scientific reports

Check for updates

OPEN

Widely targeted metabolomics study of flower buds of *Styphnolobium japonicum* from different producing areas

Leilei Zuo[1,2,3], Rundong Meng[1,3], Xiaofeng Li[1], Xiao Meng[1], Ying Zhang[1], Dayi Chen[1]✉ & Waralee Watcharin[2]✉

The flower buds of *Styphnolobium japonicum (FBSJ)*, a medicinal and food plant, exhibit broad application prospects and diverse therapeutic effects, including cooling the blood, stopping bleeding, lowering blood sugar, and exerting antioxidant and anti-inflammatory activities. To provide a reference for the identification, resource development, and utilization, a widely targeted metabolomics study was conducted on *FBSJ* samples from four production regions in China: Shandong (SJsd), Anhui (SJah), Hebei (SJhb), and Henan (SJhn). In this research, the metabolomic analysis provided the possibility of identifying metabolites to differentiate plants from different origin to broaden a limited metabolome coverage and unclear origin-metabolite-efficacy correlations. A UPLC-MS/MS-based metabolomics analysis was selected to investigate the metabolites of *FBSJ* from different geographic origins, utilizing the MWDB database and the multi-reaction monitoring mode of triple-quadrupole mass spectrometry (TQMS). A total of 1,559 metabolites were identified, of which 294 were annotated as key active ingredients in traditional Chinese medicine (TCM), and 193 as active pharmaceutical components against 11 diseases, including cancer and diabetes. Among the 708 differential metabolites identified across the four different origin groups, 176 were annotated by KEGG and distributed across 87 metabolic pathways. The most significantly altered metabolic pathway was isoflavonoid biosynthesis ($p < 0.01$), where SJsd samples exhibited an upregulation trend in key differential metabolites compared to the other groups. Among the 193 disease-active metabolites, 13 (e.g., 3′-Methoxydaidzein, Maackiain) were at least twofold more abundant in SJsd compared to any other origin, highlighting their critical role in explaining SJsd's superior pharmacological potential. This study provides a comprehensive molecular-level analysis of *FBSJ* metabolites, offering useful insights into its pharmacological analysis across different origins. These findings enhance our understanding of *FBSJ* metabolites annotated as anti-disease-related components in public databases and offer a valuable reference for investigating their potential in human health applications and the development of novel functional foods.

Keywords Metabolomics, Styphnolobium japonicum, Differential metabolites, Metabolic pathways, anti-disease-related components

Styphnolobium japonicum (SJ; synonym: *Sophora japonica L.*) is a tree species from the butterfly flower family, originating in China[1]. It is widely cultivated across various regions, particularly in northern China and the Loess Plateau. The species is also distributed in Japan, Vietnam, Korea, Europe, and several Western countries, including the United States[2]. The flower bud of *SJ* (*FBSJ*) is a major traditional medicinal material in China, Japan, and Korea, commonly used for hemostasis and cooling the blood to treat heat-related blood disorders. It is also frequently incorporated into medicinal diets and general food products[3]. Owing to its rich composition of bioactive compounds, *FBSJ* serves as a prominent active ingredient in Traditional Chinese Medicine (TCM)[4]. According to data from Tian Di Yun Tu, a big data platform dedicated to the TCM industry, China's *FBSJ* production reached 1,800 tons in 2022[5]. Its notably high rutin content further establishes *FBSJ* as the primary raw material for rutin extraction[6]. To date, an increasing number of studies have investigated *FBSJ* in

[1]College of Public Health, Chengdu University of Traditional Chinese Medicine, Chengdu 611137, PR China. [2]Theophane Venard School of Biotechnology, Assumption University, Bangkok 10240, Thailand. [3]Leilei Zuo and Rundong Meng contributed equally to this work. ✉email: chendayi00523@sina.com; waraleewtc@au.edu

various aspects, including vegetative propagation[7]; the effects of different environmental impacts on growth and nitrogen metabolism[8]; and the characterization of bioactive components, such as proteins, gene expression profiles, and flavonoids including rutin and quercetin[9–11]. Additional research has examined its lectin activity[12], aqueous extract[10], polysaccharides' antioxidant effects,[11] uric acid modulation,[12] and its pharmacological effect when combined with other Traditional Chinese medicines[13]. Metabolites are highly responsive to environmental factors and vary accordingly. However, they are also produced by plants as adaptations to specific ecological conditions, with secondary metabolites usually considered the material basis of bioactivity[14]. Previous studies have mainly focused on partial components of *FBSJ*, while research on full metabolite profiles (including non-flavonoid metabolites) at the molecular level, particularly in *Sophora* species, remains limited[15,16]. Comprehensive studies on the types of secondary metabolites and associated metabolic pathways of *FBSJ* across different geographic origins remain limited, hindering the elucidation of origin-dependent quality variations. Furthermore, the therapeutic potential of *FBSJ* against 11 major diseases (e.g., cancer, diabetes) has yet to be systematically validated at the metabolite level[17].

Currently, major diseases threatening human health worldwide include cancer[18], diabetes[19], hypertension[20], cardiovascular diseases[21], atherosclerosis[22], and thrombotic diseases[23]. Modern clinical studies have shown that *FBSJ* exhibits various pharmacological effects, including hemostasis, hypoglycemia, antioxidant activity, stomach protection, immunity enhancement, antiviral effects, blood pressure reduction, and antitumor activity[24]. Based on the CancerHSP and TCMSP databases, six major diseases—namely cancer/tumor, diabetes, hypertension, cardiovascular diseases, atherosclerosis, and thrombotic disorders—were identified as primary targets associated with *FBSJ*. In addition, five other conditions potentially linked to *FBSJ* were recognized, including osteoporosis, liver ischemic injury, inflammation, infectious diseases, and hemorrhage. These disease categories reflect the spectrum of pharmacological activities that *FBSJ*-derived metabolites may exert. Due to variations in plant origin and extraction conditions, *FBSJ*-derived compounds may exhibit distinct biological activities and health-promoting effects across different countries. In China, Shandong, Anhui, Henan, and Hebei are the main production regions of *FBSJ.* Although the flowers from these four provinces exhibit no significant morphological differences, market feedback indicates that there are certain variations in quality. However, the specific differences remain unclear, and relevant studies are scarce.

In this context, a metabolomic analysis of *FBSJ* from these four regions is necessary to identify differences in metabolite composition, explore potential new pharmacological effects, predict variations in therapeutic properties, and provide a reference for the screening and deep processing of *FBSJ*. To achieve the metabolites of *FBSJ* more comprehensively, this study employs a widely targeted metabolomics approach using UPLC-MS/MS to analyse *FBSJ* samples collected from diverse geographic origins. This strategy enables precise identification of secondary metabolites at the molecular level. Additionally, key active ingredients and disease-resistant ingredients associated with the eleven diseases were identified based on network pharmacology databases—this step specifically addresses the lack of metabolite-disease linkage in existing research. Differential metabolite screening enabled the identification of potential marker metabolites in each group (e.g., 13 high-abundance metabolites in SJsd), providing theoretical support for the establishment of quality standards. Furthermore, an analysis of significant KEGG pathways (e.g., isoflavonoid biosynthesis) among *FBSJ* from different origins offers insights for future cultivation, material basis screening, and the pharmacological mechanisms of disease-resistant ingredients, directly advancing the field's understanding of *FBSJ*'s origin-related metabolism. These findings contribute to human health promotion and support the development of *FBSJ*-based functional foods and pharmaceuticals.

Materials and methods

Materials

Methanol and acetonitrile were purchased from Shanghai Merck Chemical Technology Co., Ltd., while formic acid was obtained from Shanghai Aladdin Biochemical Technology Co., Ltd. All chemicals were of chromatographic grade and used without further purification. *FBSJ* samples were collected from four geographic origins in China: SJsd from Shandong (115°46′E, 35°16′N, Heze), SJah from Anhui (117°18′E, 31°30′N, Lujiang), SJhn from Henan (114°28′E, 33°23′N, Shangcai), and SJhb from Hebei (115°04′E, 38°59′N, Dingzhou). These samples were sourced from Yutang Food Department Store, Xinzhou High-tech Industrial Development Zone, Anhui Baitang Pharmaceutical Co., Ltd., Chengde Biological Technology Co., Ltd., and China E-commerce Co., Ltd., respectively. All samples were harvested in July 2022 and were identified by Assistant Professor Jihai Gao from Chengdu University of TCM as the dried flower buds of *Styphnolobium japonicum L.* (*Sophora japonica L.*). Voucher specimens of these materials were deposited in the State Bank of Chinese Drug Germplasm Resources (Voucher number:522230190125HM001, 522230190125HM002, 522230190125HM003, 522230190125HM004). Climatic data (temperature, rainfall, and sunshine duration) for the four production regions from February to May 2022 were obtained from publicly available sources: temperature and rainfall data from https://www.tianqi24.com and sunshine duration from https://www.ceicdata.com/zh-hans/china/sunshine-hours.

Methods

Sample Preparation and extraction

Each sample was freeze-dried and ground into a fine powder. A 50 mg aliquot of *FBSJ* powder was accurately weighed and extracted with 1.2 mL of pre-cooled (− 20 ℃) 70% methanol, which served as the internal standard extraction solution. The mixture was vortexed for 30 s every 30 min, repeated six times. The sample was then centrifuged at 1,200 rpm for 3 min, and the supernatant was filtered through a microporous membrane for subsequent UPLC-MS/MS analysis.

LC-MS/MS analysis conditions
Metabolite identification was performed using a UPLC-ESI-MS/MS system consisting of a UPLC unit (ExionLC™ AD, https://sciex.com.cn/) coupled with an MS (Applied Biosystems 4500 Q TRAP, https://sciex.com.cn/).

Chromatography-mass spectrometry acquisition conditions
Liquid-phase conditions The data acquisition was conducted using a UPLC system coupled with MS/MS. The liquid phase conditions included an Agilent SB-C18 column (1.8 μm, 2.1 mm × 100 mm), with phase A consisting of ultra-pure water with 0.1% formic acid and phase B consisting of acetonitrile with 0.1% formic acid. The elution gradient started with 5% B at 0.0 min, increased linearly to 95% B within 9.0 min, maintained 95% B from 9.0 to 10.0 min, decreased to 5% B from 10.0 to 11.1 min, and maintained 5% B from 11.1 to 14.0 min. The flow rate was set to 0.35 mL/min, the column oven temperature was maintained at 40 °C, and the injection volume was 4 μL.

MS conditions The effluent was analysed using linear ion trap (LIT) and triple quadrupole (QQQ) scans. The analytical conditions were as follows: Electrospray ionization (ESI) temperature set to 550 °C, with an ion spray voltage of 5500 V (positive ion mode) or − 4500 V (negative ion mode). The ion source gas (GSI) was set to 50 psi, gas II (GSII) to 60 psi, and the curtain gas (CUR) to 25 psi. Collision-induced ionization parameters were set to "high". The QQQ scan used Multiple Reaction Monitoring (MRM) mode, with the collision gas (nitrogen) set to medium. The declustering potential (DP) and collision energy (CE) were optimized for each MRM ion pair. A specific set of MRM ion pairs was monitored during each period based on the metabolites eluted during that time.

