Insights into the nutritional properties and molecular basis of biosynthesis of amino acids and vitamins of *Gastrodia elata* offered by metabolomic and transcriptomic analysis

Yunsheng Wang[1]* and Muhammad Qasim Shahid[2,3,4]*

[1]School of Health and Life Science, Kaili University, Kaili City, Guizhou, China, [2]State Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, South China Agricultural University, Guangzhou, China, [3]Guangdong Provincial Key Laboratory of Plant Molecular Breeding, South China Agricultural University, Guangzhou, China, [4]College of Agriculture, South China Agricultural University, Guangzhou, Guangdong, China

Gastrodia elata Blume (*GE*), a traditional and precious Chinese medicinal material, has been approved as a functional food. However, understanding *GE*'s nutritional properties and its molecular basis remains limited. Here, metabolomic and transcriptomic analyses were performed on young and mature tubers of *G. elata.f.elata* (*GEEy* and *GEEm*) and *G. elata.f.glauca* (*GEGy* and *GEGm*). A total of 345 metabolites were detected, including 76 different amino acids and their derivatives containing all human essential amino acids (e.g., l-(+)-lysine, l-leucine), 13 vitamins (e.g., nicotinamide, thiamine), and 34 alkaloids (e.g., spermine, choline). *GEGm* has higher amino acid accumulation than *GEE*y, *GEEm* and GEGy, and vitamin contents were also slightly different in all four samples. Implying that *GE*, especially *GEGm*, is a kind of excellent complementary food as amino acid nutrition provider. From assembled 21,513 transcripts (genes) based on the transcriptome, we identified many genes that encode enzymes (e.g., pfkA, bglX, tyrAa, lysA, his B, aroA), which are responsible for the biosynthesis of amino acids and enzymes (e.g., nadA, URH1, NAPRT1, punA, rsgA) that related to vitamins metabolism. A total of 16 pairs of the differentially expressed genes (DEG) and differentially accumulated metabolites (DAM) (e.g., gene-tia006709 coding GAPDH and l-(+)-arginine, and gene-tia010180 coding tyrA and l-(+)-arginine) and three DEG-DAM pairs (e.g., gene-tia015379 coding NadA and nicotinate d-ribonucleoside) show significant similar positive or negative correlation based on three, and two comparisons of *GEEy vs. GEGy*, *GEGy vs. GEGm*, GEEy *vs. GEGy* and *GEEm vs. GEGm*, which involved into amino acid biosynthesis, and nicotinate nicotinamide metabolism, respectively. These results prove that the enzyme coded by these DEG promotes (positive correlation) or inhibits (negative correlation) the biosynthesis of parallel DAM in *GE*. Overall, the data and corresponding analysis in this study provide new insights into the nutritional properties of *GE and the* related molecular basis.

KEYWORDS

differentially expressed genes, differentially accumulated metabolites, *Gastrodia elata*, metabolome, transcriptome

1 Introduction

Gastrodia elata Blume (*GE*) (Orchidaceae) has been used to promote health and treat various neurological disorders such as convulsions, migraines, dizziness, seizures, paralysis, and mood disorders for thousands of years in East Asia (Lin et al., 2018). Modern medical research has found that *GE* has a variety of healthful activities, including anti-depressant (Qiu et al., 2014), antioxidation (Liu and Mori, 1992; Jung et al., 2007), anti-inflammation (Hwang et al., 2009), anti-obesity (Sun et al., 2012) activities, and has been found to improve memory (Hsieh et al., 1997). In addition, *GE* is a high-grade supplementary food in some Chinese dishes and has recently been approved as a functional food by the Chinese government (Shu et al., 2013; Jaswir et al., 2017; Lu et al., 2020; Ma et al., 2021). *GE*'s healthful and nutritional properties are undoubtedly endowed by its unique bioactive compounds that were biosynthesized in tuber tissues. Many studies have been executed to identify the active compounds in *GE* tissues responsible for their beneficial health effects. The first reported biochemical compound of *GE* was vitamin A, identified in 1936 (Liu, 1961). Then, vanilla alcohol was identified in the middle of the last century (Liu and Yang, 1958). Gastrodin (p-hydroxymethylphenyl-β-d-glucopyranoside), *GE*'s most important medicinal compound, was first identified in 1979 (Feng et al., 1979). In addition, at least sixteen gastrodin derivatives were identified in the tubers of *GE* (Xu et al., 2019). So far, approximately 100 bioactive compounds of *GE* have been identified, and these are primarily phenolics (Ojemann et al., 2006; Li et al., 2016a; Zhan et al., 2016; Liu and Huang, 2017; Xu et al., 2019).

The molecular basis of the biosynthesis of bioactive ingredients in medicinal plants has attracted broad interest in recent years. Transcriptome sequencing and genome sequencing are common techniques used for this purpose (Han et al., 2016; Yang et al., 2016a; Saito, 2018; Trócsányi et al., 2020). Phenol-based compounds, especially gastrodin, are the dominant active components of *GE*. Pharmacokinetic studies show that these compounds play a vital role in the curative effect of *GE* (Ha et al., 2000; Hsieh et al., 2000; An et al., 2003; Yu et al., 2005; Kim et al., 2007; Makni et al., 2011; Tang et al., 2015; Tsai et al., 2016). By sequencing and analyzing the transcriptome of *GE*, the genes coding for enzymes that play a crucial role in the biosynthesis of gastrodin have been identified (Tsai et al., 2016; Yang et al., 2020; Chen and Mi, 2021). In addition, the genes associated with mannose-binding lectin antifungal proteins and the genes tightly related to the biosynthesis of the phenolics of *GE* have also been identified by transcriptome sequencing (Yang et al., 2020; Shan et al., 2021).

As mentioned above, the bioactive compounds related to healthful properties and their molecular basis of *GE* have been studied widely and understood deeply. However, which related to the nutritional properties of *GE* are still limited. As a result, our understanding of the molecular nutriology of *GE* is still far from comprehensive, which hinders the further utilization of this important functional food. So exploiting the abovementioned concerns is meaningful for developing the *GE* industry and a more comprehensive understanding of *GE* biology. For a complementary food, the type and content of amino acids and vitamins are undoubtedly important indicators of nutritional value. *GE* is a "polytypic species" composed of five forms according to morphological traits (Zhou et al., 1987). *Gastrodia. elata.f. elata* (*GEE*) and *Gastrodia. elata.f. glauca* (*GEG*) are two forms that were widely cultivated and commercially produced as medicinal materials or food ingredients. In this study, a UPLC-MS/MS platform was used to determine metabolome from four different tuber tissues of *GE*, each with three biological repeats. These consist of young and mature tubers of *G. elata f. elata* (*GEEy* and *GEEm*) and *G. elata f. glauca* (*GEGy* and *GEGm*), representing two common *GE* types and their different development stages. Meanwhile, the transcriptomes of the above samples were generated using a Hiseq 4000 sequencer. The metabolome and transcriptome of the above samples were performed as four pairwise (*GEEy* and *GEEm*, *GEGy* and *GEGm*, *GEEy* and *GEGy*, and *GEEm* and *GEGm*) analyses combined and separately. The specific aims of this study were to (1) reveal the similarities and differences in the types and contents of the metabolic components, especially the amino acid and vitamins between *GEE* and *GEG, and* their different development stages; (2) identify the functional genes involved in the biosynthesis and regulation of nutrient compounds, mainly amino acids, and vitamins, of *GE*. Overall, we want to provide insight into the nutritional properties of *GE and* its corresponding molecular basis, which will help deepen the biological understanding of *GE*.

2 Results

2.1 Pearson correlation assessment and principal component analysis of metabolomes and transcriptomes among samples

Pearson's correlation analysis showed that the metabolites of the four tuber tissues of *GE* were highly similar, with the lowest correlation coefficient (r^2) observed between samples being 0.81 (Dataset S1). The metabolite and transcript profiles in the sample repeats showed the highest correlation, with r^2 values reaching a minimum of 0.98 and 0.99, respectively, which is higher than those between non-biological repeat samples (Figures 1A, B; Dataset S1). The first and second principal components of metabolite and transcript profiles accounted for 34.81% and 39.73%, and 26.62% and 25.91% of the total variance, respectively. Meanwhile, the metabolome and transcriptome of biological repeats were distinctly clustered together compared to samples from different tissues (Figures 1C, D). The results of Pearson's correlation analysis and principal component analysis (PCA) suggest that this study's metabolite and transcript profile data are highly reliable.

