Integrated Analysis of Transcriptome and Metabolome in *Cirsium japonicum* Fisch ex DC

Yun Ji Park, Seung-A Baek, Jae Kwang Kim, and Sang Un Park*

**ABSTRACT:** *Cirsium japonicum* Fisch ex DC belongs to the Compositae family and has been used as a folk remedy source in Asian countries because of its health-promoting properties. It is known that *C. japonicum* contains flavonoids, furans, long-chain alcohols, sterols, and volatile oils. Nevertheless, the molecular mechanism of secondary metabolite biosynthesis remains poorly understood. Therefore, transcriptome analysis and metabolic profiling were performed using different parts of *C. japonicum* to investigate phenylpropanoid metabolism. Based on the BLASTX search results, we identified 29 orthologs of enzymes responsible for phenylpropanoid biosynthesis. Additionally, 75 metabolites were identified in *C. japonicum*. Most of the flavonoid biosynthetic genes were significantly expressed ranging from 2.6- to 500-fold higher in the flowers than those in the leaves. Correspondently, the total content of flavonols was 21-fold higher in the flowers than in the roots. However, the total level of flavones showed 58-fold higher amounts in the leaves than in the flowers. Additionally, the total content of flavonols was 19-fold higher in the leaves than in the roots. The results of this study provide transcriptomic and metabolic information to elucidate the tissue-specific phenylpropanoid metabolism of *C. japonicum*.

**INTRODUCTION**

*Cirsium japonicum* Fisch ex DC, belonging to the Compositae family, is a perennial plant abundantly found in Northeast Asia. More than 200 *Cirsium* species have already been identified worldwide. This plant is used as a folk remedy in Asian countries and as a diuretic, antiphlogistic, hemostatic, and detoxifying agent in Oriental medicine. The dried root of *Cirsium* species, which is one of the botanical origins of the crude drug "Wazokudan", is used for the treatment of neuralgia and rheumatism in Japan. The essential oil components of the rhizomes of *C. japonicum* are also used as flavor ingredients. Also, some *Cirsium* species have been consumed as functional food. *C. japonicum* possesses several phytochemical constituents, such as flavonoids, furans, long-chain alcohol, sterols, and volatile oils. Moreover, many pharmacological effects including antitumor, anti-diabetic, antioxidant, anti-inflammatory, and anti-fungal properties of this plant have been investigated. However, the molecular mechanism of the secondary metabolite biosynthesis remains poorly understood because of the shortage of genomic or transcriptomic resources for *C. japonicum*. Transcriptome analysis is an effective tool to identify potential candidate genes or gene families encoding enzymes related to various metabolic pathways. It leads to a deeper investigation of biosynthesis, chemical diversity, and regulation of secondary metabolites in plant species. Next-generation sequencing (NGS) technology has emerged as a cost-effective and rapid sequencing approach for high-throughput sequence determination and is extensively used. This technology has advanced transcriptomics, particularly with the emergence of RNA-sequencing (RNA-Seq), and allowed integrated analysis of RNA expression in non-model plant species in which no molecular genetic research has been previously conducted. Whole-transcriptome expression by RNA-Seq permits the comprehensive transcriptomic and metabolic analysis in plant species. Because of advances in NGS technology, comprehensive genomic and transcriptomic information has allowed the exploration of various biological, cellular, and molecular processes, especially in non-model and non-sequenced organisms. In recent years, non-model plants, such as *Daucus*, *Cicer arietinum*, *Hevea brasiliensis*, and *Sesamum indicum* have been sequenced.

Phenylpropanoids are widely distributed secondary metabolites derived from primary metabolites, phenylalanine, or tyrosine through successive enzymatic reactions in plants. Phenylpropanoid can be sorted into 10 different classes, including coumarin, flavonoids, monolignols, phenolic acids, etc. This study provides transcriptomic and metabolic information to elucidate the tissue-specific phenylpropanoid metabolism of *C. japonicum*.
and stilbenes, according to their basic chemical structures. These compounds play crucial roles in plant development by acting as important components of supportive or protective tissues, defense strategies, and signaling. Additionally, several studies have investigated the biological activities of these compounds that are beneficial to human bodies. To date, many enzymes involved in the phenylpropanoid biosynthesis, such as oxygenases, ligases, oxidoreductases, and transferases, have been discovered. The starting steps of the phenylpropanoid biosynthetic pathway, catalyzed by phenylalanine ammonia lyase (PAL), cinnamate 4-hydroxylase (C4H), and 4-coumaroyl CoA ligase (4CL), offer a crucial branch point, which leads to various resulting metabolites. The flavonoid, monolignol, phenolic acid, and stilbene biosynthetic pathways have been extensively identified; however, the pathways of species- or lineage-specific compounds are not fully elucidated.

In the present study, we employed the Illumina/ Solexa HiSeq2000 platform to establish a transcriptome database and performed de novo assembly of the C. japonicum transcriptome and functional annotation. Based on the database, the genes encoding enzymes involved in phenylpropanoid biosynthesis were identified. Moreover, metabolites responsible for phenylpropanoid biosynthesis and polar metabolites in different organs were identified. We conducted correlation studies of gene expression and metabolite quantities across selected organs of C. japonicum to investigate the tissue-specific phenylpropanoid metabolism. Therefore, our findings will enhance the understanding of phenylpropanoid biosynthesis in C. japonicum. To the best of our knowledge, this study is the first report of the transcriptome of C. japonicum Fisch ex DC and illustrates the distribution of metabolites, especially phenylpropanoids, in various organs of this plant. With this work, we provide novel insights into the medicinal plant by a system analysis of the transcriptome and metabolome. The results support our understanding of relation between transcript levels of phenylpropanoid biosynthetic genes and metabolite production. It also indicates a valuable reference for future study at developing an indigeneous plant, which would be beneficial for both the pharmacological and nutraceutical fields.

## RESULTS AND DISCUSSION

**Illumina Sequencing and De Novo Assembly.** We performed RNA-Seq for the whole-transcriptome sequencing of C. japonicum. In total, 64 Mb raw reads were obtained with an average length of 151 nucleotides per read (Table 1). After eliminating the low-quality reads, adapter sequences, and ambiguous reads, 149,313 clean reads with a total length of 105 Mb bases were obtained. Subsequently, the contigs were integrated and assembled, using the Trinity program, into 104,890 unigenes (67 Mb bases) with an average length of 643 bases. All unigenes ranged from 224 to 15,640 bases with 50% of the total sequence assembled into contigs of 888 bases or larger. The length distribution of the assembled transcripts in the C. japonicum transcriptome is shown in Figure S1. In total, 17,673 unigenes had a length longer than 1000 bp. Of these, 8671 unigenes had a length ranging from 1000 to 1500 bp. All unigene data and their information, annotated by BLAST, are available in Supporting Information 2 and 3, respectively.