Qualitative and quantitative determination of metabolites
For qualitative analysis of metabolites, the data analysis method was adapted from Wang et al.[25]. The substances were identified based on fragmentation patterns, retention time, and m/z values established in the Metware database (MWDB), provided by Metware Biotechnology Co., Ltd. (Wuhan, China). Using secondary mass spectra information, the obtained metabolite data were compared with the MWDB database to obtain structural information and classifications. Differential metabolites in the *FBSJ* samples were then studied. Isotopic signals, repetitive signals containing K^+, Na^+, NH_{4+}, and fragment ions derived from larger molecular species were excluded from the analysis.

Quantitative analysis was performed using the multi-reaction monitoring (MRM) mode of triple quadrupole mass spectrometry. After obtaining the mass spectrometry data for the different samples, the peak area normalization method was applied to calculate the relative content of *FBSJ* samples from different producing areas. Each sample was analysed in triplicate, and the average data were calculated.

Identification of key active ingredients in traditional Chinese medicine (TCM)
UPLC-MS/MS was used to analyse *FBSJ*, and all metabolites were queried in the TCM systems pharmacology database and analysis platform (TCMSP). Metabolites with oral bioavailability (OB) ≥ 5% and drug-likeness (DL) ≥ 0.14 were identified as key active ingredients. The relevant targets of the identified metabolites, along with associated disease information, were then obtained[26].

Identification of the pharmaceutical ingredients for human diseases-resistance
To identify disease-resistant annotation information, the TCMSP and CancerHSP databases (https://www.tcmsp-e.com/#/database) were used. Pharmaceutical ingredients corresponding to multiple diseases, such as diabetes, cardiovascular diseases, cancer, hypertension, atherosclerosis, thrombotic diseases, and five other related conditions including osteoporosis, liver ischemic injury, inflammation, infectious diseases, and hemorrhage, were identified. Finally, by comparing the metabolites identified through UPLC-MS/MS analysis with these disease-resistant components, the active drug ingredients in *FBSJ* were determined.

Mass spectrometry data analysis and metabolic mechanism analysis
Principal component analysis (PCA), hierarchical cluster analysis (HCA), and orthogonal partial least squares discriminant analysis (OPLS-DA) were performed on the identified metabolites using relevant software. Differential metabolites were selected based on the variable importance in projection (VIP) obtained from the OPLS-DA model, with criteria of VIP ≥ 1, Fold Change (FC) ≥ 2, or FC ≤ 0.5.

Kyoto encyclopedia of genes and genomes (KEGG) annotation and further enrichment analysis
Related metabolic pathways were screened for metabolites from the four *FBSJ* groups with different origins using the KEGG database[27]. Differential metabolites were identified based on two criteria: VIP ≥ 1 and P-value < 0.05 (ANOVA) from the OPLS-DA model. These differential metabolites were aligned and annotated using the KEGG metabolic library and then mapped to the KEGG Pathway database[28]. Further enrichment analysis of the mapped pathways was conducted using the network-based server Metabolite Sets Enrichment Analysis (MSEA). Pathways with $P \leq 0.05$ were considered significantly enriched[29].

Data analysis method
Statistical analysis and figure generation were performed using the software listed in Table 1.

Analysis	Software	Version	Data processing methods
PCA	R (base package)	3.5.1	UV(unit variance scaling)
Heatmap	R (ComplexHeatmap)	2.8.0	UV(unit variance scaling)
OPLS-DA S-plot	R (corrplot)	0.84	–
OPLS-DA	R (MetaboAnalystR)	1.0.1	log2 conversion + centralization

Table 1. Software list information.

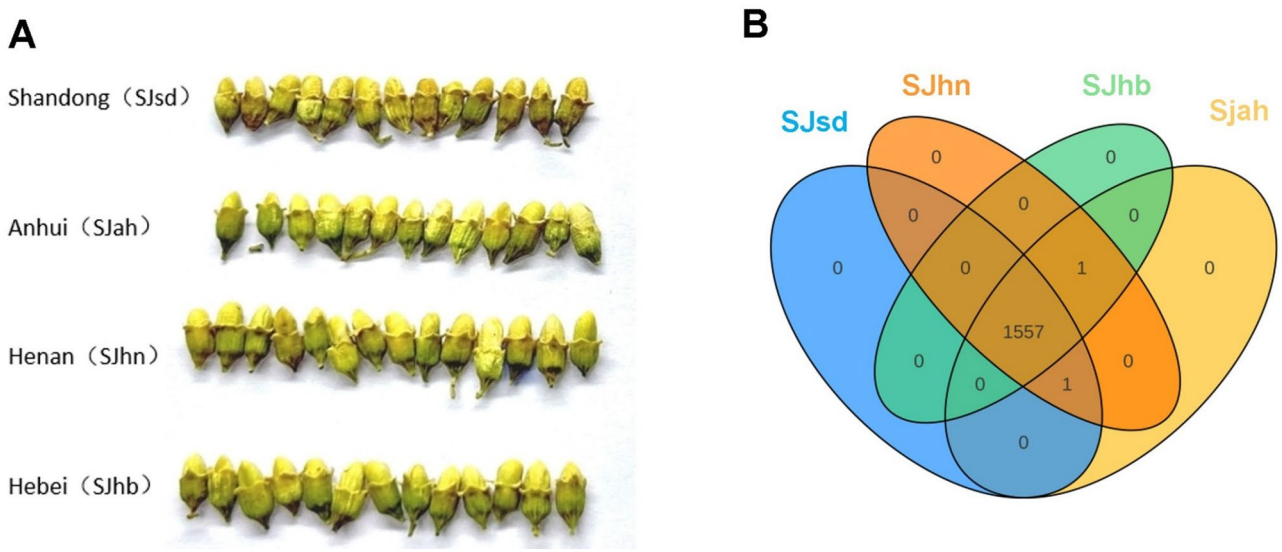

Fig. 1. *FBSJ* from four different origins (**A**) and the Venn diagram of their total metabolites (**B**).

Results and discussion

Overall analysis of the widely targeted metabolites of *FBSJ*

FBSJ, also known as *Sophora japonica (L.)*, were collected from four different geographic origins in China: SJsd from Shandong, SJah from Anhui, SJhn from Henan, and SJhb from Hebei. Figure 1A shows their shape, color, and relative size. Apart from the overall darker color of SJsd, there were no significant differences in appearance among the samples.

In this study, a metabolomics approach was employed to investigate the dynamic changes in metabolites in *FBSJ*. The widely targeted metabolite analysis, conducted using UPLC-MS/MS, identified a total of 1,559 metabolites using the Venn daigram (Fig. 1B). This is the largest number of components detected in *FBSJ* to date, surpassing the previous detection of 331 secondary metabolites[16]. Among these, 1557 metabolites were found to co-exist in SJsd, SJah, SJhb, and SJhn. Additionally, two metabolites, methyl syringate, and His-Thr-Lys-Lys, were found exclusively in specific origins: methyl syringate in SJah, SJhn, and SJhb, and His-Thr-Lys-Lys in SJah, SJhn, and SJsd.

Among all the metabolites detected, 399 flavonoids, 224 phenolic acids, 186 amino acids and their derivatives, 172 lipids, 86 alkaloids, 111 organic acids, 97 terpenoids, 74 nucleotides and their derivatives, 50 lignans and coumarins, 11 tannins, 10 quinones, and 139 other metabolites were identified. In terms of species abundance, flavonoids accounted for approximately 25.5%, phenolic acids for 14.3%, amino acids and their derivatives for 11.9%, and lipids for 11% (see Supplementary Fig. S1 online). In both variety and quantity, this analysis significantly exceeds previously detected secondary metabolites in *FBSJ*. Previous studies identified 173 flavonoids, 53 phenolic acids, 29 organic acids, 23 alkaloids, 15 terpenoids, 11 lignans, coumarins, 6 tannins, and 21 other metabolites[16].

Identification of the key active ingredients in TCM

FBSJ contains 11 types of ingredients, including flavonoids, phenolic acids, organic acids, amino acids and their derivatives, terpenoids, and tannins[30,31]. Although flavonoids such as rutin, quercetin, kaempferol, along with certain saponins, are well-characterized constituents of *FBSJ*, other bioactive compounds with potential health-promoting effects in *FBSJ* remain insufficiently defined. As a result, the functional roles of non-flavonoid metabolites are not yet fully understood, despite their possible contributions in promoting human health.

The identified metabolites were queried in the TCMSP database to identify key active components in *FBSJ* with health-promoting functions for the human body. Among the 1,559 identified metabolites—the largest number reported for *FBSJ* to date—this study significantly expands metabolome coverage, addressing the limitations of previous research and providing a more comprehensive biochemical profile. Out of the 1,559 metabolites identified, 294 were classified as active components of TCM. Using OB ≥ 5% and DL ≥ 0.14 as screening criteria, 135 key active components were identified (Table 2). Although rutin (OB = 3.2%, DL = 0.68) did not meet the predefined screening thresholds, it was nonetheless considered as a key active component of *FBSJ* due to its well-established health-promoting properties.

Among the 135 key active ingredients of TCM, flavonoids were the most abundant, comprising 87 types and accounting for 64.4% of all key active ingredients. Additionally, there were 14 lipids, eight phenolic acids, seven terpenoids, six nucleotides and their derivatives, one tannin, one quinone, and five vitamins, flavonoids, and other bioactive substances. Apart from biochanin A, 3'-methoxydaidzein, quercetin, genistein, glycitein, acacetin, catechin, 6"-O-acetylgenistin, and 3',4',7-trihydroxyflavone, which have been reported in previous studies, the remaining ingredients have not been documented in prior related metabolomics research[16]. Notably, 12 metabolites, despite lacking relevant protein target information, exhibited very high DL values (DL > 0.65). These included 6'-O-malonyl bebeoside, conazole H, rescinin A-7-O-glucoside (Indian santalin), pyrin, terpenoid tritol, kaempferol-7-O-rhamnoside, 3,24-dihydroxypier-12-ene-22 ketol (soy alcohol E), dehydrogated soy saponin I, and apigenin. This indicates that these metabolites may possess significant health-promoting potential and could be further explored for the development of new foods and medicines.

