BMC Plant Biology

RESEARCH Open Access

Integrated metabolomics and transcriptomics reveal differences in terpenoids and the molecular basis among the roots of three *Bupleurum* species

Dong Wen[1], Hongliang Ji[1], Mei Rong[1], Yang Liu[1], Jiemei Jiang[1], Xinwei Guo[1], Zhihui Gao[1], Yanhong Xu[1*] and Jianhe Wei[1,2*]

Abstract

Background *Radix Bupleuri* is a popular traditional Chinese medicinal plant. Its root contains saikosaponin and volatile oil compounds with antipyretic, anti-inflammatory, and hepatoprotective pharmacological effects. However, there are differences in the content and type of main chemical components in the roots of three *Bupleurum* species: *Bupleurum chinense* DC. (Bchi), *Bupleurum scorzonerifolium* Willd. (Bsco), and *Bupleurum marginatum* var. *stenophyllum* (Wolff) Shan et Y.Li (Bmar). The molecular mechanism behind these differences is still unclear. The present study used integrated metabolome and transcriptome analyses to uncover the differences in metabolites and expressed genes among the three *Bupleurum* species.

Results Metabolomics results revealed that Bmar contained more saikosaponins than Bchi and Bsco. Conversely, Bsco had the highest content of volatile oil monoterpenes but a lower sesquiterpene content than Bchi and Bmar. Transcriptome analysis showed that several genes were highly expressed in Bchi, Bsco, or Bmar, demonstrating the molecular mechanism responsible for the differences in their metabolic components. We combined the metabolomics and transcriptomics data to investigate the relationship between metabolites and genes. The results showed a high correlation between CYP450, UGT, and β-AS genes and 6″-acetyl-saikosaponins A, saikosaponins B1, C, and D. The subcellular localization of the two P450 genes (*Bc087391* and *Bc036879*) in the endoplasmic reticulum suggests that they may be involved in saikosaponin biosynthesis.

Conclusion We performed an integrated transcriptome and metabolome analysis to investigate the diversity of the terpenoid biosynthetic pathway in three *Bupleurum* species. The study provides new insights into the molecular basis of the metabolic differences between the three *Bupleurum* species. It also serves as a theoretical basis for the clinical application and breeding of *Bupleurum* resources.

Keywords *Bupleurum*, Metabolome, Transcriptome, Terpenoids biosynthesis

*Correspondence:
Yanhong Xu
xuyanhong99@163.com
Jianhe Wei
wjianh@263.net

Full list of author information is available at the end of the article

Introduction

Bupleurum, a perennial herb, is widely distributed in China, Russia, Mongolia, Korea, and Japan. It was initially documented in the Shennong Traditional Herbal Scriptures and has been used medicinally in China for over 2000 years [1, 2]. The dried roots are used as herbal medicine worldwide, with several species officially included in Chinese, Japanese, Korean, British, and European pharmacopoeias [3]. It can disperse and reduce fever, relieve liver depression, and lift yang-qi. It mainly cures colds and fevers, cold and heat exchange, thoracolumbar pain, irregular menstruation, and uterine and rectal prolapse [4–6]. Several chemical constituents of the *Bupleurum* plant have been identified, including triterpenoids (saikosaponins), flavones, coumarins, lignans, monoterpenes and sesquiterpenes (volatile oils) [7]. Among them, saikosaponins (SSs) and volatile oils are the primary bioactive components [3]. SSs mainly include saikosaponin A (SSa), saikosaponin C (SSc), and saikosaponin D (SSd). More than 100 SSs monomers have already been isolated. Although the *Bupleurum* plant contains the above-listed compounds, their types, contents, and pharmacological effects vary significantly between species.

Pharmacological research has demonstrated the anti-inflammatory, hepatoprotective, and immunomodulatory activities of SSs [8]. Volatile oil compounds have antipyretic and anti-inflammatory effects [9, 10]. While 3-carene and β-pinene (monoterpene) have anti-inflammatory effects [11], the active ingredient responsible for the antipyretic activity in volatile oil remains unknown. The Chinese Pharmacopoeia 2020 lists Bchi and Bsco as the two authentic sources of *Bupleurum*, while Bmar is omitted. However, Bmar has higher production and SSs content and is frequently used as a market substitute for Bchi and Bsco, resulting in increased cultivation and research. A comparison of the pharmacological effects of three types of *Bupleurum* showed that Bsco has a stronger antipyretic effect than Bchi, and Bchi has a greater anti-inflammatory effect than Bsco. The hepatoprotective effect of Bchi saponin is stronger than that of Bsco, while there is no significant difference in the hepatoprotective effect between Bmar and Bchi [12]. The chemical composition of medicinal plants provides the material basis for their efficacy. The pharmacological effects of the three kinds of *Bupleurum* vary according to their chemical composition, content, and type. The SSs content of Bmar is higher than that of Bchi and Bsco. Conversely, Bsco has the highest volatile oil content, far exceeding other varieties. Compared to Bchi, Bmar has a higher volatile oil yield [13–17]. Various factors, including the environment, agronomic measures, and germplasm, regulate the accumulation of effective components in secondary metabolites. However, a systematic assessment of the differences in the components of the three *Bupleurum* species and the specific reasons for these discrepancies is not known. Analyzing the molecular mechanism underlying the differences among their metabolic profiles can provide a theoretical framework for the clinical application of *Bupleurum*.

The biosynthesis process for terpenoids, the primary active components in *Bupleurum*, is well understood. The synthesis of SSs occurs through the mevalonate (MVA) pathway. Cytochrome P450s (CYP450s) and uridine diphosphate glycosyltransferases (UGTs) oxidize and glycosylate β-amyrin to create various forms of SSs [18–20]. Sesquiterpenoids are synthesized by sesquiterpene synthase using farnesyl diphosphate (FPP) as a substrate through the MVA pathway. Monoterpenoids can be generated via MVA and the methylerythritol phosphate pathway (MEP) using geranyl diphosphate (GPP) as the substrate and catalyzed by monoterpene synthase (MPS) [21, 22]. Studies using transcriptome analysis have examined the expression of genes related to terpenoid metabolism in *B. chinense*, focusing on the SSs biosynthesis pathway [23, 24]. A study by Qu et al. compared the differences in metabolites between the roots, stems, leaves, and flowers of Bchi and Bsco [17]. Another group investigated the metabolic composition and expressed genes involved in saikosaponin biosynthesis pathways in four organs (the root, flower, stem, and leaf) of Bchi to discover genes associated with saikosaponin biosynthesis [25]. Wan et al. also analyzed the metabolites and differentially expressed genes (DEGs) of two varieties of *Bupleurum* through transcriptome and metabolome, focusing on flavonoids and their regulatory genes [26]. The current emphasis of *Bupleurum* transcriptome and metabolome studies is to identify differential genes in the terpenoid and flavonoid synthesis pathways, and chemical composition differences between various types of *Bupleurum*. There has been little research into the correlation between different compounds and specific genes in distinct *Bupleurum* types. Thus, elucidating the molecular mechanism behind the variations in chemical composition among these three kinds of *Bupleurum* will help raise its SSs content and quality.

Our research group has previously discovered that the peak of SSa and SSd levels in *Bupleurum* roots occurs during the flowering stage. This stage is critical for the accumulation of dry matter and the improvement of root quality [27]. Therefore, the ideal time to sample the roots of a two-year-old *Bupleurum* for transcriptome and metabolomic analysis is during the flowering stage. We determined the reasons for synthesizing differential compounds in the three kinds of *Bupleurum* by combining transcriptome and metabolome analyses. In addition, we screened some candidate genes involved in SSs production. Our research will lay the groundwork for using

Bupleurum in Chinese medicine and can help guide future directional breeding of high SSs in *Bupleurum*.

Materials and methods

Plant material

The three *Bupleurum* species (Bchi, Bsco, and Bmar) were planted in April 2020 at the Institute of Medicinal Plant Development (IMPLAD), Chinese Academy of Medical Sciences & Peking Union Medical College in Beijing, China (40°03′ N and 116°27′ E). The seeds were sown in a plot measuring 30 m × 18 m with 30 cm intervals between plants. We harvested the roots (two-year-old) during the flowering stage in May 2022, cleaned them with distilled water, and dried them with blotting paper (three biological replicates per sample type). These roots were immediately frozen in liquid nitrogen and stored at − 80 °C to facilitate comparative transcriptomics and metabolomics analysis. Professor Jianhe Wei at IMPLAD bred and identified all the plants used in this study.

Sample extraction and UPLC-ESI-QTOF-MS/MS (LC-MS) profiling

Nine samples were ground using liquid nitrogen, each containing approximately 50 mg of material. Subsequently, 1000 µL of methanol: acetonitrile: water (2:2:1 v/v/v) was added, with 2-chloro-L-phenylalanine acting as the internal standard (20 µL, 1 mg/mL). Metabolomic analysis was conducted using the Waters Acquity I-Class PLUS ultra-high performance liquid chromatography-tandem Waters Xevo G2-XS QTOF high-resolution mass spectrometer and a Waters Acquity UPLC HSS T3 column (2.1 mm × 100 mm, 1.8 µm) (Waters, USA). The specific extraction methodology and LC-MS determination method have been documented previously [28].

Sample extraction, GC-MS/MS profiling, and metabolome data analysis

Approximately 500 mg of each sample was weighed and placed in a 20 mL headspace bottle (Agilent, Palo Alto, CA, USA). 1 mL of saturated NaCl solution and 10 µL of 2-octanol internal standard solution (50 µg/mL) were added sequentially. The samples were then extracted using automatic headspace solid-phase microextraction (HS-SPME) for gas chromatography-tandem mass spectrometry (GC-MS) analysis. The identification and quantification of volatile oil were conducted using an Agilent Model 8890 GC and a 7000D mass spectrometer (Agilent), equipped with a 30 m × 0.25 mm × 0.25 µm DB-5MS (5% phenyl-polymethylsiloxane) capillary column (Agilent J&W Scientific, Folsom, CA, USA). The specific extraction process and GC-MS determination method follow the previously described protocols [17].

All metabolite identification and quantification were conducted at Biomarker Technologies Corporation (Beijing, China). The metabolites detected by LC-MS and GC-MS were annotated using the Kyoto Encyclopedia of Genes and Genomes (KEGG) (https://www.kegg.jp/) [29], the Human Metabolome Database (HMDB) (http://www.hmdb.ca/) [30], and the Lipid Maps data base (http://www.lipidmaps.org/) [31]. The differential metabolites of the three *Bupleurum* species were analyzed using principal component analysis (PCA) and orthogonal partial least squares discriminant analysis (OPLS-DA). The filtering criteria for differentially accumulated metabolites (DAMs) were as follows: fold change (FC) ≥ 2, p-value < 0.05, and variable importance in projection (VIP) ≥ 1.