**Functional Annotation.** The contigs were integrated and assembled into 104,890 unigenes. After the assembly, all the unigenes were annotated via a BLAST search in the public databases, such as non-redundant (NR), nucleotide (NT), Swiss-Prot, Brassica database (BRAD), the Arabidopsis Information Resource (TAIR), Clusters of Orthologous Group (COG), and Gene Ontology (GO) (Table 2). The results revealed that 46,211 unigenes (44.06%) were significantly matched in the NR databases and 46,112 (43.96%) in the BRAD database.

### Table 2. Functional Annotations of C. japonicum Unigenes against the Public Database

| Annotations                          | Number of BLASTed Unigenes | Ratio (%) |
|--------------------------------------|----------------------------|-----------|
| All unigenes                          | 104,890                    | 100       |
| Unigenes BLASTed against NR           | 46,211                     | 44.06     |
| Unigenes BLASTed against NT           | 21,294                     | 20.3      |
| Unigenes BLASTed against Swiss-Prot   | 37,418                     | 35.67     |
| Unigenes BLASTed against TAIR         | 46,112                     | 43.96     |
| Unigenes BLASTed against COG          | 13,888                     | 13.24     |
| Unigenes BLASTed against GO           | 36,716                     | 35        |
| Unigenes CDS predicted*               | 37,524                     | 35.77     |
| All annotated unigenes               | 67,847                     | 64.68     |

*CDSs were predicted by Augustus S/W.*

For all databases together, 67,847 unigenes (64.68%) were successfully annotated. The E-value distribution of the top hits in the NR database showed that 44.06% of the sequences (46,211) were mapped to the known genes in plants with the best hits (Figure 1a). For the sequence similarity distribution analysis, 44% of matches showed high similarity (60–80%) with sequences in the NR databases (Figure 1b). For species distribution, 13% of the distinct sequences have top matches with genes from Vitis vinifera followed by Sesamum indicum (7.2%), Coffea canephora (6.8%), Nicotiana sylvestris (4.7%), N. tomentosiformis (4.4%), Theobroma cacao (3.4%), Citrus sinensis (3.3%), and Nelumbo nucifera (3.0%) (Figure 1c). The five largest categories were “translation, ribosomal structure, and biogenesis” (1390); “post-translational modification, protein turnover, and chaperones” (1169); “carbohydrate transport and metabolism” (1054); “general functional prediction only” (829); and “signal transduction mechanisms” (757). The “nuclear structure” (0) and “RNA processing and modification” (9) were the smallest COG categories (Figure 2).

As shown in Figure 3, genes assigned to the “cell” (29,864), “cell part” (29,827), and “organelle” (20,468) constituted the largest proportion in the Cellular component category. Additionally, we noticed a significant percentage of genes from “binding” (23,385) and “catalytic” (19,601) within the Molecular function category. Furthermore, the “cellular process”
The Compositae or Asteraceae is one of the largest and most successful flowering plant families, composed of 24,000 species worldwide. Of these, there are economically valuable species that are cultivated by humans with various degrees of domestication. To date, several studies have investigated genomic resources for various species belonging to the Compositae family. In our study, we obtained 104,890 unigenes with an average length of 643 bp, ranging from 224 to 15,640 bp, from 64,741,290 raw reads using the Illumina/Solexa HiSeq 2000 platform. Similarly, 51,133 unigenes from 200 to 15,402 bp with an average length of 648 bp were observed from de novo transcriptome analysis of *C. japonicum* var. *spinossimum*. Additionally, functional annotation and classification using GO analysis produced similar results. However, the highest species distribution of *C. japonicum* Fisch ex DC was obtained with *V. vinifera*, whereas the highest percentage of *C. japonicum* var. *spinossimum* sequences were matched to *Cynara cardunculus* var. *scolymus*. Based on our results, 58,679 (55.94%) unigenes had no resemblance to genes from public sources other than *C. japonicum* Fisch ex DC. These results suggest that these numerous unigenes, specific to this plant, are functionally unknown, and hence further studies are warranted.

![Figure 1](https://dx.doi.org/10.1021/acsomega.0c04001)

**Figure 1.** Classification of NR annotation results of the *C. japonicum* unigenes.

![Figure 2](https://dx.doi.org/10.1021/acsomega.0c04001)

**Figure 2.** COG functional classification of the *C. japonicum* unigenes.
Gene Expression Analysis of Key Genes Involved in Phenylpropanoid Biosynthesis. Based on the BLASTX search results, we identified 29 orthologs of enzymes, which were responsible for phenylpropanoid biosynthesis (Table S1). The candidate genes, which are responsible for the phenylpropanoid pathway, were discovered from the transcriptome dataset, including CjPALs, CjC4Hs, Cj4ClS, CjCHSs, CjCHI, CjF3’H, CjF3H, CjFS II, CjFLS, CjDFRs, CjHCTs, CjC3H, CjHQTs, and CjCOMTs. Most major enzymes responsible for phenylpropanoid metabolism in C. japonicum were consistent with those of other plants in the same family.16
Table 3. Contents of Metabolites in the Different Organs from *C. japonicum*

