



OPEN ACCESS

EDITED BY

Xu Lu,
China Pharmaceutical University, China

REVIEWED BY

Chunjuan Pu,
China Academy of Chinese Medical Sciences,
China
Jagdish Chand,
JSS Academy of Higher Education and
Research, India

*CORRESPONDENCE

Tingyu Shan
✉ ShanTY1610@126.com
Shuangying Gui
✉ guishy0520@ahtcm.edu.cn
Liangping Zha
✉ zlp_ahtcm@126.com

[†]These authors have contributed equally to this work

RECEIVED 30 June 2025

ACCEPTED 09 September 2025

PUBLISHED 30 September 2025

CITATION

Wang H, Yu S, Yu H, Fang Q, Liang J, Zhi N, Peng C, Shan T, Gui S and Zha L (2025) Integrated transcriptome and metabolome provide insight into phenolics and soluble sugar variation in the different varieties of *Gastrodia elata* Blume from different areas in China. *Front. Plant Sci.* 16:1656554. doi: 10.3389/fpls.2025.1656554

COPYRIGHT

© 2025 Wang, Yu, Yu, Fang, Liang, Zhi, Peng, Shan, Gui and Zha. This is an open-access article distributed under the terms of the [Creative Commons Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Integrated transcriptome and metabolome provide insight into phenolics and soluble sugar variation in the different varieties of *Gastrodia elata* Blume from different areas in China

Haixia Wang^{1†}, Shuo Yu^{1†}, Hanwen Yu¹, Qingying Fang¹, Juan Liang¹, Nannan Zhi¹, Chengjun Peng¹, Tingyu Shan^{2*}, Shuangying Gui^{1,3,4,5,6,7*} and Liangping Zha^{1,3,4,8*}

¹College of Pharmacy, Anhui University of Chinese Medicine, Hefei, China, ²College of Traditional Chinese Medicine, Bozhou University, Bozhou, Anhui, China, ³Center for Xin'an Medicine and Modernization of Traditional Chinese Medicine, Anhui University of Chinese Medicine, Hefei, China, ⁴MOE-Anhui Joint Collaborative Innovation Center for Quality Improvement of Anhui Genuine Chinese Medicinal Materials, Anhui University of Chinese Medicine, Hefei, China, ⁵Institute of Pharmaceuticals, Anhui Academy of Chinese Medicine, Hefei, China, ⁶Anhui Province Key Laboratory of Pharmaceutical Preparation Technology and Application, Hefei, China, ⁷Anhui Engineering Research Center for Quality Improvement and Utilization of Genuine Chinese Medicinal Materials, Anhui University of Chinese Medicine, Hefei, China, ⁸Institute of Conservation and Development of Traditional Chinese Medicine Resources, Anhui Academy of Chinese Medicine, Anhui University of Chinese Medicine, Hefei, China

Gastrodia elata Blume is known for its "medicinal food homology", its chemical components include phenols, glycosides and polysaccharides. In China, the Anhui, Hubei, Guizhou, and Yunnan provinces are the primary *G. elata*-producing areas, among which *Gastrodia elata* Bl. f. *elata* (GR), *Gastrodia elata* Bl. f. *glauca* S. Chow (GB), and *Gastrodia elata* Bl. f. *glauca* S. Chow and *Gastrodia elata* Bl. f. *elata* (GR×GB) are essential varieties. This research employed metabolomics and Ultra Performance Liquid Chromatography (UPLC) to quantify 19 soluble sugars and six phenolic compounds in *G. elata* from various origins and varieties. Transcriptome sequencing was performed on 24 *G. elata* samples to identify differentially expressed genes (DEGs). Key enzyme genes involved in the phenolic biosynthesis pathways were further analyzed through phylogenetic analysis, structural modeling and molecular docking. Sucrose, glucose, D-fructose, and D-xylose were identified as the major soluble sugar components in *G. elata*. The phenolic content of the same variety exhibited significant regional differences. In the GR, the largest number of differentially expressed genes (DEGs) was identified in the AH-GR vs. GZ-GR comparison (4,866). In addition, 96 genes encoded 11 key enzymes in the phenolic biosynthesis pathways. In the phenolic synthesis pathway, we identified several alcohol dehydrogenase (ADH) and glucosyltransferase (GT) genes at the downstream level that potentially contribute to the variation in phenolic

metabolism. Phylogenetic, structural modeling, and molecular docking analyses suggested that six GTs catalyze the production of gastrodin from p-hydroxybenzyl alcohol. This study provides a fundamental theoretical basis and data support for the selective breeding of *G. elata* varieties and aids in elucidating the regulatory mechanisms of phenolic active compounds.

KEYWORDS

Gastrodia elata Blume, gastrodin, metabolome, transcriptome, molecular mechanism

Introduction

Gastrodia elata Blume is a highly specialized perennial mycoheterotrophic herbaceous plant belonging to the Orchidaceae family. *G. elata* is characterized by the absence of chlorophyll due to its lack of green leaves and primarily derives its nutrients through a symbiotic relationship with *Armillaria mellea* (Vahl) P. Kumm (Yu et al., 2022). *G. elata* is rich in nutrients and various medicinal ingredients, with more than 210 metabolites having been reported in recent years, including phenolic compounds, polysaccharides, glycosides, organic acids, and sterols (Su et al., 2023). Clinically, the dried tubers of *G. elata* are extensively used for the treatment of convulsions, epilepsy, mental disorders, cerebrovascular diseases, inflammation, amnesia, and metabolic disorders (Huang et al., 2024; Zhang et al., 2023; Zeng et al., 2021). Phenolic substances, which are abundant and diverse in *G. elata*, exhibit significant antioxidant activity and are often considered primary contributors to the plant's pharmacological and therapeutic properties (Heese, 2020). Additionally, polysaccharides derived from *G. elata* have been shown to exert immunomodulatory, antitumor, hypotensive, antibacterial, and free radical-scavenging effects (Zhou et al., 2025; Guan et al., 2022).

G. elata has been approved in China as a functional food (Ma et al., 2021). Artificial cultivation technologies have been developed in China, Japan, South Korea, and India (Guo et al., 2016). According to the Flora of China, there are five primary varieties of *G. elata*, with *Gastrodia elata* Bl. f. *elata* (GR), *Gastrodia elata* Bl. f. *glauca* S. Chow (GB) and *Gastrodia elata* Bl. f. *glauca* S. Chow and *Gastrodia elata* Bl. f. *elata* (GR×GB) being the most commonly cultivated (Wu et al., 2009). Moreover, reliable and stable differentiation markers among the GR, GB, and GR×GB have been successfully identified using liquid chromatography-mass spectrometry (Li et al., 2025). The concentration of primary active compounds in a single variety of *G. elata* is influenced by the cultivation region, which includes the geographical environment, soil conditions, and variations in fungal symbiosis and processing methods (Liu et al., 2024; Kang et al., 2017).

Previous studies have applied transcriptome analysis and metabolomics to investigate several metabolic pathways of secondary metabolites in medicinal plants (Saito, 2018), such as *Platycodon grandiflorus* (Jacq.) A. DC (Yu et al., 2021), *Lycium*

barbarum L. (Ma et al., 2023), and *Taraxacum officinale* (Wiegand et al., 2022), and compare, using transcriptomics, the differences between the vegetative reproductive corms and white corms of *G. elata* (Tsai et al., 2016). The integration of genome and transcriptome sequencing has established a foundation for investigating secondary metabolic pathways, such as gastrodin (Yuan et al., 2018). Current research on the transcriptome of *G. elata* has primarily focused on the influence of fungal symbiosis on its metabolites (Yu et al., 2022; Zheng et al., 2023). The biological components and nutrients of *G. elata* are influenced by environmental factors and cultivar. Research conducted on the metabolic characteristics of three *G. elata* varieties from the same region revealed that parishins are more prevalent in *G. elata* f. *glauca*, whereas gastrodin, gastrol, and 3,4-dihydroxybenzaldehyde are more abundant in *G. elata* f. *elata* and *G. elata* f. *viridis* (Zeng et al., 2023). The polysaccharide properties of four different *G. elata* species have been extensively studied (Ji et al., 2022). However, data on the differences in gene expression levels in the biosynthetic pathways of phenolics among *G. elata* of different producing areas and varieties remain lacking. Therefore, this study sought to integrate metabolomics and transcriptomics to identify phenolic compounds, soluble sugars, and regulatory genes in three commercially available *G. elata* varieties from four major regions of China, namely Anhui, Hubei, Guizhou, and Yunnan provinces. The findings of this study will enhance our understanding of the biosynthesis of phenolic compounds and soluble sugars in these *G. elata* varieties at the metabolite and gene levels and provide novel insights for the cultivation of *G. elata*.

Materials and methods

Plant materials

Two-year-old *Gastrodia elata* Bl. f. *elata* (red stem, GR), *Gastrodia elata* Bl. f. *glauca* S. Chow (black stem, GB), and *Gastrodia elata* Bl. f. *glauca* S. Chow and *Gastrodia elata* Bl. f. *elata* (hybridization of GR and GB, GR×GB) were collected across four Chinese provinces (Anhui, Hubei, Guizhou, and Yunnan), a total of 8 groups, each included three replicates, and each replicate contained one plant. Detailed sampling information is presented in

Supplementary Table S1. All samples were collected by the same individual in November and were identified by Prof. Huasheng Peng (Anhui University of Chinese Medicine). After collection, they were cleaned of surface soil, blotted dry with filter paper, sliced, immediately flash-frozen in liquid nitrogen, and subsequently stored at -80°C for further analysis.

Determination of soluble sugar content

The soluble sugar content and composition were determined using MetWare (Wuhan, China, <http://www.metware.cn/>). A total of 20 mg powder was diluted to 500 μL methanol: isopropanol: water (3:3:2, v/v/v), vortexed for 3 min, and sonicated for 30 min. The extract was centrifuged at 12,000 rpm under 4°C for 3 min. The supernatant (50 μL) was mixed with 20 μL internal standard (1000 $\mu\text{g/mL}$) and evaporated under a nitrogen gas stream. The evaporated samples were transferred to a lyophilizer for freeze-drying. The residue was used for further derivatization. The derivatization method was as follows: the sample was mixed with 100 μL solution of methoxyamine hydrochloride in pyridine (15 mg/mL), incubated at 37°C for 2 h, and 100 μL of N,O-bis(trimethylsilyl) trifluoroacetamide (BSTFA) was added. The mixture was incubated at 37°C for 30 min after vortex-mixing. The mixture was analyzed using gas chromatography-mass spectrometry (GC-MS) following dilution to an appropriate concentration (Gómez-González et al., 2010; Medeiros and Simoneit, 2007; Chen et al., 2016).

