C, G) increased and the Simpson indices (see Figure 2D, H) decreased, which indicated that the richness and diversity of bacterial and fungal communities decreased with increasing fermentation time. During the drying process after the milk fan was made (R0–R24), these indices of the samples showed opposite trends, which indicated that the richness and diversity of bacterial and fungal communities had increased.

The site-to-site variability in species composition, known as β diversity, is crucial to understanding spatiotemporal patterns of species diversity and the mechanisms controlling community composition and structure (Zhou et al., 2013). A principal coordinates analysis of bacterial and fungal community structures in the samples was performed on the basis of the unweighted (Figure 3B and D) and weighted UniFrac distances (Figure 3A and C). In bacteria, the weighted UniFrac distance distribution among
the samples was large (Figure 3A), which indicated that the bacterial community structure of milk fan was significantly different between the production stages \((P < 0.001)\). As is shown in Figure 3C and D, in fungi, no obvious difference was seen in the microbial community structures of sour juice samples \((S0, S3, \text{and } S6)\) at different fermentation periods of sour juice. However, the fungal community structure of milk fan samples \((R0 \text{ and } R24)\) was significantly different over the drying period \((P < 0.001)\), whereas that of sour juice samples \((S0, S3 \text{ and } S6)\) was almost similar (Figure 3C and D).

**Composition and Succession of the Microbial Community During the Production of Milk Fan**

We further examined the changes in microbial composition at the phylum and genus level during the production of milk fan by investigating the community distributions of bacteria and fungi in sour juice and milk fan samples (Figure 4). Microbial taxonomic identification showed that there were 4 bacterial phyla and 2 fungal phyla and 10 bacterial genera and 15 fungal genera with an abundance of \(\geq 0.1\%\) in all samples during milk fan production. The dominant bacterial genera \((>1\%\)}
The relative abundance of *Lactobacillus* (76.93%) was the highest at the beginning of sour juice fermentation (S0), and *Acetobacter* (18.73%) and *Leuconostoc* (0.58%) accounted for only a small part of the initial fermentation community. During S0 to S3, the abundances of *Acetobacter* (from 18.73 to 48.51%) and *Leuconostoc* (from 0.58 to 15.00%) rapidly increased, but that of *Lactobacillus* (from 76.93 to 32.43%) was inhibited. Subsequently, the relative abundances of *Acetobacter* and *Leuconostoc* gradually decreased during S3 to S6 and also during R0 to R24. Additionally, the relative abundance of *Lactobacillus* increased to become the dominant genus in the samples.

The dominant fungal genera (>1% abundance) were *Issatchenkia*, *Geotrichum*, *Dipodascus*, *Rhodotorula*, *Cutaneotrichosporon*, and *Candida*, which belong to *Ascomycota* and *Basidiomycota*. During S0 to S3, the relative abundances of *Geotrichum* and *Dipodascus* significantly increased by 62.52 and 25.14%, respectively, while the abundance of *Issatchenkia*, which initially had the highest relative abundance, significantly decreased by 12.29% (P < 0.001). The relative abundance...
of each fungal genus changed during S3 to S6, but the amplitude was small. During R0 to R24, the relative abundance of *Geotrichum* and *Dipodascus* significantly decreased by 18.35 and 8.15%, respectively (*P* < 0.001), and the relative abundance of *Issatchenkia* increased to become the dominant genera. Notably, *Rhodotorula*, which was the core functional microorganism found in our previous studies (Chen et al., 2021), showed a low relative abundance in the sour juice fermentation stage, and began to proliferate in the drying stage of milk fan.

**Comparative Analysis of Functions of Microorganisms Obtained by Metatranscriptomics in the Milk Fan Drying Process**

As shown in the previous section, microbial community richness was the highest, variable abundances were found in the same dominant microorganisms, and new minor ones were detected in the drying stage of milk fan. Therefore, the milk fan samples at R0 and R24 were further analyzed using metatranscriptome technology. The metatranscriptomic sequencing data quality is shown in Supplemental Table S2 (https://doi.org/10.6084/m9.figshare.23893725.v1, Tian, 2023). A total of 70.29 Gbp of raw data were generated, from which 56.11 Gbp of clean data were obtained after screening and filtering. The clean data acquisition rate of each sample was >92% after removing the low-quality sequences, which was sufficient for the requirements of the analysis. The unigene N50 length ranged from 629 to 831 bp, which indicated a good assembly and splicing effect, and the metatranscriptomic sequence was sufficient for subsequent analysis.