Abbreviations: DEGs, differentially expressed genes; DAMs, Differentially accumulated metabolites; GE, *Gastrodia elata*. Blume; GEEm, mature tuber of *Gastrodia elata*.f.elata; GEEy, young tuber of *Gastrodia elata*.f.elata; GEGm, mature tuber of *Gastrodia elata*.f.glauca; GEGy, young tuber of *Gastrodia elata*.f.glauca.
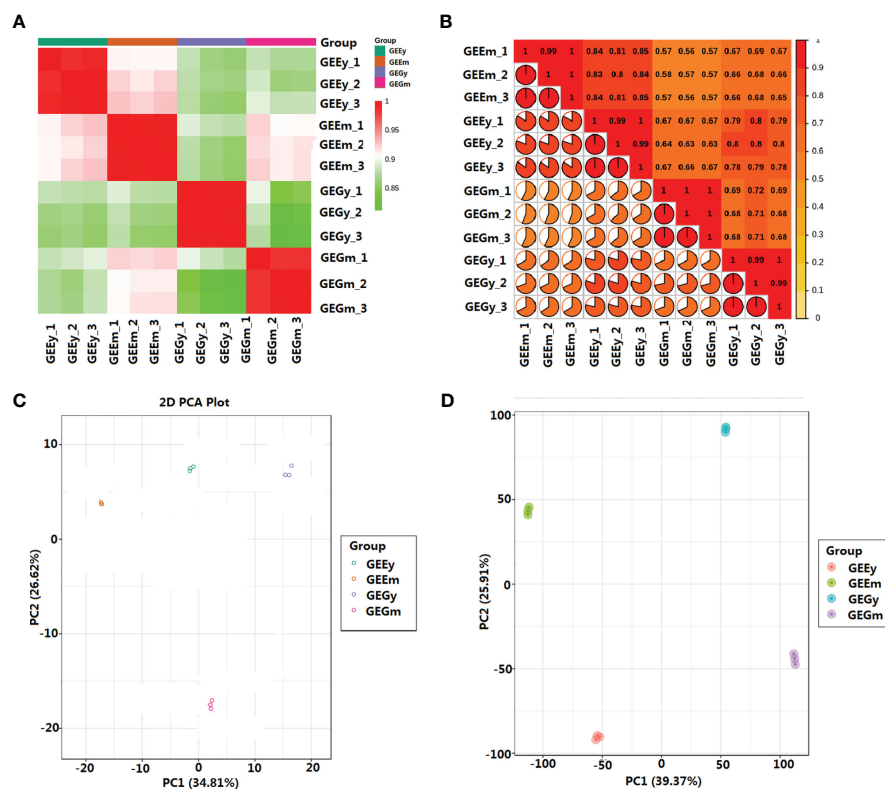
FIGURE 1
Relationships between the metabolites and transcripts of the samples. **(A)** Pearson's correlation analysis of metabolites; **(B)** Pearson's correlation analysis of transcripts; **(C)** Principal component analysis of metabolites; **(D)** Principal component analysis of transcripts. *GEEy* and *GEEm* indicate young and mature tubers of *G.elata.f.elata*, while *GEGy* and *GEGm* represent young and mature tubers *G.elata.f.glauca*, respectively.

2.2 Metabolites and transcripts profiles in four tuber tissues of GE

Three hundred forty-five unique metabolites were detected from twelve samples, including 75 types of amino acids and their derivatives (e.g., l-(+)-lysine, l-(+)-arginine, l-leucine, l-isoleucine, tryptophan, l-valine, l-(-)-threonine, l-phenylalanine, and l-methionine), and 13 kinds of vitamins, (e.g., nicotinamide, nicotinic acid, biotin, and riboflavin) (Table 1; Dataset S2). In addition, there are 45 Nucleotides and derivatives, 40 lipids, 18 organic acids, 39 types of phenolic acids, 34 types of alkaloids, 15 types of flavonoids, one tannin, five lignans and coumarin, three kinds of terpenoids also detected in young and mature tuber tissues of *GE* (Table 1; Dataset S2). Between 43,211,372 and 58,292,014 short raw reads were generated from the 12 samples using the Illumina Hiseq 4000 sequencer. After filtering out contamination, 42,333,902 to 57,576,240 clean reads, composed of 6.35 to 8.64 Gb, were obtained. More than 98% of the reads obtained for the samples were high-quality (> Q20) (Table S1). More than 90% of high-quality reads for each sample could be mapped on the released reference genome (Table S2). The mapping results identified 21,513 transcripts (genes), including 16,541 genes consistent with the sequences of the predicted protein-coding genes in the reference genome. In addition, 4,972 novel genes were also identified (Table S3).

2.3 Functional annotation of metabolome and transcriptome and identification of genes related to amino acid and vitamin metabolism

Of the 345 metabolites, 175 could be annotated by the Kyoto Encyclopedia of Genes and Genomes Database (KEGG) (https://www.kegg.jp/), but only 140 could be assigned to one or more of 82 specific metabolism pathways (Figure S1; Dataset S3). Of these, 126 (90%) were involved in various metabolic pathways, and 55 (39.29%) were involved in pathways related to the biosynthesis of secondary metabolites, including 25 metabolites were involved in pathways related to the biosynthesis of amino acids, four, three, five, seven, seven and two were related to the pathways "thiamine metabolism", "riboflavin metabolism", "vitamin B6 metabolism", "nicotinate and nicotinamide metabolism", "pantothenate and CoA biosynthesis" and "biotin metabolism" respectively (Figure S1; Dataset S3). Of the 21,513 identified genes, 16,453, (76.48%) could be annotated using at least one of the Nr, SwissPort, KOG, GO and KEGG databases (Table S3; Dataset S4). Of the 16,426 genes annotated using the Nr database, 13,585 (82.68%) were matched using the homologous sequences of *Dendrobium catenatum*, *Phalaenopsis* equestris, and *Apostasia shenzhenica*, which are also types of orchids (Table S3; Figure S2; Dataset S4). Of 12,223 GO annotated genes, the GO terms mainly fall into

TABLE 1 Counts of the metabolites in four tuber tissues of *G. elata* (*GE*).

Class	GE	GEEy	GEEm	GEGy	GEGm
Amino acids and derivatives	76	76	76	76	75
Nucleotides and derivatives	45	43	44	45	45
Lipids	40	40	40	40	40
Organic acids	18	18	18	18	18
Others	69	69	67	69	66
Phenolic acids	39	38	37	38	37
Flavonoids	15	15	15	13	14
Alkaloids	34	33	33	34	34
Terpenoids	3	3	3	3	2
Tannins	1	1	1	1	1
Lignans and Coumarins	5	3	5	3	3
Total	345	339	339	340	335

GE indicates Gastrodia elata; GEEy and GEEm indicate young and mature tubers of G. elata.f.elata (GEE), while GEGy and GEGm represent young and mature tubers of G. elata.f.glauca (GEG), respectively.

"cellular process" and "metabolic process" of biological process, and "cellular anatomical entity" of cell component, and "binding", "catalytic activity" of molecular function category (Figure 2; Dataset S4). A total of 5,720 genes were assigned to 138 pathways using KEGG annotation (Figure S3; Dataset S4). Of these, 17 genes were identified to code for eight enzymes (lysC, asd, hom, dapA, dapB, dapL, dapF, and lysA) involved in the lysine biosynthesis, which uses l-aspartate as the initial substrate. LysA, coded by the gene Gene-tia007535, played a direct role in the biosynthesis of lysine and was identified in this analysis (Figure S4; Dataset S4). 24 genes were found to involve in the pathway of histidine metabolism, including enzyme his G coding gene (gene-tia009269), his E coding genes (e.g., gene-tia012550, his I coding genes (e.g., gene-tia009135), his F coding genes (e.g., gene-tia011633), his B coding

gene (gene-tia009637), his C coding genes (gene-tia009637), and his D coding genes (e.g., gene-tia00517), which constitutes a complete metabolic pathway and catalyzes the substrate PRPP into l-histidine (Figure S5; Dataset S4). Forty-five genes (e.g., gene-tia006823, gene-tia009761, gene-tia011646) were identified to be related to phenylalanine, tyrosine, and tryptophan biosynthesis. These separately coded whole set enzymes (aroG, aroB, aroD, aroE, aroK, aroA, aroC, TRP3, trpD, TRP1, trpA, chorismate mutase, PAT, AAT, tyrAa, ADT, PDT, eta) that catalyze the substrate d-erythrose 4-phosphate into above three amino acids (Figure S6; Dataset S4). Among the 23 genes encoding 11 enzymes (SIR2, NMNAT, URH1, NAPRT1, 5`nucleotidase, nadD, CD203, nudC, nadE, nadD, nadK, nadB, and NADK) that are involved in the nicotinate and nicotinamide metabolism pathway, URH1 (which is
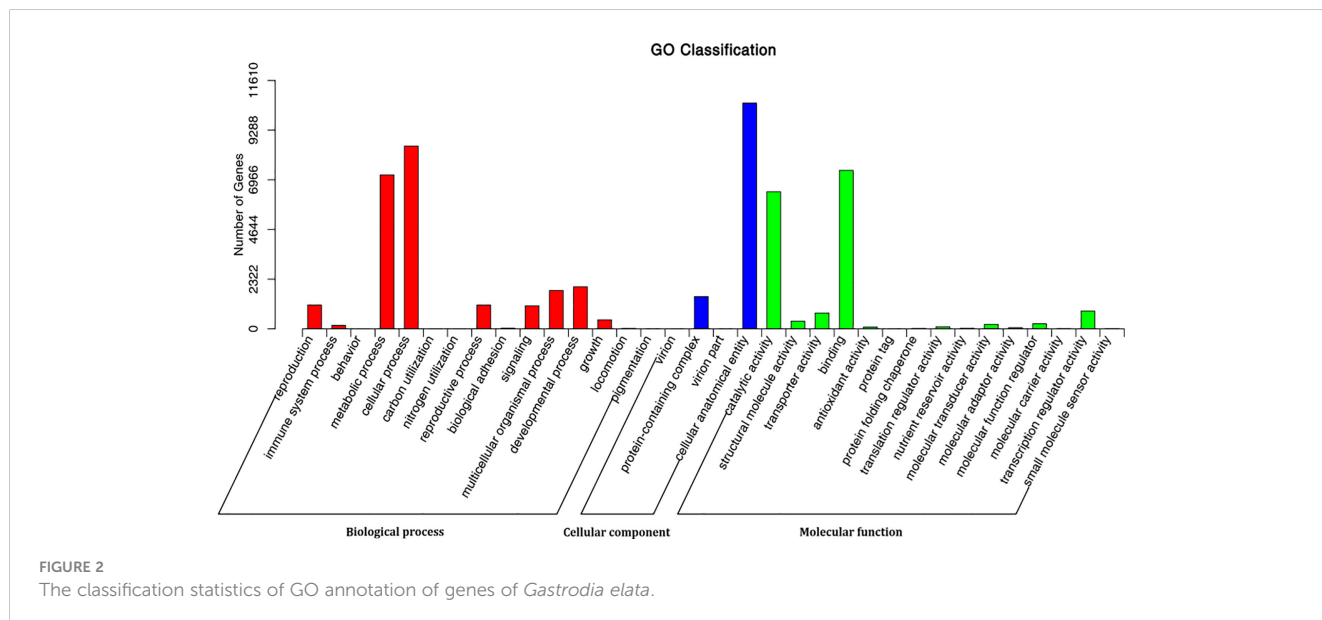

