separations

MDPI

Article

UPLC-MS/MS-Based Metabolite Profiling Reveals Origin-Dependent Chemical Variation in *Curcuma longa* and Its Relationship with Shuangjiang Weitong Pills

Wuxiang Fu [1,2,3], Qihuan Ding [1], Rongxu Zhou [1], Shanshan Wu [1], Maroof Ali [2,3,*] and Zhanxia Ma [1,4,*]

[1] School of Pu'er Tea, West Yunnan University of Applied Sciences, Pu'er 665000, China; fuwuxiang@xtbg.ac.cn (W.F.)
[2] Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Menglun 666303, China
[3] Southwest Cross-Border Biodiversity and Ecological Security Team, Key Laboratory of Tropical Forest Ecology, Xishuangbanna Tropical Botanical Garden, Chinese Academy of Sciences, Mengla 666303, China
[4] College of Life Sciences, Zhejiang University, Hangzhou 310058, China
* Correspondence: maroof@xtbg.ac.cn (M.A.); mazhanxia@xtbg.ac.cn (Z.M.)

Abstract

Curcuma longa L. is a principal botanical component of Shuangjiang Weitong Pills (SWPs), a Dai medicine. We used ultra-high-performance liquid chromatography–tandem mass spectrometry (UPLC-MS/MS)-based metabolite profiling to characterize origin-associated variations among 16 *C. longa* samples from Xishuangbanna, Dehong, Gengma, and Menglian, Yunnan, and to compare their relative profiles with five independently prepared composite replicates from one SWP batch. After quality control filtering, 1853 putatively annotated features were retained, comprising 948 positive-ion and 905 negative-ion features annotated independently. An origin-only PERMANOVA indicated multivariate origin association ($R^2 = 0.595$, $p = 0.0001$), accompanied by unequal dispersions; five of six pairwise tests remained significant after BH correction. XSBN was nearest SWP in positive-ion PCA and clustering, but mode-specific similarity measures gave different perspectives: GM had the highest Spearman correlation in positive (rho = 0.6650) and negative (rho = 0.6954) modes. DH had the lowest untransformed Bray–Curtis dissimilarity, but this result was dominated by intense signals and the positive-mode ranking changed after logarithmic transformation. Comparisons with SWP yielded 562, 545, 497, and 477 differentially abundant annotated features for XSBN, DH, GM, and ML, respectively, including 167 shared annotated features. Plant-reference pathway analysis found no significant enrichment after BH correction (FDR < 0.05), limiting pathway interpretation to descriptive annotation. Metric-dependent rankings and representation-dependent DH correlation precluded designation of a uniquely closest origin. These findings describe the measured global chemical relationships between the sampled *C. longa* materials and a single SWP production batch. Consequently, the observed similarities cannot identify the *C. longa* source used in the formulation or establish manufacturing provenance, source superiority, or therapeutic equivalence. The shared differential features remain exploratory candidates requiring further validation. This study provides a reference for characterizing origin-associated chemical variation in *C. longa* and its relationship with the overall chemical profile of the formulation, supporting future targeted validation of candidate quality control features and studies across multiple production batches.

Keywords: *Curcuma longa*; Shuangjiang Weitong Pills; UPLC-MS/MS; metabolite profiling; geographical origin; Dai medicine; chemical similarity

Check for updates

Academic Editor: Filomena Lelario

Received: 17 October 2026
Revised: 17 November 2026
Accepted: 17 November 2026
Published: 20 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

Curcuma longa L. (turmeric) is a medicinal and food plant with a long history of use in Asian medical systems, including Ayurveda and traditional Chinese medicine [1,2]. Its rhizomes contain curcuminoids, sesquiterpenoids, phenolic compounds, and other specialized metabolites [3]. These constituents have been investigated for biological activities, including antioxidant and anti-inflammatory effects [4]. Selected turmeric extracts have undergone clinical evaluation [5], drying can affect their metabolite profiles [6], and the bioactive effects and safety of turmeric and curcumin have been reviewed [7].

Dai medicine is practiced in Yunnan Province, where its modernization and prescription resources have been investigated [8,9]. Shuangjiang Weitong Pills are a marketed Dai medicinal formulation containing *C. longa*, *Zingiber purpureum*, *Stephania epigaea*, *Acorus tatarinowii*, and *Brassica integrifolia*. Zingiberaceae species are widely used in Dai medicine, with diverse traditional applications and chemical constituents [10]. Because *C. longa* is a principal plant material in SWPs, variation in its chemical composition is relevant to raw material consistency and formulation manufacture [11,12].

The chemical composition of *C. longa* can vary with germplasm, geographical origin, cultivation conditions, and post-harvest processing. Metabolomics and spectroscopic studies have discriminated turmeric samples by origin and have shown variations in curcuminoids, terpenoids, and other metabolite classes [13–15]. Interpretation is further complicated by taxonomic and nomenclatural overlap among *Curcuma* medicines used under different traditional names, including Yujin and Ezhu [16–18]. These sources of variation can affect the chemical reproducibility of raw materials supplied for complex herbal products.

Most previous research has examined *C. longa* as an individual medicinal plant, whereas direct comparisons between multiple geographical sources and a finished Dai formulation remain limited. Such comparisons require caution because SWPs contain several plant species and undergo pharmaceutical processing. Similarity between a *C. longa* sample and SWPs therefore represents resemblance of their measured global chemical profiles; it does not establish that a particular raw material origin produced the formulation or that chemical proximity is equivalent to clinical efficacy. Within these limits, metabolomics can describe the measured chemical relationships between plant materials and complex formulations [6,19,20]. Here, the analysis provides a descriptive reference framework for comparing candidate *C. longa* supplies against the formulation profile, rather than a criterion for raw material acceptance or source identification. The contributions of the other ingredients and processing may influence or mask *C. longa*-associated signals, and this design does not isolate the *C. longa* component within SWPs.

In this study, UPLC-MS/MS-based metabolite profiling used MS1 peak areas for relative comparisons and MS/MS spectra for putative annotation. Feature extraction, quality control filtering, and spectral library annotation yield relative abundance profiles suitable for multivariate comparison [21–23]. This approach is particularly useful for chemically complex plant materials and multi-herb formulations, although structural confidence varies among annotations and the resulting measurements are generally relative rather than absolute [20].

We therefore applied UPLC-MS/MS-based metabolite profiling to characterize origin-associated chemical variation among *C. longa* samples from four Dai-inhabited regions of Yunnan. We also compared their relative profiles with the complete formulation profile of one SWP production batch, identified differentially abundant annotated features shared across the four origin-versus-formulation comparisons, and explored pathways associated with their putative annotations. An additional curcuminoid-focused comparison examined relative signals without assigning exclusive botanical sources within SWPs.

2. Materials and Methods

2.1. Sample Sources and Preparation

Fresh *C. longa* rhizomes were collected from four Dai-inhabited regions of Yunnan Province, China: Xishuangbanna Dai Autonomous Prefecture (XSBN), Dehong Dai and Jingpo Autonomous Prefecture (DH), Gengma Dai and Wa Autonomous County (GM), and Menglian Dai, Lahu and Wa Autonomous County (ML) in mid-December 2024 (Figure S1 and Table S13).

Three to five distinct sites were sampled within each region. At each site, rhizomes from 10 plants were pooled to form one composite sample. The sampling site was the independent sampling unit, and each site-level composite sample was treated as one biological replicate. Replicates within each region originated from distinct sampling sites rather than subsampling or repeated preparation of the same pooled material. A total of 16 composite biological replicates were analyzed: XSBN, $n = 5$; DH, $n = 5$; GM, $n = 3$; and ML, $n = 3$.

Rhizomes were harvested at physiological maturity, approximately 9 months after planting. After harvest, they were washed with distilled water, rinsed with ultrapure water, and cut into slices approximately 3 mm thick. The slices were sun-dried for 1 week, ground into powder, and stored at −80 °C until extraction and analysis.

SWPs (National Medical Products Administration approval no. Z20026657) were purchased from Yunnan Pharmaceutical Co., Ltd., located in Kunming, Yunnan Province, China (Figure S1). Five preparation replicates were generated from a single production batch. For each replicate, five subsamples from that batch were combined in equal proportions. These replicates capture within-batch preparation and analytical variation only; they are not biological replicates or independent production batches.

Approximately 200 mg of powdered *C. longa* or SWP material was transferred to a 2 mL centrifuge tube [21]. Methanol (600 µL) containing 2-chloro-L-phenylalanine at 4 ppm was added as the extraction solvent and internal reference, and the mixture was vortexed for 30 s. After addition of a steel bead, samples were homogenized at 55 Hz for 60 s, sonicated at room temperature for 15 min, and centrifuged at 12,000 rpm and 4 °C for 10 min. The supernatant was passed through a 0.22 µm membrane before instrumental analysis.

A pooled quality control (QC) sample was prepared by combining equal aliquots of all extracts. Study samples were injected in randomized order. Pooled QC samples were used for initial system conditioning and were subsequently injected after every five study samples to monitor analytical stability.

2.2. UPLC-MS Data Acquisition

Chromatographic separation was performed on a SCIEX ExionLC AD UHPLC system (SCIEX, Framingham, MA, USA) equipped with an ACQUITY UPLC HSS T3 column (2.1 mm × 100 mm, 1.8 µm; Waters, Milford, MA, USA) at 40 °C, with a flow rate of 0.30 mL/min and an injection volume of 2 µL. For positive-ion analysis, mobile phases A and B were water and acetonitrile, respectively, each containing 0.1% formic acid. For negative-ion analysis, mobile phase A was 5 mM ammonium formate in water and mobile phase B was acetonitrile. The gradient was 8% B at 0–1 min, 8–98% B at 1–8 min, 98% B at 8–10 min, 8–98% B at 10–10.1 min, and 8% B at 10.1–12 min.