The Agilent 8890 gas chromatograph coupled to a 5977B mass spectrometer with a DB-5MS column (30 m length $\times$ 0.25 mm i.d. $\times$ 0.25 μm film thickness; J&W Scientific, USA) was employed for GC-MS analysis of soluble sugars. Helium was used as a carrier gas at a flow rate of 1 mL/min. Injections were made, with an injection volume of 1 μL , in split mode with a split ratio of 5:1. The oven temperature was held at 160°C for 1 min, and then raised to 200°C at $6^{\circ}\text{C}/\text{min}$, 270°C at $10^{\circ}\text{C}/\text{min}$, 300°C at $5^{\circ}\text{C}/\text{min}$, and finally 320°C at $20^{\circ}\text{C}/\text{min}$ and held for 5.5 min. All samples were analyzed in the selective ion monitoring mode. The ion source and transfer line temperatures were 230°C and 280°C , respectively (Medeiros and Simoneit, 2007; Chen et al., 2016). Principal component analysis (PCA) and hierarchical cluster analysis (HCA) were used to analyze the accumulation patterns of the metabolomic data. The HCA results of the samples and metabolites are presented as heat maps accompanied by dendrograms. The fold change (FC) ≥ 2 or fold change (FC) ≤ 0.5 was set as the screening condition for differentially accumulated metabolites (DAMs).

Determination of phenolic content

As previously described, with minor modification, the phenolic content was extracted from dried *G. elata* samples and measured (Chen et al., 2016). Briefly, the dried powder (0.5 g) of each *G. elata* sample was extracted in 25 mL of 50% methanol via ultrasonication for 2 h. The phenols in the extract were chromatographed and detected using

an Acquity-UPLCTM (Waters, MA, USA) system equipped with a quaternary pump, degasser, autosampler, and photodiode array detector (PDA). Phenols were separated on an ACQUITY UPLC[®] T3 column (1.8 μm , 2.1 mm $\times$ 100 mm) (Waters, MA, USA). The mobile phase consisted of water containing 0.05% phosphoric acid (A) and acetonitrile (B) with a flow rate of 0.2 mL/min at 25°C . The linear gradient elution program was as follows: 0–2 min, 5%–7% B; 2–3 min, 7%–10% B; 5–7 min, 20%–24% B; 7–9 min, 15%–18% B; 9–12 min, 18%–20% B; 12–14 min, 80%–76% B; 14–19 min, 24%–27% B; 19–21 min, 27%–15% B; 21–24 min, 15%–5% B. The injection volume was 2 μL , and the chromatograms were acquired at 220 nm.

RNA extraction and illumina sequencing

AH-GR, HB-GR, GZ-GR, AH-GR $\times$ GB, HB-GR $\times$ GB, YN-GR $\times$ GB, GZ-GB and YN-GB were used for RNA-seq analysis. Total RNA was extracted by ethanol precipitation and CTAB-PBIOZOL, and genomic DNA was removed from the total RNA using DNase I. Subsequently, total RNA was identified and quantified using a Qubit fluorescence quantifier and Qsep400 high-throughput biofragment analyzer. The cDNA libraries were constructed using MetWare Biotechnology Co., Ltd. (Wuhan, China). The final library was amplified with phi29 before being uploaded to make DNA nanoballs (DNB) with more than 300 copies of a molecule, and the DNBs were loaded onto the sequencing chip and sequenced on the BGI sequencing platform.

Functional annotation

To obtain high-quality clean data, fastp software was used for strict quality control of the data (Chen et al., 2018). The adapters of reads are removed simultaneously, along with low-quality reads (including those with an N ratio greater than 10% and those with $Q \leq 20$ bases accounting for more than 50% of the entire read). The HISAT2 software was used to compare the clean reads with the reference genome to obtain localization data of the reads on the reference genome (Kim et al., 2015). StringTie was used to assemble the reads and reconstruct the transcriptomes for subsequent analyses (Pertea et al., 2015). Finally, annotation data were obtained by sequence alignment of the genes with the Kyoto Encyclopedia of Genes and Genomes (KEGG), Gene Ontology (GO), Non-Redundant Protein Database (NR), Swiss Protein Database (Swiss-Prot), Translated European Molecular Biology Laboratory Nucleotide Sequence Database (TrEMBL), and EuKaryotic Orthologous Groups (KOG) databases using Diamond, with an E-value of $1e^{-5}$.

Differential expression analysis

The Stringtie maximum flow algorithm and fragments per kilobase of transcript per million fragments mapped (FPKM)

were used to calculate gene expression levels. DESeq2 (Love et al., 2014) was used for differentially expressed genes (DEGs) screening based on the count values of genes in each sample. The false discovery rate (FDR) was obtained by correcting the probability of hypothesis testing (P-value) for multiple hypothesis testing using the Benjamini-Hochberg method. DEGs were screened for Fold Change ≥ 2 and FDR < 0.01 . The total numbers of DEGs, upregulated genes, and downregulated genes in each group were counted and visualized using OmicShare tools (<https://www.omicshare.com/tools/>) and EVenn (<http://www.ehbio.com/test/venn/#/>).

Phylogenetic analysis

Phylogenetic trees were constructed with MEGA 6.0 software using the neighbor-joining method with 1,000 bootstrap replicates. The amino acid sequence of the UGT from *Arabidopsis thaliana* was obtained from the CAZy database (<http://www.cazy.org/>) (Ross et al., 2001). Several UGT proteins that have been confirmed to encode the function of catalyzing the conversion of p-hydroxybenzyl alcohol to gasterodin in *Escherichia coli* and *Saccharomyces cerevisiae* were used to construct phylogenetic trees, namely UGT73B6 (GenBank: AY547304) (Bai et al., 2014, 2016; Zhao et al., 2021), UGT73B6FS (Bai et al., 2016), AsUGT (GenBank: Q9AR73) (Yin et al., 2022), AsUGTsyn (Yin et al., 2022), AtUGT (GenBank: KAG7614719.1) (Xia et al., 2024), CtUGT (GenBank: AJT58578.1) (Xia et al., 2024), PgUGT (GenBank: AIE12483.1) (Xia et al., 2024), and RgUGT (GenBank: KAG5565971.1) (Xia et al., 2024). The constructed phylogenetic tree was visualized and identified using the iTOL tool (<https://itol.embl.de/>).

Sequence alignment, structure modeling, and molecular docking analyses

Multiple sequence alignments were performed using the ESPript online program (<https://espript.ibcp.fr/ESPript/cgi-bin/ESPript.cgi>). A three-dimensional model of GeUGTs was generated using the SWISS-MODEL server and visualized using PyMOL 3.0. Molecular docking was performed using AutoDock 1.5.7 to predict the binding of the glycosyl donor UDPG to the GeUGT protein, thereby forming the GeUGT-UDPG complex. Subsequently, p-hydroxybenzyl alcohol was docked onto the GeUGT-UDPG complex. The protein structure modeling and molecular docking are outlined in Supplementary Tables S2 and S3.

Quantitative real-time polymerase chain reaction analysis

To validate the reliability of the RNA-seq results for determining gene expression, qRT-PCR was performed. Based on

our preliminary studies and literature survey, the β -actin gene was selected as the reference gene (Yuan et al., 2018; Shan et al., 2021). Twelve DEGs involved in phenolic compound and carbohydrate metabolism were selected for qRT-PCR validation. Gene-specific primers were designed using the Primer 5.0 software and synthesized by General Biology Co., Ltd. (Anhui, China; Supplementary Table S4). qRT-PCR analysis was conducted using the BioRad CFX96 Real-Time PCR Detection System (Bio-Rad Laboratories) following the manufacturer's protocol (LABLEAD, China), including optimized cycling conditions and reaction mixture proportions. Relative gene expression levels were quantified using the $2^{-\Delta\Delta Ct}$ method. All experiments were performed in triplicate to ensure data reproducibility.

Data statistical analysis

Microsoft Excel and GraphPad Prism 9.30 software were used to statistically analyze the data. The statistical analysis method used was one-way ANOVA. The TB tools software 2.331, Metware Cloud (<https://cloud.metware.cn>), and OmicShare tools (<https://www.omicshare.com/tools/>) were used for visualization and analysis.