| metabolites                                          | leaf                      | stem                      | flower                    | root                       |
|------------------------------------------------------|---------------------------|---------------------------|---------------------------|----------------------------|
| alanine                                              | 2.99 ± 0.19²              | 1.69 ± 0.22²              | 9.42 ± 0.89³              | 3.68 ± 0.13³               |
| asparagine                                           | 0.91 ± 0.22²              | 3.60 ± 1.11³              | 65.38 ± 6.94⁴             | 544.26 ± 33.82⁴            |
| aspartic acid                                        | 3.48 ± 0.39³              | 5.26 ± 0.62³              | 13.03 ± 1.12³             | 34.90 ± 1.56⁴              |
| glutamic acid                                        | 12.48 ± 1.49³             | 7.07 ± 0.72³              | 28.81 ± 3.11³             | 21.17 ± 0.71³              |
| glutamine                                            | 0.54 ± 0.14²              | 1.37 ± 0.32³              | 29.39 ± 3.66³             | 42.46 ± 3.06³              |
| glycine                                              | 0.39 ± 0.04³              | 0.40 ± 0.03³              | 5.72 ± 0.57³              | 1.02 ± 0.09²               |
| isoleucine                                           | 0.21 ± 0.01³              | 0.36 ± 0.02³              | 8.47 ± 1.16³              | 0.90 ± 0.06³               |
| leucine                                              | n.d.¹                     | n.d.¹                     | 7.83 ± 0.87³              | n.d.¹                      |
| methionine                                           | n.d.¹                     | n.d.¹                     | 1.04 ± 0.09³              | n.d.¹                      |
| phytosterols (μg/g)                                  |                           |                           |                           |                            |
| α- tocopherol                                         | 102.05 ± 22.36⁴           | 12.20 ± 0.59³             | 37.30 ± 5.22³             | 8.14 ± 0.38³               |
| β- tocopherol                                         | 1.97 ± 0.08³              | 1.96 ± 0.06³              | 2.24 ± 0.15³              | n.d.¹                      |
| γ- tocopherol                                         | 43.52 ± 10.50⁴            | 4.61 ± 0.30³              | 10.15 ± 1.63³             | 1.63 ± 0.05³               |
| total                                                | 147.54 ± 32.93            | 18.77 ± 0.96              | 49.69 ± 7.01              | 9.78 ± 0.43                |
| sugars, Sugar Acids, and Sugar Alcohol (ratio/g)     |                           |                           |                           |                            |
| fructose                                             | 1415.61 ± 154.74⁴         | 2781.41 ± 151.54³         | 8237.03 ± 1161.37⁵        | 36.20 ± 12.8³              |
| galactose                                            | 1.92 ± 0.22³              | 1.97 ± 0.05³              | 5.97 ± 0.60³              | 0.73 ± 0.02³               |
| glucose                                              | 547.40 ± 53.07⁴           | 528.32 ± 24.54³           | 1482.29 ± 231.69⁴         | 231.8 ± 24.0³              |
| inositol                                             | 78.71 ± 7.83³             | 73.01 ± 5.89³             | 179.04 ± 18.71³           | 16.81 ± 0.36³              |
| mannitol                                             | 0.76 ± 0.10³              | n.d.¹                     | 0.64 ± 0.17³              | 1.05 ± 0.02³               |
| mannosio                                             | 0.85 ± 0.16³              | 1.17 ± 0.10³              | 2.65 ± 0.35³              | 0.73 ± 0.11³               |
| raffinose                                            | 7.82 ± 0.58³              | n.d.¹                     | n.d.¹                     | n.d.¹                      |
| sucrose                                              | 686.62 ± 53.64³           | 2443.64 ± 161.29³         | 5390.85 ± 615.14³         | 962.36 ± 14.40³            |
| xylose                                               | n.d.¹                     | n.d.¹                     | 1.65 ± 0.15³              | n.d.¹                      |
| glycereic acid                                       | 9.56 ± 0.97³              | 2.42 ± 0.13³              | 4.18 ± 0.40³              | 1.44 ± 0.03³               |
| threonic acid                                        | 2.23 ± 0.18³              | 1.81 ± 0.21³              | 44.29 ± 5.09³             | n.d.¹                      |
| glycerol                                             | 21.39 ± 1.78²             | 18.47 ± 0.70³             | 56.26 ± 6.55³             | 29.06 ± 1.40³              |
| phytosterols (μg/g)                                  |                           |                           |                           |                            |
| α- tocopherol                                         | 102.05 ± 22.36³           | 12.20 ± 0.59³             | 37.30 ± 5.22³             | 8.14 ± 0.38³               |
| β- tocopherol                                         | 1.97 ± 0.08³              | 1.96 ± 0.06³              | 2.24 ± 0.15³              | n.d.¹                      |
| γ- tocopherol                                         | 43.52 ± 10.50³            | 4.61 ± 0.30³              | 10.15 ± 1.63³             | 1.63 ± 0.05³               |
| total                                                | 147.54 ± 32.93            | 18.77 ± 0.96              | 49.69 ± 7.01              | 9.78 ± 0.43                |
To understand the transcript levels of the phenylpropanoid biosynthetic genes, quantitative real-time PCR (qRT-PCR) was estimated in different organs of *C. japonicum*. PAL, C4H, and 4CL genes responsible for the initial steps of the pathway are key enzymes that provide the basis for all subsequent branches.14 As shown in Figure S2, the expression levels of CjPALs varied depending on isoforms of the gene. Two isoforms of Cj4H revealed similar expression patterns, which were the highest in flowers, followed by stems, roots, and leaves. The highest expression level of Cj4CL1 occurred in the leaf. The expression of Cj4CL2 and Cj4CL3 were the highest in the stems. Most genes involved in flavonoid biosynthesis, including CjCHS, CjCHI, CjF3'H, CjFS2, and CjFLS, were strongly expressed in the flowers of *C. japonicum* (Figure S3). In particular, the transcript levels of CjCHS1 and CjCHS2 were 500- and 112-fold higher in the flowers than in the leaves, respectively. Furthermore, the expression of CjF3'H, CjCHI, CjFLS, and CjFSII in flowers was 16.6-, 13-, 5.6-, and 3.3-fold higher than that in the leaves, respectively. The transcript levels of CjFSII and CjFLS were rarely expressed in the roots. However, two isoforms of CjDFR showed 22- and 2.6-fold higher expression in the roots than in the leaves. The transcript levels of CjHCTs were predominantly expressed in the roots (Figure S4). The transcript level of CjC3H was the highest in the roots where it was 3.8-fold higher than in the leaves. The high-expressed levels of two CjHQTs were observed in the roots and flowers, respectively. Of CjCOMTs, most of these genes were significantly expressed in the flowers than in other organs. The expression level of CjCOMT6 was rarely expressed in the roots. Interestingly, the transcript level of CjCOMT6 was predominantly shown in the roots.

**Accumulation of Metabolites in Different Organs of C. japonicum**. The composition and content of polar metabolites, phenylpropanoid compounds, policosanols, tocopherols, phytosterols, and amyrins in different organs of *C. japonicum* were determined by gas chromatography–mass spectrometry (GC–MS), gas chromatography time-of-flight mass spectrometry (GC-TOFMS), and high-performance liquid chromatography (HPLC) analysis (Figures S5 and S6). In this study, we determined 75 primary and secondary metabolites, including 19 amino acids, 10 organic acids, 12 sugars, 9 policosanols, 3 tocopherols, 4 phytosterols, 2 amyrins, 4 flavones, 2 flavonols, 3 flavonols, 5 hydroxycinnamic acids, and 2 hydroxybenzoic acids.

