Food Chemistry: X 35 (2026) 103756

Contents lists available at ScienceDirect

Food Chemistry: X

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

Comprehensive insights into the mechanism of flavor formation in *Cheonggukjang*: Integration of metagenomics, volatomics, and metabolomics

Shanshan Zhao, Yuhang Sai, Mengyu Jia, Yu Qiao, Weishu Guo, Wenjing Ding, Xin Shao, Yan Zheng *

College of Food Science, Shenyang Agricultural University, Shenyang 110866, China

ARTICLE INFO

Keywords:
Cheonggukjang
Flavor compounds
Flavor formation mechanism
Multi-omics
Metabolic network

ABSTRACT

Microbial metabolism shapes the unique flavor profile of *Cheonggukjang*; however, the formation pathways of characteristic flavor compounds mediated by microbiota remain unclear, hindering precise quality control. To fill this gap, this study innovatively integrated metagenoics, volatilomics, and metabolomics to systematically decode the flavor formation mechanism during *Cheonggukjang* fermentation. Volatile compound analysis defined three fermentation stages for *Cheonggukjang* (0–18, 18–60, and 60–72 h), identifying the 60–72 h period as the most critical for flavor formation. A total of 15 key flavor compounds were identified, with 10 designated stage-specific flavor markers. LefSe analysis revealed that *Bacillus subtilis*, *Bacillus velezensis*, *Caldibacillus thermoamylovorans*, and *Bacillus licheniformis* were the key biomarkers across different fermentation stages, while redundancy analysis (RDA) indicated that total sugar as the key driver of microbial succession. Additionally, this study reconstructed the metabolic network responsible for characteristic flavor formation and identified *C. thermoamylovorans*, *B. licheniformis*, *B. velezensis*, *B. subtilis*, *Bacillus paralicheniformis*, and *Caldibacillus hisashii* as core functional microbiota modulating amino acid metabolic to drive flavor development. This study lays a theoretical framework for standardizing *Cheonggukjang* production and targeted regulating its flavor quality.

1. Introduction

Cheonggukjang, a traditional fermented soybean product in South Korea and ethnic Korean communities in China, is renowned for its unique flavor and various beneficial health effects, including immunity modulation, anti-inflammatory, antioxidant, antihypertensive, and antidiabetic effects (Kim et al., 2012; Lee et al., 2024). Flavor, a critical determinant of food quality and consumer preference, arises from microbial enzymatic activity during fermentation (Regueiro et al., 2017). These microbes enzymatically degrade soybean carbohydrates, proteins, and lipids to generate flavor-enhancing metabolites (such as amino acids, fatty acids, and their metabolites) and bioactive components (including γ-aminobutyric acid and poly-γ-glutamic acid), which collectively determine the flavor and quality of *Cheonggukjang* (Cho et al., 2017; Jin & Mah, 2024; Nam et al., 2012; Sanjukta & Rai, 2016).

Microbial composition and dynamic interactions are pivotal in shaping the flavor profile of fermented foods (Tang & Peng, 2024). In *Cheonggukjang*, microbiota historically originate from two primary sources: natural microbiota from rice straw or water in traditional methods, and standardized starter cultures in modern industrial production (Hong et al., 2024). A diverse array of *Bacillus* species, including *Bacillus amyloliquefaciens*, *Bacillus subtilis*, *Bacillus licheniformis*, and *Bacillus velezensis*, dominate spontaneous fermentation. *B. subtilis* is widely employed as a starter culture in modern industrial production, effectively reducing biogenic amines and exhibiting antimicrobial properties (Lee & Chang, 2017; Lim, 2022). While spontaneous fermentation promotes the development of favorable flavor profiles by the interaction of various microorganisms (Wen et al., 2024), it often results in inconsistent metabolite profiles and batch variability, hindering standardization and scalability (Nam et al., 2012). To address these challenges, industrial *Cheonggukjang* production has shifted toward starter culture inoculation, ensuring more predictable fermentation outcomes (Nam et al., 2012). However, this approach lacks synergistic interactions between different microbial taxa, which are essential for producing the broad-spectrum flavor characteristic of traditional *Cheonggukjang*. Constructing synthetic microbial consortia from traditional fermentation-derived

* Corresponding author.
E-mail address: zhengyan0403@163.com (Y. Zheng).

https://doi.org/10.1016/j.fochx.2026.103756
Received 25 November 2025; Received in revised form 8 March 2026; Accepted 11 March 2026
Available online 12 March 2026

core microorganisms offers an ideal strategy to resolve this dilemma, enabling both controllable fermentation and complex flavor profiles in industrial production. The successful implementation of this strategy, however, hinges on a fundamental prerequisite: identifying the core functional microbiota responsible for flavor formation and elucidating the mechanisms by which they generate key flavor compounds during natural fermentation.

In recent years, multi-omics approaches (e.g., metagenomics and metabolomics) have provided powerful tools for elucidating complex microbial-substrate interactions and their role in shaping the flavor quality of fermented products (Jian et al., 2024; Qi et al., 2025; Wang et al., 2024; Wen et al., 2024). By integrating metagenomic, volatilomic, and metabolomic data, researchers can trace flavor-formation pathways from microbial activity to final metabolic products, facilitating targeted flavor optimization and process innovation. This holistic strategy overcomes inherent limitations of single-omics technologies, such as narrow data scope, inadequate mechanistic insight, and limited practical potential (Shahid et al., 2025). Current research on flavor and functional qualities of *Cheonggukjang* has predominantly investigated three key aspects: (1) microbial composition and functional prediction using metagenomics (Hong et al., 2024; Tamang et al., 2022), (2) relationships between microbiota and bioactive compounds such as γ-aminobutyric acid and poly-γ-glutamic acid (Jin & Mah, 2024), and (3) temporal dynamics of flavor-related metabolites during fermentation (Cho et al., 2017; Park et al., 2010). For instance, Tamang et al. (2022) used shotgun metagenomics to study the microbial community of *Cheonggukjang* from different regions, predicted the microbial functions, and revealed the genes for the synthesis and metabolism of a wide range of bioactive compounds (including various essential amino acids, conjugated amino acids, different vitamins, flavonoids, and enzymes). Cho et al. (2017) explored the influence of *B. subtilis* CSY191 on the fatty acid and volatile compound profiles of *Cheonggukjang* (prepared using new Korean brown soybean cultivars, protein-rich cultivar, and oil-rich cultivar). They found that volatile profiles in *Cheonggukjang* changed significantly over time, and different cultivars represented specific volatile profiles. However, these studies have focused either on microbial composition or on changes in metabolites, and have failed to establish a causal mechanistic link between specific microbial communities and the key flavor compounds they produce. Accordingly, the mechanism governing flavor formation and its association with microbial functions—the theoretical basis for targeted quality control—remains unclear.

To fill this gap, this study adopts a multi-omics strategy integrating metagenomics, volatilomics, and metabolomics to systematically investigate the flavor formation mechanism during *Cheonggukjang* fermentation. It (1) reveals the three-stage dynamic characteristics of *Cheonggukjang* fermentation and identifies the key stages of flavor formation, flavor markers, and microbial biomarkers, offering a new perspective for understanding the rules of flavor succession; (2) establishes the functional relationship among dominant microbial communities, characteristic flavor evolution, and differential metabolite trajectories, thereby clarifying the regulatory mechanism of the "microbiome–flavor–metabolite" axis; and (3) reconstructs the core microbial metabolic network and maps the metabolic pathways of flavor formation, providing a theoretical framework for the precise regulation of industrial fermentation. These mechanistic insights establish a fundamental scientific basis for optimizing industrial fermentation schemes and quality control, while also offering an actionable roadmap for the rational design of defined starter cultures to achieve precise flavor modulation in *Cheonggukjang*.

2. Materials and methods

2.1. Materials and reagents

Soybeans were procured from a local market, Dong Market (Yanji, Jilin, China). The C7-C40 saturated alkane standard mixture and DNS chromogenic reagent (CAS: 609–99-4) were purchased from Macklin Inc. (Shanghai, China) and Aladdin Biochemical Technology Co., Ltd. (Shanghai, China), respectively. Analytical grade glucose (CAS: 50–99-7), petroleum ether (CAS: 64742–49-0), sodium chloride (CAS: 7647-14-5), agarose (CAS: 9012-36-6), formaldehyde (CAS: 50–00-0), sodium hydroxide (CAS: 1310-73-2), N-(1-naphthyl) ethylenediamine dihydrochloride (CAS: 1465-25-4), and zinc acetate dihydrate (CAS: 557–34-6) were obtained from Sinopharm Chemical Reagent (Shanghai, China). Chromatographic grade 2-octanol (CAS:123–96-6) was supplied by Acmec Biochemical Co., Ltd. (Shanghai, China), while chromatographic grade methanol (CAS:67–56-1), formic acid (CAS:64–18-6), ammonium formate (CAS: 540–69-2), and acetonitrile (CAS: 75–05-8) were purchased from Merck KGaA (Darmstadt, Germany). The Fast DNA Spin Kit for Soil and Illumina TruSeq Nano DNA LT Library Preparation Kit were obtained from MP Biomedicals (Irvine, CA, USA) and Illumina, Inc. (San Diego, CA, USA), respectively.

2.2. The preparation and sample collection of Cheonggukjang

Soybeans were procured from a local supermarket in Yanji, Jilin Province, China. *Cheonggukjang* was prepared according to previous research (Cho et al., 2017), with minor modifications. Before fermentation, soybeans underwent the following steps: washing, soaking for 12 h, boiling for about 3.5 h, draining, and cooling to 40 °C for 1 h. The cooled soybeans were transferred to a bamboo basket with rice straw lining the bottom, which was covered with a layer of cotton cloth. The basket was then wrapped in a cotton blanket and kept at around 30 °C for 72 h to facilitate natural fermentation. *Cheonggukjang* samples were collected at 0, 3, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33, 36, 39, 42, 45, 48, 51, 54, 57, 60, 63, 66, 69, and 72 h during fermentation using aseptic sampling bags. At each time point, samples were taken from three independent batches to ensure replicate analysis. All collected *Cheonggukjang* samples were immediately dispensed and sealed in a − 80 °C refrigerator for storage.

2.3. Fermentation parameters analysis

The pH of *Cheonggukjang* was determined following the method described by Tang et al. (2013). Briefly, 10 g of fully ground *Cheonggukjang* was thoroughly mixed with 20 mL of distilled water, and the pH of the mixture was then determined using a pH meter (PHB-1, Hangzhou Qi Wei Instrument Co., Ltd.). The contents of amino acid nitrogen (AAN) and reducing sugars (RS) were determined following the method of Liang et al. (2019). Total acids content was measured following Chinese national standard GB 12456–2021. Both reducing sugars and total sugars were analyzed using the 3,5-dinitrosalicylic acid (DNS) colorimetric method (Liang et al., 2019). Moisture content was measured via the 105 °C drying method (Zhang et al., 2019). Color parameters were determined with a colorimeter (CR-400, Konica Minolta, Shanghai, China) calibrated with standard white tiles. The chromaticity of *Cheonggukjang* was evaluated referring to an earlier study (Li et al., 2017). Additionally, fat and nitrite contents were measured using Soxhlet extraction (Martineau-Côté et al., 2023) and the hydrochloric acid naphthyl ethylenediamine method (An et al., 2021), respectively.

