Food Chemistry 379 (2022) 132078

Contents lists available at ScienceDirect

Food Chemistry

journal homepage: www.elsevier.com/locate/foodchem

Identification of key genes contributing to amino acid biosynthesis in *Torreya grandis* using transcriptome and metabolome analysis

Heqiang Lou [a,1], Yi Yang [a,1], Shan Zheng [a,1], Zhenmin Ma [a], Wenchao Chen [a], Chenliang Yu [a,*], Lili Song [a,b,*], Jiasheng Wu [a,b,*]

[a] State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou, Zhejiang 311300, China
[b] NFGA Engineering Research Center for Torreya Grandis 'Merrillii', Zhejiang A&F University, Hangzhou 311300, China

ARTICLE INFO

Keywords:
Torreya grandis
Amino acid
Regulatory mechanism
Anthranilate synthase
3-Deoxy-D-arabino-heptulosonate-7-phosphate synthase

ABSTRACT

Torreya grandis has high economic and nutritional value due to the high nutrients in its kernels. The kernels of different development stages vary enormously in their amino acids content. However, the molecular basis and the regulatory mechanism of amino acid biosynthesis remain unclear. Here, transcriptome and metabolome analysis were performed. Correlation analysis result showed that 4 unigenes were significantly and positively correlated with at least 10 amino acids. The full length CDS of 2 unigenes (*TgDAHP2* and *TgASA1*) were successfully cloned from the 4 unigenes for DAHP, ASA and CITS. Subcelluar localization analysis showed that both TgDAHP2 and TgASA1 were localized to the chloroplast. Overexpression of *TgDAHP2* and *TgASA1* in Arabidopsis can greatly increase the content of most amino acids. Moreover, 3 transcription factors were found to positively regulate the expression of *TgASA1*. This research contributes to understand the molecular regulatory mechanisms of amino acid biosynthesis in *T. grandis*.

1. Introduction

Torreya grandis (*T. grandis*) 'Merrilli' belongs to the *Torreya* genus of Taxaceae. The seeds of this genus are beloved by consumers owing to their unique flavor and high nutritional and medicinal value (He et al., 2016; Yu et al., 2016). It possesses dioecious flowers and stone fruit with nut seeds (He et al., 2016). The seed development of *T. grandis* can be divided into four stages: slow-growing stage, expansion stage, material transformation stage (filling stage) and mature stage. Among them, the material transformation stage and mature stage are the key periods for accumulation of nutrients, which determines the quality of the kernels. The mature kernels are rich in oil, protein, carbohydrate, calcium, zinc, iron and other nutrients. There are 17 kinds of amino acids were identified in *T. grandis* kernels with a total content of 118.1 g/kg, and 38.61% of which are the essential amino acids (Chen et al., 1998). Therefore, the *T. grandis* is an important source of essential amino acids.

Amino acids are the basic substances that make up the proteins required for animal nutrition. There are 20 basic amino acids involved in protein synthesis, and more than 1,000 non-protein amino acids have been found in nature. It plays an extremely important role in human and animal nutrition (Rodgers, 2014). Adequate protein and essential amino acids are important for biological health. Except used as energy sources, amino acids are precursors of neurotransmitters, porphyrins, polyamines, and nitric oxide biosynthesis as well (Kaspar et al, 2009). Amino acid analysis has a wide range of applications in biochemistry, proteomics, clinical medicine, food science, chemical engineering and other fields, and is a research hotspot in analytical chemistry (Pott et al., 2014).

The content of amino acids in different plants changes significantly with different developmental stages. In the metabolome of *Arabidopsis thaliana* at different stages, the contents of glycolipids, aromatic amino acids, branched-chain amino acids and stress-induced amino acids in senescent leaves changed significantly, and asparagine/asparagine was detected in the tip region of the leaves (Watanabe et al., 2013; Jones et al., 2016). In developing and aging pea seedlings, the expression of genes associated with amino acid transport and synthesis is different, and their transport mode is also changed, resulting in significant changes in amino acids in different pea tissues (Ninan et al., 2019). A comprehensive analysis of the accumulation of metabolites in tobacco senescence leaves and changes in the expression of enzyme-encoded

* Corresponding authors at: State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou, Zhejiang 311300, China (C. Yu, L. Song, J. Wu).
E-mail addresses: YuChenliang@zafu.edu.cn (C. Yu), lilisong@zafu.edu.cn (L. Song), wujs@zafu.edu.cn (J. Wu).
[1] These authors contributed equally to the work.

https://doi.org/10.1016/j.foodchem.2022.132078
Received 30 September 2021; Received in revised form 4 January 2022; Accepted 4 January 2022
Available online 7 January 2022
0308-8146/© 2022 Elsevier Ltd. All rights reserved.

genes in the corresponding metabolic pathways showed that the related metabolism of amino acids in the leaves was significantly up-regulated (Li et al., 2017). The principal component analysis of the metabolic components of peaches showed that there were obvious metabolic changes from early to late and throughout the process after harvest mature. The contents of raffinose and galactitol, tricarboxylic acid cycle intermediates and total amino acids in young leaves of jujube are higher than those in mature leaves (Du et al., 2019).

The synthesis of amino acids in plants involves many enzymes, and different enzymes play different roles. The identification of key enzymes involved in amino acid synthesis is of great significance to our research. The shikimate pathway is related to the production of aromatic amino acids, such as phenylalanine, tyrosine and tryptophan (Tzin & Galili, 2010), which are essential for the metabolism of plants, bacteria and fungi. 3-deoxy-D-*arabino*-heptulosonate-7-phosphate synthase (DAHP) acts a vital part in regulating carbon into the shikimate pathway (Mir et al., 2015). The high protein concentration in soybean seeds and the increased nitrogen flux were related to the transcriptional expression of specific asparaginase gene *ASPGB1* (Pandurangan et al., 2012). Evidence shows that glutamine-dependent asparaginase, as the main enzyme for asparagine synthesis in higher plants, plays a key role in the deployment of N resources especially during the germination of angiosperm seeds (Cañas et al., 2007). In rice, *OsASN1* plays an important role in the asparagine metabolic pathway, and its mutants cause a large decrease in amino acid content (Luo et al., 2018). Glutamine synthetase (GS) has two main isoenzymes, which are located in different subcellular compartments and show non-overlapping effects. *GS1* is supposed to be involved in the biosynthesis of amino acid in both photosynthetic and non-photosynthetic cells (Keiki et al., 2004). During somatic embryo development of the spruce, the changes of amino acid metabolism are related to the increase of glutamine and glutamate and the decrease of arginine and alanine. At the same time, the activity of the GS/GOGAT enzyme raised during the transition from filamentous to cotyledon, confirming that the GS/GOGAT cycle is the preferred pathway for nitrogen metabolism (Joy IV et al., 1997). However, the key factors involved in amino acid synthesis in *Torreya grandis* remain unclear.

In this study, we use metabolite profiling and transcriptomic analysis to investigate the key genes and regulatory mechanism of amino acid biosynthetic pathway. We identified the candidate genes that may be key players in the amino acid biosynthesis in *T. grandis* by evolutionary tree analysis and correlation analysis between transcriptome and metabolome. We carried out transgenic experiments on key factors and analyzed the amino acids content of the transgenic lines to verify their functions. In general, our research first identified the pathway and the key genes of amino acid biosynthesis and provided the molecular basis of regulatory mechanisms for amino acid biosynthesis in *T. grandis*.

2. Materials and methods

2.1. Plant materials

Three *T. grandis* cv. 'Merrillii' grafting trees (cultivated at Dongqiao Town, Fuyang City, Zhejiang Province, China) with similar growth conditions were used in this study. The seeds were collected at three typical stages during the seed development of *T. grandis*, including the early stage of material transformation (stage 1, mid-July), the late stage of material transformation (stage 2, mid-August), and the mature stage (stage 3, mid-September). After collection, the arils and seed coats were removed and the remaining kernels were immediately frozen in liquid nitrogen and stored at −80 °C for further use. In each biological replicate, 10 seeds were ground into powder together in liquid nitrogen, and then divided into three parts, which were used for RNA sequencing, metabolite profiling and physiological assay, respectively. Three biological replicates were used for each experiment.

2.2. Metabolite extraction and profiling

Metabolites analysis was performed by the Metware Biotechnology Co. Ltd. (Wuhan, China). Extraction was carried out following the description (Zhu et al., 2018). After vacuum freeze-drying, the samples were ground to a powder. Then, 100 mg of powder was dissolved in 1.0 ml methanol, and the dissolved sample was placed in a refrigerator at 4℃ overnight. Following centrifugation at 10,000 × *g* for 10 min, the extracts were absorbed and filtrated before LC-MS analysis.

An LC-ESI-MS/MS system (UPLC, Shim-pack UFLC SHIMADZU CBM30A system, www.shimadzu.com.cn/; MS/MS, Applied Biosystems 6500 Q TRAP, www.appliedbiosystems.com.cn/) was used to analyze the extracts. A Waters ACQUITY UPLC HSS T3 C18 (1.8 μm, 2.1 mm × 100 mm) was used for chromatographic analyses. The mobile phase was acetonitrile (contained 0.04% acetic acid) (solvent B) and 0.04% acetic acid in Milli-Q water (solvent A). The elution gradient was 95:5 solvent A:B (V/V) for 0 min; 5:95 solvent A:B for 11.0 min; 5:95 solvent A:B for 12.0 min; 95:5 solvent A:B for 12.1 min; 95:5 solvent A:B for 15.0 min. The flow rate was as maintained at 0.4 ml/min, and the column temperature was 40 °C.