Identification of active pharmaceutical ingredients for 11 human diseases

Given the substantial chemical diversity of *FBSJ*, directly predicting its pharmacological spectrum presents significant challenges[32]. To address this complexity, we systematically selected 11 disease targets for component screening through an integrative strategy. This approach combined global common diseases (such as cancer, cardiovascular disease, and diabetes mellitus)[33], dominant flavonoid constituents (e.g., rutin and quercetin)[34,35], and TCM property-flavor theory documented in the Chinese Pharmacopoeia[36]. This tripartite strategy identified pathologies, including hepatic dysfunction, hemorrhage, and inflammatory bowel diseases. The selection rationale was based on *FBSJ*'s meridian tropism (liver/large intestine) and its dual therapeutic functions: blood-cooling hemostasis and hepatic fire-clearing[36]. In the CancerHSP and TCMSP databases, the top six diseases—cancer/tumor, diabetes, hypertension, cardiovascular disease, atherosclerosis, and thrombotic diseases—along with five other related conditions, including osteoporosis, liver ischemic injury, inflammation, infectious diseases, and hemorrhage, were selected as representative pharmacological roles that metabolites from *FBSJ* extract may contribute to. The active ingredients in *FBSJ* that target these 11 diseases were identified to provide additional information for evaluating the preventive and therapeutic effects of *FBSJ* on these conditions. From the 294 TCM ingredients identified, a total of 193 metabolites were found to be associated with the diseases mentioned above. Most of the identified metabolites were flavonoids, comprising 90 flavonoids, 37 phenolic acids, 15

Classification	Substance name
Phenolic acids	6-O-galloyl-beta-D-glucos, isitagioside, audrosin, neochlorogenic acid (5-O-caffeoyl quinic acid), digallic acid, chlorogenic acid (3-O-caffeoyl quinic acid), di-(2-ethylhexyl) phthalat, allochthyoside
Nucleotides and their derivatives	2'-deoxyadenosine, vernine, cordycepin (3'-deoxyadenosine), uridine 5'-diphosphate-D-glucose, uridine 5'-monophosphate, inosine
Flavone	Ononin 7-O-glucoside (ononin), hispidulin, diosmetin, isorhamnetin, kaempferol-3-o-robinoside 7-o-rhamnoside (robinoside), afromosin, naringenin, pectolinarigenin, calycosin-7-glucoside, biochanin A, 3'-methoxy daidzein, sophora flavanone G, 5,7,4' -trihydroxy-3 '-methoxy isoflavones; 3' -o-methyl vanillin, homoplantine-7-o-glucoside; homoplantagin, sophorin, tectorigenin, baicalein (5,7-dihydroxy-8-methoxy-flavonoids), luteolin, apigenin, quercetin, silylutein (4',5-dihydroxy-6,7-dimethoxy-flavonoids), genistein, morin, isoluteolin, donkey eats phenol, glycitein, 7,4'-di-O-methyl daidzein, cyindigo glycoside, astrapterocarpan irisolidone, acacetin, chrysosine-7-O-glucoside, glycitein 7-O-glucoside, 7-hydroxy-8,4' -dimethoxy-isoflavones, genistin, 6 "-o-malonyl daidatin, kaempferol-7-O-rhamnoside, isoliquiritin, pterocarpin, glycyrrhizin, apigenin-7,4'-dimethyl ether, ononin, kaempferol, eriodictyol, genistein 8-C-glucoside, 3,5,7,4'-tetrahydroxy 8-methoxy-flavone, 6-hydroxyluteolin, catechin, daidzein 7-O-glucoside (daidzein), 3',4',7-trihydroxyflavone, quercetin 3-O-glucoside (isoquercetin), 7-hydroxy-4 '-methoxyisoflavanone, chickpea A-7-O-glucoside (Indian pteroside), okanin, glabridin, 3,5,6,7,8,3',4'-hepmethoxy-flavone, epigallocatechin gallate, aromadendrin (dihydrokaempferol), chickpea alcohol, 3,4',7-trihydroxyflavanones, citrin, medicarpin, liquiritigenin, naringenin-7-O-glucoside (cherry glycosides), quercetin-3-O-(2'-O-galactose) glucoside, kaempferol 3,7-O-diglucoside, 5,2'-dihydroxy-7,8-dimethoxyflavone, jujube side, isobavachin, liquiritigenin-4 '-o-glucoside (liquiritin), isorhamnein-3,7-O-digluboside, trifolirhizin, sophoricoside, 4,4'-dihydroxy-2-methoxychalcone, echinatin, apigenin-8-C-glucoside (vitexin), apigenin-7-O-glucoside, zeeland flavone-7-O-glucoside (nepitrin), 7-O-methylsericin, 3',4',7-trihydroxyflavone, conazol H, acetyldaidzein, luteolin-7,3'-o-diglucoside, trifid carnosin (5-hydroxy-4',6,7-trimethoxy-flavone), eupatilin, kaempferol, acacine-7-O-rutin (linarin), vitexin-2 "-o-rhamnoside, quercetin-3-o-rutin (rutin)
Quinones	Embelin
Others	Capillarisin, riboflavin (vitamin B_2), icariin E5, methyl 14-methylpentadecanoate, 5-O-methylvisamitol glycoside
Lignans and coumarins	Scopolactone-7-O-glucoside (scopolin), stevenin, esculine, 6,7-dihydroxycoumarin-6-O-glucoside, dehydrobisconiferol, eugenolol-4 '-o-glucoside, acanthoside B, olivine resin-4 '-o-glucoside
Tannin	Procyanidin B1
Terpene	3-hydroxylupino-20(29)-ene-28-acid (betulinic acid), soybean saponin βb, methyl ester of 2,3-dihydroxy-arbucarp-12-ene-28-acid (methyl corosolate), 3-hydroxy carpo-12-ene-28-acid (ursolic acid), 3,24-dihydroxyoleanol-12-ene-22-one (soya saponol E), dehydrosoybean saponin I, alizarin terpentriol
Lipid	Ricinoleic acid, 9-hydroxy-10,12,15-octadecatrienoic acid, punicic acid (9Z,11E, 13Z-octadecatrienoic acid), n-eicosanol, α-linolenic acid, hexadecanedioic acid, octadecenoic acid, petroselic acid, methyl linolenate, gingeripid B, 1-oleylglyceride, tetracosanoic acid (lignoceric acid), gingeripid A, eicosenoic acid

Table 2. Key active ingredients of TCM in *FBSJ* with OB ≥ 5% and DL ≥ 0.14.

organic acids, 12 lipids, nine coumarins and lignans, six terpenoids, four alkaloids, four amino acids and their derivatives, three nucleotides and their derivatives, three chromones, three vitamins, two proanthocyanidins, one quinone, one ketone, one alcohol, and two sugars (see Supplementary Tab. S1 online). Based on this, the potential preventive and therapeutic effects of *FBSJ* on the 11 diseases were speculated, or alternatively, the material basis for its therapeutic effects on these diseases was proposed. These 193 metabolites were found to have effects associated with 306 target proteins linked to 335 diseases, indicating that *FBSJ* possesses extensive pharmacological potential contributing to the treatment of the 11 diseases.

Differential metabolites in *FBSJ* from different origins

PCA and HCA cluster analysis

To identify and better understand the differences in metabolites of *FBSJ* from different origins, PCA and HCA were performed. As shown in Fig. 2A, PCA analysis of *FBSJ* from four different origins revealed clear distinctions, with the PC1 and PC2 contribution rates being 49.78% and 14.17%, respectively. The PCA results effectively reflected the metabolic differences between Shandong, Anhui, Henan, and Hebei Provinces. SJsd from Shandong showed the greatest difference from the other three origins in PC1. Although SJhn from Henan and SJah from Anhui appeared relatively similar, the 3D plot in Fig. 2B indicated that these two groups were clearly separated in terms of PC3. Biological replicates from the same origin clustered well together in the PCA 3D plot. However, this unsupervised PCA could not fully account for within-group variations and random errors unrelated to the research[37], making it less effective for identifying group differences. Further research will be conducted using supervised methods. To minimize the influence of quantity on the recognition pattern, the peak area of each metabolite in the sample was transformed logarithmically before performing HCA. As shown in Fig. 2C, the HCA results clearly distinguished the samples by color. A noticeable separation was observed between SJsd and the other samples. The metabolites from SJhn and SJah initially clustered together, then gradually grouped with SJhb, and ultimately converged with SJsd. This clustering pattern reflects the phenotypic differences between the four groups, which align with the findings from the PCA analysis.

The metabolic profiles of SJah, SJhn, and SJhb exhibited some differences compared to SJsd. Both PCA and HCA analyses revealed that the four *Styphnolobium japonicum* plants formed distinct clusters, each characterized by unique metabolic signatures. Among them, SJsd showed the most pronounced divergence from the other three origins.

Cluster analysis was performed to examine the differences in metabolites of *FBSJ* from various origins, resulting in a cluster heat map of the four sample groups (Fig. 2D). The Z-scores of the differential metabolites in *FBSJ* from each geographic origin showed notable variation. Interestingly, the flower buds from the Shandong origin (SJsd) exhibited significantly lower Z-scores for over 100 metabolites compared to the other three origins. Conversely, more than 100 metabolites in SJsd showed markedly higher Z-scores, indicating distinct metabolic enrichment patterns relative to the other samples.

OPLS-DA analysis

OPLS-DA was used to filter out orthogonal variables unrelated to the categorical variables in the metabolites. This method allows for separate analysis of non-orthogonal and orthogonal variables, providing more reliable information about the inter-group differences in metabolites and the correlation degree of experimental groups[38]. To identify the specific differential metabolites responsible for the separation of *FBSJ* from four different origins in China, four comparative OPLS-DA models were established (Fig. 2E). The R^2X and R^2Y values reflect the model's explanatory power for the X and Y variables, respectively, while the Q^2 value indicates the model's predictive ability. A Q^2 value > 0.9 indicates an excellent model, and a Q^2 value > 0.5 indicates an effective model. The model's prediction ability (Q^2) in this study was 0.727, with a P-value of 0.005. This indicates that the OPLS-DA model was well-constructed, reliable, and meaningful. The R^2X and R^2Y values of the OPLS-DA models were 0.549 and 0.996, respectively, showing that the explanatory power for X was good, while the interpretation rate of Y was excellent. Furthermore, the P-value of R^2Y was less than 0.005, indicating that no random grouping model in the permutation test had a better interpretation rate of the Y matrix than the present OPLS-DA model.

The results (see Supplementary Fig. S2 online) demonstrated clear separation among the four sample groups, indicating substantial variation in the first principal predicted component (Y, 33.8%) across the groups. In contrast, no significant differences were observed between SJsd and SJah along the second principal component (Y, 21.2%), which was not associated with the primary discriminating axis. These findings are consistent with the clustering patterns revealed by both PCA and HCA analyses.

Screening of differential metabolites of FBSJ from different origins

Based on the results of OPLS-DA, a pair-based comparison of different producing areas was conducted under the conditions of VIP > 1 and $p < 0.05$ to identify the differential metabolites (Table 3). The analysis revealed that the number of differential metabolites identified between SJsd and the other origins was significantly greater than those observed in comparisons among the remaining origins. A total of 708 differential metabolites, classified into 12 major categories, were identified across the four sample groups. These included 183 flavonoids (comprising 51 flavonoids, 48 isoflavones, 43 flavonols, 10 other flavonoids, and additional subtypes), 108 phenolic acids, 88 amino acids and their derivatives, 81 lipids (including 46 free fatty acids, 11 lysophosphatidylethanolamines, 12 lysophosphatidylcholines, and others), 45 alkaloids (8 indole alkaloids, 3 piperidine alkaloids, 2 pyridine alkaloids, and others), 45 organic acids, 36 terpenoids (30 triterpenoid saponins and 6 triterpenoids), 26 nucleotides and their derivatives, 21 lignans and coumarins (11 lignans and 10 coumarins), 6 tannins, 5 quinones, and 64 metabolites classified under other categories.