2.4. Volatile flavor components analysis

The volatiles were extracted and identified following the method of Jia et al. (2020). A mixture of 2 g of *Cheonggukjang*, 2 g of NaCl, 6 mL of CO_2-removed distilled water, and 5 μL of internal standard 2-octanol (25 mg/L), was prepared in a 20 mL extraction vial. A DVB/CAR/PDMS fiber (2 cm, 50/30 μm) (Supelco, Bellefonte, Pennsylvania, USA) was used to extract the volatiles following our previous study (Zhao et al., 2023). The volatiles were identified by headspace solid-phase micro-extraction (HS-SPME) combined with gas chromatography–mass spectrometry (GC–MS) (7980 A, 5975C, Agilent Technologies, Santa

Clara, California, USA) equipped with a DB-5MS column of 60 m × 0.25 mm × 0.25 μm. The instrument parameters for GC–MS were set according to the method of Zhao et al. (2023). The identification of volatiles was conducted by comparing the mass spectra with the NIST 11.0 library (spectrum match >75%). Retention indices (RIs) were calculated according to the retention times of an alkane mixture (C_7-C_{40}). A semi-quantitative approach based on an internal standard was employed to determine the content of volatile compounds (Zhao et al., 2023).

2.5. Odor activity value (OAV) calculation

OAV was calculated using the following formula: $OAV_i = C_i/OT_i$, where Ci and OT_i were the content and odor threshold values in water of the identified substances, respectively.

2.6. Untargeted metabolomics analysis

A quantity of 0.5 g of *Cheonggukjang* was mixed with 600 μL of methanol containing 2-chloro-*L*-phenylalanine (4 ppm) and vortexed for 30 s using a mixer. A grinding bead with a diameter of 6 mm were added to the sample mixture, which was then ground in a frozen tissue grinder at 55 Hz for 60 s. After grinding, the mixture was ultrasonicated for 15 min at room temperature and centrifuged (10 min, 4 °C, 12000 rpm). Subsequently, the supernatant was filtered through a 0.22 μm membrane prior to LC-MS analysis. Vanquish Ultra-high performance liquid chromatography system (Thermo Fisher Scientific, USA) and ACQUITY UPLC ® HSS T3 (2.1 × 100 mm, 1.8 μm) (Waters, Milford, MA, USA) were used to separate the metabolic extract. The column temperature, flow rate, and inject volume were 40 °C, 0.3 mL/min, and 2 μL, respectively. The separation gradient and mobile phase in positive and negative ion modes are presented in **Table S1**. The non-volatile metabolites were identified using an Orbitrap Exploris 120 (Thermo Fisher Scientific, USA) equipped with an electrospray ionization (ESI) source, detected in both positive and negative ion modes. The sheath gas, auxiliary gas, capillary temperature, and positive and negative ion spray voltages were set at 40 arb, 10 arb, 325 °C, 3.50 kV, and − 2.5 kV, respectively. Samples were scanned with a range of 100–1000 *m/z*, and MS1 and MS2 resolutions were 70,000 and 17,500, respectively.

2.7. DNA extraction and metagenomic sequencing

Genomic DNA was extracted using Fast DNA™ Spin Kit for Soil DNA. DNA concentration, purity and quality were assessed using a Filter Max F5 (Shanghai Molecular Devices), Nanodrop 2000 (Beijing Keyu Xingye Technology Development Co., Ltd), and 1% agarose gel electrophoresis, respectively. The microbial DNA of metagenomic sequencing libraries with 350 bp was constructed using the Illumina TruSeq Nano DNA LT Library Preparation Kit. The Illumina HiSeq 2500 (Illumina, USA) was employed to sequence all libraries. The relative abundance of microorganisms in *Cheonggukjang* was obtained by sequence comparison of optimized quality control (QC) data with a reference database in NCBI, and gene function annotation was performed by comparing the gene set obtained by sequencing to the KEGG gene database.

2.8. Statistical analysis

Metabolomics analysis was conducted with six biological replicates, whereas the other tests were subjected to three replicates. IBM SPSS statistics (Inc., Chicago, USA) and Origin (OriginLab, USA) were employed for data analysis and plotting, respectively. Significant differences ($P < 0.05$) among groups were judged using Duncan's multiple comparison tests and one-way analysis of variance (ANOVA). Partial least-squares discrimination analysis (PLS—DA) and two-way Orthogonal Partial Least Squares (O2PLS) were performed using SIMCA14.1. RDA analyses were conducted using CANOCO 5.

3. Results and discussion

3.1. Dynamics of fermentation parameters during Cheonggukjang fermentation

Fermentation parameters serve as critical modulators of microbial dynamics and metabolic flux in food fermentation systems (Xiong et al., 2024). The dynamic changes in fermentation parameters of *Cheonggukjang* are presented in Fig. 1 and **Table S2**. Specifically, *Cheonggukjang* exhibited triphasic physicochemical evolution: (1) moisture (64.79 ± 0.29% → 55.92 ± 0.23%), total acidity (19.41 ± 0.78 → 5.65 ± 0.07 g/kg), and RS (7.97 ± 0.10% → 5.85 ± 0.05%) displayed sequential attenuation patterns—initial gradual decline, accelerated reduction, and eventual stabilization. Conversely, pH, AAN, and lipid content demonstrated inverse kinetic trajectories, progressing from incremental elevation to rapid augmentation before reaching biochemical equilibrium during fermentation. AAN is a crucial indicator of the maturity of fermented soybean products, which is associated with the catabolism of soybean protein by microbial and endogenous enzymes. The level of AAN reached a maximum value of 0.55 ± 0.03 g/100 g (international standard 0.5%) at 54 h of fermentation, after which it did not fluctuate notably. The contents of RS and nitrites in *Cheonggukjang* displayed an initial increase, followed by a decrease, and ultimately stabilized. The nitrite content increased during the first 39 h of fermentation, peaking at 1.74 ± 0.05 g/100 g at 39 h, after which it gradually declined and stabilized after 48 h. As the fermentation proceeded, chromatic parameters showed marked decreases in *L** (lightness) and *b** (yellowness) values, stabilizing after 54 h of fermentation. Since the drastic changes in various fermentation parameters mainly occurred from 0 h to 45 h, we inferred that the changes in the bacterial flora and metabolites during *Cheonggukjang* fermentation were similar. Therefore, the *Cheonggukjang* fermented for 0, 9, 18, 27, 36, 45, 60, and 72 h was selected for the subsequent experiments.

3.2. The microbial diversity and composition

The metagenomics sequencing was utilized to investigate the microbial diversity and composition. A total of 873,698,264 reads were generated from the samples, averaging 78,261,832.83 reads per sample. Microbial α-diversity dynamics during *Cheonggukjang* fermentation are shown in Fig. 2A-D. Both microbial richness (Chao1 and Abundance-based Coverage Estimator (ACE) indices) and diversity (Shannon and Simpson indices) were highest in the unfermented samples (0 h). The diversity of microbial communities in *Cheonggukjang* decreased rapidly during the initial fermentation (0–9 h), followed by fluctuating trend. The lowest microbial richness and diversity were observed at 45 h and 18 h of fermentation, respectively, showing significantly differences ($P < 0.05$) compared to the 0 h control. To elucidate the difference in microbiota during *Cheonggukjang* fermentation, principal coordinate analysis (PCoA) was carried out. The PCoA plot suggested that a clear separation was found among various *Cheonggukjang* samples from non-fermentation stage (0 h), the initial fermentation stage (9–18 h), and the middle to late fermentation stage (27–72 h).

A total of 26 phyla (including 23 bacterial and 3 fungal phyla) were detected during the fermentation of *Cheonggukjang*, and *Firmicutes* (87.37 ± 26.47%) was the absolute dominant phylum. At the genus level, 535 bacteria and 14 fungal genera were detected during *Cheonggukjang* fermentation, of which dominant genera (mean relative abundance >1%) included *Bacillus* (63.62 ± 27.72%), *Caldibacillus* (17.57 ± 15.54%), *Dietzia* (2.63 ± 6.06%), *Aeromonas* (1.78 ± 5.09%), *Weissella* (1.67 ± 5.27%), *Leclercia* (1.58 ± 4.19%), *Aeromicrobium* (1.44 ± 3.30%), *Rhodococcus* (1.34 ± 3.19%), and *Brevibacillus* (1.02 ± 0.88%). *Bacillus* was the keystone taxon throughout the fermentation process, with its relative abundance reaching a peak of 92.36 ± 0.60% at 9 h, followed by a wave-like fluctuations. The relative abundance of *Cladibacillus* increased rapidly during the middle and later fermentation