Linear ion trap (LIT) and triple quadrupole (QQQ) scans were acquired on a triple quadrupole-linear ion trap (Q TRAP) mass spectrometer (API 6500 Q TRAP LC/MS/MS System) equipped with an ESI Turbo Ion-Spray interface, operating in a positive ion mode and controlled by Analyst 1.6.3 software (AB Sciex). The parameters for ESI source were as follows: ion source, turbo spray; source temperature 500 °C; ion spray voltage, 5500 V; ion source gas I, gas II, and curtain gas were set at 55, 60, and 25.0 psi, respectively; the collision gas was high. Instrument tuning and mass calibration were performed with 10 and 100 μmol/L polypropylene glycol solutions in QQQ and LIT modes, respectively. QQQ scans were acquired as multiple reaction monitoring (MRM) experiments with the collision gas (nitrogen) set to 5 psi. DP and CE for individual MRM transitions were determined after further DP and CE optimization. A specific set of MRM transitions was monitored for each period based on the eluted metabolites.

2.3. MS data analysis

The metabolite content data were normalized with the range method by R software (www.r-project.org). Cluster analysis was used to analyze the accumulation patterns of metabolites among different samples. Fold change (FC) and variable importance in projection (VIP) value of orthogonal partial least squares discriminant analysis (OPLS-DA) model were used to screen the differential metabolites. Screening criteria were as follows: when $2 \leq FC \leq 0.5$ and $VIP \geq 1$ occurred, the difference was significant. The Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database was used to annotate the different metabolites. The web-based sever Metabolite Sets Enrichment Analysis (http://www.msea.ca) was used for the pathway enrichment analysis. A P-value$<$0.05 was considered to indicate significant enrichment.

2.4. Transcriptome analysis

Total RNA was extracted using total RNA extraction kit (DP441, Tiangen). The mRNA was converted into single-stranded cDNA after purified and fragmented. Next, the second strand was generated to create double-stranded cDNA. The purified double-stranded cDNA was end repaired, a tail added and sequencing adaptors connected. The 200-bp cDNA was screened using AMPure XP beads and PCR amplified. The PCR products were purified with AMPure XP beads, and finally, the library was obtained. Then each library was loaded onto the Illumina Hiseq2000 platform and high-throughput sequencing was performed.

High-quality reads were obtained after filtering the original sequencing data by the fastp software, and the transcriptome was obtained by splicing high-quality sequencing data with Trinity. The differential expression of the samples was analyzed using the DESeq2

software. The false discovery rate (FDR) was obtained by using the Benjamin–Hochberg method to correct the hypothesis test probability (P value). The screening criteria for differential gene were | log2FC |≥1 and FDR < 0.05. BLAST software was used to compare cDNA or protein sequences to the KEGG database to obtain the KO number of the genes, and then, the KEGG pathways for the genes were determined. The screening criterion for the KEGG enrichment analysis was a corrected P-value < 0.05.

2.5. Subcellular localization

Each of the full-length CDS of *TgDAHP2* and *TgASA1* without stop codon was fused to the *N*-terminus of the green fluorescent protein (GFP) gene between the BamI and *Sal*I sites under the control of the 35S promoter to generate the expression vectors. Primers used are listed in Table S1. The vectors were transiently expressed into a 45-day-old tobacco plant with a 1 ml syringe (with the needle removed) for 84 h. Fluorescent signals were detected by confocal laser scanning microscopy (LSM510, Karl Zeiss) at room temperature.

2.6. Generation of the transgenic Arabidopsis plants

The overexpression vectors 1300-ASA1 and 1300-DAHP2 were constructed and then introduced into the *Agrobacterium tumefaciens* strain GV3101, respectively. Primers used are listed in Table S1. The dipping flower method was applied to transform Arabidopsis and putatively transgenic Arabidopsis seeds were sown on MS medium with hygromycin for selecting transgenic plants.

2.7. Measurement of amino acid contents in transgenic plants

Amino acids of *Arabidopsis thaliana* seeds were determined by Nanjing Ruiyuan Biotechnology Co., Ltd. Each 0.5 g sample was added to 3 ml of extraction buffer (20% ethanol solution containing 0.001 mol/L HC1) and mixed. Amino acid was extracted by sonication at 4 °C for 30 min. Following centrifugation at 12,000 × g for 5 min, the supernatant was collected. Then the precipitate was repeatedly extracted twice with 1 ml of the extraction buffer. The supernatants were pooled and diluted to 5 ml. The extraction was filtered with aqueous membrane (0.22 mm pore size). The amino acid concentration was detected by HPLC-MS/MS (AGLIENT1260 and AB4000, America). The HPLC conditions were as follows: Column, Information-HILICZ (2.7 μm, 3.0 mm*100 mm); mobile phase, 75% acetonitrile (V/V) (solvent A) and 0.1 mol/L sodium acetate (solvent B); flow rate, 0.3 ml/min; column temperature, 35 °C; injection volume, 1 μl. All the experiments were conducted on three replicates.

2.8. Dual luciferase assay

In order to demonstrate the transactivation activity of transcription factors to *TgASA1* promoter, a transient dual luciferase assay was carried out. Under the control of CaMV 35S promoter as an effector, the selected transcription factors were cloned into pCAMBIA1300 vector. The promoter fragments were introduced into the pGreenII 0800-LUC dual reporter vector, so that these promoter fragments can be cloned as transcriptional fusions with the firefly luciferase gene (LUC). The constructed effector plasmid and reporter plasmid were transformed into *Agrobacterium tumefaciens* (strain GV1301), and then co-transformed into tobacco. All primers used are listed in Table S1. According to the user manual, use the dual luciferase assay kit (Promega) to measure the LUC and REN luciferase activities. Result was calculated by the ratio of LUC to REN.

2.9. Statistical and sequence analyses

Pearson's correlation coefficients (*r*) were calculated using SPSS, version 16.0 (SPSS, Inc., Chicago, IL, U.S.A.). Phylogenetic analysis was conducted using a bootstrap neighbor-joining evolutionary tree by MEGA 7.0 software with 1000 bootstrap replicates. Amino acid sequence alignment was conducted by DNAMAN software (version 9).

3. Results

3.1. Metabolome analysis of T. Grandis seeds

During ripening, the fruit gradually enlarge (Fig. 1A and B). Nine samples from three developmental stages were selected to study the changes in metabolites using a UPLC-MS/MS analysis. On the basis of widely targeted metabolomics, a total of 612 metabolites were detected in the metabolic analysis (Table S2). To check the quality of the acquired MS data, a MRM metabolite detection multi-peak map and total ion chromatograms (TICs) were generated and analyzed, revealing a high degree of overlap and a good sample preparation process (Fig. 1C and D; Fig. S1). An overview of the metabolite profiling and principal component analysis (PCA) of *T. grandis* seed in the three maturation stages is shown in Fig. 1E and Fig. 1F. All the identified metabolites were classified into 31 different categories (Table S2). Among them, organic acids (72 metabolites), amino acid derivatives (57 metabolites), and nucleotides and their derivatives (55 metabolites) represented the largest numbers of metabolites.

3.2. Identification of differentially accumulated metabolites (DAMs) of T. Grandis seeds in the three maturation stages

To reveal the variation in metabolite abundance during seed maturation, a quantitative analysis and inter-group comparisons of all the identified metabolites were carried out. The metabolomics data were analyzed using the OPLS-DA model, and the score map of each group was constructed to further show the differences among the groups (Fig. S2A). R2X, R2Y, and Q2 were close to 1, indicating the model was stable and reliable. To identify differential metabolites, we selected metabolites with a FC ≥ 2 (up-regulated) or ≤ 0.5 (down-regulated) with a VIP value ≥ 1 from the OPLS-DA model. The statistical analysis identified 233 DAMs, including 84 up- and 149 down-regulated metabolites, in stage 2 compared with in stage 1 seeds (Fig. S2B; Table S3). In the stage 3 vs stage 1 comparison group, 311 DAMs, including 98 up- and 213 down-regulated metabolites, were identified in the former (Fig. S2B; Table S4). In the stage 3 vs stage 2 comparison group, 188 DAMs, including 59 up- and 129 down-regulated metabolites, were identified (Fig. S2B; Table S5). Furthermore, the number of main accumulated metabolite characteristics belonging to each category was analyzed (Fig. S2C). Interestingly, most amino acids, carbohydrates, flavonoids, and organic acids differentially accumulated in stage1 rather than in stage 3 seeds. Additionally, lipids predominantly accumulated in stage 3 seeds.