Among these differential metabolites in *FBSJ*, flavonoids, phenolic acids, amino acids and their derivatives, and lipids accounted for 25.85%, 15.25%, 12.43%, and 11.44%, respectively. The results indicated 143

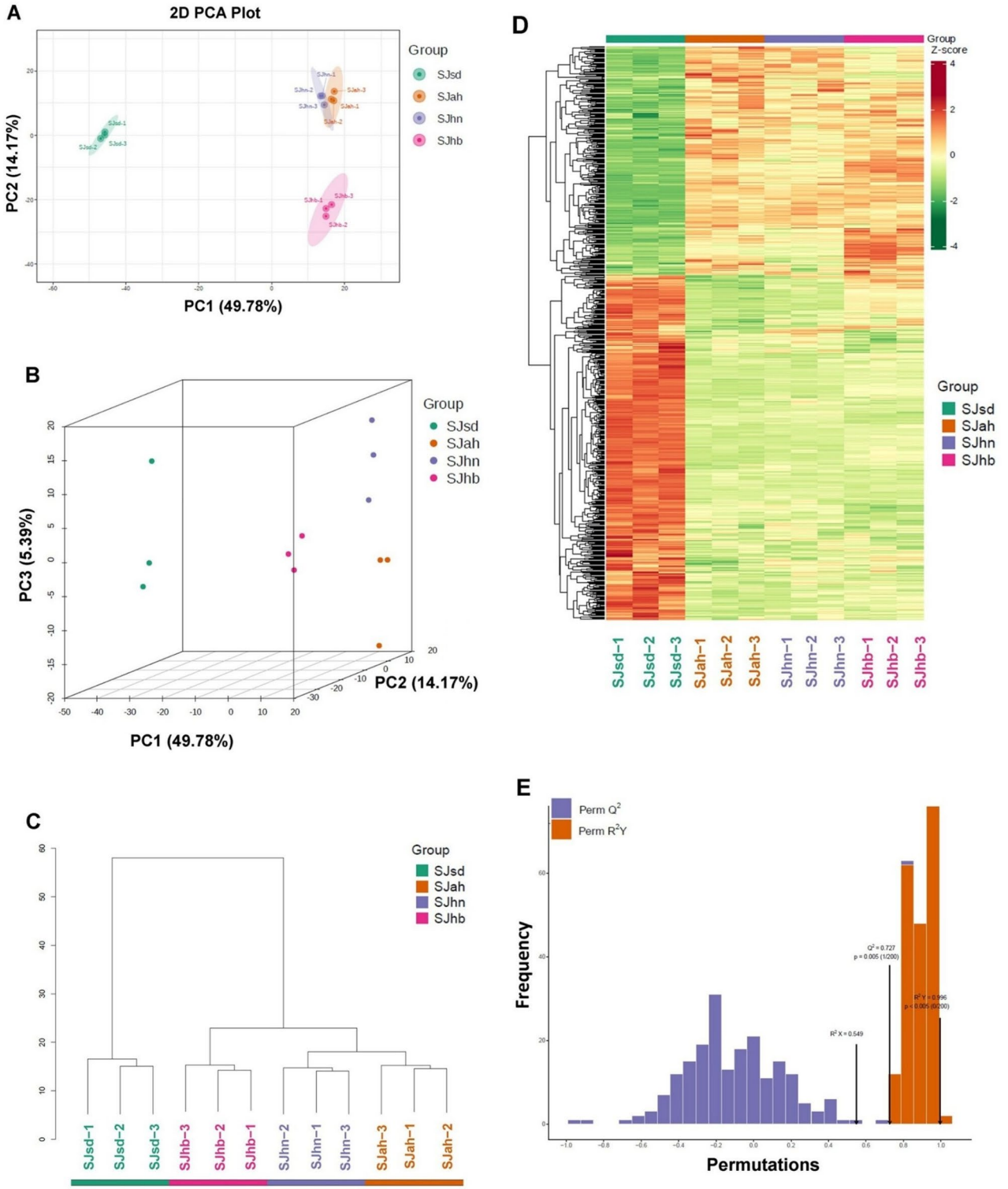

Fig. 2. 2D PCA plot (**A**), 3D PCA plot (**B**), HCA diagram (**C**), Heatmap (**D**), and OPLS-DA model verification diagrams (E) of *FBSJ* from four different origins.

downregulated metabolites, with the highest number observed in the SJsd vs. SJhb comparison. In contrast, 160 metabolites were upregulated, most prominently in the SJsd vs. SJah comparison. Minimal changes were detected between SJah and SJhn, with only two downregulated metabolites, and between SJhn and SJhb, with just three upregulated metabolites. The SJsd group demonstrated the most pronounced metabolic alterations, with 143 metabolites significantly downregulated in comparison to SJhb and 160 metabolites upregulated relative to SJah. These significant metabolic shifts establish SJsd as a key candidate for further mechanistic investigation, as

Group comparison	All differential metabolites (VIP > 1 and $p < 0.05$)			135 Key Active Ingredients (FC ≥ 2 and FC ≤ 0.5)		
	All sig diff	Upregulated	Downregulated	All sig diff	Upregulated	Downregulated
SJah vs. SJhb	41	13	28	7	2	5
SJah vs. SJhn	6	4	2	0	0	0
SJhn vs. SJhb	25	3	22	3	2	1
SJsd vs. SJah	260	160	100	35	35	0
SJsd vs. SJhb	253	110	143	21	17	4
SJsd vs. SJhn	236	147	89	30	30	0
SJsd vs. SJah vs. SJhn vs. SJhb	708	0	0	135	0	0

Table 3. Count of significantly different metabolites by two-by-two comparisons of *FBSJ* in four groups.

its distinct metabolic profile may underlie the observed phenotypic differences. In contrast, minimal intergroup variability was observed among SJah, SJhn, and SJhb, with only slight metabolic divergence between SJah and SJhn. Specifically, only two metabolites exhibited reduced abundance between SJah and SJhn, suggesting that the observed metabolic divergence may be attributable to technical variation rather than biological differences.

Among the 708 differential metabolites in *FBSJ*, 135 metabolites were identified as key active ingredients of TCM. Of these, 102 metabolites were anti-disease active ingredients for 11 diseases, including four phenolic acids, three nucleotides and their derivatives, 73 flavonoids, eight lipids, five lignans and coumarins, three terpenoids, two chromones, one quinone, one vitamin, one ketone, and one tannin. These metabolites represented potential key health-promoting compounds that differentiate *FBSJ* samples according to their geographic origin in this study. The disease spectrum includes six diseases: cancer/tumor, diabetes, hypertension, cardiovascular diseases, atherosclerosis, and thrombotic diseases, alongside five additional diseases: osteoporosis, liver ischemic injury, inflammation, infectious diseases, and hemorrhage. To further elucidate the content variation trends of potential marker compounds in *FBSJ* across four geographic origins, the dominant differential metabolites in *FBSJ* from different origins were screened based on the relative expression ratios among the sample groups. The fold change (FC) values of the metabolites in the comparison group were calculated as FC ≥ 2 or FC ≤ 0.5, resulting in a total of 46 dominant metabolites. These included 35 flavonoids, four coumarins and lignans, three nucleotides and their derivatives, two phenolic acids, one terpenoid, and one tannin. As shown in Table 3, SJsd exhibited a greater number of dominant upregulated metabolites, with 35, 17, and 30 species showing elevated expression in comparisons with SJah, SJhb, and SJhn, respectively. These differential metabolites were predominantly enriched in flavonoids (82.9%), reflecting region-specific biosynthetic specialization. In contrast, SJah, SJhn, and SJhb displayed minimal intergroup variability, with fewer than five differentially abundant metabolites identified between any two groups. These low-magnitude differences reinforce the functional homogeneity among SJah, SJhn, and SJhb, sharply contrasting with SJsd's distinct metabolic identity. These results strongly highlight the distinct metabolic profile of SJsd, which is likely driven by a combination of environmental stressors, genetic adaptations, and specific regulatory mechanisms. In contrast, the other three groups, particularly SJah and SJhn, exhibited minimal metabolic divergence, suggesting a higher degree of similarity in their biochemical composition[39,40].

Moreover, the expression levels of 13 differential metabolites, which could serve as potential anti-disease active ingredients for 11 diseases, were significantly higher in SJsd compared to the other three groups (FC ≥ 2) (Fig. 3). These metabolites included 12 flavonoids and one terpenoid: Diosmetin, 3'-methoxydaidzein, ononin, pectolinarigenin, 7,4'-di-O-methyldaidzein, apigenin-7,4'-dimethyl ether, luteolin, maackiain, isoliquiritin, medicarpin, liquiritigenin, 3,5,6,7,8,3',4'-heptamethoxyflavone, and corosolic acid methyl ester. In the other three groups, differential metabolite expression was largely consistent, with the exception of two metabolites; maackiain and medicarpin in the SJah/SJhb comparison (FC < 0.5). Overall, the FC values for metabolites in SJah, SJhb, and SJhn did not exhibit significant variation.

These ingredients have been shown to possess notable pharmacological activities, including anti-oxidative, anti-inflammatory[41–45], neuroprotective[46,47], anticancer[48–50], blood sugar-regulating[43,46], cardiovascular and cerebrovascular vessel-protecting[43,46,51], antiviral[46], antifungal[46,52], antibiosis[53], anti-allergic[54], osteoporosis-related[46,55,56], PDE4 inhibition[57], analgesic[58], as well as preventive and therapeutic potential for skin disorders[59]. The broad spectrum of pharmacological activities observed in SJsd suggests that it may exert more favorable therapeutic effects compared to the other three groups in these respective aspects.

Differential metabolite KEGG metabolic pathway analysis

Pathway enrichment analysis of 708 different metabolites across the four *FBSJ* groups was performed using the KEGG database based on the characteristics of the differential metabolites. In total, 382 metabolites were annotated by KEGG, with 176 showing significant differences. These metabolites were distributed across 87 metabolic pathways. Among these, the most significant enrichment was observed in the isoflavone biosynthesis pathway ($p < 0.01$), followed by neomycin, kanamycin, and gentamicin biosynthesis ($p < 0.05$), starch and sucrose metabolism, aminoacyl-tRNA biosynthesis, linoleic acid metabolism, and indole alkaloid biosynthesis pathways ($p > 0.05$). The observed upregulation of key metabolites within the isoflavonoid biosynthesis pathway in SJsd suggests a potentially enhanced isoflavone synthesis capacity relative to other origins. However, transcriptomic or proteomic validation is required to confirm the underlying enzymatic mechanisms driving this metabolic distinction.