RNA extraction, illumine sequencing, and transcriptome data analysis

Total RNA was isolated from nine samples using TRIzol reagent (Life Technologies, Carlsbad, CA, USA) according to the manufacturer's instructions. RNA concentration was determined spectrophotometrically at A260/A280 using a Nanodrop 2000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). RNA integrity was assessed by the Agilent 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA). cDNA libraries were constructed with a NEBNext® Ultra™ RNA Library Prep Kit for Illumina® (New England Biolabs, Ipswich, MA) following the manufacturer's protocol. The cDNAs were purified using Beckman AMPure XP beads (Beckman Coulter, Brea, CA) and analyzed on an Agilent Bioanalyzer 2100 system (Agilent 2100) to detect the inserted cDNA fragments. The cDNA libraries were sequenced using an Illumina NovaSeq 6000 platform, and the clean data was subjected to sequence assembly to obtain a unigene library. After that, we could proceed with differential expression analyses, gene functional annotation, and functional enrichment. The cDNA library construction and RNA-seq were performed at Biomarker Technologies Corporation (Beijing, China).

Gene functions were annotated using the Nr (NCBI non-redundant protein sequence) [32], Pfam (protein family) (http://pfam.xfam.org/), KOG/COG/eggNOG (Clusters of Orthologous Groups of proteins) [33], Swiss-Prot (manually annotated and reviewed protein sequence) [34], KEGG [27], and GO (Gene Ontology) [35] databases. It was followed by differential expression analysis of the *Bupleurum* species using the DESeq R package (version 1.10.1). DEGs were identified using screening parameters of FC ≥ 2 and false discovery rate (FDR) < 0.01 [36, 37]. GO term enrichment analysis of DEGs was performed using the GOseq R package, followed by KEGG pathway enrichment analysis using KOBAS (2.0) software, which identifies DEG-enriched KEGG pathways [38]. The raw data from the study were deposited at NCBI under the bioproject

accession codes: PRJNA1088560 (biosample accession numbers SRR28365472, SRR28365473, SRR28365474, SRR28365475, SRR28365476, and SRR28365477) and PRJNA1137573 (biosample accession numbers SRR29898470, SRR29898471, and SRR29898472). In the bioproject codes PRJNA1137573, the Latin name of *B. marginatum* was utilized in the uploaded data instead of *B. marginatum* var. *stenophyllum*.

Metabolome-transcriptome correlation analysis

We integrated the metabolome and transcriptome data. DEGs involved in SSs biosynthesis were subjected to correlation analysis with four types of differentially accumulated saikosaponin metabolites (6"-acetyl-saikosaponin A (6"-acetyl-SSa), SSb1, SSc, and SSd). Correlation between the DEGs involved in the MEP pathway and monoterpenes (3-carene and β-pinene) was also analyzed. The Pearson correlation coefficient (PCC) was computed to assess the correlation between the relative content of saikosaponin and transcriptional alterations. Screening criteria included a correlation coefficient (|CC|) of at least 0.80 and a p-value less than 0.05 (Table S11-12). Cytoscape (version 3.9.1) was used to visualize the final interaction network [39].

Quantitative real-time polymerase chain reaction (qRT-PCR) analysis

RNA was extracted from the RNA-Seq analysis samples using an RNA extraction kit (Takara, Dalian, Liaoning, China). Twelve genes involved in saikosaponin synthesis were selected and used for expression analysis. The primers used in this study were designed with Primer Premier 5 (Premier Biosoft, USA) and are listed in Table S13. First-strand cDNA synthesis was performed using 1 μg of total RNA and the PrimeScript™ RT reagent kit (Takara). For qRT-PCR, we used the CFX96 Touch Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA) and the Green qPCR SuperMix UDG kit (TransStart, Beijing, China). The *β-tubulin* gene (GenBank accession number: FJ389750) was used as an internal control gene [40]. The PCR cycling protocol was as follows: 95 °C for 2 min, followed by 40 cycles of 95 °C for 5 s, 60 °C for 10 s, and 72 °C for 10 s. The dissolution curve signals were collected from 60 ℃ to 95 ℃. The relative gene expression levels were calculated using the $2^{-\Delta\Delta CT}$ method. The analysis used three biological and three technical replicates [41].

Subcellular localization

The fusion expression vector of *Bc087391* and *Bc036879* (CYP450 genes) was constructed by seamless cloning. We designed primers with homologous arms based on the *Bc087391* and *Bc036879* gene sequences (Table S13). *B. chinensis* cDNA was utilized as a template to amplify the two genes' full-length coding sequences (CDS). We used the restriction endonuclease enzyme SmaI (Takara) to linearize the PBI121-enhanced green fluorescent protein (eGFP) vector. The target fragments were ligated to the linear vector using the seamless cloning kit (Mei5 Biotechnology, Beijing, China), generating Bc087391-eGFP and Bc036879-eGFP fusion proteins. The two recombinant constructs were introduced into *Agrobacterium tumefaciens* GV3101 (pSoup) strains and suspended in a mixture solution containing 10 mM $MgCl_2$, 10 mM MES, and 100 μM acetosyringone at pH 5.6. After three hours at room temperature, the bacterial solution was injected into the *Nicotiana benthamiana* leaves. After three days, the eGFP signals were visualized with a laser scanning confocal microscope (Leica TCS SMD FCS, Germany) using excitation and emission wavelengths of 484 nm and 507 nm, respectively.

Results

Metabolic profiling

Using LC-MS and GC-MS technologies, we conducted qualitative and quantitative metabolomic analyses of nine samples to assess the variations in metabolite composition between Bchi, Bsco, and Bmar roots. The LC-MS platform discovered 1,567 metabolites, including 359 lipids, 240 organic acids, 170 terpenoids, 87 flavonoids, and other compounds (Fig. 1A, Table S1). Terpenoids were the primary therapeutic compounds identified, including 24 monoterpenoids, 28 sesquiterpenoids, 34 diterpenoids, 45 triterpenoids, and 8 tetraterpenoids (Table S1). The GC-MS platform detected 183 metabolites, including 37 terpenoids, 18 esters, 11 hydrocarbons, and other compounds (Fig. 1B, Table S2). The terpenoids included 13 monoterpenes and 22 sesquiterpenes (Table S2). PCA results showed that the three *Bupleurum* PC1 × PC2 score plots were well separated. In LC-MS detection, PC1 × PC2 was 41.57% and 21.68%, while in GC-MS detection, PC1 × PC2 was 53.83% and 38.00%, indicating good homogeneity and data reliability among bioreplicates (Fig. 1C, D). The OPLS-DA score plots convincingly distinguish between the three *Bupleurum* species, demonstrating satisfactory goodness of fit and statistical significance (Fig. S1). These results indicated that the three *Bupleurum* species had distinct metabolic profiles.

Identifying dams among the Bchi, Bsco, and Bmar

We performed DAM screening using FC ≥ 2, VIP ≥ 1, and a p-value < 0.05 as the criterion to gain a better understanding of the DAMs in three different types of *Bupleurum* roots. The screening results are presented as Venn diagrams (Fig. 2A, B). A comparison of the Bchi and Bsco groups identified 398 DAMs. Of these, 282 were detected by LC-MS (133 upregulated and 149 downregulated) and 116 by GC-MS (53 upregulated and 63 downregulated)

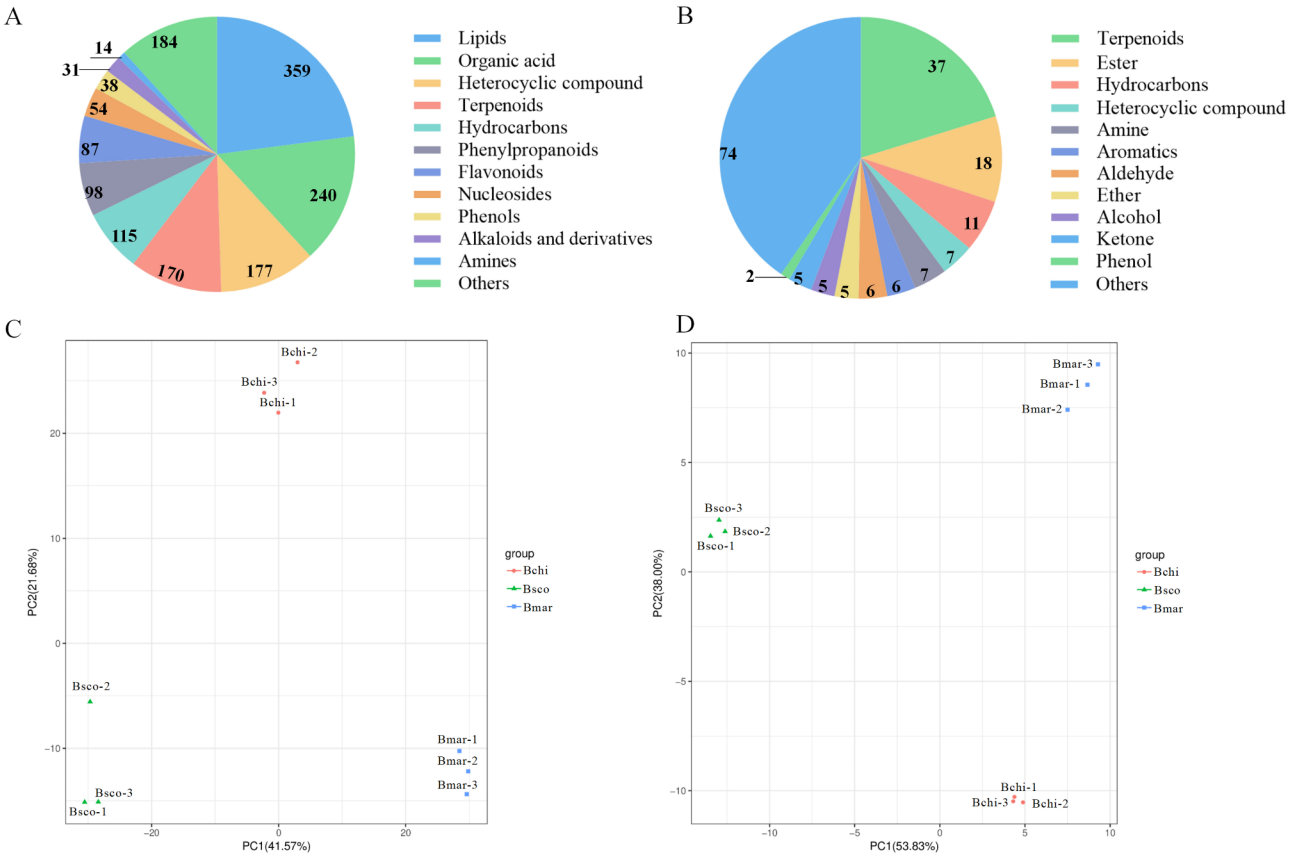

Fig. 1 Preliminary analysis of metabolomics data. Classification pie chart of metabolites detected by LC-MS (**A**) and GC-MS (**B**). PCA score plots were generated for all LC-MS (**C**) and GC-MS (**D**) samples

(Fig. 2C, Table S3). A comparison of the Bchi and Bmar groups revealed 436 DAMs, 319 of which were detected by LC-MS (160 upregulated and 159 downregulated) and the remaining 117 by GC-MS (50 upregulated and 67 downregulated) (Fig. 2C, Table S4). A comparison of the Bsco and Bmar groups revealed 652 DAMs. LC-MS detected 528 DAMs (247 upregulated and 281 downregulated), while GC-MS identified 124 DAMs (63 upregulated and 61 downregulated) (Fig. 2C, Table S5).