3.3. Amino acids present different contents at different developmental stages

The key period for the material transformation and accumulation of *T. grandis* was from July to September in *T.grandis* seeds. Therefore, our study is devoted to focus on the changes of amino acid content during these stages. Using the LC-MS/MS-based metabolic profiling method, 28 kinds of amino acids in *T. grandis* kernels were identified (Fig. 2A). We found that the contents of L-Homoglutamic acid, L-Leucine, L-Methionine, L-Tryptophan, L-Isoleucine, L-Theanine, L-Glutamine, L-Histidine, L-Phenylalanine, L-(-)-Tyrosine, L-Citrulline,L-(+)-Arginine were decreased from July to September. The L-Glutamic acid was improved from stage1 to stage3. Besides, the contents of L-Alanine, L-Lysine, L-(-)-Cystine, L(+)-Ornithine reached the highest value in August. The contents of L-Valine, L-Asparagine, L-Homocitrulline, DL-homocysteine, L-Serine, L-Proline, Dl-Norvaline, and L-Homocystine were changed

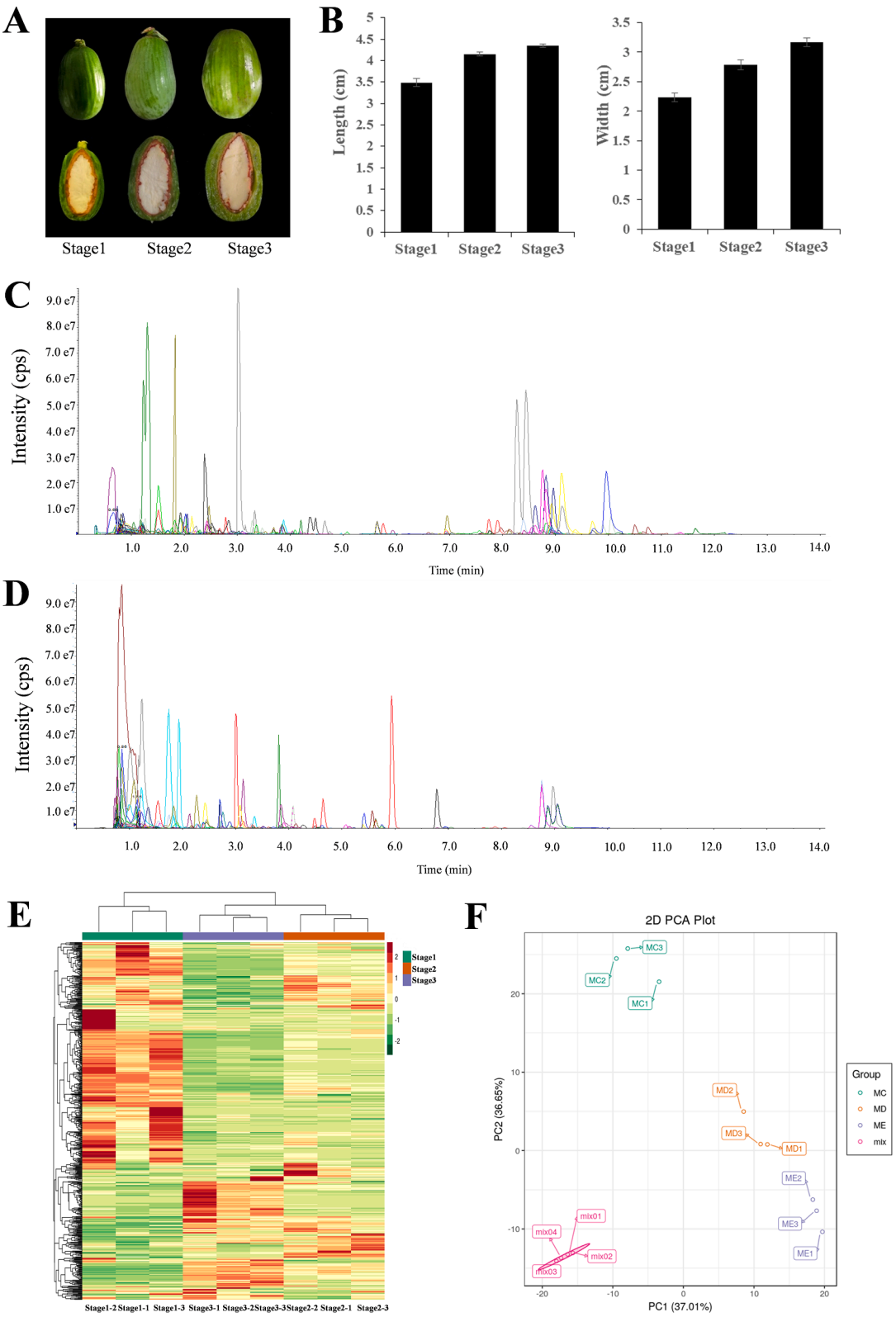
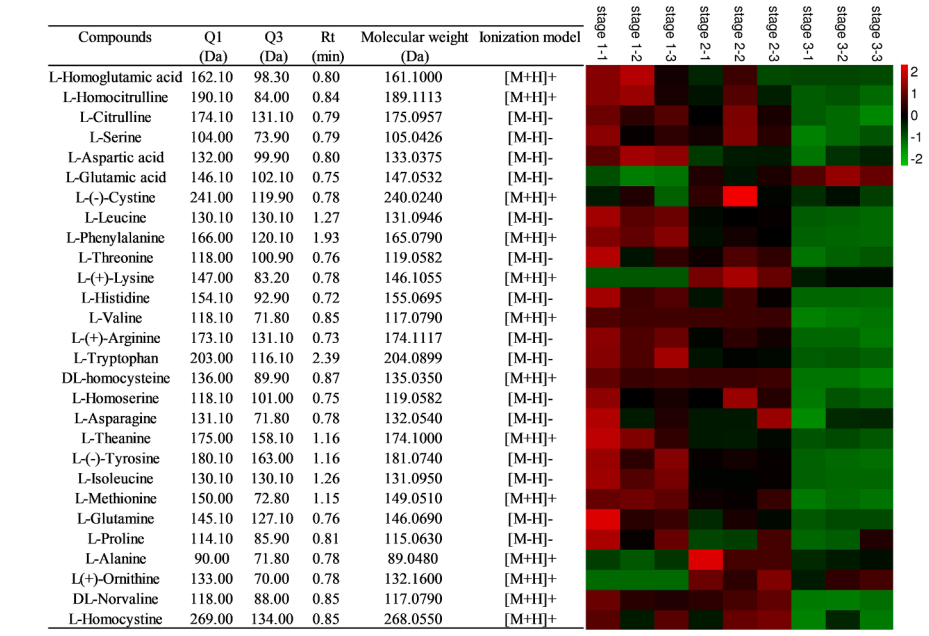


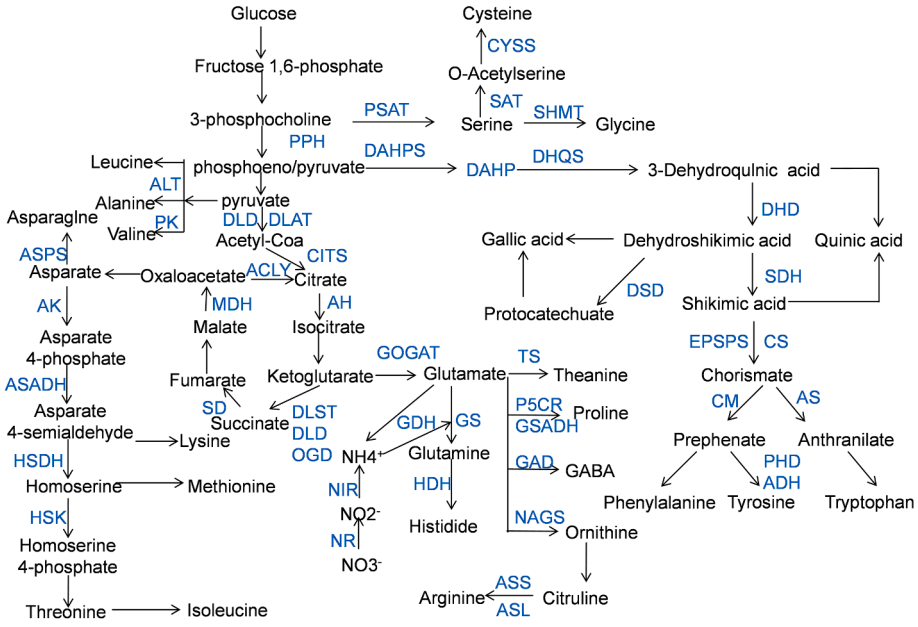
Fig. 1. Widely targeted metabolite profiling reveals the metabolites in *Torreya grandis* seed maturation. (A) Pictures of *T.grandis* seed in three maturation phases. (B) The Length and width of seed in three maturation phases. (C) MRM metabolite detection multi peak graph in ESI^+ mode. (D) MRM metabolite detection multi peak graph in ESI^- mode. The MRM metabolite detection multi peak diagram in the figure shows the substances that can be detected in the sample, and each mass spectrum peak with different colors represents one metabolite detected. (E) Hierarchical cluster analysis of the metabolites identified in the metabolomes of the three seed samples. The colour indicates the level of accumulation of each metabolite, from low (green) to high (red). The Z-score represents the deviation from the mean by standard deviation units. (F) Principal component analysis (PCA) of all identified metabolites. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