High-resolution MS and MS/MS spectra were acquired in separate positive- and negative-ion modes on a SCIEX TripleTOF 5600+ mass spectrometer (AB SCIEX, Concord, ON, Canada) using information-dependent acquisition. The acquisition ranges were segmented at m/z 100–300, 300–450, 450–600, 600–750, and 750–1200. Ion-source settings were as follows: ion-spray voltage, +5500/−4500 V; curtain gas, 35 psi; source temperature, 600 °C; ion-source gases 1 and 2, 60 psi; and declustering potential, ±100 V. The

results presented here are based on TripleTOF MS1 peak areas and MS/MS-supported putative annotations.

2.3. Data Processing and Metabolite Annotation

TripleTOF raw files were converted to mzXML format with MSConvert in ProteoWizard version 3.0.8789 [24]. Feature detection, retention time correction, and alignment were performed using R XCMS version 3.12.0 and the centWave algorithm [25]. The principal settings were ppm = 15, peakwidth = c(5, 30), and mzdiff = 0.01. The QC-corrected peak areas in Table S1 were derived from TripleTOF data and were used for the reported feature-level comparisons. The workflow therefore provides broad relative profiling rather than absolute concentration measurements.

Metabolite annotation used accurate precursor mass, adduct information, and available MS/MS spectral matches against HMDB [26], LIPID MAPS [27], KEGG [28], ChEBI [29], and a proprietary spectral database maintained by Panomix Biomedical Tech Co., Ltd. (Suzhou, China).

Systematic signal drift was corrected by support vector regression (SVR) based on the pooled QC samples, and features with QC relative standard deviation (RSD) > 30% were excluded [30]. No authentic reference standards were analyzed under the study conditions. Compound names are therefore reported as putative annotations rather than confirmed identities. Available experimental and reference MS/MS spectral comparisons are provided in the identi_plot directory for feature-level evaluation.

2.4. Data Preprocessing and Multivariate Analysis

Sample labels, species codes, replicate identifiers, ionization modes, and feature identifiers were checked before analysis, and missingness was summarized by sample and feature. Sample total signal, missingness, pairwise correlations, QC clustering, and PCA were examined jointly to identify potential outliers. For PCA and OPLS-DA, remaining zero or missing entries were replaced with half the smallest positive value in the corresponding input matrix, followed by logarithmic transformation, mean centering, and unit-variance scaling [31]. Missing values were not replaced with zero.

PCA and orthogonal partial least squares discriminant analysis (OPLS-DA) were performed using ropls version 1.22.0 [32]. Separate OPLS-DA models compared XSBN, DH, GM and ML with SWPs using the combined matrix of 1853 features after logarithmic transformation, mean centering and unit-variance scaling. Hierarchical clustering and heatmaps used Z-score-standardized relative abundances. Q^2 was estimated by seven-fold cross-validation, and model performance was examined using 200 permutations. R^2Y describes model fit, whereas Q^2 estimates cross-validated predictive performance; neither statistic was interpreted as independent external validation.

For the separate relative-profile similarity analysis, QC-corrected positive- and negative-ion peak-area matrices were aligned by sample identifier and normalized independently to each sample's total signal within each mode. Spearman correlation and Bray–Curtis dissimilarity were calculated separately for the 948-feature positive-mode matrix, the 905-feature negative-mode matrix, and their 1853-feature concatenation. Group profiles were defined by feature-wise medians after sample normalization and were not re-normalized; logarithmic transformation and unit-variance scaling were not applied. Uncertainty was estimated using 2000 two-stage stratified bootstrap iterations, first resampling replicates within groups and then features within modes. Sample indices were paired across modes, feature indices were shared across groups, and the SWP resample was shared across origin comparisons. Normalization preceded resampling; denominators were not recalculated after feature resampling. The 2.5th and 97.5th percentiles and first-rank frequencies are conditional on the

single analyzed SWP production batch. For a sensitivity analysis, each sample-normalized signal p was transformed as $\log2(1 + 10^6\ p)$, after which feature-wise group medians and Bray–Curtis dissimilarities were recalculated without further normalization (Table S12).

2.5. Differential Feature Screening and Complementary Analyses

For each origin-versus-SWP comparison, two-sided Welch independent-samples *t*-tests were applied to QC-corrected TripleTOF peak areas without logarithmic transformation or total-signal normalization. *p* values were adjusted across all 1853 features within each comparison using the Benjamini–Hochberg method. Fold change was the mean peak area in the *C. longa* origin group divided by that in SWPs; VIP values were obtained from the corresponding origin-versus-SWP comparison. Differentially abundant annotated features (DAMs) were required to satisfy all three criteria: variable importance in projection > 1.2, fold change ≥ 2.0 or ≤ 0.5, and a Benjamini–Hochberg false-discovery-rate-adjusted q-value < 0.05 [33]. The 10 exploratory candidates were selected from the 167 shared DAMs based on consistent change directions across the four origin-versus-SWP comparisons and representation of organic acid, phenolic, flavonoid, and terpenoid annotations. The shortlist was used for hypothesis generation, not as a set of validated quality control markers. This descriptive shortlist does not establish specificity to the *C. longa* component of SWPs. A separate annotation-led comparison of curcumin, demethoxycurcumin, bisdemethoxycurcumin and hexahydrocurcumin used these same Welch tests and all-feature BH corrections; inclusion in this focused panel did not require the DAM screening thresholds.

Plant-reference over-representation analysis used the KEGG Arabidopsis thaliana catalogue (ath; retrieved 10 November 2026) [28], with compound memberships from corresponding reference maps. We retained 102 individual metabolic pathways with numerical IDs below 01100, excluding global/overview and non-metabolic pathways. The 1853 measured features contained 1371 unique KEGG compound IDs; 616 mapped to this library and formed the background. Compound IDs were counted once. One-sided hypergeometric tests [34] were BH-corrected across all 89 pathways containing at least one background ID, including zero-hit pathways, separately within each comparison. Arabidopsis serves as a plant proxy; these annotation-based comparisons do not establish *C. longa*-specific pathways, pathway activity or ingredient-specific sources.

Origin-associated variation was tested separately in the 16 *C. longa* samples, excluding SWPs. The QC-corrected and QC-RSD-filtered matrix was natural-log transformed, centered, and scaled to unit variance; total-signal normalization was not applied. Overall differences were tested by Euclidean-distance PERMANOVA with 9999 permutations using vegan::adonis2 [35]. Six exact pairwise tests enumerated all allocations that preserved group sizes, and their *p* values were adjusted by the Benjamini–Hochberg method. Exact *p* values have finite resolution determined by the available label allocations and were not interpreted as effect-size estimates.

Homogeneity of multivariate dispersion was assessed with vegan: betadisper using centroid distances, small-sample bias adjustment, and 9999 permutations [36]. PCA was used to visualize the same origin-only matrix. Site-level composite samples were the permutation units; these analyses test association with origin labels and do not establish causal geographical effects.

3. Results

3.1. Quality Assessment of the Metabolite Profiling Data

Representative base peak chromatograms show the profiles of the five study groups in positive- and negative-ion modes (Figure 1A,B); these panels are not a sequence-wide

overlay of pooled QC injections. Pooled QC and study samples were included together in PCA for both ionization modes (Figure 1C,D). The RSD distributions showed that 81.6% of positive-ion features and 83.1% of negative-ion features had QC RSD values < 30% (Figure 1E,F); features with QC RSD > 30% were excluded before statistical analysis. These distributions summarize QC variability before feature filtering, without establishing absolute quantification or external model validity.