FIGURE 2
The classification statistics of GO annotation of genes of *Gastrodia elata*.

coded for by gene-tia004446 and gene-tia004985) and NAPRT1 (which is coded for by gene-tia008609) are key enzymes for the mutual transformation of nicotinate and nicotinate D ribonucleotide. The enzyme punA also catalyzes this mutual transformation, but no gene was found that codes for this enzyme (Figure S7; Dataset S4). Among 14 genes involved in the pathway of thiamine metabolism, the gene-tia010644 and gene-tia016175 codeTHI20, gene-tia010644 and gene-tia016175 code thiE, gene-tia002595 and gene-tia003843 code rsgA, and these three enzymes can catalyze the 4-amino-5-hydroxymethyl-2-methylpyrimidine phosphate into thiamine) (Figure S8; Dataset S4).

2.4 Metabolite accumulation similarity and differences among samples

Significant differences in the accumulation of metabolites (DAMs) were observed between *GEEm* and *GEGm for* 106 metabolites, *GEEy* and *GEGy for* 90 metabolites, and 36 metabolites in the above two pairwise samples (Figures 3A, B; Dataset S5). Of these, nine were up-regulated, including one amino acid (l-tyramine) and twenty-seven metabolites, including five amino acids and their derivatives (l-asparagine, l-ornithine, l-(+)-arginine l-(-)-tyrosine and 2,6-diaminooimelic acid) were down-regulated based on pairwise analysis of *GEEm vs. GEGm*, and *GEEy vs. GEGy* (Dataset S5). 92, 105, and 33 DAMs were identified from pairwise studies of *GEEy vs. GEEm, GEGy vs. GEGm, and* both *GEEy vs. GEE*m and *GEGy vs. GEGm* (Figures 3C, D; Dataset S5). Nine were down-regulated, including one amino acid derivative (e.g., alanylleucine). Twenty-four metabolites, including nine amino acids and their derivatives (l-(-)-threonine, α-aminocaproic acid, l-isoleucine, l-allo-isoleucine, l-leucine, glutamic acid, Cys-Gly, l-tryptophan, glutathione reduced form) were up-regulated based on pairwise analysis of *GEEy vs. GEEm* and *GEGy vs. GEGm* (Dataset S5). According to four comparisons, most metabolites accumulate with non-significant differences between samples (*GEEy vs. GEEm, GEGy vs. GEGm, GEEy vs. GEGy* and *GEEm vs. GEGm*). 139 metabolites, including 22 amino acids and derivatives (e.g., l-phenylalanine, l-methionine, histamine, l-2-chlorophenylalanine, d-serine, n-α-acetyl-l-glutamine), 10 vitamins (e.g., nicotinamide, thiamine, and biotin) and 15 alkaloids (e.g., agmatine, spermidine, and spermine)
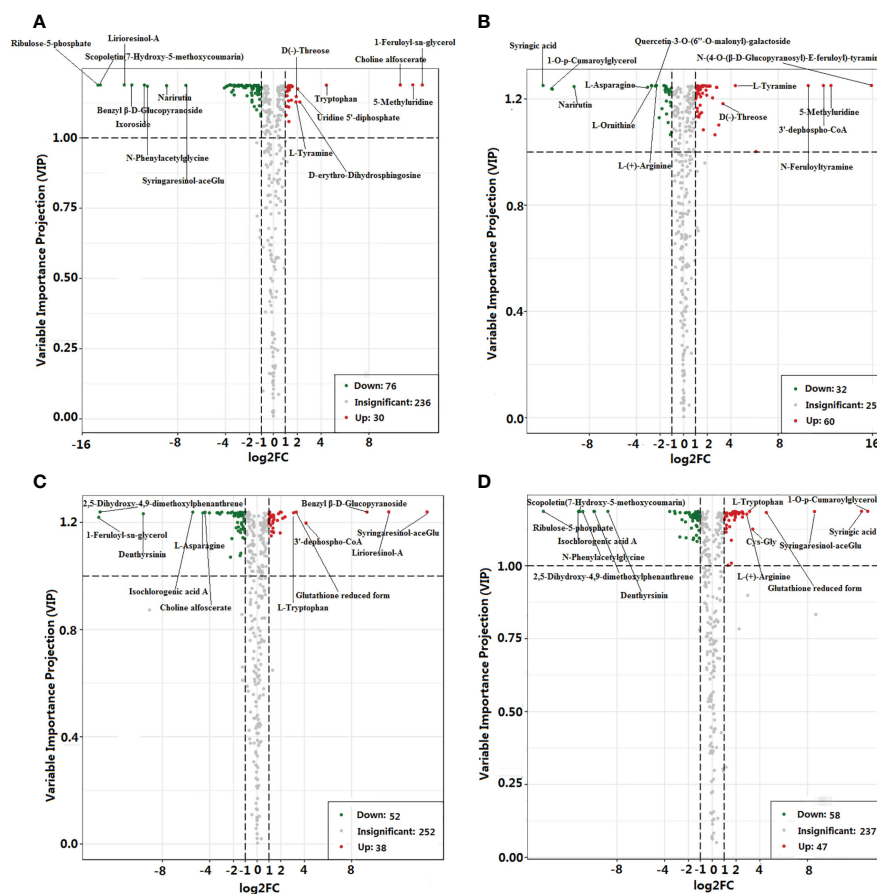

FIGURE 3
Statistics of different and stable accumulation of metabolites between comparisons. **(A)** *GEEm vs. GEGm*; **(B)** *GEEy vs. GEEm*; **(C)** *GEEy vs.* **(D)** *GEGy*; *GEGy vs. GEGm. GEEy* and *GEEm* indicate young and mature tubers of *G. elata.f.elata*, while *GEGy* and *GEGm* indicate young and mature tubers of *G. elata.f.elata*.

contents were found to be similar in all four samples (Dataset S5). Ten metabolites, including l-asparagine and gastrodin, are DAMs based on all the above four pairwise analyses (Dataset S5).

2.5 Differentially expressed genes among samples and these related to amino acids and vitamin metabolism

There were 2853 differentially expressed genes (DEGs), 1979 down-regulated and 874 up-regulated, in *GEEy* compared with *GEEm*, and 4030 DEGs, 1778 down-regulated and 2252 up-regulated, in *GEGy* when compared with *GEGm*. Based on pairwise analyses, 153 down-regulated and 76 up-regulated DEGs appeared between *GEEy vs. GEEm* and *GEGy vs. GEGm* (Figures 4A, B; Dataset S6). The DEGs from the two pairwise analyses are significantly enriched in several KEGG pathways, including starch and sucrose metabolism, flavonoid biosynthesis, and phenylpropanoid biosynthesis pathways (Figure S9; Dataset S7). Of these DEGs from the two pairwise analyses, 18 (e.g., gene-