Microbial functions were identified using KEGG annotation of metatranscriptome data in the milk fan drying stage (Kanehisa et al., 2002). The primary metabolic pathways could be divided into 6 categories.
(Figure 5A), namely metabolism, genetic information processing, human diseases, environmental information processing, cellular processes, and organismal systems. Among them, the number of genes annotated by the metabolism category was the largest, followed by genetic information processing and environmental information processing. The secondary metabolic pathways with the largest number of enriched genes in these 3 functional annotations were carbohydrate metabolism, translation, and signal transduction, respectively. Metabolism-related genes were the most abundant, with an annotation abundance of 35.89%, whereas genes associated with cellular processes were the least abundant. Furthermore, the relative abundance of genes in most metabolism-related pathways, such as carbohydrate metabolism and metabolism of other AA, gradually increased during the drying process (Figure 5B). These findings suggest that carbohydrates and AA have the most important effect on the formation of flavor in milk fan.

The COG functional classification can be used to determine the specific reactions and potential functions of related genes in the milk fan microbiota by annotating unknown sequences to known proteins (Wu et al., 2015). The predominant COG categories in the milk fan samples included translation, ribosomal structure and biogenesis (J); carbohydrate transport and metabolism (G); posttranslational modification, protein turnover, and chaperones (O); AA transport and metabolism (E); nucleotide transport and metabolism (F); and replication, recombination and repair (L; Figure 6). This finding suggests that cell growth, energy production and metabolism of microbial communities are active during the milk fan drying period.

Microbial Function and Metabolic Activity in the Milk Fan Drying Process

To compare the expression and distribution of differentially expressed genes between the samples in the drying stage of milk fan, the expression levels of all genes in the 2 samples (R0 and R24) were visualized by scatter and volcano plots (Supplemental Figure S1; https://doi.org/10.6084/m9.figshare.23893725.v1, Tian, 2023). We found 35,470 genes in the metatranscriptomic sequencing results of all milk fan samples, of which 15,711 genes were differentially expressed (14,187 upregulated genes, 1,524 downregulated genes). Additionally, 108 genes were significantly differentially expressed, with 97 upregulated genes and 11 downregulated genes ($P < 0.05$). During the drying stage of milk fan, the number of upregulated differentially expressed genes was significantly higher than that of downregulated genes ($P < 0.05$). This finding indicated that the functional diversity and metabolic activity of microorganisms gradually increases during the drying stage of milk fan.

The main public database in pathways is KEGG, which divides biological metabolic pathways into 7 categories (Kanehisa et al., 2008). Figure 7 shows that the metabolic pathways with enriched genes included ribosome and oxidative phosphorylation, RNA transport, glycolysis/gluconeogenesis, biosynthesis of AA, carbon metabolism, fatty acid biosynthesis and fatty acid metabolism. The enrichment pathways with a significant difference between the 2 samples were ribosomes from genetic information processing and oxidative phosphorylation from metabolism ($P < 0.0001$). The differentially expressed genes were mainly from Rho-
Among them, *Rhodotorula* and *Yarrowia* were the main active genera, corresponding to 30 and 29 differentially expressed genes, respectively (accession number: PRJNA884265).