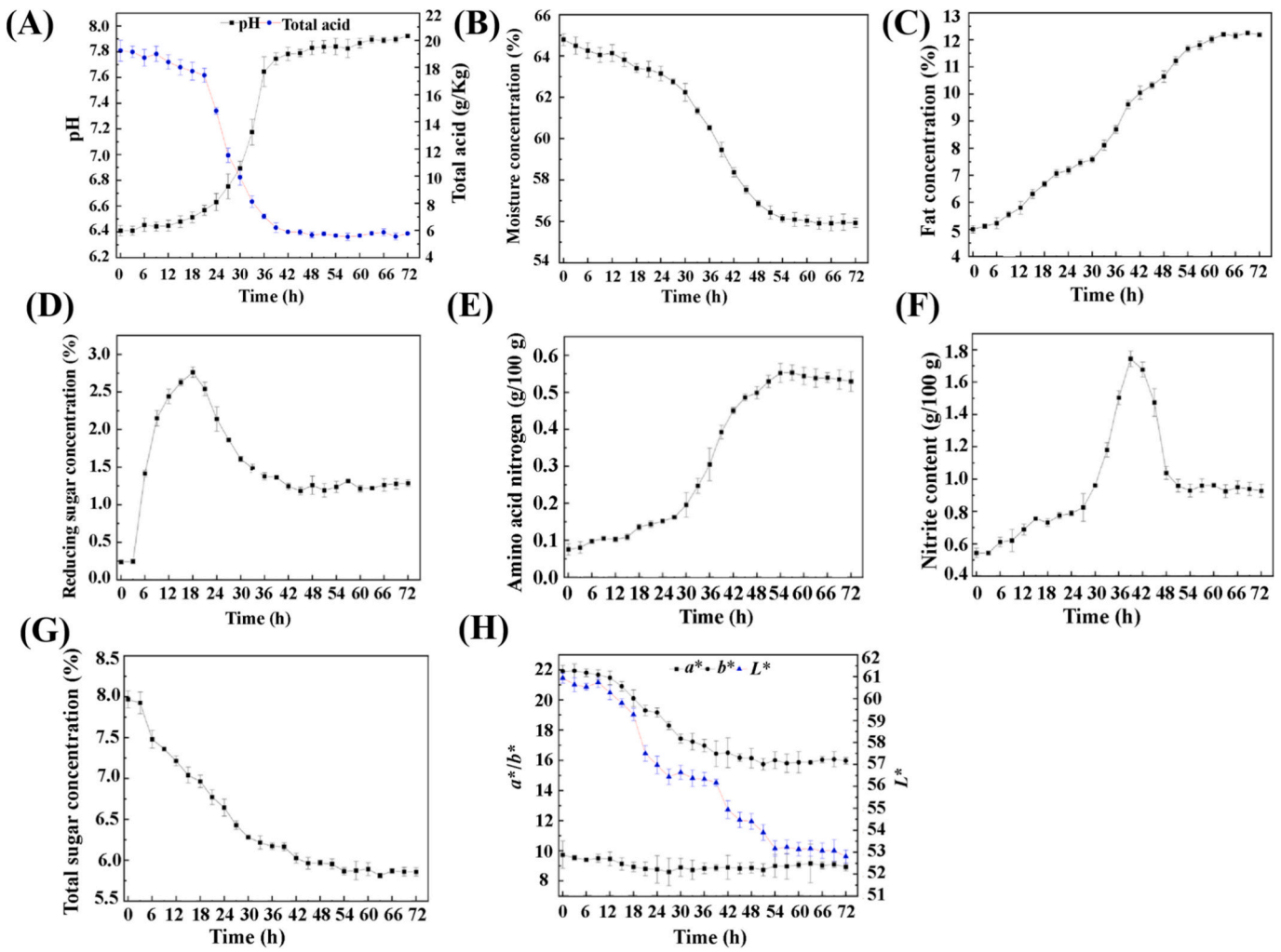

Fig. 1. Fermentation parameters of *Cheonggukjang* at different fermentation stages. (A) represents the pH and total acid of *Cheonggukjang* at different fermentation stages. (B), (C), (D), (E), (F) and (G) represent the moisture, fat, reducing sugar (RS), amino acid nitrogen (AAN), nitrite and total sugar concentration, respectively. (H) represents the color properties of *Cheonggukjang* at different fermentation stages.

stages (27–72 h) and peaked at 43.11 ± 1.32% at 36 h, indicating it may play pivotal roles in the mid-late fermentation phases. A total of 1902 species were identified during *Cheonggukjang* fermentation, including 1889 bacterial species and 13 fungal species. *B. licheniformis* maintained a relatively high abundance across all fermentation stages, demonstrating its essential role in *Cheonggukjang* fermentation.

Linear discriminant analysis effect size (LEfSe) analysis was performed to identify the biomarkers for each fermentation phase, with the Linear discriminant analysis (LDA) score threshold set at 4.0 (Fig. 2I). The results suggested that *B. subtilis*, *B. velezensis*, *C. thermoamylovorans*, and *B. licheniformis* were the biomarkers for the 9, 18, 36, and 60 h fermentation periods of *Cheonggukjang*, respectively. Collectively, these findings demonstrated that *B. subtilis*, *B. velezensis*, *C. thermoamylovorans*, and *B. licheniformis* may tremendously affect the quality formation of *Cheonggukjang*.

3.3. Volatile compounds of Cheonggukjang by different fermentation times

The flavor properties of fermented foods arise from a vast array of diverse flavor metabolites, which collectively define their sensory signature and commercial value (Zhao et al., 2024). HS-SPME-GC–MS was employed to explore the changes in volatiles during *Cheonggukjang* fermentation. Sixty-one volatiles were identified, including 12 alcohols, 9 ketones, 7 acids, 4 aldehydes, 7 pyrazines, 5 phenols, 5 esters, and 12 other categories (Fig. 3 and **Table S3**). The results revealed significant changes in volatiles of *Cheonggukjang* across different fermentation stages, with marked variations observed in pyrazines.

The production of volatile compounds has been attributed to a series of biochemical reactions. At 72 h of fermentation, the content of volatile substances in *Cheonggukjang* reached a maximum of 7909.90 ± 47.74 μg/kg, with pyrazines constituting the predominant fraction (73.59 ± 7.52%). Both of pyrazines and aldehydes content exhibited a gradual increase throughout fermentation, with a marked metabolic surge observed after 60 h of fermentation. At the terminal fermentation phase (72 h), aldehyde and pyrazine levels reached 1218.69 ± 180.66 and 5820.83 ± 594.70 μg/kg, representing 81.6 and 2163.9 times higher than initial values (0 h), respectively. The concentrations of alcohols, acids, esters, and ketones all exhibited an initial rise followed by a decline during *Cheonggukjang* fermentation. Their contents reached the highest values of 175.69 ± 24.67, 139.48 ± 7.80, 8.97 ± 1.97, and 1030.07 ± 182.20 μg/kg at 18, 27, 18, and 9 h, respectively. Phenolics in *Cheonggukjang* showed progressive accumulation during fermentation, rising from 12.62 ± 1.87 μg/kg at 0 h to 212.33 ± 18.46 μg/kg at 72 h. In addition, other flavor components such as furans, thiazoles, and indoles were identified in *Cheonggukjang*, which can influence its flavor through their own odors.

To explore the differences in volatile profiles of *Cheonggukjang*

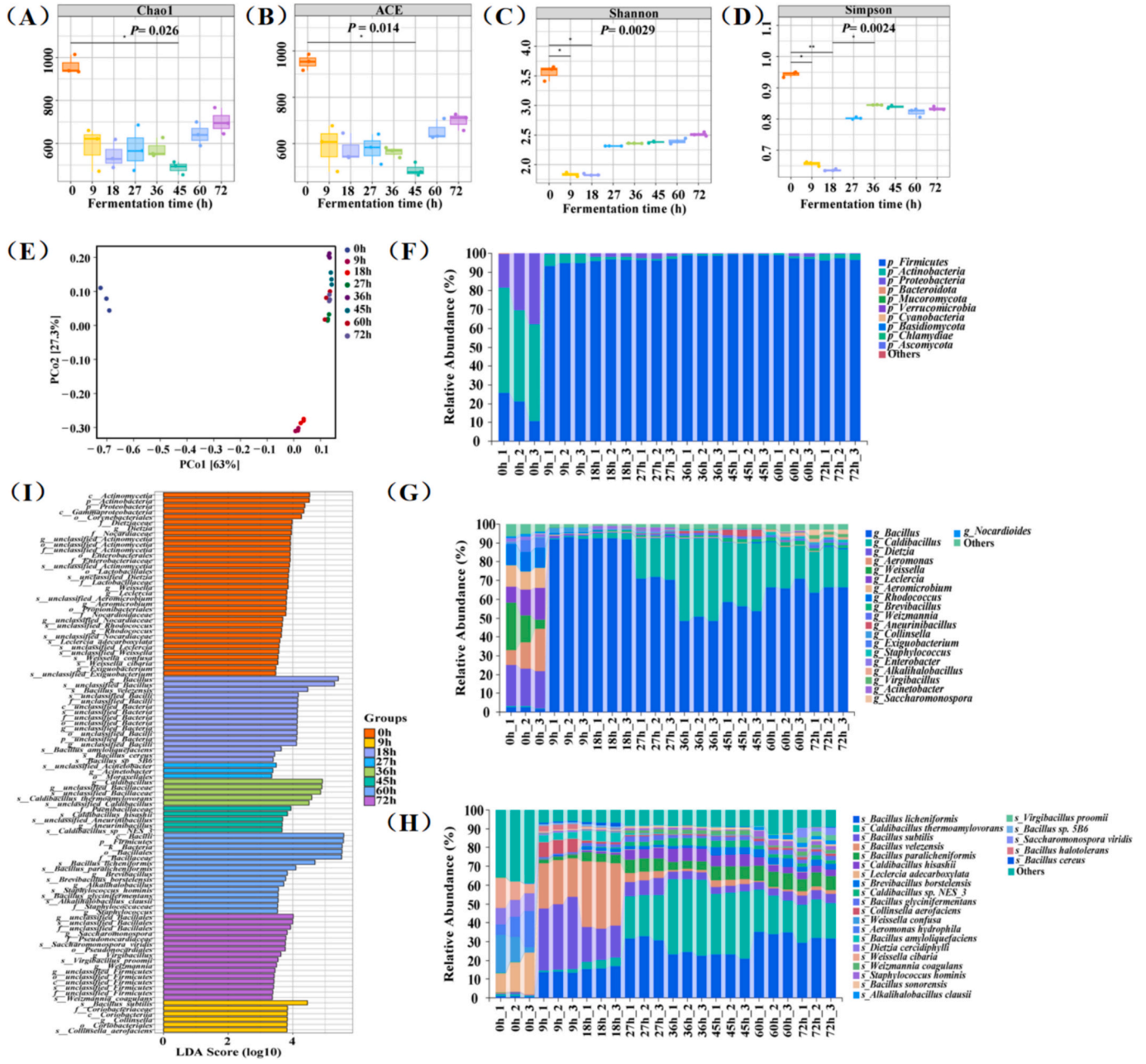

Fig. 2. The microbial diversity and composition of *Cheonggukjang* at different fermentation stages. Alpha diversity indices, including (A) Chao 1 index, (B) Abundance-based coverage estimator (ACE) index, (C) Shannon index, and (D) Simpson index. (E) Principal coordinate analysis (PCoA). Relative abundance of dominant microorganisms on different taxonomic classes, (F) phylum, (G) genus, and (H) species. (I) The microbial biomarker by Lefse analysis.

among different fermentation stages, PLS-DA and cluster analysis were performed in this study. PLS-DA analysis showed significant differentiation among samples of different fermentation durations (Fig. 3C). The 200-permutation test yielded a lower Q2 value for the PLS-DA model compared to the original model, indicating that the original model was not over-fitted and showed good predictive performance (Fig. 3D). Cluster analysis of volatile profile classified the *Cheonggukjang* fermentation process into three key stages: the early stage (0–18 h), the middle stage (18–60 h), and the late stage (60–72 h) (Fig. 3E), implying that 18–27 h and 60–72 h are key transitional periods for volatile compound evolution. Variable importance in the projection (VIP) values of volatile compounds at different fermentation stages are displayed in Fig. 3F. The VIP values of the volatile components reflect their relative importance in discriminating the flavor profiles of *Cheonggukjang* across different fermentation stages. A total of 25 volatile components with VIP value>1, including 5 pyrazines, 4 alcohols, 4 ketones, 3 phenols, 2 aldehyde,1 esters, and 6 other categories, were identified as potential volatile markers in *Cheonggukjang* at different fermentation stages. Hence, these substances can be utilized to distinguish *Cheonggukjang* at different fermentation stages based on volatile profile.