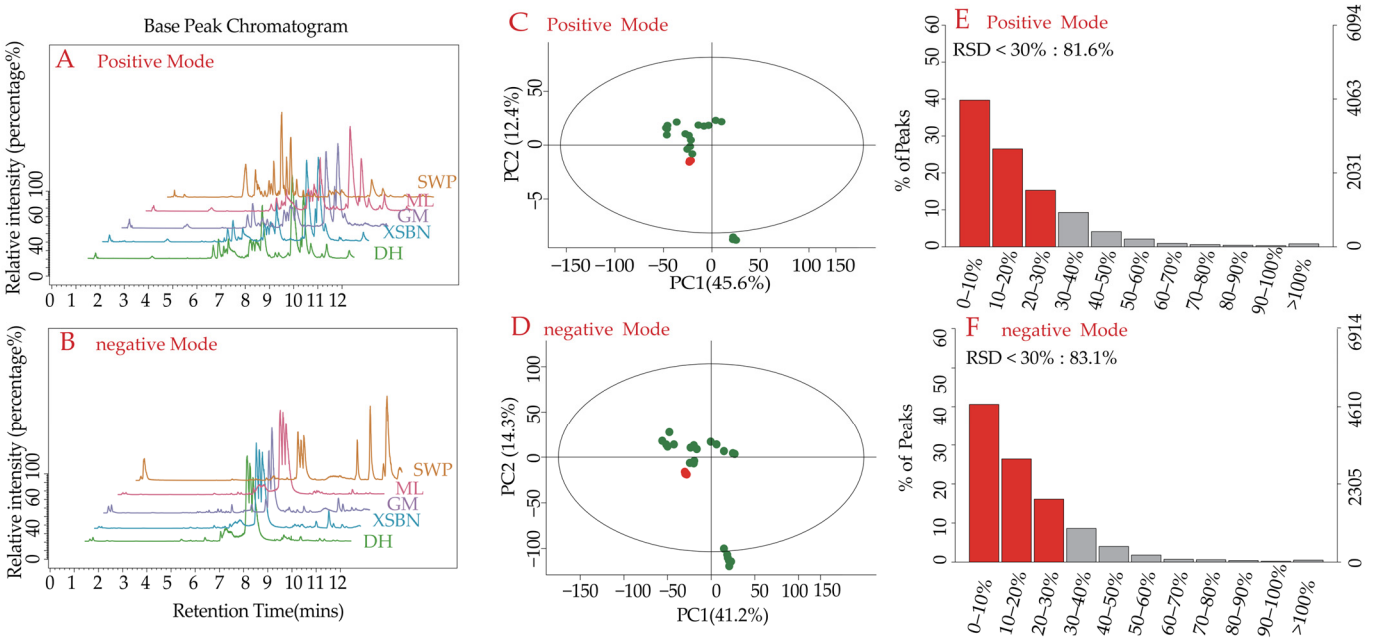

Figure 1. Quality assessment of the UPLC-MS/MS-based metabolite profiling data. (**A**,**B**) Representative base peak chromatograms of study samples in positive- and negative-ion modes, respectively. (**C**,**D**) PCA score plots of study samples and pooled QC samples in positive- and negative-ion modes, respectively; red dots indicate QC samples and green dots indicate study samples. (**E**,**F**) Distribution of relative standard deviations (RSDs) of detected peaks in positive and negative ion modes, respectively. Red bars indicate peaks with RSD < 30%, whereas gray bars indicate peaks with RSD ≥ 30%. SWPs, Shuangjiang Weitong Pills; XSBN, Xishuangbanna; DH, Dehong; GM, Gengma; ML, Menglian.

3.2. Global Chemical Variation Among C. longa Origins and SWPs

A total of 1853 putatively annotated features were retained across the analyzed samples, comprising 948 positive-ion and 905 negative-ion features (Table S1). Annotation was performed independently in each ionization mode, and features were not merged or deduplicated across modes; this total is therefore not a count of unique metabolites. In positive-ion mode, PC1, PC2, and PC3 explained 35.98%, 17.81%, and 12.50% of the variance, respectively, for a cumulative 66.29% (Figure 2A). The corresponding components in negative-ion mode explained 31.86%, 19.37%, and 11.06%, for a cumulative 62.29% (Figure 2B).

The four geographical groups and SWPs occupied distinct regions of the PCA space, indicating origin-associated variation and pronounced differences between the single herb and the multi-herb product. In positive-ion mode, XSBN samples were positioned closer to SWPs than the other raw material groups (Figure 2A). Hierarchical clustering (Figure S2) and the positive-ion heatmap showed the same relative tendency (Figure 2C). In negative-ion mode, however, SWPs formed a separate cluster from all four *C. longa* groups (Figure 2B,D). The observed XSBN-SWP proximity was therefore mode dependent and relative rather than evidence of chemical equivalence.

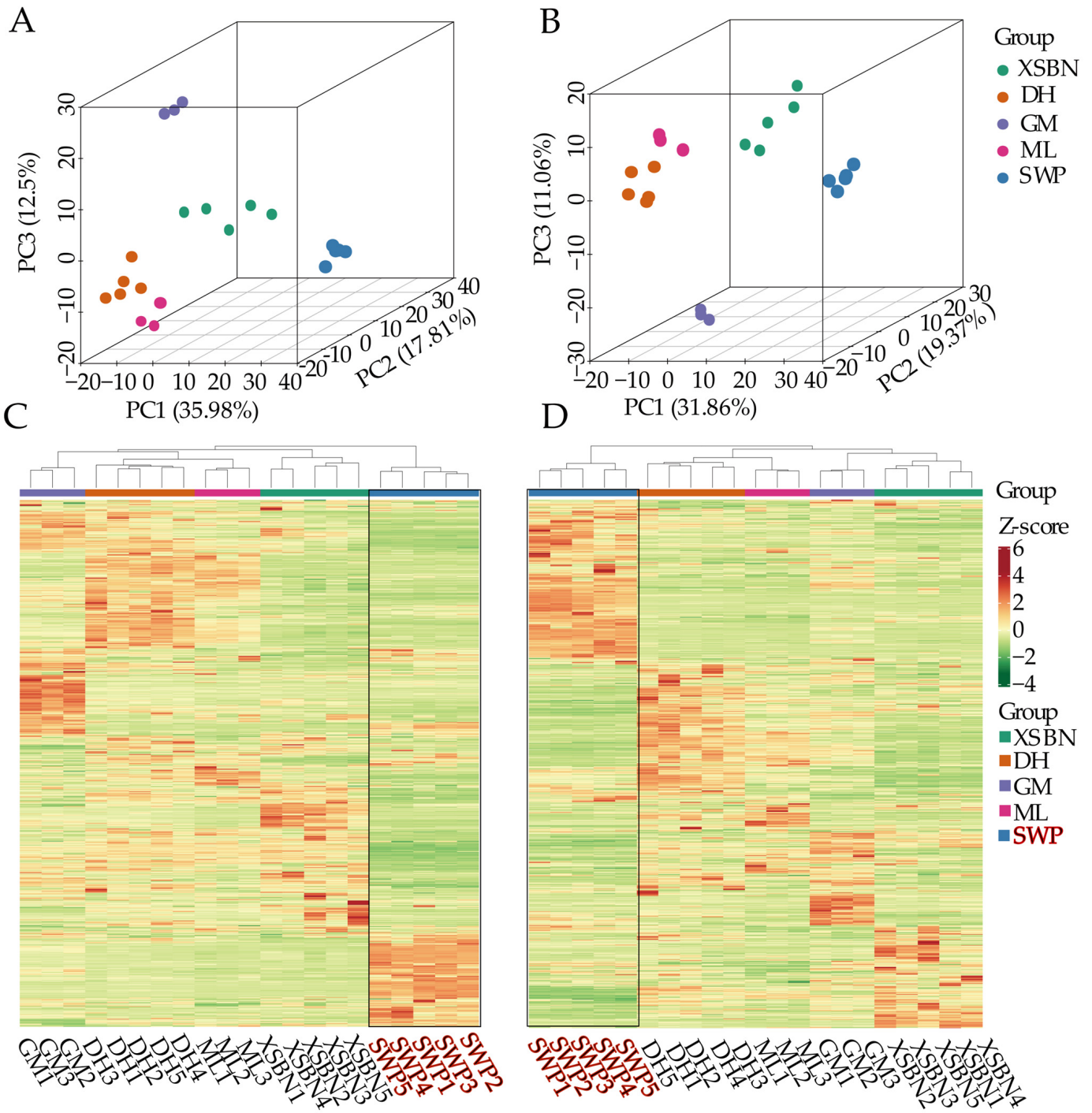

Figure 2. Global metabolic relationships among *C. longa* samples from four geographical origins and Shuangjiang Weitong Pills. (**A**,**B**) Three-dimensional PCA score plots in positive- and negative-ion modes, respectively. (**C**,**D**) Heatmaps of Z-score-standardized annotated feature abundances in positive- and negative-ion modes, respectively; red denotes higher and green denotes lower relative abundance.

A separate analysis restricted to the 16 *C. longa* samples directly tested variations among the four origins. PCA of the combined log/UV feature matrix explained 28.23% and 20.86% of the variances on PC1 and PC2, respectively (Figure S3). Overall PERMANOVA indicated an association between origin labels and multivariate profiles (pseudo-F(3,12) = 5.8751, R^2 = 0.5949, permutation p = 0.0001; Table S2). However, multivariate dispersion also differed among origins (PERMDISP F(3,12) = 11.9142, p = 0.0001; Figure S4B), so the

PERMANOVA result cannot be interpreted solely as separation of group centroids. Five pairwise comparisons had BH-adjusted $p = 0.0214$, whereas GM versus ML had exact $p = 0.1000$ and BH-adjusted $p = 0.1000$ (Table S3; Figure S4A). For sample sizes 5 versus 5, 5 versus 3 and 3 versus 3, the exact tests enumerate 252, 56 and 20 label allocations, with minimum raw p values of 2/252 (0.00794), 1/56 (0.01786) and 2/20 (0.1000), respectively. The five identical BH-adjusted values of 0.0214 reflect these discrete limits and correction across six comparisons, not equal effect sizes. Pairwise R^2 ranged from 0.4057 to 0.6639 (Table S3); these values require interpretation alongside the dispersion differences.

Four binary origin-versus-SWP OPLS-DA models were fitted using the combined feature matrix (Figure 3). For XSBN-SWP, DH-SWP, GM-SWP and ML-SWP, R^2X(cum) values were 0.648, 0.668, 0.719 and 0.716, respectively; R^2Y(cum) was reported as 1.000 for all four models; and Q^2(cum) values were 0.988, 0.993, 0.988 and 0.988, respectively. The corresponding RMSEE values were 0.0157, 0.0074, 0.0113 and 0.0145. Score plots are shown in Figure 3A–D and permutation tests in Figure 3E–H. The retained five-group score plots (Figure 3I,K) and their permutation tests (Figure 3J,L) are separate exploratory displays and are not the four binary models. These models summarize origin-versus-formulation contrasts and do not test differences among the four origins.