A

Compounds	Q1 (Da)	Q3 (Da)	Rt (min)	Molecular weight (Da)	Ionization model
L-Homoglutamic acid	162.10	98.30	0.80	161.1000	[M+H]+
L-Homocitrulline	190.10	84.00	0.84	189.1113	[M+H]+
L-Citrulline	174.10	131.10	0.79	175.0957	[M-H]-
L-Serine	104.00	73.90	0.79	105.0426	[M-H]-
L-Aspartic acid	132.00	99.90	0.80	133.0375	[M-H]-
L-Glutamic acid	146.10	102.10	0.75	147.0532	[M-H]-
L-(-)-Cystine	241.00	119.90	0.78	240.0240	[M+H]+
L-Leucine	130.10	130.10	1.27	131.0946	[M-H]-
L-Phenylalanine	166.00	120.10	1.93	165.0790	[M+H]+
L-Threonine	118.00	100.90	0.76	119.0582	[M-H]-
L-(+)-Lysine	147.00	83.20	0.78	146.1055	[M+H]+
L-Histidine	154.10	92.90	0.72	155.0695	[M-H]-
L-Valine	118.10	71.80	0.85	117.0790	[M+H]+
L-(+)-Arginine	173.10	131.10	0.73	174.1117	[M-H]-
L-Tryptophan	203.00	116.10	2.39	204.0899	[M-H]-
DL-homocysteine	136.00	89.90	0.87	135.0350	[M+H]+
L-Homoserine	118.10	101.00	0.75	119.0582	[M-H]-
L-Asparagine	131.10	71.80	0.78	132.0540	[M-H]-
L-Theanine	175.00	158.10	1.16	174.1000	[M+H]+
L-(-)-Tyrosine	180.10	163.00	1.16	181.0740	[M-H]-
L-Isoleucine	130.10	130.10	1.26	131.0950	[M-H]-
L-Methionine	150.00	72.80	1.15	149.0510	[M+H]+
L-Glutamine	145.10	127.10	0.76	146.0690	[M-H]-
L-Proline	114.10	85.90	0.81	115.0630	[M-H]-
L-Alanine	90.00	71.80	0.78	89.0480	[M+H]+
L(+)-Ornithine	133.00	70.00	0.78	132.1600	[M+H]+
DL-Norvaline	118.00	88.00	0.85	117.0790	[M+H]+
L-Homocystine	269.00	134.00	0.85	268.0550	[M+H]+



B



C

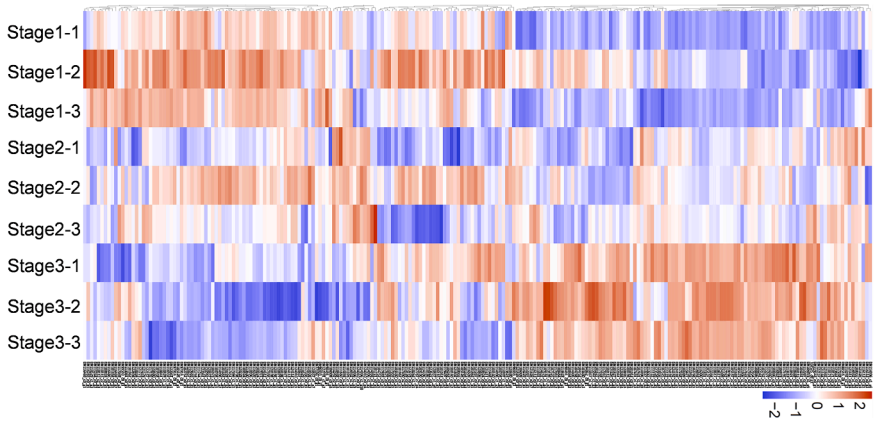


Fig. 2. Gene expression and amino acids analysis for amino acid synthesis pathway of *T. grandis*. (A) The relative contents of amino acids in *T. grandis* kernels in the three mature stages. (B) Schematic diagram of the tricarboxylic acid cycle (TCA) pathway and amino acid biosynthesis in *T. grandis*. (C) The heat map of the expression pattern of DEG in the amino acid metabolism pathway analyzed by KEGG database.

slightly from July to August, while decreased significantly from stage2 to stage3. The contents of L-Homoserine and L-Aspartic acid were increased from stage1 to stage2, and then kept unchanging. These showed that the content of amino acids vary enormously with the development time in *T. grandis* kernels.

3.4. Identification of amino acid metabolism pathway

Amino acids are synthesized from glucose through tricarboxylic acid (TCA) cycle and the oxidized pentose phosphate pathway (OPPP). Based on the amino acid synthesis pathways reported by other species and the dates of transcriptome and metabolome in *T. grandis kernels*, we first proposed the biosynthetic pathway of amino acids in *T. grandis* (Fig. 2B). In total, there are 267 unigenes encoding 28 enzymes from transcriptome and 28 metabolites (including amino acids and their derivatives) from metabolome involved in amino acid synthesis were identified. Of the 267 unigenes and 28 metabolites, there are 67 unigenes and 1 metabolome related to tricarboxylic acid (TCA) cycle and 17 unigenes and 2 metabolomes related to OPPP. The expression level and annotation of all the unigenes and the unigenes involved in amino acid biosynthetic pathway was shown in Table S6 and Fig. 2C.

3.5. Correlation analysis between gene expression of amino acid biosynthesis and metabolites of amino acid synthesis pathway

In Fig. 2C, we found that there are 28 enzymes were involved in amino acid biosynthesis, most of which have multiple unigenes. Therefore, we performed phylogenetic tree analysis on the unigenes encoding proteins from *T. grandis* and their homologes from Arabidopsis or rice to identify possible candidate genes (Fig. S3). The unigenes most closed to the function known protein were indicated by red box. To further uncover the key genes controlling amino acid biosynthesis, the pearson correlation analysis was performed between the expression levels of the candidate genes and the content of amino acids and their derivatives in the amino acid synthesis pathway (Fig. 3). The result showed that TRINITY_DN111744_c0_g1 and TRINITY_DN111744_c0_g2 encoding DAHP, TRINITY_DN119683_c1_g1 encoding anthranilate synthase (ASA), TRINITY_DN118841_c0_g1 encoding citrate synthase (CITS) were positively and significantly correlated to at least 10 amino acids, suggesting that these genes play key roles in regulating amino acid biosynthesis. Thereafter, the CDSs of two unigenes (TRINITY_DN111744_c0_g2 named as *TgDAHP2* and TRINITY_DN119683_c1_g1 named as *TgASA1*) were successfully cloned and selected for further analysis.

3.6. Amino acid sequence alignments and subcellular localization of TgDAHP2 and TgASA1

The results of multiple sequence alignments depicted that the amino acid sequences of *TgDAHP2* and *TgASA1* in *T. grandis* are highly similar to their homologes in Arabidopsis (Fig. 4A). The TgDAHP2 of *T. grandis* and the DAHP of Arabidopsis have the same DAHP_synth_2 superfamily conserved domains. The *TgASA1* and the ASA of Arabidopsis have the same PLN02445 superfamily conserved domains.

To investigate the subcellular localization of TgDAHP2 and TgASA1, TgDAHP2-GFP and TgASA1-GFP fusion plasmids were constructed, which were then transiently expressed into *Nicotiana benthamiana* leaves driven by the CaMV35S promoter, respectively. Confocal microscopy examination of the epidermal cells indicated that strong GFP signals of both TgDAHP2-GFP and TgASA1-GFP were co-localized with auto fluorescence of chloroplast, suggesting that TgDAHP2 and TgASA1were both localized to chloroplast (Fig. 4B and C).

3.7. Amino acid contents in transgenic Arabidopsis

The TgDAHP2 and TgASA1 overexpression vectors were transformed into *Arabidopsis thaliana*, respectively (Fig. 5A). The amino acid content of transgenic Arabidopsis and wild type (WT) seeds were measured using HPLC-MS. The findings revealed that the content of the most of amino acid in the seeds overexpressing *TgDAHP2* (DAHP2-OE) and *TgASA1* (ASA1-OE) were significantly higher compared with that of the WT (Fig. 5B and C). In the seeds, the total amino acid contents of DAHP2-OE and ASA1-OE are 802.83 μg/g and 647.45 μg/g, respectively. The total amino acid content of WT is 478.49 μg/g. For its single amino acid analysis, the results demonstrated that the content of aspartic acid, glutamic acid, histidine, arginine, alanine, tyrosine, valine, methionine, isoleucine, tryptophan, phenylalanine and Lysine in the ASA1-OE and DAHP2-OE were significantly higher than that of WT.