3.4. Key volatile flavor compounds

Despite the numerous volatiles in *Cheonggukjang*, only a limited number of key components define its sensory profile, contributing its characteristic flavor. To investigate these key components, the OAV of volatile substances was calculated based on their content and threshold values (Huang et al., 2022; Zhao et al., 2025). In general, volatiles with OAV ≥ 1 are regarded as key flavor compounds. Fifteen volatiles with OAV > 1 were obtained, including 3 alcohols, 4 pyrazines, 3 phenols, 2

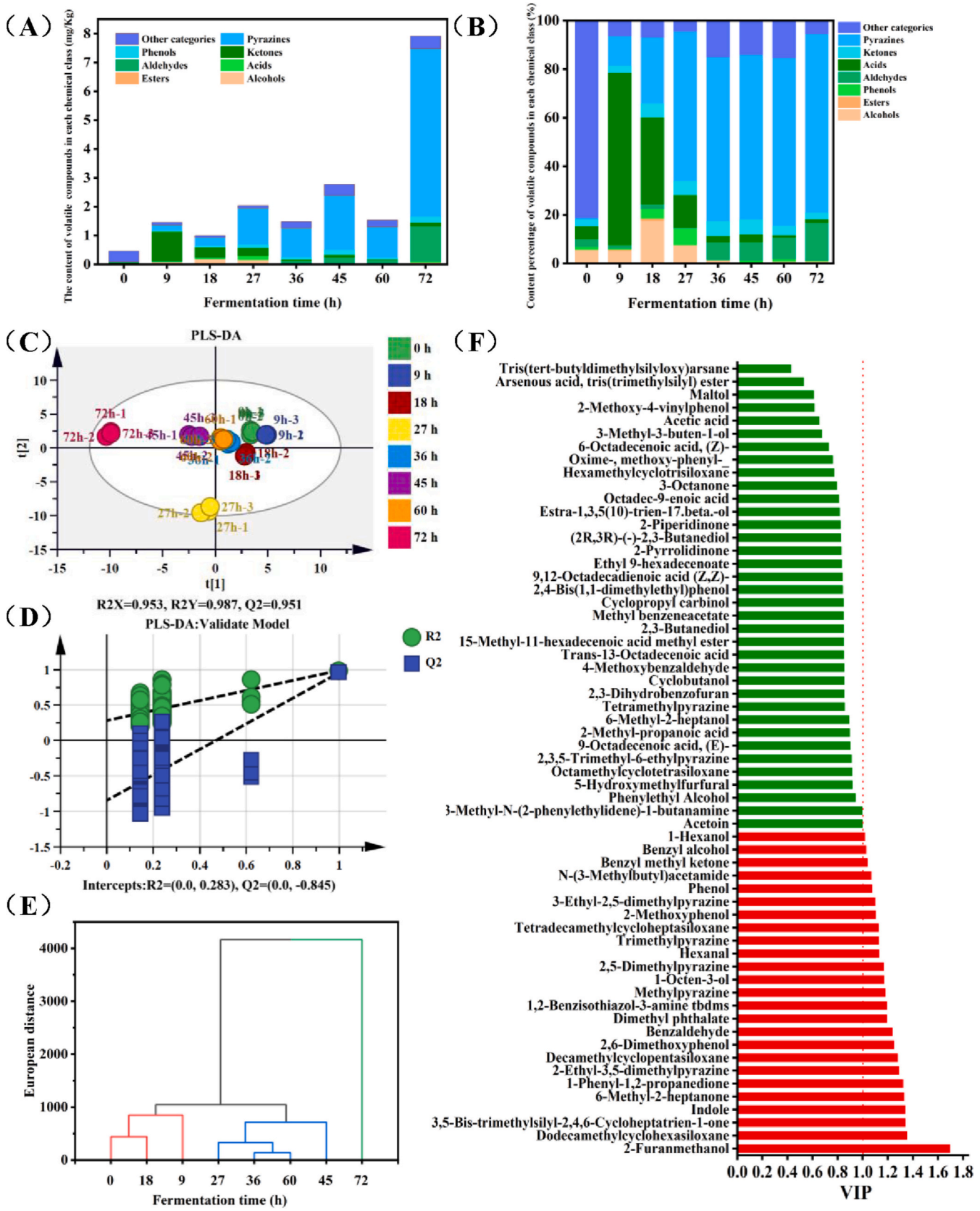

Fig. 3. Dynamic changes of volatile flavor of *Cheonggukjang* at different fermentation stages. The content (A) and relative concentration (B) of different types of volatile compounds. (C) and (D) were the score plot of volatile compounds of PLS-DA (Partial least squares discrimination analysis) and validation plot of 200 permutation tests of the model, respectively. (E) Hierarchical clustering analysis diagram of volatile components based on Euclidean distance. (F) Variable importance in the projection (VIP) of different fermentation stages.

ketones, 2 aldehydes, and 1 acid (Table 1**, Fig. S1**), which primarily accumulated during the critical fermentation stage between 60 and 72 h. This provides a precise time window for targeted flavor interventions, facilitating a modular management approach for fermentation optimization.

During the initial fermentation stage (0–18 h), 1-octen-3-ol exhibited the highest OAV, establishing mushroom-like odor as the dominant aromatic note. As fermentation progressed, pyrazines and benzaldehyde (OAV = 406.23) gradually became the primary aroma contributors, with 3-ethyl-2,5-dimethylpyrazine (OAV = 351.72) and 2-ethyl-3,5-dimethylpyrazine (OAV = 128.5) serving as the key pyrazine components. Cross-referencing the data from this section with Section 3.3 identified ten key flavor markers for distinct fermentation stages (Table 1).

3.5. Non-volatile substances analysis

LC-MS technique was utilized to analyze the non-volatile metabolites at each fermentation stage. PCA results indicated that the QC samples exhibited excellent aggregation characteristics in both positive and negative ion modes, demonstrating that the detection was reproducible and highly reliable (**Fig. S2**). A total of 581 non-volatile metabolites were obtained after QC processing of the raw data, and the numbers and relative abundances of each class of non-volatile compounds were presented in Fig. 4A-B. The largest number of non-volatile metabolites was associated with amino acid metabolism (125), followed by organic acids (89) and fatty acids (76). The relative abundance of non-volatile metabolites associated with amino acid metabolism was the highest during *Cheonggukjang* fermentation, followed by that of ketones and fatty acids.

To further investigate differential non-volatile metabolites during *Cheonggukjang* fermentation, we performed a PLS-DA analysis **(**Fig. 4C). The R2X and R2Y values were 0.834 and 0.995, respectively, demonstrating that the model exhibited excellent predictive ability and did not appear to be overfitted **(**Fig. 4D). The non-volatile compounds at various fermentation stages could be separated, again confirming that there were noteworthy differences in metabolite composition during *Cheonggukjang* fermentation. VIP analysis was employed to study the different metabolites during *Cheonggukjang* fermentation. 216 differential non-volatile substances were identified based on VIP $\geq$ 1 and $P < 0.05$, which were primarily enriched in 20 pathways based on KEGG (Kyoto Encyclopedia of Genes and Genomes, http://www.genome.jp/kegg/). (Fig. 4E**, Table S4**). The main metabolic pathways related to flavor formation in *Cheonggukjang* were amino acid metabolism. Specifically, the main pathways included arginine biosynthesis; alanine, aspartate and glutamate metabolism; phenylalanine, tyrosine and tryptophan biosynthesis; tyrosine metabolism; phenylalanine metabolism; as well as valine, leucine and isoleucine biosynthesis.

3.6. Correlation analysis of microbial community with fermentation parameters, key volatile flavor compounds, and different non-volatile metabolites

The pivotal role of microorganisms in shaping the flavor of fermented foods is modulated by a complex interplay of environmental factors, nutritional conditions, and microbial interactions, which collectively influence their growth, metabolic activity, and community composition (Xiong et al., 2024). To delineate the association between environmental variables and microbial community structure, we performed redundancy analysis (RDA) (Fig. 5A). The model showed high reliability, as it explained 94.15% of the total variation in RDA results. *B. licheniformis*, *Bacillus glycinifermentans*, *B. paralicheniformis*, *Brevibacillus borstelensis*, *Caldibacillus hisashii*, *Caldibacillus* sp. NES_3, and *C. thermoamylovorans* have arrows in the same direction with pH, fat, nitrite, and AAN, and in the opposite direction with moisture, color parameters, and total sugar, suggesting that they may be responsible for the metabolism of fats, sugars, and the production of AAN and nitrite. Additionally, *B. subtilis*, *B. velezensis*, and RS concentration displayed a strong correlation, suggesting that these bacteria may help promote the transformation of RS. As shown in the RDA bioplot, total sugar content exhibited the strongest vector magnitude, followed by RS and pH, indicating that total sugar content is the key fermentation parameter driving microbial community succession during *Cheonggukjang* fermentation.

Microbial interactions play a pivotal role in sustaining community stability and ensuring fermented food quality (Melkonian et al., 2023). We explored the correlations among the 50 most abundant microbial taxa during *Cheonggukjang* fermentation (Fig. 5B**).** A total of 48 nodes and 290 edges were obtained, reflecting the complexity of co-occurrence networks. The network patterns demonstrate distinct ecological interactions: positive correlations between microorganisms imply potential mutualistic interactions (e.g., symbiosis, metabolic cooperation, or synergistic relationships) that stabilize the community structure, while negative correlations reflect competitive exclusion or antagonistic interactions that drive ecological niche specialization. In this network, positive correlations were significantly more common than negative correlations. Microbial networks exhibited a characteristic scale-free topology, with the majority of microorganisms possessing few connections and only a minority serving as highly connected hub nodes. Three microbial species, including *B. licheniformis*, *B. glycinifermentans*, and *Bacillus sonorensis*, exhibited the highest connectivity (connectivity degree>20, $|r| > 0.6$, $P < 0.05$) and demonstrated predominantly

Table. 1
The OAV of *Cheonggukjang* during fermentation.

Volatile compounds	Thresholds(μg/kg)	Fermentation time/h							
		0	9	18	27	36	45	60	72
1-Octen-3-ol	1	17.11	65.33	31.74	0.00	0.00	0.00	0.00	0.00
Benzyl alcohol	3	0.00	0.00	2.08	3.22	1.51	2.3	1.78	6.35
Phenethyl alcohol	3.17	0.00	0.00	1.59	4.68	2.83	0.71	2.03	4.67
Isobutyric acid	10	0.00	0.00	1.54	2.32	0.05	2.18	1.62	3.5
Benzaldehyde	3	4.15	4.51	5.93	0.00	36.27	70.24	45.23	406.23
Hexanal	5	0.5	1.33	0.00	0.00	0.00	0.00	0.00	0.00
Acetoin	14	0.00	70.73	23.18	14.62	0.00	0.00	0.00	0.00
3-Octanone	21.4	0.00	0.00	0.00	0.00	0.00	1.24	0.00	1
2-Methoxyphenol	1.6	0.00	7.48	20.43	44.76	32.12	62.16	22.73	77.64
Phenol	31	0.00	0.09	0.52	1.05	0.85	1.61	0.44	2.26
2-Methoxy-4-vinylphenol	3	1.04	5.97	1.53	2.52	2.02	2.76	0.91	1.92
2,5-Dimethylpyrazine	80	0.03	0.64	1.83	8.45	8.28	16.93	9.75	54.52
Trimethylpyrazine	23	0.00	0.77	4.1	16.52	12.29	20.39	10.55	54.65
3-Ethyl-2,5-dimethylpyrazine	0.4	0.00	0.00	61.82	220.99	108.81	111.79	76.8	351.72
2-Ethyl-3,5-dimethylpyrazine	0.16	0.00	0.00	0.00	0.00	0.00	0.00	8.82	128.5

Note: Bolded entries in the table denote key flavor markers for different fermentation stages. Thresholds of different volatile flavor compounds were determined in water media and obtained from relevant literature and websites.