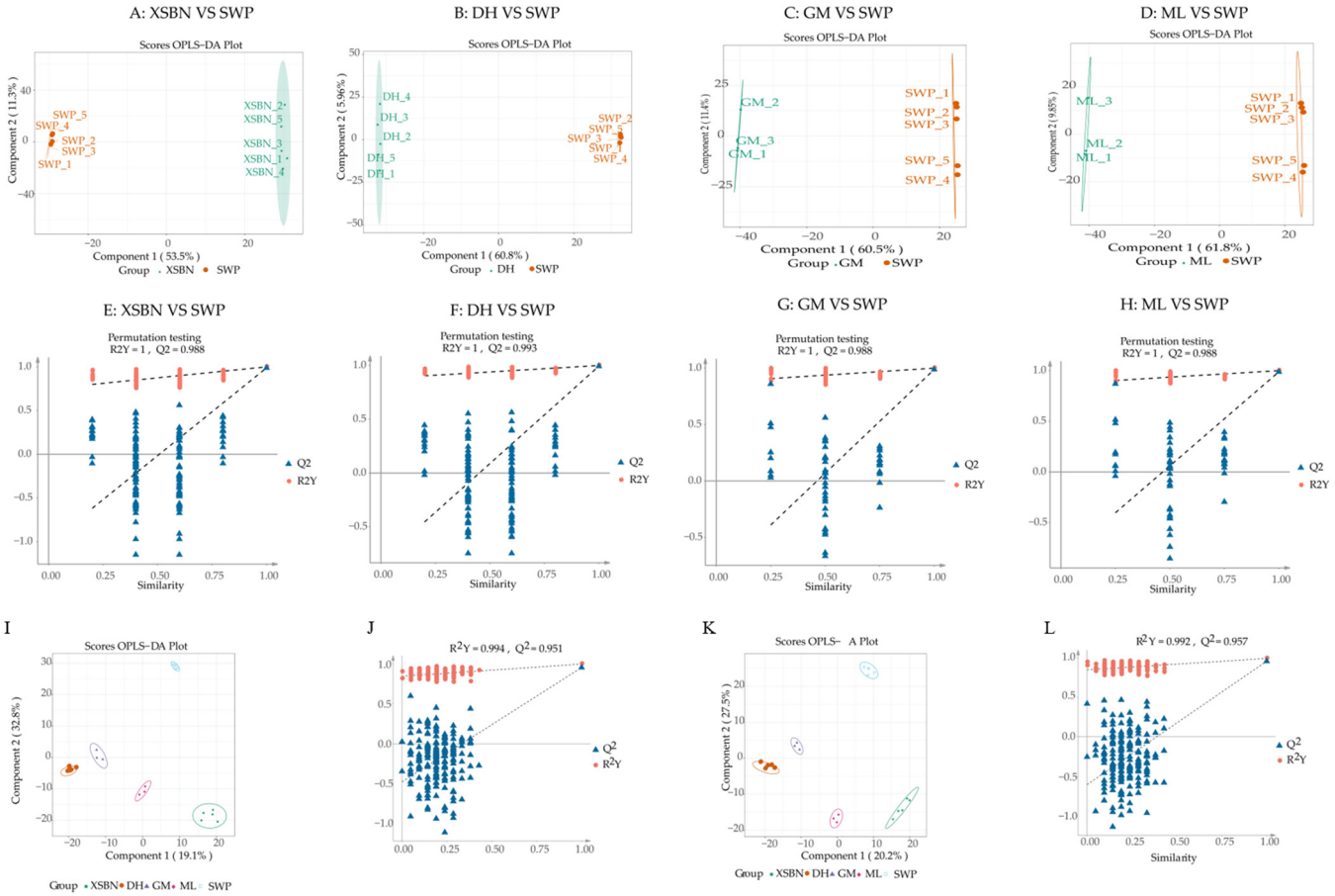

Figure 3. OPLS-DA score plots and permutation tests for metabolomic comparisons among the five sample groups. (**A**–**D**) Pairwise OPLS-DA score plots comparing XSBN, DH, GM, and ML with SWP, respectively. The percentages on the axes indicate the variance explained by the corresponding components. (**E**–**H**) Permutation test results for the four pairwise OPLS-DA models, showing the distributions of Q^2 (blue triangles) and R^2Y (red circles) against permutation similarity. The original model statistics are indicated in each panel. (**I**,**K**) OPLS-DA score plots showing the distribution of samples from the five groups (XSBN, DH, GM, ML, and SWP). The percentages on the axes represent the variances explained by the respective components. (**J**,**L**) Permutation test results for

the corresponding five-group OPLS-DA models. The Q^2 and R^2Y values of the original models are indicated in the respective panels. The shaded ellipses in the OPLS-DA score plots represent the group-specific sample distribution regions. In the permutation test plots, the solid horizontal line indicates the zero reference line ($y = 0$), while the dashed lines represent the linear regression lines of R^2Y and Q^2 against permutation similarity.

Each pairwise model was assessed using seven-fold cross-validation and 200 label permutations. Q^2 regression intercepts were −0.889, −0.797, −0.594 and −0.634 for XSBN-SWP, DH-SWP, GM-SWP and ML-SWP, respectively; the corresponding R^2Y intercepts were 0.751, 0.876, 0.891 and 0.884. Although the Q^2 intercepts were negative, the high original R^2Y/Q^2 values and relatively high R^2Y intercepts do not establish external predictive validity or exclude optimistic bias. With only 3–5 *C. longa* samples per origin and five preparation replicates from one SWP batch, these models remain exploratory; no independent external validation was performed.

3.3. Differentially Abundant Annotated Features in Origin-Versus-SWP Comparisons

DAMs between each *C. longa* origin and SWPs were screened using variable importance in projection greater than 1.2, fold change at least 2.0 or no more than 0.5, and FDR-adjusted *q*-value less than 0.05 (Tables S4–S7). Their overlap and direction of change are summarized by the Venn diagram and volcano plots in Figure 4A–E.

Relative to the single analyzed SWP batch, 562 DAMs were identified for XSBN (256 higher and 306 lower; Figure 4E), 545 for DH (337 higher and 208 lower; Figure 4B), 497 for GM (267 higher and 230 lower; Figure 4C), and 477 for ML (212 higher and 265 lower; Figure 4D). A total of 167 DAMs were shared across the four comparisons. These features were putatively assigned to flavonoids, phenylpropanoids, terpenoids, fatty acyls, and other chemical classes (Table S8).

Plant-reference over-representation analysis tested 89 pathways in each origin-versus-SWP comparison; none reached BH FDR < 0.05 (Figure 4F–I; Table S10). The numbers with raw $p < 0.05$ were 3, 0, 2 and 4 for XSBN, DH, GM and ML, respectively. Examples included tyrosine metabolism in XSBN (13 hits among 24 background pathway compounds; $p = 0.008777$, FDR = 0.7812), stilbenoid, diarylheptanoid and gingerol biosynthesis in GM (5/8; $p = 0.02573$, FDR = 1.000), and butanoate metabolism in ML (5/6; $p = 0.003916$, FDR = 0.3485). These are descriptive annotation patterns rather than statistically supported pathway enrichment. The 10 descriptive putative candidates were 3-dehydroshikimate, 4-hydroxyphenylpyruvic acid, isopiperitenol, thymol, isorhamnetin-3-O-glucoside, flavanone, coumarin, isoimperatorin, 1′-acetoxychavicol, and β-bisabolene (Table S9). Isopiperitenol and thymol had higher relative abundances in each of the four *C. longa* groups than in SWP, whereas the other eight candidates had lower relative abundances in all four comparisons.

3.4. Relative Metabolomic Profile Similarity Between C. longa Origins and SWPs

Across the 1853 putatively annotated features, GM had the highest Spearman correlation with the SWPs in positive mode (rho = 0.6650) and negative mode (rho = 0.6954); DH values were 0.6347 and 0.6317, respectively (Table 1). The secondary concatenated representation gave GM rho = 0.6829 and DH rho = 0.3845. The lower combined DH value does not indicate poorer within-mode rank agreement: separate mode normalization changes cross-mode relative ranks after concatenation. Combined estimates and their bootstrap rankings are retained in Table 1 as representation-dependent summaries, rather than evidence for a biological source-matching order.