Results

Morphological variations and soluble sugar analysis in *G. elata*

The morphology and appearance of *G. elata* tubers vary considerably across different cultivars. The GR exhibits an elongated cylindrical shape, whereas the GB is characterized by a distinct oval form. The GR $\times$ GB effectively combines the favorable morphological traits of both parental lines, demonstrating a more robust and plump tuber morphology than its parents (Figure 1A). Soluble sugars in *G. elata* tubers from different producing areas and cultivars were analyzed using GC-MS. The total ion overlap diagram chromatograms demonstrated excellent instrument stability (Supplementary Figure S1A). PCA revealed a clear separation among the samples (Supplementary Figure S1B). A total of 19 soluble sugar components were identified (Supplementary Table S5), among which sucrose, glucose, D-fructose, and D-xylose were the major soluble sugar components in *G. elata*, collectively accounting for over 98% of the total soluble sugar content. Sucrose (50.91–80.81 mg/g) was the most abundant across all samples (Supplementary Table S5). Comparatively, the content of glucose, D-fructose, and D-xylose in GR and GB was higher than that in GR $\times$ GB (Figure 1B). DAMs primarily consisted of monosaccharides and disaccharides. Specifically, 3, 5, 3, 1, 3, and 3 DAMs were detected in the comparison groups HB-GR vs. AH-GR, GZ-GR vs. AH-GR, GZ-GR vs. HB-GR, HB-GR $\times$ GB vs. AH-GR $\times$ GB, AH-GR $\times$ GB vs. YN-GR $\times$ GB, and YN-GB vs. GZ-GB, respectively (Supplementary Table S6), and identified a DAM

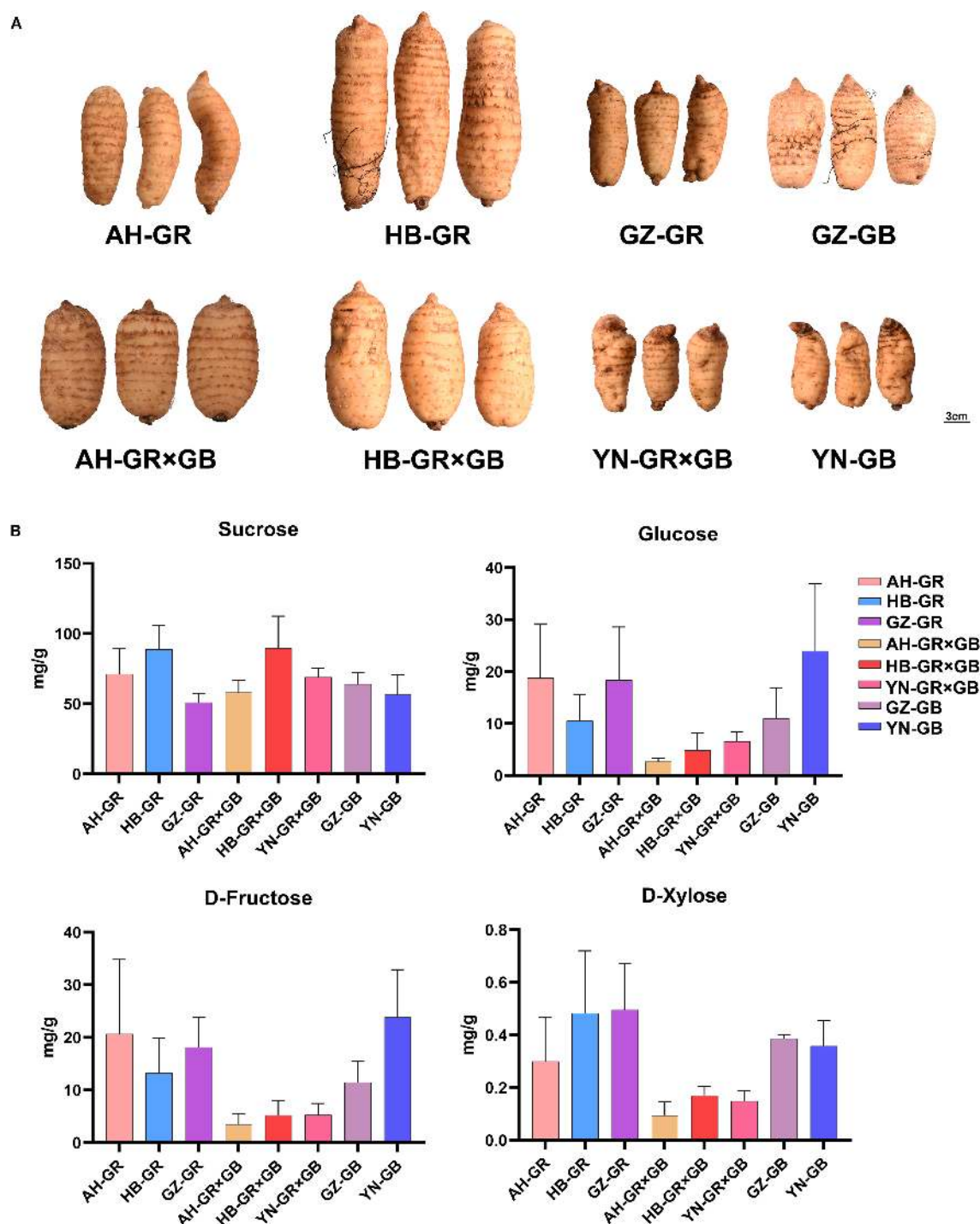


FIGURE 1

Morphological characteristics of three *G. elata* cultivars and the statistical analysis of soluble sugar contents from *G. elata*. (A) The morphological characteristics of *G. elata*. Scale bars = 3 cm. (B) Sucrose, Glucose, D-fructose, and D-xylose contents from different producing areas and varieties of *G. elata*.

(xylitol) in the GR from three regions (Supplementary Figure S2). Additionally, in the comparative groups from the same regions of Anhui, Hubei, Guizhou, and Yunnan, there were 7, 3, 3, and 8 DAMs (Supplementary Table S6), respectively.

Analysis of phenolic metabolites in *G. elata*

Phenolic compounds are the major class of bioactive constituents in *G. elata* (Shan et al., 2021). In all examined

samples, parishin A, parishin B, and parishin E were present at relatively high concentrations, whereas p-hydroxybenzyl alcohol and parishin C were detected at comparatively lower levels (Supplementary Table S7). In the GR group, the gastrodin content in AH-GR and GZ-GR was significantly higher ($p < 0.05$) than that in HB-GR. The p-hydroxybenzyl alcohol content in AH-GR was significantly higher than that in GZ-GR, while the parishin E content in HB-GR was significantly higher ($p < 0.05$) than that in AH-GR. In the GR×GB group (Figure 2), significant differences were observed in the contents of parishin A and parishin E between the HB-GR×GB and YN-GR×GB. In the GB group, the content of p-hydroxybenzyl alcohol in YN-GB accessions was higher than that in GZ-GB (Figure 2). Differences were also observed in the content of specific compounds among different varieties of *G. elata* from the same origin. The gastrodin and p-hydroxybenzyl alcohol content in AH-GR were significantly higher ($p < 0.05$) than those in AH-GR×GB, the gastrodin content in GZ-GR was significantly higher than that in GZ-GB, and the p-hydroxybenzyl alcohol content in YN-GB was significantly higher than that in YN-GR×GB (Figure 2).

Transcriptomic differences reflect regional and varietal influences

A total of 24 cDNA libraries were subjected to transcriptome sequencing and yielded between 43,248,992 and 56,780,560 clean reads; the Q30 and Q20 percentages for all libraries exceeded 92.54% and the GC content percentages were $> 45.43\%$ (Supplementary Table S8). The correlation coefficient between

samples exceeded 0.85 (Supplementary Figure S3). The PCA plot shows that PC1 and PC2 account for 26.4% and 18.1% of the variance, respectively (Supplementary Figure S4). Overall, the results of transcriptome analysis were of high quality and suitable for further research. Of the 23,162 identified genes, 18,415 (79.51%) were successfully annotated using at least one of the following databases: Nr, SwissPort, KOG, Trembl, Pfam, GO, or KEGG (Supplementary Table S9; Supplementary Figure S5). Among the 18,347 genes annotated in the NR database, 87.88% showed homologous sequence matches with eight species of the Orchidaceae family, with the highest homology observed with *Dendrobium catenatum* (Supplementary Figure S6). Among the 15,504 genes annotated with GO terms, the categories were predominantly “cellular process” and “metabolic process” in biological processes, “cellular anatomical entity” in cellular components, and “binding” and “catalytic activity” in the molecular function categories (Supplementary Figure S7). Additionally, 14,161 genes were mapped to 147 pathways using KEGG annotation (Supplementary Figure S8).

The analysis also revealed that the number of DEGs in the three pairwise comparisons within the GR group from different producing areas was substantially higher than that in the GR×GB and GB groups, reaching 3,255, 4,866 and 4,037, respectively. The GR×GB groups contained 1,791, 608 and 2,983 DEGs, while the GB group showed the lowest number, with a total of only 617 (Figure 3A). Furthermore, the number of DEGs in the four comparison groups of different varieties within the same origin was generally low. Particularly, the comparison between HB-GR×GB and HB-GR yielded only common 50 DEGs

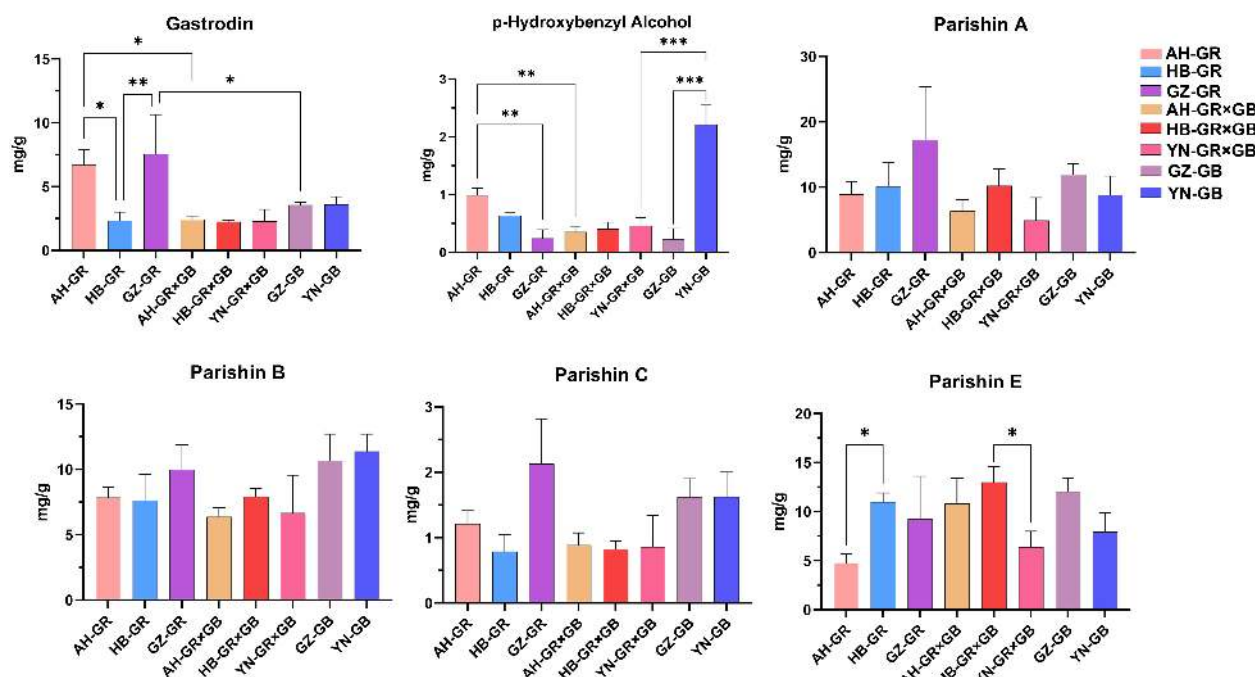


FIGURE 2

Contents of phenolic in *G. elata* cultivars. ‘*’ represents statistical differences in the same content between different cultivars, with $p < 0.05$ being a significant difference (* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). Error bars represent standard deviations.