Table 3. continued

| metabolites                  | leaf           | stem           | flower         | root            |
|------------------------------|----------------|----------------|----------------|-----------------|
| campestester                 | 377.59 ± 87.03| 139.76 ± 11.70| 158.75 ± 34.31| 239.50 ± 20.86  |
| cholesterol                  | 2.29 ± 0.34    | n.d.           | 2.18 ± 0.17    | n.d.            |
| β-sitosteryl                 | 211.30 ± 50.74 | 96.96 ± 6.56   | 300.00 ± 74.17| 132.76 ± 12.12  |
| stigmasteryl                 | 242.24 ± 52.88 | 100.49 ± 10.22 | 160.18 ± 30.29 | 205.79 ± 19.46  |
| total                        | 833.43 ± 190.99| 337.21 ± 28.48 | 621.11 ± 138.94| 578.05 ± 52.44  |
| amyrins (mg/g)               |                |                |                |                 |
| α-aminorufin                 | 2.24 ± 0.54    | 0.35 ± 0.02    | 1.30 ± 0.21    | 0.03 ± 0.00     |
| β-aminorufin                 | 0.74 ± 0.18    | 0.17 ± 0.01    | 0.90 ± 0.16    | 0.01 ± 0.00     |
| total                        | 2.98 ± 0.72    | 0.32 ± 0.02    | 2.21 ± 0.36    | 0.03 ± 0.00     |
| flavonoids (mg/g)            |                |                |                |                 |
| flavones                     |                |                |                |                 |
| 16.549 ± 0.002               | 1.765 ± 0.004  | 2.417 ± 0.008  | 0.016 ± 0.001  |                 |
| apigenin                     | 0.188 ± 0.001  | 0.154 ± 0.004  | 0.219 ± 0.008  | 0.326 ± 0.004   |
| hispidulin                   | 1.121 ± 0.003  | 0.005 ± 0.002  | 1.037 ± 0.046  | 0.040 ± 0.001   |
| cirsimarin                   | 23.943 ± 0.004 | 2.874 ± 0.004  | 1.537 ± 0.010  | 0.336 ± 0.002   |
| total                        | 41.802 ± 0.009 | 4.798 ± 0.014  | 5.208 ± 0.072  | 0.718 ± 0.008   |
| rutin                        | 1.92 ± 0.02    | 5.57 ± 1.16    | 8.19 ± 0.59    | n.d.            |
| epicatechin                  | 17.82 ± 2.38   | n.d.           | n.d.           | n.d.            |
| total                        | 19.74 ± 2.40   | 5.57 ± 1.16    | 8.19 ± 0.59    | n.d.            |
| flavonoids (mg/g)            |                |                |                |                 |
| total                        | 18.24 ± 3.41   | 39.21 ± 2.23   | 75.84 ± 14.64  | 360 ± 0.14      |
| hydroxycinnamic acids (mg/g)|                |                |                |                 |
| chlorogenic acid             | 83.81 ± 14.30  | 35.85 ± 7.21   | 41.17 ± 3.25   | 56.73 ± 11.61   |
| caffeic acid                 | 5.17 ± 0.51    | 86.78 ± 6.69   | 11.51 ± 0.85   | 313.04 ± 2.54   |
| p-coumaric acid              | 4.76 ± 0.38    | 0.50 ± 0.02    | 7.70 ± 1.02    | 2.07 ± 2.06     |
| ferulic acid                 | 8.16 ± 0.53    | 2.99 ± 0.01    | 3.45 ± 0.53    | 3.21 ± 0.24     |
| trans-cinnamic acid          | 1541.52 ± 32.12| 171.60 ± 2.28  | 157.61 ± 4.09  | 0.28 ± 0.03     |
| total                        | 1643.42 ± 47.84| 297.73 ± 10.21| 221.44 ± 6.75  | 375.33 ± 16.49  |
| hydroxybenzoic acids (mg/g)  |                |                |                |                 |
| 4-hydroxybenzoic acid        | 10.87 ± 1.54   | n.d.           | n.d.           | n.d.            |
| benzoic acid                 | 35.48 ± 4.48   | 0.23 ± 0.05    | 1.34 ± 0.68    | 2.48 ± 1.54     |
| total                        | 46.35 ± 6.02   | 0.23 ± 0.05    | 1.34 ± 0.68    | 2.48 ± 1.54     |

*Each value is the mean of three biological replicates ± standard deviation. Means followed by different letters are significantly different at P < 0.05 by Duncan’s multiple range tests. n.d, not detectable.*
Additionally, we identified their diverse content patterns among these compounds in different organs (Table 3). Based on the data, most of the amino acids and their derivatives showed higher levels in the flowers than in the other organs. Especially, the content of phenylalanine, the precursor of phenylpropanoid metabolism, was 14.8-fold higher in the flowers than in the leaves. Organic acid content was varied to different organs. Most of the sugars, sugar acids, and sugar alcohols also significantly accumulated in the flowers. However, the level of raffinose was only detected, and the level of glyceric acid was the highest in the leaves. The level of total policosanols was the highest in the leaves followed by the flowers, stems, and...
roots. Among policosanols, triacontanol was predominantly accumulated in most of the organs, except for roots. The content of total tocopherols indicated 15-fold higher amount in the leaves than in the roots. \(\alpha\)-Tocopherol was shown as a major tocopherol in all organs. The level of total phytosterols was the highest in the leaves followed by the flowers, roots, and stems. Similarly, the level of total amyrins was accumulated approximately 99-fold higher in the leaves than in the roots.

The leaf tissue contained the highest amounts of flavones, such as cirsimarin (16.55 mg/g dry weight (dw) \(\pm\) 0.00), hispidulin (1.12 mg/g dw \(\pm\) 0.00), and cirsimaritin (23.94 mg/g dw \(\pm\) 0.00). The contents of catechin hydrate and epicatechin were the highest in the flowers and leaves, respectively, while those were not detected in the roots. Among flavonols, rutin was present at similar levels in the leaves, stems, and flowers, whereas it was not detected in the roots. Quercetin and kaempferol were predominantly present in the flowers. The patterns of individual hydroxycinnamic acids were different in all organs. The chlorogenic acid, ferulic acid, and trans-cinnamic acid were highly accumulated in the leaves compared to other organs. However, the level of caffeic acid was 60-fold higher in the roots than in the leaves. \(p\)-Coumaric acid content was the highest in the flowers followed by leaves, roots, and stems. Two of the hydroxybenzoic acids were predominant in the leaves.

**Visualization of Metabolic Profiling Data Using Principal Component Analysis (PCA), Heat Map, and PathVisio.** PCA was performed to overview the multivariate data obtained from metabolic profiling. The score plot of PCA showed clear separation among the organs (Figure 4a). Flowers were separated from the other organs by the first component (principal component 1; PC1). The stems and roots were separated from the leaves by principal component 2 (PC2). The metabolites in the loading plot were related to the separation of organs on the score plot (Figure 4b). Amino acids and their

**Figure 5.** Heat map of the contents of metabolites in *C. japonicum*. The standardized data were represented as red and blue as a bar at the top right of the heat map. C20-ol, eicosanol; C21-ol, heneicosanol; C22-ol, docosanol; C23-ol, tricosanol; C24-ol, tetracosanol; C26-ol, hexacosanol; C27-ol, heptacosanol; C28-ol, octacosanol; and C30-ol, triacontanol
derivatives, except for asparagine and aspartic acid, had negative eigenvector values in PC1. This indicates that the flowers contained relatively higher levels. However, the organic acids, except for pyruvic acid and malic acid, had positive eigenvector values in PC1. This indicated that the organic acid levels were relatively higher in the leaves, stems, and roots than in the flowers. Due to the negative eigenvector values in PC2, the levels of citric acid, fumaric acid, malic acid, and succinic acid were higher in the stems and roots than in the leaves.