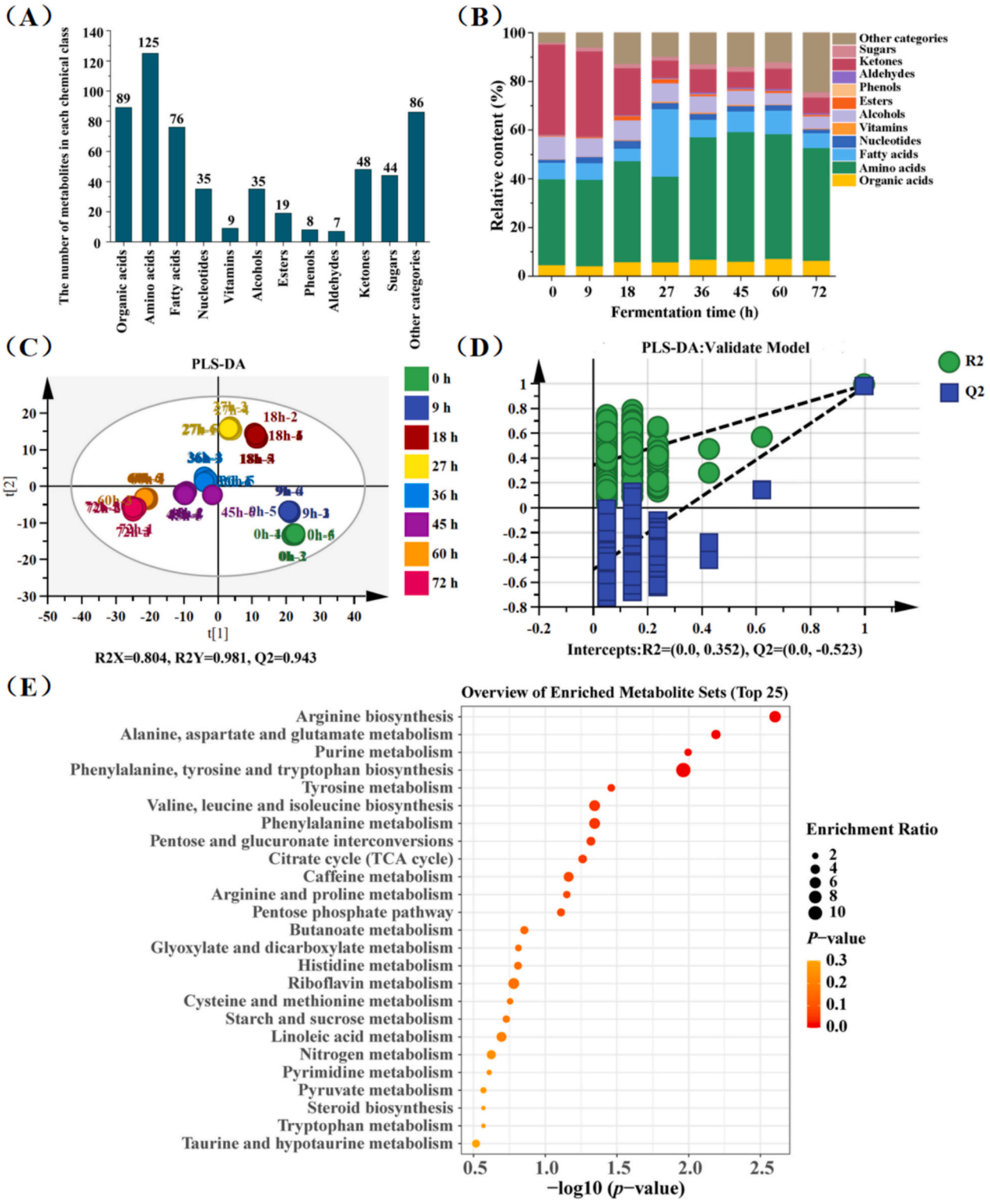

Fig. 4. Dynamic changes of non-volatile metabolites of *Cheonggukjang* at different fermentation stages. The number (A) and relative content (B) of different types of non-volatile metabolites. (C) Partial least squares discrimination analysis (PLS-DA) plot of non-volatile metabolites. (D) Validation plot of 200 permutation tests of the model (E) Classification and statistics of KEGG functions in differential non-volatile metabolites during *Cheonggukjang* fermentation.

negative correlations with other bacterial taxa, suggesting competitive interactions for nutrients and spatial resources.

O2PLS is commonly employed for describing the liaisons between two groups of data, which contribute to reflecting the overall impact between different data sets and screening key variables influencing the other set (Rajasundaram & Selbig, 2016). In this study, O2PLS analysis was used to investigate the liaisons between microorganisms and key flavors (Fig. 5C). The high values of R2 (0.924) and Q2 (0.878) suggest that the model possesses excellent explanatory capacity and reliable predictive performance (**Table S6**). The permutation test results (200 iterations) indicated that the model was not over-fitted (Fig. 5D). The VIP scores for the analyzed microbiota ranged from 0.14 to 2.05, with 12

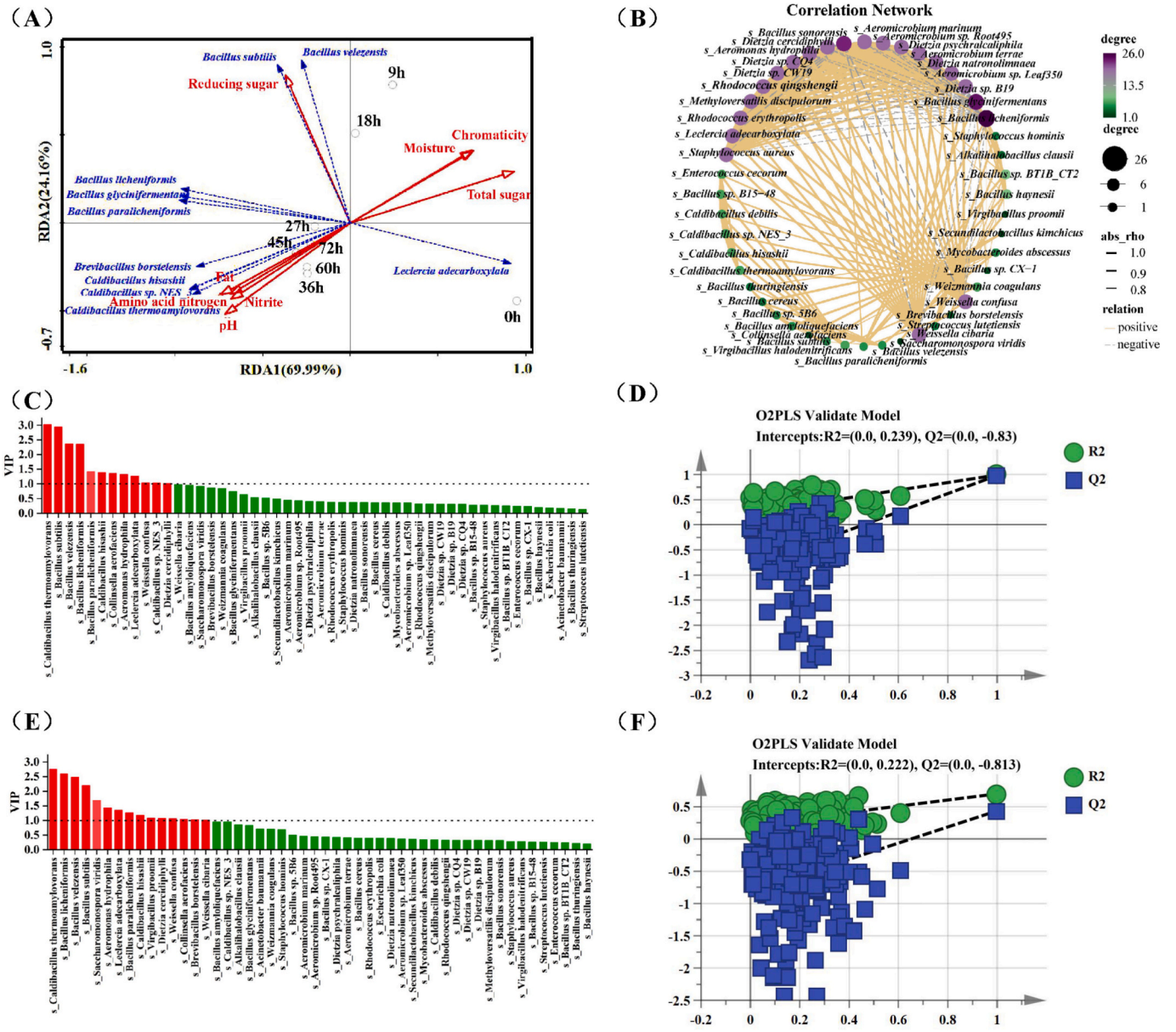

Fig. 5. Association analysis of species-level microorganisms with fermentation parameters (A). The red and blue arrows represent fermentation parameters and microbial species, respectively. (B) Correlation network between microbial communities ($P < 0.05$, $r > 0.6$). (C) and (E) were VIP_{pre} (VIP value for predictive variables) plots of O2PLS analyses. (D) and (F) represent the cross-validation with 200 permutation tests. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

microbial species (VIP > 1.0) identified as key contributors to characteristic flavor compounds. Notably, *C. thermoamylovorans* exhibited the highest VIP score (3.02), followed closely by *B. subtilis* (VIP = 2.94), strongly suggesting that they play the most critical role in defining the key flavor profile of *Cheonggukjang*. The links between microorganisms and key flavor compounds during *Cheonggukjang* fermentation were also evaluated using Pearson correlation analysis (**Fig. S3**). The analysis revealed a statistically significant positive correlation between the 8 microbial species and the characteristic flavor compounds, except for negative correlations with 1-octecn-3-ol and hexanal. *Bacillus paralicheniformis* exhibited the most extensive correlations with characteristic flavor compounds, followed by *B. licheniformis*. Both of *B. paralicheniformis* and *B. licheniformis* were associated with several characteristic flavor compounds including 2-methoxyphenol, benzyl alcohol, and isobutyric acid derived from amino acid metabolism.