**Dynamic Changes in Volatile Compounds During the Production of Milk Fan**

To further study the effect of microbial community structure on the formation of characteristic aroma components of milk fan, qualitative and quantitative analyses of volatile compounds in milk fan samples at different production stages were carried out (Supplemental Table S3; https://doi.org/10.6084/m9.figshare.23893725.v1, Tian, 2023). Sixteen volatile compounds with an odor activity value (OAV) >1 were selected by calculating the OAV values of each compound (Table 1). Most of the volatile compounds, such as butyric acid, hexanoic acid, ethyl butyrate, and isobutyl butyrate, are similar with our previous study (Chen et al., 2021). These compounds contributed to the overall aroma at least in one milk fan sample. The concentrations of ethyl butyrate and ethyl hexanoate gradually decreased in the fermentation stage of sour juice (Table 1). The concentrations of ethyl acetate, 2-heptanone, dipentene, ethyl caprylate, butyric acid, isovaleric acid, and 4'-methylacetophenone gradually increased. In particular, 4 aldehydes (isovaleraldehyde, hexanal, heptaldehyde, and 1-nonanal) appeared in the drying stage of milk fan, and their concentrations gradually increased. This finding illustrated that the microbial activity in the drying stage of milk fan was conducive to the release of aldehydes. Aldehydes are formed from AA by transamination or oxidative deamination of amines (Collins et al., 2003; Fox and McSweeney, 2017). Aldehydes have low perception thresholds and are characterized by an herbaceous and green-grass-like aroma (Ozturkoglu-Budak et al., 2016). Among them, isovaleraldehyde and 1-nonanal endow milk fan with a malt and citrus aroma, respectively.

To analyze the effect of the succession of microbial community structure on the aroma of milk fan during production, the Pearson correlation coefficients between 16 aroma-active compounds (OAV >1) and 10 dominant microorganisms (4 bacteria and 6 fungi) were calculated. The interactions with coefficients >0.6 were selected as strongly correlated nodes. Using a co-occurrence network analysis, 23 nodes and 48 edges were obtained (|ρ| > 0.6, P < 0.05, Figure 8). Figure 8 shows that the predominant microorganisms were *Lactococcus* and *Rhodotorula*, which were strongly positively correlated with a large quantity of aroma-active compounds, such as hexanal, 1-nonanal, heptaldehyde, butyric acid, isovaleraldehyde, dipentene, and 2-heptanone. For fungi, *Candida*, *Rhodotorula*, and *Cutaneotrichosporon* were significantly positively correlated with aroma-active compounds, which mainly comprised ethyl acetate (0.65–0.66), 2-heptanone (0.93–0.97), dipentene (0.71–0.90), isovaleraldehyde (0.98–0.99), butyric acid (0.88–0.96), nonanal (0.74–0.93), and hexanal (0.91–0.99;
These results were partly similar with our previous study (Chen et al., 2021), in which *Rhodotorula*, *Debaryomyces*, and *Candida* were positively correlated with heptanoic acid, octanoic acid, 1-nonanal, 2-heptanone, heptanoic acid, octanoic acid, isovaleraldehyde, and isobutyl butyrate. For bacteria, *Lactococcus* were significantly positively correlated with aroma-active compounds, which mainly comprised ethyl acetate ($r = 0.66$), 2-heptanone ($r = 0.72$), dipentene ($r = 0.64$), isovaleraldehyde ($r = 0.86$), butyric acid ($r = 0.66$), nonanal ($r = 0.81$), and hexanal ($r = 0.87$; $P < 0.001$). Some negative correlations were also observed: *Geotrichum* and *Dipodascus* were negatively correlated with ethyl hexanoate and ethyl butyrate,
Table 1. Changes in volatile compounds (odor activity value >1) in the production of milk fan