The heat map clearly showed the tissue-specific metabolites (Figure 5). The standardized metabolite content data were represented by red and blue colors as a bar at the top right of the heat map. Similar results between PCA and heat map were obtained. Additionally, the heat map showed the separation in metabolites between the stems and roots, which was not clearly shown in the PCA results.

PathVisio is a free-downloaded program and can aid in the visualization of the expression of genes, proteins, metabolites, and interactions between them based on biological pathways. To compare the tissue-specific metabolites on the pathway, the standardized data were visualized using PathVisio 3.3.0 (Figure 6). The red color indicates high levels of metabolites, and green indicates low levels according to the scale bar at the bottom right. Tocopherols were biosynthesized through the methylerythritol 4-phosphate (MEP) pathway, and phytosterols and amyrins were biosynthesized through the mevalonate (MVA) pathway. High accumulation of tocopherols, phytosterols, and amyrins indicated activation of the MEP and MVA pathways in the leaves compared to other organs. The levels of fumaric acid and malic acid, which are intermediates of the tricarboxylic acid (TCA) cycle, were higher in the stems. Flowers contained higher levels of sugars, pyruvic acid, and amino acids than the other organs. The flower secretes nectar to attract a pollinator through nectaries, which are specialized organs for nectar production, secretion, and release. The nectar consists mainly of sugars (sucrose, glucose, and fructose) and water. Glucose and fructose are produced by the hydrolysis of sucrose through various biochemical pathways before being secreted. Amino acids and lipids are also minor constituents of nectar. This indicated that sugar and amino acid biosynthetic metabolism was higher in the flowers than in the other organs. Glutamine is produced from ammonium, and then glutamine is used to make glutamic acid by glutamate synthetase in the roots. α-Ketoglutaric acid, one of the intermediates in the TCA cycle, is also used in the glutamine synthetase/glutamate synthase (GS/GOGAT) cycle. The high levels of glutamic acid, glutamine, succinic acid, and citric acid in the roots could explain the role of roots in the assimilation of nitrates from the soil. Because asparagine and aspartic acid are made from glutamic acid, and pyroglutamic acid is a cyclized derivative of glutamic acid, these metabolites were also at higher levels in the roots than in other organs. Additionally, glutamic acid is used to synthesize amino acids, including 4-aminobutyric acid and

Figure 6. Schematic metabolite pathway comparing primary and secondary metabolites of C. japonicum. The unit variance scaling data is represented by the intensity of the green or red color as the box at the bottom right. The gray boxes represent the metabolites that could not be detected. G6P, glucose 6-phosphate; F6P, fructose 6-phosphate; F1,6BP, fructose 1,6-bisphosphate; 3PG, 3-phosphoglycerate; PEP, 2-phosphoenolpyruvate; MEP, methylerythritol 4-phosphate; and MVA, mevalonate.
proline, indicating that the levels of glutamic acid and glutamine were also higher in the flowers than in other organs.

Polyphenol compounds, derived from the phenylpropanoid pathway, are a massive group of secondary metabolites in plants. The phenylpropanoid pathway serves as a rich source of metabolites in plants and an initial point for the production of various important compounds, including coumarin, flavonoids, and lignans. In the present study, putative unigenes corresponding to the phenylpropanoid biosynthetic pathway were isolated from the C. japonicum Fisch ex DC transcripts. The profile and concentration of the phenolic compounds may vary depending on different tissues or populations of the same plant species. The different composition of metabolites among organs relies on de novo biosynthesis and catabolism or can result from compounds being re-allocated. Several studies have investigated whether flavonoid accumulation depends on the expression patterns of PAL, C4H, and 4CL in many species. However, based on our results, most flavonoids were highly accumulated in the leaves, whereas different expression profiles of CjPALs, CjC4Hs, and Cj4CLs were observed according to the plant organ (Figure 7). The organ-specific accumulation of flavonols could be explained by the expression pattern of CjFLS, which was relatively high in the flowers and not detected in the roots in C. japonicum. The FLS gene is a key biosynthetic gene to produce flavonols, including quercetin and kaempferol. Furthermore, the highest expression level of CjC3H, which directs flux into caffeic acid, might affect its abundant biosynthesis in the roots of C. japonicum. The expression patterns of CjCHSs were significantly high in the flowers of C. japonicum. Previous studies have revealed that the CHS gene is spatially responsible for anthocyanin biosynthesis. Therefore, we suggest that the expression patterns of several phenylpropanoid biosynthetic genes correspond to organ-specific distribution in C. japonicum. However, the accumulation patterns of flavones were not correlated with the expression levels of CjF3’H and CjFSII, which were highly expressed in the flowers. Phenylpropanoid metabolism in plants is a complicated network involving numerous enzymes. Multiple isoforms and paralogous genes might have redundant functions, and these synergistic regulations might complicate the elucidation of phenylpropanoid metabolism.

According to HPLC analysis, we have identified that flavones in leaves were the main components in C. japonicum Fisch ex DC. In a previous study, flavone glycosides, linarin, cirsimarin, and hispidulin were isolated from the aerial part of C. ussuriensis. Moreover, their pharmacological activities, such as hepatopro-

Figure 7. Schematic phenylpropanoid biosynthetic pathway comparing gene expression and metabolites content of C. japonicum. The gene expression is represented by the intensity of the blue or red color. The unit variance scaling data is represented by the intensity of the blue or red (gene expression) and green or red (metabolites content) color as the box at the bottom right. The gray boxes represent the metabolites that could not be detected. 4CL1, 4-coumarate-CoA ligase-like 6; 4CL2, 4-coumaroyl CoA ligase; 4CL3, 4-coumarate-CoA ligase-like 5; C3H, p-coumarate 3-hydroxylase; C4H, cinnamate 4-hydroxylase; CHI, chalcone isomerase; CHS, chalcone synthase; COMT, caffeic acid 3-O-methyltransferase; DFR, dihydroflavanol 4-reductase; F3H, flavanone 3-hydroxylase; F3’H, flavonoid 3’-hydroxylase; FLS, flavonol synthase; FS II, flavone synthase II; HCT, shikimate O-hydroxycinnamoyl transferase; HQT, hydroxycinnamoyl CoA quinate transferase; and PAL, phenylalanine ammonia-lyase.

https://dx.doi.org/10.1021/acsomega.0c04001
ACS Omega 2020, 5, 29312−29324
ective, adenosine-antagonistic, and antilipogenic effects, have been investigated.\textsuperscript{29} Additionally, metabolic profiling has been performed using GC–MS analysis in this study. Policosanol is components of plant waxes, which are present in fruits, leaves, seeds, and surfaces.\textsuperscript{27} They have health-improving properties, including lipid-lowering effects, blood-pressure-lowering effects, anti-cancer effects, and liver protection.\textsuperscript{28} Tocopherols are a precursor of vitamin E and exhibit antioxidant activity. Vitamin E is an essential dietary component for human nutrition and can be elucidated. Overall, our results provide valuable information on phenylpropanoid and the gene-to-metabolite network in \textit{C. japonicum}. Furthermore, the natural distribution of primary and secondary metabolites within the different organs of \textit{C. japonicum} were discovered to high-throughput sequencing data, 29 orthologous genes and secondary metabolites within the different organs of \textit{C. japonicum} were elucidated for the first time in our study.\textsuperscript{32–35} This result also provides basic information for the utilization of \textit{C. japonicum} as a pharmaceutical supplement.