Terpenoids are the principal pharmacodynamic components of *Bupleurum* and were discovered as the dominant DAMs among the three kinds of *Bupleurum* (Table S3-S5). Further examination of the terpenoids detected by LC-MS revealed that Bsco had the highest relative content of monoterpenes and tetraterpenes. At the same time, in LC-MS, Bmar exhibited a higher relative content of sesquiterpenes, diterpenes, and triterpenes. Bchi occupied an intermediate position between them (Table S6). Heatmap cluster analysis revealed that Bchi had the highest relative SSd content, Bsco had the highest SSc content, and Bmar had the highest levels of 6”-acetyl-SSa and SSb1. Bmar also had the highest total relative SSs content (Figs. 2E and 3, and Table S6). All these results were consistent with previous studies [17, 24]. GC-MS primarily detects volatile oils, with monoterpenes and sesquiterpenes being the most commonly found terpenoid molecules (Table S7). Heatmap cluster analysis showed that nearly all monoterpenes had a higher relative abundance in Bsco compared to Bchi and Bmar, except for (1R)-2,6,6-Trimethylbicyclo[3.1.1]hept-2-ene ((1R)-(+)-α-Pinene), especially 3-carene and β-pinene, which have anti-inflammatory properties. Comparatively, Bsco had lower sesquiterpene content than Bchi and Bmar. Sesquiterpene levels vary between Bchi and Bmar; some compounds are more prevalent in Bmar, while others are in Bchi. However, monoterpene content was slightly higher in Bchi than in Bmar (Fig. 2D, Table S7). Previous studies have shown that 1-dodecanol and decanal have antibacterial, antiviral, tyrosinase inhibition, anxiolytic, and sedative activities [42, 43]. These compounds were highest in Bsco, almost nonexistent in Bmar, and intermediate in Bchi in our study. Moreover, Bsco had a higher aldehydes content than Bchi and Bmar (Table S2).

Thus, Bsco had more volatile oil compounds with definite pharmacological activities than Bchi and Bmar. Saikosaponin and sesquiterpenoid levels were high in Bmar and modest in Bchi. It suggests that Bmar tends to accumulate saikosaponin, whereas Bsco tends to

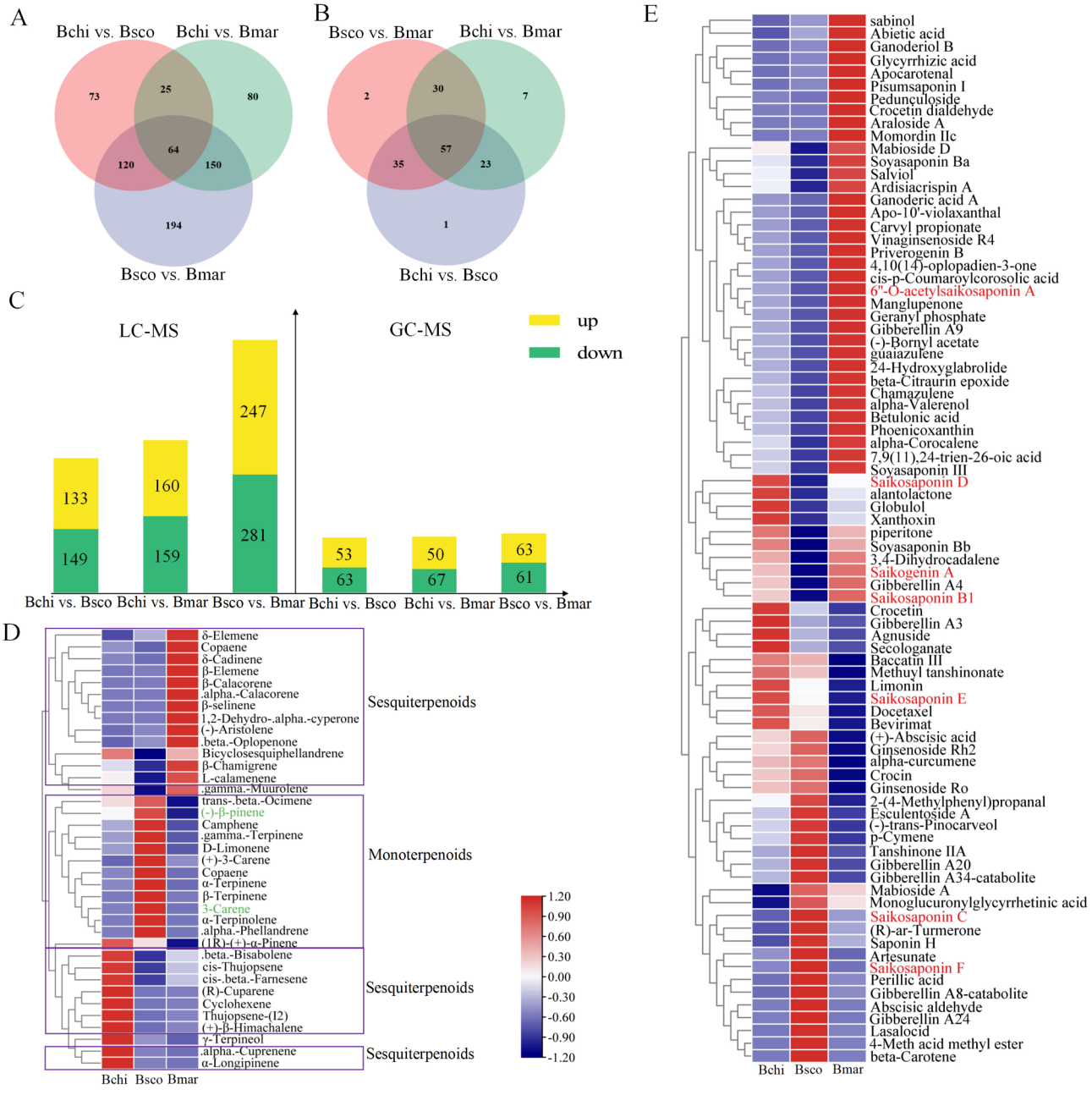

Fig. 2 DAMs analysis. (**A**) Venn diagram of DAMs detected by LC-MS. (**B**) Venn diagram of DAMs detected by GC-MS. (**C**) Statistical analysis of the number of differential metabolites for all different groups in LC-MS and GC-MS detection, with yellow color indicating upregulated metabolites and green color indicating downregulated metabolites. (**D**) Heat map clustering analysis of the different terpenoids identified from three *Bupleurum* species by GC-MS. The relative contents of two monoterpenoids were significantly different in the three *Bupleurum* species, as indicated by the green color. (**E**) Heat map clustering analysis of the different terpenoids identified from three *Bupleurum* species by LC-MS. The SSs are shown in red

accumulate volatile oil compounds. This provides more evidence that distinct *Bupleurum* species have varied clinical indications.

Functional annotation and enrichment analysis of dams

KEGG pathway analysis was performed for DAMs in Bchi vs. Bmar, Bchi vs. Bsco, and Bsco vs. Bmar to identify significantly enriched metabolic pathways. As Fig. 4 illustrates, in Bchi vs. Bsco, the LC-MS detected DAMs were primarily enriched in "carotenoid biosynthesis," "indole alkaloid biosynthesis," and "diterpenoid biosynthesis" (Fig. 4A). In contrast, the GC-MS detected DAMs were enriched in "monoterpenoid biosynthesis" (Fig. 4B). In Bchi vs. Bmar, the LC-MS detected DAMs were mainly enriched in "biosynthesis of unsaturated fatty acid," "terpenoid skeleton biosynthesis," and