<table>
<thead>
<tr>
<th>Compound</th>
<th>RI²</th>
<th>Identification method²</th>
<th>S0</th>
<th>S3</th>
<th>S6</th>
<th>R0</th>
<th>R24</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ethyl acetate</td>
<td>884</td>
<td>RI, MS</td>
<td>ND</td>
<td>ND</td>
<td>58.15 ± 13.22</td>
<td>ND</td>
<td>71.19 ± 16.46</td>
</tr>
<tr>
<td>Isovaleraldehyde</td>
<td>916</td>
<td>RI, MS</td>
<td>ND</td>
<td>ND</td>
<td>2 ± 2.35</td>
<td>ND</td>
<td>7.38 ± 5.45</td>
</tr>
<tr>
<td>Ethyl butyrate</td>
<td>1,036</td>
<td>RI, MS</td>
<td>7 ± 0.74</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Hexanal</td>
<td>1,106</td>
<td>RI, MS</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Heptaldehyde</td>
<td>1,175</td>
<td>RI, MS</td>
<td>ND</td>
<td>ND</td>
<td>13.09 ± 18.32</td>
<td>ND</td>
<td>10.12 ± 6.42</td>
</tr>
<tr>
<td>2-Heptanone</td>
<td>1,193</td>
<td>RI, MS</td>
<td>ND</td>
<td>ND</td>
<td>4.04 ± 0.65b</td>
<td>4.51 ± 3.56b</td>
<td>19.32 ± 5.46a</td>
</tr>
<tr>
<td>3-Methyl-1-butanol</td>
<td>1,221</td>
<td>RI, MS</td>
<td>41.04 ± 4.04ab</td>
<td>58.62 ± 1.84a</td>
<td>46.45 ± 3.05b</td>
<td>38.39 ± 15.39b</td>
<td>36.36 ± 6.32b</td>
</tr>
<tr>
<td>Ethyl hexanoate</td>
<td>1,240</td>
<td>RI, MS</td>
<td>38.43 ± 1.71a</td>
<td>22.11 ± 1.29b</td>
<td>12.98 ± 0.82a</td>
<td>4.44 ± 3.75d</td>
<td>24.2 ± 6.62a</td>
</tr>
<tr>
<td>Dipentene</td>
<td>1,263</td>
<td>RI, MS</td>
<td>ND</td>
<td>ND</td>
<td>17.56 ± 2.25a</td>
<td>46.78 ± 42.19a</td>
<td>53.78 ± 12.64a</td>
</tr>
<tr>
<td>2-Octanone</td>
<td>1,288</td>
<td>RI, MS</td>
<td>379.37 ± 29.05b</td>
<td>813.41 ± 9.6a</td>
<td>352.42 ± 13.76b</td>
<td>392.91 ± 30.15b</td>
<td>371.12 ± 42.43a</td>
</tr>
<tr>
<td>2-Heptanol</td>
<td>1,317</td>
<td>RI, MS</td>
<td>1.82 ± 0.66a</td>
<td>3.59 ± 0.97a</td>
<td>2.85 ± 0.71a</td>
<td>ND</td>
<td>3.91 ± 3.28a</td>
</tr>
<tr>
<td>1-Nonanal</td>
<td>1,397</td>
<td>RI, MS</td>
<td>ND</td>
<td>ND</td>
<td>93.62 ± 4.8</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>Ethyl caprylate</td>
<td>1,434</td>
<td>RI, MS</td>
<td>ND</td>
<td>ND</td>
<td>11.91 ± 1.5b</td>
<td>11.59 ± 1.66b</td>
<td>27.05 ± 4.08b</td>
</tr>
<tr>
<td>Butyric acid</td>
<td>1,659</td>
<td>RI, MS</td>
<td>11.59 ± 1.66b</td>
<td>11.59 ± 1.66b</td>
<td>17.29 ± 2.08b</td>
<td>ND</td>
<td>9.99 ± 4.94b</td>
</tr>
<tr>
<td>Isovaleric acid</td>
<td>1,691</td>
<td>RI, MS</td>
<td>ND</td>
<td>ND</td>
<td>169.47 ± 21.39</td>
<td>84.89 ± 22.65b</td>
<td>47.11 ± 11.14b</td>
</tr>
</tbody>
</table>

¹Values in the same row with different superscripts are significantly different (P < 0.05).
²Aroma compounds detected in the samples by GC-MS.
³Retention index of volatile compounds on an HP-Innowax column (Agilent, Santa Clara, CA).
⁴RI = retention index calculation in agreement with a value from the literature. MS = mass spectrum comparison using the NIST17.L library (http://webbook.nist.gov/).
⁵Values are the mean ± SD. S0, S3, S6 = sour juice samples collected at 0 (S0), 3 (S3), and 6 (S6) d after fermentation; R0, R24 = milk fan samples collected at 0 (R0) and 24 (R24) h after the drying stage.
⁶ND indicates not detected.
DISCUSSION