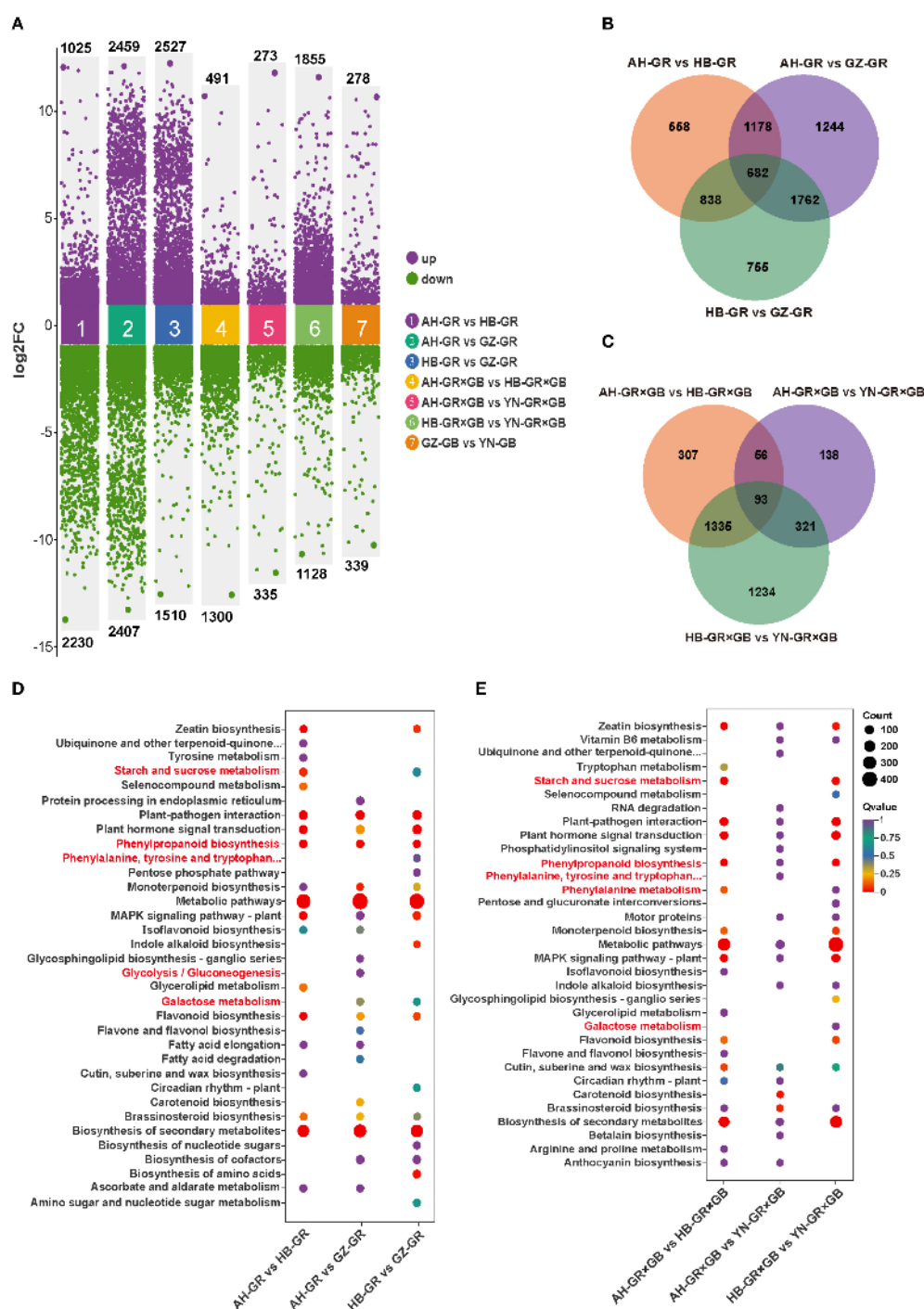


FIGURE 3

Different expression genes (DEGs) of three commercially available *G. elata* varieties from four major regions. (A) DEGs statistics of AH-GR vs. HB-GR, AH-GR vs. GZ-GR, HB-GR vs. GZ-GR, AH-GRxGB vs. HB-GRxGB, AH-GRxGB vs. YN-GRxGB, HB-GRxGB vs. YN-GRxGB, and GZ-GB vs. YN-GB. Purple represents upregulated DEGs, green represents downregulated DEGs. (B) The Venn diagram of DEGs in GR. (C) The Venn diagram of DEGs in GRxGB. (D) The top 20 enriched KEGG pathways of DEGs in AH-GR vs. HB-GR, AH-GR vs. GZ-GR, and HB-GR vs. GZ-GR. (E) The top 20 enriched KEGG pathways of DEGs in AH-GRxGB vs. HB-GRxGB, AH-GRxGB vs. YN-GRxGB, and HB-GRxGB vs. YN-GRxGB. The size of the dots represented the number of DEGs. Red and purple represented high and low expression levels, respectively.

(Supplementary Figure S9). Based on the identified differences from the comparisons, we conducted further analysis on *G. elata* of the same variety from different producing areas. Venn diagrams illustrated 682 common DEGs in the three GR comparison

groups (Figure 3B), and 93 in the three GRxGB comparison groups (Figure 3C). Further analysis of the DEGs across different groups was conducted to investigate their functional categories using the KEGG database. The highest number of DEGs was

identified in the comparison AH-GR_vs._GZ-GR within the GR group (Figure 3D), with an enrichment of 139 KEGG pathways. Conversely, the fewest DEGs were identified in the AH-GR×GB_vs._YN-GR×GB within the GR×GB group (Figure 3E), with 159 DEGs enriched in 93 KEGG pathways. In the GB comparison group, GZ-GB_vs._YN-GB (Supplementary Figure S10) revealed enrichment of 177 DEGs across 102 KEGG pathways. Seven comparison groups demonstrated significant enrichment in the “metabolic pathway,” “secondary metabolite biosynthesis,” and “plant-pathogen interaction” pathways. Notably, pathways related to phenolic metabolism, such as phenylpropanoid biosynthesis, phenylalanine metabolism, and the biosynthesis of phenylalanine, tyrosine, and tryptophan, as well as glucose metabolic pathways, including starch and sucrose metabolism, galactose metabolism, and glycolysis/gluconeogenesis, were enriched.

Unraveling the phenolic biosynthesis in *G. elata* through a multi-omics approach

In our previous study, we postulated the biosynthetic pathway of phenolic compounds in *G. elata* (Shan et al., 2021) (Figure 4B). Phenolic metabolite analysis in *G. elata* revealed significant variations in phenolic content across different producing areas and varieties. Based on these findings, we next investigated the molecular mechanisms underlying the differential accumulation of phenolic compounds in *G. elata*. Gastrodin, p-hydroxybenzaldehyde, and p-hydroxybenzyl alcohol are synthesized from phenylalanine via the phenylpropanoid pathway, whereas the citric acid biosynthesis pathway is implicated in the biosynthesis of parishin, we identified a total of 11 enzymes: Phenylalanine ammonia-lyase (PAL), Cinnamate 4-hydroxylase (C4H), 4-coumarate coenzyme A ligase (4CL), Shikimate O-hydroxycinnamoyltransferase (HCT), 5-O-(4-coumaroyl)-D-quinic acid 3'-monooxygenase (C3H), Caffeoyl-CoA O-methyltransferase (CCoAOMT), Alcohol dehydrogenase (ADH), Glycosyltransferases (GT), Pyruvate dehydrogenase/pyruvate dehydrogenase complex (aceE), Dihydrolipoyllysine-residue acetyltransferase (DLAT) and Citrate synthase (CS) corresponding to 96 genes (Supplementary Table S10). The phenylpropanoid biosynthesis pathway serves as the fundamental route for the synthesis of various phenolic compounds in plants, with its upstream steps primarily responsible for generating universal precursors. For our target metabolites, that is, gastrodin, p-hydroxybenzyl alcohol, and parishin-type compounds, the downstream ADH and GT genes represent more critical targets for investigation. The expression of *ADH1* and *ADH4* was lower in AH-GR than AH-GR×GB, whereas the expression of *ADH5*, *ADH6*, and *ADH11* was higher (Figure 4B), which may be attributed to the higher content of p-hydroxybenzyl alcohol and gastrodin in AH-GR. Correlation clustering analysis between six phenolic compounds and eleven ADH genes indicated a significant positive correlation ($p < 0.05$) of *ADH5* and *ADH11* with the accumulation of p-hydroxybenzyl alcohol and gastrodin (Figure 4C), suggesting that