tia002357, gene-tia002852, gene-tia003785, gene-tia004085, gene-tia005120) were identified to be involved in pathways biosynthesis of amino acids, one (gene-tia010644), two (gene-tia001539, gene-tia003140), two (gene-tia014555, gene-tia000824), one (gene-tia004985), two (gene-tia009761, gene-tia002852) and one (gene-tia009822) were identified to be related to thiamine metabolism, riboflavin metabolism, vitamin B6 metabolism, nicotinate, and nicotinamide metabolism, pantothenate and CoA biosynthesis and biotin metabolism, respectively (Dataset S7). There were 6221 DEGs, 2854 down-regulated and 3367 up-regulated, found in *GEEy* compared with *GEGy*, and 7016 DEGs, 2935 down-regulated and 4081 up-regulated, found in *GEEm* when compared with *GEGm*. Among these DEGs, 3766 appeared in both pairwise analyses (Figures 4C, D; Dataset S6). The DEGs between *GEEy* and *GEEm* and between *GEGy* and *GEGm* are significantly enriched in several KEGG pathways, including galactose metabolism, flavonoid biosynthesis, and phenylpropanoid biosynthesis (Figure S9; Dataset S7). Among these DEGs from the two pairwise analyses, 28 (e.g., gene-tia001033, gene-tia001508, gene-tia001739, gene-tia002208, gene-tia004922) were related to the biosynthesis of
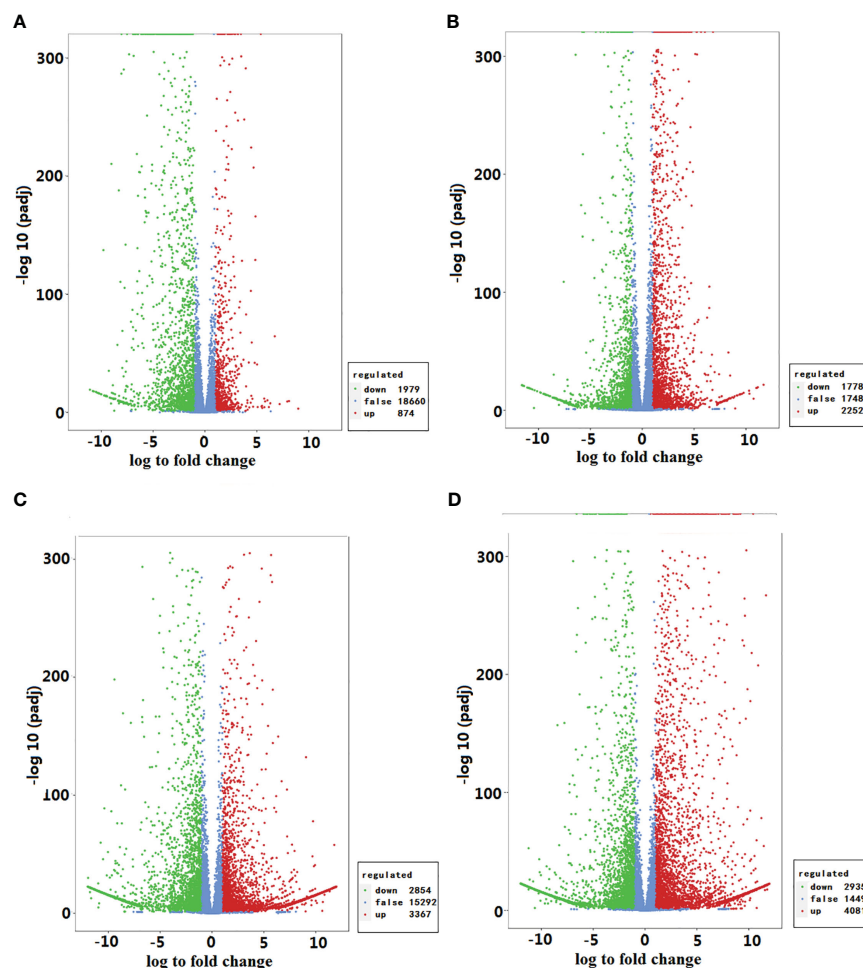

FIGURE 4
Statistics of different and stable expressions of genes between different comparisons. **(A)** *GEEy vs. GEEm*; **(B)** *GEEy vs. GEGy*; **(C)** *GEEm vs. GEGm*; **(D)** *GEGy vs. GEGm GEEy* and *GEEm* indicate young and mature tubers of *G. elata.f.elata*, while *GEGy* and *GEGm* represent young and mature tubers *G. elata.f.glauca*, respectively.

amino acids, two (gene-tia003026, gene-tia014391), four (gene-tia003899, gene-tia001539, gene-tia001540, gene-tia003140), three (gene-tia015010, gene-tia000824, gene-tia000825), three (gene-tia009709, gene-tia011772, gene-tia01282), four (gene-tia009761, gene-tia001787, gene-tia012005, gene-tia005298) and five (gene-tia017675, gene-tia017698, novel.2534, novel.2739, gene-tia015278) were identified to be related to thiamine metabolism, riboflavin metabolism, vitamin B6 metabolism, nicotinate, and nicotinamide metabolism, pantothenate and CoA biosynthesis and biotin metabolism, respectively (Dataset S6).

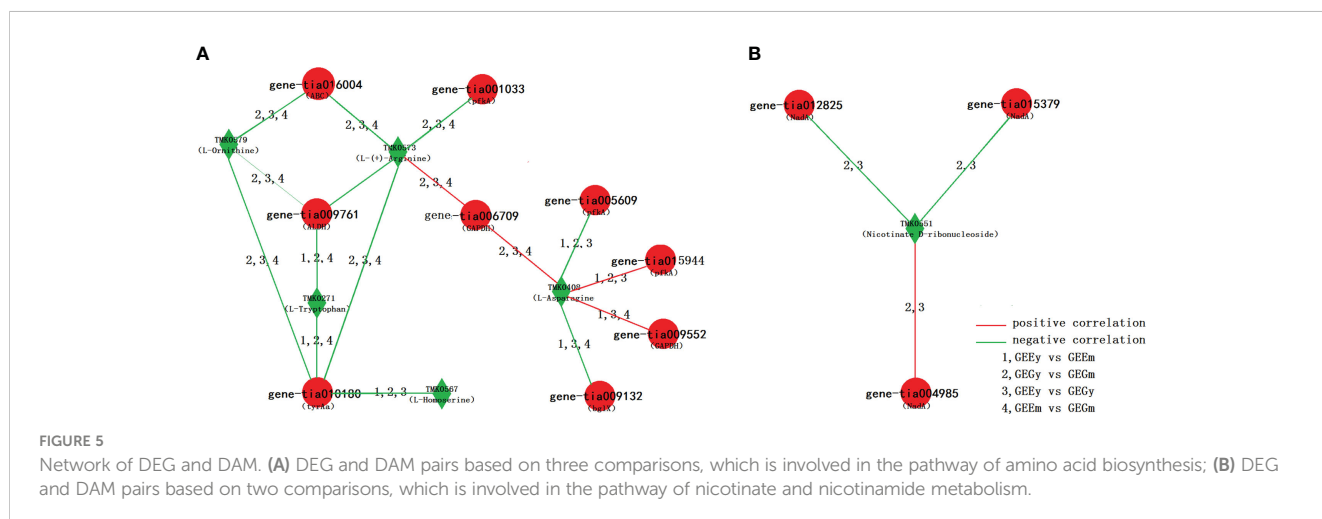## 2.6 Correlation and network expression of DAMs and DEGs

Many pairs of DEGs and DAMs, based on the comparisons of *GEGy* and *GEGy*, *GEEm* and *GEGm*, *GEEy* and *GEEm*, and *GEGy* and *GEGm*, were found to be involved in the same pathways and showed a significant relationship (Dataset S8). 732, 76, and one pair appeared in two, three, and four comparative pairwise analyses. It showed the same correlation trend, implying a stable one-way or two-way direction in the positive or negative expression regulation between them in the tubers of *GE* (Dataset S9). Of 76 pairs based on three comparisons, 16 DEG-DAM pairs consisted of five metabolites (TMK0579:l-ornithine, TMK0408:l-asparagine, TMK0573:l-(+)-arginine; TMK0271:l-tryptophan; TMK0567:l-homoserine) and nine genes (gene-tia001033, gene-tia005609 and gene-tia015944 coding pfkA, gene-tia009552 and gene-tia006709 coding GAPDH, gene-tia009132 coding bglX, gene-tia009761 coding ALDH, gene-tia010180 coding tyrAa, gene-tia016004 coding ABC) were identified to be related to the biosynthesis of amino acid, four pairs, including gene-tia006709-TMK0408, gene-tia006709-TMK0573, gene-tia009552-TMK0408 and gene-tia015944-TMK0408 show positive correlation and other 12 pairs (e.g., gene-tia001033-TMK0579, gene-tia005609-TMK0408, gene-tia009761-TMK0271, gene-tia010180-TMK0567, gene-tia016004-TMK0573) show a negative correlation (Figure 5A; Dataset S9). Of 732 pairs based on two comparisons, Three pairs of DEG-DAM consisted of one metabolite (TMK0551:nicotinate d-

ribonucleoside) and three genes (gene-tia015379 and gene-tia012825 code NadA and gene-tia004985 coding uridine nucleosidase) based on two comparisons were found to involved in nicotinate and nicotinamide metabolism, pair of gene-tia004985 and TMK0551 shows a positive correlation, pairs of gene-tia015379 and TMK0551, and gene-tia012825 and TMK0551 show a negative correlation (Figure 5B; Dataset S9).