Milk fan is a traditional milk product of Yunnan province, China, with strong national characteristics. Similar to many other fermented dairy products (De Filippis et al., 2016), a variety of microorganisms (e.g., bacteria and fungi), which are abundant in the production process of milk fan, play an important role in the maturation and flavor formation of milk fan (Yohan et al., 2016). In our previous research, we identified the key aroma compounds of milk fan (Tian et al., 2019) and analyzed their correlations with the microbial community (Chen et al., 2021). However, no study has focused on the succession of microorganisms during milk fan fermentation. Recently, high-throughput sequencing approaches have been successfully applied to determine the relationship between a food product and its microbes in more detail (Liu et al., 2019; Chen et al., 2020). Therefore, the aroma profiles associated with microbial succession during milk fan fermentation were determined by high-throughput sequencing and flavor profiling in the present study.

The production of milk fan involves sour juice fermentation and milk fan drying. In the process of sour juice fermentation, *Lactobacillus* and *Geotrichum* are dominant bacteria and yeast, respectively. Some bacteria and yeast, such as *Acetobacter*, *Geotrichum*, and *Dipodascus*, proliferated, while the number of other microorganisms, such as *Issatchenka*, gradually declined. As for flavor compounds, ethyl caprylate, isovaleric acid, and 4’-methylacetophenone are shown in a relatively high level (Table 1). *Lactobacillus* is generally considered to have strong esterase activity (Fox and Wallace, 1997; Collins et al., 2003). Esterases of *Lactobacillus* can hydrolyze the di- and monoglycerides in milk fat to release short-chain fatty acids (such as isovaleric acid), synthesize esters (such as ethyl caprylate) from glycerides, and alcohols via a transferase reaction. There was also almost no succession of fungal community structure during the stage of sour juice fermentation. The weighted UniFrac distance distribution among the samples indicated that the change of bacterial community presented more obvious changes than fungal community during the sour juice fermentation stage.

During the milk fan drying stage, a relatively lower Aw and moisture was found in R24 than R0. Such envi-

Figure 8. Correlation network between the core microorganisms and volatile compounds (odor activity value >1) during milk fan production. The blue, green, and yellow nodes represent bacteria, fungi, and volatile compounds, respectively.

and *Lactococcus* were negatively correlated with ethyl caprylate.
environment enabled microorganisms, such as *Lactobacillus, Issatchenkia, Lactococcus*, and *Rhodotorula*, to proliferate in large quantities. This result is partly consistent with our previous study, in which the microbial communities found in all milk fan samples were dominated by *Lactococcus, Lactobacillus*, and *Rhodotorula* (Chen et al., 2021). The different sampling locations may contribute to the various dominant microbial communities. The relative abundance of *Rhodotorula* and *Lactococcus* were observed to be positively related to 2-Heptanone, which was significantly \( P < 0.05 \) increased in milk fan samples compared with sour juice. Study has shown that β-oxidation or autooxidation of fatty acids can produce methyl ketones in cheese (Hanko et al., 2018). Furthermore, the richness and diversity of microbial communities changed with the drying time, and it is similar to that of a previous study on Porc and Corsican varieties of cheeses, in which microorganisms initially showed decreasing trends then increased (Casalta et al., 2009; Aldrete-Tapia et al., 2014). In addition, the β diversity analysis showed that there was a significant difference in the microbial community structure between milk fan samples and sour juice samples \( P < 0.001 \). And this finding demonstrated that the drying process of the milk fan greatly changed the microbial community structure in the samples.

Similar to a previous study (Dugat-Bony et al., 2015), we found that carbohydrate metabolism, AA metabolism, energy metabolism, transport and catabolism, and translation and signal transduction accounted for the most abundant transcripts throughout the drying period of milk fan. Carbohydrates play an important role in regulating cell activity. Carbohydrates yield carbon dioxide, water, energy, ethanol, acetic acid, lactic acid, and other substances through glycolysis, pyruvate metabolism and the tricarboxylic acid cycle. In the drying stage of milk fan, where a low Aw results in an adverse environment, cellular metabolism is directed toward multiplication and energy production from carbohydrates (De Filippis et al., 2016). In this study, the largest increased metabolic class involved carbohydrates, especially secondary ascorbate and aldarate metabolism, glycosylation or gluconeogenesis, and glyoxylate and dicarboxylate metabolism, with average increases in abundance of 117.48, 47.61, and 38.36%, respectively (Supplemental Table S4; https://doi.org/10.6084/m9.figshare.23893725.v1, Tian, 2023). The abundance of fatty acid degradation was increased by 28.14%. Therefore, the microorganisms in milk fan might be beneficial for the above-mentioned metabolic processes in the drying stage of milk fan. Bacterial gene expression associated with ripening activities may be associated with metabolite profiles, such as proteolysis and lipolysis in the process of ripening (De Filippis et al., 2016).

