Food Chemistry: X 24 (2024) 101877

Contents lists available at ScienceDirect

Food Chemistry: X

journal homepage: www.sciencedirect.com/journal/food-chemistry-x

Machine learning discrimination and prediction of different quality grades of sauce-flavor *baijiu* based on biomarker and key flavor compounds screening

Shuai Li[a,b], Tao Li[a,b], Yueran Han[c], Pei Yan[c], Guohui Li[c], Tingting Ren[c], Ming Yan[c], Jun Lu[c,**], Shuyi Qiu[a,b,*]

[a] College of Liquor and Food Engineering, Key Laboratory of Fermentation Engineering and Biological Pharmacy of Guizhou Province, Guizhou University, Guiyang, Guizhou 550025, China
[b] Key Laboratory of Fermentation Engineering and Biological Pharmacy of Guizhou Province, China
[c] Guizhou Guotai Distillery Co. Ltd., Renhuai, Guizhou 564501, China

ARTICLE INFO

Keywords:
Machine learning
Sauce-flavor *baijiu*
Quality grade
Flavor compounds
Biomarker

ABSTRACT

The quality grade of base *Baijiu* directly determines the final quality of sauce-flavor *Baijiu*. However, traditional methods for assessing these grades often rely on subjective experience, lacking objectivity and accuracy. This study used GC-FID, combined with quantitative descriptive analysis (QDA) and odor activity value (OAV), to identify 27 key flavor compounds, including acetic acid, propionic acid, ethyl oleate, and isoamyl alcohol etc., as crucial contributors to quality grade differences. Sixteen bacterial biomarkers, including *Komagataeibacter* and *Acetobacter* etc., and 7 fungal biomarkers, including *Aspergillus* and *Monascus* etc., were identified as key microorganisms influencing these differences. Additionally, reducing sugar content in *Jiupei* significantly impacted base *Baijiu* quality. Finally, 11 machine learning classification models and 9 prediction models were evaluated, leading to the selection of the optimal model for accurate quality grade classification and prediction. This study provides a foundation for improving the evaluation system of sauce-flavor *Baijiu* and ensuring consistent quality.

1. Introduction

Chinese *Baijiu* is one of the six most renowned distilled spirits worldwide (Qiao et al., 2023), with sauce-flavor *Baijiu* being particularly noted for its complex and diverse aromatic profile, characterized by a prominent sauce aroma, elegance, finesse, and a lingering fragrance in an empty glass (Duan et al., 2022). However, these qualities also complicate the determination of its quality grade.

The brewing process of sauce-flavor *Baijiu* involves two rounds of raw material addition, nine steaming cycles, eight fermentations, and seven rounds of alcohol extraction (Liu & Sun, 2018). The final product of sauce-flavor *Baijiu* is created by blending the alcohol extracted during seven rounds in varying proportions. This alcohol, referred to as base *Baijiu,* serves as the foundation for the finished product (Wu et al., 2023). The base *Baijiu* from different rounds varies in production cycles due to the eight fermentation stages of the *Jiupei* (Fermented grains), leading to distinct microbial compositions across the rounds. For example, Yang et al. revealed the spatiotemporal differences in microbial community structure, succession, and environmental drivers during the stacked fermentation of different rounds of sauce-flavor *Baijiu* (Yang et al., 2023). Previous studies have shown that the microbial community composition and metabolism of fermented *Jiupei* influence the flavor profile of sauce-flavor *Baijiu* (Du, Ji, et al., 2021; Du, Song, et al., 2021; Zhang, Wang, Tan, et al., 2021), leading to distinct flavor characteristics in base *Baijiu* from different rounds. Liu et al. found that Grain aroma and Sour taste were characteristic of the first and second rounds of base *Baijiu*, while Sauce aroma and Roasted aroma were typical of the third to fifth rounds, and Burnt aroma characterized the sixth and seventh rounds (Liu, He, et al., 2023). However, studies have shown that abnormal fermentation during the first round of sauce-flavor *Baijiu*

Abbreviations: QDA, Quantitative descriptive analysis; OAV, odor activity value; MLP, Multilayer Perceptron; SVM, Support Vector Machine; KNN, K-Nearest Neighbor; LDA, Linear Discriminant Analysis; VIP, Variable Importance in Projection..

* Correspondence to: Shuyi Qiu, College of Liquor and Food Engineering, Guizhou University, Guiyang, Guizhou, China.
** Correspondence to: Jun Lu, Guizhou Guotai Distillery Co. Ltd., Renhuai, Guizhou, China.
E-mail addresses: lujun_0316@hotmail.com (J. Lu), syqiu20067474@163.com (S. Qiu).

https://doi.org/10.1016/j.fochx.2024.101877
Received 21 August 2024; Received in revised form 25 September 2024; Accepted 3 October 2024
Available online 5 October 2024
2590-1575/© 2024 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY-NC license (http://creativecommons.org/licenses/by-nc/4.0/).

production could disrupt subsequent pit fermentation (Chen et al., 2022; Zhang, Wang, Tan, et al., 2021). Since pit fermentation is a closed anaerobic process, once the *Jiupei* enters the pit, no further adjustments can be made regardless of the fermentation status. Therefore, interventions must occur before the *Jiupei* enters the pit to ensure the quality grade of the base *Baijiu* (Yang et al., 2024). Currently, industry research on base *Baijiu* quality grades primarily focuses on flavor defects. For example, Wang et al. characterized the pickle-like off-flavors in sauce-flavor *Baijiu*, which are mainly caused by sulfur compounds like 2-methyl-3-furanthiol and dimethyl disulfide(Wang et al., 2020). However, few studies have analyzed the relationship between base *Baijiu* quality grades and the microbial composition of the corresponding *Jiupei*. More importantly, current industry methods for determining base *Baijiu* quality grades rely heavily on traditional sensory evaluations based on subjective experience (He et al., 2022; Li et al., 2023), making the results susceptible to external factors such as the physical condition of the tasters and the tasting environment. Therefore, the industry not only needs to understand the factors that contribute to differences in base *Baijiu* quality grades but also requires the development of objective and accurate methods for quality grade determination.

Instrumental analysis, as a rapid and accurate detection method, is increasingly applied in the analysis of *Baijiu* and other flavor-rich foods (He et al., 2023; Zhou et al., 2023). However, modern analytical instruments generate extensive data, some of which may include parameters unrelated to quality grading. This redundant information can affect the accuracy of *Baijiu* quality discrimination (Liu, Zhang, et al., 2023). To address these challenges, machine learning has demonstrated significant advantages in denoising, parameter selection, model discrimination, and prediction (Menichetti et al., 2023). It has already achieved notable success in disease prediction and food quality grading (Catacutan et al., 2024; Schreurs et al., 2024). Nevertheless, the application of machine learning in *Baijiu* research is still rare. Therefore, constructing predictive models based on machine learning algorithms combined with multidimensional data from modern instruments offers a more comprehensive and accurate approach to reflecting sensory quality differences in *Baijiu* compared to traditional statistical methods and subjective sensory evaluations.

Addressing the challenges in determining base *Baijiu* quality grades, this study focused on the first-round base *Baijiu* from sauce-flavor *Baijiu* production. The study characterized sensory differences in *Baijiu* style across different quality grades by analyzing physicochemical indicators (total acids and total ester) combined with quantitative descriptive analysis (QDA). Representative *Jiupei* samples corresponding to different quality grades of base *Baijiu* were subjected to metagenomic sequencing and physicochemical analysis to identify key indicators and microorganisms responsible for the observed quality differences. Representative *Jiupei* samples corresponding to different quality grades of base *Baijiu* were subjected to metagenomic sequencing and physicochemical analysis to identify key indicators and microorganisms responsible for the observed quality differences. GC-FID analysis identified 52 major flavor components in base *Baijiu* of different quality grades, and OPLS-DA and odor activity value (OAV) were used to elucidate the key flavor compounds contributing to these quality differences. Finally, 11 machine learning classification models and 9 regression models were employed to construct predictive models for classifying base *Baijiu* quality grades. The optimal model, evaluated by model performance metrics, was further refined using the SHAP method for feature interpretation and final model construction. This study not only provides an efficient and accurate method for improving the evaluation system of base *Baijiu* quality grades in Sauce-flavor *Baijiu*, but also offers a scientific foundation for preventing and controlling fermentation issues to ensure consistent base *Baijiu* quality.

2. Materials and methods

2.1. Sample collection

The base *Baijiu* and *Jiupei* (Fermented grain) samples were obtained from Guizhou Guotai Liquor Group Co., Ltd. in Maotai Town, Guizhou, China. The base *Baijiu* samples were collected during the first round of sauce-flavor *Baijiu* winter production (December 2023–January 2024). A total of 133 base *Baijiu* samples were collected, including 42 first-grade, 42 s-grade, and 49 third-grade samples, labeled as A, B, and C, respectively. The base *Baijiu* samples were stored at room temperature (25 °C). Sampling of *Jiupei* after the completion of stacked fermentation was conducted following the method described by Wang et al. (Wang, Huang, & Huang, 2021) (Fig. S1). A total of 134 *Jiupei* samples were collected. Samples for physicochemical analysis and volatile compound detection were stored at 4 °C, while those for metagenomic sequencing were stored at −80 °C.