In the present study, a transcriptome analysis of \textit{C. japonicum} using the Illumina/Solexa HiSeq2000 platform obtained 104,890 unigenes with an average length of 643 bp. According to high-throughput sequencing data, 29 orthologous genes correlated with phenylpropanoid biosynthesis were discovered in \textit{C. japonicum}. Furthermore, the natural distribution of primary and secondary metabolites within the different organs of \textit{C. japonicum} was elucidated. Overall, our results provide valuable information on phenylpropanoid and the gene-to-metabolite network in \textit{C. japonicum}.

**METHODS**

**Plant Materials.** \textit{C. japonicum} Fisch ex DC seeds were collected from the National Institute of Biological Resources. \textit{C. japonicum} plants were grown in vitro under light/dark (16/8 h) conditions at 25 °C. Whole plantlets were harvested after 3 months for sequencing. Additionally, \textit{C. japonicum} plants were cultured under greenhouse conditions and exposed to outdoor conditions for 2 years from April 2014. Temperature conditions and illumination were not controlled in the cultures. The different organs including leaves, stems, flowers, and roots were harvested from three individual plants on May 2016. The plant materials were washed thoroughly, immediately frozen in liquid nitrogen, and stored at −80 °C for further use. All experiments for downstream analyses were carried out using biological triplicate.

**RNA Isolation and Sequence Analysis.** Total RNA was obtained from \textit{C. japonicum} using the Total RNA mini kit (Geneaid, Taiwan) and Trizol (Invitrogen, USA) according to the manufacturer’s protocol with some modifications. The concentration of RNA and its integrity was assessed using a Nanodrop spectrophotometer (GE Healthcare Bio-Science Crop, USA) and 1.0% denaturing agarose gel. The \textit{C. japonicum} mRNA was separated and purified using Sera-Mag Magnetic Oligo (dT) beads to obtain transcriptome data. Paired-end sequencing with a read length of 151 bp was performed on an Illumina/Solexa HiSeq2000 platform (NICEM, Seoul, Korea). Subsequently, raw-read-quality evaluation using FastQC (http://www.bioinformatics.babraham.ac.uk/project/fastqc) with default options and adapter sequence trimming using CutAdapt (http://pypi.python.org/pypi/cutadapt) were conducted to obtain high-quality adapter-free reads. Finally, in silico normalization was conducted to collect high-quality reads from \textit{C. japonicum}.

**Transcriptome Assembly and Functional Annotation.** For de novo assembly, the resulting clean reads were assembled using Trinity package (http://trinityrnaseq.github.io) using a k-mer length of 25, minimum contig length of 100 bp, and average fragment length of 300 bp.\textsuperscript{36} Assembled contigs were clustered, and representative sequences for each cluster were chosen as unigenes using CD-Hit-EST (http://weizhogli-lab.org/cd-hit).\textsuperscript{37} Distinct sequences were mapped against various protein and nucleotide databases, such as NR, NT, BRAD, TAIR, COG, and GO using NCBI BLAST (http://blast.ncbi.nlm.nih.gov/Blast.cgi) to identify the functional annotation of unigenes, and an E-value cutoff of 1 × 10\textsuperscript{−5} was set. GO annotation for each unigene was transferred from BLASTX hits to the UniProt protein database. Then, the gene ontology functional classification in the GO hierarchy level 2 was obtained using WeGo (http://wego.genomics.org.cn).\textsuperscript{38}

**cDNA Synthesis and qRT-PCR.** Total RNA was reverse transcribed using a ReverTra Ace-α kit from Toyobo. Primer sets for qRT-PCR were designed by GenScript Real-time PCR (TaqMan) Primer Design (www.genscript.com) (Table S2). The qRT-PCR reactions were assigned on a BIO-RAD CFX (Agilent, CA, USA) equipped with LECO Pegasus HT (Leco Corp., MI, USA) using 2X Real-Time PCR smart mix (Solgent Co., Ltd. Korea) according to the manufacturer’s instructions. The quantification method of qRT-PCR was followed by a previous study.\textsuperscript{39} The transcript levels were normalized relative to the housekeeping gene elongation factor-1 alpha (CjEF1α). All reactions were conducted with triplicate.

**GC–TOFMS Analysis.** Samples were dried in a freeze dryer at −80 °C for 3 days and ground into a fine powder. Ten milligrams of fine powder was added to 1 mL of extraction solvent (methanol/water/chloroform, 2.5:1:1, v/v) with 60 μL of ribitol (200 μg/mL) as the internal standard (IS) and incubated at 37 °C for 30 min at 20 Hz. The crude mixture was centrifuged at 166.6 Hz for 3 min. An 800 μL sample of the upper layer was mixed with 400 μL of water and incubated at 166.6 Hz for 3 min. In total, 900 μL of the upper layer was concentrated using a centrifugal evaporator for 3 h. The sample was freeze dried for 16 h to dry completely. An 80 μL sample of 2% (w/v) methoxyamine hydrochloride in pyridine was added and incubated at 30 °C for 90 min at 20 Hz. An 80 μL sample of M-nethyl-N-(trimethylsilyl) trifluoroacetamide (MSTFA) was added and incubated again at 37 °C for 30 min at 20 Hz. A 1 μL sample of the derivatized substance was injected with a split 25:1 ratio into a CP-SIL 8 CB-MS column (30 μm × 0.25 mm, 0.25 μm thickness) in the Agilent 6890N network GC system (Agilent, CA, USA) equipped with LECO Pegasus HT (LECO Corp., MI, USA). The injection, transfer line, and ion source temperature were set at 230, 250, and 200 °C, respectively. Helium was used as the carrier gas, and the gas flow rate was set at 1.0 mL/min. The column temperature program was as follows: hold 80 °C for 2 min, ramp 15 °C/min to 320 °C, and hold 320 °C for 10 min. The detector voltage was set to 1700 volts, and a mass range of 85–600 was used. Prior to quantitative analysis, we used the ChromaTOF software to support peak findings and automated deconvolution of reference mass spectra. The polar metabolites were identified using an in-house library.\textsuperscript{40} Relative quantification was calculated from peak area ratios relative to the peak area of the IS.