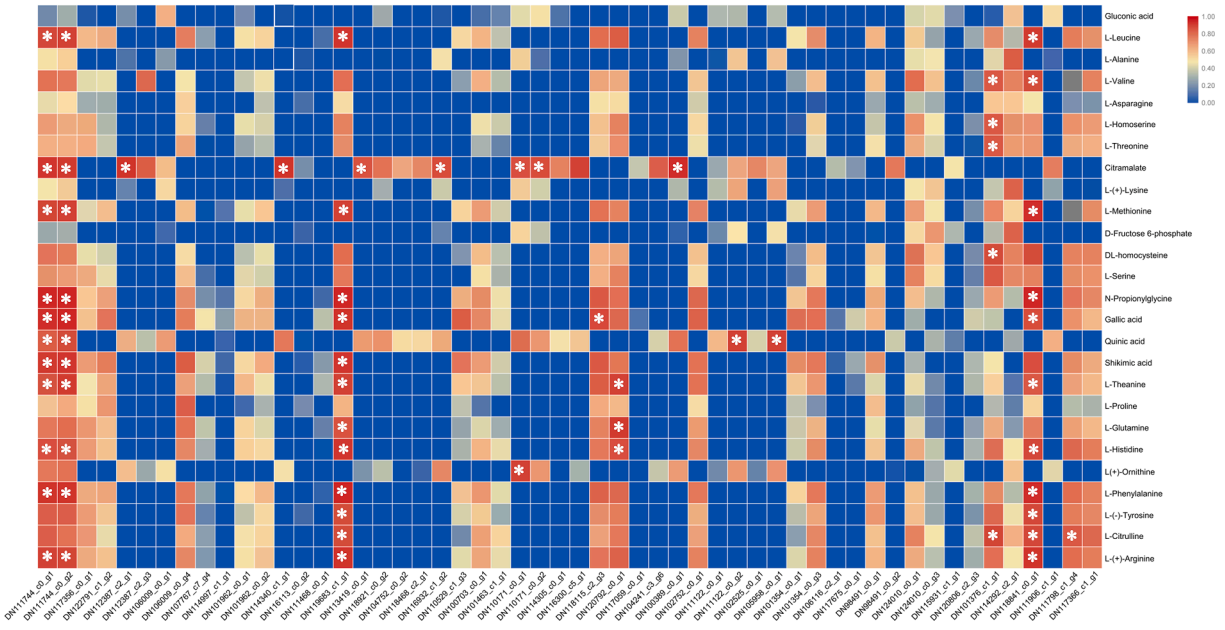


Fig. 3. Pearson correlation analysis between amino acid coding genes and metabolites in amino acid synthesis pathway. The Pearson correlation analysis between the expression levels of the possible coding genes selected by the phylogenetic tree analysis and the metabolites on the amino acid synthesis pathway in the three mature stage of *T. grandis* seeds. The significant ones are marked with * ($p < 0.05$).

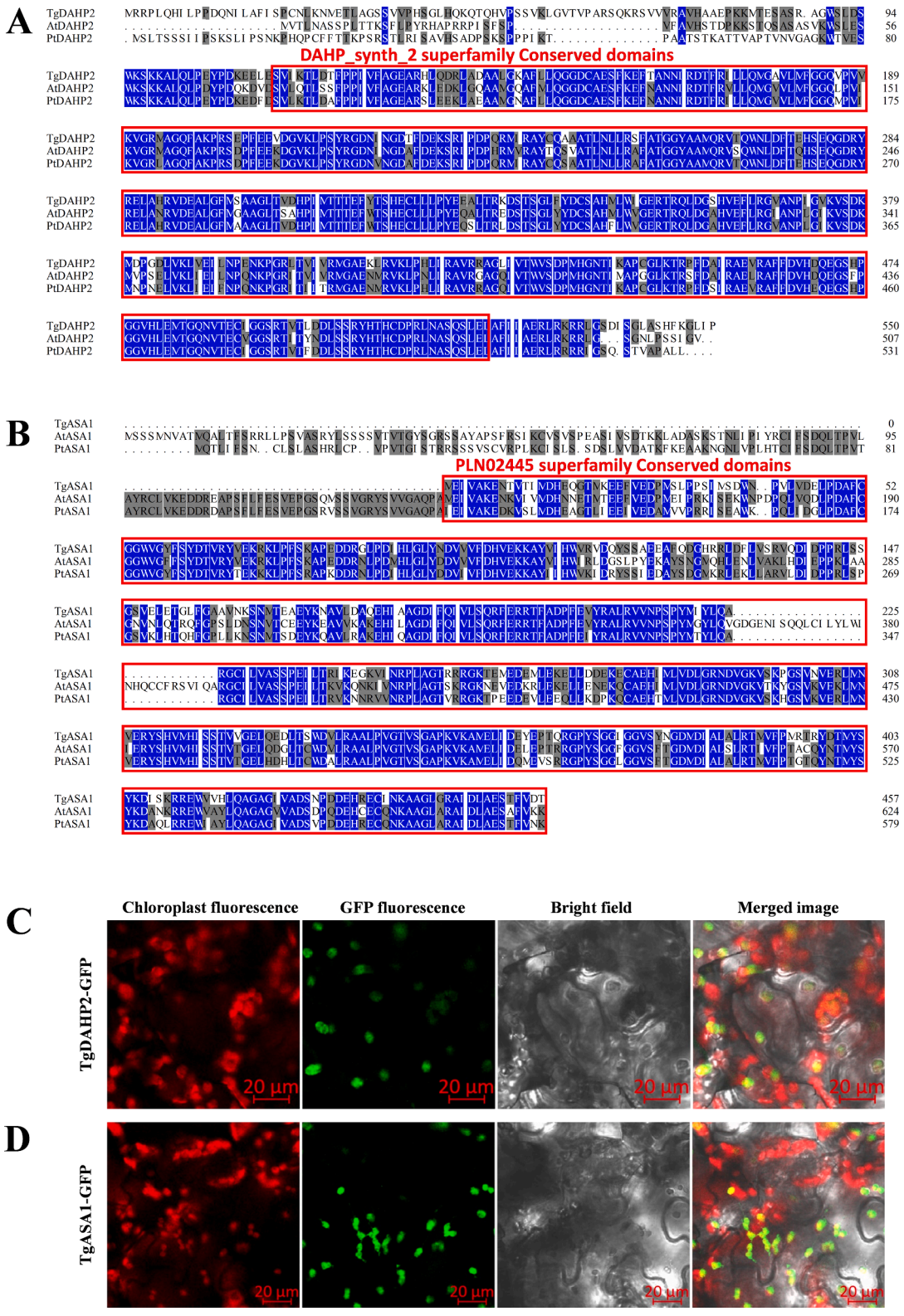


Fig. 4. Amino acid sequence alignment and subcellular location of proteins involved in amino acid biosynthesis of *T. grandis* and other plants. (A) Alignment of the deduced amino acid sequences of TgDAHP2 (TRINITY_DN111744_c0_g2) AtDAHP2 and PtDAHP2. The DAHP_synth_2 superfamily Conserved domains site indicated by red box. (B) Alignment of the deduced amino acid sequences of TgASA1 (TRINITY_DN119683_c1_g1), AtASA1 (AT2G29690) and PtASA1 (XP_002316223.2). The PLN02445 superfamily Conserved domains site indicated by red box. (C) Subcellular localization of TgDAHP2. TgDAHP2-GFP fusion plasmids, as well as ER marker were transiently expressed in leaves of *N. benthamiana*. (D) Subcellular localization of TgASA1. TgASA1-GFP fusion plasmids, as well as ER marker were transiently expressed in leaves of *N. benthamiana*. Proteins were localized with confocal fluorescence microscopy. Scale bars: 20 μm. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