2.2. Quantitative descriptive analysis

The methodology was adapted and modified based on previous literature (Liu, He, et al., 2023). The quality grade of base *Baijiu* was evaluated based on three dimensions: aroma, taste, and mouthfeel, following the GB/T33404–2016 standard and the international ISO11035 standard for selecting sensory descriptive language. Grain aroma was selected as the aroma characteristic, Sour taste as the taste characteristic, and Mellowness, Durability, and Cleanliness as the mouthfeel characteristics. Each panel member had passed basic olfactory and taste tests and training (GB/T15549–2022) and had at least 2 years of evaluation training experience. The panel consisted of 10 members aged 25 to 45, all of whom had over 2 years of experience in evaluating sauce-flavor *Baijiu*. All participants were informed of all the details of the sensory experiment and their consent was obtained before the experiment began. All participants provided written informed consent prior to study participation, and participation was completely voluntary.Samples (20 mL each) were presented in standard *Baijiu* tasting glasses and randomly coded. Sensory evaluation results were scored on a scale of 0 to 10, where 0 indicated no odor and 10 indicated the strongest odor. Each panel member conducted sensory evaluations three times, and the average score was recorded as the final result.

2.3. Odor activity value

The OAV was used to determine whether a compound contributes to the aromatic profile, following the method reported by Wang et al. (Wang et al., 2020) with slight modifications. The OAV is calculated by dividing the concentration of the compound (C) by its odor threshold (OT). Compounds with an OAV ≥ 1 are considered to contribute to the overall aroma characteristics (Gong et al., 2023).

2.4. Physicochemical analysis

The physicochemical properties of *Jiupei*, including temperature, moisture, titratable acidity, reducing sugars, and starch content, were measured following the method reported by Yang et al. (Yang et al., 2023). Lactic acid and ethanol contents were determined using an M-100 biosensor (Xilman Technology, Beijing, China). The total acids and total esters in the base *Baijiu* samples were determined according to the method reported by Zhang et al. (Zhang, Wang, Wang, et al., 2021).

2.5. Metagenomic sequencing

From the 134 *Jiupei* samples, 9 were selected for metagenomic sequencing, representing different quality grades of base *Baijiu*: 3 samples corresponding to A-grade base *Baijiu*, 3 to B-grade, and 3 to C-grade, designated as A-JP, B-JP, and C-JP, respectively. Metagenomic sequencing of the *Jiupei* samples was conducted by Personal Biotechnology Co., Ltd. (Shanghai, China). DNA was extracted from the samples using the OMEGA Mag-Bind Soil DNA Kit (M5635–02) (Omega Bio-Tek, Norcross, GA, USA). All sequencing analyses were performed on the Illumina NovaSeq 6000 platform, following the method outlined in previous studies (Cao et al., 2024).

2.6. Qualitative and quantitative analysis of 52 major flavor compounds in base baijiu using GC-FID

GC-FID analysis was performed using a Shimadzu instrument equipped with a flame ionization detector (GC-2030). The method was adapted from previous literature (Gong et al., 2023; Liu, He, et al., 2023) with slight modifications, using a DB-Wax column (30 m × 0.25 mm × 0.25 μm) for separation. The injection temperature was set to 250 °C, and high-purity nitrogen was used as the carrier gas. The flow rate was set to 1 mL/min with a split ratio of 37:1 and a tail purge of 20 mL/min. Hydrogen and air flow rates were adjusted to 40 mL/min and 400 mL/min, respectively, with the detector temperature maintained at 250 °C. The temperature program was held at 60 °C for 3 min, then increased at a rate of 5 °C/min. The temperature was further increased to 230 °C at a rate of 10 °C/min and held for 5 min. Standard compounds were dissolved in a 56 % (*v*/v) ethanol solution prepared with ethanol and ultrapure water to obtain a mixed standard stock solution. This solution was then diluted into a series of different concentration gradients. The standard solution (1.0 mL) was mixed with a reference solution containing isoamyl alcohol, n-pentyl acetate, and 2-ethylbutyric acid, and the mixture was analyzed using GC-FID. Qualitative analysis of the compounds was performed by comparing the mass spectra with information from the NIST 17 databases. Accurate quantification was achieved by establishing standard curves for 52 major flavor compounds (Table S1).

2.7. Machine learning model construction

The base *Baijiu* quality grade classification and prediction models were constructed as shown in Fig. 1. To avoid model overfitting, Lasso feature selection was applied, following the method reported by Shaon et al. (Shaon et al., 2024). Data preprocessing involved standardizing the transformed features using the "StandardScaler" function. The dataset was split into training and test sets (20 % of the data) using the train_test_split function, following the method reported by Al-Zaiti et al. (Al-Zaiti et al., 2023). Stratified sampling (stratify = labels) was applied to maintain consistent class distribution. SMOTE (Synthetic Minority Over-sampling Technique) was used to oversample the training set, ensuring balanced sample data for model training and prediction. Eleven machine learning classifiers, including Logistic Regression, K-Nearest Neighbo (KNN), Random Forest, Linear Discriminant Analysis (LDA), Multilayer Perceptron (MLP), Support Vector Machine (SVM), Decision Tree, Extra Trees, Gradient Boosting, Bagging, and AdaBoost, were selected for base *Baijiu* quality grade classification. Nine machine learning regression models, including Linear Regression, KNN, Random Forest, Decision Tree, SVM, Extra Trees, Gradient Boosting, Bagging, and AdaBoost, were used for quality grade score prediction. SMOTE was again applied to the training set to ensure data balance during model training and prediction. The classification models were evaluated using AUC, F1-Score, Accuracy, Recall, and Precision, while the prediction models were evaluated using R^2 and RMSE (Liu, Zhang, et al., 2023; Schreurs et al., 2024). The SHAP algorithm was used for global interpretation, following the method reported by Chen et al. (Chen et al., 2023). Shapley values for each feature were calculated across various models to rank feature importance, thereby enhancing the credibility and transparency of the

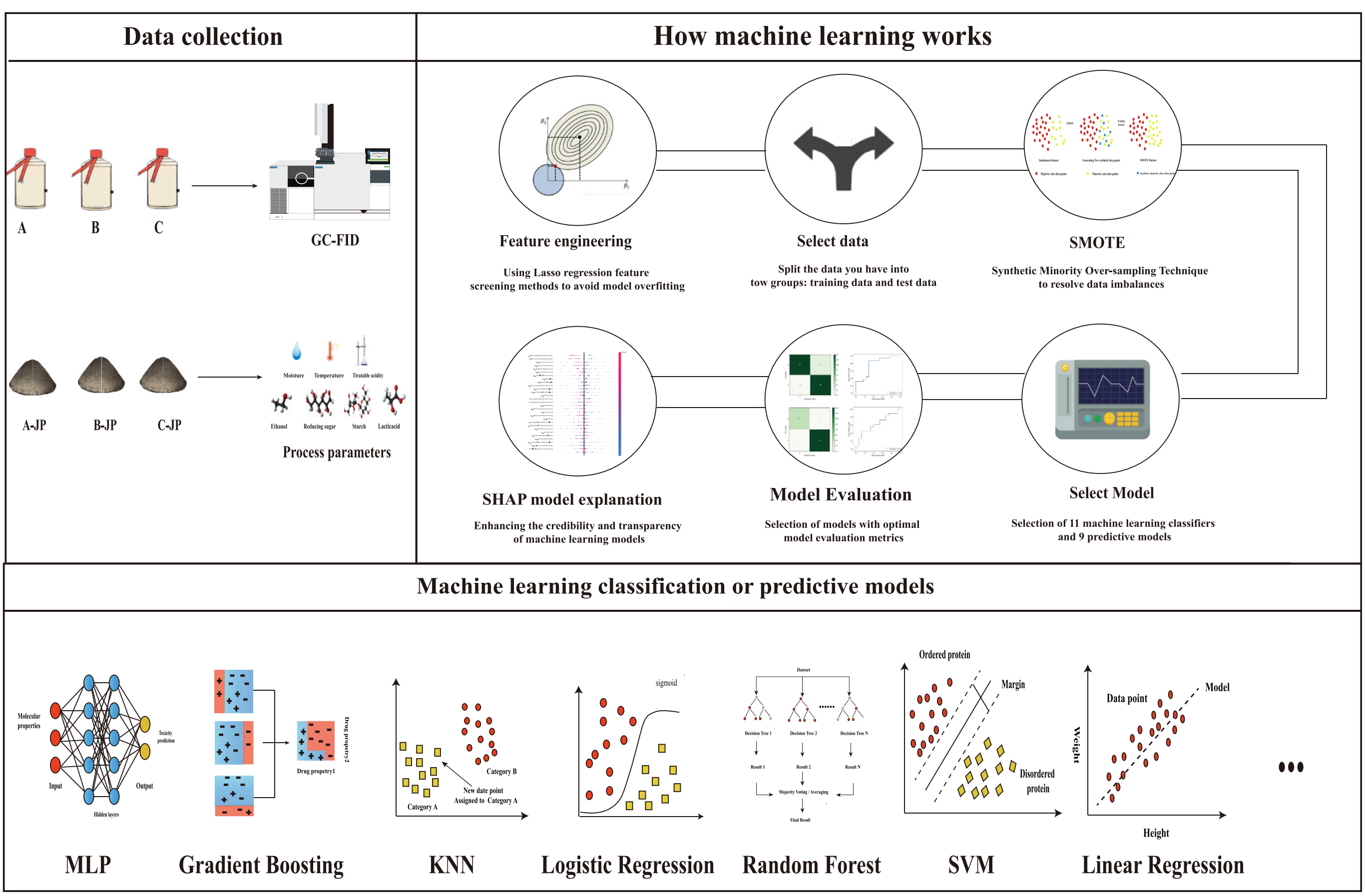

Fig. 1. Workflow for constructing machine learning models.

