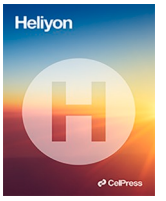
Heliyon 10 (2024) e28636

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Integrated transcriptomic and metabolomic analysis reveals the regulation mechanism of early bolting and flowering in two cultivars of *Angelica sinensis*

Chenghao Zhu [a], Yu Bai [a], Yuan Jiang [a], Yuanfan Zhang [a], Shangtao Wang [a], Fusheng Wang [b,*], Zhirong Sun [a,**]

[a] *School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing, 102488, China*
[b] *Dingxi Academy of Agricultural Sciences, Dingxi, 743000, Gansu, China*

ARTICLE INFO

Keywords:
Angelica sinensis
Early bolting and flowering
Transcriptomics
Metabolomics
Gene-regulatory networks

ABSTRACT

The root of *Angelica sinensis* is utilized in Traditional Chinese medicine to enhance blood replenishment and facilitate blood circulation. The early bolting and flowering (EBF) of *A. sinensis*, however, compromises the quality of the roots and restricts the yield of medicinal substances. The study was conducted to compare the transcriptomic and metabolomic profiles between EBF plants and normal plants of two cultivars of *A. sinensis*, followed by validation of the transcriptome results using qRT-PCR. There were 3677 DEGs in EBF plants compared to normal plants of cultivar 2 (Mingui No.2), and cultivar 4 (Mingui No.4) was 3354. The main differential metabolites in the EBF and normal plants were phenolic acids, flavonoids, lignans, and coumarins. The analysis of 5 EBF-related pathways revealed 28 genes exhibiting differential expression and 5 metabolites showing differential accumulation. The expression of the *Lhcb5*, *Lhcb2*, *Lhcb6*, *Lhcb1*, *Lhca4*, *ATPG1*, *EGLC*, *CELB*, *AMY*, *glgA*, *CYCD3*, *SnRK2*, *PYL*, *AHK2*, *AUX1*, *BSK*, *FabI/K*, *ACACA* and *FabV* decreased and the expression of the *PsbR*, *PsbA*, *LHY*, *FT*, *CO*, *malQ*, *HK*, *GPI* and *DELLA* increased in EBF plants. In addition, the Abscisic acid, D-Glucose-6P, α-D-Glucose-1P, NADP+, and ADP were more significantly enriched in EBF plants. The findings offer novel perspectives on the EBF mechanisms in *A. sinensis* and other medicinal plants of the Apiaceae family.

1. Introduction

The perennial herb *Angelica sinensis* (Oliv.) Diels, belonging to the Apiaceae family, is primarily cultivated in northwestern regions of China such as Gansu, Sichuan, and Yunnan provinces. Traditionally utilized for its dried root, this herb is known for its blood replenishing properties, ability to enhance blood circulation and regulate menstruation, alleviate pain, and provide intestinal moisture for relieving constipation [1,2]. The majority of the market supply of wild *A. sinensis* comes from artificial cultivation due to increasing demand, as it is scarce and grows at high altitudes [3]. For industrialized cultivation, seeds are sown in May, and seedlings are collected in autumn and overwintered indoors. Subsequently, the seedlings are transplanted for vegetative growth and either harvested in the autumn of the second year to obtain non-lignified roots or left in the field until June [4]. The early bolting and flowering (EBF) in

* Corresponding author. Dingxi Academy of Agricultural Sciences, Dingxi, Gansu 743000, China.
** Corresponding author. School of Chinese Materia Medica, Beijing University of Chinese Medicine, Beijing 102488, China.
E-mail addresses: wangfs1974@163.com (F. Wang), szrbucm67@163.com (Z. Sun).

https://doi.org/10.1016/j.heliyon.2024.e28636
Received 24 September 2023; Received in revised form 12 March 2024; Accepted 21 March 2024
Available online 27 March 2024

A. sinensis occurs during the second year, resulting in a significant reduction in both root yield and quality due to root lignification and degradation of bioactive compounds [5].

With the development of the *A. sinensis* industry, controlling the EBF becomes an urgent issue. Plants must undergo vernalization and long-day conditions to bolt and flower. Refraining from exposure to vernalization or long-day conditions can diminish the EBF rate of *A. sinensis* [6,7]. The EBF can be influenced by various factors, such as germplasm selection, seedling age and weight, geographical coordinates (latitude and longitude), and soil conditions [4]. The previous studies on the EBF of *A. sinensis* have concentrated on its physiology and ecology [8,9]. Some studies have already reported that many genes and pathways can affect the EBF of *A. sinensis*, for example, *COL3*, *AS1*, *SPL5*, and *HDR1* in the photoperiodic pathway, *GA2OX8* in the GA pathway, and *WIN1* and *PAS2* in the VLCFA biosynthesis [4–7,10,11]. How to effectively explore the essential genes from these large numbers of genes remains a scientific issue.

In this study, two *A. sinensis* cultivars were selected as experimental materials. Among them, cultivar 2 (Mingui No.2) was obtained through the systematic selection method of the wild plant. The stem at the flowering and seed setting stage exhibited a tender green color, with a stable EBF rate of 43%. Cultivar 4 (Mingui No.4) was derived by subjecting the original cultivar to irradiation using a medium energy ion beam of 55MeV/u 40Ar17+ at a dose of 2.5Gy. The stem at the flowering and seed setting stage displayed a dark purple hue, while maintaining a consistent EBF rate of 14% [12] (Supplementary Figs. S1 and S2). The EBF rate of the two cultivars varies in the same region and growth stage. One hypothesis of the difference is the genetic variation within the genes.

We conducted an analysis of the transcriptome and metabolome to explore the disparities between EBF plants and normal plants in these two cultivars. One hypothesis is the mutations occurred in the EBF related genes might be the internal cause of the different EBF rates among cultivars and could potentially assist with mining the relevant genes that control EBF. We summarized five biosynthetic pathways related to EBF, and comprehensively proposed a hypothesis on the regulatory pathways affecting EBF in *A. sinensis*. One good attempt is to find differential genes between EBF plants and normal plants of different germplasms to evaluate the key genes of EBF, followed a gene editing to solve the problem of EBF in *A. sinensis* through molecular breeding.

2. Materials and methods

2.1. Plant material

The two-year-old *A. sinensis* cultivars utilized in this study were identified by Director Fusheng Wang of Dingxi Academy of Agricultural Sciences. The experiment samples were collected at the production base of Dingxi in Gansu Province (35°23′53.25″ N, 105°1′37.98″E, altitude 2390 m) in August, China. We randomly collected fresh tender leaves from 20 EBF plants and 20 normal plants (Mingui No.2 and No.4). After thorough mixing, the sample was carefully weighed and transferred into a 50 mL centrifuge tube before being stored in liquid nitrogen. We collected 4 g samples for each treatment, which were divided into two parts and used for RNAseq and metabolomics respectively. The samples included EBF plants (EZ) and normal plants (EW) of Mingui No.2, EBF plants (SZ) and normal plants (SW) of Mingui No.4. Each group had three biological replicates.

2.2. RNA extraction, Illumina library construction, and sequencing

The Spin Column Plant Total RNA Purification Kit (Sangon Biotech, Shanghai, China) was employed for total RNA extraction. The purity of the extracted RNA was evaluated using a NanoPhotometer spectrophotometer (IMPLEN, Los Angeles, CA, USA). For quantification of RNA, the Qubit RNA Assay Kit (Life Technologies, Carlsbad, CA, USA) was utilized [13]. RNA integrity was assessed using an Agilent Bioanalyzer 2100 system (Agilent Technologies, Santa Clara, CA, USA). The results of RNA integrity testing were shown in Supplementary Table S1. As previously described, Illumina sequencing libraries were constructed [13]. The cDNA libraries were sequenced using the Illumina HiSeq platform (Illumina Inc., San Diego, CA, USA) by Wuhan MetWare Biological Science and Technology Co., Ltd. (Wuhan, China). The index-coded samples were clustered on a cBot Cluster Generation System using TruSeq PE Cluster Kit v3-cBot-HS (Illumia), following the manufacturer's instructions. Subsequently, library preparations were subjected to sequencing on an Illumina platform, generating 150 bp paired-end reads.

2.3. Transcriptome data analysis

The sequences were subjected to adapter trimming, and reads with $\geq$5 uncertain bases or over 50% of bases with Qphred $\leq$20 were filtered out using fastp(v0.23.2) in order to obtain high-quality clean reads. Subsequently, the GC content of the resulting clean reads was calculated. The Fast QC tool also generated Q20 and Q30 values for assessing the quality of bases. Trinity(v2.13.2) was used to concatenate 12 high-quality clean reads to obtain the transcriptome sequence of *A. sinensis*(Trinity.fa), then we used Corset(v1.09) to remove redundancy to obtain the Unigene sequence. The Unigene sequence (Unigene.fa) was used to align the high-quality reads, and subsequently, the gene expression level was determined.

TransDecoder(v5.3.0) was utilized to identify potential coding regions within transcript sequences obtained from de novo RNA-Seq transcript assembly using Trinity(v2.13.2). The gene function was annotated using Diamond (v2.0.9) based on the following databases: Nr, Swiss-Prot, Trembl, KEGG, GO, KOG/COG, and Pfam. The annotation process utilized a BLAST + algorithm with a threshold E-value $<1.0 \times 10^{-5}$ [14,15]. The sequenced reads were aligned to the unigene sequence using Bowtie(v2.4.5), and gene expression levels were quantified using RSEM(v1.3.1). Gene expression levels were measured in FPKM values. To identify differentially expressed genes (DEGs), DESeq2 (v1.22.2) was employed with the following criteria: $P < 0.005$, $|\mathrm{Log_2FC}| \geq 1$, and false discovery rate (FDR) $<$ 0.05 [16,17]. The GO enrichment analysis was conducted using the P.A Wallenius non-central hypergeometric distribution, and

KOBAS 2.0(v2011) was employed for analyzing the KEGG pathway enrichment of DEGs [13].