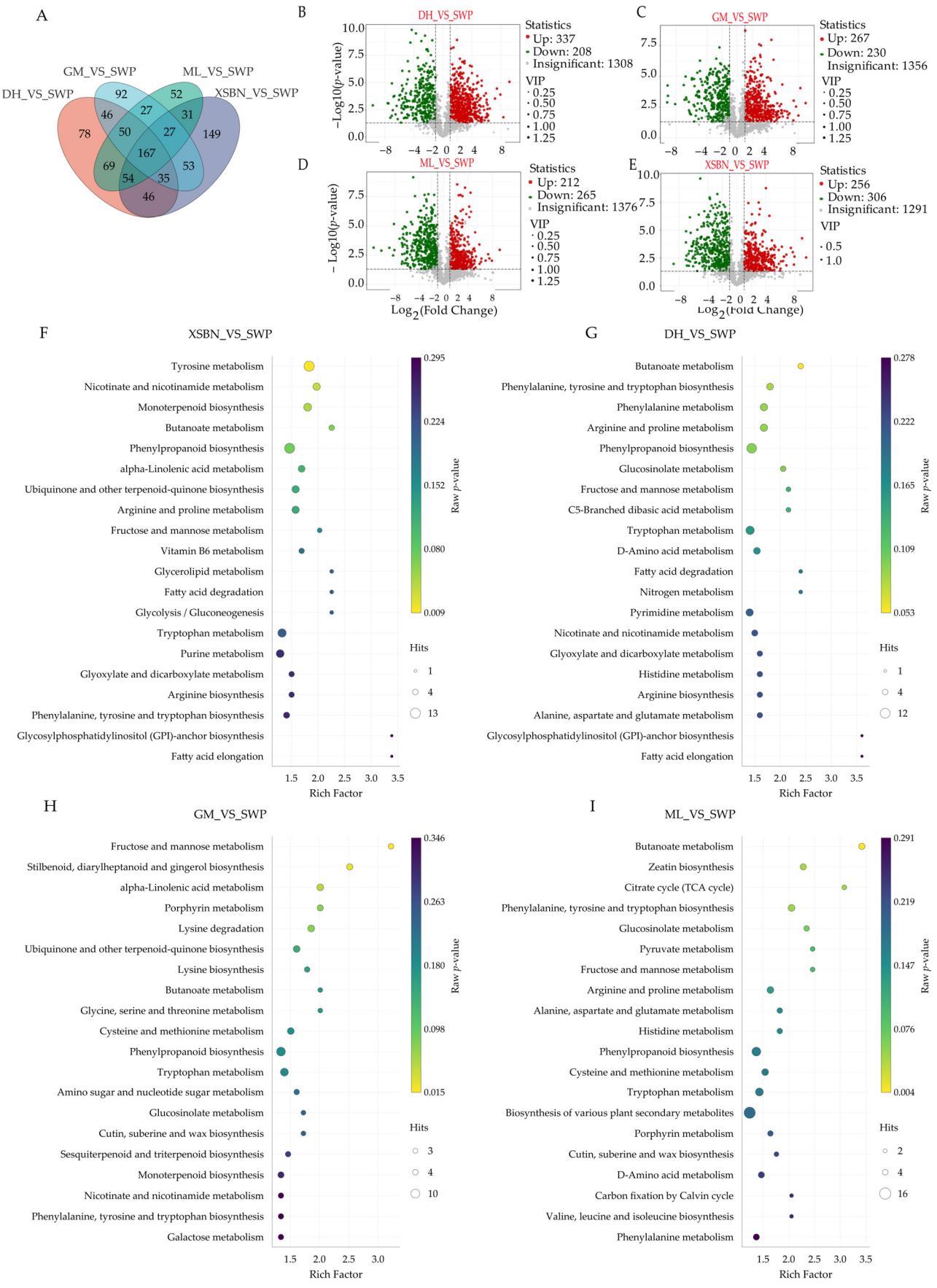

Figure 4. Differentially abundant putatively annotated features (DAMs) and plant-reference pathway summaries among the four pairwise comparisons. (**A**) Venn diagram showing the overlap of

differentially accumulated metabolites (DAMs) identified in the four comparisons (DH vs. SWP, GM vs. SWP, ML vs. SWP, and XSBN vs. SWP). Different colors represent the four comparison groups, as indicated by the corresponding labels. (**B**–**E**) Volcano plots showing the differential metabolite profiles for DH vs. SWP, GM vs. SWP, ML vs. SWP, and XSBN vs. SWP, respectively. Red and green dots represent upregulated and downregulated metabolites, respectively, while gray dots indicate metabolites that did not meet the differential screening criteria. The numbers of upregulated, downregulated, and non-significant metabolites are provided in each panel. (**F**–**I**) Plant-reference KEGG over-representation summaries of DAMs for XSBN vs. SWP, DH vs. SWP, GM vs. SWP, and ML vs. SWP, respectively. The *x*-axis represents the Rich Factor, bubble size indicates the number of matched metabolites (Hits), and bubble color represents the raw *p* value. No displayed pathway reaches BH FDR < 0.05; raw-P coloring is descriptive (Table S10).

Table 1. Relative-profile similarity between *C. longa* origin groups and SWPs based on mode-specific and combined data.

Mode	Origin	*n*	Spearman Rho (95% CI)	Highest Rho (%)	Bray–Curtis (95% CI)	Lowest BC (%)
Positive	XSBN	5	0.6092 (0.5361–0.6504)	0.00	0.7530 (0.0106–0.9494)	33.10
	DH	5	0.6347 (0.5856–0.6803)	6.65	0.7349 (0.0152–0.9483)	44.80
	GM	3	0.6650 (0.6174–0.7065)	80.90	0.9031 (0.0159–0.9578)	19.55
	ML	3	0.6426 (0.5871–0.6873)	12.45	0.8953 (0.7557–0.9612)	2.55
Negative	XSBN	5	0.6448 (0.5925–0.6898)	1.25	0.5306 (0.0169–0.9239)	25.85
	DH	5	0.6317 (0.5747–0.6739)	0.05	0.0301 (0.0188–0.8769)	69.30
	GM	3	0.6954 (0.6434–0.7302)	98.65	0.5300 (0.3611–0.9367)	0.55
	ML	3	0.6239 (0.5613–0.6713)	0.05	0.4274 (0.2829–0.9123)	4.30
Combined	XSBN	5	0.6209 (0.3684–0.6551)	18.15	0.6398 (0.0229–0.9099)	27.95
	DH	5	0.3845 (0.2933–0.6001)	1.05	0.3659 (0.0801–0.9165)	63.20
	GM	3	0.6829 (0.5016–0.7024)	72.55	0.7169 (0.2033–0.9329)	6.50
	ML	3	0.5872 (0.5368–0.6297)	8.25	0.6653 (0.4091–0.9218)	2.35

Notes: Positive- and negative-ion peak areas were normalized separately by sample total signal and analyzed both separately (948 and 905 features) and after concatenation (1853 features). Values compare feature-wise group-median profiles with the SWP group-median profile; group medians were not re-normalized. The combined Bray–Curtis value is the weighted average of mode-specific distances, with weights proportional to the sums of the two compared group profiles within each mode. Higher Spearman rho and lower Bray–Curtis values indicate greater similarity. Confidence intervals and first-rank frequencies were estimated from 2000 two-stage stratified bootstrap iterations. Sample resampling used site-level biological replicates for *C. longa* (XSBN, $n = 5$; DH, $n = 5$; GM, $n = 3$; ML, $n = 3$) and five preparation replicates from one SWP batch. SWP resampling was confined to those five preparation replicates from the single analyzed batch. The intervals include sample- and annotated-feature-resampling variability conditional on that batch and do not estimate SWP between-batch uncertainty. Mode-specific results are primary; combined-data rankings are secondary and sensitive to cross-mode ranks. All values in this table use untransformed normalized signals; the log-transformed sensitivity results are reported separately in Table S12.

Untransformed Bray–Curtis dissimilarity produced a different ordering. In the secondary combined representation, DH had the lowest point estimate (0.366; 95% bootstrap confidence interval, 0.080–0.916) and ranked first in 63.2% of iterations, followed by XSBN (0.640; 0.023–0.910), ML (0.665; 0.409–0.922), and GM (0.717; 0.203–0.933). DH also had the lowest Bray–Curtis point estimates in positive (0.7349) and negative (0.0301) modes (Table 1; Figure S6). The putatively annotated curcumin feature M367T447_3 accounted for 98.26% and 96.47% of the sums of the negative-mode DH and SWP group-median profiles, respectively, showing that the very low negative-mode distance was dominated by this intense feature. After $\log2(1 + 10^6\, p)$ transformation, the DH negative-mode distance was 0.3135; values for XSBN, GM and ML were 0.3779, 0.3927 and 0.3817, respectively. In positive mode, the minimum transformed distance was GM (0.4615), rather than DH (0.4727); XSBN and ML were 0.4751 and 0.5346, respectively (Table S12). These transformed distances use

a different weighting scale and are not directly interchangeable with the untransformed estimates. Spearman correlation and Bray–Curtis dissimilarity capture different properties of the relative profiles. The broad intervals and metric-dependent rankings characterize uncertainty in relative profile resemblance; they do not establish which origin supplied the *C. longa* used in SWPs.

Reanalysis showed that the DH correlation structure depended on data preprocessing (Figures S5 and S7). In the QC-corrected peak-area matrix, pairwise Spearman correlations among DH samples were high (rho = 0.933–0.949). Within positive and negative modes, correlations were 0.9295–0.9490 and 0.9205–0.9436, respectively, and were unchanged by total-signal normalization. After total-signal normalization separately within the positive- and negative-ion modes and subsequent concatenation, DH_1, DH_4, and DH_5 formed one highly correlated subset (rho = 0.964–0.974), and DH_2 and DH_3 formed another (rho = 0.955), with correlations between the subsets ranging from −0.148 to 0.022. Separate normalization preserves within-mode ranks but can change cross-mode relative ranks upon concatenation, accounting for the low correlations between subsets. Thus, the two-subcluster pattern was sensitive to data representation and cannot, by itself, establish species differences or sample mix-ups.

3.5. Focused Comparison of Putatively Annotated Curcuminoids

Curcumin, demethoxycurcumin and bisdemethoxycurcumin are characteristic curcuminoids reported from *C. longa* rhizomes [37]. The present data included features putatively annotated as curcumin (M367T447_3), demethoxycurcumin (M337T387_3) and hexahydrocurcumin (M357T364). M309T367_2, recorded in Table S1 under a systematic enol-form name, was included as putative bisdemethoxycurcumin because its formula C19H16O4, CAS 24939-16-0 and KEGG C17743 were consistent with this assignment [38]. Hexahydrocurcumin was considered an extended related feature. These annotations were supported by the available MS/MS mirror plots but were not confirmed using authentic standards under the study conditions.