machine learning models. All machine learning models were implemented using the Scikit-learn library in Python 3.8 on the Jupyter Notebook platform. SHAP values for the samples were calculated using the SHAP library.

2.8. Statistical analysis and visualization

Statistical differences between and within sample groups were evaluated using one-way ANOVA combined with *t*-tests in SPSS 22.0, with $p < 0.05$ considered statistically significant(Yang et al., 2024). Principal component analysis (PCA) was conducted using the "FactoMine" package in R (V.4.3.2), combined with Permutational Multivariate Analysis of Variance (Permanova) using the adonis function from the "vegan" package to determine differences in microbial community structure, the R value in the PCA plot refers to the ANOSIM R statistic, calculated using the anosim function from the vegan package in R (v.4.3.2), which measures the degree of separation between different sample groups. LEfSe analysis was performed using the "tidyverse," "microeco," "magrittr," and "patchwork" packages in R (V.4.3.2), with LDA > 2 and $p < 0.05$ considered indicative of key microorganisms in *Jiupei* (Cao et al., 2024). Mantel-test analysis was conducted using the "vegan" and "corrplot" packages in R (V.4.3.2), with $p < 0.05$ and $r > 0.6$ considered significant, the r value in the Mantel test refers to the Spearman correlation coefficien (Yang et al., 2023). OPLS-DA analysis was conducted using the opls function from the "ropls" package in R (V.4.3.2), with Variable Importance in Projection (VIP) scores used to assess the significance of flavor compound differences, where VIP > 1 was considered significant (Gong et al., 2023). Other statistical analyses and visualizations were generated using OriginPro 2023 (Version 2023, OriginLab Corporation), RStudio (V.4.3.2), and Adobe Illustrator CC 2018 (Version 2018, Adobe Systems Incorporated).

3. Results and discussion

3.1. Quantitative descriptive analysis and physicochemical differences of base baijiu across quality grades

Base *Baijiu* of different quality grades exhibits distinct sensory profiles, with grade A outperforming others across various dimensions. ANOVA analysis of Grain aroma, Sour taste, Mellowness, Durability, and Cleanliness revealed significant differences between quality grades across all five dimensions (Fig. 2a) ($p < 0.001$). However, *t*-test analysis showed no significant differences between grades A and B in Mellowness and Grain aroma ($p > 0.05$), but significant differences were observed in Sour taste, Durability, and Cleanliness ($p < 0.05$). Grade C exhibited significantly greater differences compared to A and B across all dimensions ($p < 0.001$) (Fig. S2a-e). In summary, high-quality base *Baijiu* is characterized by prominent Grain aroma, distinct Sour taste, Smoothness, Cleanliness, and a long aftertaste. The most typical features of first-round base *Baijiu* are its pronounced Sour taste and Grain aroma.

Previous studies have shown that first-round base *Baijiu* is marked by prominent Sour taste and Grain aroma (Liu, He, et al., 2023; Wu et al., 2023), likely due to the limited steaming cycles in the early production stage, resulting in high tannin content in *Jiupei*. During fermentation, microorganisms degrade tannins, producing significant amounts of ferulic acid, acetic acid, and other acidic compounds (Zhao et al., 2024). Lower-quality first-round base *Baijiu* typically lacks distinct Sour taste and is affected by off-flavors such as pickled vegetables, bran, and rust, which diminish the overall quality of the liquor (Wang et al., 2020; Wu et al., 2021). Significant differences were also found between quality grades in terms of total acids (B: 3.679 ± 1.20 g /L > A: 3.352 ± 0.65 g/L > C: 2.684 ± 0.48 g/L) and total esters content (B: 8.360 ± 2.40 g /L > A: 7.768 ± 1.63 g /L > C: 6.204 ± 1.11 g /L) ($p < 0.001$) (Fig. 2b, c).

3.2. Quantitative analysis of 52 major flavor compounds in base baijiu of different quality grades by GC-FID and identification of differential compounds

Building on previous studies (Duan et al., 2022; Wang et al., 2024), 52 major flavor compounds were identified in sauce-flavor *Baijiu*, including 20 esters, 11 alcohols, 9 acids, 6 aldehydes, 3 pyrazines, 2 ketones, and 1 phenol. Ethyl acetate, n-propanol, ethyl lactate, and acetic acid were identified as dominant compounds in first-round base *Baijiu* (Fig. S3b), consistent with earlier findings (Liu, He, et al., 2023; Wu et al., 2023). To identify the key flavor compounds contributing to differences in quality grades, an OPLS-DA model was applied (Fig. 3a). The model demonstrated a high explanatory power of 71.55 % and exhibited strong stability and reliability (Fig. S3a) (Gong et al., 2023). Additionally, compounds with VIP > 1 in the OPLS-DA were identified as the key contributors to flavor differences across quality grades (Fig. 3b). A total of 27 flavor compounds were identified, including 9 acids: acetic acid, propionic acid, valeric acid, heptanoic acid, caprylic acid, succinic acid, isovaleric acid, linoleic acid, and oleic acid. The alcohols included 2,3-butanediol, n-propanol, methanol, 2-phenylethanol, 1,2-propanediol, and isoamyl alcohol. The aldehydes included benzaldehyde, acetaldehyde, furfural, and isobutyraldehyde. The esters included ethyl acetate, diethyl succinate, ethyl myristate, ethyl caprylate, ethyl butyrate, ethyl oleate, ethyl valerate, and ethyl linoleate. Other compounds included tetramethylpyrazine, acetoin, and acetal. To further analyze the key aroma compounds contributing to the flavor differences among quality grades, OAV and aroma descriptors were calculated for the VIP >1 compounds (Table 1) (Dong et al., 2024; Wang et al., 2024). Notably, the OAV of acetic acid (A: 1491.3 > B: 1350.0 > C: 923.67) showed a positive correlation with the quality grade of base *Baijiu*, while propionic acid had an OAV < 1 in grade C, indicating no aroma contribution (Table 1). Therefore, differences in the concentrations of acetic acid and propionic acid may be the main factors contributing to the sour taste differences among quality grades of first-round base *Baijiu.* Furthermore, ethyl oleate, which has a soapy and fatty odor (C: 74.1 > B: 49.8 > A: 43.0), ethyl linoleate with a plastic

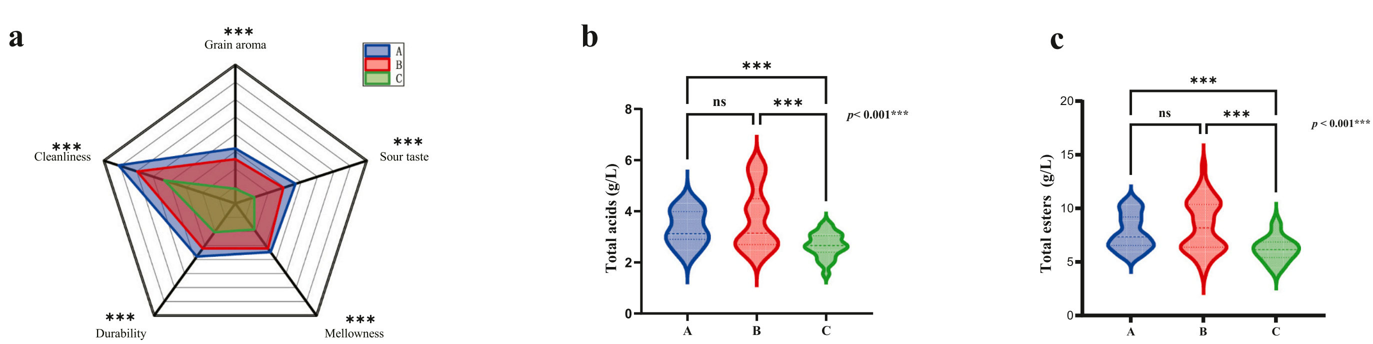

Fig. 2. (a) Radar chart of quantitative descriptive analysis (QDA) for different quality grades of base *Baijiu*. Physicochemical differences among base *Baijiu* quality grades: (b) total acids, (c) total esters. The *, ** and *** indicate statistical significance at $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively.