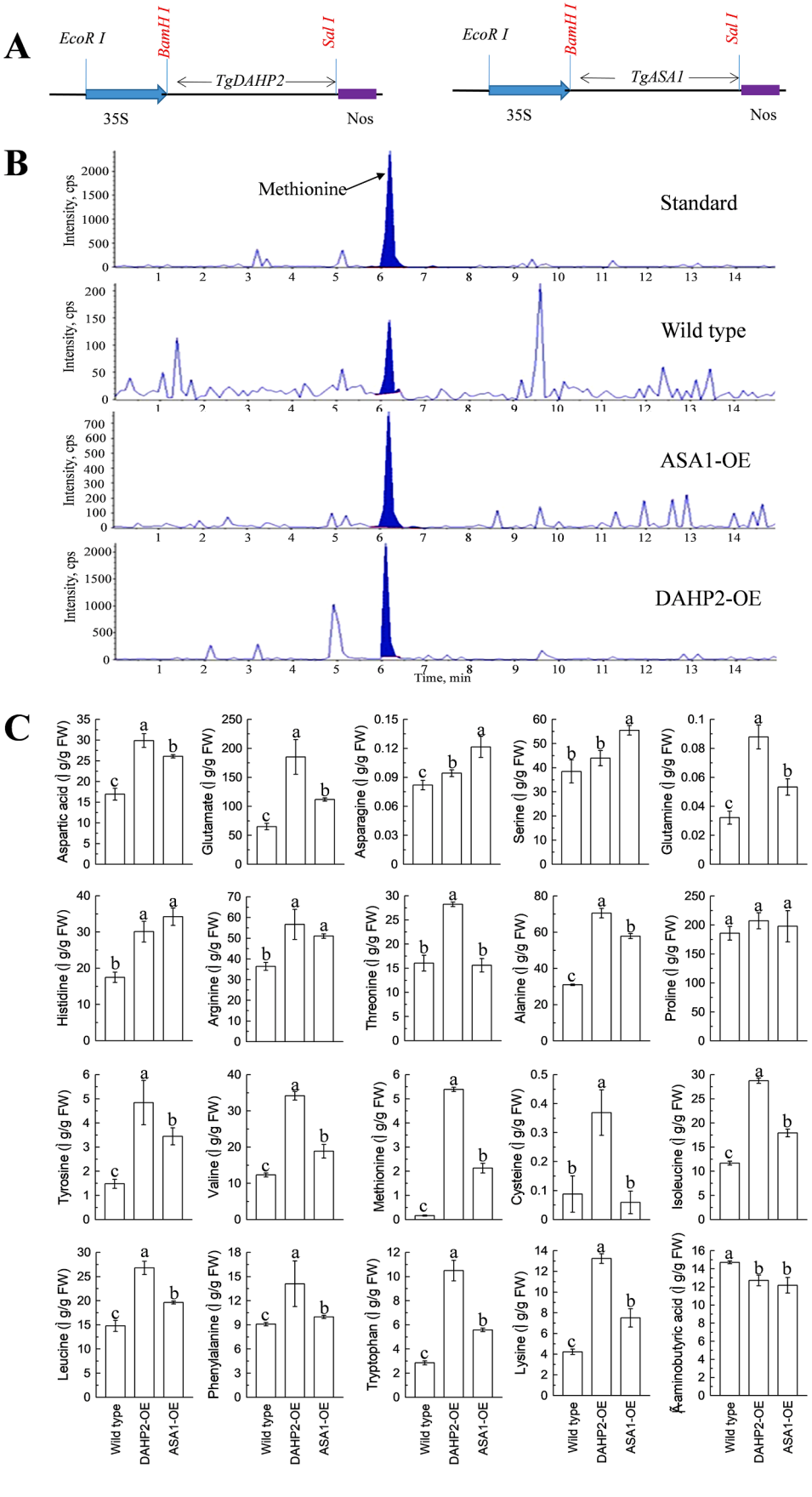


Fig. 5. Overexpression of *TgDAHP2* and *TgASA1* in Arabidopsis affects its amino acid content. (A) Each of the full-length CDS of *TgDAHP2* and *TgASA1* was inserted into the pCAMBIA1300 expression vector under control of the 35S promoter. (B) Representative chromatograms of methionine from standard, wild type and *TgDAHP2* and *TgASA1* overexpressed Arabidopsis seeds. (C) Amino acid content in the wild type and *TgDAHP2* and *TgASA1* overexpressed Arabidopsis seeds. FW, fresh weight. Data are means ± SD (n = 3). Different letters indicate significant differences between samples at $p < 0.05$.

However, serine only increased significantly in ASA1-OE. Threonine, cysteine, and leucine only increased significantly in DAHP2-OE. The amino acid content of asparagine, proline and GABA in the ASA1-OE and DAHP2-OE did not change significantly compared to WT. Therefore, *TgDAHP2* and *TgASA1* play key roles in amino acid synthesis.

3.8. Identification of candidate transcriptional modulators involved in amino acid biosynthesis pathway

It is well established that transcription factors are essential for regulating gene expression. To explore transcription factors that involved in regulating amino acid biosynthesis, co-expression analysis was conducted (Fig. 6A). The gene expression was clustered into 12 clusters. *TgASA1* was found in cluster 7, however *TgDAHP2* was not

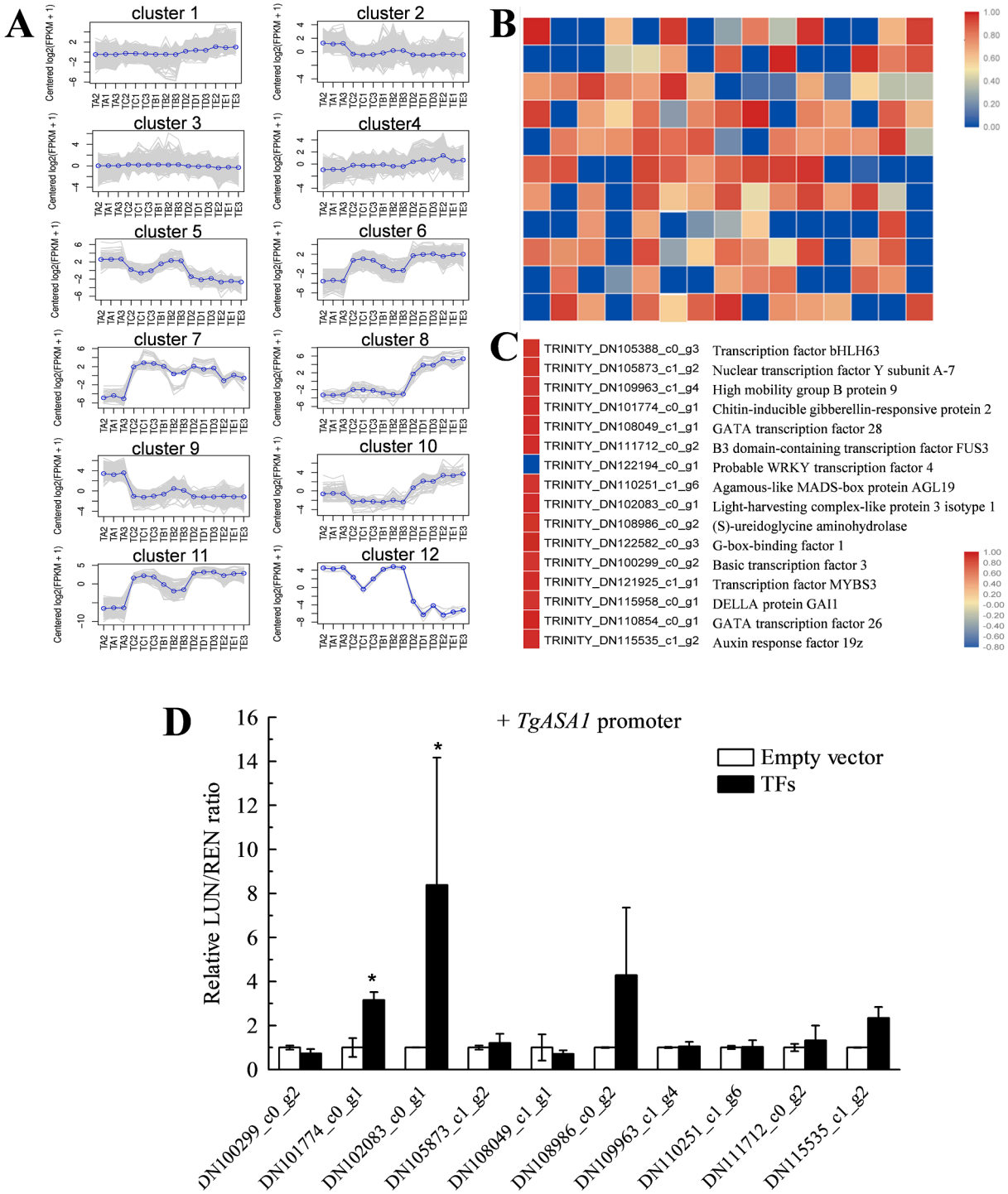


Fig. 6. Screening of transcription factors co-expressed with genes encoding enzymes involved in the biosynthesis of amino acid from transcriptome. (A) Relative expression levels of 12 transcription factor libraries. (B) Pearson's correlation analysis between the expression levels of genes encoding enzymes involved in the biosynthesis of amino acid and all transcription factors from transcriptome. (C) The transcription factors significantly correlated with *TgASA1* with the Pearson's correlation coefficients ≥ 0.90 were selected. (D) We cloned 10 transcription factors and *TgASA1* promoters for dual-molecule luciferase experiments. *, $p < 0.05$.

found from all the 12 clusters. Then Pearson correlation analysis between the transcripts annotated as transcription factors from the cluster 7 and *TgASA1* was conducted. The correlation coefficients were presented as heatmap (Fig. 6B). With Pearson correlation coefficient ≥ 0.90 and FPKM ≥ 10, the transcription factors that are significantly correlated to *TgASA1* were selected (Fig. 6C). Further dual-luciferase analysis was carried out to detect the transcription factors which can activate *TgASA1* promoter (Fig. 6D). The result showed that TRINITY_DN101774_c0_g1, TRINITY_DN108986_c0_g2 and TRINITY_DN115535_c1_g2 could activate the expression of *TgASA1*. TRINITY_DN101774_c0_g1 encodes a gibberellin-responsive protein, TRINITY_DN108986_c0_g2 encodes a (S)-ureidoglycine aminohydrolase, and TRINITY_DN115535_c1_g2 encodes an auxin response factor.

4. Discussion

Research on the genus *Torreya* is still in its primary stage. At present, the main focus has been on the composition and content of *Torreya*. The material extracted from *Torreya nucifera* has strong antibacterial and antioxidative effects, indicating that this species can be used as a natural antibacterial with antioxidative effects (Rhim & Choi, 2015). In total, 86 compounds were found in *T. grandis* cv. Merrilli arils, including some important oils, terpenes, esters, ketones, and aldehydes (Yu et al., 2016). Widely targeted metabolite profiling using MS/MS data has been widely applied to large-scale metabolic profiling and comparative metabolomics of many plant species (Wang et al., 2018; Zou et al., 2020). To date, the changes in metabolites during *T. grandis* seed development have not been thoroughly studied. Here, we found a great deal of metabolites belonging to different categories using LC-MS/MS-based widely targeted metabolomics. This provided an opportunity to elucidate the regulatory mechanisms of *T. grandis* seeds' nutrient accumulation. We identified a total of 612 metabolites, 233, 311, and 188 of which were differentially accumulated in stage 2 vs stage 1, stage 3 vs stage 1, and stage 3 vs stage 2 comparison groups, respectively.