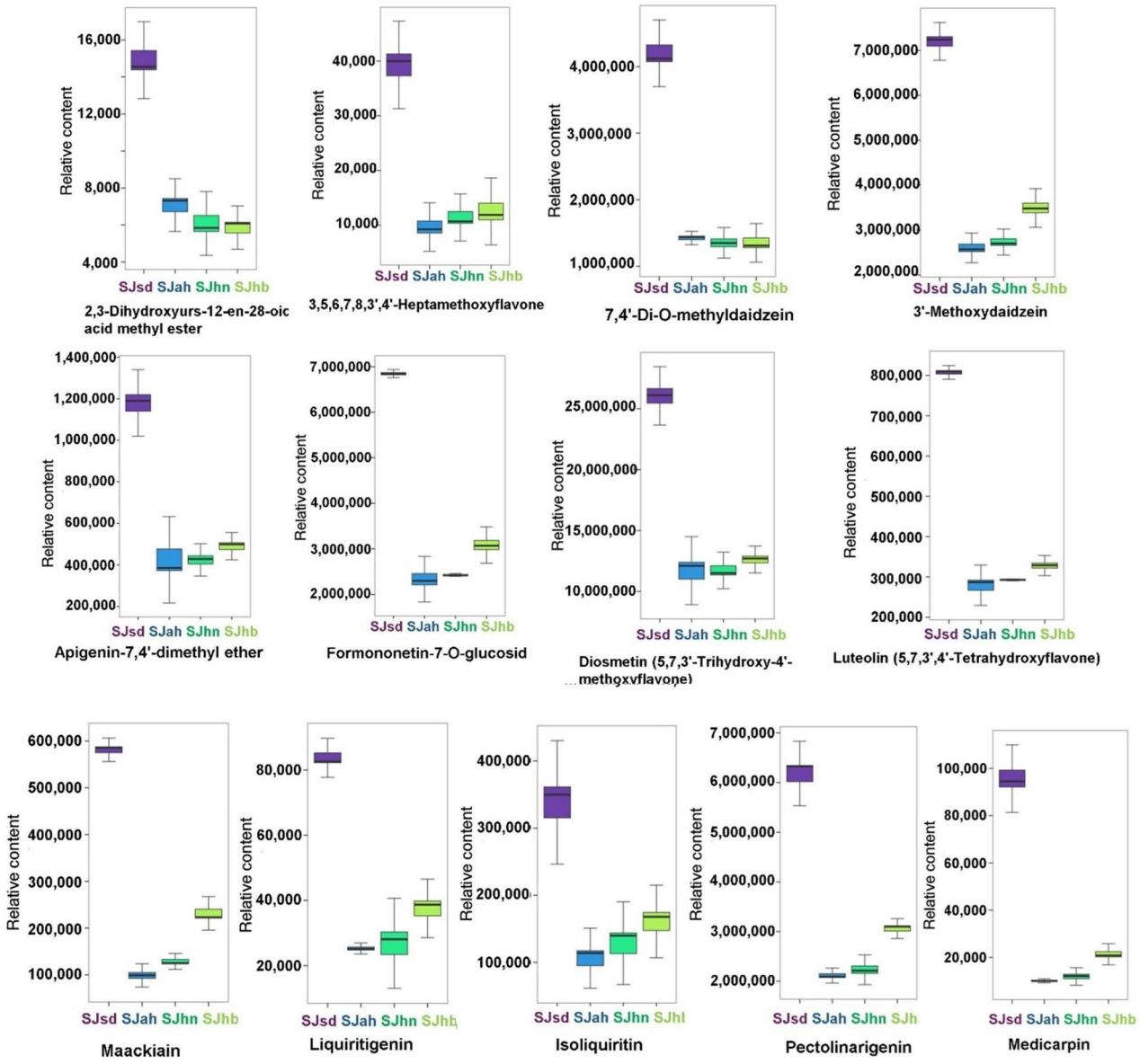

Fig. 3. Boxplots of the 13 dominant metabolites, comparing their relative contents in the flower buds of SJsd, SJah, SJhn and SJhb.

Based on the KEGG database, metabolic pathway enrichment analysis was conducted for the three contrast combinations—SJsd vs. SJah, SJsd vs. SJhb, and SJsd vs. SJhn—using a large number of differential metabolites. This analysis helped to understand the mechanisms behind changes in differential metabolites within metabolic pathways. The Differential Abundance (DA) Score is a pathway-based metabolic change analysis method that captures the overall change of all metabolites in a pathway[60]. Unlike the KEGG enrichment bubble map, the DA score map includes line segments, with the length representing the absolute value of the DA Score. The size of the dots at the end of the line segments indicates the number of differential metabolites in the pathway. The dots are positioned to the left of the central axis, and the longer the line segment, the more the overall expression of the pathway tends to be downregulated compared to the previous state.

In the SJsd vs. SJah comparison, a total of 260 different metabolites were annotated across 51 metabolic pathways. Notably, the isoflavonoid biosynthesis pathway exhibited a high degree of metabolite enrichment and significant variation ($p < 0.05$) (Fig. 4A), highlighting its potential role in the metabolic divergence between these two origins. 21 significant differential metabolites were involved in this pathway, including 7,4'-dihydroxyflavone, isoformononetin, 6"-O-malonyldaidzin, pseudobaptigenin, daidzein, 6"-O-malonylglycitin, formononetin-7-O-glucoside (ononin), formononetin-7-O-(6"-malonyl)glucoside, glycitein, genistein, daidzein-7-O-glucoside (daidzin), 2'-hydroxygenistein, calycosin, 4',5,7-trihydroxyflavone (apigenin), biochanin A, liquiritigenin, maackiain, 5,4'-dihydroxy-7-methoxyisoflavone (prunetin), 3,9-dihydroxypterocarpan, medicarpin, and

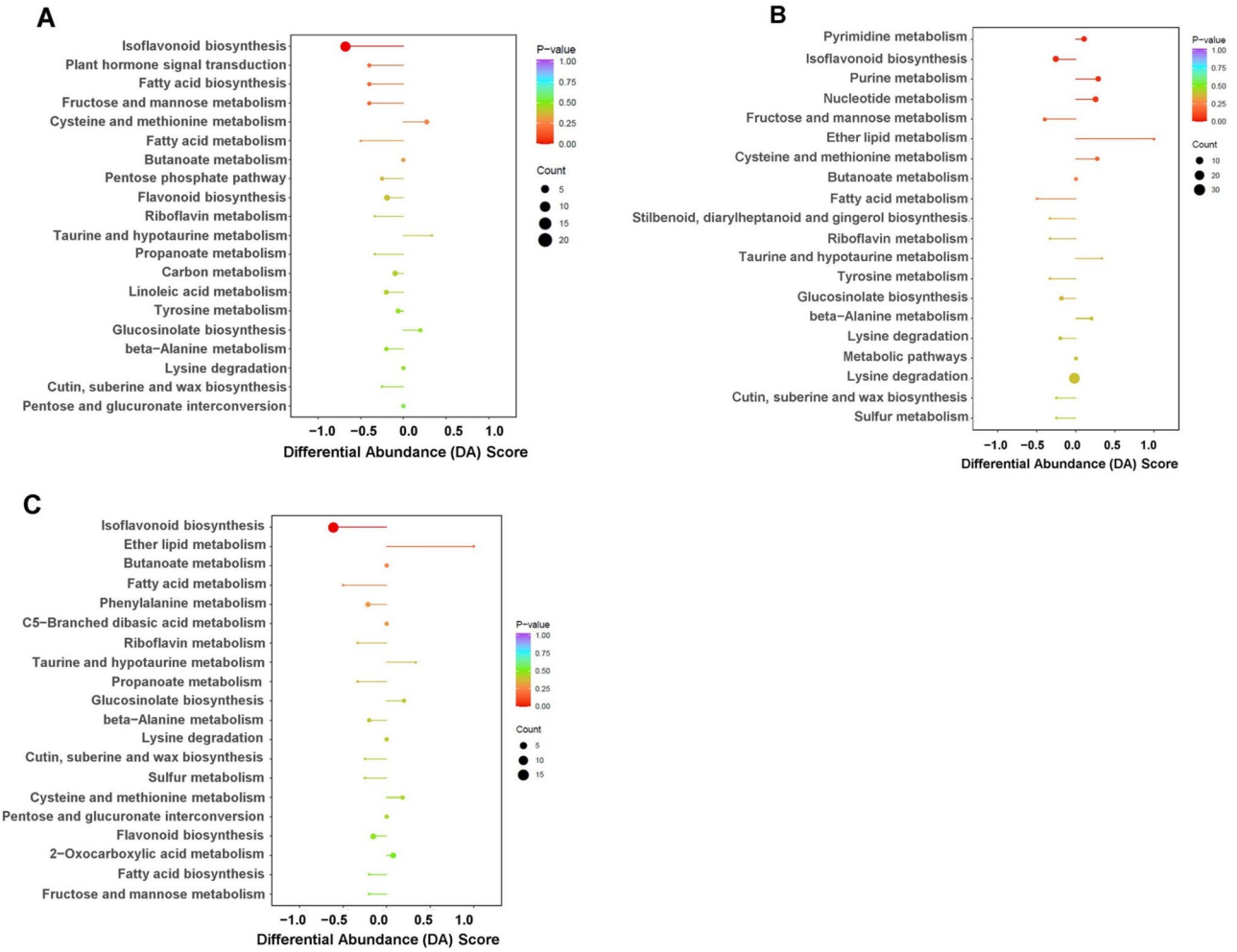

Fig. 4. KEGG enrichment of common different metabolites (DA score) between SJsd and SJah (**A**), SJsd and SJhb (**B**), and SJsd and SJhn (**C**).

7-hydroxy-4'-methoxyisoflavone (formononetin). The overall expression of SJah in this metabolic pathway was downregulated compared to SJsd, indicating a reduced biosynthetic activity associated with this metabolic route.

In the SJsd vs. SJhb comparison, a total of 253 different metabolites were annotated in the KEGG database and mapped to 57 distinct metabolic pathways (Fig. 4B). The metabolic pathways exhibiting high concentrations of metabolites and significant differences ($p < 0.05$) included pyrimidine metabolism, isoflavonoid biosynthesis, and purine metabolism. The overall expression of metabolites involved in isoflavonoid biosynthesis was more markedly downregulated in SJah compared to SJsd, with an even greater downregulation trend observed in the SJsd vs. SJhb comparison. The differential metabolites involved in isoflavonoid biosynthesis included 7,4'-dihydroxyflavone, medicarpin, formononetin-7-O-glucoside (ononin), 3,9-dihydroxypterocarpan, daidzein, 6"-O-malonylglycitin, maackiain, and liquiritigenin. In addition, the overall expression of SJhb in pyrimidine metabolism, purine metabolism, and nucleotide metabolism was relatively upregulated compared to that of SJsd. These three pathways are associated with nucleotide metabolism. The differential metabolites involved in pyrimidine metabolism included nucleotides and their derivatives, as well as organic acids. The organic acids consisted of malonic acid and 3-hydroxypropanoic acid, while the nucleotides and derivatives included 2'-deoxycytidine, uridine 5'-monophosphate, barbituric acid, malonylurea, 2,4,6-pyrimidinetrione, and cytidine 5'-monophosphate (cytidylic acid). The differential metabolites in purine metabolism were nucleotides and derivatives, including guanosine 5'-monophosphate, cyclic 3',5'-adenylic acid, 2'-deoxyadenosine, 2'-deoxyinosine, guanosine, and guanosine 3',5'-cyclic monophosphate. Differential metabolites involved in nucleotide metabolism contributed to guanosine 5'-monophosphate, cyclic 3',5'-adenylic acid, 2'-deoxyadenosine, 2'-deoxyinosine, guanosine, and guanosine 3',5'-cyclic monophosphate. These findings suggest that SJhb's metabolic profile reflects elevated cellular activity[61] (e.g., proliferation, stress adaptation), which may be influenced by environmental factors or regulated through genetic mechanisms involving enzymes like PRPP synthetase[62].