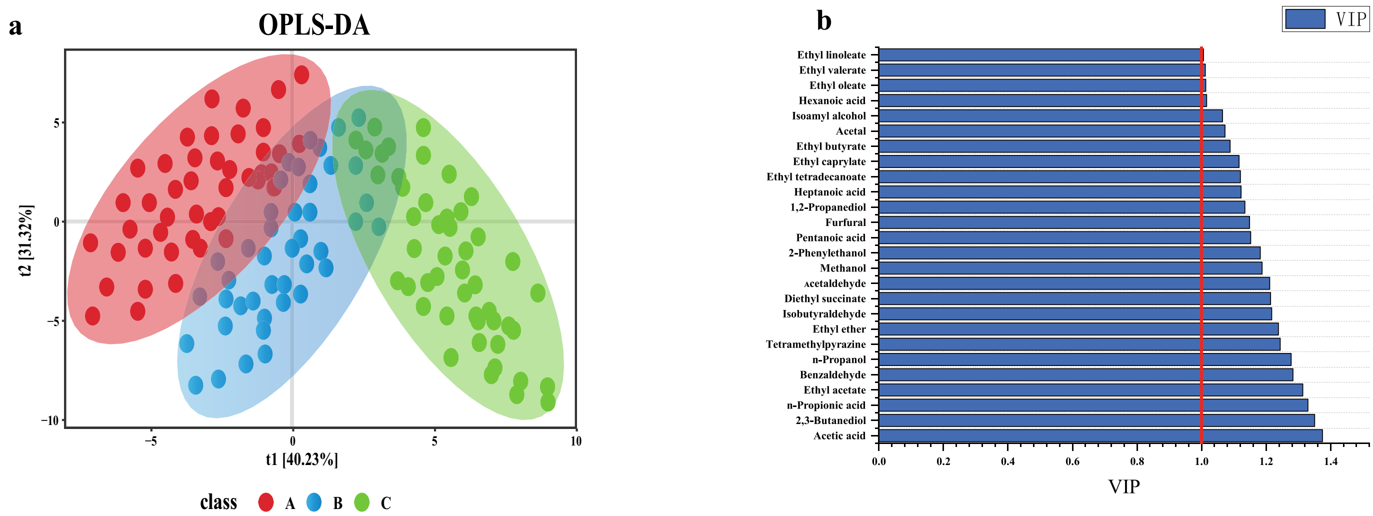

Fig. 3. (a) OPLS-DA analysis of different base *Baijiu* quality grades; (b) Key flavor compounds (VIP > 1) in different base *Baijiu* quality grades.

Table 1
OAV of key flavor compounds in base *Baijiu* of different quality grades.

CAS	Flavor components	Threshold(mg/L)	Average Concentration(mg/L)			OAV			Aroma Description
			A	B	C	A	B	C	
75–07-0	Acetaldehyde	1.2	110.87	128.18	155.32	92.39	106.82	129.43	Green leaves, grassy notes
78–84-2	Isobutyraldehyde	0.006	10.86	7.76	5.15	1810.45	1293.71	858.50	Floral
141–78-6	Ethyl acetate	17	4578.63	5253.92	3437.85	269.33	309.05	202.23	Apple
105–57-7	Acetal	50	70.88	86.52	99.78	1.42	1.73	2.00	Grassy, Fruity
67–56-1	Methyl Alcohol	100	189.43	233.70	151.85	1.89	2.34	1.52	Alcohol
105–54-4	Ethyl butyrate	0.15	32.85	40.22	19.36	218.98	268.16	129.03	Pineapple
71–23-8	n-Propanol	720	3674.90	4482.13	1753.97	5.10	6.23	2.44	Alcoholic, Fruity
539–82-2	Ethyl valerate	0.02678	5.08	5.27	2.52	189.73	196.61	94.02	Fruity
123–51-3	Isopentyl alcohol	6.5	242.76	223.20	275.14	37.35	34.34	42.33	Motor oil
513–86-0	Ethylene glycol	0.259	6.12	13.05	10.35	23.65	50.39	39.96	Creamy
106–32-1	Ethyl caprylate	0.24	1.29	1.50	1.85	5.36	6.23	7.70	Fruity Pear, Pineapple
367–64-6	Acetic acid	2.6	3097.50	3510.25	2401.56	1191.35	1350.10	923.68	Acetic acid
98–01-1	Furfural	0.15	78.84	91.87	51.88	525.60	612.48	345.89	Almond scent
1124-11-4	Tetramethylpyrazine	80.073	19.14	13.89	7.77	0.24	0.17	0.10	Nutty
100–52-7	Benzaldehyde	4.203	0.79	0.92	0.50	0.19	0.22	0.12	Bitter almond flavor
79–09-4	n-Propionic acid	20	33.82	51.10	17.65	1.69	2.55	0.88	Sour flavor
123,513–85-9	2,3-Butanediol	0.141	25.03	36.65	28.35	177.49	259.96	201.06	Fruity
57–55-6	1,2-Propanediol	–	106.88	102.02	46.58	–	–	–	–
123–25-1	Diethyl succinate	353.193	2.67	2.92	3.92	0.01	0.01	0.01	Fruity
64,118–37-2	Valeric acid	0.5	1.95	3.13	1.71	3.90	6.25	3.41	Fatty odor, slightly sour
142–62-1	Hexanoic acid	8.6	3.95	5.09	2.86	0.46	0.59	0.33	Sweat odor, slightly sour
98–85-1	2-Phenylethanol	0.01	10.09	9.95	12.54	1008.50	995.17	1253.71	Rose scent, honey odor
111–14-8	Heptanoic acid	70.5	0.19	0.79	0.29	0.00	0.01	0.00	Pungent, sour, fatty odor
124–06-1	Ethyl Tetradecanoate	1.89	1.34	1.53	1.77	0.71	0.81	0.94	Sweet
111–62-6	Ethyl oleate	0.43	18.52	21.42	31.89	43.06	49.81	74.16	Soap odor
544–35-4	Ethyl linoleate	0.45	62.96	75.41	97.26	139.92	167.58	216.14	Plastic

odor (C: 97.2 > B: 75.4 > A: 62.9), and isoamyl alcohol with an oily odor (C: 42.3 > B: 36.3 > A: 32.3) all showed a negative correlation between their OAV and base *Baijiu* quality grades (Table 1). Hence, ethyl oleate, ethyl linoleate, and isoamyl alcohol might be key aroma compounds responsible for the decrease in cleanliness scores of base *Baijiu*.

3.3. Development of classification and prediction models for base baijiu quality grades using machine learning algorithms

A total of 52 major flavor compounds from first-round base *Baijiu* and 7 physicochemical parameters of *Jiupei* were selected as candidate features for Lasso regression feature selection. The optimal λ was determined to be 0.06, with non-zero λ values considered the best features (Fig. 4a, b) (Shaon et al., 2024). Fifteen optimal features were identified, including 13 flavor compounds: isobutyraldehyde, acetal, sec-butanol, isobutanol, isoamyl acetate, ethyl valerate, ethyl lactate, 2,3,5-trimethylpyrazine, acetic acid, ethyl decanoate, diethyl succinate, caprylic acid, and ethyl linoleate. Additionally, 2 physicochemical parameters, reducing sugars and starch, were selected (Fig. 4c). Eleven classification models, including MLP, Gradient Boosting, Bagging, and Extra Trees, were used to construct models for different quality grades of base *Baijiu* (Fig. 1). The classification performance was evaluated using AUC, F1-Score, Accuracy, Recall, and Precision (Table S2). Gradient Boosting (AUC = 0.79), Extra Trees (AUC = 0.75), and MLP (AUC = 0.68) etc. showed no signs of overfitting (Fig. 4e). Based on the evaluation scores, Gradient Boosting was selected as the optimal model for classifying different quality grades of base *Baijiu*. SHAP analysis ranked isobutyraldehyde, ethyl lactate, acetal, isobutanol, and reducing sugars as the most important features (Fig. 4f). Isobutyraldehyde, acetal, reducing sugars, and acetic acid had the greatest impact on predicting grade A, while isobutanol was the key compound influencing predictions for grades B and C (Fig. 4g). Previous studies have shown that