3 Discussion

3.1 GE is an excellent complementary food, but it should not be consumed excessively

A total of 345 metabolites were found in *GEEy, GEEm, GEGy,* and *GEGm*. Relative to other plant tissues, the number of metabolites found in the tuber tissue of *GE* is lower when determined using the same method. For example, there are approximately 600 metabolites found in the leaf, flower, root, and fruit of *Eriobotrya japonica* (Wang, 2021), and 661 are found in the different development stages of *Chrysanthemum morifolium* (Wang et al., 2019). In chestnuts (*Castanea mollissima* Bl.), 611 metabolites were found to have varying degrees of calcification (Xiao et al., 2021). In the young shoots of Albino Tea cultivars, 839 metabolites have been found (Wang et al., 2020), with 499 metabolites found in *Allium fistulosum–A. cepa* (Abdelrahman et al., 2019), and 637 were found in different tissues of *Salsola collina* Pall (Li et al., 2021). We speculate that this is because *GE* adopts a heterotrophic lifestyle (Park and Lee, 2013), and some metabolic pathways in autotrophic plants fail to function in *Gastrodia elata*, so the corresponding metabolites are also reduced. Seventy-six types of amino acids, including eight kinds of amino acids (l-valine, l-(+)-lysine, l-leucine, l-isoleucine, tryptophan, l-(-)-threonine, l-phenylalanine, and l-methionine) that the human body cannot synthesize and that must be supplemented from the diet, and 13 type vitamins were found in this study, which revealed that *GE* contained excellent auxiliary ingredients. Besides 39 phenolic acids and 15 flavonoids containing potential healthful ingredients (e.g., gastrodin, cinnamic acid, narirutin, tangerine), 34 alkaloids were also identified in the

FIGURE 5
Network of DEG and DAM. **(A)** DEG and DAM pairs based on three comparisons, which is involved in the pathway of amino acid biosynthesis; **(B)** DEG and DAM pairs based on two comparisons, which is involved in the pathway of nicotinate and nicotinamide metabolism.

tuber of *GE*, which contain some higher accumulated metabolites, such as choline and spermine. Choline is a very effective healthful-promoting compound (Wecker et al., 1982; Leermakers et al., 2015), and spermine is a necessary component of all vertebrate cells, but toxic effects will be produced if it accumulates in cells (Seiler et al., 2000). This investigation demonstrates that the spermine contents in all four samples were relatively higher. As a result, the *GE* might not be suitable for excessive food consumption.

3.2 *GEGm* excels over *GEEy, GEEm,* and *GEGy* for supplementing amino acids

In total, 335 to 339 metabolites were identified from young and mature tubers of *GEE* and *GEG*, respectively, suggesting slight variations in the metabolites among different *GE* varieties and the various development stages of *GE*. This is similar to the results of Xiong et al. (2014) and Yang et al. (2016b). The quality of *GE* is mainly reflected in the content of specific nutritional and medicinal components. The types and abundance of amino acids are important indicators of the nutritional quality of *GE* (Lei et al., 2015). There are conflicting views on whether *GE, GEG* or *GEE* present better medicinal and nutritional qualities (Wu et al., 2011; Tan et al., 2012; Li et al., 2016a; Yang et al., 2016b; Wang et al., 2018). In this study, nine different amino acids, including glutamic acid, Cys–Gly, l-tryptophan, and glutathione (reduced form), showed higher content in mature tubers than in young tubers. Only one amino acid, alanyl-leucine, showed lower abundance in mature tubers than in young tubers. Therefore, if the nutrient quality of *GE* is judged based on the content of amino acids, the mature tuber of *GEGm* excels over that of the other tuber samples. Another key nutrient element in supplementary food is vitamins. Of identified 13 type vitamins in four samples of *GE*, 10 are no-DAM metabolites based on four comparisons, only nicotinic acid in *GEGm* is higher than that of *GEEm*, and nicotinate d-ribonucleoside in *GEGy* is higher than that of *GEEy*, nicotinate d-ribonucleoside and riboflavin in *GEGy* are higher than that in *GEGm*. These results show that *GEG*, especially *GEGy*, has a slight advantage over *GEE*.

3.3 First insights into the molecular basis of amino acids and vitamins that related to the nutritional properties

The main objective of pharmacognosy is to uncover the molecular basis and genetic mechanism of bioactive ingredients with potential health values in medicinal plants. In recent years, attempts have been made to uncover the genetic mechanism of phenols, including gastrodin found in *GE*, by comparing the transcriptomes from different developmental stages of *GE* tubers. For example, Tsai et al. (2016) identified two unigenes encoding glycosyltransferase and monooxygenase as the molecular basis of gastrodin biosynthesis in the tubers of *GE*. Genes code eleven key enzymes, including CM, PDT, PAL, 4CL, CHS, F5H, F3'M, C4H,

CCR, and CAD, are involved in the phenylpropanoid metabolic pathway and are most likely responsible for the molecular mechanism of the medicinal quality formation of *GE*, were detected (Wen et al., 2017). Seventy-six candidate genes encoding eight key enzymes involved in the *GE* phenolics biosynthesis pathway were identified (Shan et al., 2021). Using co-expression analysis, Chen and Mi (2021) showed that 4-hydroxybenzyl alcohol mainly originates from phenylalanine metabolism. Polysaccharides are another group of critical active ingredients found in *GE* (Zhu et al., 2019) and have been found to exhibit various activities, including anti-cancer (Zhao et al., 2005), antioxidation, and anti-aging (Liu and Mori, 1992; Xie et al., 2010) activities; immunomodulatory effects (Li et al., 2016b); neuroprotection (Zhou et al., 2018); and cardiovascular system activities (Kim et al., 2012). Previous studies show that the polysaccharide composition of *GE* included xylose, glucose, galactose, rhamnose, and mannose, with glucose accounting for the vast majority (Zhao et al., 2005; Qiu et al., 2007; Li et al., 2008; Chen et al., 2011; Lee et al., 2012; Ming et al., 2012; Bao et al., 2017). These monosaccharides are linked by 1→6 and 1→4 glycosidic bonds, or 1→2 glycosidic bonds (Zhou et al., 2009), with 1→4 glycosidic bonds being the most common (Zhao et al., 2005; Qiu et al., 2007; Ming et al., 2012). Previous studies on *GE*'s bioactive substances and corresponding molecular basis mainly focused on the health-promoting components.

Until now, investigations on nutrient components and molecular mechanisms underlying *GE* have been limited. Under the new background that *GE* was officially listed drug and food source by the Chinese government, it is particularly urgent to reveal the molecular basis related to the nutritional characteristics of *GE*. Here, we detected 76 different amino acids and their derivatives, 13 vitamins, and some other metabolism endow the nutritional properties. We also identified many genes that code enzymes responsible for the biosynthesis or metabolism of some amino acids (lysine, histidine, phenylalanine, tyrosine, tryptophan) and vitamins (nicotinate, nicotinamide, thiamine). Meanwhile, comparing the integrative analyses of the metabolomes and transcriptomes of different tuber samples pairwise, many DEG and DAM pairs were annotated into the same metabolism pathways. Some of these pairs exhibit a strong positive or negative correlation with a similar trend, indicating that there is a consistent positive or negative regulation between them. Additionally, it was found that certain vitamins and amino acids directly promote or inhibit the expression of the genes involved in the biosynthesis of these metabolites.

4 Conclusions

A total of 345 metabolites and 21,513 transcripts were detected in *GEEy, GEEm, GEGy* and *GEGm*. These metabolites show that fresh *GE* tissues make a great supplemental diet since they contain 76 amino acids and their derivatives, which contain all of the human necessary amino acids (such as l-(+)-lysine and l-leucine) and 13 vitamins. However, consuming too much *GE* could be

dangerous because of its high spermine content. Of *GEEy*, *GEEm*, *GEGy*, and *GEGm*, *GEGm* has higher amino acid ingredients. Meanwhile, some genes metabolizing some amino acids and vitamins were identified. There were also a few whose up- or down-regulated expression would considerably encourage or prevent the accumulation of various vitamins and amino acids. In general, the data and analysis in this work represent a significant step forward in understanding the molecular basis, genetic basis, and regulatory mechanism of GE nutritional characteristics, which will aid in the rational consumption of *GE* industry and deepen our understanding of *GE* biology.

5 Materials and methods

5.1 Plant material sampling and tissue pretreatment

Young (length=~3cm, diameter=~2cm, the terminal bud is not apparent) and mature (length=~10cm, diameter=~25cm, the terminal bud is obvious) tubers of a GEG strain, grown at an altitude of around 2300 meters, were taken from Xiaocaoba town, Yiliang county, Yunnan province. The young (length=~3cm, diameter=~2cm, the terminal bud is not obvious) and mature (length=~12cm, diameter=~20cm, the terminal bud is obvious) tuber of a GEE strain, grown at an altitude of around 1500 meters, were taken from Niuzhuang town, Yiling county, Hubei Province. Both strains were cultivated in sandy soil rich in humus under broad-leaved forests before sampling. Both sample collection sites belong to the humid climate zone of the central subtropical zone, with mountainous terrain. The average annual temperature and precipitation in Xiaocaoba are 15°C and 1100 mm, while in Niuzhuang they are 10°C and 1600 mm, respectively. In order to avoid contamination from the tissues of symbiotic Armillaria mellea, the tubers were stripped of the surface cortex before the tissue within the epidermal layer was freeze-dried and ground into powder. One hundred milligrams of powder from each sample was weighed and used for total RNA extraction. Then, another 100 mg of powder from each sample was dissolved in 0.6 ml of 70% methanol extract. The dissolved samples were refrigerated overnight at 4°C and centrifuged at 10,000 rpm. The supernatant was filtered with a microporous membrane (0.22 μm pore size) and stored in an injection bottle, ready for UPLC-MS/MS analysis. Three biological repeats were carried out for each tissue sample. The remaining sample tissues not utilized in this experiment were kept at Kaili University's -80°C refrigerator.