In addition to carbohydrate metabolism and fatty acid degradation, AA metabolism was a main metabolic activity in the milk fan drying process (Figure 5). This finding is similar to that of a study by Jiang et al. (2021), and it may be due to the microorganisms in milk fan encoding enzymes related to AA. Furthermore, the high relative abundance of *Lactobacillus* might be associated with AA metabolism in the milk fan, which was also reported by Jiang et al. (2021). Generally, AA metabolism in cheese mainly involves 2 pathways. One pathway is catalyzed by AA lyases, which cleave the side chains of AA and finally produce phenol, indol, and methanethiol (Yvon and Rijnen, 2001). Another pathway involves the production of α-keto acid intermediates from AA under the action of aminotransferases, and then the production of aldehydes under the action of α-keto acid decarboxylase (de la Plaza et al., 2009). Further metabolism of aldehydes leads to the production of acids or alcohols by various cellular dehydrogenases, all of which are important in the development of flavor.

In this study, we identified significant correlations between the abundances of particular microbial community and the levels of different volatile compounds \( P < 0.05 \). This finding indicated that the microbes in milk fan had an important effect on the production of these compounds. Notably, *Lactococcus, Cutaneotrichosporon*, *Candida*, and *Rhodotorula* were correlated with butyric acid, which is associated with cream flavors. Similar findings in the fermented products paocai and rice wine have previously been reported (Li et al., 2018; Jiang et al., 2021; Zhao et al., 2022). Additionally, *Rhodotorula*, which was the main active genera and the main source of the differentially expressed genes, was significantly positively correlated with dipentene, 2-heptanone, isovaleraldehyde, and esters in our study \( P < 0.05 \). We also found that *Lactobacillus* and *Issatchenkia* were strongly associated with ester substances such as ethyl hexanoate and ethyl heptanoate, which suggests that both bacteria and fungi are crucial for formation of the volatile compounds of milk fan.

Similar with our previous study, *Lactobacillus* was correlated with ethyl hexanoate, *Rhodotorula* was positively correlated with isobutyl butyrate (Chen et al., 2021).

**CONCLUSIONS**

In this study, the dynamic changes in volatile compounds and the regulation of succession and metabolic activity of microbial communities during the production of milk fan were examined. We found that microbial
richness and diversity were more abundant during the drying stage than during sour juice fermentation. Lactobacillus and Issatchenkia were the dominant microorganisms in milk fan. Lactococcus, Rhodotorula, Candida, Cutaneotrichosporon, and Yarrowia were identified as important genera in the process of milk fan fermentation and were significantly positively correlated with the main aroma-active compounds. Carbohydrate metabolism, translation, and signal transduction were the main metabolic functions of microbial communities. This study may provide a theoretical reference for understanding the growth and activity of microorganisms associated with the flavor of milk fan during fermentation and the drying process. Furthermore, our findings may guide for achieving desirable flavors by selecting a starter culture for the development and industrial production of milk fan.

ACKNOWLEDGMENTS

This work was financially supported by the National Natural Science Foundation of China (31771943), and the Shanghai Engineering Technology Research Center of Shanghai Science and Technology Commission (No. 20DZ2255600; Shanghai, China). The authors have not stated any conflicts of interest.

REFERENCES

Tian, H. 2023. Supplemental Material 1.do

ORCIDS
Chen Chen https://orcid.org/0000-0002-4885-393X
Haiyan Yu https://orcid.org/0000-0001-6472-1366
Huixiang Tian https://orcid.org/0000-0007-1194-3812