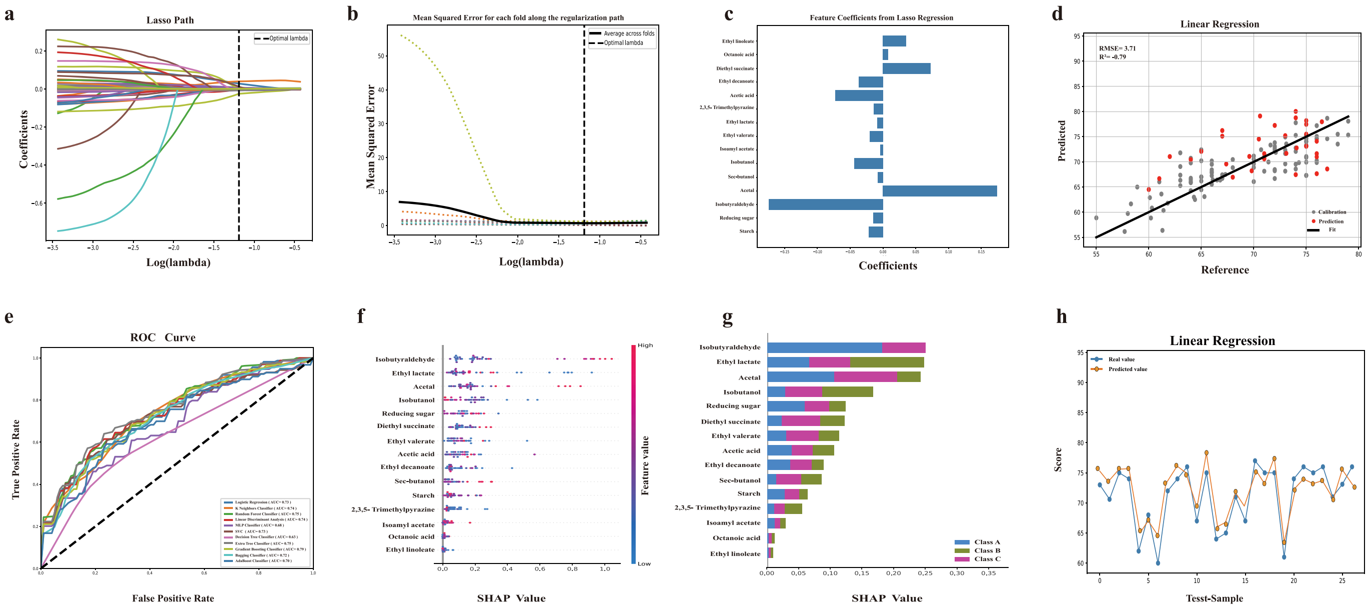

Fig. 4. Lasso feature selection for different base *Baijiu* quality grades: (a) coefficient path plot; (b) optimal coefficient distribution plot; (c) bar chart of optimal feature coefficients. Regression prediction models for different base *Baijiu* quality grades: (d) goodness-of-fit plot for the linear regression model; (h) comparison of predicted versus actual values in regression prediction models. Classification models for different base *Baijiu* quality grades: (e) ROC curves; (g), (f) SHAP feature importance ranking plots.

aldehydes in *Baijiu* are primarily formed by the oxidation of alcohols, and that excessive higher alcohols can produce fusel oil off-flavors, negatively impacting *Baijiu*'s aroma quality (Wei et al., 2024). Therefore, we hypothesize that the quality grade of base *Baijiu* may be related to its ability to oxidize alcohols into aldehydes. Extending storage time and using more breathable ceramic jars in production could enhance oxidation and improve the quality grade of base *Baijiu* (Wei et al., 2023). Furthermore, machine learning methods have reinforced that acetic acid, with its volatile sour taste, is likely a key compound influencing the sourness scores across different quality grades of base *Baijiu* (Dong et al., 2024). This analysis also confirmed that reducing sugar content in *Jiupei* is fundamental to the high quality of base *Baijiu* (Fig. 4g). Finally, 9 machine learning regression models, including Linear Regression, Gradient Boosting, Bagging, and Extra Trees etc., were employed to predict the quality grades of base *Baijiu*. The models were evaluated using RMSE and R^2 (Fig. S4a-h). Based on the evaluation metrics, Linear Regression ($R^2 = -0.79$, RMSE = 3.71) was selected as the best model for predicting quality grades (Fig. 4d, h).

3.4. Analysis of microbial community composition and physicochemical differences in Jiupei across different quality grades of base baijiu

3.4.1. Physicochemical differences in Jiupei

The quality differences in base *Baijiu* are closely related to the microbial community structure and metabolic regulation during the *Jiupei* fermentation process (Wang, Wu, et al., 2021; Zhang, Wang, Tan, et al., 2021). ANOVA was used to analyze physicochemical differences in *Jiupei* from 134 base *Baijiu* samples of varying quality grades. Significant differences were found in ethanol content (C-JP: 5.33 ± 3.04 mg/g > B-JP: 4.85 ± 3.41 mg/g > A-JP: 4.51 ± 3.05 mg/g) and titratable acidity (B-JP: 1.06 ± 0.35 mmol/10 g > A-JP: 0.92 ± 0.17 mmol/10 g > C-JP: 0.79 ± 0.21 mmol/10 g) ($p < 0.05$) (Fig. 5e, d). Previous studies have indicated that the acidity of *Jiupei* is a primary source of the Sour taste in base *Baijiu*, which may explain the prominent Sour taste in higher-quality grades (Deng et al., 2020). Additionally, significant differences in reducing sugar content were observed in *Jiupei* across different quality grades of base *Baijiu* (A-JP: 13.84 ± 3.80 mg/g > B-JP: 13.34 ± 3.55 mg/g > C-JP: 11.59 ± 3.93 mg/g) ($p < 0.05$) (Fig. 5b). Correlation analysis using the Mantel test revealed a strong positive correlation between reducing sugar content and base *Baijiu* quality grades ($r > 0.6$, $p < 0.05$) (Fig. 5h). Previous studies suggest that reducing sugar content during fermentation reflects the saccharification and fermentation rates of *Jiupei* (Wang, Wu, et al., 2021). High reducing sugar content in A-JP ensures microbial utilization of fermentation substrates during subsequent pit fermentation, contributing to the production of aroma compounds and stable yields (Wang, Wu, et al., 2021; Yang et al., 2024). Therefore, differences in reducing sugar content may lead to variations in base *Baijiu* quality grades.

3.4.2. Microbial community composition differences in Jiupei

Metagenomic analysis detected a total of 475 bacterial genera and 135 fungal genera in *Jiupei* across different quality grades of base *Baijiu*. Permutational multivariate analysis (Permanova) combined with unsupervised principal component analysis (PCA) revealed significant differences in the bacterial ($p = 0.005$, $R = 0.8864$) and fungal ($p = 0.004$, $R = 0.7534$) communities of *Jiupei* across different quality grades. PCA explained 89.98 % and 89.24 % of the variance, respectively, indicating distinct microbial community structures in *Jiupei* associated with different quality grades (Fig. 6c, f).

At the bacterial genus level, *Acetobacter*, *Bacillus*, *Komagataeibacter*, *Kroppenstedtia*, *Staphylococcus*, *Lactobacillus*, *Corynebacterium*, *Weizmannia*, *Pediococcus*, and *Thermoactinomyces* were the dominant bacterial genera in first-round *Jiupei* (relative abundance >1 %) (Fig. 6b). *Acetobacter* was the most abundant genus in first-round *Jiupei* (relative abundance >60 %). *Acetobacter* oxidizes ethanol to acetic acid, which contributes to the acidity of sauce-flavor *Baijiu* (He et al., 2023). Therefore, we hypothesize that the typical Sour taste of first-round base *Baijiu* is related to the dominance of *Acetobacter*. Notably, *Acetobacter* had the highest relative abundance in A-JP (A-JP: 71.13 % > C-JP: 64.73 % > B-JP: 61.13 %), which may explain the higher Sour taste rating in A-JP (Fig. 2a). LEfSe analysis identified 16 bacterial biomarkers in *Jiupei* associated with different quality grades of base *Baijiu* (LDA > 2, $p < 0.05$). Eight bacterial biomarkers were found in A-JP, including *Corynebacterium*, *Weizmannia*, *Lactiplantibacillus*, *Levilactobacillus*, *Limosilactobacillus*, *Komagataeibacter*, *Mixta*, and *Acetobacter*. B-JP had four biomarkers: *Mammaliicoccus*, *Lactobacillus*, *Lentilactobacillus*, and