Compound-specific differences were observed (Table 2). Curcumin origin/SWP mean peak-area ratios were 0.352, 2.547, 0.0357 and 0.0492 for XSBN, DH, GM and ML, respectively. After BH correction, curcumin signals were lower in GM and ML, whereas demethoxycurcumin signals were higher in DH, GM and ML. Only GM showed significant differences for putative bisdemethoxycurcumin and hexahydrocurcumin. The curcumin DH/GM mean ratio was 71.30, but DH signals varied 579.51-fold between the lowest and highest samples; the DH-SWP comparison was not significant after correction ($q = 0.2524$). These findings describe heterogeneous relative signals rather than absolute concentrations or source matching. Individual QC-corrected peak areas for all four curcuminoids in DH1-DH5 are provided in Table S11. The high-curcumin samples DH1, DH4 and DH5 are the same samples as subset A in the normalized concatenated correlation display; DH2 and DH3 form subset B.

Table 2. Exploratory comparisons of putatively annotated curcuminoids between *C. longa* origins and SWP.

Putatively Annotated Compound (Feature ID)	Origin	FC	*p*	BH *q*
Curcumin	XSBN	0.352	0.0496	0.0893
	DH	2.547	0.206	0.252
	GM	0.036	1.7×10^{-5}	0.000366
	ML	0.049	3.5×10^{-5}	0.000558

Table 2. *Cont.*

Putatively Annotated Compound (Feature ID)	Origin	FC	*p*	BH *q*
Demethoxycurcumin	XSBN	0.927	0.66	0.716
	DH	3.980	0.0005	0.00238
	GM	2.652	0.00227	0.00966
	ML	3.086	0.00377	0.0139
Bisdemethoxycurcumin *	XSBN	2.479	0.0925	0.144
	DH	6.003	0.0585	0.0872
	GM	8.229	0.00507	0.0169
	ML	1.160	0.0319	0.0644
Hexahydrocurcumin	XSBN	0.819	0.342	0.421
	DH	0.770	0.177	0.223
	GM	4.317	3.7×10^{-7}	6.85×10^{-5}
	ML	0.750	0.136	0.196

Notes: FC is the arithmetic mean peak area in each origin divided by that in SWPs. Two-sided Welch tests used QC-corrected peak areas without logarithmic transformation or total-signal normalization; BH *q* values were adjusted across all 1853 features within each comparison. Sample sizes were XSBN = 5, DH = 5, GM = 3, ML = 3 and SWP = 5 preparation replicates from one batch. Curcumin and demethoxycurcumin were negative-mode features; bisdemethoxycurcumin and hexahydrocurcumin were positive-mode features. * M309T367_2 was assigned putatively to bisdemethoxycurcumin from its systematic enol-form name, formula, CAS 24939-16-0 and KEGG C17743. Values describe relative signals, not absolute concentrations or exclusive botanical sources.

4. Discussion

The UPLC-MS/MS-based metabolite profiling dataset distinguished *C. longa* samples by geographical origin and separated all raw material groups from SWPs. Positive-ion PCA and hierarchical clustering placed XSBN relatively closer to SWPs than the other origins, but the primary mode-specific and combined relative-profile analyses showed that the ranking depended on the similarity measure and intensity transformation. GM had the highest Spearman point estimate in each ion mode. Although DH had the lowest untransformed Bray–Curtis distance, its negative-mode value was dominated by the putative curcumin signal, and positive-mode ranking changed after logarithmic transformation. Thus, no representation-independent abundance-similarity ordering was established. These results describe relative profile resemblance between the sampled single-herb materials and the complete formulation, without identifying the source of its *C. longa* component. The central finding is origin-associated variation within *C. longa* together with metric- and ion-mode-dependent chemical relationships to the formulation.

Variation in plant-specialized metabolites can arise from genetic background, local environment, cultivation practice, harvest stage, and processing [22]. Turmeric contains curcuminoids, phenolic acids, flavonoids, and volatile terpenoids, and previous metabolomic studies have used these compound classes to distinguish *Curcuma* species and geographical sources [3,6,13–15]. The observed differences in phenylpropanoids, flavonoids, terpenoids, lipids, organic acids, and related classes are consistent with this established chemical plasticity. Because environmental and genetic variables were not independently measured here, the data support an origin-associated pattern but do not isolate its causal drivers.

The additional origin-only analysis provides a formal test of association beyond visual separation in PCA or clustering. Nevertheless, the significant PERMDISP result indicates that location and dispersion effects cannot be cleanly separated in this small, unbalanced dataset. Pairwise significance should therefore be interpreted as evidence of differences in multivariate profile distributions rather than proof of isolated geographical effects. In particular, GM-ML had $R^2 = 0.6639$ but exact $p = 0.1000$, the smallest attainable value among 20 label allocations at $n = 3$ per group. This test cannot reach $p < 0.05$ and cannot

demonstrate equivalence. Sampling origin can also be associated with cultivation, genotype, harvest, and handling; these potential drivers were not independently manipulated. The origin-only results do not identify the *C. longa* source used in SWPs, because the other four ingredients were not profiled separately and only one SWP production batch was examined.

SWPs contain four botanical ingredients in addition to *C. longa* and undergo pharmaceutical processing. Its global chemical profile therefore reflects the combined contributions of these ingredients and potential processing-related extraction, degradation, or transformation. Because the other four ingredients were not profiled separately, their contributions cannot be distinguished from those of *C. longa* in the present dataset. The numerous DAMs in each origin-versus-SWP comparison and the separate negative-ion cluster must therefore be interpreted in the context of this compositional difference. The positive-ion PCA proximity of XSBN, the highest mode-specific Spearman point estimates for GM, and the transformation-sensitive Bray–Curtis results show that apparent similarity changes with the analytical representation and metric. These relationships describe the sampled *C. longa* materials relative to the analyzed SWP batch. They cannot establish which sampled origin resembles the *C. longa* component within the pills, identify the raw material source used in manufacture, or demonstrate superior pharmacological quality or therapeutic efficacy. Curcumin and bisdemethoxycurcumin have also been isolated from *Z. cassumunar* [39], which Kew recognizes as a synonym of *Z. purpureum* subsp. *purpureum* [40]. Thus, restricting comparisons to these curcuminoids still cannot establish an exclusive *C. longa* contribution to SWPs.

The plant-reference analysis found no pathway enrichment after BH correction; nominal raw-P associations therefore provide only descriptive chemical context. The Arabidopsis reference library is a proxy and may omit *C. longa*-specific specialized chemistry; pathway membership alone does not establish the occurrence or activity of a complete pathway in turmeric. However, the mapped compound identities were not confirmed using authentic reference standards under the study conditions. Annotation uncertainty therefore limits the specificity of the pathway interpretation. Moreover, the analysis compared a single plant material with a multi-herb product, so pathway associations cannot be attributed specifically to the *C. longa* component of SWPs. These results should be read as exploratory organization of the chemical data, not as evidence of pathway activation, altered metabolic flux, or causal mechanisms. Accordingly, phenylpropanoid, monoterpenoid or aromatic-amino-acid pathway labels are not interpreted as the principal biochemical differences supported by this dataset.

The 167 shared DAMs constitute an exploratory candidate pool rather than validated quality control markers. Relative peak-area differences cannot establish absolute content limits or fingerprint acceptance criteria. No authentic reference standards were analyzed under the study conditions for identity confirmation. The 10 prioritized candidates remain putative annotations. Their spectral-match quality and structural assignments require individual review. Only one SWP production batch and a limited number of *C. longa* samples were studied; between-batch consistency, extraction recovery, analytical stability, and inter-instrument or inter-laboratory reproducibility were not evaluated. This shortlist is intended to guide identity confirmation and targeted studies. Development of usable markers requires authentic-standard confirmation, fit-for-purpose quantitative or semiquantitative method validation, and assessment in independent samples spanning multiple origins and production batches.

This study has several limitations. SWPs were represented by a single production batch, and the five preparation replicates captured only within-batch preparation and analytical variation. The other four botanical ingredients were not profiled separately, preventing separation of ingredient-composition and processing effects or identification of the geo-

graphical origin of the *C. longa* used in the formulation. The *C. longa* groups were small and unequal, and the OPLS-DA models were not independently validated using external samples; high R^2Y and Q^2 values therefore do not establish generalizability. The correlation structure among DH samples was sensitive to preprocessing because concatenation changed cross-mode relative ranks; any biological explanation requires further investigation using sample records and species authentication evidence. This structure alone cannot establish species differences or sample mix-ups. Furthermore, relative abundance does not represent absolute concentration, and compound identities were not confirmed using authentic standards under the same experimental conditions. Uncertainty in the putative annotations limits compound-specific and pathway interpretations. Accordingly, the conclusions are restricted to origin-associated chemical variation among the sampled *C. longa* materials and exploratory chemical-profile comparisons with a single SWP batch. These findings may inform subsequent targeted studies, but their applicability to raw material source identification, quality control, or pharmacological evaluation requires independent samples and further experimental validation.