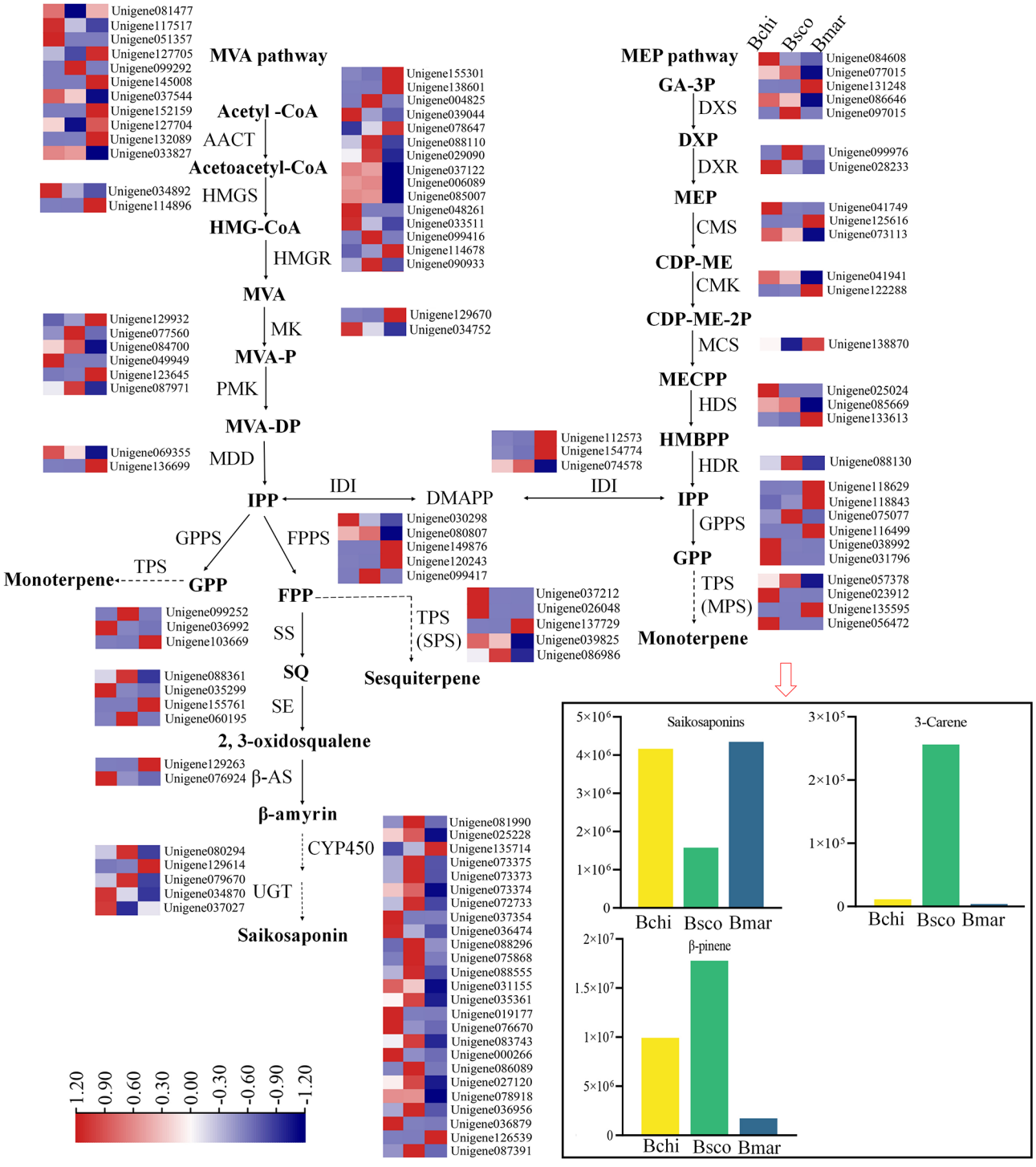

Fig. 3 Pathways and genes involved in terpenoids biosynthesis in *Bupleurum*. The colored cell on each gene's left or right side is the expression heat map of the critical enzyme genes for terpenoid synthesis in three *Bupleurum* species (Bchi, Bsco, and Bmar). The histogram on the right side is the metabolite's relative content in the three *Bupleurum* species

"diterpenoid biosynthesis" (Fig. 4C), and the GC-MS detected DAMs were enriched in "monoterpenoid biosynthesis" (Fig. 4D). Lastly, in Bsco vs. Bmar, the DAMs detected by LC-MS were primarily enriched in the "one carbon pool by folate," "tyrosine metabolism," and "diterpenoid biosynthesis" (Fig. 4E), while those detected by GC-MS were enriched in the "monoterpenoid biosynthesis" (Fig. 4F). Certain metabolic processes, such as diterpenoid and monoterpenoid biosynthesis, are shared across these comparative groups. These findings indicate

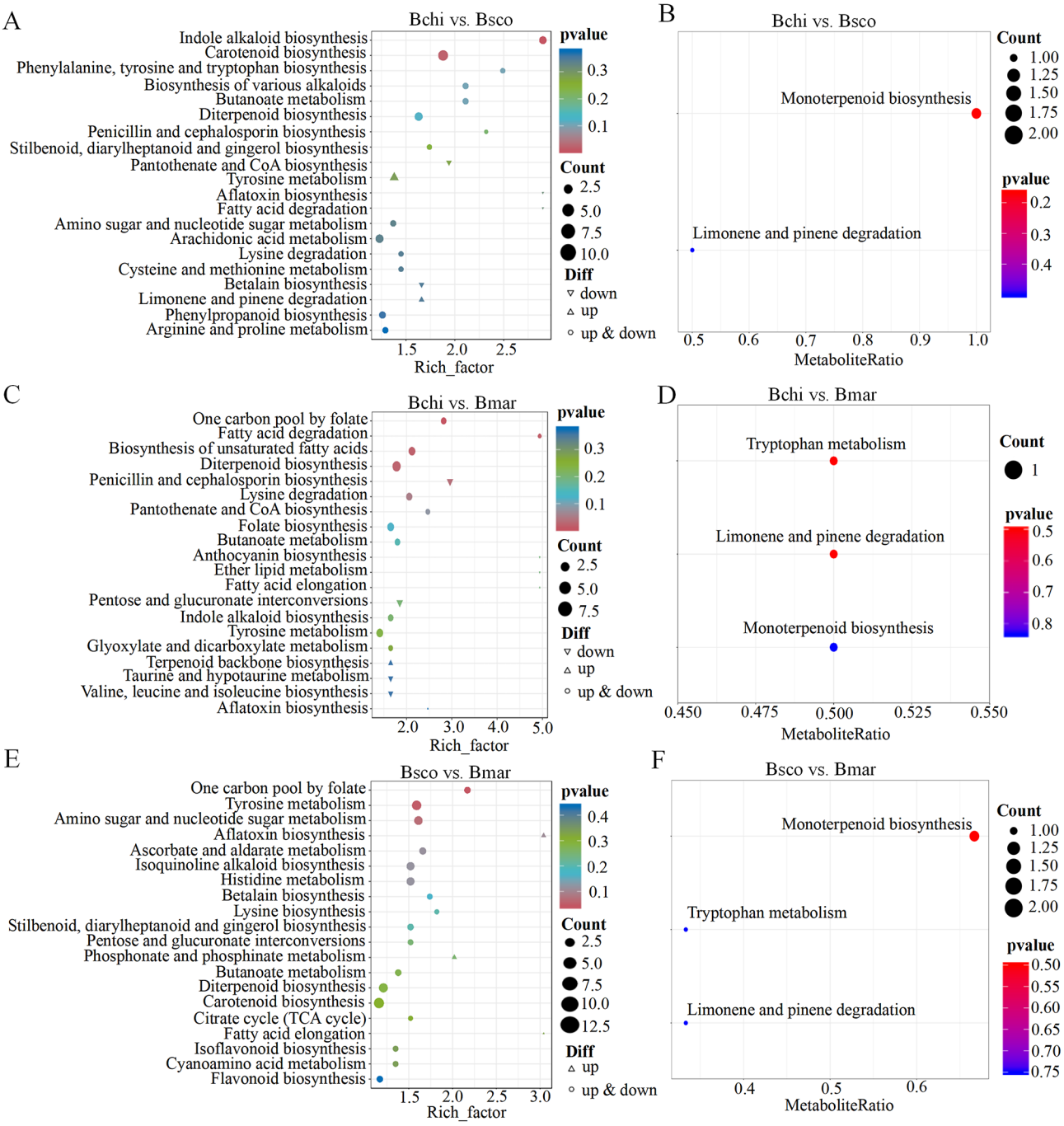

Fig. 4 Differential metabolite pathway bubble plots. (**A-B**) KEGG pathway analysis of Bchi vs. Bsco DAMs detected by LC-MS and GC-MS, respectively. (**C-D**) Bchi vs. Bmar. (**E-F**) Bsco vs. Bmar

that the DAMs of the three *Bupleurum* species were mainly enriched in the terpenoid biosynthesis pathway. We can thus infer that variations in terpenoids' biosynthesis account for the difference in the functions of the three *Bupleurum* species.

Transcriptome analysis of the roots of three *Bupleurum*

Nine samples from three *Bupleurum* roots were sequenced, yielding 56.90 Gb of clean data. Every sample reached the 5.81 Gb of clean data, with a Q30 base percentage above 90.98%. GC content ranged from 42.54 to 42.87% (Table S8). A total of 99,974 unigenes were assembled (Table S9).

Unigenes sequence similarity comparisons were performed using BLASTx (E-value < $1e^{-10}$) against the Nr, Swiss-Prot, and Pfam protein databases. The results revealed that 59,465 transcripts (97.84%), 39,528 transcripts (65.04%), and 43,076 transcripts (70.87%) were significantly similar to the Nr database, Swiss-Prot database, and Pfam database, respectively. The screening criteria for analyzing differential genes in the three groups (Bchi vs. Bsco, Bchi vs. Bmar, and Bsco vs. Bmar) were FC ≥ 2 and FDR < 0.01. As shown in Figs. 5A, 28 and 239 DEGs (13,369 upregulated and 14,870 downregulated), 7,002 DEGs (2,953 upregulated and 4,049 downregulated), and 29,442 DEGs (13,991 upregulated and 15,451 downregulated) were detected in Bchi vs. Bmar, Bchi vs. Bsco, and Bsco vs. Bmar, respectively (Fig. 5A). The transcriptomic analysis results indicated that the metabolic profiles of the three *Bupleurum* species differed due to the combined effect of several DEGs. The Venn diagram (Fig. 5B) illustrates the overlap of DEGs among the three *Bupleurum* species. Bchi vs. Bsco had the least number of DEGs, while Bsco vs. Bmar had the highest number (Fig. 5A, B). Despite the differences in the root differential compounds between Bchi and Bsco, their similar genetic backgrounds might have contributed to the similarity in their gene expression patterns. Therefore, comparing these three *Bupleurum* species can help us identify the factors that cause this variation in compounds.

KEGG enrichment analyses of DEGs were performed to analyze the functions of DEGs in the three types of *Bupleurum* species. DEGs of Bchi vs. Bsco and Bchi vs. Bmar were enriched in sesquiterpenoid and triterpenoid biosynthesis (ko00909), terpenoid backbone biosynthesis (ko00900), monoterpenoid biosynthesis (ko00902), and flavonoid biosynthesis (ko00941) (Fig. 5C, D). This is consistent with the results of metabolome analysis, which further demonstrated the significance of the terpenoid synthesis pathways in the primary metabolic pathways of the three *Bupleurum* roots. Bchi vs. Bmar and Bsco vs. Bmar were significantly enriched for ubiquitin-mediated proteolysis (ko04120) (Fig. 5D, E). As the primary regulator and terminator of gene and protein functions, the ubiquitin-mediated protein degradation pathway controls the selective degradation of most proteins in eukaryotic cells [44]. Thus, we hypothesize significant metabolic variations across the three species of *Bupleurum*, perhaps due to variations in their DEGs related to these metabolic pathways.