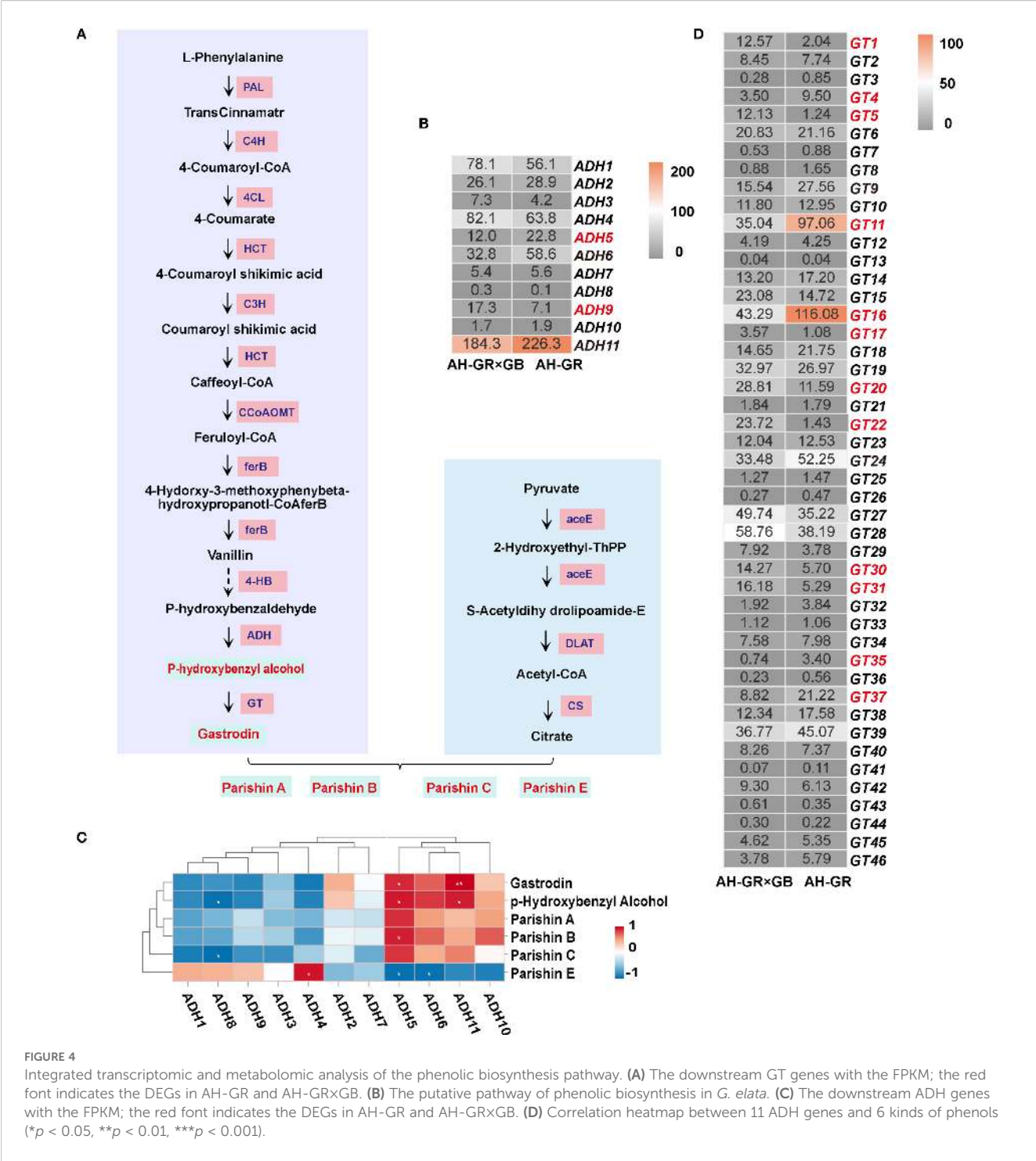
ADH5 and *ADH11* are likely key genes involved in the biosynthesis of p-hydroxybenzyl alcohol. The content of gastrodin and p-hydroxybenzyl alcohol in AH-GR was significantly higher than that in AH-GR×GB (Figure 2). We identified a total of 46 GT genes; there were 12 differentially expressed GTs (*GT1*, *GT4*, *GT5*, *GT11*, *GT16*, *GT17*, *GT20*, *GT22*, *GT30*, *GT31*, *GT35*, and *GT37*) between AH-GR×GB and AH-GR (Figure 4D).

Correlation and phylogenetic analysis of GeUGTs

Correlation clustering analysis between six phenolic compounds and 46 GT genes revealed (Figure 5A) that Group A (*GT3*, *GT38*, *GT11*, *GT16*, *GT24*, *GT7*, *GT39*, *GT9*, *GT25*, *GT41*, *GT45*, *GT4*, *GT8*, *GT35*, *GT37*, *GT46*, and *GT18*) and Cluster I (parishin A, parishin B, parishin C, p-hydroxybenzyl alcohol, and gastrodin), which showed positive correlations ($p < 0.05$), negatively correlated with Cluster II (parishin E). In Group C, *GT19* and *GT1* significantly positively correlated with parishin E ($p < 0.05$). These findings provide key candidate genes and a solid basis for further functional studies on glycosyltransferases in *G. elata*. To gain a comprehensive understanding of the evolutionary relationships and classify the 46 UGT proteins in *G. elata*, a phylogenetic tree including 25 UGT proteins from *Arabidopsis thaliana* (Groups A to N) and 5 from *Zea mays* (Groups O to Q) was constructed using MEGA 7.0. The phylogenetic tree revealed that the UGT proteins in *G. elata* could be classified into 13 groups, with no sequences identified in groups B, H, N, or P (Figure 5B). Among them, group D contained the largest number of UGT proteins of *G. elata*, with a total of 15 members, followed by group E, which contained 9 members. Groups G, L, and A consisted of five, four, and three members, respectively. Groups C, F, K, M, and O each contained only one member. Notably, group D contained two UGT members (CtUGT and UGT73B6FS), whereas group E contained 4 UGT members (AtUGT, AsUGTsyn, PgUGT, and RgUGT), all of which have been shown to catalyze the conversion of p-hydroxybenzyl alcohol to gastrodin *in vitro*. The 24 UGT genes in groups D and E of *G. elata* may also have similar catalytic functions.

Sequence alignment, structure modeling and molecular docking analysis

Based on the phylogenetic analysis, four GeUGTs (*GT16*, *GT17*, *GT18*, and *GT44*) from group E and two GeUGTs (*GT24* and *GT42*) from group D, which showed close evolutionary relationships with functionally characterized UGT genes, were selected for sequence and structural analyses. Multiple sequence alignments revealed that these six GeUGTs shared 48.81% sequence identity with the known functional UGTs. All identified GeUGTs contained a highly conserved 44-amino acid sequence at their C-termini, designated the plant secondary product glycosyltransferase (PSPG) motif, which is characteristic of plant secondary



metabolism glycosyltransferases. Notably, the catalytic residues His (H) and Asp (D) are highly conserved among the UGTs (Supplementary Figure S11). Subsequent tertiary structure modeling demonstrated that all six GeUGT structures were monomeric proteins. In the structural models, the PSPG box is highlighted in red, while the catalytic residues are represented as spheres, with histidine (His) in blue and aspartate (Asp) in yellow (Figure 6).

Molecular docking was used to investigate the interaction between the protein and ligand and predicted the stability and binding affinity of the docked complex (Supplementary Table S11). The molecular docking analysis revealed that p-hydroxybenzyl alcohol binds to the GT16-UDPG complex with a binding energy of -5.043 kcal/mol. Additionally, His48 formed hydrogen bonds with UDPG at a distance of 3.3 Å (Figure 7A). Similarly, p-hydroxybenzyl alcohol exhibited stronger binding affinity (-5.326

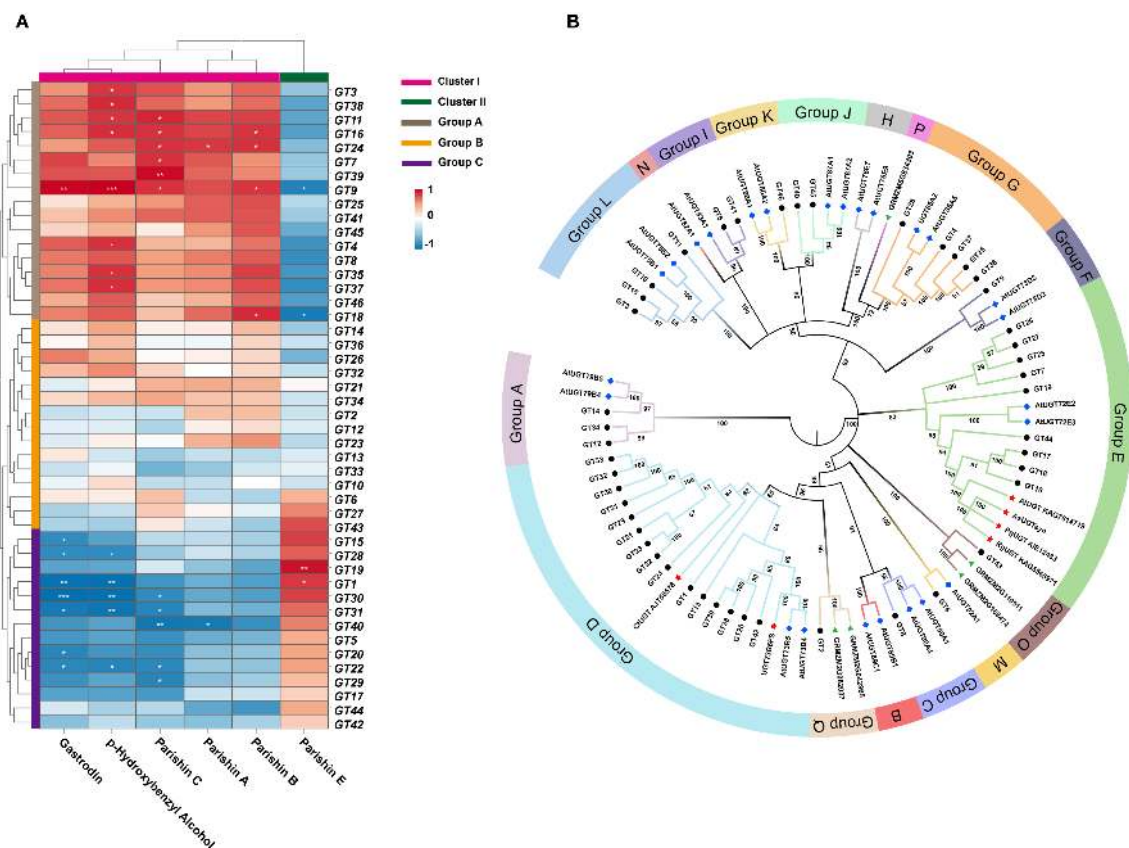


FIGURE 5 (A) Correlation clustering analysis between 6 phenolic compounds and 46 GT genes ($p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$). (B) Phylogenetic tree of the UGT family members. Black circular dots represent *G. elata* proteins, blue rhombuses represent *Arabidopsis* proteins, green triangles represent maize proteins, and red five-pointed stars represent proteins with known function.