5. Conclusions

UPLC-MS/MS-based metabolite profiling revealed clear origin-associated chemical variation among *C. longa* samples from four regions of Yunnan and pronounced differences between the raw material and SWPs. Positive-ion PCA and clustering placed XSBN relatively closer to SWPs, whereas the mode-specific and combined relative-profile analyses found the highest Spearman correlations for GM (positive rho = 0.6650; negative rho = 0.6954). The low untransformed Bray–Curtis value for DH was strongly influenced by intense signals, and the positive-mode minimum changed after logarithmic transformation. Combined metrics are secondary and sensitive to cross-mode ranks. These metric-dependent rankings, together with the representation-dependent correlation structure within DH, do not support designation of a uniquely closest or preferred origin. Direct origin-only PERMANOVA supported multivariate origin association, but significant dispersion differences prevent a purely centroid-based interpretation. Five pairwise comparisons survived BH correction; GM versus ML did not. The representation-dependent DH pattern does not establish distinct biological subgroups. The 167 shared DAMs summarize chemical contrasts common to the four origin-versus-formulation comparisons. No pathway in the plant-reference analysis of the individual comparison DAM sets remained enriched after BH correction; pathway mappings rely on putative annotations and do not establish pathway activity, metabolic flux, or causal mechanisms. The origin-versus-formulation results are limited to a descriptive comparison of the sampled *C. longa* materials with one SWP production batch. This comparison cannot identify the *C. longa* source used in SWPs, because the other four ingredients were not profiled separately. Their contributions and those of processing may influence or mask *C. longa*-associated signals. The absence of additional production batches also prevents generalization across batches. These data do not establish manufacturing provenance, a matching *C. longa* source, source superiority, chemical equivalence, or therapeutic equivalence. The metric- and ion-mode-dependent rankings require cautious interpretation, and the 167 shared DAMs remain exploratory candidates for further targeted evaluation rather than validated quality control markers. The focused curcuminoid results likewise remain exploratory and do not establish *C. longa*-specific markers in the formulation.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/separations13110348/s1, Figure S1. Representative *Curcuma longa* materials from DH, XSBN, GM and ML and the Shuangjiang Weitong Pills (SWP) sample. XSBN, Xishuangbanna; DH, Dehong; GM, Gengma; ML, Menglian. Figure S2. Hierarchical clustering of the study samples. Sample labels identify the four *Curcuma longa* origins and SWP; the colour strip

indicates group membership. Clustering is descriptive and does not establish the botanical source of SWP constituents. Figure S3. PCA restricted to the 16 site-level composite *C. longa* samples and 1853 putatively annotated features after logarithmic transformation and unit-variance scaling. PC1 and PC2 explain 28.23% and 20.86% of variance. Origin-only PERMANOVA: pseudo-F = 5.8751, $R^2 = 0.5949$, $p = 0.0001$ (9999 permutations). Figure S4. Origin-only multivariate comparisons. (A) Pairwise PERMANOVA R^2 values and BH-adjusted exact p values. The five values of 0.0214 and GM–ML value of 0.1000 reflect the finite resolution of size-preserving label allocations, rather than effect size. (B) Bias-adjusted distances to group centroids for PERMDISP. Overall dispersion differed among origins ($F = 11.9142$, $p = 0.0001$); location and dispersion effects therefore cannot be separated by these tests. Figure S5. Pairwise Spearman correlations among the five DH samples using QC-corrected peak areas (left) and separately mode-normalized concatenated profiles (right). The latter representation produces two correlated subsets. This change is sensitive to preprocessing and does not establish biological subgroups or sample misidentification. Figure S6. Similarity of each *C. longa* origin to the SWP group-median profile in positive mode (948 features), negative mode (905 features) and the concatenated data (1853 features). Results show Spearman correlation and Bray–Curtis dissimilarity with 95% percentile intervals from 2000 two-stage stratified bootstrap iterations. Samples were normalized within mode before group medians were calculated. Sample draws were paired across modes, and the SWP resample was shared among origin comparisons. Higher Spearman and lower Bray–Curtis indicate greater resemblance under the respective metric. Intervals and first-rank frequencies are conditional on the sampled materials and one SWP batch. Figure S7. DH correlation diagnostics across data representations. Within-mode total-signal normalization leaves Spearman correlations unchanged: 0.9295–0.9490 in positive mode and 0.9205–0.9436 in negative mode. Direct concatenation of unnormalized peak areas gives 0.9328–0.9486; concatenation after separate mode normalization gives between-subset correlations of −0.1481–0.0222. Normalization changes relative ranks across modes, despite preserving ranks within each mode. Table S1. Putatively annotated features across *C. longa* samples and SWP. Table S2. Overall origin-only PERMANOVA and multivariate dispersion tests. Table S3. Exact pairwise PERMANOVA among the four *C. longa* origins. Table S4. Differentially abundant annotated features between XSBN and SWP. Table S5. Differentially abundant annotated features between DH and SWP. Table S6. Differentially abundant annotated features between GM and SWP. Table S7. Differentially abundant annotated features between ML and SWP. Table S8. Shared differentially abundant annotated features across the four origin-versus-SWP comparisons. Table S9. Exploratory candidate compounds shared across four origin-versus-SWP comparisons. Table S10. Plant-reference pathway over-representation analysis of differentially abundant putatively annotated features in each *C. longa* origin versus SWP. Table S11. Individual QC-corrected peak areas of four putatively annotated curcuminoid-related features in DH samples. Table S12. Bray–Curtis sensitivity to logarithmic transformation of sample-normalized signals. Table S13. Sample collection and provenance information for *Curcuma longa* materials from four regions of Yunnan, China.

Author Contributions: Conceptualization, Q.D. and Z.M.; methodology, W.F., M.A., R.Z., S.W., Q.D. and Z.M.; software, W.F.; validation, R.Z. and S.W.; formal analysis, W.F. and Z.M.; investigation, W.F. and Z.M.; resources, Q.D.; data curation, R.Z.; writing—original draft preparation, W.F. and M.A.; writing—review and editing, Q.D. and M.A.; visualization, Z.M. and M.A.; supervision, Q.D.; project administration, Z.M.; funding acquisition, Z.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Special Basic Cooperative Research Programs of Yunnan Provincial Undergraduate Universities' Association, grant number 202301BA070001-066. The APC was funded by the Special Basic Cooperative Research Programs of Yunnan Provincial Undergraduate Universities' Association (grant number 202301BA070001-066), Yunnan Provincial Department of Education Scientific Research Fund Project (2023J1277) and School of Pu'er Tea, West Yunnan University of Applied Sciences Industry-Education Integration Project (2025YJXM05).

Institutional Review Board Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article/Supplementary Material. Further inquiries can be directed to the corresponding authors.