2.4. Metabolome analysis

The biological samples were subjected to freeze-drying using a vacuum freeze-dryer (Scientz-100F). Subsequently, the freeze-dried sample was pulverized for 1.5 min at 30 Hz using a mixer mill (MM 400, Retsch) with a zirconia bead. To dissolve the lyophilized powder, 50 mg was mixed with 1.2 mL of a 70% methanol solution and vortexed for 30 s every 30 min, repeating this process six times in total. After centrifugation at 12,000 rpm for 3 min, the extracts were filtered through a SCAA-104 membrane with a pore size of 0.22 μm (ANPEL, Shanghai, China) prior to UPLC-MS/MS analysis. The sample extracts were then analyzed using a UPLC-ESI-MS/MS system (UPLC, ExionLC™ AD; MS, Applied Biosystems 4500 Q TRAP). The analytical conditions were as follows: UPLC column - Agilent SB-C18 (1.8 μm, 2.1 mm × 100 mm), and the mobile phase consisted of solvent A (pure water with 0.1% formic acid) and solvent B (acetonitrile with 0.1% formic acid) [18]. The acquisition of LIT and triple quadrupole scans was performed on a QTRAP-MS equipped with an ESI Turbo Ion-Spray interface, utilizing the AB Sciex QTRAP 4500 System in positive ion mode. The instrument control was facilitated by Analyst software (v1.6.1). The solvent system, gradient program, and ESI source operation parameters were implemented following previously established methodologies [18].

The qualitative analysis of primary and secondary MS data was conducted by querying the internal database using a self-compiled database provided by Wuhan MetWare Biological Science and Technology Co., Ltd. The Variable Importance of Projection (VIP) score from the OPLS-DA model was utilized to identify the most significant metabolites that differentiate between treatments. Metabolites with significant differences in content were set with thresholds of VIP $\geq$1 and fold change $\geq$2 or $\leq$0.5. A Principal component analysis (PCA) was used to analyze the variability between groups and within groups. The functional annotation of differentially accumulated metabolites (DAMs) was performed based on KEGG pathways [13].

2.5. Co-joint analysis of the transcriptome and metabolome

Based on the results of differential metabolite analysis in this experiment, we integrated the results from differential gene analysis to map the differentially expressed genes and metabolites onto the KEGG pathway map for the same comparative group. Correlation analysis was performed using quantitative values of genes and metabolites across all samples. The correlation analysis was performed using the Cor function in R(v3.6.3) to calculate the Pearson correlation coefficient between genes and metabolites. A correlation coefficient greater than 0.80 with a P-value less than 0.05 was considered statistically significant. Subsequently, all correlation calculations for differentially expressed genes and metabolites were selected, and a correlation clustering heatmap was generated [13].

2.6. qRT-PCR verification

The validation of ten DEGs associated with EBF was performed using qRT-PCR [19]. The Actin gene was utilized for internal control, and the relative expression levels were determined using a $2^{-\Delta\Delta Ct}$ method. For each gene, three technical repeats were used, and three independent experiments were conducted. A previous sequencing sample was utilized for RNA validation. The HiScript® II Q RT SuperMix for qPCR (+gDNA wiper) (Vazyme, Nanjing, China) was employed in all reactions, in conjunction with the CFX96 Real-Time PCR System (Bio-Rad, CA, USA) [20]. Supplementary Table S2 contains the primer sequences.

2.7. Statistical analysis

The validation experiment for qRT-PCR consisted of three biological and three technical replicates to ensure accurate determination of PCR efficiency. Statistical analysis was conducted using ANOVA and Duncan multiple comparison tests, as determined by SPSS(v22.0).

Table 1
Overview of the sequencing data obtained from the 12 samples.

Sample	Clean Reads	Clean Base(G)	Q20(%)	Q30(%)	GC Content(%)
EW1	42,696,226	6.4	97.23	92.02	41.96
EW2	46,785,338	7.02	98.33	94.67	41.89
EW3	49,853,364	7.48	98.16	94.24	42.15
EZ1	45,625,468	6.84	97.95	93.77	42.49
EZ2	44,926,932	6.74	98.24	94.47	42.44
EZ3	47,006,504	7.05	98.32	94.68	42.11
SW1	49,282,886	7.39	98.27	94.59	42.16
SW2	43,635,338	6.55	97.94	93.75	42.38
SW3	42,178,954	6.33	98.27	94.66	42.3
SZ1	48,909,926	7.34	98.26	94.5	42.41
SZ2	48,385,720	7.26	98.06	94.04	43.29
SZ3	45,397,206	6.81	98.09	94.11	42.67

3. Results

3.1. Overview of the A. sinensis transcriptomic analysis

The sequencing of 12 constructed libraries yielded a total of 554.7 million high-quality clean reads. The Q20, Q30, and GC content ranged from 97.23% to 98.33%, 92.02%–94.68%, and 41.89%–43.29%, respectively. The GC content of EBF plants was generally observed to be higher compared to that of normal plants (Table 1). A total of 207,777 transcripts were generated from the clean reads, with an average contig size of 1234 bp and an N50 contig size of 2223 bp (Supplementary Table S3). The functional annotations for all unigenes were provided in Supplementary Table S4. *A. sinensis* exhibited a total of 113,404 annotated unigenes (54.58%). The NR database identified 82,457 (72.71%) significant matches. *A. sinensis* transcript sequences were 68.16% similar to that of *Daucus carota* subsp. sativus, 19.65% similar to *Quercus suber*, and 0.55% similar to *Hordeum vulgare* subsp. vulgare (Supplementary Fig. S3).

The transcripts of *A. sinensis* were categorized into three distinct groups of Gene Ontology (GO) terms, specifically biological process (BP) (195,646), cellular component (CC) (68,927), and molecular function (MF) (93,127), as presented in Supplementary Table S5. The terms were subsequently segregated into 48 subcategories. Within the BP category, the most frequently observed terms were cellular process, metabolic process, and responses to stimuli. In the CC category and MF category, cellular anatomical entities and protein-containing complexes were the predominant terms for describing binding and catalytic activities respectively, as depicted in Supplementary Fig. S4. There were 47,626 unigenes assigned to 139 KEGG pathways and 51,947 unigenes were assigned to 25 KOG pathways in *A. sinensis*. In KOG pathways, 593 genes belonged to nucleotide transport and metabolism, 4746 genes were signal transduction mechanisms and 3064 genes were energy production and conversion. Analyzing metabolic pathways in *A. sinensis* was made possible by these annotation results.

The Pearson's Correlation Coefficient was utilized as an evaluation metric to assess the biological repeated correlations. A Pearson's correlation coefficient ranging from 0.39 to 0.79 was identified between the EZ and EW samples (Fig. 1A). A Pearson's correlation coefficient ranging from 0.65 to 0.80 was observed between the SZ and SW samples, while a Pearson's correlation coefficient ranging from 0.67 to 0.97 was found between the EZ and SZ samples. According to these findings, there were significant disparities in gene expression between EBF plants and normal plants of the same cultivar, as well as variations observed among EBF plants from different cultivars. Fig. 1B illustrated distinct gene expression profiles between EBF and normal plants across both cultivars (PCA).

3.2. Different expression genes analysis of A. sinensis

Four pairwise comparisons were established (EW vs. SW, EZ vs. EW, EZ vs. SZ, and SZ vs. SW) to investigate the DEGs in *A. sinensis* for EBF and normal plants. The number of transcripts that had significant changes was determined by volcano plots. The comparison between SZ and SW identified a total of 3354 differentially expressed genes (DEGs), with 1326 up-regulated and 2028 down-regulated genes. Similarly, the comparison between EZ and EW revealed a total of 3677 DEGs, including 1970 up-regulated and 1707 down-regulated genes. Additionally, the comparison between EW and SW showed that there were 268 up-regulated DEGs and 191 down-regulated DEGs. Conversely, the comparison between EZ and SZ indicated that there were 1955 up-regulated DEGs and 300 down-regulated DEGs (Fig. 2A). By searching for literature reports on pathways related to EBF, 189 transcription factors (115 up-regulated and 74 down-regulated) were selected from 3354 differential transcription factors in SZ vs. SW, and EZ vs. EW was 191 up-regulated and 92 down-regulated, while EZ vs. SZ was 38 up-regulated and 28 down-regulated (Fig. 2B–D). A Venn diagram showed that 7652 DEGs were obtained across the four treatments, 4 of which were shared, the functions of four common DEGs were NPH3 family protein expression, fatty acid dehydrogenase synthesis, V-ATPase subunit H synthesis, and D-mannose binding lectin expression. In the comparison of EBF plants of the same cultivar and the normal plants, there were 2283 and 1960 independent genes of Mingui

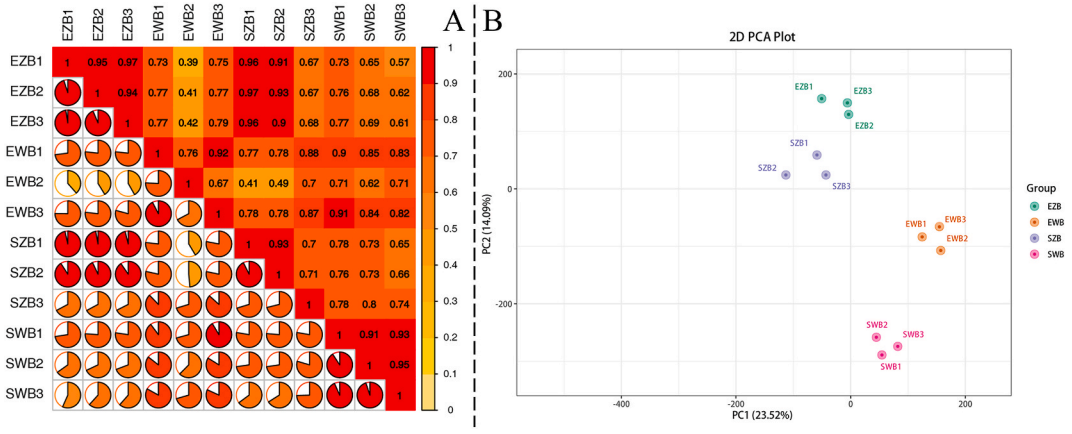


Fig. 1. Analyses of *A. sinensis* transcriptomic parameters. (A) Pearson's correlations between EBF and normal plants of two cultivars; (B) PCA analysis.