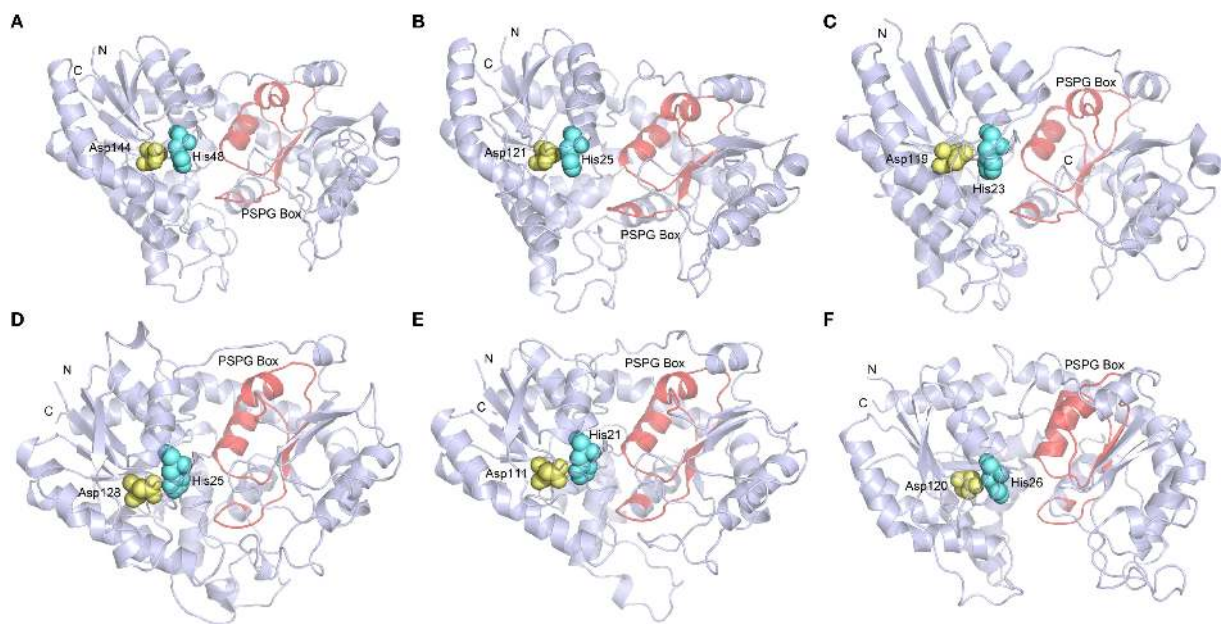


FIGURE 6 Spatial structure model of GT16 (A), GT17 (B), GT18 (C), GT24 (D), GT42 (E), and GT44 (F).

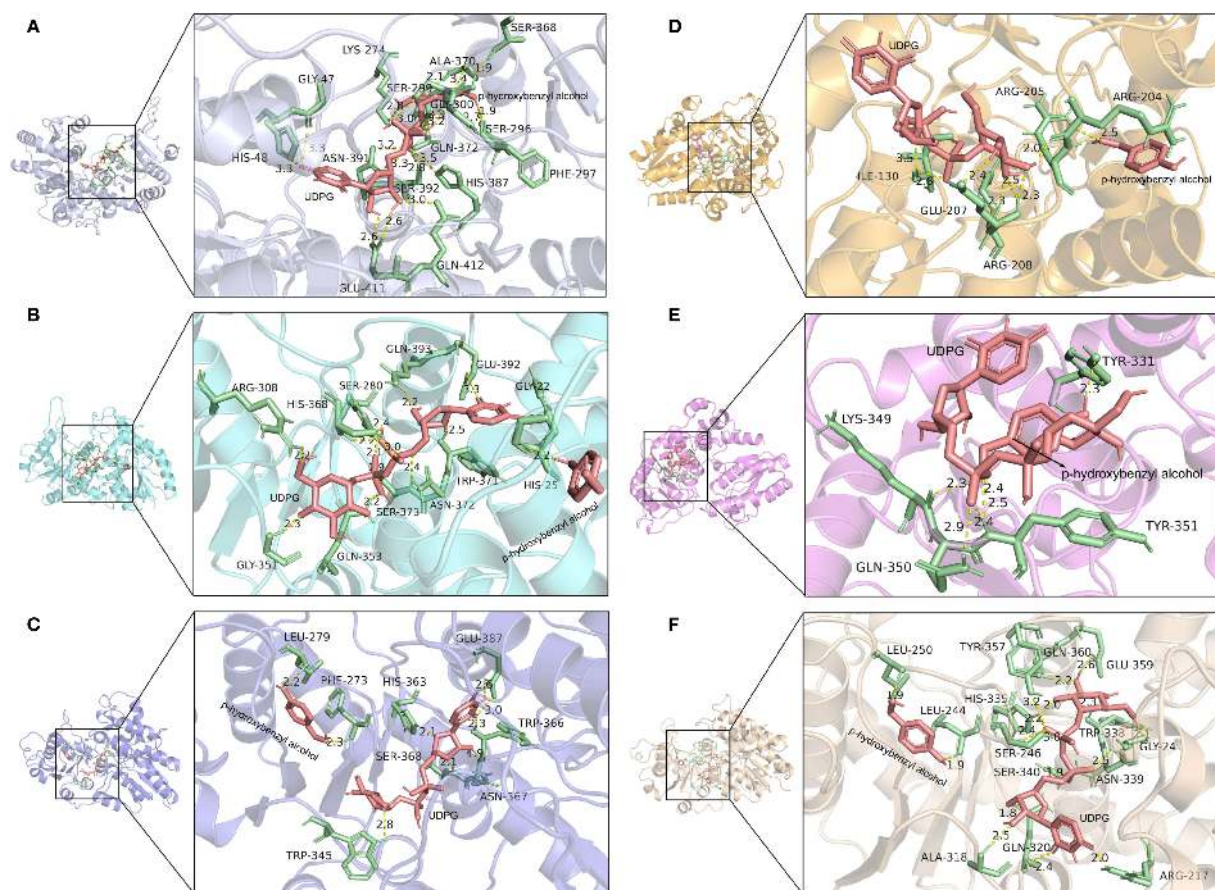


FIGURE 7

Visualization of the molecular docking results of GT16 (A), GT17 (B), GT18 (C), GT24 (D), GT42 (E), and GT44 (F).

kcal/mol) to the GT17-UDPG complex, where the catalytic residue His25 formed shorter hydrogen bonds (2.2 Å) with UDP-glucose (Figure 7B). The key conserved residues in the PSPG box of GT16, GT17, GT18, and GT44 formed hydrogen bonds with UDP-glucose (UDPG), facilitating substrate recognition and binding (Figures 7A–C, F). However, in GT24 and GT42, neither the catalytic residue His nor the key residues within the PSPG box formed interactions with UDPG and p-hydroxybenzyl alcohol (Figures 7D, E).

qRT-PCR analysis

To confirm the accuracy of the transcriptome sequencing results, six DEGs (*PAL4*, *CCoAMT1*, *C4H*, *4CL2*, *ADH4*, *UGE1*, and *INV8*) involved in phenolic and soluble sugar metabolic pathways, and six GT genes (*GT16*, *GT17*, *GT18*, *GT24*, *GT42*, and *GT44*) were randomly selected for validation using qRT-PCR analysis. As shown in Supplementary Figure S12, the expression profiles of the 12 genes were consistent in the qRT-PCR and FPKM values, confirming the accuracy and repeatability of the transcriptome analysis.

Discussion

The accumulation of secondary metabolites and quality traits in medicinal plants is co-regulated by the genetic characteristics of germplasm resources and geo-ecological factors, such as climate conditions, soil physicochemical properties, and altitude gradients (Li et al., 2020). Regional environments, cultivation techniques, and management practices significantly influence the growth of *G. elata*. With breakthroughs in planting technologies and the rapid development of the health industry, the cultivation area of *G. elata* has expanded, both globally and nationally, garnering significant attention for the breeding of varieties adapted to different regions (Fan et al., 2025). Zhong collected *G. elata* from Yunnan, Guizhou, and Hubei provinces to investigate its volatile components. The authors showed that the flavor characteristics of *G. elata* from Yunnan were significantly superior to those from Guizhou and Hubei (Zhong et al., 2025). In this study, the samples were sourced from partial traditional producing areas spanning two major regions: Southwest China (Yunnan and Guizhou) and Central China (Hubei and Anhui). Soluble sugars are important nutritional components in *G. elata*. Sucrose, glucose, D-fructose,

and D-xylose were the most abundant in all *G. elata* samples (Figure 1B), however, their variations across different producing areas and varieties were not significant. Therefore, we focused more on the differences in medicinal components, specifically phenolic compounds. Among different varieties of *G. elata*, the phenolic content was highest in the GB, followed by GR, and lowest in the GR×GB. From a geographical perspective, the phenolic content of the same variety shows significant regional differences, consistent with previous findings (Zhu et al., 2024).

Previous studies compared the transcriptomic differences between the tubers and immature tubers of *G. elata*, identifying two unigenes potentially involved in the hydroxylation and glycosylation steps of gastrodin biosynthesis (Tsai et al., 2016). Transcriptomic analysis has also been conducted on the flowers, stems, and tubers of *G. elata*, along with a comparison of the content of six major phenolic compounds (Shan et al., 2021). Another study performed metabolomic and transcriptomic analyses on immature and mature tubers of *G. elata*, detecting 76 different amino acids and their derivatives as well as 34 vitamins in the metabolome. A total of 16 DEGs were identified in the transcriptome, and the enzymes encoded by these DEGs were found to promote or inhibit the biosynthesis of amino acids and vitamins in *G. elata* (Wang and Shahid, 2023). Additionally, research on the transcriptome of *G. elata* seeds during germination in symbiosis with mycorrhizal fungi revealed the involvement of clathrin-mediated endocytosis in seed germination and elucidated the molecular mechanisms of the symbiotic germination between endogenous fungi and *G. elata*. These findings provide significant guidance for the innovative development of *G. elata* cultivation techniques (Zeng et al., 2017). However, transcriptomic data for *G. elata* from different regions and varieties remain lacking. Some studies have attributed the regional differences in metabolic accumulation and antioxidant capacity of mature *G. elata* tubers to the genetic characteristics of germplasm resources and geoeological factors (Zhang et al., 2024). Our transcriptome analysis of eight groups of *G. elata* from different origins and varieties showed a lower number of DEGs within the same production area, and a higher number for the same variety across different production areas (Supplementary Figure S9). Furthermore, 93 common DEGs were identified in the GR×GB comparison across three production areas (Figure 3B), indicating genetic stability in the selected hybrid variety during evolution (Yoldi-Achaladabaso et al., 2025).