For the shared differential metabolites between SJsd and SJhn, a total of 236 different metabolites were mapped to 54 metabolic pathways (Fig. 4C). Among these pathways, isoflavonoid biosynthesis was the pathway with a high concentration of metabolites and a significant difference ($p < 0.05$). In this pathway, 19

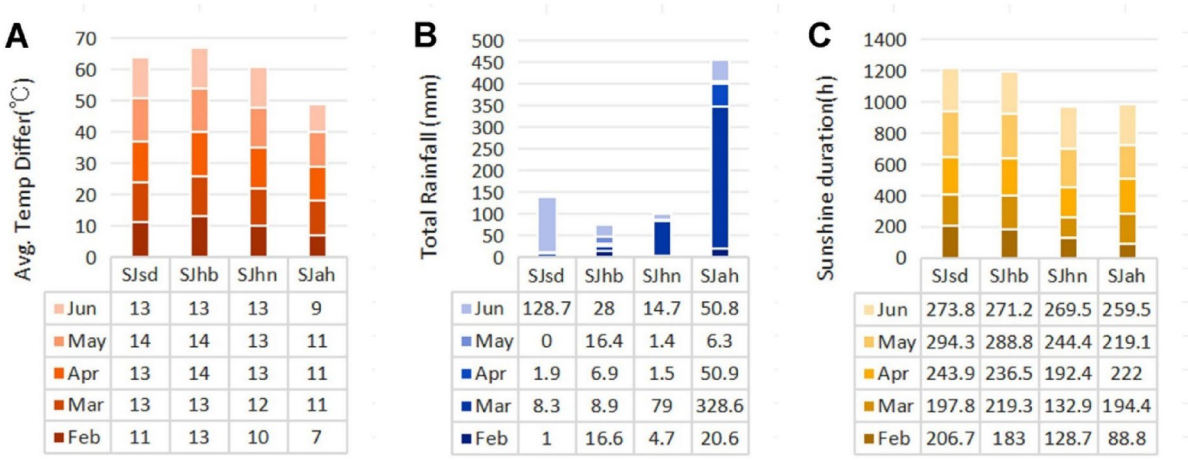

Fig. 5. Column chart of temperature difference(**A**), total rainfall(**B**) and sunshine duration(**C**) accumulation in four regions (Feb to Jun 2022). Note: sunshine duration data were based on data from nearby stations in Shandong Jinan, Hebei Baoding, Henan Zhengzhou, and Anhui Hefei.

common differential metabolites were identified, including 7,4'-dihydroxyflavone, genistein, formononetin-7-O-(6'-malonyl)glucoside, 3,9-dihydroxypterocarpan, formononetin (7-hydroxy-4'-methoxyisoflavone), prunetin (5,4'-dihydroxy-7-methoxyisoflavone), formononetin-7-O-glucoside (ononin), maackiain, apigenin, 4',5,7-trihydroxyflavone, medicarpin, daidzein-7-O-glucoside (daidzin), 2'-hydroxygenistein, calycosin, 6"-O-malonylglycitin, pseudobaptigenin, daidzein, liquiritigenin, biochanin A, and 6"-O-malonyldaidzin. The overall expression of metabolites within this pathway was more substantially downregulated in SJhn compared to SJsd, indicating a lower biosynthetic activity associated with this metabolic route.

Across the SJsd vs. SJhn, SJhb, and SJah comparisons, the isoflavone metabolic pathway emerged as a common differential metabolic pathway, encompassing a total of 21 distinct metabolites. A systematic analysis of the KEGG isoflavone metabolic pathway (ko00943) revealed that three components—7,4'-dihydroxyflavone, liquiritigenin, and 4',5,7-trihydroxyflavone (apigenin)—were significantly upregulated in SJsd compared to SJah, indicating enhanced pathway activity in the SJsd group. These components are primarily regulated by FNS I and CYP93B2_16 in flavonoid metabolism. The remaining 18 differentially accumulated compounds were identified as isoflavones, including isoformononetin, 6"-O-malonyldaidzin, pseudobaptigenin, daidzein, 6"-O-malonylglycitin, formononetin-7-O-glucoside (ononin), formononetin-7-O-(6"-malonyl)glucoside, glycitein, genistein, daidzein-7-O-glucoside (daidzin), 2'-hydroxygenistein, calycosin, biochanin A, maackiain, 5,4'-dihydroxy-7-methoxyisoflavone (prunetin), 3,9-dihydroxypterocarpan, medicarpin, and 7-hydroxy-4'-methoxyisoflavone (formononetin). These isoflavone components mainly fall into four functional categories: ① precursors (e.g., genistein); ② O-methylated products (e.g., biochanin A); ③ glycosylated derivatives (e.g., daidzin); and ④ acylated modifications. Metabolic regulatory network analysis revealed that the synergistic action of key enzymes, including core skeleton synthases (CYP93C, HIDH), hydroxylation modification enzymes (CYP81E9, CYP81E1/E7), O-methyltransferases (7-IOMT, HI4OMT), glycosylation/acylation enzymes (IF7GT, IF7MAT), and specific product synthases (CYP93A1), was the core mechanism driving the differential accumulation of metabolites.

Notably, the differential metabolites, including formononetin, pseudobaptigenin, 2'-hydroxygenistein, 6"-O-malonyldaidzin, daidzein-7-O-glucoside (daidzin), formononetin-7-O-glucoside (ononin), calycosin, apigenin, genistein, prunetin, biochanin A, glycitein, and isoformononetin, are common in the comparisons of SJsd vs. SJah and SJsd vs. SJbn, but absent in SJsd vs. SJhb. This differential trend aligns with variations in temperature, precipitation, and sunshine hours. Thus, it can be speculated that the distinct metabolic profile of SJsd and SJhb may be associated with their unique geographic origins in Shandong and Hebei Province. Environmental stressors prevalent in these regions could drive adaptive phytochemical responses[63]. Greater fluctuations in average temperature (Fig. 5A), reduced precipitation Fig. 5B), and prolonged sunshine duration (Fig. 5C) may serve as key factors influencing the biosynthesis of these components by isoflavone metabolic regulatory enzymes. Although this association is speculative and based on observed correlations between climatic variables and metabolite content, establishing a causal relationship between environmental factors and metabolite biosynthesis will require further validation through controlled experimental studies. This finding aligns with a wide range of existing literature. For instance, controlled experiments have demonstrated the regulatory influence of temperature and drought stress on soybean isoflavone content, supporting the hypothesis that climate factors (such as temperature fluctuations, and reduced precipitation) modulate isoflavone biosynthesis *via* metabolic enzyme regulation[64,65]. Additionally, prunetin and biochanin A levels in *Sophora alopecuroides L.* have been shown to significantly increase under mild drought stress, whereas isoflavonoids content markedly decrease under moderate and severe drought conditions[66].

In contrast, while Hebei Province shares broadly similar environmental conditions with Shandong, SJhb specimens exhibited lower isoflavone content than SJsd, despite exceeding the levels observed in samples from Henan and Anhui Provinces (Fig. 5). Differential metabolites identified across the SJsd vs. SJhn, SJhb,

and SJah comparisons include maackiain, medicarpin, 3,9-dihydroxypterocarpan, 6”-O-malonylglycitin, 7,4’-dihydroxyflavone, daidzein, liquiritigenin, and formononetin-7-O-(6”-malonyl) glucoside. All of them are associated with isoflavonoid biosynthesis and contribute to the metabolic divergence among these origins. These observations suggest that, beyond climatic influences, additional factors significantly affect the regulation of key enzymes within the isoflavone metaboloc pathway—particularly core skeleton synthases (HIDH), glycosylation and acylation enzymes (IF7GT, IF7MAT), specific product synthases (CYP93A1), and flavone synthases I and II. These findings align with existing research. Under UV-B stress, the levels of total isoflavones and aglycones (genistein, daidzein, and glycitein) increased. When irradiated with UV-B, H_2 treatment suppressed the activity and gene expression of IF7GT and IF7MaT. Additionally, appropriate line spacing has been shown to enhance daidzein accumulation in soybean seeds[67]. Collectively, these environmental and agronomic factors offer a foundational reference for the cultivation, screening, development, and application of *FBSJ*.

However, it should be noted that, beyond climatic influences, additional variables likely contribute to the differential accumulation of metabolites in *FBSJ*, potentially through modulation of enzymatic activity and regulatory pathways. These variables include the composition of the soil microbiome[39,68], epigenetic regulation[68], and anthropogenic influences[69,70], which likely modulate differential accumulation of metabolites in *FBSJ*. These factors act synergistically to finely regulate isoflavone biosynthesis. In particular, the complex interactions between rhizosphere microorganisms and plants may influence the plant’s metabolic processes, thereby promoting the synthesis and accumulation of isoflavones[71]. The metabolic pathway of isoflavones is highly complex and involves the coordinated action of multiple enzymes. In addition to the previously mentioned enzymes, several upstream biosynthetic enzymes—including phenylalanine ammonia-lyase (PAL), cinnamate-4-hydroxylase (C4H), and 4-coumarate-CoA ligase (4CL), along with various transcription factors, also play critical roles in regulating isoflavone biosynthesis[72]. Moreover, genetic variation among different cultivars or individual plants can result in differential expression of isoflavone metabolism-related genes, enzyme activities, and variability in the overall efficiency of the metabolic pathway[73,74]. Transcriptomic analysis further revealed that the upregulation of upstream synthetic genes (e.g., PAL, CHS, F3H) and transcription factors (e.g., MYB, bHLH) in flavonoid biosynthesis may be mediated by epigenetic regulation[74]. These genetic adaptations, combined with localized abiotic pressures, likely act synergistically to enhance the production of secondary metabolites, including bioactive compounds with potential health-promoting properties.