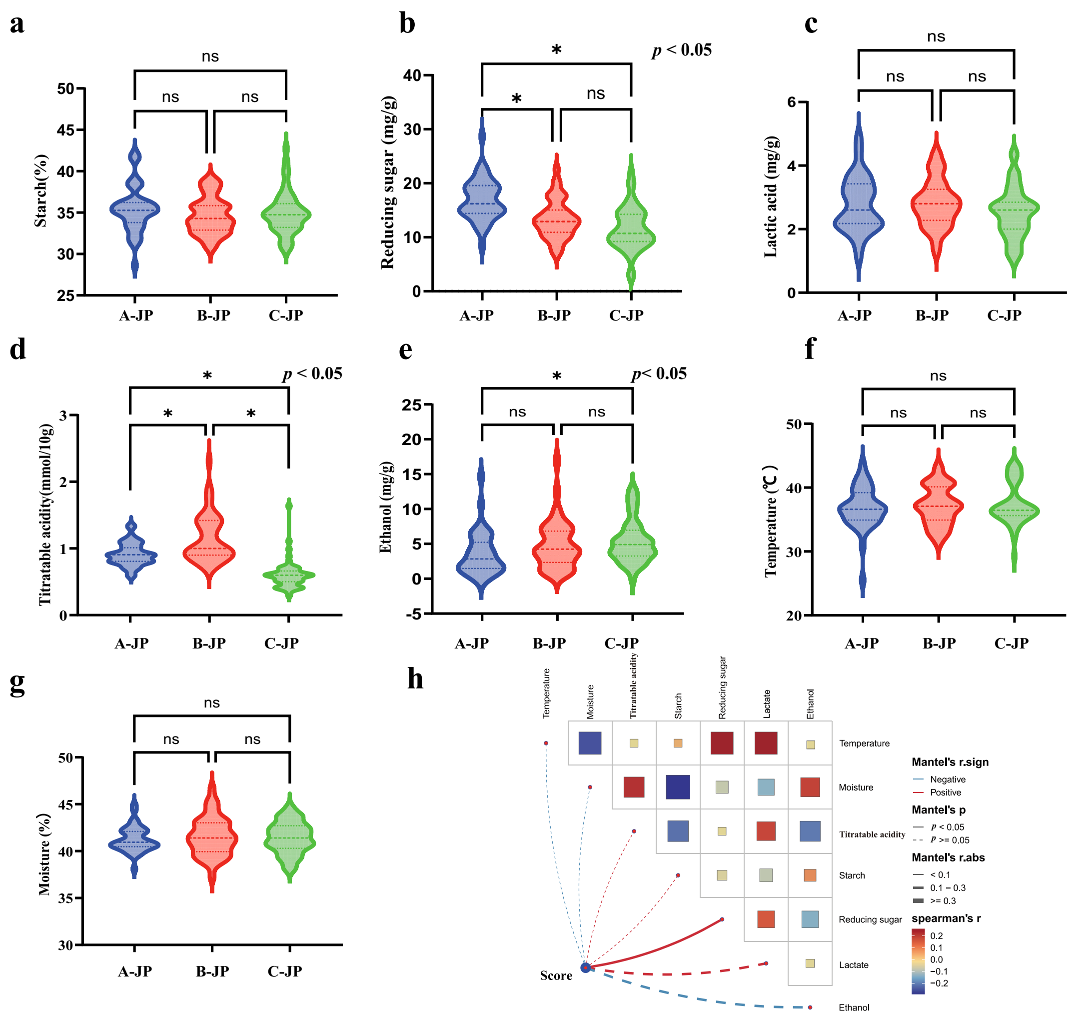

Fig. 5. Analysis of physicochemical differences in *Jiupei* across different base *Baijiu* quality grades: (a) starch; (b) reducing sugar; (c) lactic acid; (d) titratable acidity; (e) ethanol; (f) temperature; (g) moisture. (h) Mantel test analysis of the correlation between *Jiupei* physicochemical properties and quality scores across different *Baijiu* grades. The *, ** and *** indicate statistical significance at $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively.

Ligilactobacillus. C-JP had four biomarkers: *Caldibacillus*, *Leuconostoc*, *Lactococcus*, and *Saccharomonospora* (Fig. 6a). Notably, the bacterial biomarkers in *Jiupei* were dominated by acid-producing bacteria such as *Lactiplantibacillus*, *Lactobacillus*, and *Lactococcus*, which may be the primary contributors to the typical Sour taste of first-round base *Baijiu*. Additionally, previous studies have reported that *Komagataeibacter* enhances the stability and functionality of microbial ecosystems in traditional solid-state fermentation (Peng et al., 2021). In summary, the bacterial community, particularly *Acetobacter*, *Lactobacillus*, and *Lactococcus*, plays a key role in shaping the flavor profile of sauce-flavor *Baijiu*. These acid-producing bacteria contribute significantly to the production of acetic acid, which drives the sour taste in high-quality *Baijiu*. Additionally, *Komagataeibacter* enhances fermentation stability, ensuring consistent flavor compound production. The dominance of these bacteria in A-JP ensures a higher quality *Baijiu* with distinct sourness and cleanliness, critical characteristics that differentiate superior *Baijiu* from lower grades.

At the fungal genus level, *Pichia*, *Saccharomyces*, *Schizosaccharomyces*, *Zygosaccharomyces*, *Monascus*, *Torulaspora*, *Aspergillus*, *Lichtheimia*, *Wickerhamomyces*, and *Rhizopus* were dominant in *Jiupei* across different quality grades of base *Baijiu* (relative abundance >1 %) (Fig. 6e). Previous studies have identified *Pichia* and *Saccharomyces* as key ethanol-producing genera during sauce-flavor *Baijiu* fermentation (Kang et al., 2024; Wang, Huang, & Huang, 2021). The combined relative abundance of these genera was lowest in A-JP (A-JP: 71.6 % < B-JP: 94.63 % < C-JP: 96.11 %), which may account for the significant differences in ethanol content among the quality grades (C-JP: 5.33 ± 3.04 mg/g > B-JP: 4.85 ± 3.41 mg/g > A-JP: 4.51 ± 3.05 mg/g) (Fig. 5e). High ethanol concentrations in fermented *Jiupei* are likely to inhibit the growth and metabolism of other microorganisms, leading to a simplified microbial community structure (Pan et al., 2023). This is consistent with the observation that A-JP exhibits significantly higher α-diversity in both bacterial and fungal communities compared to other quality grades (Fig. S5a, b). In A-JP, *Schizosaccharomyces* (A-JP: 4.73 % > C-JP: 2.90 % > B-JP: 2.52 %) and *Zygosaccharomyces* (A-JP: 7.10 % > B-JP: 1.28 % > C-JP: 0.94 %) had the highest relative abundances. Studies have shown that these non-Saccharomyces yeasts play crucial roles in sauce-flavor *Baijiu* fermentation, enhancing the production of aromatic compounds and ensuring stable yields (Du, Song, et al., 2021; Xu et al., 2017). LEfSe analysis identified seven fungal biomarkers associated with different quality grades of base *Baijiu* (LDA > 2, $p < 0.05$). A-JP was characterized by *Aspergillus*, *Torulaspora*, and *Monascus*; B-JP had *Pichia* and *Schizosaccharomyces*; C-JP was defined by *Saccharomyces* and *Lichtheimia* (Fig. 6d). Research has demonstrated that *Aspergillus* and *Monascus* in A-JP are prolific producers of amylase during sauce-flavor *Baijiu* fermentation, which breaks down starch into reducing sugars(Pan et al., 2023; Wang, Wu, et al., 2021). This likely explains the significantly higher reducing sugar content in A-JP

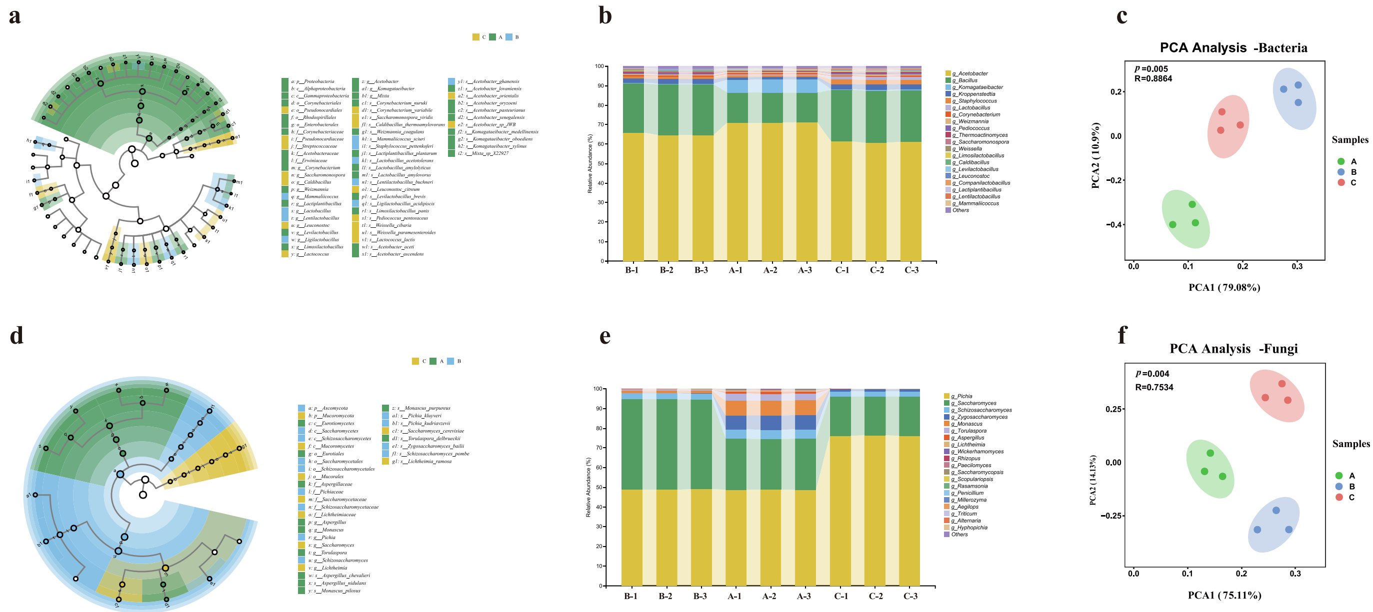

Fig. 6. LEfSe analysis identifying key microorganisms in *Jiupei* of different base *Baijiu* quality grades: (a) bacteria; (d) fungi. Microbial community structure of *Jiupei* in different *Baijiu* quality grades: (b) bacteria; (e) fungi. Principal component analysis of microbial communities in *Jiupei* across different Baijiu quality grades: (c) bacteria; (f) fungi.