Research on gastrodin in plants, especially regarding the phenolic synthesis pathway, is still relatively limited. In particular, the detailed biosynthetic pathway, regulatory mechanisms, and physiological importance need further confirmation through molecular biology and metabolomics studies. Within the proposed biosynthetic pathway, we have identified potential key regulatory genes: *ADH5* and *ADH11*, and 19 GT genes (*GT3*, *GT38*, *GT11*, *GT16*, *GT24*, *GT7*, *GT39*, *GT9*, *GT25*, *GT41*, *GT45*, *GT4*, *GT8*, *GT35*, *GT37*, *GT46*, *GT18*, *GT19* and *GT1*). These genes are hypothesized to play central regulatory roles in the synthesis of phenolic compounds. Glycosylation modifications are highly effective in improving the solubility and enhancing the stability of natural products and are one

of the most common and important modifications in nature (Moradi et al., 2016). Glycosidic natural products generated following glycosylation can exhibit significantly enhanced water solubility, pharmacological activity, bioavailability, and thus, higher medicinal value (Ren et al., 2023). The GT family is a large and diverse gene family (Yang et al., 2023). Correlation analysis among 46 GT genes and 6 phenolic compounds revealed significant positive associations between *GT16*, *GT18*, and *GT24* and gastrodin accumulation, suggesting that these genes play key catalytic roles in gastrodin biosynthesis. Several UGTs from groups D and E in phylogenetic analysis can catalyze the conversion of p-hydroxybenzyl alcohol into gastrodin (Figure 5A). For instance, UGT73B6 from group D can convert 4-hydroxybenzoic acid into gastrodin in *E. coli*, and an engineered variant, UGT73B6FS, showed more than a 7-fold increase in catalytic efficiency for gastrodin production (Bai et al., 2016). Similarly, AsUGTsyn from group E efficiently catalyzes the transformation of p-hydroxybenzyl alcohol into gastrodin in yeast, achieving a yield of 2.1 g/L (Yin et al., 2022). Additionally, AtUGT from *Arabidopsis thaliana* demonstrated high efficiency in gastrodin synthesis with a conversion rate of 94.34% (Xia et al., 2024). Based on these findings, we performed protein structure modeling and molecular docking analysis on several GeUGTs from groups D and E, namely *GT16*, *GT17*, *GT18*, *GT24*, *GT42*, and *GT44* (Figures 6 and 7), to predict their potential roles. In *GT16*, *GT17*, *GT18*, and *GT44*, the catalytic residue His and the key residues within the PSPG box formed interactions with UDPG and p-hydroxybenzyl alcohol. Therefore, we hypothesized that p-hydroxybenzyl alcohol can produce gastrodin under the catalysis of the enzymes encoded by *GT16*, *GT17*, *GT18*, and *GT44*.

In summary, GR, GB, and GR×GB are the primary commercially circulating varieties in the Chinese market, with distinct appearances and morphological characteristics among different *G. elata* varieties. GR×GB's superior morphological traits are favored by consumers; however, its content of carbohydrates and phenolic compounds is slightly lower than its parent strains. For medicinal markets and scientific research, traditional GR and GB varieties align more closely with the demands of authentic medicinal material markets and extraction-specific applications due to their higher bioactive compound content. In future studies, we intend to conduct *in vitro* enzymatic assays on GT to fully elucidate the catalytic mechanism underlying gastrodin biosynthesis and enhance the medicinal components in *G. elata*.

Conclusion

In this study, we systematically analyzed the characteristics of 6 phenolic and 19 sugar metabolic pathways and their molecular regulatory mechanisms in *G. elata* from different origins by integrating metabolomic and transcriptomic analyses. Our results indicate that the accumulation of secondary metabolites (e.g., gastrodin, parishin, p-hydroxybenzyl alcohol, and soluble sugars) was significantly higher in the GR and GB than GR×GB. We identified 96 genes related to the phenolic synthesis pathway, and key functional genes, such as *GT16*, *GT17*, *GT18*, *GT24*, *GT42*, and

GT44, were analyzed for sequence and structural analyses. Our findings also demonstrated the effects of geographic differences and genetic background on the accumulation of secondary metabolites, providing valuable insights into the metabolites and molecular regulatory mechanisms of *G. elata* metabolites in different varieties from different regions and facilitating further research in this field.

Data availability statement

The RNA-seq data has been submitted to NCBI SRA: PRJNA1249630. Data will be made available on request.

Author contributions

HW: Writing – original draft. SY: Writing – review & editing. HY: Data curation, Writing – review & editing. QF: Software, Writing – review & editing. JL: Validation, Writing – review & editing. NZ: Supervision, Writing – review & editing. CP: Supervision, Writing – review & editing. TS: Writing – review & editing. Resources. SG: Funding acquisition, Writing – review & editing. LZ: Project administration, Writing – review & editing.

Funding

The author(s) declare financial support was received for the research and/or publication of this article. This work was supported by the Research Funds of Center for Xin'an Medicine and Modernization of Traditional Chinese Medicine of IHM (2023CXMMTCM008), National Natural Science Foundation of China (U21A20406), the Excellent Young Scholars Project of Natural Science Foundation of Anhui Province in China (2208085Y30), the Key Project Foundation of Support Program for the Excellent Young Faculties in Universities of Anhui Province in China (gxyqZD2022051), Science Research Project at the Universities of Anhui Province for Distinguished Young Scholars (2023AH020036), Young Elite Scientists Sponsorship Program by CACM (CACM-2023-QNRC2-B23), Traditional Chinese Medicine

high-level key discipline construction project of National Administration of Traditional Chinese Medicine Science of Chinese Medicinal Material Resources (pharmaceutical botany) (zyyzdxk-2023095). MOE-Anhui Joint Collaborative Innovation Center for Quality Improvement of Anhui Genuine Chinese Medicinal Materials(xtcx202403).

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that no Generative AI was used in the creation of this manuscript.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fpls.2025.1656554/full#supplementary-material>

References

- Bai, Y., Bi, H., Zhuang, Y., Liu, C., Cai, T., Liu, X., et al. (2014). Production of salidroside in metabolically engineered *Escherichia coli*. *Sci. Rep.* 4, 6640. doi: 10.1038/srep06640
- Bai, Y., Yin, H., Bi, H., Zhuang, Y., Liu, T., and Ma, Y. (2016). *De novo* biosynthesis of gastrodin in *Escherichia coli*. *Metab. Eng.* 35, 138–147. doi: 10.1016/j.mbs.2016.01.002
- Chen, S., Liu, J. Q., Xiao, H., Zhang, J., and Liu, A. (2016). Simultaneous qualitative assessment and quantitative analysis of metabolites (phenolics, nucleosides and amino acids) from the roots of fresh *Gastrodia elata* using UPLC-ESI-triple quadrupole ion MS and ESI-linear ion trap high-resolution MS. *PLoS One* 11, e0150647. doi: 10.1371/journal.pone.0150647
- Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560
- Fan, Y., Tian, M., Hu, D., and Xiong, Y. (2025). Exploring the biocultural nexus of *Gastrodia elata* in Zhaotong: A pathway to ecological conservation and economic growth. *Biology* 14, 846. doi: 10.3390/biology14070846
- Gómez-González, S., Ruiz-Jiménez, J., Priego-Capote, F., and Luque de Castro, M. D. (2010). Qualitative and quantitative sugar profiling in olive fruits, leaves, and stems by gas chromatography-tandem mass spectrometry (GC-MS/MS) after ultrasound-assisted leaching. *J. Agric. Food Chem.* 58, 12292–12299. doi: 10.1021/jf102350s
- Guan, H., Ling, X., Xu, J., Zhu, Y., Zhang, J., and Liu, X. (2022). Structural characterization of polysaccharide derived from *Gastrodia elata* and its