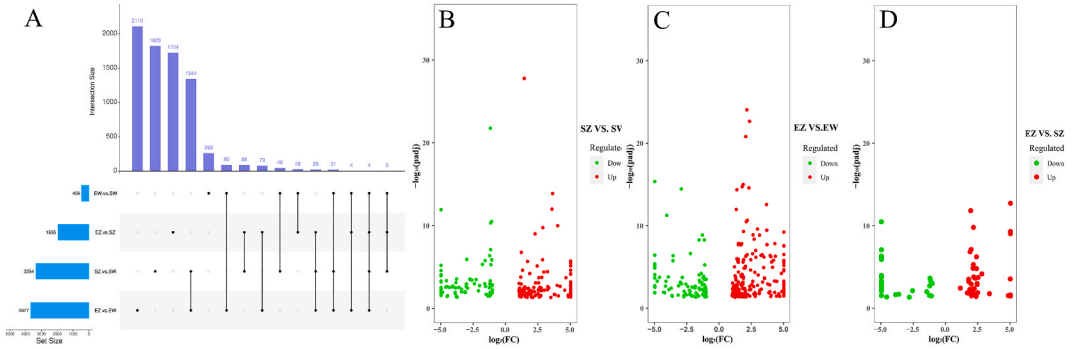


Fig. 2. Differential transcription factors among different groups. (A) Overall situation, (B–D) Transcription factors related to early bolting in each group.

No.2 and Mingui No.4 respectively, as well as 1394 shared genes.

The GO enrichment analysis results revealed that the four comparisons' DEGs were significantly enriched in 1,689, 4,037, 3,749, and 4076 GO terms, respectively. The cellular component (CC) category exhibited plant-type cell walls (GO: 0009505) as the most enriched components, whereas peroxidase activity (GO: 0004601) was the top-ranking terms in the molecular function (MF) category. The biological process (BP) category was characterized by the highly represented terms of response to oxygen levels (GO: 0070,482) and decreased oxygen levels (GO: 0036,293) (Fig. 3, SZ vs. SW).

For the EW vs. SW, the most enriched term of BP, CC, MF was plant organ senescence (GO: 0090,693), nuclear chromosome (GO: 0000228), and the MF category was oxidoreductase activity (GO: 0016,717) respectively. Under the comparison of EZ vs. EW, in the BP category, the most enriched terms were secondary metabolite biosynthetic (GO: 0044,550), the CC category was plant-type cell wall (GO: 0009505), and the MF category was UDP-glucosyltransferase activity (GO: 0035,251). For the comparison between EZ and SZ, the BP category exhibited significant enrichment in terms related to translation within the cytoplasm (GO: 0002181), the CC category was categorized into peptidase complex (GO: 1905368), and the MF category was oxidoreductase activity (GO: 0016,717) (Supplementary Figs. S5–S7).

An analysis of KEGG revealed the key biological pathways of *A. sinensis*, and the variation in EBF rate among different cultivars could provide insights into the changes occurring within these pathways. The six most significantly enriched pathways for SZ vs. SW (Fig. 4) included metabolic pathway, biosynthesis of secondary metabolites, ribosome, carbon metabolism, plant-pathogen interaction, and plant hormone signal transduction. The EZ group demonstrated significant enrichment in various pathways, including secondary metabolite biosynthesis, carbon metabolism, ribosomes, and the MAPK signaling pathway, when compared to the SZ group. Similarly, the metabolic pathway, biosynthesis of secondary metabolites, plant-pathogen interaction, signal transduction, and ribosome were found to be the top five pathways enriched in EZ vs. EW. In contrast, the top five significantly enriched pathways for EW vs. SW included the metabolic pathway, biosynthesis of secondary metabolites, plant hormone signal transduction, biosynthesis of amino acids, and starch along with the MAPK signaling pathway (Supplementary Figs. S8–S10).

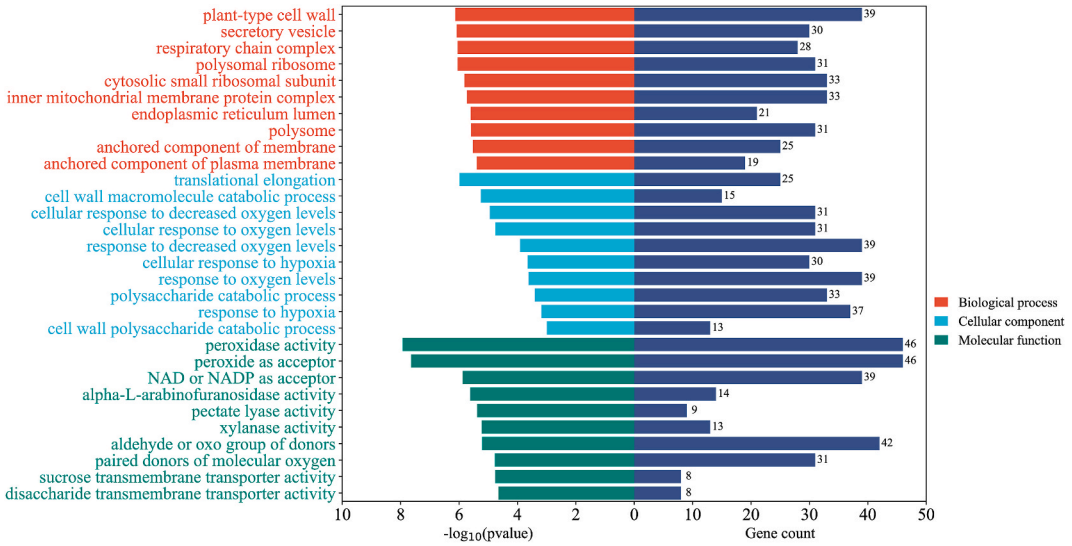


Fig. 3. Functional classifications of genes that are differentially expressed in SZ compared to SW according to gene ontology.

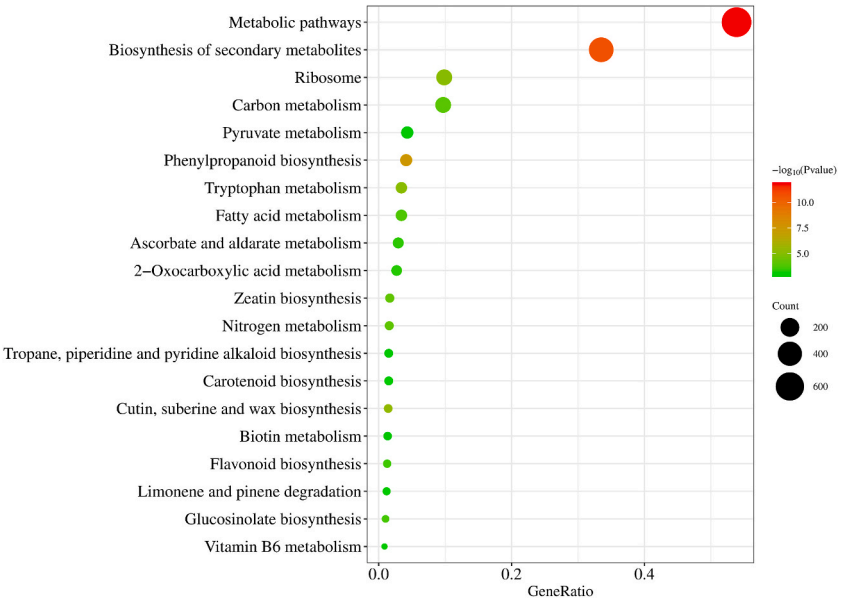


Fig. 4. Statistics of KEGG enrichment of SZ vs. SW.

3.3. Different expressed transcription factor analysis

There were 167 TFs differentially expressed between SZ and SW (Supplementary Table S6). TFs in this study were divided into 46 families, with C_2H_2 (17), bHLH (12), bZIP (11), MADS-MIKC (10), and AP2/ERF (10) being the major groups. With respect to type of regulation, 7C_2H_2s and 8 bHLHs were up-regulated, while 10 WRKYs and 4 bHLHs were down-regulated. Most TFs were up-regulated in SZ as compared with SW.

The significant TFs identified included AP2/ERF-AP2(23), C_2H_2 (16), bHLH (14), NAC (13), MADS-MIKC (11), bZIP (9), and more than half of TFs were up-regulated in 235 TFs in EZ vs. EW (Supplementary Table S7). The significant TFs identified in EZ vs. SZ (Supplementary Table S8) included C_2H_2 (13), bHLH (4), HMG (4), and 43 TFs were up-regulated in 58 TFs. The major TFs identified included C_2H_2 (3), Jumonji (3), NAC (3), and 23 TFs were up-regulated in 33 TFs in EW vs. SW (Supplementary Table S9).

3.4. Different accumulated metabolites analysis

Prior to conducting an analysis on the differentially accumulated metabolites (DAMs), the level of variability among and within groups was assessed using PCA. There was good separation between the 12 samples (Fig. 5A). A total of 38.54 percent of the variability was accounted for by PC1, while 13.77% was accounted for by PC2. Flavonoids (19.07%), phenolic acids (15.58%), and lipids (12.84%) were the most prevalent metabolites in all samples (Fig. 5B).

A total of 403 differentially expressed genes (219 up-regulated and 184 down-regulated) were identified in the comparison between SZ and SW. They were 72 phenolic acids, 66 flavonoids, 51 lignans and Coumarins, 33 amino acids and their derivatives, 32 lipids, 25 alkaloids, 24 organic acids, 22 nucleotides and derivatives, 17 terpenoids, 2 quinones, and 59 other metabolites (Supplementary Table S10). In NCBI, 10 DAMs exhibiting large differential multiples and high content related to EBF were found to be

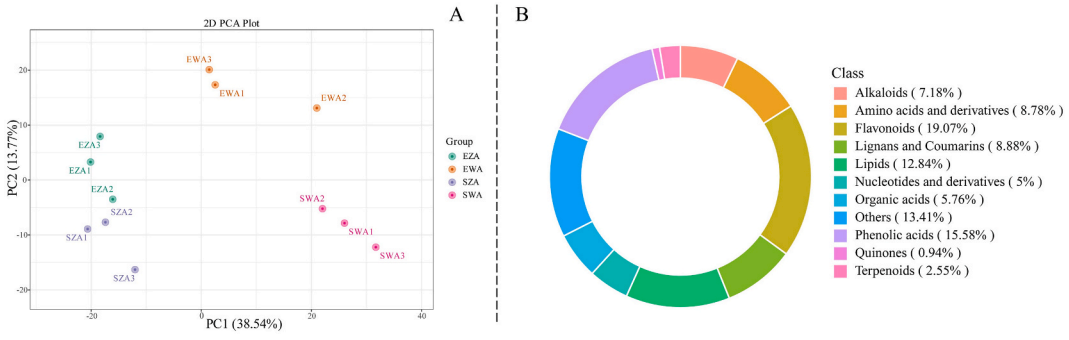


Fig. 5. Metabolics analysis using principal component analysis (PCA) and category of metabolites.