compared to other *Jiupei* samples (A-JP: 13.84 ± 3.80 mg/g > B-JP: 13.34 ± 3.55 mg/g > C-JP: 11.59 ± 3.93 mg/g) (Fig. 5b). *Torulaspora* has been reported to enhance fruity and nutty aromas in alcoholic beverages (Pan et al., 2023). Overall, the fungal community, particularly *Aspergillus*, *Monascus,* and *Torulaspora,* plays a vital role in shaping the flavor profile of sauce-flavor *Baijiu*. These fungi are key contributors to amylase production, which increases reducing sugars and enhances fermentation efficiency. *Torulaspora* contributes to fruity and nutty aromas, while *Aspergillus* and *Monascus* support the production of key flavor compounds. The higher abundance of these genera in A-JP leads to improved aroma complexity, distinguishing higher-quality *Baijiu* from lower grades.

4. Conclusion

This study comprehensively reveals the factors contributing to the quality grade differences in first-round base *Baijiu* of sauce-flavor *Baijiu*. QDA analysis indicated significant differences across five dimensions, including Sour taste and Cleanliness, among base *Baijiu* of different quality grades ($p < 0.05$). GC-FID combined with OAV analysis identified acetic acid (A: 1491.3 > B: 1350.0 > C: 923.67) and propionic acid (B: 2.55 > A: 1.69 > C: 0.88) as the key flavor compounds responsible for the Sour taste differences among base *Baijiu* of different quality grades. Ethyl linoleate (C: 97.2 > B: 75.4 > A: 62.9), ethyl oleate (C: 97.2 > B: 75.4 > A: 62.9), and isoamyl alcohol (C: 42.3 > B: 36.3 > A: 32.3) were identified as key compounds contributing to the decline in Cleanliness across different quality grades. Additionally, the reducing sugar content in *Jiupei* was found to significantly impact the quality of base *Baijiu* ($p < 0.05$, $r > 0.6$), which correlates with the presence of high amylase-producing *Aspergillus* and *Monascus* in A-JP as revealed by LEfSe analysis (LDA > 2, $p < 0.05$). Finally, 11 machine learning classification models and 9 regression models were used to develop classification and prediction models for base *Baijiu* quality grades. Gradient Boosting (AUC = 0.79) and Linear Regression ($R^2 = -0.79$) were selected as the best-performing models. SHAP model revealed that isobutyraldehyde and acetal were the most influential features for predicting first-grade base *Baijiu*, while isobutanol was the key compound for distinguishing other grades. Therefore, we recommend extending the storage time of base *Baijiu* and using more breathable ceramic jars during production to enhance the oxidation of higher alcohols, thereby improving the quality grade. This study aims to improve the quality grade of sauce-flavor *Baijiu* base *Baijiu* and provide a scientific basis for establishing a more objective and accurate evaluation system.

Ethical statement

Ethical review and approval were waived for this study, because experimental samples used in the study are consumed in daily life.

CRediT authorship contribution statement

Shuai Li: Data curation, Formal analysis, Methodology, Software, Visualization, Writing - original draft. **Tao Li:** Writing – review & editing, Visualization, Software, Data curation. **Yueran Han:** Writing – review & editing, Project administration, Funding acquisition. **Pei Yan:** Writing – review & editing. **Guohui Li:** Writing – review & editing. **Tingting Ren:** Writing – review & editing. **Ming Yan:** Supervision, Funding acquisition. **Jun Lu:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition. **Shuyi Qiu:** Writing – review & editing, Supervision, Resources, Project administration, Methodology, Funding acquisition, 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.

Data availability

Data will be made available on request.

Acknowledgment

Guizhou Provincial Science and Technology Project (Qiankehe Achievement -ZK [2023] General 149), Guizhou Provincial Science and Technology Project (Qiankehe Achievement - ZK[2023] General 150).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochx.2024.101877.

References

Al-Zaiti, S. S., Martin-Gill, C., Zègre-Hemsey, J. K., Bouzid, Z., Faramand, Z., Alrawashdeh, M. O., Gregg, R. E., Helman, S., Riek, N. T., Kraevsky-Phillips, K., Clermont, G., Akcakaya, M., Sereika, S. M., van Dam, P., Smith, S. W., Birnbaum, Y., Saba, S., Sejdic, E., & Callaway, C. W. (2023). Machine learning for ECG diagnosis and risk stratification of occlusion myocardial infarction. *Nature Medicine, 29*(7), 1804–1813. https://doi.org/10.1038/s41591-023-02396-3

Cao, R., Ma, Y., Zhou, Q., Du, H., Jin, G., Yu, X., & Xu, Y. (2024). Mucorpepsin from rhizomucor pusillus relates the quality of medium-temperature daqu. *Food Bioscience, 59*, Article 104027. https://doi.org/10.1016/j.fbio.2024.104027

Catacutan, D. B., Alexander, J., Arnold, A., & Stokes, J. M. (2024). Machine learning in preclinical drug discovery. *Nature Chemical Biology, 20*(8), 960–973. https://doi.org/10.1038/s41589-024-01679-1

Chen, H., Covert, I. C., Lundberg, S. M., & Lee, S. (2023). Algorithms to estimate shapley value feature attributions. *Nature Machine Intelligence, 5*(6), 590–601. https://doi.org/10.1038/s42256-023-00657-x

Chen, X., Wu, Y., Zhu, H., Wang, H., Lu, H., Zhang, C., Li, X., Xu, Y., Li, W., & Wang, Y. (2022). Turning over fermented grains elevating heap temperature and driving microbial community succession during the heap fermentation of sauce-flavor baijiu. *Lwt, 172*, Article 114173. https://doi.org/10.1016/j.lwt.2022.114173

Deng, N., Du, H., & Xu, Y. (2020). Cooperative response of Pichia kudriavzevii and Saccharomyces cerevisiae to lactic acid stress in baijiu fermentation. *Journal of Agricultural and Food Chemistry, 68*(17), 4903–4911. https://doi.org/10.1021/acs.jafc.9b08052

Dong, W., Dai, X., Jia, Y., Ye, S., Shen, C., Liu, M., … Deng, B. (2024). Association between baijiu chemistry and taste change: Constituents, sensory properties, and analytical approaches. *Food Chemistry, 437*, Article 137826. https://doi.org/10.1016/j.foodchem.2023.137826

Du, H., Ji, M., Xing, M., Wang, X., & Xu, Y. (2021). The effects of dynamic bacterial succession on the flavor metabolites during baijiu fermentation. *Food Research International, 140*, Article 109860. https://doi.org/10.1016/j.foodres.2020.109860

Du, H., Song, Z., Zhang, M., Nie, Y., & Xu, Y. (2021). The deletion of Schizosaccharomyces pombe decreased the production of flavor-related metabolites during traditional baijiu fermentation. *Food Research International, 140*, Article 109872. https://doi.org/10.1016/j.foodres.2020.109872

Duan, J., Yang, S., Li, H., Qin, D., Shen, Y., Li, H., Sun, J., Zheng, F., & Sun, B. (2022). Why the key aroma compound of soy sauce aroma type baijiu has not been revealed yet? *Lwt, 154*, Article 112735. https://doi.org/10.1016/j.lwt.2021.112735

Gong, J., Ma, Y., Li, L., Cheng, Y., & Huang, Y. (2023). Comparative characterization and contribution of key aroma compounds in the typical base liquor of Jiang-flavor baijiu from different distributions in the Chinese Chishui River basin. *Food Chemistry: X, 20*, Article 100932. https://doi.org/10.1016/j.fochx.2023.100932