- immunostimulatory effect on RAW264.7 cells. *Molecules* 27, 8059. doi: 10.3390/molecules27228059
- Guo, T., Wang, C., Xue, Q., Zhao, J., and Yang, L. (2016). Phylogenetic analyses of *Armillaria* reveal at least 15 phylogenetic lineages in China, seven of which are associated with cultivated *Gastrodia elata*. *PLoS One* 11, e0154794. doi: 10.1371/journal.pone.0154794
- Heese, K. (2020). *Gastrodia elata* Blume (Tianma): Hope for brain aging and dementia. *Evidence-Based Complement. Altern. Med.* 2020, 8870148. doi: 10.1155/2020/8870148
- Huang, C., Huang, C., Liao, W., Kao, S., and Cheng, C. (2024). Neuroprotective effects of *Gastrodia elata* Blume on promoting M2 microglial polarization by inhibiting JNK/TLR4/T3JAM/NF- κ B signaling after transient ischemic stroke in rats. *Front. Pharmacol.* 15. doi: 10.3389/fphar.2024.1469602
- Ji, N., Liu, P., Zhang, N., Yang, S., and Zhang, M. (2022). Comparison on bioactivities and characteristics of polysaccharides from four varieties of *Gastrodia elata* Blume. *Front. Chem.* 10. doi: 10.3389/fchem.2022.956724
- Kang, C., Lai, C., Zhao, D., Zhou, T., Liu, D., Lv, C., et al. (2017). A practical protocol for comprehensive evaluation of sulfur-fumigation of *Gastrodia* Rhizoma using metabolome and health risk assessment analysis. *J. Hazard. Mater.* 340, 221–230. doi: 10.1016/j.jhazmat.2017.07.003
- Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317
- Li, Y., Kong, D., Fu, Y., Sussman, M. R., and Wu, H. (2020). The effect of developmental and environmental factors on secondary metabolites in medicinal plants. *Plant Physiol. Biochem.* 148, 80–89. doi: 10.1016/j.plaphy.2020.01.006
- Li, N., Zheng, M., Kong, X., Zhou, Y., Yang, X., He, Q., et al. (2025). Unlocking the chemical blueprint of *Gastrodia elata* Blume (Orchidaceae): A pseudotargeted LC-MS/MS metabolomics approach for cultivar authentication. *J. Pharm. Biomed. Anal.* 264, 116954. doi: 10.1016/j.jpba.2024.116954
- Liu, J., Yang, X., Li, Z., Miao, J., Li, S., Zhang, W., et al. (2024). The role of symbiotic fungi in the life cycle of *Gastrodia elata* Blume (Orchidaceae): a comprehensive review. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1309038
- Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15, 550. doi: 10.1186/s13059-014-0550-8
- Ma, J., Deng, Y., Wang, Y., Liu, Q., An, J., Li, M., et al. (2021). A comparative study on ingredient and efficiency difference between fresh and steamed *Gastrodia elata* Blume: An herbal material to a novel functional food. *J. Funct. Foods* 82, 104512. doi: 10.1016/j.jff.2021.104512
- Ma, R., Sun, X., Yang, C., and Fan, Y. (2023). Integrated transcriptome and metabolome provide insight into flavonoid variation in goji berries (*Lycium barbarum* L.) from different areas in China. *Plant Physiol. Biochem.* 199, 107722. doi: 10.1016/j.plaphy.2023.107722
- Medeiros, P. M., and Simoneit, B. R. T. (2007). Analysis of sugars in environmental samples by gas chromatography-mass spectrometry. *J. Chromatogr. A* 1141, 271–278. doi: 10.1016/j.chroma.2006.12.017
- Moradi, S., Hussein, W., Varamini, P., Simerska, P., and Toth, I. (2016). Glycosylation, an effective synthetic strategy to improve the bioavailability of therapeutic peptides. *Chem. Sci.* 7, 2492–2500. doi: 10.1039/c5sc04392a
- Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122
- Ren, J., Barton, C. D., and Zhan, J. (2023). Engineered production of bioactive polyphenolic O-glycosides. *Biotechnol. Adv.* 65, 108146. doi: 10.1016/j.biotechadv.2023.108146
- Ross, J., Li, Y., Lim, E., and Bowles, D. J. (2001). Higher plant glycosyltransferases. *Genome Biol.* 2, REVIEWS3004. doi: 10.1186/gb-2001-2-2-reviews3004
- Saito, K. (2018). Development of plant metabolomics and medicinal plant genomics. *Yakugaku Zasshi* 138, 1–18. doi: 10.1248/yakushi.17-00193
- Shan, T., Yin, M., Wu, J., Yu, H., Liu, M., Xu, R., et al. (2021). Comparative transcriptome analysis of tubers, stems, and flowers of *Gastrodia elata* Blume reveals potential genes involved in the biosynthesis of phenolics. *Fitoterapia* 153, 104988. doi: 10.1016/j.fitote.2021.104988
- Su, Z., Yang, Y., Chen, S., Tang, Z., and Xu, H. (2023). The processing methods, phytochemistry and pharmacology of *Gastrodia elata* Bl.: A comprehensive review. *J. Ethnopharmacol.* 314, 116467. doi: 10.1016/j.jep.2023.116467
- Tsai, C., Wu, K., Chiang, T., Huang, C., Chou, C., Li, S., et al. (2016). Comparative transcriptome analysis of *Gastrodia elata* (Orchidaceae) in response to fungus symbiosis to identify gastrodin biosynthesis-related genes. *BMC Genomics* 17, 212. doi: 10.1186/s12864-016-2508-6
- Wang, Y., and Shahid, M. Q. (2023). Insights into the nutritional properties and molecular basis of biosynthesis of amino acids and vitamins of *Gastrodia elata* offered by metabolomic and transcriptomic analysis. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1183139
- Wieghaus, A., Roelfs, K., Twyman, R., Prüfer, D., and Schulze Gronover, C. (2022). Comparative transcriptome analysis in *Taraxacum koksaghyz* to identify genes that determine root volume and root length. *Front. Genet.* 12. doi: 10.3389/fgenet.2021.784883
- Wu, Z., Raven, P. H., and Hong, D. (2009). *Flora of China* Vol. 25 (Beijing: Science Press), 203.
- Xia, H., Zhang, Z., Ding, J., Jiang, K., and Xue, F. (2024). A glycosyltransferase from *Arabidopsis thaliana* enables the efficient enzymatic synthesis of gastrodin. *Catalysis Lett.* 154 (4), 1558–1566. doi: 10.1007/s10562-023-04406-y
- Yang, Q., Zhang, Y., Qu, X., Wu, F., Li, X., Ren, M., et al. (2023). Genome-wide analysis of UDP-glycosyltransferases family and identification of UGT genes involved in abiotic stress and flavonol biosynthesis in *Nicotiana tabacum*. *BMC Plant Biol.* 23 (1), 204. doi: 10.1186/s12870-023-04208-9
- Yin, H., Hu, T., Zhuang, Y., and Liu, T. (2022). Correction to: Metabolic engineering of *Saccharomyces cerevisiae* for high-level production of gastrodin from glucose. *Microb. Cell Fact.* 21, 18. doi: 10.1186/s12934-022-01747-y
- Yoldi-Achalandabaso, A., Agirresarobe, A., Katamadze, A., Burini, G., Vergara-Díaz, O., Mota, M., et al. (2025). Tritordeum, barley landraces and ear photosynthesis are key players in cereal resilience under future extreme drought conditions. *Plant Stress* 15, 100765. doi: 10.1016/j.stress.2025.100765
- Yu, E., Gao, Y., Li, Y., Zang, P., Zhao, Y., and He, Z. (2022). An exploration of mechanism of high quality and yield of *Gastrodia elata* Bl. f. *glauca* by the isolation, identification and evaluation of *Armillaria*. *BMC Plant Biol.* 22, 621. doi: 10.1186/s12870-022-04007-8
- Yu, H., Liu, M., Yin, M., Shan, T., Peng, H., Wang, J., et al. (2021). Transcriptome analysis identifies putative genes involved in triterpenoid biosynthesis in *Platycodon grandiflorus*. *Planta* 254, 34. doi: 10.1007/s00425-021-03677-2
- Yuan, Y., Jin, X., Liu, J., Zhao, X., Zhou, J., Wang, X., et al. (2018). The *Gastrodia elata* genome provides insights into plant adaptation to heterotrophy. *Nat. Commun.* 9, 1615. doi: 10.1038/s41467-018-03423-5
- Zeng, Y., Gu, J., Chen, L., Zhang, T., and Zhou, X. (2021). Gastrodin as a multi-target protective compound reverses learning memory deficits and AD-like pathology in APP/PS1 transgenic mice. *J. Funct. Foods* 77, 104324. doi: 10.1016/j.jff.2020.104324
- Zeng, X., Li, J., Chen, T., Li, Y., and Guo, S. (2023). Global metabolic profile and multiple phytochemicals in the different varieties of *Gastrodia elata* Blume. *Front. Plant Sci.* 14. doi: 10.3389/fpls.2023.1249456
- Zeng, X., Li, Y., Ling, H., Liu, S., Liu, M., Chen, J., et al. (2017). Transcriptomic analyses reveal clathrin-mediated endocytosis involved in symbiotic seed germination of *Gastrodia elata*. *Bot. Stud.* 58, 31. doi: 10.1186/s40529-017-0185-7
- Zhang, J., Li, L., Liu, Q., Zhao, Z., Su, D., Xiao, C., et al. (2023). Gastrodin programs an Arg-1⁺ microglial phenotype in hippocampus to ameliorate depression and anxiety-like behaviors via the Nrf2 pathway in mice. *Phytomedicine* 113, 154725. doi: 10.1016/j.phymed.2023.154725
- Zhang, Z., Li, X., Zhang, Y., Yin, N., Wu, G., Wei, G., et al. (2024). Ecological factors impacting genetic characteristics and metabolite accumulations of *Gastrodia elata*. *Chin. Herbal Medicines* 17, 562–574. doi: 10.1016/j.chmed.2024.09.002
- Zheng, H., Zhang, P., Qin, J., Guo, J., and Deng, J. (2023). High-throughput sequencing-based analysis of the composition and diversity of endophytic bacteria community in tubers of *Gastrodia elata* f. *glauca*. *Front. Microbiol.* 13, 1092552. doi: 10.3389/fmicb.2022.1092552
- Zhao, M., Hong, X., Abdullah, A., Yao, R., and Xiao, Y. (2021). Rapid biosynthesis of phenolic glycosides and their derivatives from biomass-derived hydroxycinnamates. *Green Chem.* 23. doi: 10.3389/fmicb.2022.1092552
- Zhong, Y., Li, J., Liu, H., and Wang, Y. (2025). Rapid traceability of *Gastrodia elata* Blume origins and analysis of key volatile organic components using FTIR and HS-SPME-GC-MS combined with chemometrics. *Food Chem.: X* 29, 102770. doi: 10.1016/j.fochx.2025.102770
- Zhou, D., Hu, Y., Li, R., He, J., Zhao, M., and Tang, Q. (2025). Optimization of ultrasonic-assisted compound enzyme extraction process, structural characterization, and antioxidant activity of *Gastrodia elata* polysaccharides. *J. Mol. Struct.* 1327, 141214. doi: 10.1016/j.molstruct.2024.141214
- Zhu, Y., Zhou, X., Xu, L., Wang, C., Ma, J., Ruan, P., et al. (2024). Comprehensive quality evaluation of *Gastrodia elata* from different producing areas based on membership function method. *J. Chin. Med. Mater.* 47, 409–414. doi: 10.13863/j.issn1001-4454.2024.02.024