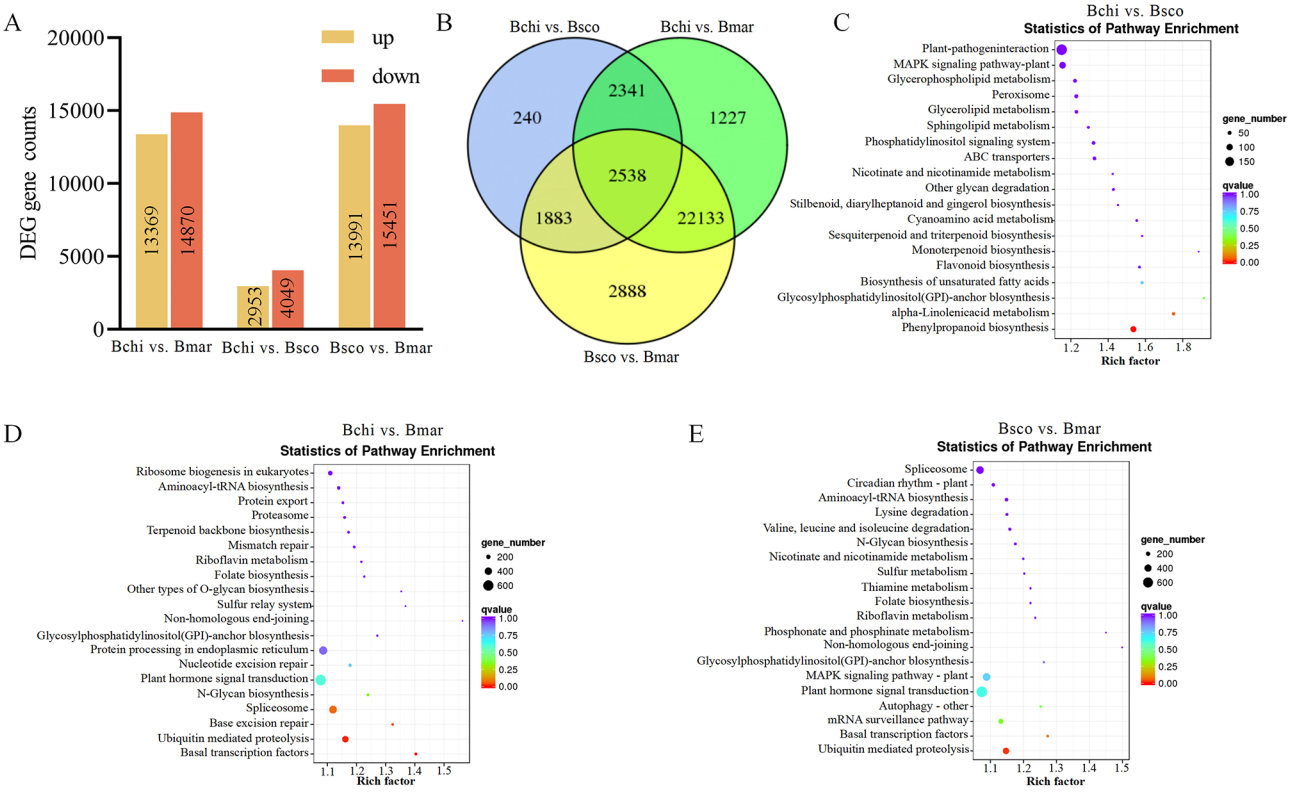

Fig. 5 Transcriptome sequencing analyses of Bchi, Bsco, and Bmar roots. (**A**) DEGs statistics. (**B**) Venn diagram showing the number of DEGs in Bchi vs. Bsco, Bchi vs. Bmar, and Bsco vs. Bmar. (**C**) Top 20 KEGG pathways with the most significant enrichment in Bchi vs. Bsco. (**D**) Top 20 KEGG pathways with the most significant enrichment in Bchi vs. Bmar. (**E**) Top 20 KEGG pathways with the most significant enrichment in Bsco vs. Bmar. The smaller the q-value, the more reliable the enrichment significance of DEGs in this pathway. The size of the circle represents the number of DEGs

Gene expression patterns associated with the terpenoid biosynthesis pathways

We analyzed the gene expression patterns of 119 genes associated with terpenoid biosynthesis pathways in the three *Bupleurum* roots, focusing on the MEP and MVA pathways, to better understand the relationships between the different metabolic networks of the three *Bupleurum* roots (Fig. 3, Table S10). A combined analysis of the metabolomic and transcriptomic data showed substantial expression variations in the genes involved in terpenoid biosynthesis. As shown in Fig. 3, the the β-AS gene (Unigene129263) is expressed significantly higher in Bmar than in Bchi and Bsco. The high expression of the β-AS gene may increase the synthesis of the precursor compound β-amyrin, resulting in a higher relative content of SSs in Bmar (Fig. 3, Table S10). In addition, the HMGS gene (Unigene114896) and the HMGR gene (Unigene155301) had the highest expression in Bmar. The MPS gene (Unigene057378) had a high expression in Bsco, while the FPPS gene (Unigene120243) and SPS gene (Unigene137729) exhibited high expression in Bmar. The KEGG annotations of these genes are terpene-related (monoterpene, terpenoid backbone, and sesquiterpene and triterpene biosynthesis pathways). Consequently, high expression of synthase genes, including HMGS, HMGR, β-AS, and FPPS upstream of the terpenoid pathways, can impact downstream metabolite synthesis.

We further selected 12 genes involved in the terpenoid biosynthesis pathway (AACT, HMGS, HMGR, MK, PMK, MDD, SS, SE, β-AS, CYP450, UGT, and TPS) to confirm the transcriptome results. Except for MDD, most genes exhibited a significant correlation between the qRT-PCR and transcriptomic datasets. The qRT-PCR results showed that Bmar had much higher β-AS genes expression than Bchi and Bsco, in line with our expectations based on metabolome data. However, the qRT-PCR results of the CYP450 and UGT genes showed high CYP450 expression in Bsco and high UGT gene expression in Bmar. This suggests that the effects of modified genes on saikosaponin synthesis may be more complex (Fig. 6). Thus, we hypothesized that the elevated levels of saikosaponin and sesquiterpene in Bmar are triggered by the upregulation of essential enzyme genes in the corresponding pathway, increasing the content of precursor compounds for saikosaponin and sesquiterpene synthesis.

Correlation between terpenoid content and the expression of terpenoid biosynthesis genes

We conducted Pearson correlation analyses on the relative contents of four metabolites (6”-acetyl-SSa, SSb1, SSc, and SSd) with the DEGs (AACT, HMGR, HMGS, β-AS, CYP450, and UGT) involved in the saikosaponin synthesis pathway in *Bupleurum* to better understand the connection between genes and metabolites in the terpenoid biosynthesis pathway in Bchi, Bsco, and Bmar. Correlation analysis with terpenoid skeleton synthesis genes (AACT, HMGR, HMGS, and β-AS) showed that 7 HMGR genes in Bchi correlated with 6”-acetyl-SSa, SSb1, SSc, and SSd. In addition, 2 AACT, 2 HMGR, and 1 β-AS gene in Bsco correlated with 6”-acetyl-SSa, SSc, and SSd. In Bmar, 1 HMGR gene correlated with 6”-acetyl-SSa and SSd, 3 AACT genes and 1 β-AS gene correlated with SSb1 and SSd, and 4 HMGR genes correlated with 6”-acetyl-SSa, SSb1, SSc, and SSd (Table S11). CYP450s and UGTs are primarily involved in the post-modification process of triterpenoid biosynthesis, which is critical in producing different triterpenoids [45]. We identified 25 CYP450 genes in our study based on KEGG annotation results. Bchi possessed 9 CYP450 genes closely linked to SSd, while Bsco had 10 CYP450 genes highly related to SSc. Bmar has 7 CYP450 genes closely associated with SSb1 and 1 CYP450 related to 6”-acetyl-SSa (Fig. 7). Furthermore, we simultaneously found a strong correlation between two P450 genes (*Bc087391* and *Bc036879*) and three saikosaponin compounds (6”-acetyl-SSa, SSb1, and SSd) in Bchi. The KEGG annotation assigned these two genes to triterpenoid and sesquiterpenoid biosynthesis pathways (Table S9). Thus, the synthesis of saikosaponin may be mediated by these two genes.

Concurrently, we examined the relationship between monoterpene compounds (3-carene and β-pinene) and MEP pathway genes. Despite the high relative levels of 3-carene and β-pinene in Bsco, the genes associated with the *Bupleurum* species were essentially the same, with Bsco having the fewest genes (Table S12). This may be because the synthesis of terpenoid compounds is a complex process that involves not only skeleton genes but also other modification genes, which needs further investigation.

Subcellular localization

Almost all CYP450s are membrane-associated proteins located in the endoplasmic reticulum (ER), with only a few localized in chloroplasts and mitochondria [46]. In this study, both CYP450s (*Bc087391* and *Bc036879*) are predicted to localize in the ER. We used confocal microscopy to verify this prediction. The results showed that the control group expressing the 35s-eGFP construct showed fluorescent labeling in the endoplasmic reticulum and nucleoplasm. In tobacco epidermal cells, the fluorescence signals of the fusion proteins Bc087391-eGFP and Bc036879-eGFP were likewise concentrated in the ER, consistent with the predictions of their localization in the ER (Fig. 8).