Conclusions

This study systematically characterized the chemical diversity of *FBSJ* from four geographic origins using UPLC-MS/MS-based widely targeted metabolomics. Multivariate analyses (PCA, HCA, and OPLS-DA) revealed significant geographic variation in metabolic profiles. Beyond flavonoids—the dominant class of compounds—other metabolite categories such as lipids, phenolic acids, lignans, coumarins, nucleotides, and terpenoids showed origin-specific patterns of accumulation. These findings suggest their potential as novel chemotaxonomic markers for quality control of *FBSJ*. Among the 1,559 identified metabolites, the largest number reported for *FBSJ* to date, 135 metabolites were recognized as key active ingredients used in TCM screened via TCMSP database, and 193 metabolites were annotated as active pharmaceutical components linked to 11 major diseases (e.g., cardiovascular disorders, diabetes) via TCMSP and CancerHSP databases. Comparative metabolomic profiling revealed that SJsd specimens contained higher concentrations of both TCM-associated and disease-targeting metabolites compared to SJhb, SJah, and SJhn, suggesting superior therapeutic potential and biochemical richness. Notably, 13 critical pharmacological compounds (e.g., 3’-Methoxydaidzein, Maackiain, Medicarpin) were found to be over twofold more abundant in SJsd highlighting its potential as a superior source material for investigating the health-promoting properties of *FBSJ*, particularly in relation to analgesic, anti-inflammatory, and bone-conserving effects. This finding provides empirical support for the origin–quality correlation, addressing a gap previously unclarified in the literature. KEGG pathway analysis identified isoflavonoid biosynthesis (ko00943) as the most differentially regulated pathway. SJsd showed enriched expression of metabolites associated with key biosynthetic enzymes, (e.g., CYP93A1, IF7GT), which may be influenced by Shandong’s favorable climate conditions, such as moderate temperature fluctuations, lower precipitation, and extended sunlight exposure. Combined with epigenetic factors, these environmental elements likely enhance isoflavone synthesis through effects on gene expression, metabolic regulation, and microbial community structure.

This work represents a comprehensive metabolic analysis of *FBSJ*, offering new insights into metabolite variability across geographic origins. These findings not only resolve the ambiguity surrounding origin–quality correlations but also provide pharmacological assessments and elucidate metabolite–disease associations linked to differential metabolite accumulation in *FBSJ*. It establishes a foundational framework for germplasm evaluation, geographic origin traceability, and the strategic development of functional food products derived from *FBSJ*. Future research could integrate transcriptomic data and soil metagenomics to elucidate gene–environment interactions contributing to chemotypic variation in *FBSJ*, thereby advancing our understanding of how genetic regulation and microbial ecology jointly influence secondary metabolite biosynthesis.

Data availability

Data is provided within the manuscript or supplementary information files.The experimental data and the simulation results that support the findings of this study are the original research. For the qualitative analysis of substances based on the criteria of fragmentation pattern, retention time, Metware database (MWDB) V2.0 and public database established by m/z and Metware Biotechnology Co., Ltd (Wuhan, China). The statistical analysis and figures were carried out by software in the provided in the list: R program Version 2.8.0 for Complex Heatmap, R program Version 0.84 for Corrplot, R program Version 0.84 for MetaboAnalystR. All data supporting in this study related to TCMSP and CancerHSP database were provided in https://www.tcmsp-e.com/#/home. All

data supporting in this study related to KEGG database were provided in http://www.kegg.jp/kegg/compound/ and http://www.kegg.jp/kegg/pathway.html. Climatic data (temperature, rainfall, and sunshine duration) for the four production regions from February to May 2022 were obtained from publicly available sources: temperature and rainfall data from https://www.tianqi24.com and sunshine duration from https://www.ceicdata.com/zh-hans/china/sunshine-hours.

Received: 12 August 2024; Accepted: 28 October 2025
Published online: 19 November 2025

References

1. Su, F. *Study on the extraction and purification of the secondary metabolites with biological activities from Sophora Japonica Roots* Master thesis, Tianjin University of Science and Technology, (2017).
2. Miao, M. S., Cheng, B. L. & Jiang, N. Effect of sophora Japonica total flavonoids on mouse models of hyperglycemia and diabetes model. *Appl. Mech. Mater.* **664**, 397–401. https://doi.org/10.4028/www.scientific.net/AMM.664.397 (2014).
3. Younsook, K. et al. Prediction of the therapeutic mechanism responsible for the effects of sophora Japonica flower buds on contact dermatitis by network-based Pharmacological analysis. *Journal Ethnopharmacology* **271** (2021).
4. Wang*, C., An, R., Li, Z. & Chen, M. in *International Conference of National Product and Traditional Medicine.* 326–329.
5. Jung, C. H. et al. Antihyperglycemic activity of herb extracts on streptozotocin-induced diabetic rats. *Biosci. Biotechnol. Biochem.* **70**, 2556–2559 (2006).
6. Cao, X. & Yang, H. Study on enzymatic-ultrasonic assisted extraction of polysaccharides from sophora Japonica Linn. And its free radical scavenging activity. *Northern Horticulture*, 141–146 (2016).
7. Shlapak, V. P. & Pukas, S. S. Vegetative reproduction of sophora Japonica L. with stem grafts. *Plant Introduction* **36** (2007).
8. Jing, T., Yue, P. & Zhong, Z. Drought, Salinity, and Low Nitrogen Differentially Affect the Growth and Nitrogen Metabolism of Sophora japonica (L.) in a Semi-Hydroponic Phenotyping Platform. *Frontiers in Plant Science* 12 (2021).
9. Li, L. et al. Protective effect of polysaccharide from sophora Japonica L. flower buds against UVB radiation in a human keratinocyte cell line (HaCaT cells). *Journal Photochem. & Photobiology B: Biology* **191** (2018).
10. Xie, Z., Lam, S. & Wu, J. Chemical fingerprint and simultaneous determination of flavonoids in Flos sophorae Immaturus by HPLC-DAD and HPLC-DAD-ESI-MS/ MS combined with chemometrics analysis. *Analytical Methods* **6** (2014).
11. Zha, X. Q. et al. Antioxidant properties of polysaccharide fractions with different molecular mass extracted with hot-water from rice Bran. *Carbohydr. Polym.* **78**, 570–575. https://doi.org/10.1016/j.carbpol.2009.05.020 (2009).
12. Wang, F. et al. Isorhamnetin, the Xanthine oxidase inhibitor from *Sophora japonica*, ameliorates uric acid levels and renal function in hyperuricemic mice. *Food Funct.* **12**, 12503–12512. https://doi.org/10.1039/D1FO02719K (2021).
13. Yiqing, G.et al. The Combination of Scutellaria baicalensis Georgi and Sophora japonica L. ameliorate Renal Function by Regulating Gut Microbiota in Spontaneously Hypertensive Rats.*Frontiers in Pharmacology* 11 (2021).
14. Xiangcai, M., Xiaoying, L., Jie, Y., Ling, K. & Yu, G. Mechanism of ecological stress enhancing quality of genuine medicinal materials and its quality evaluation Idea. *Chin. Traditional Herb. Drugs.* **53**, 1587–1594 (2022).
15. Xirui He, Y. B. et al. Local and traditional uses, phytochemistry, and Pharmacology of sophora Japonica L.: A review. *J. Ethnopharmacol.* **187**, 160–182. https://doi.org/10.1016/j.jep.2016.04.014 (2016).
16. Wang, J. R. et al. Metabolomic analysis reveals dynamic changes in secondary metabolites of sophora Japonica L. during flower maturation. *Front. Plant Sci.* **13** https://doi.org/10.3389/fpls.2022.916410 (2022).
17. M, K. & S, G. KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research* 28 (2000).
18. Hongren, Z. et al. An end-to-end weakly supervised learning framework for cancer subtype classification using histopathological slides. *Expert Syst. Applications* **237** (2024).
19. A., J. P. et al. Raman spectroscopy analysis of plasma of diabetes patients with and without retinopathy, nephropathy, and neuropathy. *Spectrochimica Acta Part. A: Mol. Biomol. Spectroscopy* **304** (2024).
20. Richa, A., Thapa, M. P. & Kumar, K. R. Fok I and Bsm I gene polymorphism of vitamin D receptor and essential hypertension: a mechanistic link. *Clin. Hypertens.* **29**, 5 (2023).
21. Audesh, B., Arti, D., Kaushik, D. & Editorial Endocrine-related cardiovascular diseases: recent advances in diagnosis and treatment. *Front. Endocrinol.* **14**, 1147752 (2023).
22. Fengge, W. et al. Dendritic cell-expressed IDO alleviates atherosclerosis by expanding CD4 + CD25 + Foxp3 + Tregs through IDO-Kyn-AHR axis. *Int. Immunopharmacol.* **116**, 109758 (2023).
23. Zhongyu, H. et al. The role of monocytes in thrombotic diseases: a review. *Front. Cardiovasc. Med.* **10**, 1113827 (2023).
24. Lin, G., Zhifeng, X., Huachao, Y. & Honglin, W. Apoptosis effect and mechanism study of Quercetin on the breast cancer MCF-7 and MDA-MB-435 cells. *Progress Mod. Biomed.* **21**, 3638–3644 (2021).
25. Wang, P. et al. Widely targeted metabolomic and transcriptomic analyses of a novel albino tea mutant of Rougui. *Forests* **11** (2020).
26. Xiaohu, W. et al. Bioinformatics-based screening of hub genes in colon adenocarcinoma and active ingredients of anti-cancer herbal traditional Chinese medicine. *Chin. Remedies Clin.* **22**, 614–617 (2022).
27. Kanehisa, M. & Goto, S. K. E. G. G. Kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* **28**, 27–30. https://doi.org/10.1093/nar/28.1.27 (2000).
28. Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M. & Tanabe, M. KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* **44**, D457–D462. https://doi.org/10.1093/nar/gkv1070 (2016).
29. Hongyou, L. et al. Comparative metabolomics study of Tartary (Fagopyrum Tataricum (L.) Gaertn) and common (Fagopyrum esculentum Moench) buckwheat seeds. *Food Chemistry* 371 (2021).
30. Fengping, F. et al. Comparison of quality differences of flavonoids in *Sophora japonica L.* flowers and *Roinia pseudoacacia L.* flowers. *Hubei Agricultural Sci.* **61**, 108–112. https://doi.org/10.14088/j.cnki.issn0439-8114.2022.07.020 (2022).
31. Xiao, W., Yu, W. & Bing, Z. Zhi-jian, L. Research progress on herbaceous,chemical constituents and Pharmacological effects of different medicinal parts of sophora Japonica. *Chin. Traditional Herb. Drugs.* **49**, 4461–4467 (2018).
32. Newman, D. J. & Cragg, G. M. Natural products as sources of new drugs over the nearly four decades from 01/1981 to 09/2019. *J. Nat. Prod.* **83**, 770–803. https://doi.org/10.1021/acs.jnatprod.9b01285 (2020).
33. Rana, J. S., Khan, S. S., Lloyd-Jones, D. M. & Sidney, S. Changes in mortality in top 10 causes of death from 2011 to 2018. *J. Gen. Intern. Med.* **36**, 2517–2518 (2021).
34. Zhao, X. et al. Sophoridine from sophora flower attenuates ovariectomy induced osteoporosis through the RANKL-ERK-NFAT pathway. *J. Agric. Food Chem.* **65**, 9647–9654 (2017).
35. Li, Y. et al. Inflammation and immunity. *Nutrients* **8**, 167. https://doi.org/10.3390/nu8030167 (2016). Quercetin.
36. Chinese Pharmacopoeia Commission. *Pharmacopoeia of the People's Republic of China (Vol. I)* (China Medical Science, 2020).
37. Sven Serneels, T. V. Principal component analysis for data containing outliers and missing elements. *Comput. Stat. Data Anal.* **52**, 1712–1727 (2008).
38. Chuanhe, L. et al. Diversity analysis of differential expressed genes and differential metabolites between two pineapple cultivars with different stress resistance. *Acta Bot. Boreali-Occidentalia Sinica.* **42**, 1514–1522 (2022).