significantly up-regulated or down-regulated. The results indicated that lignans, such as Syringaresinol-4′-*O*-(6″-acetyl) glucoside, Medioresinol-4,4′-di-*O*-glucoside and Syringaresinol-4′-*O*-glucoside, were significantly up-regulated in EBF plants. However, 4,5-*O*-Decaffeoylquinic Acid Methyl Ester (Phenolic acids) and Naringenin-7-*O*-glucoside (Flavanones) were significantly down regulated (Fig. 6A). In EZ vs. EW, the 325 DAMs (198 up-regulated and 127 down-regulated) were identified, they were 60 phenolic acids, 58 flavonoids, 40 lignans and coumarins, 32 lipids, 27 amino acids and derivatives, 23 alkaloids, 18 organic acids, 12 nucleotides and derivatives, 12 terpenoids, 1 Quinones and 42 other metabolites (Supplementary Table S11). In the comparison between SW and EW, a total of 212 DAMs (147 up-regulated and 65 down-regulated) were identified. These metabolites comprised 36 phenolic acids, 32 flavonoids, 24 lignans and coumarins, 27 lipids, 11 amino acids and derivatives, 20 alkaloids, 11 organic acids, 5 nucleotides and derivatives, 10 terpenoids, 3 quinones, as well as an additional set of other metabolites (Supplementary Table S12). The comparison between EZ and SZ revealed the identification of 94 DAMs (89 up-regulated and 5 down-regulated). These included a variety of compounds, such as 12 phenolic acids, 13 flavonoids, 30 lignans and coumarins, 7 lipids, 5 amino acids and derivatives, 11 alkaloids, 3 organic acids, 2 nucleotides and derivatives, 2 terpenoids, 1 quinone, and an additional set of eight other metabolites (Supplementary Table S13).

The 403 DAMs were classified and assigned to 25 metabolic pathways using a combined KEGG enrichment analysis in the comparison between SZ and SW (Fig. 6B). The primary pathways identified were metabolic pathways (ko01100), followed by the biosynthesis of secondary metabolites (ko01110) and cofactors (ko01240). Additionally, variations were observed in plant hormone signal transduction (ko04075), carbon metabolism (ko01200), and photosynthesis (ko00195). The pathways that demonstrated the most significant enrichment in EZ compared to SZ were those associated with the biosynthesis of diverse plant secondary metabolites (ko00999) and tryptophan metabolism (ko00380), as illustrated in Supplementary Figs. S11–14.

The DEGs and DAMs were selected from the overall correlation results for further correlation analysis, based on a Pearson correlation coefficient greater than 0.80 and a P-value less than 0.05. The network relationship between genes and metabolites was established through KGML analysis. The findings revealed a consistent trend in the majority of cases, where gene expression and metabolite changes were observed to align in pathways such as photosynthesis, photosynthesis-antenna proteins, starch and sucrose metabolism, as well as plant hormone signal transduction associated with bolting and flowering. The genes expression and the change of metabolites had the same trend in most cases, such as *PsbA* and NADP, *AHP* and Abscisic acid (Fig. 7). This finding underscored the importance of genes and metabolites interactions in biological processes. Additionally, the metabolism pathways of starch and sucrose in SZ demonstrated upregulation of alpha-D-glucose 1-phosphate and *AMY* compared to SW, as depicted in Supplementary Figs. S15–S18.

3.5. *Integrated analysis of DEGs and DAMs in the photosynthesis-related biosynthesis pathway*

In the comparison of DEGs and DAMs in the photosynthesis-related biosynthesis pathway between EBF and normal plants(Fig. 8), we found that compared with normal plants, the *psbR* and *psbA* in EBF plants were significantly up-regulated, while *ATPG1*, *Lhca4*, *Lhcb1*, *Lhcb6*, *Lhcb2*, and *Lhcb5* were otherwise significantly down-regulated (genes with the same expression in both cultivars). In addition, the NADP+ and ADP were significantly up-regulated in EBF plants. The process of photosynthesis involves two photochemical reactions. One is known as the Photosystem II, which absorbs short-wavelength red light at 680 nm. The other is called Photosystem I, which captures long-wavelength red light at 700 nm. The PS II complex is composed of essential components, namely the chlorophyll-binding inner peripheral antenna proteins (CP47 and CP43) and two core polypeptides (D1 and D2), which are the largest proteins in the complex located on the thylakoid membrane of the chloroplast [21]. The outer layer of PS II is a light-harvesting fragment complex (LHCII) formed by combining photosynthetic pigments and proteins. In our study, we observed a significant up-regulation of *PsbA* in D1 and *PsbR* in D2, along with a remarkable down-regulation of *Lhcb1*, *Lhcb6*, *Lhcb2*, and *Lhcb5* in LHCII. Based on these findings, it could be inferred that these changes were conducive to water decomposition and facilitate the transfer of electrons from water to PS I. Besides, notably down-regulated *Lhca4* in LHCI and *ATPG1* in ATP synthase might increase the production of ADP and NADP + to provide energy for EBF.

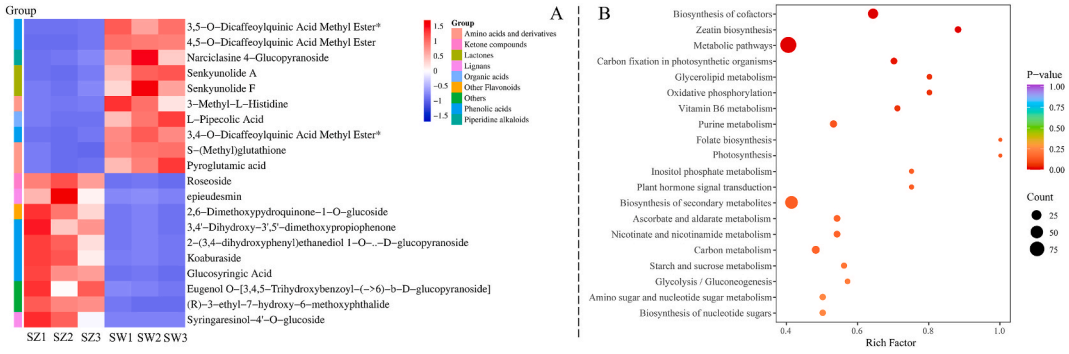


Fig. 6. The key DAMs and their enriched KEGG pathways of SZ and SW.

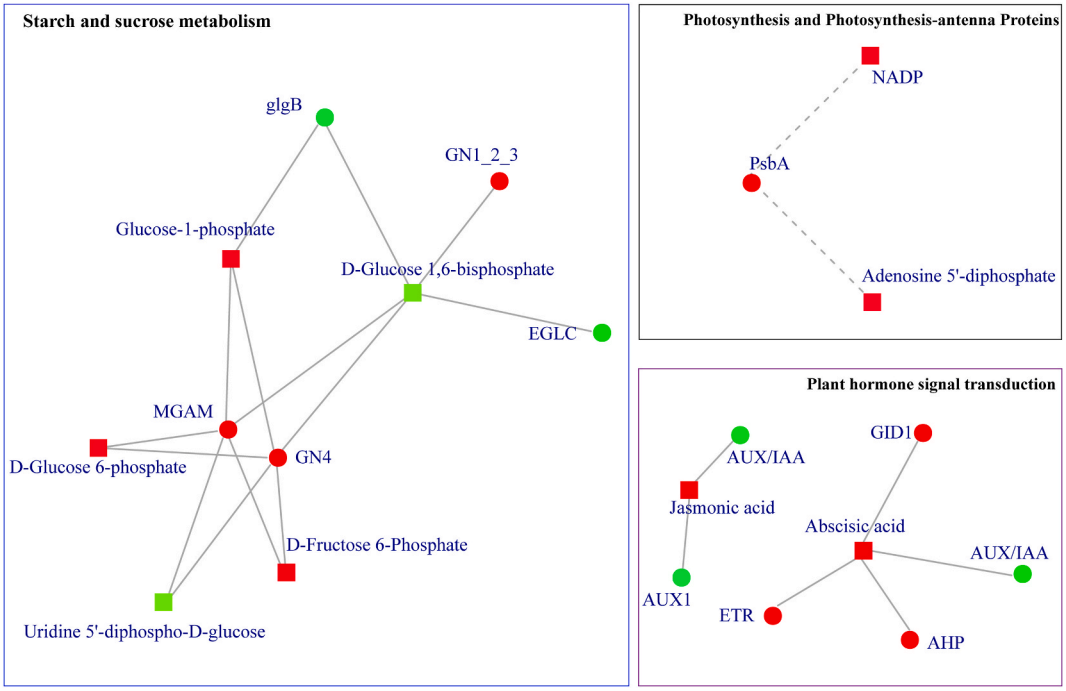


Fig. 7. Interactions between genes and metabolites in pathways related to EBF.

3.6. Integrated analysis of DEGs and DAMs in the circadian rhythm pathway

Plant flowering in photoperiodic induction requires an appropriate circadian rhythm and the formation of certain balance between photoreceptors (photosensitive pigments, cryptocyanins, and ultraviolet B receptors) or some related substances. If the length of day and night changes, the balance will be disrupted, leading to the expression of some genes that promote or inhibit flowering, thereby initiating or inhibiting flowering [22]. CONSTANS(*CO*) protein is an output channel regulated by the circadian rhythm of the photoperiod-biological clock. *CO* encodes a B-type zinc finger domain protein, and its expression is regulated by the circadian rhythm of the biological clock. The receptor and the biological clock interact under long sunlight conditions to activate *CO* production in phloem companion cells, the *CO* then activates TWINSISTER OF FT (*TSF*) and FLLOWERING LOCUS T (*FT*) genes, which are downstream targets of *CO*. Flower meristems are eventually induced to promote flowering [23,24]. In the circadian rhythm pathway between EBF and NP of two cultivars (Fig. 9), we found that *LHY*, *CO,* and *FT* were significantly up-regulated in EBF plants and possibly were the key genes inducing EBF. Under this pathway, we also made a discovery that significantly up-regulated *CDF1* between the EBF of the two cultivars (EZ vs. SZ) could possibly be one of the genes leading to the difference in EBF rates.