5.2 Metabolite data acquisition using UPLC-MS

Ultra-high Performance Liquid Chromatography (UHPLC) (Shim-pack UFLC SHIMADZU CBM30A; https://www.shimadzu.com.cn/) and tandem mass spectrometry (Applied

Biosystems 4500 qtrap; http://www.appliedbiosystems.com.cn/) were used to collect metabolite data. The UHPLC conditions were as follows: chromatographic column (Waters ACQUITY UPLC HSS T3 C18 1.8 μm, 2.1 mm * 100 mm); mobile phase A (ultra-pure water with 0.04% acetic acid); mobile phase B (acetonitrile with 0.04% acetic acid); elution gradient (the starting concentration of phase B was 5%; phase B was increased linearly to 95% within 10.00 min and maintained at 95% for 1.0 min; phase B decreased to 5% at 11.00–11.10 min and maintained at 5% for 14 min. Flow rate = 0.35 ml/min; column temperature = 40°C; and injection volume = 4 μL). The mass spectrum conditions were set as follows: the temperature of electrospray ionization was set to 550°C, the voltage was set to 5500 V, the curtain gas pressure was fixed at 30 psi, and the parameter of collision-activated dissociation was set to high. In the triple quadrupole (QQQ), each ion pair was scanned and detected according to the optimized declustering potential and collision energy (Chen et al., 2013).

5.3 Qualitative and quantitative analysis of metabolites

The self-built metware database (MWDB) by Metware Biotechnology Co., Ltd. (Wuhan, China) was used to qualitatively analyze the metabolites based on secondary spectrum information. Repeated signals, including $NH4^+$ ion, Na^+ ion and K^+ ion, and the repeated signal of fragment ions for larger molecular weight substances were removed. Metabolite quantification was performed via multiple reaction monitoring mode analysis using triple quadrupole mass spectrometry (Fraga et al., 2010).

5.4 Metabolite data processing and analysis

The mass spectrum data were handled using Software Analyst 1.6.3. PCA, clustering, repetitive correlation, and differential expression analysis were performed using the MetaboAnalystR software package (Chong and Xia, 2018). The KEGG database was used to annotate and display the biosynthetic pathways of different metabolites.

5.5 RNA extraction, library construction, and sequencing

The total RNA from tuber tissues was extracted using an RNA kit (Thermo Fisher Scientific Inc, Waltham, MA, USA) following the guidelines provided by the manufacturer. The concentration and quality of the extracted total RNA were determined using Nanodrop 2000 (Thermo Fisher Scientific, USA). The integrity of the RNA samples was evaluated by agarose gel electrophoresis. The total RNAs with concentrations ≥ 300ng/μL, OD260/280 = 1.8–2.0 were used to construct the sequencing library as follows: (1) mRNA was extracted from the total RNA using Oligo (dT) magnetic beads;

(2) approximately 300 bp length mRNA fragment sequences were obtained, and first strand of cDNA was synthesized; (3) the RNA strand was degraded by RNaseH before the second strand of cDNA was synthesized with dNTPs with the help of a DNA polymerase I system; (4) purified double-stranded cDNA gel electrophoresis was performed, and the end of the cDNA was repaired; (5) cDNA was screened for 250–300 bp lengths, and polymerase chain reaction (PCR) products were refined using Ampure XP beads; (6) and finally, the library was obtained. More than 6G of raw data per library were generated using a Hiseq 4000 sequencer (Illumina, San Diego, CA, USA).

5.6 RNA sequencing data filtering and statistics

The raw sequencing data were filtered by removing the reads containing an adaptor sequence, paired-end reads with an N content exceeding 10% of the length proportion in either single read and low-quality (≤ 5) base numbers exceeding 50% of the length proportion in either single read. The remaining high-quality clean data were used for analyses. Sequencing reads the sequencing error rate, the Q20 ratio, Q30 ratio, and GC content of the raw and clean data outputs were counted using a custom Pearl script developed by Maiwei Co., Ltd (https://www.metware.cn/). The raw transcriptome data of twelve samples have been deposited in GenBank (https://www.ncbi.nlm.nih.gov/) (Accessible ID: PRJNA825667).

5.7 Clean RNA sequencing data mapping and gene expression statistics

Clean RNA reads with > Q20 for *GEEm*, *GEEy*, *GEGm*, and *GEGy* were mapped onto the assembled reference genome of *G. elata Bl.f.elata* assembled (Wang and Shahid, 2023) using HISAT v2.0.5 (https://arc-ts.umich.edu/software-item/hisat2/) to generate reference transcripts. Subsequently, the clean reads were mapped onto the transcripts using Bowtie2 v2.3.4.1 (https://wiki.rc.usf.edu/index.php/Bowtie2#Version). If a transcript sequence mapped onto the predicted protein-coding region of the reference genome, this was regarded as an existing gene; otherwise, the sequence was regarded as a new gene. Based on the mapping results produced by Bowtie2, the number of reads mapped on each transcript in each sample was counted. Then, the FPKM (fragments per kilobase per million bases)-transformed results were generated using RSEM software (Li and Dewey, 2011). The expression abundance of protein-coding genes or their transcripts was obtained. The PCA and Pearson's correlation analysis of the biological repeats based on total gene expression was performed using the R software package (https://www.r-project.org/). DEGs between sample groups were analyzed using the software DESeq2 (Love et al., 2014).

5.8 Functional annotation of transcriptome and DEGs by KEGG and GO enrichment analysis

The transcriptome, including DEGs, were annotated for molecular function using various databases, including Nr (https://www.ncbi.nlm.nih.gov/refseq/), SwissPort (http://www.ebi.ac.uk/swissprot/), KOG (https://www.ncbi.nlm.nih.gov/COG/), GO (http://geneontology.org/), and KEGG using BLAST v2.10.0 software (https://www.ncbi.nlm.nih.gov/books/NBK131777/) with $1e^{-5}$ as the threshold of the e-value. Based on the GO and KEGG annotation, GO enrichment and KEGG enrichment analyses of DEGs were further carried out using GOseq (Young et al., 2010) and KOBAS (Mao et al., 2005) software, respectively.

5.9 Combined analysis of transcriptomes and metabolomes

Firstly, a combined KEGG annotated pathway was produced according to this experiment's differential metabolite analysis results and the transcriptome differential gene analysis results. In other words, the differential genes and metabolites of the same paired sample groups for comparison were simultaneously mapped to the KEGG pathway map to understand the expression relationship between metabolites and genes. Secondly, the cor program in R software was used to estimate Pearson's correlation coefficient of differential genes and differential metabolites. Finally, the metabolites and genes with Pearson correlation coefficients greater than 0.8 in each different group in each pathway were analyzed by canonical correlation analysis using R software (https://www.r-project.org/).

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found at the NCBI, accession number: PRJNA825667.

Author contributions

YW designed the experiments and collected samples. YW and MS analyzed data and wrote the paper. All authors contributed to the article and approved the submitted version.

Funding

Acknowledgments

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.

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.2023.1183139/full#supplementary-material

SUPPLEMENTARY DATA SHEET 7
List of differentially expressed genes (DEGs).

SUPPLEMENTARY DATA SHEET 9
Network correlation of DEGs and DAMs involved in same KEGG pathways.

References

Abdelrahman, M., Hirata, S., Sawada, Y., Hirai, M. Y., Sato, S., Hirakawa, H., et al. (2019). Widely targeted metabolome and transcriptome landscapes of *Allium fistulosum–A. cepa* chromosome addition lines revealed a flavonoid hot spot on chromosome 5A. *Sci. Rep.* 9, 3541. doi: 10.1038/s41598-019-39856-1

An, S. J., Park, S. K., Hwang, I. K., Choi, S. Y., Kim, S. K., Kwon, O. S., et al. (2003). Gastrodin decreases immunoreactivities of gamma-aminobutyric acid shunt enzymes in the hippocampus of seizure-sensitive gerbils. *Neurosci. Res.* 71, 534_543. doi: 10.1002/jnr.10502

Bao, Q. W., Qian, L., Gong, C., and Shen, X. Z. (2017). Immune-enhancing activity of polysaccharides from *Gastrodia elata*. *J. Food. Process Pres.* 41, e13016. doi: 10.1111/jfpp.13016

Chen, X., Cao, D. X., Zhou, L., Jin, H. Y., Dong, Q., Yao, J., et al. (2011). Structure of a polysaccharide from *Gastrodia elata*., and oligosaccharides prepared thereof with anti-pancreatic cancer cell growth activities. *Carbohydr. Polym.* 86, 1300_1305. doi: 10.1016/j.carbpol.2011.06.029

Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., et al. (2013). A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. *Mol. Plant* 6, 1769_1780. doi: 10.1093/mp/sst080

Chen, C., and Mi, H. (2021). Analysis of genes related to biosynthesis of gastrodin. *Plant Physiol. J.* 57, 1819_1828.