39. Wang, J. R. et al. Variations in the components and antioxidant and tyrosinase inhibitory activities of styphnolobium Japonicum (L.) Schott extract during flower maturity stages. *Chem. Biodivers.* **16**, e1800504. https://doi.org/10.1002/cbdv.201800504 (2019).
40. Trush, K. Pal'ove-Balang, P. Biosynthesis and role of isoflavonoids in legumes under different environmental conditions. *Plant. Stress*. **8**, 100153. https://doi.org/10.1016/j.stress.2023.100153 (2023).
41. Yang, Y. et al. Diosmetin exerts anti-oxidative, anti-inflammatory and anti-apoptotic effects to protect against endotoxin-induced acute hepatic failure in mice. *Oncotarget* **8**, 30723–30733 (2017).
42. Huh, J. W. et al. Maackiain, a compound derived from sophora flavescens, increases IL-1β production by amplifying nigericin-mediated inflammasome activation. *FEBS Open. Bio*. **10**, 1482–1491 (2020).
43. Krishna, M. S., Joy, B. & Sundaresan, A. Effect on oxidative stress, glucose uptake level and lipid droplet content by apigenin 7, 4'-dimethyl ether isolated from Piper longum L. *J. Food Sci. Technol.* **52**, 3561–3570 (2015).
44. Okuyama, S. et al. 3′, 4′-heptamethoxyflavone, a citrus polymethoxylated flavone, attenuates inflammation in the mouse hippocampus. *Brain Sci.*. **3, 5**, (6), 118–129 (2015).
45. Omasa, T., Sawamoto, A., Nakajima, M. & Okuyama, S. Anti-Inflammatory and neurotrophic factor production effects of 3, 5, 6, 7, 8, 3′, 4′-Heptamethoxyflavone in the hippocampus of Lipopolysaccharide-Induced inflammation model mice. *Molecules* **29**, 5559 (2024).
46. Bhuia, M. S. et al. Therapeutic potentials of Ononin with mechanistic insights: A comprehensive review. *Food Bioscience*. **56**, 103302. https://doi.org/10.1016/j.fbio.2023.103302 (2023).
47. Ramalingam, M., Kim, H., Lee, Y. & Lee, Y. I. Phytochemical and Pharmacological role of Liquiritigenin and Isoliquiritigenin from radix glycyrrhizae in human health and disease models. *Front. Aging Neurosci.* **10** https://doi.org/10.3389/fnagi.2018.00348 (2018).
48. Imran, M. et al. Luteolin, a flavonoid, as an anticancer agent: A review. *Biomed. Pharmacother.* **112**, 108612 (2019).
49. Zhou, Y. & Ho, W. S. Combination of liquiritin, isoliquiritin and isoliquirigenin induce apoptotic cell death through upregulating p53 and p21 in the A549 non-small cell lung cancer cells. *Oncol. Rep.* **31**, 298–304 (2014).
50. Iwase, Y. et al. Cancer chemopreventive activity of 3, 5, 6, 7, 8, 3′, 4′-heptamethoxyflavone from the Peel of citrus plants. *Cancer Lett.* **163**, 7–9 (2001).
51. Wang, Y. et al. Medicarpin protects cerebral microvascular endothelial cells against oxygen-glucose deprivation/reoxygenation-induced injury via the PI3K/Akt/FoxO pathway: a study of network Pharmacology analysis and experimental validation. *Neurochem. Res.* **47**, 347–357 (2022).
52. Cheriet, T., Ben-Bachir, B., Thamri, O., Seghiri, R. & Mancini, I. Isolation and biological properties of the natural flavonoids Pectolinarin and Pectolinarigenin—A review. *Antibiotics* **9**, 417. https://doi.org/10.3390/antibiotics9070417 (2020).
53. Pandey, M. K., Pandey, R., Singh, V. P. & Pandey, V. B. Singh, U. P. Antifungal activity of 4′,7-Dimethoxyisoflavone against some fungi. *Mycobiology* **30**, 55–56 (2002).
54. Seelinger, G., Merfort, I. & Schempp, C. M. Anti-oxidant, anti-inflammatory and anti-allergic activities of Luteolin. *Planta Med.* **74**, 1667–1677 (2008).
55. Su, Z. et al. Isoliquiritin treatment of osteoporosis by promoting osteogenic differentiation and autophagy of bone marrow mesenchymal stem cells. *Phytother Res.* **38**, 214–230 (2024).
56. Tyagi, A. M. et al. Medicarpin inhibits osteoclastogenesis and has nonestrogenic bone conserving effect in ovariectomized mice. *Mol. Cell. Endocrinol.* **325**, 101–109 (2010).
57. Tan, B. X. et al. Bioactive triterpenoids from the leaves of eriobotrya Japonica as the natural PDE4 inhibitors. *Nat. Prod. Res.* **31**, 2836–2841 (2017).
58. Run-Jia, X. et al. 3′-Methoxydaidzein exerts analgesic activity by inhibiting voltage-gated sodium channels. *Chin. J. Nat. Med.* **17**, 413–423 (2019).
59. Tan, H., Sonam, T. & Shimizu, K. The potential of triterpenoids from Loquat leaves (Eriobotrya japonica) for prevention and treatment of skin disorder. *Int. J. Mol. Sci.* **18**, 1030 (2017).
60. Chen, L. et al. Comparative metabolic study of two contrasting Chinese cabbage genotypes under mild and severe drought stress. *Int. J. Mol. Sci.* **23**, 5947. https://doi.org/10.3390/ijms23115947 (2022).
61. Stasolla, C., Katahira, R., Thorpe, T. A. & Ashihara, H. Purine and pyrimidine nucleotide metabolism in higher plants. *J. Plant Physiol.* **160**, 1271–1295. https://doi.org/10.1078/0176-1617-01169 (2003).
62. Yao, Y. et al. Nitrogen fixation capacity and metabolite responses to phosphorus in soybean nodules. *Symbiosis* **88**, 21–35. https://doi.org/10.1007/s13199-022-00882-9 (2022).
63. Yao, P. et al. Assessment of the combined vulnerability to droughts and heatwaves in Shandong Province in summer from 2000 to 2018. *Environ. Monit. Assess.* **196**, 464. https://doi.org/10.1007/s10661-024-12637-8 (2024).
64. Kassem, M. A. in *In Soybean Seed Composition: Protein, Oil, Fatty Acids, Amino Acids, Sugars, Mineral Nutrients, Tocopherols, and Isoflavones*. 497–511 (eds Kassem, M. A.) (Springer International Publishing, 2021).
65. Caldwell, C. R., Britz, S. J. & Mirecki, R. M. Effect of temperature, elevated carbon dioxide, and drought during seed development on the isoflavone content of Dwarf soybean [Glycine max (L.) Merrill] grown in controlled environments. *J. Agric. Food Chem.* **53**, 1125–1129. https://doi.org/10.1021/jf0355351 (2005).
66. Huang, X. et al. Integrated metabolomic and transcriptomic analysis of specialized metabolites and isoflavonoid biosynthesis in sophora alopecuroides L. under different degrees of drought stress. *Ind. Crops Prod.* **197**, 116595. https://doi.org/10.1016/j.indcrop.2023.116595 (2023).
67. Ragin, B. et al. Effect of row spacing on seed isoflavone contents in soybean [Glycine max (L.) Merr]. *Am. J. Plant. Sci.* **5**, 4003 (2014).
68. Tian, J., Pang, Y., Zhao, Z. & Drought Salinity, and low nitrogen differentially affect the growth and nitrogen metabolism of sophora Japonica (L.) in a Semi-Hydroponic phenotyping platform. *Front. Plant Sci.* **12** https://doi.org/10.3389/fpls.2021.715456 (2021).
69. Zhou, F., Wang, J. & Yang, N. Growth responses, antioxidant enzyme activities and lead accumulation of sophora Japonica and Platycladus orientalis seedlings under Pb and water stress. *Plant. Growth Regul.* **75**, 383–389. https://doi.org/10.1007/s10725-014-9927-7 (2015).
70. Abd-Alla, H. I., Souguir, D. & Radwan, M. O. Genus sophora: a comprehensive review on secondary chemical metabolites and their biological aspects from past achievements to future perspectives. *Arch. Pharm. Res.* **44**, 903–986. https://doi.org/10.1007/s12272-021-01354-2 (2021).
71. Liu, C. W. & Murray, J. D. The role of flavonoids in nodulation Host-Range specificity: an update. *Plants* **5**, 33. https://doi.org/10.3390/plants5030033 (2016).
72. Du, H., Huang, Y. & Tang, Y. Genetic and metabolic engineering of isoflavonoid biosynthesis. *Appl. Microbiol. Biotechnol.* **86**, 1293–1312 (2010).
73. Kim, J. M., Lee, J. W., Seo, J. S., Ha, B. K. & Kwon, S. J. Differentially expressed genes related to isoflavone biosynthesis in a soybean mutant revealed by a comparative transcriptomic analysis. *Plants* **13**, 584. https://doi.org/10.3390/plants13050584 (2024).
74. Wu, M. et al. Full-Length transcriptome sequencing and comparative transcriptomic analyses provide comprehensive insight into molecular mechanisms of flavonoid metabolites biosynthesis in styphnolobium Japonicum. *Genes* **15** https://doi.org/10.3390/genes15030329 (2024).

Acknowledgements

This work was supported by Xinglin Scholar Research Promotion Project of Chengdu University of TCM

(XCZX2022014), and Sichuan Huaijin Agricultural Development Company.The authors acknowledge Mogo Internet Technology Co., LTD. for professional English language editing.

Author contributions

L.Z.: Writing – original draft, Data curation, Formal analysis, Investigation, Writing – review & editing, Project administration. R. M.: Formal analysis, Investigation, Writing – original draft, Writing – review & editing. X. L.: Formal analysis, Investigation, Writing – original draft, Writing – review & editing, review & editing. X. M.: Writing – review & editing, Supervision, Project administration. Y. Z.: Project administration. D. C.: Resources, Writing – review & editing, Supervision, Project administration, Funding acquisition. W. W.: Validation, Writing – review & editing, Supervision.

Declarations

Competing interests

The authors declare no competing interests.

Plant experiment legality statement

This plant experiment complied with relevant institutional, national, and international guidelines and legislation.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-025-26305-5.

Correspondence and requests for materials should be addressed to D.C. or W.W.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.