**GC–MS Analysis.** The method used for extraction and analysis has been previously described.\textsuperscript{41} Samples (50 mg) were
mixed with 3 mL of ethanol, added with 0.1% (w/v) ascorbic acid and 0.5 μg of 5α-cholestanol (IS), then incubated at 85 °C for 5 min. Subsequently, 120 μL of 80% (w/v) potassium hydroxide was added, and the mixture was further incubated at 85 °C for 10 min for saponification. The cooled mixture was mixed with 1.5 mL of water and 1.5 mL of hexane. After centrifugation at 20 Hz for 5 min, the supernatant was transferred to a new tube, and 1.5 mL of hexane was added to the mixture again. After being vortexed and centrifuged, the supernatant was collected and dried completely using nitrogen gas. For derivatization, 30 μL of MSTFA and 30 μL of pyridine were added and incubated at 60 °C for 30 min at 20 Hz. Policosanols, tocopherols, phytosterols, and amyrins were separated using an Rtx-5MS column (30 cm × 0.25 mm, 0.25 μm thickness) equipped with GC−MS (GCMS-QP2010 Ultra system, Shimadzu, Kyoto, Japan). A 1.0 μL sample was injected with a split 10:1 ratio at 290 °C injection temperature. Helium gas was used as the carrier gas with a flow rate of 1.0 mL/min. The column temperature program was as follows: hold at 150 °C for 5 min, ramp at 15 °C/min to 320 °C, and hold at 320 °C for 10 min. The interface and ion source temperatures were set at 280 and 230 °C, respectively. Content of policosanols, tocopherols, phytosterols, and amyrins in C. japonicum was identified based on their retention times, correlations of mass spectral information, and peak names and calculated using calibration curves according to the peak area ratios observed from the IS.

**HPLC Analysis.** For flavones, the samples (20 mg) were covered with 1 mL of 70% (v/v) acetonitrile, sonicated for 45 min, then centrifuged at 200 Hz for 5 min. Prior to analysis, all samples were filtered through a PVDF membrane filter. Samples were detected using an Agilent 1260 Infinity II Quat pump (CA, USA) with a DAD WR detector (CA, USA). Chromatographic separation was made on an INNO C18 column (25 cm × 4.6 mm, 5 μm) operated at an injection volume of 10 μL, UV length at 270 nm, column temperature set to 25 °C, and a flow rate of 1.0 mL/min. The binary mobile phase comprised (A) water/acetic acid (95.5:0.5, v/v) and (B) acetonitrile. All reference compounds were purchased from the Natural Product Institute of Science and Technology (NIST, Anseong, Korea).

Approximately 100 mg of samples was used for the determination of other phenolic compounds. The samples were extracted in 3 mL of 80% (v/v) methanol at 25 °C for 1 h. Homogenates were centrifuged at 200 Hz for 10 min. The supernatant was used to measure the other phenolic compounds. Flavonoid and phenolic acid contents were estimated as described in a previous study.42

**Statistical Analysis.** Data were subjected to a one-way ANOVA to determine the statistical significance of the detected differences using SPSS version 24. Duncan’s multiple range test was carried out. P < 0.05. The data obtained from the metabolic profiling were normalized with unit variance (UV) scaling. SIMCA-P (version 14.1; Umetrics, Umeå, Sweden) was used for PCA analysis. MetaboAnalyst 4.0 (https://www.metaboanalyst.ca/) was used for Pearson’s correlation analysis and heat maps. PathVisio 3.0.0 (https://pathvisio.github.io) was used for visualization of metabolic pathways based on the Arabidopsis metabolic pathway downloaded from WikiPathways [AtMetExpress overview (Arabidopsis thaliana)].

### ASSOCIATED CONTENT

* Supporting Information

The Supporting Information is available free of charge at https://pubs.acs.org/doi/10.1021/acsomega.0c04001.

Comparison of phenylpropanoid biosynthetic genes for C. japonicum with the most orthologous genes; qRT-PCR primers used in this study; length distribution of contigs and unigenes in C. japonicum transcriptome; qRT-PCR analysis of upstream genes involved in phenylpropanoid biosynthesis; qRT-PCR analysis of flavonoid biosynthetic genes; qRT-PCR analysis of phenolic acid biosynthetic genes; GC-TOFMS chromatogram of polar metabolites in the leaf and flower of C. japonicum; and GC-MS chromatogram of policosanols, tocopherols, phytosterols, and amyrins in the leaf of C. japonicum (PDF).

Unigene data annotated by BLAST (TXT). Information about all the unigene data (XLSX)

### AUTHOR INFORMATION

* Corresponding Authors

Jae Kwang Kim — Division of Life Sciences, College of Life Sciences and Bioengineering, Incheon National University, Incheon 22012, Republic of Korea; ∗ orcid.org/0000-0003-2692-5370; Phone: +82-32-835-8241; Email: ljkkpj@inu.ac.kr.; Fax: +82-32-835-0763

Sang Un Park — Department of Crop Science and Department of Smart Agriculture Systems, Chungnam National University, Daejeon 34134, Republic of Korea; orcid.org/0000-0003-2157-2246; Phone: +82-42-821-5730; Email: supark@cnu.ac.kr.; Fax: +82-42-822-2631

* Authors

Yun Ji Park — Department of Crop Science, Chungnam National University, Daejeon 34134, Republic of Korea

Seung-A Baek — Division of Life Sciences, College of Life Sciences and Bioengineering, Incheon National University, Incheon 22012, Republic of Korea

Complete contact information is available at: https://pubs.acs.org/10.1021/acsomega.0c04001

* Author Contributions

Y.J.P. and S.-A.B. contributed equally to this work.

* Notes

The authors declare no competing financial interest.

### ACKNOWLEDGMENTS

This research was supported by the Bio & Medical Technology Development Program of the National Research Foundation (NRF) funded by the Ministry of Science, ICT & Future Planning (2016M3A9A5919548).

### REFERENCES

1. Lee, W. B.; Kwon, H. C.; Cho, O. R.; Lee, K. C.; Choi, S. U.; Baek, N. I.; Lee, K. R. Phytochemical constituents of Cirsium setidens Nakai and their cytotoxicity against human cancer cell lines. Arch. Pharmacal Res. 2002, 25, 628−635.

2. Kim, S.-J.; Kim, G.-H. Identification for flavones in different parts of Cirsium japonicum. J Food Sci Nutr 2003, 8, 330−335.