Chong, J., and Xia, J. (2018). MetaboAnalystR: an r package for flexible and reproducible analysis of metabolomics data. *Bioinformatics* 34, 4313_4314. doi: 10.1093/bioinformatics/bty528

Feng, X. Z., Chen, Y. W., and Yang, J. S. (1979). Chemical constituents of *Gastrodia elata* blume. *ATCA Chimca Sin.* 37, 175_181.

Fraga, C. G., Clowers, B. H., Moore, R. J., and Zink, E. M. (2010). Signature-discovery approach for sample matching of a nerve-agent precursor using liquid chromatography-mass spectrometry, XCMS, and chemometrics. *Anal. Chem.* 82, 4165_4173. doi: 10.1021/ac1003568

Ha, J. H., Lee, D. U., Lee, J. T., Kim, J. S., Yong, C. S., Kim, J. A., et al. (2000). 4-hydroxybenzaldehyde from *Gastrodia elata* B1. is active in the antioxidation and GABAergic neuromodulation of the rat brain. *J. Ethnopharmacol.* 73, 329_333. doi: 10.1016/s0378-8741(00)00313-5

Han, R., Rai, A., Nakamura, M., Suzuki, H., Takahashi, H., Yamazaki, M., et al. (2016). *De novo* deep transcriptome analysis of medicinal plants for gene discovery in biosynthesis of plant natural products. *Methods Enzymol.* 576, 19_45. doi: 10.1016/bs.mie.2016.03.001

Hsieh, C. L., Chang, C. H., Chiang, S. Y., Li, T. C., Tang, N. Y., Pon, C. Z., et al. (2000). Anticonvulsive and free radical scavenging activities of vanillyl alcohol in ferric chlorideinduced epileptic seizures in sprague-dawley rats. *Life Sci.* 67, 1185_1195. doi: 10.1016/s0024-3205(00)00706-2

Hsieh, M. T., Wu, C. R., and Chen, C. F. (1997). Gastrodin and p-hydroxybenzyl alcohol facilitate memory consolidation and retrieval, but not acquisition, on the passive avoidance task in rats. *J. Ethnopharmacol.* 56, 45_54. doi: 10.1016/s0378-8741(96)01501-2

Hwang, S. M., Lee, Y. J., Kang, D. G., and Lee, H. S. (2009). Anti-inflammatory effect of *Gastrodia elata* rhizome in human umbilical vein endothelial cells. *Am. J. Chin. Med.* 37, 395_406. doi: 10.1142/S0192415X09006916

Jaswir, I., Octavianti, F., Lestari, W., Hendri, R., and Ahmad, H. (2017). Some characteristics and functional properties of chunma (*Gastrodia elata*) as a food supplement: a short review. *Int. Food Res. J.* 24, S274_S280.

Jung, T. Y., Suh, S. I., Lee, H., Kim, I. S., Kim, H. J., Yoo, H. S., et al. (2007). Protective effects of several components of *Gastrodia elata* on lipid peroxidation in gerbil brain homogenates. *Phytother. Res.* 21, 960_964. doi: 10.1002/ptr.2193

Kim, H. J., Hwang, I. K., and Won, M. H. (2007). Vanillin, 4-hydroxybenzyl aldehyde and 4-hydroxybenzyl alcohol prevent hippocampal CA1 cell death following global ischemia. *Brain Res.* 1181, 130_141. doi: 10.1016/j.brainres.2007.08.066

Kim, K. J., Lee, O. H., Han, C. K., Kim, Y. C., and Hong, H. D. (2012). Acidic polysaccharide extracts from *Gastrodia rhizomes* suppress the atherosclerosis risk index through inhibition of the serum cholesterol composition in sprague dawley rats fed a high-fat diet. *Int. J. Mol. Sci.* 13, 1620_1631. doi: 10.3390/ijms13021620

Lee, O. H., Kim, K. I., Han, C. K., Kim, Y. C., and Hong, H. D. (2012). Effects of acidic polysaccharides from *Gastrodia rhizome* on systolic blood pressure and serum lipid concentrations in spontaneously hypertensive rats fed a high-fat diet. *Int. J. Mol. Sci.* 13, 698_709. doi: 10.3390/ijms13010698

Leermakers, E. T. M., Moreira, E. M., Jong, J. C. K., Darweesh, S. K. L., Visser, T., Voortman, T., et al. (2015). Effects of choline on health across the life course: a systematic review. *Nutr. Rev.* l73, 500_522. doi: 10.1093/nutrit/nuv010

Lei, Y., Li, J., Xiao, J., Yue, Y., Zhang, X., Zhang, M., et al. (2015). Correlation between contents of gastrodin and polysaccharides with grade of *Gastrodia* tuber. *Chin. Herb. Med.* 46, 418_423. doi: 10.7501/j.issn.0253-2670.2015.03.021

Li, S., Chen, Y., Duan, Y., Zhao, Y., Zhang, D., Zang, L., et al. (2021). Widely targeted metabolomics analysis of different parts of *Salsola collina* pall. *Molecules* 26, 1126. doi: 10.3390/molecules26041126

Li, B., and Dewey, C. N. (2011). RSEM accurate transcript quantification from RNA-seq data with or without a reference genome. *BMC Bioinf.* 12, 323. doi: 10.1186/1471-2105-12-323

Li, Y., Wang, Z., Geng, Y., Zhou, H., Liu, D., and Wang, X. (2016a). Determination of chemical constitutes in *Gastrodia elata. f. glauca* and *Gastrodia elata. f. elata* by HPLC-ESI-TOF/MS. *Nat. Prod. Res. Dev.* 28, 1758_1763.

Li, C., Wang, J. R., Ji, X. H., and Lu, X. L. (2008). Isolation of *Gastrodia elata* polysaccharides and analysis of its composition of monosaccharide. *Chin. Agric. Sci. Bull.* 24, 89_92.

Li, X. B., Zhan, J. P., Zhang, Y. T., Xie, Z. L., Zhu, Y. Q., Chen, Y. L., et al. (2016b). Effect of polysaccharide from *Gastrodia elata* on humoral immune function in immunosuppressed mice induced by cyclophosphamide. *Chin. J. Gerontol.* 36, 1027_1028.

Lin, Y. E., Lu, K. H., and Sheen, L. Y. (2018). Recent research progress on the antidepressant-like effect and neuropharmacological potential of *Gastrodia elata* blume. *Curr. Pharmacol. Rep.* 4, 220_237. doi: 10.1007/s40495-018-0137-2

Liu, Y., and Huang, G. (2017). The chemical composition, pharmacological effects, clinical applications and market analysis of *Gastrodia elata*. *Pharm. Chem. J.* 51, 211_215. doi: 10.1007/s11094-017-1584-5

Liu, J., and Mori, A. (1992). Antioxidant and free radical scavenging activities of *Gastrodia elata* bl. and *Uncaria rhynchophylla* (Miq.) jacks. *Neuropharmacology* 31, 1287_1298. doi: 10.1016/0028-3908(92)90058-w

Liu, X. K., and Yang, Y. (1958). Studies on the constituents of *Gastrodia elata* blume. i. extraction and identification of vanilla alcohol. *Fudan Univ. J. Med. Sci.* 1958, 67_68.

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

Lu, K. H., Ou, G. L., Chang, H. P., Chen, W. C., Liu, S. H., and Sheen, L. Y. (2020). Safety evaluation of water extract of *Gastrodia elata* blume: genotoxicity and 28-day oral toxicity studies. *Regul. Toxicol. Pharmacol.* 114, 104657. doi: 10.1016/j.yrtph.2020.104657

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

Makni, M., Chtourou, Y., Fetoui, H., Garoui, E. M., Boudawara, T., and Zeghal, N. (2011). Evaluation of the antioxidant, anti -inflammatory and hepatoprotective properties of vanillin in carbon tetrachloride-treated rats. *Eur. J. Pharmacol.* 68, 133_139. doi: 10.1016/j.ejphar.2011.07.001

Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. J. B. (2005). Automated genome annotation and pathway identification using the KEGG orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787_3793. doi: 10.1093/bioinformatics/bti430

Ming, J., Liu, J., Wu, S. R., Guo, X. H., Chen, Z. D., and Zhao, G. H. (2012). Structural characterization and hypolipidemic activity of a polysaccharide PGEB-3H from the fruiting bodies of *Gastrodia elata* blume. *Procedia. Eng.* 37, 169_173. doi: 10.1016/j.proeng.2012.04.221

Ojemann, L. M., Nelson, W. L., Shin, D. S., Rowe, A. O., and Buchanan, R. (2006). Tian ma, an ancient Chinese herb, offers new options for the treatment of epilepsy and other conditions. *J. Epilepsy Behavior.* 8, 376_383. doi: 10.1016/j.yebeh.2005.12.009