Previous studies identified 17 amino acids in *T. grandis* (Li et al., 2005). However, here using a widely targeted metabonomics analysis, 28 amino acids were identified, of which 26 were differentially accumulated during seed development. *T. grandis* seeds are rich in amino acids, however their biosynthesis pathway and regulatory mechanism remain unknown. In this research, it's the first time using the transcriptome and metabolome analysis to report the amino acid biosynthesis. In total, 28 metabolites and 267 unigenes encoding 28 enzymes related to amino acid synthesis were identified. Correlation analysis showed that unigenes for DAHP, ASA and CITS significantly correlated with at least 10 of 28 amino acids, suggesting that these enzymes may play vital roles in amino acids synthesis.

As the critical rate-limiting enzyme, DAHP is involved in the biosynthesis of indispensable aromatic compounds in microorganisms and plants (Zhao et al., 2019). In addition, it is also a crucial factor in regulating the carbon flow into the oxalate pathway (Cui et al., 2019). The cDNA of DAHP from *Solanum tuberosum* L. specified a chloroplast transport sequence near its 5′-end (Zhao et al., 2019). The AtDAHP2 has been reported to localize to the chloroplast in Arabidopsis (Kanaris et al., 2021), and this is consistent with our analyses of TgDAHP2 that were found to be chloroplast localized. Expression of a bacterial DAHPS in Arabidopsis resulted in increased levels of phenylalanine, tryptophan and other shikimate pathway-derived compounds (Tzin et al., 2012). At the same time, some scholars have discovered other functions of the DAHP gene. In morning glory, DAHP is responsible for not only the coupling of metabolites from primary metabolism to secondary metabolism, but also the final function of FVBP (Floral volatile benzenoid/phenylpropanoid biosynthesis) biosynthesis (Langer et al., 2014). Tryptophan is synthesized from chorismate in the chloroplasts (Radwanski & Last, 1995). Anthranilate synthase (ASA), as the enzyme participating in the initial reactions in the tryptophan biosynthesis pathway in many plants and microorganisms, catalyzes the conversion of chorismate to anthranilate (Srivastava & Sinha, 2017). In Arabidopsis, ASA1 and ASA2 encode the subunits of ASA, both proteins have putative chloroplast transit peptides at their amino termini (Niyogi & Fink, 1992). Our result showed that TgASA1 was localized to the chloroplast, which provides direct evidence for the chloroplast localization of ASA1. Overexpression of an ASA alpha subunit gene, *OASA1D*, in Arabidopsis accumulated higher contents of phenylalanine and tyrosine (Ishihara et al., 2006). In this study, we found that not only the content of phenylalanine, tyrosine and tryptophan, but also some other amino acids in the *TgASA1* and *TgDAHP2* overexpressed Arabidopsis seeds were significantly higher than that of WT. CITS is an enzyme that catalyzes the condensation of acetyl CoA from glycolysis or other dissimilation reactions with oxaloacetic acid to synthesize citric acid (Wu et al., 2020). CITS can perform the citric acid cleavage required for the functional ROTCA cycle (Zhang et al., 2020). In *Haloxylon salicornicum* plants, CITS, succinate dehydrogenase, and malate dehydrogenase were up-regulated under dry salt treatment, displaying that they devoted to providing energy for salt tolerance (Panda et al., 2020).

In the identification of *TgASA1* and co-expressed transcription factors and the verification of LUC, we found that TRINITY_DN101774_c0_g1, TRINITY_DN108986_c0_g2 and TRINITY_DN115535_c1_g2 interact with *TgASA1* promoters. The main function of TRINITY_DN101774_c0_g1 is to induce gibberellin-responsive protein, TRINITY_DN108986_c0_g2 is a (S)-ureidoglycine aminohydrolase, and TRINITY_DN115535_c1_g2 is an auxin response factor. Previous studies have also found many transcription factors that interact with ASA genes. In Arabidopsis, *ASA1* is up regulated by ERF1 directly, giving rise to accumulating auxin and inhibiting root growth induced by ethylene (Mao et al., 2016). During lateral root formation, AtASA1 takes part in auxin biosynthesis and transport mediated by jasmonate as well (Sun et al., 2009). To sum up, ASA performs a significant role in regulating auxin in plants (Zhang et al., 2019), which contributes a reliable basis for the transcription factors we screened. But regarding the relationship of ASA and gibberellin, the response mechanism is still few and further research is needed.

5. Conclusion

We studied the accumulations of metabolites during the seed maturation of *T. grandis* using an LC-MS/MS-based metabolic approach. The composition and content of carbohydrates, organic acids, flavonoids, amino acids, and lipids varied during seed ripening. By combining metabolome and transcriptome analysis, we first proposed the biosynthetic pathway of amino acids in *T. grandis*. Most notably, we have identified the key factors participated in the biosynthetic pathway of amino acids and verified their functions. And found four transcription factors that may interact with *TgASA1*. The identification of the enzyme-encoding genes of the amino acid biosynthetic pathway and their new transcription activators provides a basis for us to characterize the novel amino acid biosynthetic regulatory mechanism.

CRediT authorship contribution statement

Heqiang Lou: Methodology, Writing – review & editing, Supervision. **Yi Yang:** Methodology, Validation, Investigation, Resources, Writing – original draft. **Shan Zheng:** Methodology, Validation, Investigation. **Zhenmin Ma:** Methodology, Validation, Investigation. **Wenchao Chen:** Methodology, Validation, Investigation. **Chenliang Yu:**

Methodology, Software. **Lili Song:** Resources, Writing – review & editing, Supervision. **Jiasheng Wu:** Resources, Supervision, Funding acquisition.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was funded by the Key Research and Development Program of Zhejiang Province (2020C02019 and 2021C02001), the Fundamental Research Funds for the Provincial Universities of Zhejiang (2020YQ003), the National Natural Science Foundation of China (32171830), the Breeding of New Varieties of *Torreya grandis* Program (2021C02066-11), the Scientific Research Startup Fund Project of Zhejiang A&F University (2018FR028), the State Key Laboratory of Subtropical Silviculture (ZY20180312 and ZY20180209), the Young Elite Scientists Sponsorship Program by China Academy of Space Technology (CAST) (2018QNRC001).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.foodchem.2022.132078.

References

Cañas, R. A., de la Torre, F., Cánovas, F. M., & Cantón, F. R. (2007). Coordination of *PsAS1* and *PsASPG* expression controls timing of re-allocated N utilization in hypocotyls of pine seedlings. *Planta, 225*(5), 1205–1219. https://doi.org/10.1007/s00425-006-0431-9

Chen, Z., Zheng, H., Fu, Q., Zhou, Y., & Weng, Y. (1998). Determination of oil contents and fatty acids in seeds of *Torreya Arn.* in China. *China Journal of Chinese Materia Medica, 23*, 456.

Cui, D., Deng, A., Bai, H., Yang, Z., Liang, Y., Liu, Z., … Wen, T. (2019). Molecular basis for feedback inhibition of tyrosine-regulated 3-deoxy-D-arabino-heptulosonate-7-phosphate synthase from *Escherichia coli*. *Journal of Structural Biology, 206*(3), 322–334. https://doi.org/10.1016/j.jsb:2019.04.001

Du, B., Kruse, J., Winkler, J. B., Alfarray, S., Schnitzler, J. P., Ache, P., Hedrich, R., & Rennenberg, H. (2019). Climate and development modulate the metabolome and antioxidative system of date palm leaves. *Journal of Experimental Botany*, *70*, 5959-5969. https://doi.org/10.1093/jxb/erz361.

He, Z., Zhu, H., Li, W., Zeng, M., Wu, S., Chen, S., … Chen, J. (2016). Chemical components of cold pressed kernel oils from different *Torreya grandis* cultivars. *Food Chemistry, 209*, 196–202. https://doi.org/10.1016/j.foodchem.2016.04.053

Ishihara, A., Asada, Y., Takahashi, Y., Yabe, N., Komeda, Y., Nishioka, T., … Wakasa, K. (2006). Metabolic changes in *Arabidopsis thaliana* expressing the feedback-resistant anthranilate synthase α-subunit gene *OASA1D*. *Phytochemistry, 67*(21), 2349–2362. https://doi.org/10.1016/j.phytochem.2006.08.008

Jones, D. C., Zheng, W., Huang, S., Du, C., Zhao, X., Yennamalli, R. M., … Li, L. (2016). A clade-specific Arabidopsis gene connects primary metabolism and senescence. *Frontiers in Plant Science, 7*. https://doi.org/10.3389/fpls.2016.00983

Joy, R. W., Vogel, H. J., & Thorpe, T. A. (1997). Inorganic nitrogen metabolism in embryogenie white spruce cultures: A nitrogen 14/15 NMR study. *Journal of Plant Physiology, 151*(3), 306–315. https://doi.org/10.1016/S0176-1617(97)80257-X

Kanaris, M., Poulin, J., Shahinas, D., Johnson, D., Crowley, V. M., Fucile, G., … Christendat, D. (2021). Elevated tyrosine results in the cytosolic retention of 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase in *Arabidopsis thaliana*. *The Plant Journal*. https://doi.org/10.1111/tpj.15590