3.7. Integrated analysis of DEGs and DAMs in the starch and sucrose metabolism pathway

The levels of D-Glucose-6P and α-D-Glucose-1P were significantly up-regulated in the starch and sucrose metabolism pathway (Fig. 10). The expression of *EGLC*, *CELB*, *AMY*, and *glgA* in the D-Glucose synthesis pathway was significantly down-regulated in EBF plants, whereas *GPI* and *HK* involved in D-Fructose-6P synthesis were significantly up-regulated. In this pathway, we also found that *HK* and *treX* were significantly up-regulated between EBF plants of the two cultivars, while *bglX* was significantly down-regulated (EZ vs. SZ).

3.8. Integrated analysis of DEGs and DAMs in the plant hormone signal transduction pathway

It was observed that in this pathway (Fig. 11), seven DEGs and one DAM were identified. *AUX1* in tryptophan metabolism, *AHK2* in zeatin biosynthesis, *PYL* and *SnRK2* in carotenoid biosynthesis, and *BSK* and *CYCD3* in brassinosteroid biosynthesis were significantly down-regulated in EBF plants. Besides, *DELLA* involved in diterpenoid biosynthesis were significantly up-regulated, and abscisic acid was significantly up-regulated. Under this pathway, we also found that *AUX/IAA* in tryptophan metabolism and *PR-1* in phenylalanine metabolism were significantly up-regulated between EBF plants of the two cultivars (EZ vs. SZ), which might presumably contribute to the plant growth and disease resistance of cultivar 2.

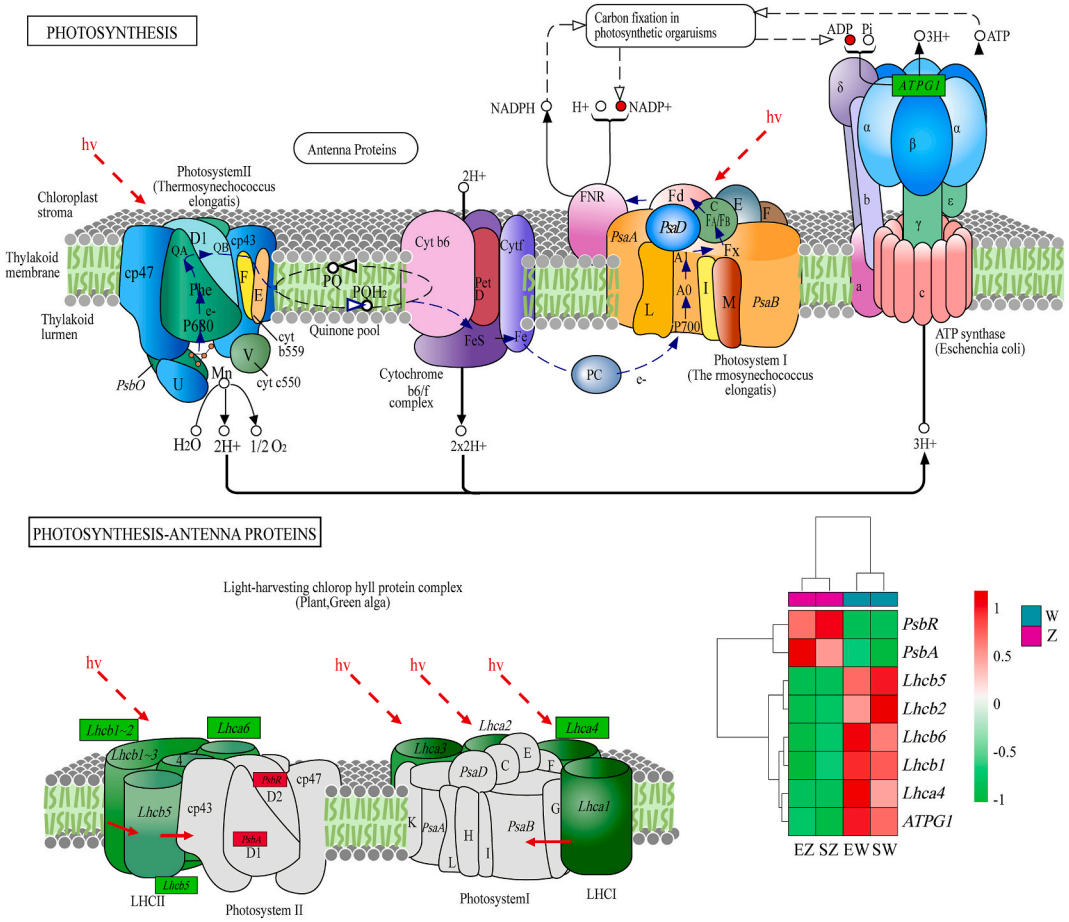


Fig. 8. The metabolic pathway map of biosynthesis related to photosynthesis. The heat map was drawn based on the FPKM values of differential expression levels. Red/green represents up/down-regulated genes, and small red circles indicated the differentially accumulated metabolites. W represented normal plants, and Z represented EBF plants, the same as below. All data in the 3.6 to 3.9 showed a common trend of E and S. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

3.9. Integrated analysis of DEGs and DAMs in the very-long-chain fatty acid biosynthesis pathway

VLCFAs present in plants serve multiple functions, including serving as an energy source for the synthesis of glycerides, biofilm lipids, and sphingolipids in seeds. Additionally, they act as precursors for the biosynthesis of cuticular wax. Moreover, these VLCFAs are found in the epidermis of plants where they play a crucial role in regulating water retention, enhancing drought resistance, and promoting growth [25]. DEGs involved in the VLCFAs biosynthesis pathway (Fig. 12) including acetyl-CoA carboxylase (*ACACA*), enoyl-[acyl-carrier protein] reductase (*FabV*), I (*FabI*), and II (*FabK*) were down-regulated in EBF plants as compared to normal plants. A slight decrease in the content of VLCFAs could promote cell proliferation, thereby promoting the growth of buds and causing EBF. Long-chain acyl CoA synthase (*ACSL*), one of the possible genes resulting in higher EBF rate of cultivar 2 to cultivar 4, was significantly up-regulated in the EBF plants (EZ vs. SZ).

3.10. The validation of gene expression through qRT-PCR

To ensure the accuracy of the transcriptome data, we conducted a random selection of 10 genes from 5 regulatory pathways associated with EBF in *A. sinensis* (Fig. 13). The relative expression levels (RELs) were found to be in accordance with the FPKM values, with down-regulated 0.8-, 0.4-, 0.7-, 0.7-, 0.3-, 0.1-, 0.4-, and 0.3-fold for the genes *glgA*, *AMY*, *CELB*, *EGLC*, *Lhcb6*, *AHK2*, *AUX1*, and *CYCD3*, respectively, in the EBF plants compared to the normal plants (cultivar 2). And with up-regulated 1.7-, and 10.3-fold for the genes *malQ* and *CO* in the EBF plants (cultivar 2). Notably, the CO gene expression was observed to be significantly elevated in cultivar 2 compared to cultivar 4 ($P < 0.001$). This finding suggests that the higher expression of the *CO* gene in cultivar 2 may contribute to the observed disparity in EBF rate between the two cultivars.

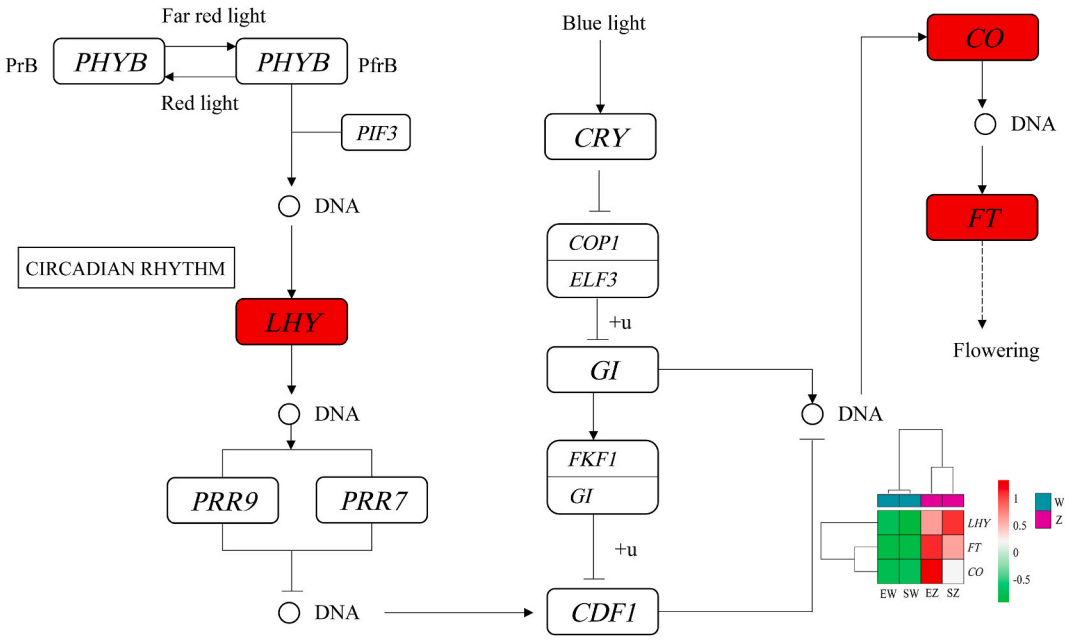


Fig. 9. Metabolic pathway map of the circadian rhythm.

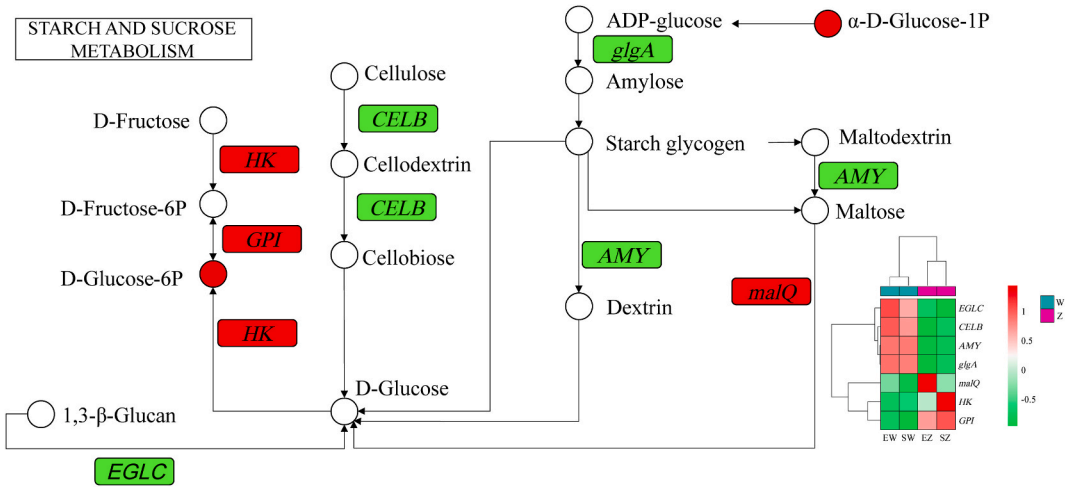


Fig. 10. Pathway map of the starch and sucrose metabolism.