Similarly, links between microbial and non-volatile differential metabolites were explored by O2PLS analysis. The model demonstrated excellent performance with R2 and Q2 values of 0.916 and 0.812 (**Table S6**). Based on 200 permutation tests, we concluded that the model was not over-fitted (Fig. 5F). VIP scores spanned from 0.20 to 2.75, of which 16 microbial species were significantly associated with metabolic activity during fermentation (Fig. 5E). Comprehensively analyzing the correlations between microorganisms and the characteristic flavor substances and differential metabolites during *Cheonggukjang* fermentation (**Fig. S3**), we found that 10 microbial species, including *B. velezensis*, *Virgibacillus proomii*, *B. subtilis*, *Saccharomonospora viridis*, *B. licheniformis*, *B. paralicheniformis*, *D. cercidiphylli*, *C. hisashii*, *C. thermoamylovorans*, and *Bre. Borstelensis*, showed associations with at least 20 differential metabolites, which were key drives of metabolic transformations during *Cheonggukjang* fermentation. Notably, *V. proomii* was the low-abundance microorganism in *Cheonggukjang*, implying that low-abundance microorganisms may play an indispensable part in flavor

formation. *B. paralichniformis* was correlated with the highest number of differential non-volatile metabolites (78), followed closely by *B. licheniformis* (61), which is consistent with their contribution to characteristic flavors.

Integrated analysis of microbial associations with characteristic flavors and differential non-volatile metabolites demonstrated that seven species (*C. thermoamylovorans*, *B. subtilis*, *B. velezensis*, *B. licheniformis*, *B. paralicheniformis*, *C. hisashii*, and *Collinsella aerofaciens*) were key contributors to flavor development through metabolic transformation during *Cheonggukjang* fermentation.

4. Discussion

This study systematically elucidated the formation mechanism of characteristic flavors during the spontaneous fermentation of *Cheonggukjang* by integrating metagenomics, volatilomics, and metabolomics. It identified the key stages of flavor formation, the core functional microorganisms, and the metabolic pathways involved, thereby filling the gap in the previously unidentified causal link between the microbial community and the characteristic flavors of *Cheonggukjang*. These findings not only advance the mechanistic understanding of microbe–metabolite–flavor interactions but also establish a theoretical basis for the targeted regulation and quality improvement of *Cheonggukjang*.

The dynamic changes in physicochemical parameters during the fermentation process reflect a substrate transformation process driven by microbial metabolism. RDA analysis identified total sugar as the core physicochemical factor driving microbial community succession, with its continuous depletion providing the carbon source and energy basis for this process. Previous studies have demonstrated that starch and RS concentrations drives microbial assembly over the temporal scale during the Stacking Fermentation (SF) process of Maotai-flavor baijiu (MFB) (Wang et al., 2021; Yang et al., 2023). Our prior research demonstrated that RS have a momentous effect on the bacterial community structure of dajiang (Zhao et al., 2025). Collectively, these findings indicate that the composition of raw materials, particularly the carbon sources for microbial growth, is a key factor driving microbial succession in fermented foods. Initially (0–18 h), reducing sugars generated from total sugar hydrolysis and inherently present in soybeans supported rapid proliferation of facultative anaerobes (e.g., *B. subtilis*, *B. velezensis*), establishing them as the initial dominant taxa. Concurrently, the rapid microbial proliferation coincided with enhanced nitrate reductase activity, promoting the conversion of nitrate to nitrite and leading to its transient accumulation. As fermentation proceeds, total sugars are continuously consumed, the temperature of the system rises, the pH increases, and the original diverse microbial niches are gradually reshaped.

Firmicutes and *Bacillus* predominated throughout the process, consistent with previous research (Hong et al., 2024; Nam et al., 2012) and further substantiating the essential ecological function of *Bacillus* in high-protein soybean fermentations. LEfSe analysis identified *C. thermoamylovorans* as a biomarker at 36 h of fermentation. During fermentation, microbial aerobic respiration rapidly degrades carbohydrates and proteins, generating substantial ATP and metabolic heat that elevates the temperature of *Cheonggukjang*. These changes create suitable conditions for subsequent proliferation of *C. thermoamylovorans* at later fermentation stages (Andersson et al., 1995). Additionally, the rise of *C. thermoamylovoran* coincided with a critical environmental transition—the depletion of total sugars and the concurrent increase in pH—confirming that carbon-source depletion drives niche differentiation. Once easily metabolized carbon sources were exhausted and the system pivoted to complex substrates like soybean proteins, bacteria (such as *C. thermoamylovoran*, and *B. licheniformis*) harboring specific hydrolase systems outcompeted other taxa. *B. licheniformis* encodes the alkaline protease AprE, which displays higher hydrolytic efficiency compared to other proteases (Zhou et al., 2025). The proteases secreted by these microbes extensively hydrolyze soybean proteins, leading to the release of free amino acids and basic amines, which in turn drives a sustained increase in both pH and amino acid nitrogen (AAN). AAN peaks at 0.55 $\pm$ 0.03 g/100 g at 54 h, marking product maturation. Concurrently, the rise in pH neutralizes organic acids, thereby reducing total acidity. During the final stage (60–72 h), all parameters stabilize. Nitrite is degraded to safe levels by nitrite reductase under acidic conditions, signifying the completion of *Cheonggukjang* fermentation.

During *Cheonggukjang* fermentation, volatile flavor compounds exhibited distinct stage-specific profiles. The initial stage was characterized by alcohols and acids, while pyrazines and aldehydes gradually increased in the middle and late stages, peaking at 72 h. A previous study did not observe significant accumulation of pyrazines, which may be attributed to differences in fermentation conditions compared to the present study (Cho et al., 2017). This dynamic closely paralleled the succession of the microbial community. During the initial stage, microorganisms—primarily *B. subtilis* and *B. velezensis*—generated abundant alcohols and acids. In the middle and late stages, proteases and lipases secreted by *B. licheniformis* and *Caldibacillus. thermoamylovorans* extensively released flavor precursors, including amino acids and fatty acids, which were subsequently converted into characteristic pyrazines and aldehydes. Notably, although pyrazines exhibited the highest total content among the volatile compounds, benzaldehyde displayed the highest OAV (406.23) among the 15 key flavor compounds with OAV $\geq$ 1 that define the characteristic flavor of *Cheonggukjang*. This finding highlights the necessity of incorporating odor activity thresholds in flavor analysis to assess the actual contribution of individual compounds. Untargeted metabolomics analysis revealed that the differential metabolites during *Cheonggukjang* fermentation were primarily enriched in amino acid metabolic pathways. The metabolites derived from these pathways serve as key precursors for the major volatile flavor compounds in *Cheonggukjang*, including pyrazines, phenols, aldehydes, and alcohols. The significant enrichment of the phenylalanine and tyrosine metabolic pathways not only provided direct precursors for aromatic aldehydes and alcohols such as benzaldehyde, phenylethyl alcohol, and benzyl alcohol, but also served as the core metabolic basis for the generation of phenolic compounds including phenol, 2-methoxyphenol, and 2-methoxy-4-vinylphenol. Furthermore, the active metabolism of threonine provided ample precursors for the synthesis of nitrogen-containing heterocyclic flavor compounds such as pyrazines, thereby enhancing the characteristic roasted and nutty aromas of *Cheonggukjang*. These findings demonstrated that the systemic reprogramming of the primary metabolic network during *Cheonggukjang* fermentation is highly coupled with the formation of volatile flavor compounds. The synergistic regulation of amino acid and fatty acid metabolic pathways collectively drives the directed conversion of non-volatile precursors into characteristic flavor components, including 2,5-dimethylpyrazine, trimethylpyrazine, 3-ethyl-2,5-dimethylpyrazine, 2-ethyl-3,5-dimethylpyrazine, phenol, 2-methoxyphenol, 2-methoxy-4-vinylphenol, benzaldehyde, phenylethyl alcohol, and benzyl alcohol.

Microorganisms were the most indispensable participants in flavor formation through their metabolic activities. We integrated metagenomics, volatilomics and metabolomics to reconstruct the microbial-mediated metabolic network (Fig. 6 **and Table S7)**, which revealed that the formation of characteristic flavors was closely associated with microbial regulation of carbohydrate metabolism, protein and lipid degradation, fatty acid oxidation, as well as amino acid metabolism and biosynthesis. Carbohydrate metabolism provided the fundamental biochemical basis for microbial proliferation and flavor components biosynthesis in *Cheonggukjang*. As one of the main components of soybeans, starch was degraded into glucose by *A. hydrophila*, *L. adecarboxylata*, and *D. cercidiphylli*. As the primary energy source for microbial proliferation, glucose was further metabolized into pyruvate by multiple strains, including *A. hydrophila*, *L. adecarboxylata*, *D. cercidiphylli*, *C. thermoamylovorans*, *C. hisashii*, *Caldibacillus* sp. NES_3, *W. coagulans*, *Caldibacillus debilis*, and *Bacillus* sp. B15–48. Pyruvate acted as a central metabolic hub linking glucose metabolism to

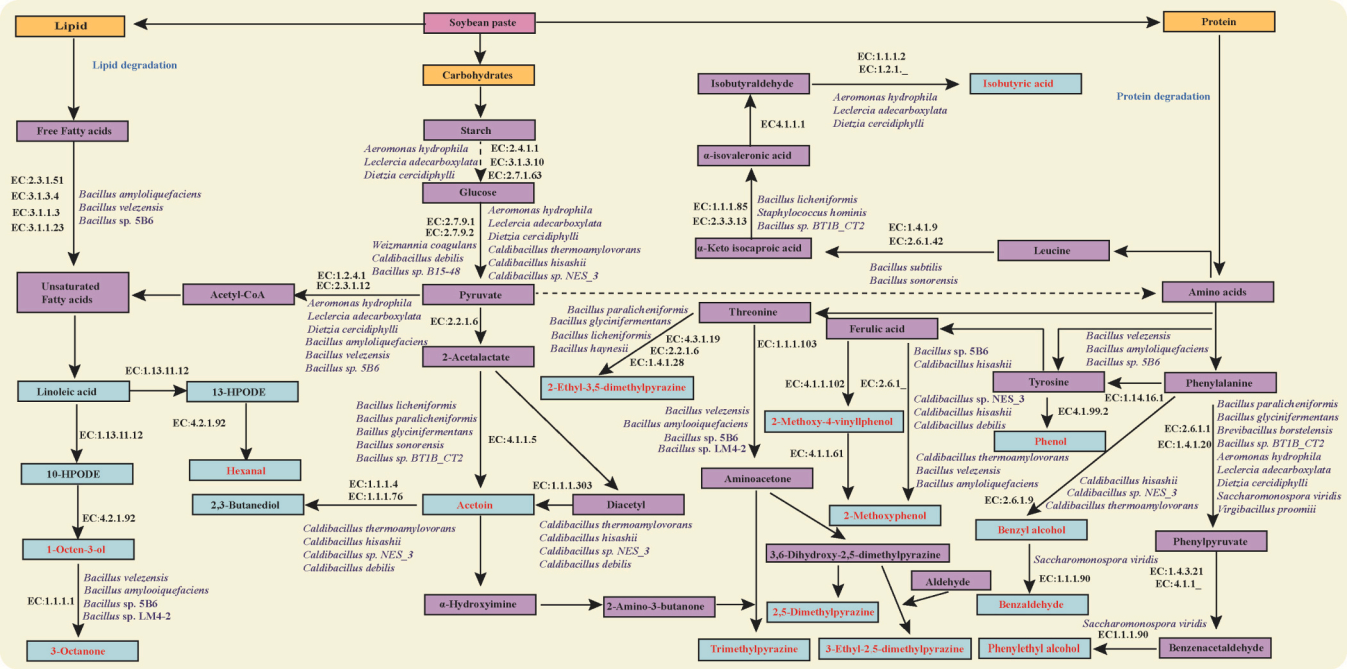

Fig. 6. Metabolic network of key flavor formation mediated by microorganism in *Cheonggukjang*. The red and black fonts in the box present the characteristic flavor components and intermediate metabolites in *Cheonggukjang*, respectively. Black and blue fonts not in the box represent enzymes and main microorganisms that produce the corresponding enzymes, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