Kaspar, H., Dettmer, K., Gronwald, W., & Oefner, P. J. (2009). Advances in amino acid analysis. *Analytical & Bioanalytical Chemistry, 393*(2), 445–452. https://doi.org/10.1007/s00216-008-2421-1

Keiki, I., Eri, I., Mayumi, T., Tomoyuki, Y., & Hideki, T. (2004). Biochemical background and compartmentalized functions of cytosolic glutamine synthetase for active ammonium assimilation in rice roots. *Plant and Cell Physiology, 45*, 1640–1647. https://doi.org/10.1093/pcp/pch190

Langer, K. M., Jones, C. R., Jaworski, E. A., Rushing, G. V., Kim, J. Y., Clark, D. G., & Colquhoun, T. A. (2014). *PhDAHP1* is required for floral volatile benzenoid/phenylpropanoid biosynthesis in *Petunia* × *hybrida* cv 'Mitchell Diploid'. *Phytochemistry, 103*, 22–31. https://doi.org/10.1016/j.phytochem.2014.04.004

Li, W., Zhang, H., Li, X., Zhang, F., Liu, C., Du, Y., … Guo, Y. (2017). Intergrative metabolomic and transcriptomic analyses unveil nutrient remobilization events in leaf senescence of tobacco. *Scientific Reports, 7*(1). https://doi.org/10.1038/s41598-017-11615-0

Li, Z. J., Luo, C. F., Chen, X. J., Feng, X. J., & Yu, W. W. (2005). Component analysis and nutrition evaluation of seeds of *Torreya grandis* 'Merrillii'. *Journal of Zhejiang Forestry College, 22*, 540–544. https://doi.org/10.1360/aps040037

Luo, L., Qin, R., Liu, T., Yu, M., Yang, T., & Xu, G. (2018). OsASN1 plays a critical role in asparagine-dependent rice development. *International Journal of Molecular Sciences, 20*, 130. https://doi.org/10.3390/ijms20010130

Mao, J. L., Miao, Z. Q., Wang, Z., Yu, L. H., Cai, X. T., Xiang, C. B., & Qu, L. J. (2016). *Arabidopsis* ERF1 mediates cross-talk between ethylene and auxin biosynthesis during primary root elongation by regulating *ASA1* expression. *PLoS Genetics, 12*(1), e1005760. https://doi.org/10.1371/journal.pgen.1005760

Mir, R., Jallu, S., & Singh, T. P. (2015). The shikimate pathway: Review of amino acid sequence, function and three-dimensional structures of the enzymes. *Critical Reviews in Microbiology, 41*(2), 172–189. https://doi.org/10.3109/1040841X.2013.813901

Ninan, A. S., Grant, J., Song, J., & Jameson, P. E. (2019). Expression of genes related to sugar and amino acid transport and cytokinin metabolism during leaf development and senescence in *Pisum sativum* L. *Plants, 8*, 76. https://doi.org/10.3390/plants8030076

Niyogi, K. K., & Fink, G. R. (1992). Two anthranilate synthase genes in Arabidopsis: Defense-related regulation of the tryptophan pathway. *The Plant Cell, 4*(6), 721–733. https://doi.org/10.1105/tpc.4.6.721

Panda, A., Rangani, J., & Parida, A. K. (2020). Comprehensive proteomic analysis revealing multifaceted regulatory network of the xero-halophyte *Haloxylon salicornicum* involved in salt tolerance. *Journal of Biotechnology, 324*, 143–161. https://doi.org/10.1016/j.jbiotec.2020.10.011

Pandurangan, S., Pajak, A., Molnar, S. J., Cober, E. R., Dhaubhadel, S., Hernandez-Sebastia, C., … Marsolais, F. (2012). Relationship between asparagine metabolism and protein concentration in soybean seed. *Journal of Experimental Botany, 63*, 3173–3184. https://doi.org/10.1016/0021-9150(95)05530-A

Pott, M., Schmidt, M. J., & Summerer, D. (2014). Evolved sequence contexts for highly efficient amber suppression with noncanonical amino acids. *ACS Chemical Biology, 9*(12), 2815–2822. https://doi.org/10.1021/cb5006273

Radwanski, E. R., & Last, R. L. (1995). Tryptophan biosynthesis and metabolism: Biochemical and molecular genetics. *The Plant Cell, 7*(7), 921–934. https://doi.org/10.1105/tpc.7.7.921

Rhim, T.-J., & Choi, M.-Y. (2015). Antimicrobial effects on food-borne pathogens and the antioxidant activity of *Torreya Nucifera* Extract. *The Korean Journal of Community Living Science, 26*(4), 697–705. https://doi.org/10.7856/kjcls.2015.26.4.697

Rodgers, K. J. (2014). Non-protein amino acids and neurodegeneration: The enemy within. *Experimental Neurology, 253*, 192–196. https://doi.org/10.1016/j.expneurol.2013.12.010

Srivastava, A., & Sinha, S. (2017). Uncoupling of an ammonia channel as a mechanism of allosteric inhibition in anthranilate synthase of *Serratia marcescens*: Dynamic and graph theoretical analysis. *Molecular Biosystems, 13*(1), 142–155. https://doi.org/10.1039/C6MB00646A

Sun, J., Xu, Y., Ye, S., Jiang, H., Chen, Q., Liu, F., Zhou, W., Chen, R., Li, X., & Tietz, O. (2009). *Arabidopsis ASA1* is important for jasmonate-mediated regulation of auxin biosynthesis and transport during lateral root formation. *The Plant Cell*, *21*, 1495-1511. https://doi.org/10.1105/tpc.108.064303.

Tzin, V., & Galili, G. (2010). New insights into the shikimate and aromatic amino acids biosynthesis pathways in plants. *Molecular Plant, 3*(6), 956–972. https://doi.org/10.1093/mp/ssq048

Tzin, V., Malitsky, S., Zvi, M. M. B., Bedair, M., Sumner, L., Aharoni, A., & Galili, G. (2012). Expression of a bacterial feedback-insensitive 3-deoxy-d-arabino-heptulosonate 7-phosphate synthase of the shikimate pathway in Arabidopsis elucidates potential metabolic bottlenecks between primary and secondary metabolism. *New Phytologist, 194*, 430–439. https://doi.org/10.1111/j.1469-8137.2012.04052.x

Wang, D., Zhang, L., Huang, X., Wang, X., Yang, R., Mao, J., … Li, P. (2018). Identification of nutritional components in black sesame determined by widely targeted metabolomics and traditional Chinese medicines. *Molecules, 23*(5), 1180. https://doi.org/10.3390/molecules23051180

Watanabe, M., Balazadeh, S., Tohge, T., Erban, A., Giavalisco, P., Kopka, J., Mueller-Roeber, B., Fernie, A. R., & Hoefgen, R. (2013). Comprehensive dissection of spatiotemporal metabolic shifts in primary, secondary, and lipid metabolism during developmental senescence in Arabidopsis. *Plant Physiology*, *162*, 1290-1310. https://doi.org/10.1104/pp.113.217380.

Wu, X., Tovilla-Coutiño, D. B., & Eiteman, M. A. (2020). Engineered citrate synthase improves citramalic acid generation in *Escherichia coli*. *Biotechnology and Bioengineering, 117*(9), 2781–2790. https://doi.org/10.1002/bit.v117.910.1002/bit.27450

Yu, Y. J., Ni, S., Wu, F., Sang, & W. G. (2016). Chemical composition and antioxidant activity of essential oil from *Torreya grandis* cv. *merrillii* arils. *Journal of Essential Oil Bearing Plants*, *19*, 1170-1180. https://doi.org/10.1080/0972060X.2014.989183.

Zhang, G., Zhao, F., Chen, L., Pan, Y., Sun, L., Bao, N., … Xu, L. (2019). Jasmonate-mediated wound signalling promotes plant regeneration. *Nature Plants, 5*(5), 491–497. https://doi.org/10.1038/s41477-019-0408-x

Zhang, T., Shi, X.-C., Ding, R., Xu, K., & Tremblay, P.-L. (2020). The hidden chemolithoautotrophic metabolism of *Geobacter sulfurreducens* uncovered by adaptation to formate. *Isme Journal, 14*(8), 2078–2089. https://doi.org/10.1038/s41396-020-0673-8

Zhao, H., Gao, H., Ji, K., Yan, B., Li, Q., Mo, S., … Jiang, C. (2019). Isolation and biochemical characterization of a metagenome-derived 3-deoxy-d-arabino-

heptulosonate-7-phosphate synthase gene from subtropical marine mangrove wetland sediments. *AMB Express, 9*(1). https://doi.org/10.1186/s13568-019-0742-4

Zhu, G., Wang, S., Huang, Z., Zhang, S., Liao, Q., Zhang, C., … Huang, S. (2018). Rewiring of the fruit metabolome in tomato breeding. *Cell, 172*(1-2), 249–261.e12. https://doi.org/10.1016/j.cell.2017.12.019

Zou, S., Wu, J., Shahid, M. Q., He, Y., Lin, S., Liu, Z., & Yang, X. (2020). Identification of key taste components in loquat using widely targeted metabolomics. *Food Chemistry, 323*, 126822. https://doi.org/10.1016/j.foodchem.2020.126822