4. Discussion

Early bolting and flowering (EBF) is the phenomenon of bolting in advance and entering reproductive growth, before the vegetative body of a plant is fully grown. Factors affecting the EBF of plants include genetic factors, ecological environmental factors, and growth materials. Currently, most studies focus on large quantity of vegetables with bolting characteristics, such as cabbage [26] and onion [27]. Despite the prevalence of perennial medicinal plants susceptible to endogenous biological factors, there remains a dearth of systematic investigation into the underlying mechanisms of this phenomenon within medicinal plant research. Among them, EBF is common phenomenon in Apiaceae medicinal plants, such as *Angelica dahurica* [28], *A. sinensis*, *Hansenia weberbaueriana* [29], and an unresolved puzzle in the production of traditional medicine. In this study, based on the investigation of the EBF problem of *A. sinensis*, we found that there were six cultivated species in actual production, of which two (Minggui No. 2 and No. 4) were the most typical. They were green and purple stems respectively, and the content of active components in the root also varies. The EBF rate of No. 2 had been steadily higher than No. 4 over the years. The transcriptomic and metabolomic analysis of these two cultivars was conducted to explore the key genes that determine their EBF. In previous studies, transcriptome analysis was conducted only on normal and EBF plants of a specific cultivar, without comparing differential metabolites. Additionally, it was found that the accumulation of

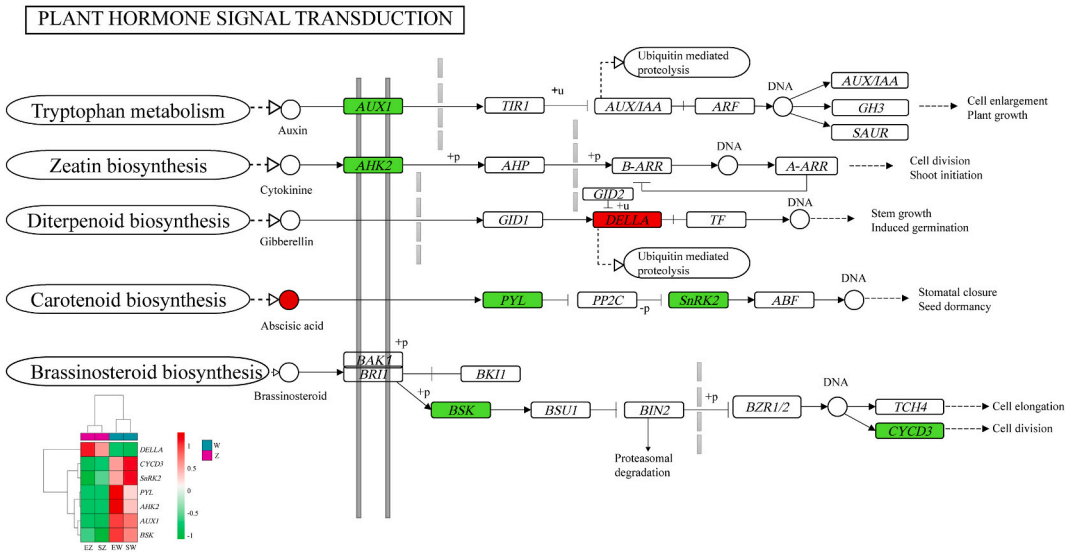


Fig. 11. Pathway map of the plant hormone signal transduction.

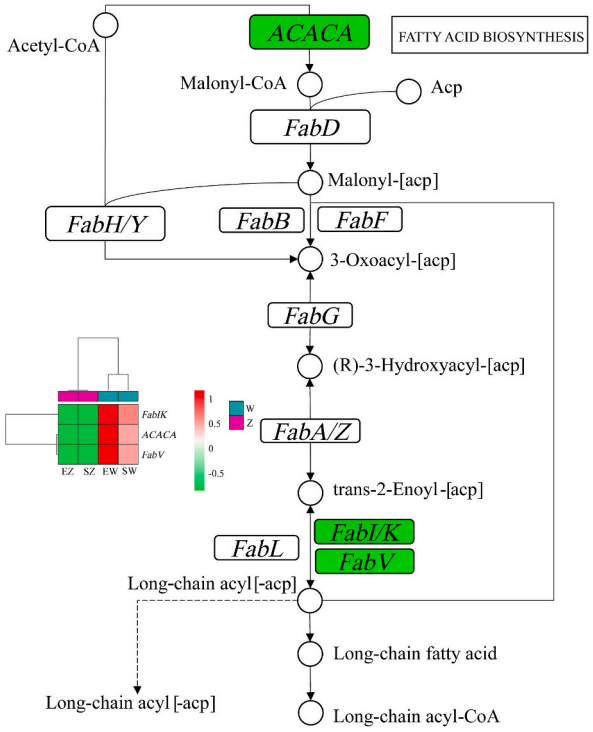


Fig. 12. Pathway map of the very-long-chain fatty acid biosynthesis.

endogenous hormones (ABA and ACC) in leaves and bulbs of Lilium formolongi can inhibit flowering [30]. Our findings indicate a significant accumulation of ABA in EBF plants, suggesting its potential role as a key metabolite in such plants. Furthermore, the expression patterns of ABA-related genes exhibited a positive correlation with stem length during the development of leaf lettuce, consistent with the outcomes of the investigation [31]. All evidence indicated that studying the changes in differential metabolites in EBF plants was necessary. In addition, the EBF rates of the two cultivars were different. From the perspective of germplasm comparison, it was inclined to refine and minimize the identification of important genes that affect EBF and combine gene editing technology to conduct germplasm innovation to reduce the occurrence of EBF in *A. sinensis* fundamentally.

Compared to the normal plants of the same period, the EBF plants of cultivar 2 and 4 had 3677 and 3354 DEGs, respectively. This

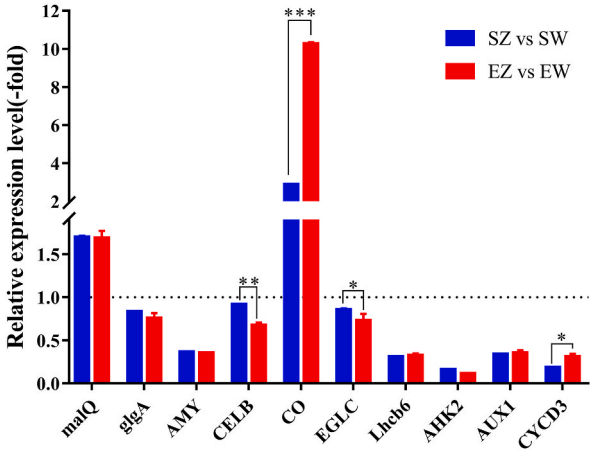


Fig. 13. The expression levels of 10 genes associated with EBF were compared between SZ and SW, as well as EZ and EW. "***" denotes a statistically significant distinction ($P < 0.001$).

indicated that multiple genes mediated the occurrence of EBF. The specific gene expression could activate or inhibit a certain biosynthetic pathway, leading to abnormal changes in metabolites and causing changes in physiological and growth indicators. The expression of RcPAP1 in the epichile before orchid flowering was found by Li to specifically activate the anthocyanin biosynthesis pathway, leading to significant accumulation of cyanin and resulting in a purple-red color [32]. The key active compounds in the traditional application of *A. sinensis* roots are Ferulic acid, Ligustilide, and Chlorogenic acid [33]. The present study revealed significant variations in the relative content of phenolic acids and flavonoids among different varieties of normal plants ($P < 0.005$). Notably, Minggui 4 exhibited a ferulic acid content that was 4.58 times higher than that observed in Minggui 2. Additionally, the differential metabolites observed in the EBF plants primarily consisted of phenolic acids, flavonoids, lignans, and coumarins when compared to the normal plants. The presence of EBF resulted in a reduction of active compounds in *A. sinensis*, consequently impacting the quality of the medicinal material.

There are at least four flowering induction pathways in *Arabidopsis thaliana*, including photoperiodic, autonomous and vernal pathways, carbohydrate or sucrose pathway, and gibberellin pathway [34,35]. The photoperiodic pathway is, to date, the most clearly explained pathway. Plant leaves sense light, in which light receptors sense light signals, triggering circadian rhythms based on length of day and night and intensity of light. A certain balance will be formed between the photoreceptor itself or some related substances. Changes in the length of day and night can disrupt this balance, leading to the expression of genes that promote or inhibit flowering, thereby initiating or inhibiting flowering [22]. The peripheral subunits (*PsbA* and *PsbR*) of the oxygen-evolving complex (OEC), which were responsible for stabilizing PSII, exhibited differential expression in EBF plants. Polymerization stability, photosynthetic protection, and water oxidation processes are all influenced by the presence of *PsbR* [36]. The process of photosynthesis consists of two distinct components, namely a reaction center and a peripheral antenna. The light-harvesting complex (LHC) primarily comprises antenna proteins, which play a crucial role in this process. The photosynthesis antenna protein pathway of C1 in EBF plants exhibited a significant enhancement, notably, accompanied by a noteworthy decrease in the expression of various genes encoding *Lhcb*(*1,2,5*, and *6*) and *Lhca4*. These genetic alterations were instrumental in bringing about the observed changes in EBF.

Photoperiod-induced flowering of plants requires an appropriate circadian rhythm. Under prolonged exposure to sunlight, the biological clock and receptors interact to induce the upregulation of CONSTANS protein (*CO*) in companion cells of leaf phloem. CO then activates downstream target genes FLOWERING LOCUS T (*FT*) and TWINSISTER OF FT (*TSF*). In the photoperiodic pathway, *FT* plays a crucial role in determining the flowering time of plants, and it is possible that "anthocyanin", a flowering-inducing substance, is an expression product of the *FT* gene. The process of photoperiod induction triggers the translocation of FT mRNA from the leaf to the apical tissue of the plant. Once translated into the FT protein, it forms an association with the transcription factor FD of the bZIP molecule, thereby facilitating the activation of SUPPRESSOR OF OVEREXPRESSION OF CO1 gene (*SOC1*), APETALA1 (*AP1*), and LEAFY (*LFY*). This activation triggers the production of flower meristem, ultimately promoting flowering [23,24]. The expression of LONGATED HYPOCOTYL (*LHY*), a transcription factor implicated in the regulation of the circadian clock, exhibits a rhythmic pattern following the circadian rhythm [37]. The up-regulation of *LHY* expression occurs in response to increased red-light signals received by phytochrome B (*PHYB*), thereby promoting the expression of CO and FT, ultimately resulting in the generation of EBF.