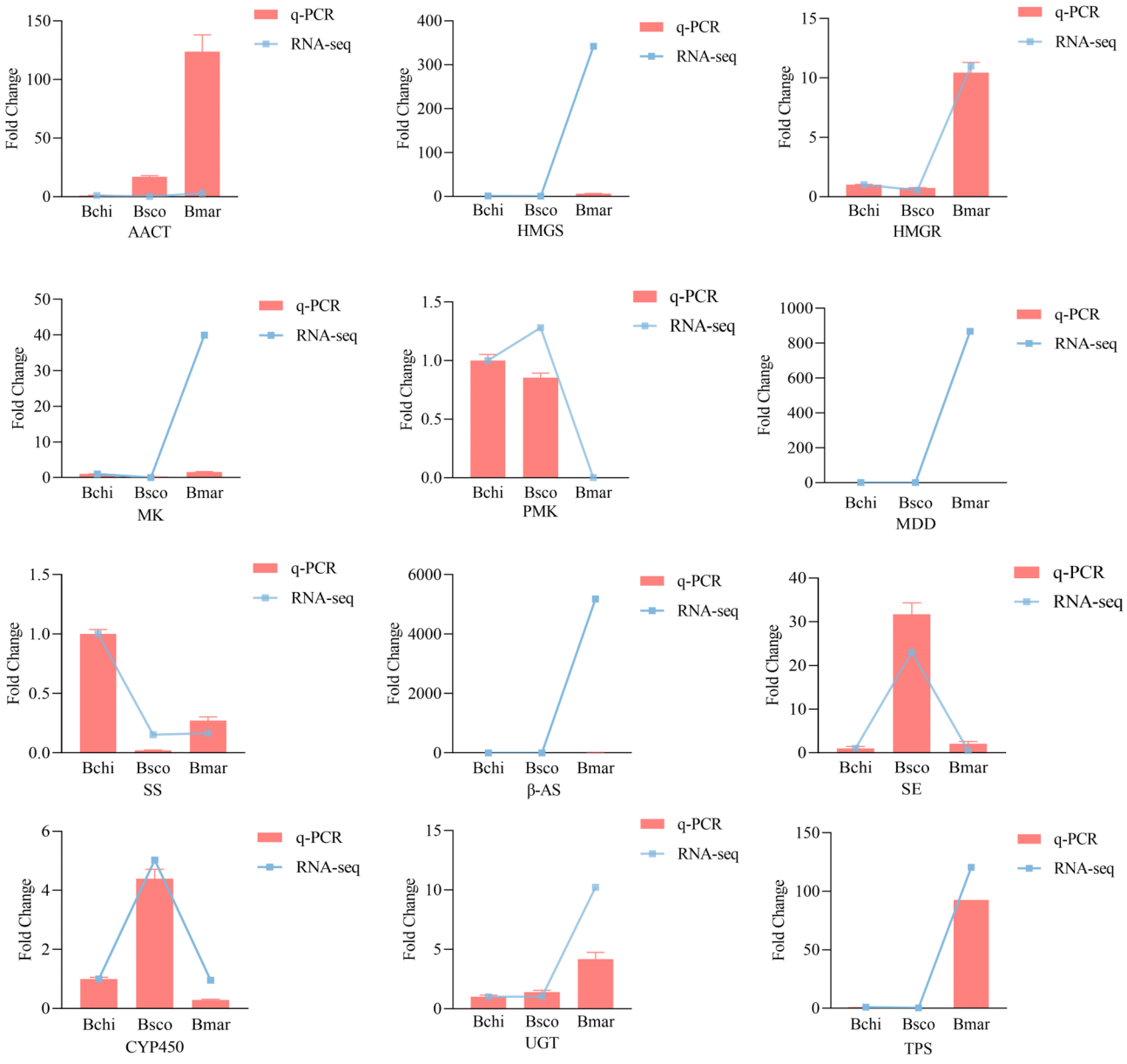

Fig. 6 Verification of saikosaponin biosynthesis-related DEGs by qRT-PCR. The relative expression levels of the candidate genes were calculated using the *β-tubulin* gene as a reference. The error bars represent the standard deviation of three replicates

Discussion

Bupleurum has a long medicinal history, and the dried roots of some varieties are recorded in the pharmacopoeias of many countries. According to the Chinese Pharmacopoeia, Bchi or Bsco roots are the official botanical source of *Bupleurum*. Bmar is not mentioned in the Chinese Pharmacopoeia but is recorded in several local Chinese standards as a superior germplasm because of its high saponin content. Therefore, we conducted a comparative analysis of the transcriptome and metabolome of the three kinds of *Bupleurum* roots to analyze their chemical differences. The study would provide theoretical support for breeding *Bupleurum* varieties with high saponin and volatile oil content and researching the corresponding metabolic pathways.

The metabolome results showed that the three types of *Bupleurum* roots differ significantly in their primary and secondary metabolites (Fig. 2C). The DAMs were predominantly terpenoids, consistent with earlier findings [17, 24]. The results allow us to conclude that variations in germplasm and genetic factors, rather than environmental and cultivation management, account for the differences between the three *Bupleurum* species. SSs are the main active ingredients of *Bupleurum*. We discovered that Bmar had a higher relative content of total SSs than Bchi and Bsco and a higher relative content of

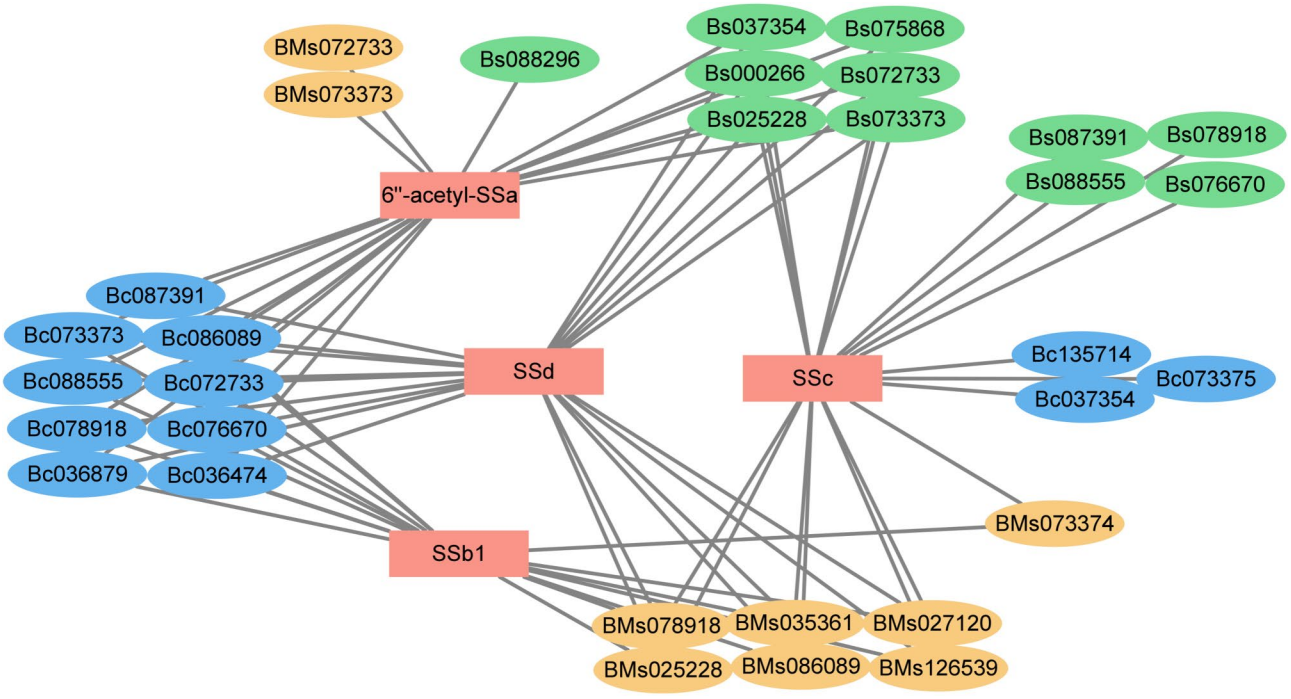

Fig. 7 Connection network between CYP450 genes and SSs (6"-acetyl-SSa, SSb1, SSc, and SSd). Bc, *B. chinense*; Bs, *B. scorzonerifolium*; BMs, *B. marginatum* var. *stenophyllum*

sesquiterpenes than Bsco. Conversely, Bsco had a higher monoterpene content than Bmar and Bchi.

Terpenoids are known to be synthesized by the MVA and MEP pathways [47]. The MVA pathway biosynthesizes triterpenoids and sesquiterpenoids, whereas the MEP pathway biosynthesizes monoterpenoids, diterpenoids, and tetraterpenoids [48]. The expression levels of genes in the terpenoid pathway significantly differed among the three kinds of *Bupleurum* root (Fig. 3). The MVA pathway genes AACT (Unigene127705) and HMGR (Unigene155301) showed the highest expression in Bmar (Fig. 3, Table S10). Similarly, Bmar had high expression levels of the β-AS gene, which encodes a crucial enzyme that catalyzes 2,3-oxidosqualene conversion into the triterpenoid backbone β-amyrin. The AACT and β-AS genes in Bmar are highly correlated with SSb1 and SSd, implying that these two genes may directly regulate the biosynthesis of SSb1 and SSd in Bmar, discovered for the first time in *Bupleurum* species. DXS has a crucial regulatory and rate-limiting role in the MEP pathway during the biosynthesis of terpenes in plastids [49]. Our study identified several upregulated DXS genes in different *Bupleurum* species. The critical enzyme genes of GPPS and FPPS provide GPP and FPP precursors for the biosynthesis of monoterpenes, sesquiterpenes, and triterpenes. MPS and SPS then convert these precursors into monoterpenes and sesquiterpenes [50, 51]. The results revealed that the FPPS gene (Unigene120243) and SPS gene (Unigene137729) showed higher expression in Bmar, whereas the GPPS gene (Unigene075077) showed higher expression in Bsco. However, no highly expressed MPS were found in Bsco. These results indicated that DEGs may be responsible for differences in terpenoids and their relative content, contributing to variances in the chemical composition of the three *Bupleurum* roots.

Multiple enzymes, gene expression regulation, and genetic diversity affect the terpenoid synthesis pathway, resulting in differences in the plant terpenoid content. The only genome information available for *Bupleurum* plants is for *B. chinense*. *Bupleurum* genome data will become more accessible in the future due to the ongoing advancements in genome sequencing. This will offer additional insights into the mechanisms underlying the variations in chemical composition among *Bupleurum* plants and the evolution of terpenoid synthesis genes. Transcription factors (TF) such as WRKY, MYB, bHLH, AP2/ERF, and bZIP could also regulate terpenoid synthesis genes and affect terpenoid content [52].