He, F., Yang, S., Zhang, G., Xu, L., Li, H., Sun, J., … Sun, B. (2022). Exploration of key aroma active compounds in strong flavor baijiu during the distillation by modern instrument detection technology combined with multivariate statistical analysis methods. *Journal of Food Composition and Analysis, 110*, Article 104577. https://doi.org/10.1016/j.jfca.2022.104577

He, G., Xie, F., Ren, X., Yin, B., Du, L., Wei, Y., & Zhou, J. (2023). Revealing the functional microbiota for acetic acid formation in daqu starter for Chinese nong-Xiang baijiu fermentation. *Food Bioscience, 53*, Article 102782. https://doi.org/10.1016/j.fbio.2023.102782

Kang, J., Huang, X., Li, R., Zhang, Y., Chen, X., & Han, B. (2024). Deciphering the core microbes and their interactions in spontaneous baijiu fermentation: A comprehensive review. *Food Research International, 188*, Article 114497. https://doi.org/10.1016/j.foodres.2024.114497

Li, H., Zhang, X., Gao, X., Shi, X., Chen, S., Xu, Y., & Tang, K. (2023). Comparison of the aroma-active compounds and sensory characteristics of different grades of light-flavor baijiu. *Foods, 12*(6), 1238. https://doi.org/10.3390/foods12061238

Liu, H., & Sun, B. (2018). Effect of fermentation processing on the flavor of baijiu. *Journal of Agricultural and Food Chemistry, 66*(22), 5425–5432. https://doi.org/10.1021/acs.jafc.8b00692

Liu, Q., He, D., Ma, Y., Wang, H., Li, Y., Cheng, Y., & Huang, Y. (2023). Sensory profile and the contribution of key aroma compounds in jiang-flavor rounded-base baijiu produced in the Chishui river basin. *Lwt, 189*, Article 115474. https://doi.org/10.1016/j.lwt.2023.115474

Liu, Q., Zhang, X., Zheng, L., Meng, L., Liu, G., Yang, T., … Xu, Z. (2023). Machine learning based age-authentication assisted by chemo-kinetics: Case study of strong-flavor Chinese baijiu. *Food Research International, 167*, Article 112594. https://doi.org/10.1016/j.foodres.2023.112594

Menichetti, G., Ravandi, B., Mozaffarian, D., & Barabási, A. (2023). Machine learning prediction of the degree of food processing. *Nature Communications, 14*(1), 2312. https://doi.org/10.1038/s41467-023-37457-1

Pan, F., Qiu, S., Lv, Y., & Li, D. (2023). Exploring the controllability of the baijiu fermentation process with microbiota orientation. *Food Research International, 173*, Article 113249. https://doi.org/10.1016/j.foodres.2023.113249

Peng, M., Zhang, X., Huang, T., Zhong, X., Chai, L., Lu, Z., Shi, J., & Xu, Z. (2021). Komagataeibacter europaeus improves community stability and function in solid-state cereal vinegar fermentation ecosystem: Non-abundant species plays important role. *Food Research International, 150*, Article 110815. https://doi.org/10.1016/j.foodres.2021.110815

Qiao, L., Wang, J., Wang, R., Zhang, N., & Zheng, F. (2023). A review on flavor of baijiu and other world-renowned distilled liquors. *Food Chemistry: X, 20*, Article 100870. https://doi.org/10.1016/j.fochx.2023.100870

Schreurs, M., Piampongsant, S., Roncoroni, M., Cool, L., Herrera-Malaver, B., Vanderaa, C., Theßeling, F. A., Kreft, A., Botzki, A., Malcorps, P., Daenen, L., Wenseleers, T., & Verstrepen, K. J. (2024). Predicting and improving complex beer flavor through machine learning. *Nature Communications, 15*(1), 2368. https://doi.org/10.1038/s41467-024-46346-0

Shaon, M. S. H., Karim, T., Shakil, M. S., & Hasan, M. Z. (2024). A comparative study of machine learning models with LASSO and SHAP feature selection for breast cancer prediction. *Healthcare Analytics, 6*, Article 100353. https://doi.org/10.1016/j.health.2024.100353

Wang, B., Wu, Q., Xu, Y., & Sun, B. (2021). Multiple sugars promote microbial interactions in Chinese baijiu fermentation. *Lwt, 138*, Article 110631. https://doi.org/10.1016/j.lwt.2020.110631

Wang, H., Huang, Y., & Huang, Y. (2021). Microbiome diversity and evolution in stacking fermentation during different rounds of Jiang-flavoured baijiu brewing. *Lwt, 143*, Article 111119. https://doi.org/10.1016/j.lwt.2021.111119

Wang, L., Fan, S., Yan, Y., Yang, L., Chen, S., & Xu, Y. (2020). Characterization of potent odorants causing a pickle-like off-odor in moutai-aroma type baijiu by comparative aroma extract dilution analysis, quantitative measurements, aroma addition, and omission studies. *Journal of Agricultural and Food Chemistry, 68*(6), 1666–1677. https://doi.org/10.1021/acs.jafc.9b07238

Wang, L., Tang, P., Zhang, P., Lu, J., Chen, Y., Xiao, D., & Guo, X. (2024). Unraveling the aroma profiling of baijiu: Sensory characteristics of aroma compounds, analytical approaches, key odor-active compounds in different baijiu, and their synthesis mechanisms. *Trends in Food Science & Technology, 146*, Article 104376. https://doi.org/10.1016/j.tifs.2024.104376

Wei, J., Du, H., & Xu, Y. (2024). Revealing the key microorganisms producing higher alcohols and their assembly processes during jiang-flavor baijiu fermentation. *Food Bioscience, 61*, Article 104569. https://doi.org/10.1016/j.fbio.2024.104569

Wei, L., Hu, J., Pan, C., Cheng, P., Zhang, J., Xi, D., Chen, M., Lu, L., Lu, H., & Hu, F. (2023). Effects of different storage containers on the flavor characteristics of jiangxiangxing baijiu. *Food Research International, 172*, Article 113196. https://doi.org/10.1016/j.foodres.2023.113196

Wu, J., Chen, R., Li, X., Fu, Z., Xian, C., Zhao, W., Zhao, C., & Wu, X. (2023). Comprehensive identification of key compounds in different quality grades of soy sauce-aroma type baijiu by HS-SPME-GC-MS coupled with electronic nose. *Frontiers in Nutrition, 10*. https://doi.org/10.3389/fnut.2023.1132527

Wu, J., Liu, Y., Zhao, H., Huang, M., Sun, Y., Zhang, J., & Sun, B. (2021). Recent advances in the understanding of off-flavors in alcoholic beverages: Generation, regulation, and challenges. *Journal of Food Composition and Analysis, 103*, Article 104117. https://doi.org/10.1016/j.jfca.2021.104117

Xu, Y., Zhi, Y., Wu, Q., Du, R., & Xu, Y. (2017). Zygosaccharomyces bailii is a potential producer of various flavor compounds in Chinese maotai-flavor liquor fermentation. *Frontiers in Microbiology, 8*. https://doi.org/10.3389/fmicb.2017.02609

Yang, L., Huang, X., Hu, J., Deng, H., He, J., & Zhang, C. (2024). The spatiotemporal heterogeneity of microbial community assembly during pit fermentation of soy sauce flavor baijiu. *Food Bioscience, 61*, Article 104438. https://doi.org/10.1016/j.fbio.2024.104438

Yang, L., Xian, C., Li, P., Wang, X., Song, D., Zhao, L., & Zhang, C. (2023). The spatio-temporal diversity and succession of microbial community and its environment driving factors during stacking fermentation of maotai-flavor baijiu. *Food Research International, 169*, Article 112892. https://doi.org/10.1016/j.foodres.2023.112892

Zhang, H., Wang, L., Tan, Y., Wang, H., Yang, F., Chen, L., … Xu, Y. (2021). Effect of pichia on shaping the fermentation microbial community of sauce-flavor baijiu. *International Journal of Food Microbiology, 336*, Article 108898. https://doi.org/10.1016/j.ijfoodmicro.2020.108898

Zhang, H., Wang, L., Wang, H., Yang, F., Chen, L., Hao, F., … Xu, Y. (2021). Effects of initial temperature on microbial community succession rate and volatile flavors during baijiu fermentation process. *Food Research International, 141*, Article 109887. https://doi.org/10.1016/j.foodres.2020.109887

Zhao, W., Liu, Y., Zhou, K., Li, K., & Shen, J. (2024). How sorghum grain composition affects the quality of Chinese baijiu - a comprehensive review. *Journal of Food Composition and Analysis, 134*, Article 106512. https://doi.org/10.1016/j.jfca.2024.106512

Zhou, R., Chen, X., Huang, M., Chen, H., Zhang, L., Xu, D., … Dai, X. (2023). ATR-FTIR spectroscopy combined with chemometrics to assess the spectral markers of irradiated baijius and their potential application in irradiation dose control. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy, 302*, Article 123162. https://doi.org/10.1016/j.saa.2023.123162