The presence of sucrose (Suc) and its hydrolysis products glucose (Glc) and fructose (Fru) is indispensable for cellular biosynthesis and signal transduction throughout the entire life cycle of a plant. Moreover, the stimulation of flowering in Arabidopsis is attributed to the increased expression of *LFY* induced by sucrose. This phenomenon underscores the significance of sucrose in regulating plant development [38]. The expression of *EGLC*, *CELB*, *AMY*, and *glgA* involved in D-Glucose synthesis was significantly down-regulated in EBF plants. Conversely, *GPI* and *HK* involved in D-Fructose-6P synthesis were significantly up-regulated. Additionally, there was a significant increase in the levels of D-Glucose-6P and α-D-Glucose-1P. Amylase is an enzyme that produces maltose by hydrolyzing amylopectin and amylose, and the synthesis of D-Glucose-6P was increased in the upregulation of *malQ* and *HK*. During the flowering transformation process, sucrose levels in the phloem and stem tips rapidly increase, stimulating EBF by enhancing LFY expression.

The regulation of plant growth and development is orchestrated by a repertoire of phytohormones, encompassing GA, IAA, JA, CTK, SA, and BR. We found seven DEGs linked to hormone signaling, such as *AUX1*, *AHK2*, *DELLA*, *PYL*, *SnRK2*, *BSK*, and *CYCD3*. It is widely believed that GA promotes flowering by enhancing the transcriptional activity of LFY or activating SOC1 expression through GAMYB [39,40]. During flower development, GA can override the inhibitory effects of DELLA (key proteins in the GA signaling pathway) proteins, particularly RGA and RGL2, thereby promoting the transcription of B and C functional genes. Additionally, GA facilitates seed germination, enhances stem growth, and regulates flowering time [41]. We found that DELLA proteins were up-regulated in EBF plants. DELLA, as a negative regulatory factor, may serve as a response regulation in plants to inhibit GA, or it positively regulates EBF in plants through interactions with other plant hormones as an integrated protein, this may be an interesting point for further research.

In the epidermis, wax compounds are synthesized directly from very-long-chain fatty acids (VLCFAs). Pollen-stigma interactions have been reported to involve them in cellular communications, which are essential for plant development [42]. In Arabidopsis shoot apex cells, VLCFAs appear to regulate cell proliferation [25]. The comparison of four transcriptome data sets revealed the presence of four commonly observed DEGs, primarily involved in unsaturated fatty acid biosynthesis and fatty acid metabolism, including acetyl-CoA carboxylase (*ACACA*), enoyl-[acyl-carrier protein] reductase (*FabV*), I (*FabI*), and II (*FabK*). The reduction in VLCFAs content can enhance cellular proliferation and induce EBF by accelerating the transition from vegetative to reproductive growth phases [10].

The occurrence of EBF in *A. sinensis* is the result of a complex interplay among various factors. The study revealed an up-regulation of positive regulators associated with photosynthesis, including the photosynthesis-antenna proteins pathway, circadian rhythm pathway, starch and sucrose metabolism pathway, plant hormone signal transduction pathway, and very-long-chain fatty acid biosynthesis pathway. Conversely, genes that inhibit EBF were found to be down-regulated. The present study proposes a regulatory pathway for the control of EBF in *A. sinensis* (Fig. 14). A crucial element of photoperiodism involves the interaction between photoreceptors and circadian rhythm, which triggers *CO* gene expression and positively regulates *FT*. The hormone signaling pathway encompasses genes involved in biosynthesis and signaling, which activate essential genes encoding MADS-box proteins, including *SOC1*. The *SOC1* protein functions as a pivotal regulator that initiates the onset of floral development. MADS-box proteins are believed to integrate signals from photoperiodism and hormone signaling pathways, thereby regulating the expression of *LFY*, which serves as a link between floral induction and floral development. Different gene expressions of sucrose and ATP synthase pathways provide energy to EBF. VLCFAs mainly affect cell proliferation on the stem to induce EBF. We integrated the changed genes in five pathways related to the EBF of *A. sinensis*. Compared with previous studies [4–7,10,11], we found that *FT*, *SOC1*, *LFY*, and Sucrose were the key factors and common characteristics affecting EBF. And we compared two cultivars with different EBF rates and found that the genes influencing the consistent EBF rate may be *CDF1*, *HK*, *treX*, *bglX*, *AUX/IAA*, *PR-1*, *ACSL*, and *CO*. These genes were different from previous studies and may also be noteworthy genes for solving the problem of EBF in *A. sinensis*, which needed further validation. After conducting a comprehensive analysis of the transcriptome and metabolome of two *A. sinensis* varieties, we observed that the expression levels of genes associated with EBF were comparatively lower in Mingui 4 when compared to Mingui 2. Consequently, it is imperative to

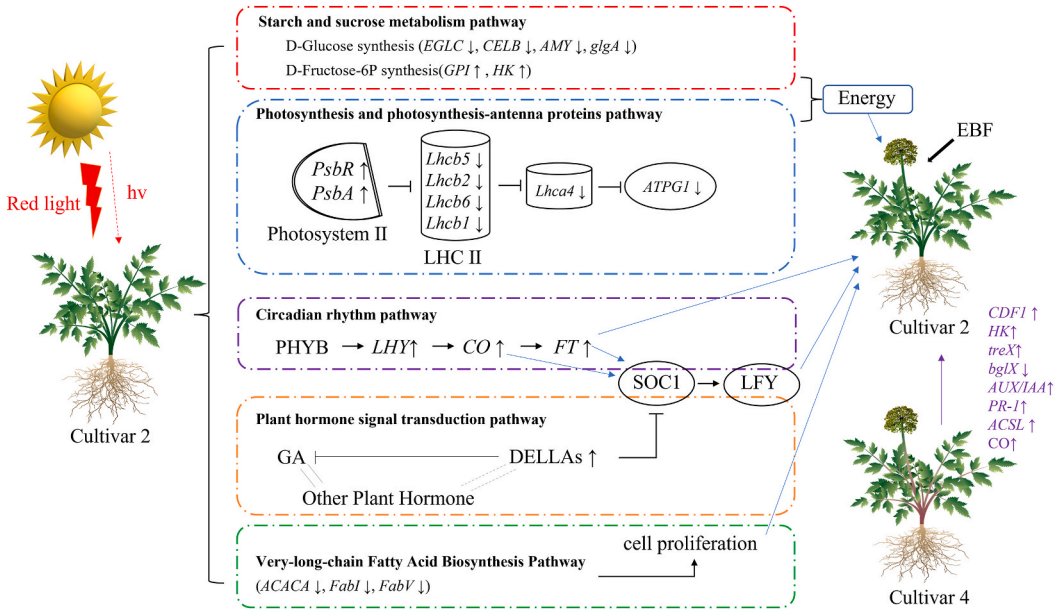


Fig. 14. Simplified representation of the regulatory networks proposed for EBF genes in *A. sinensis*. The arrows represent a promotion, while T-ends signify inhibition, and two horizontal dashed lines indicate an indirect interaction. The purple color signifies a potential genetic factor influencing the variation in EBF rates between the two cultivars. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

emphasize on enhancing the promotion and application of Mingui 4 in practical production.

5. Conclusions

The EBF and normal plants of each cultivar exhibited DEGs and DAMs, which played a pivotal role in the development of EBF. Transcriptomic and metabolomic analysis revealed five pathways associated with EBF of *A. sinensis*, namely the photosynthesis and photosynthesis-antenna proteins pathway, circadian rhythm pathway, starch and sucrose metabolism pathway, plant hormone signal transduction pathway, and very-long-chain fatty acid biosynthesis pathway. The key genes associated with EBF in *A. sinensis* were identified and genetically mapped. The utilization of gene-editing techniques, such as CRISPR/Cas 9, is warranted for the purpose of validating the regulatory role of these genes in EBF. The obtained results offer novel insights into the mechanisms underlying EBF in A. sinensis and other medicinal plants belonging to the Apiaceae family.

Data availability statement

Data associated with this study has been deposited at the sequence read archive (SRA) repository (https://dataview.ncbi.nlm.nih.gov/object/PRJNA962311). And the data (assembled transcriptome, corresponding functional annotation, predicted project and coding sequences) is available at the Figshare database (https://doi.org/10.6084/m9.figshare. 24960972. v1, 24960981, v1, 24961002. v1, and 24960996. v1).

CRediT authorship contribution statement

Chenghao Zhu: Writing – original draft, Software, Investigation, Formal analysis, Data curation. **Yu Bai:** Supervision, Software, Formal analysis. **Yuan Jiang:** Software, Resources. **Yuanfan Zhang:** Validation, Resources. **Shangtao Wang:** Data curation. **Fusheng Wang:** Validation, Resources, Funding acquisition. **Zhirong Sun:** Writing – review & editing, Supervision, Methodology, Conceptualization.

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 supported by the China Agriculture Research System of MOF and MARA (CARS-21) and Evaluation of Gansu Characteristic Traditional Chinese Medicine Germplasm and Integrated Construction and Promotion of High-Quality Seed and Seedling Production Technology System (22ZD6FA021-1).

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.heliyon.2024.e28636.