Acknowledgments: The authors thank Panomix Biomedical Technology Co., Ltd. for its technical assistance with sample preparation, UPLC-MS/MS analysis, and metabolite annotation.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Iweala, E.J.; Uche, M.E.; Dike, E.D.; Etumnu, L.R.; Dokunmu, T.M.; Oluwapelumi, A.E.; Okoro, B.C.; Dania, O.E.; Adebayo, A.H.; Ugbogu, E.A. *Curcuma longa* (Turmeric): Ethnomedicinal uses, phytochemistry, pharmacological activities and toxicity profiles—A review. *Pharmacol. Res. Mod. Chin. Med.* **2023**, *6*, 100222. [CrossRef]
2. Tian, W.-W.; Liu, L.; Chen, P.; Yu, D.-M.; Li, Q.-M.; Hua, H.; Zhao, J.-N. *Curcuma longa* (turmeric): From traditional applications to modern plant medicine research hotspots. *Chin. Med.* **2025**, *20*, 76. [CrossRef] [PubMed]
3. Li, S.; Yuan, W.; Deng, G.; Wang, P.; Yang, P.; Aggarwal, B.B. Chemical Composition and Product Quality Control of Turmeric (*Curcuma longa* L.). *Pharm. Crops* **2011**, *2*, 28–54. [CrossRef]
4. Fuloria, S.; Mehta, J.; Chandel, A.; Sekar, M.; Rani, N.N.I.M.; Begum, M.Y.; Subramaniyan, V.; Chidambaram, K.; Thangavelu, L.; Nordin, R.; et al. A Comprehensive Review on the Therapeutic Potential of *Curcuma longa* Linn. in Relation to its Major Active Constituent Curcumin. *Front. Pharmacol.* **2022**, *13*, 820806. [CrossRef] [PubMed]
5. Madhu, K.; Chanda, K.; Saji, M.J. Safety and efficacy of *Curcuma longa* extract in the treatment of painful knee osteoarthritis: A randomized placebo-controlled trial. *Inflammopharmacology* **2013**, *21*, 129–136. [CrossRef] [PubMed]
6. Salem, M.A.; El-Shiekh, R.A.; Fernie, A.R.; Alseekh, S.; Zayed, A. Metabolomics-based profiling for quality assessment and revealing the impact of drying of Turmeric (*Curcuma longa* L.). *Sci. Rep.* **2022**, *12*, 10288. [CrossRef] [PubMed]
7. Sharifi-Rad, J.; El Rayess, Y.E.; Rizk, A.A.; Sadaka, C.; Zgheib, R.; Zam, W.; Sestito, S.; Rapposelli, S.; Neffe-Skocińska, K.; Zielińska, D.; et al. Turmeric and Its Major Compound Curcumin on Health: Bioactive Effects and Safety Profiles for Food, Pharmaceutical, Biotechnological and Medicinal Applications. *Front. Pharmacol.* **2020**, *11*, 1021. [CrossRef] [PubMed]
8. Li, Z.; Li, C.; Zhang, X.; Tang, S.; Yang, H.; Cui, X.; Huang, L. Policies and Problems of Modernizing Ethnomedicine in China: A Focus on the Yi and Dai Traditional Medicines of Yunnan Province. *Evid.-Based Complement. Altern. Med.* **2020**, *2020*, 1023297. [CrossRef] [PubMed]
9. Zhang, C.; Chongsuvivatwong, V.; Keawpradub, N.; Lin, Y. Analysis of prescription database extracted from standard textbooks of traditional Dai medicine. *J. Ethnobiol. Ethnomed.* **2012**, *8*, 34. [CrossRef] [PubMed]
10. Dai, X.-X.; Tan, H.; Ma, H.-J.; Yang, H.-Y.; Jin, Q.; Tan, W.-H. Research progress on six species of Zingiberaceae Dai medicinal plants. *J. Yunnan Minzu Univ. (Nat. Sci. Ed.)* **2026**, *35*, 17–22+51. (In Chinese) Available online: https://ynmz.cbpt.cnki.net/portal/journal/portal/client/paper/da1ef42996ee738131b071f4b40c6b3f (accessed on 15 November 2026).
11. Kanase, V.; Khan, F. An overview of medicinal value of *Curcuma* species. *Asian J. Pharm. Clin. Res.* **2018**, *11*, 40–45. [CrossRef]
12. Chatzinasiou, L.; Booker, A.; MacLennan, E.; Mackonochie, M.; Heinrich, M. Turmeric (*Curcuma longa* L.) products: What quality differences exist? *J. Herb. Med.* **2019**, *17–18*, 100281. [CrossRef]
13. Lee, J.; Jung, Y.; Shin, J.-H.; Kim, H.; Moon, B.; Ryu, D.; Hwang, G.-S. Secondary Metabolite Profiling of *Curcuma* Species Grown at Different Locations Using GC/TOF and UPLC/Q-TOF MS. *Molecules* **2014**, *19*, 9535–9551. [CrossRef] [PubMed]
14. Nguyen Thi, K.-O.; Do, H.-G.; Duong, N.-T.; Nguyen, T.D.; Nguyen, Q.-T. Geographical Discrimination of *Curcuma longa* L. in Vietnam Based on LC-HRMS Metabolomics. *Nat. Prod. Commun.* **2021**, *16*, 1934578X211045479. [CrossRef]
15. Nurani, L.H.; Rohman, A.; Windarsih, A.; Guntarti, A.; Riswanto, F.D.O.; Lukitaningsih, E.; Fadzillah, N.A.; Rafi, M. Metabolite Fingerprinting Using 1H-NMR Spectroscopy and Chemometrics for Classification of Three *Curcuma* Species from Different Origins. *Molecules* **2021**, *26*, 7626. [CrossRef] [PubMed]
16. Ao, M.; Li, X.; Liao, Y.; Zhang, C.; Fan, S.; Hu, C.; Chen, Z.; Yu, L. *Curcumae Radix*: A review of its botany, traditional uses, phytochemistry, pharmacology and toxicology. *J. Pharm. Pharmacol.* **2022**, *74*, 779–792. [CrossRef] [PubMed]
17. Liu, F.; Bai, X.; Yang, F.-Q.; Zhang, X.-J.; Hu, Y.; Li, P.; Wan, J.-B. Discriminating from species of *Curcumae Radix* (Yujin) by a UHPLC/Q-TOFMS-based metabolomics approach. *Chin. Med.* **2016**, *11*, 21. [CrossRef] [PubMed]
18. Zhou, Y.; Xie, M.; Song, Y.; Wang, W.; Zhao, H.; Tian, Y.; Wang, Y.; Bai, S.; Zhao, Y.; Chen, X.; et al. Two Traditional Chinese Medicines *Curcumae Radix* and *Curcumae Rhizoma*: An Ethnopharmacology, Phytochemistry, and Pharmacology Review. *Evid.-Based Complement. Altern. Med.* **2016**, *2016*, 4973128. [CrossRef] [PubMed]
19. Wishart, D.S. Emerging applications of metabolomics in drug discovery and precision medicine. *Nat. Rev. Drug Discov.* **2016**, *15*, 473–484. [CrossRef] [PubMed]
20. Wolfender, J.-L.; Marti, G.; Thomas, A.; Bertrand, S. Current approaches and challenges for the metabolite profiling of complex natural extracts. *J. Chromatogr. A* **2015**, *1382*, 136–164. [CrossRef] [PubMed]
21. De Vos, R.C.; Moco, S.; Lommen, A.; Keurentjes, J.J.; Bino, R.J.; Hall, R.D. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nat. Protoc.* **2007**, *2*, 778–791. [CrossRef] [PubMed]

22. Fiehn, O. Metabolomics—The link between genotypes and phenotypes. In *Functional Genomics*; Town, C., Ed.; Springer: Dordrecht, The Netherlands, 2002; pp. 155–171. [CrossRef] [PubMed]
23. Rochfort, S. Metabolomics Reviewed: A New "Omics" Platform Technology for Systems Biology and Implications for Natural Products Research. *J. Nat. Prod.* **2005**, *68*, 1813–1820. [CrossRef] [PubMed]
24. Chambers, M.C.; Maclean, B.; Burke, R.; Amodei, D.; Ruderman, D.L.; Neumann, S.; Gatto, L.; Fischer, B.; Pratt, B.; Egertson, J.; et al. A cross-platform toolkit for mass spectrometry and proteomics. *Nat. Biotechnol.* **2012**, *30*, 918–920. [CrossRef] [PubMed]
25. Smith, C.A.; Want, E.J.; O'Maille, G.; Abagyan, R.; Siuzdak, G. XCMS: Processing Mass Spectrometry Data for Metabolite Profiling Using Nonlinear Peak Alignment, Matching, and Identification. *Anal. Chem.* **2006**, *78*, 779–787. [CrossRef] [PubMed]
26. Wishart, D.S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A.C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; et al. HMDB: The Human Metabolome Database. *Nucleic Acids Res.* **2007**, *35*, D521–D526. [CrossRef] [PubMed]
27. Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E.A.; Glass, C.K.; Merrill, A.H.; Murphy, R.C.; Raetz, C.R.H.; Russell, D.W.; et al. LMSD: LIPID MAPS structure database. *Nucleic Acids Res.* **2007**, *35*, D527–D532. [CrossRef] [PubMed]
28. Ogata, H.; Goto, S.; Sato, K.; Fujibuchi, W.; Bono, H.; Kanehisa, M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* **1999**, *27*, 29–34. [CrossRef] [PubMed]
29. Hastings, J.; Owen, G.; Dekker, A.; Ennis, M.; Kale, N.; Muthukrishnan, V.; Turner, S.; Swainston, N.; Mendes, P.; Steinbeck, C. ChEBI in 2016: Improved services and an expanding collection of metabolites. *Nucleic Acids Res.* **2016**, *44*, D1214–D1219. [CrossRef] [PubMed]
30. Broadhurst, D.; Goodacre, R.; Reinke, S.N.; Kuligowski, J.; Wilson, I.D.; Lewis, M.R.; Dunn, W.B. Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics* **2018**, *14*, 72. [CrossRef] [PubMed]
31. van den Berg, R.A.; Hoefsloot, H.C.; Westerhuis, J.A.; Smilde, A.K.; van der Werf, M.J. Centering, scaling, and transformations: Improving the biological information content of metabolomics data. *BMC Genom.* **2006**, *7*, 142. [CrossRef] [PubMed]
32. Thévenot, E.A.; Roux, A.; Xu, Y.; Ezan, E.; Junot, C. Analysis of the Human Adult Urinary Metabolome Variations with Age, Body Mass Index, and Gender by Implementing a Comprehensive Workflow for Univariate and OPLS Statistical Analyses. *J. Proteome Res.* **2015**, *14*, 3322–3335. [CrossRef] [PubMed]
33. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B Stat. Methodol.* **1995**, *57*, 289–300. [CrossRef]
34. Rivals, I.; Personnaz, L.; Taing, L.; Potier, M.-C. Enrichment or depletion of a GO category within a class of genes: Which test? *Bioinformatics* **2007**, *23*, 401–407. [CrossRef] [PubMed]
35. Anderson, M.J. A new method for non-parametric multivariate analysis of variance. *Austral Ecol.* **2001**, *26*, 32–46. [CrossRef]
36. Anderson, M.J. Distance-Based Tests for Homogeneity of Multivariate Dispersions. *Biometrics* **2006**, *62*, 245–253. [CrossRef] [PubMed]
37. Jiang, H.; Timmermann, B.N.; Gang, D.R. Use of liquid chromatography-electrospray ionization tandem mass spectrometry to identify diarylheptanoids in turmeric (*Curcuma longa* L.) rhizome. *J. Chromatogr. A* **2006**, *1111*, 21–31. [CrossRef] [PubMed]
38. Research Center for Medicinal Plant Resources, NIBN. Comprehensive Medicinal Plant Database: Bisdemethoxycurcumin. Available online: https://mpdb.nibn.go.jp/mpdb-bin/view_compound_data.cgi?id=23&lang=en (accessed on 11 November 2026).
39. Li, M.-X.; Bai, X.; Ma, Y.-P.; Zhang, H.-X.; Nama, N.; Pei, S.-J.; Du, Z.-Z. Cosmetic potentials of extracts and compounds from *Zingiber cassumunar* Roxb. rhizome. *Ind. Crops Prod.* **2019**, *141*, 111764. [CrossRef]
40. Royal Botanic Gardens, Kew. Plants of the World Online: *Zingiber cassumunar* Roxb. Available online: https://powo.science.kew.org/taxon/798323-1 (accessed on 11 November 2026).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.