Terpenoid biosynthesis is a complex process, with the CYP450 gene playing a significant role. CYP450s are multifunctional biocatalysts critical for terpenoid skeleton modification and structural diversity [53]. Currently, *Bupleurum* has few CYP450 genes with precise functions. Moses et al. found CYP716Y1 in *B. falcatum*, which catalyzes the hydroxylation of β-amyrin C-16. However, it only catalyzes hydroxylation in the α-configuration [54]. He et al. discovered two CYP450 genes in Bchi involved in saikosaponin biosynthesis, but their specific

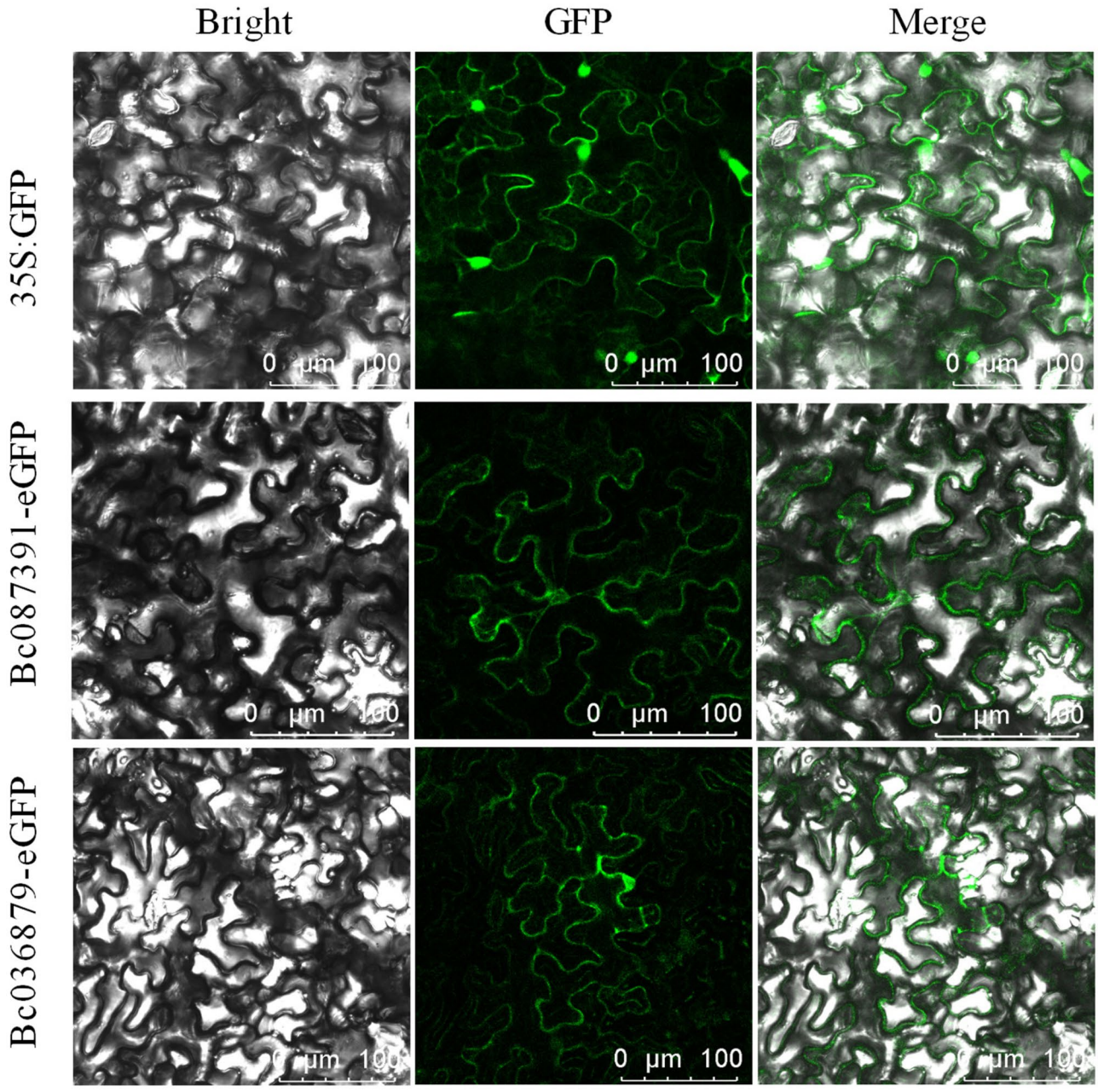

Fig. 8 Nuclear localization of the fusion proteins Bc087391-eGFP and Bc036879-eGFP

roles are unknown [25]. Our results revealed that the genes *Bc087391* and *Bc036879* were significantly associated with 6"-acetyl-SSa, SSb1, and SSd (Fig. 7). The two genes were annotated in the KEGG pathway as sesquiterpenoid and triterpenoid biosynthesis (ko00909). Gene family analysis revealed that both genes belong to the CYP71D subfamily. Members of the CYP71D subfamily are involved in the biosynthesis of indole alkaloids, flavonoids, and terpenoids [55]. Our results on subcellular localization showed that both genes are localized in the endoplasmic reticulum. Therefore, we hypothesized that these two CYP450 family candidate genes may be involved in the biosynthesis of saikosaponin triterpenoids. However, this hypothesis requires further investigation. The study provides a theoretical basis for understanding the molecular mechanisms underlying variances in active ingredient formation among different *Bupleurum* species. It could also help guide the selection and breeding of new *Bupleurum* species.

Conclusions

We compared the metabolome and transcriptome of the three *Bupleurum* roots, focusing on DEGs related to terpenoids, to comprehend each root's different compounds

and metabolic basis. High expression of the β-AS gene (Unigene129263), a key enzyme for β-amyrin production, may contribute to the high SSs content in Bmar. Conversely, high expression of the *TPS* gene (Unigene057378, MPS), a key enzyme in monoterpenoid synthesis, may lead to the high monoterpenoid content, specifically 3-carene and β-pinene in Bsco. In addition, correlation network analysis identified 29 CYPs strongly correlated with the four primary saikosaponin-related compounds (6"-acetyl-SSa, SSb1, SSc, and SSd).

Our results showed that the three *Bupleurum* species differed in their chemical composition. We also identified the critical synthase genes that may help elucidate the genetic basis for these variations. The study provides new insights into the synthesis and accumulation of saikosaponin and volatile oil in *Bupleurum*. Meanwhile, other candidate genes associated with saikosaponin and volatile oil biosynthesis are also being identified. These findings provide a foundation for future studies to elucidate the mechanism of terpenoid synthesis and enable the molecular breeding of *Bupleurum*.

Abbreviations

AACT	Acetyl-CoA acetyltransferase
HMGS	3-hydroxy-3-methylglutaryl-CoA synthase
HMGR	3-hydroxy-3-methy-lglutary-CoA reductase
MK	Mevalonate kinase
PMK	Phosphomevalonate kinase
MDD	Mevalonate diphosphate decarboxylase
IDI	Isopentenyl diphosphate isomerase
SS	Squalene synthase
SE	Squalene epoxidase
β-AS	β-amyrin synthase
CYP450	Cytochrome P450 monooxygenase
UGT	UDP-glycosyltransferase
DXS	1-deoxy-D-xylulose 5-phosphate synthase
DXR	1-deoxy-D-xylulose 5-phosphate synthase
CMS	4-diphosphocytidyl-2-C-methyl-D-erythritol synthase
CMK	4-diphosphocytidyl-2-C-methyl-D-erythritol kinase
MCS	2-C-methyl-D- erythritol 2,4-cyclodiphosphate synthase
HDS	1-hydroxy-2-methyl-2-(E)-butenyl 4- diphosphate synthase
IPK	Isopentenyl monophosphate kinase
FPPS	Farnesyl diphosphate synthase
GPPS	Geranyl diphosphate synthase
TPS	Terpene synthase
SPS	Sesquiterpene synthase
MPS	Monoterpene synthase. Dashed lines indicate unelucidated enzymes
HDR	4-hydroxy-3-methylbut-2-enyl diphosphate reductase

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06441-w.

Supplementary Material 1

Supplementary Material 2

Supplementary Material 3

Supplementary Material 4

Supplementary Material 5

Supplementary Material 6

Supplementary Material 7

Supplementary Material 8

Supplementary Material 9

Supplementary Material 10

Supplementary Material 11

Supplementary Material 12

Supplementary Material 13

Supplementary Material 14

Supplementary Material 15

Acknowledgements
Not applicable.

Author contributions
DW performed the experiments, prepared materials, analyzed the data, and wrote the manuscript. HLJ assisted with material preparation. MR, YL, JMJ, XWG, and ZHG contributed to the data analysis. YHX revised the paper. JHW revised the paper, initiated the project, and supervised it.

Funding
This work was supported by the CAMS Innovation Fund for Medical Sciences (CIFMS): 2021-I2M-1-032 and earmarked funding for CARS (CARS-21).

Data availability
Sequence data that support the findings of this study have been deposited in the National Center for Biotechnology Information with the primary accession code PRJNA1088560 and PRJNA1137573. And the data is under review.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Author details
[1]Key Laboratory of Bioactive Substances and Resources Utilization of Chinese Herbal Medicine, Ministry of Education & National Engineering Laboratory for Breeding of Endangered Medicinal Materials, Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100193, China
[2]Hainan Provincial Key Laboratory of Resources Conservation and Development of Southern Medicine & Key Laboratory of State Administration of Traditional Chinese Medicine for Agarwood Sustainable Utilization, Hainan Branch of the Institute of Medicinal Plant Development, Chinese Academy of Medical Sciences and Peking Union Medical College, Haikou 570311, China

Received: 16 March 2024 / Accepted: 21 March 2025
Published online: 31 March 2025

13 References

1. Deborde C, Jacob D. MeRy-B, a metabolomic database and knowledge base for exploring plant primary metabolism. Methods Mol Biol. 2014;1083:3–16.
2. Pavlidis DE, Mallouchos A, Nychas GJ. Microbiological assessment of aerobically stored horse fillets through predictive microbiology and metabolomic approach. Meat Sci. 2021;172:108323.