References

[1] I.L. Hook, Danggui to Angelica sinensis root: are potential benefits to European women lost in translation? A review, J. Ethnopharmacol. 152 (2014) 1–13.
[2] W.L. Wei, R. Zeng, C.M. Gu, Y. Qu, L.F. Huang, Angelica sinensis in China-A review of botanical profile, ethnopharmacology, phytochemistry and chemical analysis, J. Ethnopharmacol. 190 (2016) 116–141.
[3] H.Y. Zhang, W.G. Bi, Y. Yu, W.B. Liao, Angelica sinensis (Oliv.) Diels in China: distribution, cultivation, utilization and variation, Genet. Resour. Crop Evol. 59 (2012) 607–613.
[4] M. Li, J. Li, J. Wei, P.W. Pare, Transcriptional controls for early bolting and flowering in Angelica sinensis, Plants 10 (2021).
[5] M. Li, X. Cui, L. Jin, M. Li, J. Wei, Bolting reduces ferulic acid and flavonoid biosynthesis and induces root lignification in Angelica sinensis, Plant Physiol. Biochem. 170 (2022) 171–179.
[6] M.M. Luo, X.X. Liu, H.Y. Su, M.L. Li, M.F. Li, J.H. Wei, Regulatory networks of flowering genes in Angelica sinensis during vernalization, Plants-Basel 11 (2022).
[7] J. Li, M.L. Li, T.T. Zhu, X.N. Zhang, M.F. Li, J.H. Wei, Integrated transcriptomics and metabolites at different growth stages reveals the regulation mechanism of bolting and flowering of Angelica sinensis, Plant Biol. 23 (2021) 574–582.
[8] F.X. Guo, W.J. Xiao, Y. Chen, Y.J. Zhang, Y.Z. Chen, L.L. Liu, X. Gao, Initiation of early bolting by pre-enhancing anthocyanin and catalase activity in Angelica sinensis tender leaf during medicine formation cultivation year, Russ J Plant Physl+ 68 (2021) 763–773.
[9] J.M. Chen, W.M. Feng, H. Yan, P. Liu, G.S. Zhou, S. Guo, G. Yu, J.A. Duan, Explore the interaction between root metabolism and rhizosphere microbiota during the growth of Angelica sinensis, Front. Plant Sci. 13 (2022).
[10] X. Gao, F. Guo, Y. Chen, G. Bai, Y. Liu, J. Jin, Q. Wang, Full-length transcriptome analysis provides new insights into the early bolting occurrence in medicinal Angelica sinensis, Sci. Rep. 11 (2021) 13000.
[11] G. Yu, Y.X. Ma, J.A. Duan, B.S. Song, Z.Q. He, Identification of differentially expressed genes involved in early bolting of Angelica sinensis (Apiaceae), Genet. Mol. Res. 11 (2012) 494–502.

[12] L. Li, F. Wang, R. Yang, S. Wang, X. Pan, T. Du, L. Liu, W. Wang, Effects of varieties and seedling sizes on medicinal quality and economic benefit of Angelica sinensis during medicinal value formative period, Journal of Cold-Arid Agricultural Sciences 2 (2023) 239–245.
[13] Y. Bai, H. Liu, J. Pan, S. Zhang, Y. Guo, Y. Xian, Z. Sun, Z. Zhang, Transcriptomics and metabolomics changes triggered by inflorescence removal in panax notoginseng (burk.), Front. Plant Sci. 12 (2021) 761821.
[14] B. Buchfink, C. Xie, D.H. Huson, Fast and sensitive protein alignment using DIAMOND, Nat. Methods 12 (2015) 59–60.
[15] J. Mistry, R.D. Finn, S.R. Eddy, A. Bateman, M. Punta, Challenges in homology search: HMMER3 and convergent evolution of coiled-coil regions, Nucleic Acids Res. 41 (2013) e121.
[16] C. Trapnell, B.A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M.J. van Baren, S.L. Salzberg, B.J. Wold, L. Pachter, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation, Nat. Biotechnol. 28 (2010) 511–515.
[17] Y. Zhang, L. Yang, J. Yang, H. Hu, G. Wei, J. Cui, J. Xu, Transcriptome and metabolome analyses reveal differences in terpenoid and flavonoid biosynthesis in cryptomeria fortunei needles across different seasons, Front. Plant Sci. 13 (2022) 862746.
[18] W. Chen, L. Gong, Z. Guo, W. Wang, H. Zhang, X. Liu, S. Yu, L. Xiong, J. Luo, 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 (2013) 1769–1780.
[19] G. Qin, C. Liu, J. Li, Y. Qi, Z. Gao, X. Zhang, X. Yi, H. Pan, R. Ming, Y. Xu, Diversity of metabolite accumulation patterns in inner and outer seed coats of pomegranate: exploring their relationship with genetic mechanisms of seed coat development, Hortic. Res. 7 (2020) 10.
[20] T. Peng, Y. Wang, T. Yang, F. Wang, J. Luo, Y. Zhang, Physiological and biochemical responses, and comparative transcriptome profiling of two Angelica sinensis cultivars under enhanced ultraviolet-B radiation, Front. Plant Sci. 12 (2021) 805407.
[21] K. Morris, S.A. MacKerness, T. Page, C.F. John, A.M. Murphy, J.P. Carr, V. Buchanan-Wollaston, Salicylic acid has a role in regulating gene expression during leaf senescence, Plant J. 23 (2000) 677–685.
[22] M. Takano, N. Inagaki, X. Xie, N. Yuzurihara, F. Hihara, T. Ishizuka, M. Yano, M. Nishimura, A. Miyao, H. Hirochika, T. Shinomura, Distinct and cooperative functions of phytochromes A, B, and C in the control of deetiolation and flowering in rice, Plant Cell 17 (2005) 3311–3325.
[23] H.L. An, C. Roussot, P. Suarez-Lopez, L. Corbesler, C. Vincent, M. Pineiro, S. Hepworth, A. Mouradov, S. Justin, C. Turnbull, G. Coupland, CONSTANS acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of Arabidopsis, Development 131 (2004) 3615–3626.
[24] L. Corbesier, C. Vincent, S.H. Jang, F. Fornara, Q.Z. Fan, I. Searle, A. Giakountis, S. Farrona, L. Gissot, C. Turnbull, G. Coupland, FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis, Science 316 (2007) 1030–1033.
[25] T. Nobusawa, Y. Okushima, N. Nagata, M. Kojima, H. Sakakibara, M. Umeda, Synthesis of very-long-chain fatty acids in the epidermis controls plant organ growth by restricting cell proliferation, PLoS Biol. 11 (2013) e1001531.
[26] Y. Wang, X. Huang, X. Huang, W. Su, Y. Hao, H. Liu, R. Chen, S. Song, BcSOC1 promotes bolting and stem elongation in flowering Chinese cabbage, Int. J. Mol. Sci. 23 (2022).
[27] D.Y. Hyun, O.T. Kim, K.H. Bang, Y.C. Kim, N.H. Yoo, C.W. Kim, J.H. Lee, Genetic and molecular studies for regulation of bolting time of onion (Allium cepa L.), J. Plant Biol. 52 (2009) 602–608.
[28] L. Liu, X.J. Wang, S.T. Chen, D. Liu, C. Song, S.Y. Yi, F.C. Zhu, W. Wang, F. Wang, G.L. Wang, X.W. Song, B. Jia, C.W. Chen, H.S. Peng, L.P. Guo, B.X. Han, Fungal isolates influence the quality of Peucedanum praeruptorum Dunn, Front. Plant Sci. 13 (2022).
[29] J.T. Azietaku, H.F. Ma, X.A. Yu, J. Li, M.B. Oppong, J. Cao, M.R. An, Y.X. Chang, A review of the ethnopharmacology, phytochemistry and pharmacology of Notopterygium incisum, J. Ethnopharmacol. 202 (2017) 241–255.
[30] Y. Zhao, Q. Zhang, J. Li, X. Yan, H. He, X. Gao, G. Jia, High temperature in the root zone repressed flowering in Lilium × formolongi by disturbing the photoperiodic pathway and reconfiguring hormones and primary metabolism, Environ. Exp. Bot. 192 (2021).
[31] L. Chen, M. Xu, C. Liu, J. Hao, S. Fan, Y. Han, LsMYB15 regulates bolting in leaf lettuce (Lactuca sativa L.) under high-temperature stress, Front. Plant Sci. 13 (2022) 921021.
[32] B.-J. Li, B.-Q. Zheng, J.-Y. Wang, W.-C. Tsai, H.-C. Lu, L.-H. Zou, X. Wan, D.-Y. Zhang, H.-J. Qiao, Z.-J. Liu, Y. Wang, New insight into the molecular mechanism of colour differentiation among floral segments in orchids, Commun. Biol. 3 (2020) 89.
[33] Zhu T., Zhang M., Su H., Li M., Wang Y., Jin L., Li M., Integrated metabolomic and transcriptomic analysis reveals differential mechanism of flavonoid biosynthesis in two cultivars of Angelica sinensis, Molecules 27(2022)306.
[34] J. Putterill, R. Laurie, R. Macknight, It's time to flower: the genetic control of flowering time, Bioessays 26 (2004) 363–373.
[35] P.K. Boss, R.M. Bastow, J.S. Mylne, C. Dean, Multiple pathways in the decision to flower: enabling, promoting, and resetting, Plant Cell 16 (2004) S18–S31.
[36] J. De Las Rivas, P. Heredia, A. Roman, Oxygen-evolving extrinsic proteins (PsbO,P,Q,R): bioinformatic and functional analysis, Bba-Bioenergetics 1767 (2007) 575–582.
[37] S.X. Lu, S.M. Knowles, C. Andronis, M.S. Ong, E.M. Tobin, CIRCADIAN CLOCK ASSOCIATED1 and LATE ELONGATED HYPOCOTYL function synergistically in the circadian clock of Arabidopsis, Plant Physiol. 150 (2009) 834–843.
[38] S. Smeekens, J. Ma, J. Hanson, F. Rolland, Sugar signals and molecular networks controlling plant growth, Curr. Opin. Plant Biol. 13 (2010) 274–279.
[39] G.F. Gocal, C.C. Sheldon, F. Gubler, T. Moritz, D.J. Bagnall, C.P. MacMillan, S.F. Li, R.W. Parish, E.S. Dennis, D. Weigel, R.W. King, GAMYB-like genes, flowering, and gibberellin signaling in Arabidopsis, Plant Physiol. 127 (2001) 1682–1693.
[40] M.A. Blazquez, R. Green, O. Nilsson, M.R. Sussman, D. Weigel, Gibberellins promote flowering of arabidopsis by activating the LEAFY promoter, Plant Cell 10 (1998) 791–800.
[41] P. Hedden, V. Sponsel, A century of gibberellin research, J. Plant Growth Regul. 34 (2015) 740–760.
[42] A. Wang, Q. Xia, W. Xie, T. Dumonceaux, J. Zou, R. Datla, G. Selvaraj, Male gametophyte development in bread wheat (Triticum aestivum L.): molecular, cellular, and biochemical analyses of a sporophytic contribution to pollen wall ontogeny, Plant J. 30 (2002) 613–623.