downstream flavor synthesis, and could be converted into acetoin via 2-acetolactate, which was mediated by *B. licheniformis*, *B. paralicheniformis*, *B. glycinifermentans*, *B. sonorensis*, and *Bacillus* sp. BT1B_CT2. Acetoin further reacted with aminoacetone derived from threonine degradation to form trimethylpyrazine (Li et al., 2026). During fermentation, proteolysis contributed to the accumulation of abundant free amino acids and derivatives, such as phenylalanine, tyrosine, threonine, and leucine, which served as critical flavor precursors. These amino acids could be further converted into aldehydes, alcohols, phenols, and ketones, which collectively shape the characteristic flavor profile of *Cheonggukjang* (Zhao et al., 2023). Specifically, tyrosine was converted into phenol 2-methoxy-4-vinylphenol, and 2-methoxyphenol mainly by *B. velezensis*, *B. amyloliquefaciens*, *Bacillus* sp. 5B6, *C. thermoamylovorans*, *C. hisashii*, *Caldibacillus* sp. NES_3, and *C. debilis*, thereby contributing to the typical phenolic and smoky odors. Additionally, these phenolics can also be synthesized from ferulic acid released from soybean seed coat cell walls (Shashank et al., 2014). Phenylalanine was transformed into phenylethyl alcohol, benzyl alcohol, and benzaldehyde via decarboxylation and deamination, which endowed *Cheonggukjang* with fruity and flowery odors (Hao et al., 2023; Zhao et al., 2025). These conversions were dominated by *B. paralicheniformis*, *B. glycinifermentans*, *Bre. borstelensis*, *Bacillus* sp. BT1B_CT2, *A. hydrophila*, *L. adecarboxylata*, *D. cercidiphylli*, *S. viridis*, *V. proomii*, *C. thermoamylovorans*, *C. hisashii*, and *Caldibacillus* sp. NES_3. Meanwhile, isobutyric acid was mainly produced from leucine metabolism mediated by *B. subtilis*, *B. sonorensis*, *B. licheniformis*, *Staphylococcus hominis*, *Bacillus* sp. BT1B_CT2, *A. hydrophila*, *L. adecarboxylata*, and *D. cercidiphylli*. Fatty acid metabolism also constituted essential pathways for the formation of alcohols, aldehydes, and ketones in *Cheonggukjang*. Lipids were degraded to free fatty acids by microbial lipases, and further converted into unsaturated fatty acids (such as linoleic acid and arachidonic acid) by *B. amyloliquefaciens*, *B. velezensis*, and *Bacillus* sp. 5B6. These unsaturated fatty acids were subsequently transformed into C6/C8 aldehydes and alcohols such as and hexanal, and 1-octen-3-ol, with the latter being further oxidized into 3-octanone, which contributes a fruity odor (Mathatheeranan et al., 2024). These reactions were mediated by methylglyoxal synthase (EC: 1.1.1.1) derived from *B. amyloliquefaciens*, *B. velezensis*, *Bacillus* sp. 5B6, and *Bacillus* sp. LM 4–2.

In this study, six strains, namely *C. thermoamylovorans*, *B. licheniformis*, *B. velezensis*, *B. subtilis*, *B. paralicheniformis*, and *C. hisashii*, were identified as core functional microorganisms contributing to the characteristic flavor formation of *Cheonggukjang*, based on the three criteria of high and stable abundance, strong correlation with key flavor compounds, and possession of flavor synthesis-related enzyme genes. Therefore, the formation of *Cheonggukjang*'s characteristic flavor can be summarized as follows: total sugar, as the core driving factor, directs the succession of the microbial community. The core functional microorganisms secrete extracellular enzymes such as proteases, amylase, and lipases to degrade the soybean substrate, generating key flavor precursors including amino acids and fatty acids. These precursors (mainly threonine, phenylalanine, leucine, tyrosine, acetoin, and linoleic acid) are subsequently converted by core functional microorganisms via pathways such as amino acid metabolism, and the lipoxygenase pathway into characteristic flavor compounds, including pyrazines, aldehydes, and alcohols, ultimately shaping the unique flavor profile of *Cheonggukjang*. By establishing a comprehensive mechanism that integrates substrates, microorganisms, enzymes, precursors, and characteristic flavor compounds, this study elucidated the metabolic pathways responsible for flavor formation during *Cheonggukjang* fermentation. These findings offer a theoretical basis for the precise modulation of *Cheonggukjang* flavor and support its industrial-scale production.

This study systematically elucidated the microbial regulatory mechanisms underlying flavor formation in *Cheonggukjang* through an integrated analysis of metagenomics, volatilomics, and untargeted metabolomics. Nevertheless, several limitations should be acknowledged. First, while untargeted metabolomics enabled comprehensive metabolite profiling and the identification of core flavor-associated metabolic pathways, its inherent technical constraints prevent absolute quantification of key flavor precursors, thereby limiting accurate assessment of the conversion efficiency from precusors to flavor

compounds. Second, the core functional microorganisms identified through multi-omics association analysis require further functional validation via in vitro fermentation experiments. Future work will therefore focus on: (i) performing absolute quantification of core flavor precursors and key metabolic nodes using targeted metabolomics to precisely characterize flavor compound biosynthesis and refine a quantitative mechanistic model of flavor formation in *Cheonggukjang*; and (ii) isolating and purifying the core functional microorganisms and validating their flavor-regulating functions through single-strain or synthetic microbial community fermentation. These efforts will ultimately support the development of standardized starters capable of precisely regulating flavor, thereby advancing the industrial standardization of *Cheonggukjang* production.

5. Conclusion

This study integrated metagenomic, volatilomic, and metabolomic analyses to reveal the flavor formation mechanism of *Cheonggukjang*, elucidating the functional associations between its core microbiota and key flavor compounds to establish a scientific basis for targeted quality control. Analysis of volatile compounds revealed that the *Cheonggukjang* fermentation process comprises three distinct stages (0–18, 18–60, 60–72 h), of which the final stage was designated as the most critical for flavor development. Fifteen key flavor compounds were identified during *Cheonggukjang* fermentation, with 10 serving as key flavor markers of different fermentation stages. Amino acid metabolism mediated by microorganisms during *Cheonggukjang* fermentation was responsible for the formation of these key flavor components. LefSe analysis revealed that *B. subtilis*, *B. velezensis*, *C. thermoamylovorans*, and *B. licheniformis* were the biomarkers for the 9, 18, 36, and 60 h fermentation periods of *Cheonggukjang*, respectively. Total sugar was the main factor driving microbial community succession, followed by reducing sugar and pH. *C. thermoamylovorans*, *B. licheniformis*, *B. velezensis*, *B. subtilis*, *B. paralicheniformis*, and *C. hisashii* were the core functional microbes for characteristic flavor formation. Collectively, this research reveals the development and formation mechanisms of characteristic flavor quality mediated by microorganisms during *Cheonggukjang* fermentation, providing theoretical guidance for standardized production and quality control of *Cheonggukjang* flavor.

Achnowledgment

This work was supported by A key joint grant for regional innovation from the National Natural Science Foundation of China (Grant No. U22A20547).

CRediT authorship contribution statement

Shanshan Zhao: Writing – original draft, Visualization, Formal analysis, Conceptualization. **Yuhang Sai:** Visualization, Methodology, Investigation, Formal analysis, Data curation. **Mengyu Jia:** Software, Investigation. **Yu Qiao:** Methodology, Investigation, Formal analysis. **Weishu Guo:** Investigation, Data curation. **Wenjing Ding:** Validation, Software, Data curation. **Xin Shao:** Validation, Methodology. **Yan Zheng:** Writing – review & editing, Supervision, Resources, Project administration, 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.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.fochx.2026.103756.

Data availability

Data will be made available on request.

References

An, F., Li, M., Zhao, Y., Zhang, Y., Mu, D., Hu, X., You, S., Wu, J., & Wu, R. (2021). Metatranscriptome-based investigation of flavor-producing core microbiota in different fermentation stages of dajiang, a traditional fermented soybean paste of Northeast China. *Food Chemistry, 343*, Article 128509. https://doi.org/10.1016/j.foodchem.2020.128509

Andersson, M., Laukkanen, M., Nurmiaho-Lassila, E. L., Rainey, F. A., Niemelä, S. I., & Salkinoja-Salonen, M. (1995). *Bacillus thermosphaericus sp. nov. a new thermophilic ureolytic: Bacillus isolated from air. Systematic and Applied Microbiology, 18*(2), 203–220. doi: 10.1016/S0723-2020(11)80391-7.