3. Ashour ML, Wink M. Genus bupleurum: A review of its phytochemistry, Pharmacology and modes of action. J Pharm Pharmacol. 2011;63:305–21.
4. Moses T, Pollier J, Almagro L, Buyst D, Van Montagu M, Pedreño MA, Martins JC, Thevelein JM, Goossens A. Combinatorial biosynthesis of sapogenins and saponins in Saccharomyces cerevisiae using a C-16α hydroxylase from *Bupleurum falcatum*. PNAS. 2014;111:1634–9.
5. Shin M, Park H, Seo B, Roh S. New approach of medicinal herbs and sulfasalazine mixture on ulcerative colitis induced by dextran sodium sulfate. World J Gastroentero. 2020;26:5272–86.
6. Tian RT, Xie PS, Liu HP. Evaluation of traditional Chinese herbal medicine: Chaihu (Bupleuri Radix) by both high-performance liquid chromatographic and high-performance thin-layer chromatographic fingerprint and chemometric analysis. J Chromatogr A. 2009;1216:2150–5.
7. Jiang H, Yang L, Hou A, Zhang J, Wang S, Man W, Zheng S, Yu H, Wang X, Yang B, et al. Botany, traditional uses, phytochemistry, analytical methods, processing, Pharmacology and pharmacokinetics of bupleuri radix:a systematic review. Biomed Pharmacother. 2020;131:110679.
8. Sun P, Li Y, Wei S, Zhao T, Wang Y, Song C, Xue L, Wang F, Xiao L, Wu J, et al. Pharmacological effects and chemical constituents of bupleurum. Mini-Rev Med Chem. 2019;19:34–55.
9. Lorente I, Ocete MA, Zarzuelo A, Cabo MM, Jimenez J. Bioactivity of the essential oil of *Bupleurum fruticosum*. J Nat Prod. 1989;52:267–72.
10. Martin S, Padilla E, Ocete MA, Galvez J, Jiménez J. Anti-inflammatory activity of the essential oil of *Bupleurum fruticescens*. Planta Med. 1993;59:533–6.
11. Gil ML, Jimenez J, Ocete MA, Zarzuelo A, Cabo MM. Comparative study of different essential oils of bupleurum Gibraltaricum Lamarck. Pharmazie. 1989;44:284–7.
12. Liu XJ, Hu J, Li ZY, Zhang LZ, Qin XM. Pharmacodynamics of bupleurum Chinese DC. and B. Scorzonerifolium willd. (in Chinese). Liaoning J Traditional Chin Med. 2012;39:712–4.
13. Xia ZD, Liu X, Tong LG, Wang H, Feng M, Xi XH, He P, Qin XM. Comparison of chemical constituents of *Bupleurum marginatum* var. *stenophyllum* and *Bupleurum chinense* DC. using UHPLC-Q-TOF-MS based on a metabonomics approach. Biomed Chromatogr. 2021;35:e5133.
14. Zhang SS, Xing J, Li ZY. Metabolic essential oil of different radix bupleuri by GC-MS (in Chinese). Chin J Experimental Traditional Med Formulae. 2014;20:84–7.
15. Zhou YY, Guo CY, Cai CM, Differential. analysis of *Bupleurum chinense* DC., *Bupleurum scorzonerifolium* Willd. and *Bupleurum marginatum* var. *stenophyllum* (in Chinese). Strait Pharmaceutical Journal. 2019;31:59–62.
16. Chen RQ, Wang XH, Xu HT, Zhao RZ, Hu QH. Comparative study on volatile oils among bupleuri radix species and habitats: yields, chemical characterization and antipyretic activities. Chem Biodivers. 2022;19:e202200549.
17. Qu XJ, Hu SQ, Li T, Zhang JQ, Wang BS, Liu CL. Metabolomics analysis reveals the differences between *Bupleurum Chinense* DC. and *Bupleurum scorzonerifolium* willd. Front Plant Sci. 2022;13:13.
18. Abe I. Enzymatic synthesis of Cyclic triterpenes. Nat Prod Rep. 2007;24:1311–31.
19. Lin TY, Chiou CY, Chiou SJ. Putative genes involved in Saikosaponin biosynthesis in *Bupleurum* species. Int J Mol Sci. 2013;14:12806–26.
20. Suzuki H, Achnine L, Xu R, Matsuda SPT, Dixon RA. A genomics approach to the early stages of triterpene saponin biosynthesis in medicago truncatula. Plant J. 2002;32:1033–48.
21. McCaskill D, Croteau R. Prospects for the bioengineering of isoprenoid biosynthesis. Adv Biochem Eng Biotechnol. 1997;55:107–46.
22. Zebec Z, Wilkes J, Jervis AJ, Scrutton NS, Takano E, Breitling R. Towards synthesis of monoterpenes and derivatives using synthetic biology. Curr Opin Chem Biol. 2016;34:37–43.
23. Sui C, Chen M, Xu J, Wei J, Jin Y, Xu Y, Sun J, Gao K, Yang C, Zhang Z, et al. Comparison of root transcriptomes and expressions of genes involved in main medicinal secondary metabolites from *Bupleurum Chinense* and *Bupleurum scorzonerifolium*, the two Chinese official radix bupleuri source species. Physiol Plant. 2015;153:230–42.
24. Yu M, Chen H, Liu SH, Li YC, Sui C, Hou DB, Wei JH. Differential expression of genes involved in Saikosaponin biosynthesis between *Bupleurum Chinense* DC. and *Bupleurum scorzonerifolium* willd. Front Genet. 2020;11:583245.
25. He YL, Chen H, Zhao J, Yang YX, Yang B, Feng L, Zhang Y, Wei P, Hou D, Zhao J, et al. Transcriptome and metabolome analysis to reveal major genes of Saikosaponin biosynthesis in *Bupleurum Chinense*. BMC Genomics. 2021;22:839–52.
26. Wan HF, Zhou L, Wu B, Han WJ, Sui C, Wei JH. Integrated metabolomics and transcriptomics analysis of roots of *Bupleurum Chinense* and *B. scorzonerifolium*, two sources of medicinal Chaihu. Sci Rer. 2022;12:22335–48.
27. Wang H. Effects and mechanisms of stem removal and deftoration on root growth and quality in *Bupleurum chinense* DC. Peking Union Medical College 2020. https://doi.org/10.27648/d.cnki.gzxhu.2020.000297
28. Zhang EY, Zhu XJ, Wang WL, Sun Y, Tian XM, Chen ZY, Mou XS, Zhang YL, Wei YH, Fang ZX, et al. Metabolomics reveals the response of hydroprimed maize to mitigate the impact of soil salinization. Front Plant Sci. 2023;14:1109460.
29. Kanehisa M, Goto S, Kawashima S, Okuno Y, Hattori M. The KEGG resource for Deciphering the genome. Nucleic Acids Res. 2004;32:D277–80.
30. Wishart DS, Feunang YD, Marcu A, Guo AC, Liang K, Vázquez-Fresno R, Sajed T, Johnson D, Li C, Karu N, et al. HMDB 4.0: the human metabolome database for 2018. Nucleic Acids Res. 2018;46:D608–17.
31. Fahy E, Sud M, Cotter D, Subramaniam S. LIPID MAPS online tools for lipid research. Nucleic Acids Res. 2007;35:W606–12.
32. Li W, Jaroszewski L, Godzik A. Tolerating some redundancy significantly speeds up clustering of large protein databases. Bioinformatics. 2002;18:77–82.
33. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, et al. The COG database: an updated version includes eukaryotes. BMC Bioinformatics. 2003;4:1–14.
34. Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its supplement trembl in 2000. Nucleic Acids Res. 2000;28:45–8.
35. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, et al. Gene ontology: tool for the unification of biology. Nat Genet. 2000;25:25–9.
36. Love MI, Huber W, Anders S. Moderated Estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550–71.
37. Varet H, Brillet-Guéguen L, Coppée JY, Dillies MA, SARTools:. A DESeq2- and EdgeR-Based R pipeline for comprehensive differential analysis of RNA-Seq data. PLoS ONE. 2016;11:e157022.
38. Mao XZ, Cai T, Olyarchuk JG, Wei L. Automated genome annotation and pathway identification using the KEGG orthology (KO) as a controlled vocabulary. Bioinformatics. 2005;21:3787–93.
39. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, Amin N, Schwikowski B, Ideker T. Cytoscape: A software environment for integrated models of biomolecular interaction networks. Genome Res. 2003;13:2498–504.
40. Yang LL, Sun Z, Han M. Screening of reference genes in *Bupleurum scorzonerifolium* and tissue expression analysis of key enzyme genes (in Chinese). Chin Traditional Herb Drugs. 2018;49:3651–8.
41. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using Real-Time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods. 2001;25:402–8.
42. Murray AF, Satooka H, Shimizu K, Chavasiri W, Kubo I. Polygonum odoratum essential oil inhibits the activity of mushroom derived tyrosinase. Heliyon. 2019;5:e02817.
43. Togashi N, Shiraishi A, Nishizaka M, Matsuoka K, Endo K, Hamashima H, Inoue Y. Antibacterial activity of long-chain fatty alcohols against Staphylococcus aureus. Molecules. 2007;12:139–48.
44. Ciechanover A, Schwartz AL. The ubiquitin-mediated proteolytic pathway: mechanisms of recognition of the proteolytic substrate and involvement in the degradation of native cellular proteins. FASEB J. 1994;8:182–91.
45. Zheng XY, Li P, Lu X. Research advances in cytochrome P450-catalysed pharmaceutical terpenoid biosynthesis in plants. J Exp Bot. 2019;70:4619–30.
46. Neve EP, Ingelman-Sundberg M. Cytochrome P450 proteins: retention and distribution from the Endoplasmic reticulum. Curr Opin Drug Discov Devel. 2010;13:78–85.
47. Pichersky E, Raguso RA. Why do plants produce so many terpenoid compounds? New Phytol. 2018;220:692–702.
48. Abe I, Rohmer M, Prestwich GD. Enzymatic cyclization of squalene and oxidosqualene to sterols and triterpenes. Chem Rev. 1993;93:2189–206.
49. Lange BM, Wildung MR, McCaskill D, Croteau R. A family of transketolases that directs isoprenoid biosynthesis via a mevalonate-independent pathway. Proc Natl Acad Sci USA. 1998;95:2100–4.
50. Jiang SY, Jin J, Sarojam R, Ramachandran S. A comprehensive survey on the terpene synthase gene family provides new insight into its evolutionary patterns. Genome Biol Evol. 2019;11:2078–98.
51. Jiang H, Wang X. Biosynthesis of monoterpenoid and sesquiterpenoid as natural flavors and fragrances. Biotechnol Adv. 2023;65:108151–73.
52. Li X, Xu Y, Shen S, Yin X, Klee H, Zhang B, Chen K, Hancock R. Transcription factor CitERF71 activates the terpene synthase gene CitTPS16 involved in the synthesis of E-geraniol in sweet orange fruit. J Exp Bot. 2017;68:4929–38.
53. Zheng X, Li P, Lu X. Research advances in cytochrome P450-catalysed pharmaceutical terpenoid biosynthesis in plants. J Exp Bot. 2019;7018:4619–30.

54. Moses T, Pollier J, Almagro L, Buyst D, Van Montagu M, Pedreño MA, Martins JC, Thevelein JM, Goossens A. Combinatorial biosynthesis of sapogenins and saponins in Saccharomyces cerevisiae using a C-16α hydroxylase from *Bupleurum falcatum*. Proc Natl Acad Sci USA. 2014;111:1634–9.
55. Hamberger B, Bak S. Plant P450s as versatile drivers for evolution of species-specific chemical diversity. Philos Trans R Soc Lond B Biol Sci. 2013;368:1612–27.

Publisher's note