Park, E., and Lee, W. (2013). *In vitro* symbiotic germination of myco-heterotrophic *Gastrodia elata* by *Mycena* species. *Plant Biotech. Rep.* 7, 185_191. doi: 10.1007/s11816-012-0248-x

Qiu, F., Liu, T. T., Qu, Z. W., Qiu, C. Y., Yang, Z., and Hu, W. P. (2014). Gastrodin inhibits the activity of acid-sensing ion channels in rat primary sensory neurons. *Eur. J. Pharmacol.* 731, 50_57. doi: 10.1016/j.ejphar.2014.02.044

Qiu, H., Tang, W., Tong, X. K., Ding, K., and Zuo, J. P. (2007). Structure elucidation and sulfated derivatives preparation of two α-d-glucans from *Gastrodia elata* and their anti-dengue virus bioactivities. *Carbohydr. Res.* 342, 2230_2236. doi: 10.1016/j.carres.2007.06.021

Saito, K. (2018). Development of plant metabolomics and medicinal plant genomics. *Yakugaku Zasshi* 138, 1_18. doi: 10.1248/yakushi.17-00193

Seiler, N., Duranton, B., Gossé, F., and Raul., F. (2000). Spermine cytotoxicity to human colon carcinoma-derived cells (CaCo-2). *Cell Biol. Toxicol.* 16, 117_30. doi: 10.1023/a:1007642126765

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

Shu, G., Yang, T., Wang, C., Su, H., and Xiang, M. (2013). Gastrodin stimulates anticancer immune response and represses transplanted H22 hepatic ascitic tumor cell growth: involvement of NF-κB signaling activation in CD4+ T cells. *Toxicol. Appl. Pharmacol.* 269, 270_279. doi: 10.1016/j.taap.2013.02.019

Sun, W., Miao, B., Wang, X. C., Duan, J. H., Ye, X., Han, W. J., et al. (2012). Gastrodin inhibits allodynia and hyperalgesia in painful diabetic neuropathy rats by decreasing excitability of nociceptive primary sensory neurons. *PloS One* 7, e396476. doi: 10.1371/journal.pone.0039647

Tan, B., Kong, H., and Tao, J. (2012). Chromatographic classification & quality assessment of *Gastrodia* tuber. *Chem. Bioengeering* 29, 85_90.

Tang, C., Wang, L., Liu, X., Cheng, M., and Xiao, H. (2015). Pharmacokinetic study of *Gastrodia elata* in rats. *Anal. Bioanal. Chem.* 407, 8903_8910. doi: 10.1007/s00216-015-9054-y

Trócsányi, E., György, Z., and Zámboriné-Németh, É. (2020). New insights into rosmarinic acid biosynthesis based on molecular studies. *Curr. Plant Biol.* 23, 100162. doi: 10.1016/j.cpb.2020.100162

Tsai, C. C., Wu, K. M., Chiang, T. Y., Huang, C. Y., Chou, C. H., Li, S. J., 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. S. (2021). A draft genome, resequencing, and metabolomes reveal the genetic background and molecular basis of the nutritional and medicinal properties of loquat (*Eriobotrya japonica* (Thunb.) lindl). *Hortic. Res.* 8, 231. doi: 10.1038/s41438-021-00657-1

Wang, Q., Li, D., Chen, A., Hou, K., Li, X., and Wu, W. (2018). Comparison on main chemical composition contents in three variants of *Gastrodia elata* from southwest China. *Zhong Cao Yao* 49, 2646_2652.

Wang, Y. S., and Shahid, M. Q. (2023). Genome sequencing and resequencing identified three horizontal gene transfers and uncovered the genetic mechanism on the intraspecies adaptive evolution of gastrodia elata blume. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.1035157

Wang, P., Zheng, Y., Guo, Y., Liu, B., Jin, S., Liu, S., et al. (2020). Widely targeted metabolomic and transcriptomic analyses of a novel albino tea mutant of "Rougui". *Forests* 11, 229. doi: 10.3390/f11020229

Wang, T., Zou, Q., Guo, Q., Yang, F., Wu, L., and Zhang, W. (2019). Widely targeted metabolomics analysis reveals the effect of flooding stress on the synthesis of flavonoids in *Chrysanthemum morifolium*. *Molecules* 24, 3695. doi: 10.3390/molecules24203695

Wecker, L., Flynn, C. J., Stouse, M. R., and Trommer, B. A. (1982). Choline availability: effects on the toxicity of centrally active drugs. *Drug Nutr. Interact.* 1, 125_130.

Wen, H., Zhang, D., Peng, C., Wang, W., Xie, X., Rao, C., et al. (2017). Transcriptional analysis of phenylpropanoid metabolic pathway in *Gastrodia elata*. *J. Chin. Med. Mat.* 40, 789_795.

Wu, Z. H., Liu, L. M., Pan, J. R., and Wang, S. B. (2011). Determination of gastrodin in hybrid varieties of *Gastrodia elata* by HPLC. *Chin. J. Ethnomed. Ethnopharm.* 20, 41_41.

Xiao, J., Gu, C., He, S., Zhu, D., Huang, Y., and Zhou, Q. (2021). Widely targeted metabolomics analysis reveals new biomarkers and mechanistic insights on chestnut (*Castanea mollissima* bl.) calcification process. *Food Res. Int.* 141, 110128. doi: 10.1016/j.foodres.2021.110128

Xie, X. Y., Chao, Y. M., Du, Z., and Zhang, Y. (2010). Effects of polysaccharides from *Gastrodia elata on* anti-aging of ageing mice. *Pharm. J. Chin. PLA.* 26, 206_209.

Xiong, L., Zhao, F., Wang, R., Qi, J., Zhang, Z., and Luo, Y. (2014). Analysis of volatile components in four *Gastrodia elata* varieties. *J. Zhejiang Agr. Sci.* 348, 1364_1367.

Xu, C. B., Guo, Q. L., Wang, Y. N., Lin, S., Zhu, C. G., and Shi, J. G. (2019). Gastrodin derivatives from *Gastrodia elata*. *Nat. Prod. Bioprospect.* 9, 393_404. doi: 10.1007/s13659-019-00224-1

Yang, Z., Deng, W., Zong, L., Xu, D., Li, Z., and Jiang, Y. (2016b). HPLC fingerprint of the different producing areas and different varieties of *Gastrodia elata*. *West China J. Pharm. Sci.* 32, 188_191.

Yang, J., Xiao, Q., Xu, J., Da, L., Guo, L., Huang, L., et al. (2020). GelFAP: gene functional analysis platform for *Gastrodia elata*. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.563237

Yang, L., Yang, C. Q., Li, C. Y., Zhao, Q., Liu, L., Fang, X., et al. (2016a). Recent advances in biosynthesis of bioactive compounds in traditional Chinese medicinal plants. *Sci. Bull.* 61, 3_17. doi: 10.1007/s11434-015-0929-2

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. J. G. B. (2010). Gene ontology analysis for RNA-seq: accounting for selection bias. *Genome Biol.* 11, R14–R14. doi: 10.1186/gb-2010-11-2-r14

Yu, S. J., Kim, J. R., Lee, C. K., Han, J. E., Lee, J. H., Kim, H. S., et al. (2005). Gastrodia elata blume and an active component, p-hydroxybenzyl alcohol reduce focal ischemic brain injury through antioxidant related gene expressions. *Biol. Pharm. Bull.* 28, 1016–1020. doi: 10.1248/bpb.28.1016

Zhan, H. D., Zhou, H. Y., Du, X. L., Sui, Y. P., Wang, W., Dai, L., et al. (2016). The rhizome of gastrodia elata blume - an ethnopharmacological review. *J. Ethnopharmacol.* 189, 361–385. doi: 10.1016/j.jep.2016.06.057

Zhao, G. H., Kan, J. Q., Li, Z. X., and Chen, Z. D. (2005). Characterization and immunostimulatory activity of an (1→6)-α-d-glucan from the root of ipomoea batatas. *Int. Immunopharmacol.* 5, 1436–1445. doi: 10.1016/j.intimp.2005.03.012

Zhou, B. H., Tan, J., Zhang, C., and Wu, Y. (2018). Neuroprotective effect of polysaccharides from gastrodia elata blume against corticosterone-induced apoptosis in PC12 cells via inhibition of the endoplasmic reticulum stress-mediated pathway. *Mol. Med. Rep.* 17, 1182–1190. doi: 10.3892/mmr.2017.7948

Zhou, X., Yang, X. H., Liang, H. X., and Liu, C. Y. (1987). *Morphology of gastrodia elata* (Beijing: Science Press).

Zhou, B. H., Yang, L., Yuan, Y., Shen, H., Feng, Q., Guo, Z. L., et al. (2009). Isolation and structure identification of an acidic and heteropolysaccharide from gastrodia elata blume. *Chin. J. Hosp. Pharm.* 29, 2002–2006. doi: 10.1016/j.carbpol.2015.03.066

Zhu, H., Liu, C., Hou, J., Long, H., Wang, B., Guo, D., et al. (2019). Gastrodia elata blume polysaccharides: a review of their acquisition, analysis, modification, and pharmacological activities. *Molecules.* 24, 2436. doi: 10.3390/molecules24132436