Cho, K. M., Lim, H.-J., Kim, M.-S., Kim, D. S., Hwang, C. E., Nam, S. H., … Shin, E. C. (2017). Time course effects of fermentation on fatty acid and volatile compound profiles of *Cheonggukjang* using new soybean cultivars. *Journal of Food and Drug Analysis, 25*(3), 637–653. https://doi.org/10.1016/j.jfda.2016.07.006

Hao, Z., Feng, J., Chen, Q., Lin, H., Zhou, X., Zhuang, J., … Yu, B. (2023). Comparative volatiles profiling in milk-flavored white tea and traditional white tea Shoumei via HS-SPME-GC-TOFMS and OAV analyses. *Food Chemistry: X, 18*, Article 100710. https://doi.org/10.1016/j.fochx.2023.100710

Hong, J. H., Jin, Y. H., Pawluk, A. M., & Mah, J. H. (2024). Metagenomic and culturomic analyses of bacterial species contributing to tyramine formation in *Cheonggukjang*. *LWT- Food Science and Technology, 201*, Article 116265. https://doi.org/10.1016/j.lwt.2024.116265

Huang, Y., Wan, J., Wang, Z., Sun, M., Feng, T., Ho, C.-T., & Song, S. (2022). Variation of volatile compounds and corresponding aroma profiles in Chinese steamed bread by various yeast species fermented at different times. *Journal of Agricultural and Food Chemistry, 70*(12), 3795–3806. https://doi.org/10.1021/acs.jafc.2c00550

Jia, Y., Niu, C. T., Lu, Z. M., Zhang, X. J., Chai, L. J., Shi, J. S., … Li, Q. (2020). A bottom-up approach to develop a synthetic microbial community model: Application for efficient reduced-salt broad bean paste fermentation. *Applied and Environmental Microbiology, 86*(12), Article e00306–00320. https://doi.org/10.1128/AEM.00306-20

Jian, C., Sun, M., Ma, T., Wang, W., Lv, B., Wang, J., Su, X., Li, S., & Guo, Y. (2024). Revealing the formation mechanisms of key flavor components during the fermentation of bamboo shoots by combining flavoromics and metagenomics. *Food Research International, 198*, 115361. https://doi.org/10.1016/j.foodres.2024.115361

Jin, Y. H., & Mah, J. H. (2024). Culture-dependent and -independent analyses of bacterial compositions and its contributions to formation of γ-aminobutyric acid and poly-γ-glutamic acid in *Cheonggukjang*. *Food Research International, 179*, Article 114026. https://doi.org/10.1016/j.foodres.2024.114026

Kim, M. H., Han, S. Y., Ko, J. M., & Kim, Y. S. (2012). Degradation characteristics of proteins in *Cheonggukjang* (fermented unsalted soybean paste) prepared with various soybean cultivars. *Food Science and Biotechnology, 21*(1), 9–18. https://doi.org/10.1007/s10068-012-0002-8

Lee, H. Y., Park, Y. M., Shin, D. Y., Hwang, H. M., Jeong, H. N., Park, H. Y., … Kim, J. G. (2024). Immune-enhancing effect of fermented soybean food, *Cheonggukjang* on cyclophosphamide-treated immunosuppressed rat. *Heliyon, 10*(18), Article e37845. https://doi.org/10.1016/j.heliyon.2024.e37845

Lee, S. G., & Chang, H. C. (2017). Assessment of *Bacillus subtilis* SN7 as a starter culture for Cheonggukjang, a Korean traditional fermented soybean food, and its capability to control *Bacillus cereus* in Cheonggukjang. *Food Control, 73*, 946–953. https://doi.org/10.1016/j.foodcont.2016.10.015

Li, B., Tang, J., Zhu, C., Gong, J., Wu, Q., Yang, Y., … Huang, Y. (2026). Analyzing the metabolic mechanisms and regulation of flavor compounds in baijiu from a microbial sociology perspective. *Food Bioscience*. , Article 108546. https://doi.org/10.1016/j.fbio.2026.108546

Li, Z., Rui, J., Li, X., Li, J., Dong, L., Huang, Q., … Chen, F. (2017). Bacterial community succession and metabolite changes during doubanjiang-meju fermentation, a Chinese traditional fermented broad bean (*Vicia faba* L.) paste. *Food Chemistry, 218*, 534–542. https://doi.org/10.1016/j.foodchem.2016.09.104

Liang, J., Li, D., Shi, R., Wang, J., Guo, S., Ma, Y., & Xiong, K. (2019). Effects of microbial community succession on volatile profiles and biogenic amine during sufu fermentation. *LWT- Food Science and Technology, 114*, Article 108379. https://doi.org/10.1016/j.lwt.2019.108379

Lim, E. S. (2022). Influence of bacteriocin-producing *Bacillus* strains on quality characteristics of fermented soybean product with biogenic amine-forming lactic acid bacteria. *Applied Biological Chemistry, 65*(1), Article 5. https://doi.org/10.1186/s13765-021-00664-9

Martineau-Côté, D., Achouri, A., Pitre, M., Wanasundara, J., Karboune, S., & L'Hocine, L. (2023). Investigation of the nutritional quality of raw and processed Canadian faba bean (*Vicia faba* L.) flours in comparison to pea and soy using a human in vitro gastrointestinal digestion model. *Food Research International, 173*, Article 113264. https://doi.org/10.1016/j.foodres.2023.113264

Mathatheeranan, P., Wongprasert, T., Fang, M., Lu, T. J., & Suppavorasatit, I. (2024). Effect of drying techniques on the stability of volatile compounds and antioxidative activities of dried Thai fermented black soybean (thua nao). *Journal of Agriculture and Food Research, 18*, Article 101305. https://doi.org/10.1016/j.jafr.2024.101305

Melkonian, C., Zorrilla, F., Kjærbølling, I., Blasche, S., Machado, D., Junge, M., … Zeidan, A. A. (2023). Microbial interactions shape cheese flavour formation. *Nature. Communications, 14*(1), Article 8348. https://doi.org/10.1038/s41467-023-41059-2

Nam, Y. D., Yi, S. H., & Lim, S. I. (2012). Bacterial diversity of *cheonggukjang*, a traditional Korean fermented food, analyzed by barcoded pyrosequencing. *Food Control, 28*(1), 135–142. https://doi.org/10.1016/j.foodcont.2012.04.028

Park, M. K., Cho, I. H., Lee, S., Choi, H. K., Kwon, D. Y., & Kim, Y. S. (2010). Metabolite profiling of *Cheonggukjang*, a fermented soybean paste, during fermentation by gas chromatography-mass spectrometry and principal component analysis. *Food Chemistry, 122*(4), 1313–1319. https://doi.org/10.1016/j.foodchem.2010.03.095

Qi, C., Wen, B., Zheng, Y., Wang, S., Zhao, Y., Liu, Y., … Wang, W. (2025). Combining untargeted metabolomics and high-throughput sequencing to explore the dynamics of small-molecular metabolites in the fermentation stage of inner Mongolian cheese. *Journal of Agricultural and Food Chemistry, 73*(32), 20341–20351. https://doi.org/10.1021/acs.jafc.5c03520

Rajasundaram, D., & Selbig, J. (2016). More effort — More results: Recent advances in integrative "omics" data analysis. *Current Opinion in Plant Biology, 30*, 57–61. https://doi.org/10.1016/j.pbi.2015.12.010

Regueiro, J., Negreira, N., & Simal-Gandara, J. (2017). Challenges in relating concentrations of aromas and tastes with flavor features of foods. *Critical Reviews in Food Science and Nutrition, 57*(10), 2112–2127. https://doi.org/10.1080/10408398.2015.1048775

Sanjukta, S., & Rai, A. K. (2016). Production of bioactive peptides during soybean fermentation and their potential health benefits. *Trends in Food Science & Technology, 50*, 1–10. https://doi.org/10.1016/j.tifs.2016.01.010

Shahid, A., Fan, Z., Su, K., Zhao, A., Mehmood, M. A., Chang, J.-S., … Xu, J. (2025). Sensory chemistry of Spirulina: Unveiling trends and innovations in aromatic volatile organic compound biosynthesis in off-flavors and odor mitigation strategies. *Trends in Food Science & Technology, 156*, Article 104886. https://doi.org/10.1016/j.tifs.2025.104886

Shashank, M., Ashish, S., Ambarish, S., & Ghosh, S. (2014). Transformation of ferulic acid to 4-vinyl guaiacol as a major metabolite: A microbial approach. *Reviews in Environmental Science and Biotechnology, 13*, 377–385. https://doi.org/10.1007/s11157-014-9348-0

Tamang, J. P., Das, S., Kharnaior, P., Pariyar, P., Thapa, N., Jo, S.-W., … Shin, D.-H. (2022). Shotgun metagenomics of *Cheonggukjang*, a fermented soybean food of Korea: Community structure, predictive functionalities and amino acids profile. *Food Research International, 151*, Article 110904. https://doi.org/10.1016/j.foodres.2021.110904

Tang, A., & Peng, B. (2024). Metatranscriptomics reveals microbial community function succession and characteristic flavor formation mechanisms during black rice wine fermentation. *Food Chemistry, 457*, Article 140428. https://doi.org/10.1016/j.foodchem.2024.140428

Tang, H., Han, C., Liu, Y., & Wang, D. (2013). Analysis of nutrient changes in traditional fermented soybean paste. *Food and Fermentation Industries, 39*(4), Article 4. https://doi.org/10.13995/j.cnki.11-1802/ts.2013.04.028

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

Wang, J., Wang, X. S., Zhang, Z., Zhou, D. Y., Huang, X. H., & Qin, L. (2024). Comprehensive insights into the organoleptic characteristics attributes of "HuangYuXiang": Integration of volatilomics, sensomics, macrogenomics, lipidomics, and metabolomics. *Food Chemistry, 460*, Article 140409. https://doi.org/10.1016/j.foodchem.2024.140409

Wen, L., Lei, J., Yang, L., Kan, Q., Wang, P., Li, J., Chen, C., He, L., Fu, J., Ho, C.-T., Huang, Q., & Cao, Y. (2024). Metagenomics and untargeted metabolomics analyses to unravel the formation mechanism of characteristic metabolites in Cantonese soy sauce during different fermentation stages. *Food Research International, 181*, Article 114116. https://doi.org/10.1016/j.foodres.2024.114116

Xiong, S., Xu, X., Zhang, L., Du, T., Huang, T., Huang, J., Ren, H., Xiong, T., & Xie, M. (2024). Integrated metatranscriptomics and metabolomics reveal microbial succession and flavor formation mechanisms during the spontaneous fermentation of Laotan Suancai. *Food Research International, 177*, Article 113865. https://doi.org/10.1016/j.foodres.2023.113865

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, P., Zhang, P., Wu, J., Tao, D., & Wu, R. (2019). Effects of *Leuconostoc mesenteroides* on physicochemical and microbial succession characterization of soybean paste, Da-jiang. *LWT - Food Science and Technology, 115*, Article 108028. https://doi.org/10.1016/j.lwt.2019.04.029

Zhao, S., Yue, J., Wang, Y., Shao, J., Li, Z., & Li, M. (2024). The regulation of volatile flavor compounds in fermented meat products mediated by microorganisms: A review. *Food Bioscience, 62*. https://doi.org/10.1016/j.fbio.2024.105180

Zhao, S., Zhang, Z., Qiao, Y., Li, M., Chen, J., Zheng, Y., & Yue, X. (2025). Influence of environmental factors on microbial community and flavor profile of dajiang of Liaoning. *Food Bioscience, 63*, Article 105661. https://doi.org/10.1016/j.fbio.2024.105661

Zhao, S. S., Sai, Y. H., Liu, W. T., Zhao, H. W., Bai, X., Song, W. Y., … Yue, X. Q. (2023). Flavor characterization of traditional fermented soybean pastes from Northeast China and Korea. *Foods, 12*(17), Article 3294. https://doi.org/10.3390/foods12173294

Zhou, C., Kong, Y., Zhang, N., Zhang, X., Qin, W., Zhang, L., … Lu, F. (2025). Transcriptomic analysis of *Bacillus licheniformis* 2709 reveals the molecular mechanism of alkaline protease biosynthesis regulated by the DegS/DegU two-component system. *International Journal of Biological Macromolecules, 306*, Article 140868. https://doi.org/10.1016/j.ijbiomac.2025.140868