3. Miyazawa, M.; Yamafuji, C.; Kurose, K.; Ishikawa, Y. Volatile components of the rhizomes of Cirsium japonicum DC. Flavour Frag J 2003, 18, 15−17.
Lee, O.-H.; Kim, J. H.; Kim, Y. H.; Lee, Y. J.; Lee, J. S.; Jo, J. H.; Kim, B. G.; Lim, J. K.; Lee, B.-Y. Nutritional components and physiological activities of Cirsium setidens Nakai. *Korean J. Food & Nutr.* 2014, 43, 791−798.

Zhang, X.; Liao, M.; Cheng, X.; Liang, C.; Diao, X.; Zhang, L. Ultra-high-performance liquid chromatography coupled with triple quadrupole and time-of-flight mass spectrometry for the screening and identification of the main flavonoids and their metabolites in rats after oral administration of Cirsium japonicum DC. extract. *Rapid Commun. Mass Spectrom.* 2018, 32, 1451−1461.

Reddy, N. R. R.; Mehta, R. H.; Soni, P. H.; Makasana, J.; Gajbhiye, N. A.; Ponnuchamy, M.; Kumar, J. Next generation sequencing and transcriptome analysis predicts biosynthetic pathway of sennosides from Senna (Cassia angustifolia Vahl), a non-model plant with potent laxative properties. *PLoS One* 2015, 10, e0119242.

Tang, Q.; Ma, X.; Mo, C.; Wilson, I. W.; Song, C.; Zhao, H.; Yang, Y.; Fu, W.; Qiu, D. An efficient approach to finding Sirtuins in the transcriptome of Cirsium setidens Nakai. *ACS Omega* http://pubs.acs.org/journal/acsodf

Barker, M. S.; Matvienko, M.; Yu, Q.; Turner, K. G.; Pearl, S. A.; Bell, G. Triterpene biosynthetic genes by RNA-seq and digital gene expression analysis. *Rapid Commun. Mass Spectrom.* 2018, 32, 1451−1461.

Vogt, T. Phenylpropanoid biosynthesis. *Mol. Plant* 2010, 3, 2−20.

Lai, Z.; Kane, N. C.; Kozik, A.; Hodgins, K. A.; Dlugosch, K. M.; Barker, M. S.; Matvienko, M.; Yu, Q.; Turner, K. G.; Pearl, S. A.; Bell, G. D. Genomics of Compositae weeds: EST libraries, microarrays, and evidence of intronogenesis. *Am. J. Bot.* 2012, 99, 209−218.

De Paolis, A.; Pignone, D.; Morgese, A.; Sonnante, G. Characterization and differential expression analysis of artichoke phenylalanine-ammonia-lyase-coding sequences. *Characterization and differential expression analysis of artichoke phenylalanine-ammonia-lyase-coding sequences. Phytochem. 2008, 78, 33−43.*

Kutmon, M.; van Iersel, M. P.; Bohler, A.; Kelder, T.; Nunes, N.; Pico, A. R.; Evelo, C. T. PathVisio 3: an extendable pathway analysis toolbox. *PLoS Comput. Biol.* 2015, 11, e1004085.

Somme, L.; Vanderplanck, M.; Michex, D.; Lombaerde, I.; Moerman, R.; Watthelet, B.; Wattiez, R.; Lognay, G.; Jacquemart, A.-L. Pollen and nectar quality drive the major and minor floral choices of bumble bees. *Apidologie* 2015, 46, 92−106.

Brandenburg, A.; Dell’Olivo, A.; Bshary, R.; Kuhlemeier, C. The sweetest thing: advances in nectar research. *Curr. Opin. Plant Biol.* 2009, 12, 486−490.

Forde, B. G.; Lea, P. J. Glutamate in plants: metabolism, regulation, and signalling. *J. Exp. Bot.* 2007, 58, 2339−2358.

Fraser, C. M.; Chapple, C. The phenylpropanoid pathway in Arabidopsis. *Arabidopsis Book/ASPB* 2011, 9, No. e01522.

Francisco, M.; Alli, M.; Ferreres, F.; Moreno, D. A.; Velasco, P.; Soengas, P. Organ-specific quantitative genetics and candidate genes of phenylpropanoid metabolism in Brassica oleracea. *Front. Plant Sci.* 2016, 6, 1240.

Kim, Y. J.; Kim, Y. B.; Li, X.; Choi, S. R.; Park, S.; Park, J. S.; Lim, Y. P.; Park, S. U. Accumulation of phenylpropanoids by white, blue, and red light irradiation and their organ-specific distribution in Chinese cabbage (*Brassica rapa* ssp. *pekinesis*). *J. Agric. Food Chem.* 2015, 63, 6772−6778.

Wellmann, F.; Lukačin, R.; Moriguichi, T.; Britsch, L.; Schulte, E.; Matern, U. Functional expression and mutational analysis of flavonol synthase from *Citrus unshiu*. *Eur. J. Biochem.* 2002, 269, 4134−4142.

Zhou, L.; Wang, Y.; Peng, Z. Molecular characterization and expression analysis of chalcone synthase gene during flower development in *Fluonicia transversa* (Fagopyrum tataricum Gaertn.). *Afr. J. Biotech.* 2011, 10, 1275−1284.

Han, H.-S.; Shin, J.-S.; Lee, S.-B.; Park, J. C.; Lee, K.-T. *Cirsimarin*, a flavone glucoside from the aerial part of *Cirsium japonicum* var. *ussurienne* (Regel) Ktmt. exp Ohwi, suppresses the JAK/STAT and IRF-3 signaling pathway in LPS-stimulated RAW 264.7 macrophages. 2018, 293, 38−47, DOI: 10.1016/j.cbio.2018.07.024.

Irmak, S.; Dunford, N. T.; Milligan, J. Policosanol contents of beeswax, sugar cane and wheat extracts. *Food Chem.* 2006, 95, 312−315.

Shen, J.; Luo, F.; Lin, Q. Policosanol: Extraction and biological functions. *J. Funct. Foods* 2019, 57, 351−360.

Döremann, P. Functional diversity of tocochromanols in plants. *Planta* 2007, 225, 269−276.

Gabay, O.; Sanchez, C.; Salvat, C.; Chevy, F.; Breton, M.; Nourissat, G.; Wolf, C.; Jacques, C.; Berenbaum, F. Stigmastanol: a phytosterol with potential anti-osteoarthritic properties. *Osteoarthr. Cartil.* 2010, 18, 106−116.

Chen, D.; Xu, F.; Zhang, P.; Deng, J.; Sun, H.; Wen, X.; Liu, J. Practical Synthesis of α-Amyrin, β-Amyrin, and Lupeol: The Potential Natural Inhibitors of Human Oidosesuqale Cyclase. *Arch. Pharm. (Weinheim)* 2017, 350, 1700178.

Hasanlou, T.; Bahmani, M.; Sepelhrifar, R.; Kalantari, F. Determination of tocopherols and fatty acids in seeds of *Silybum marianum* (L.) Gaerth. *J. Medicinal Plants 2008, 7, 69−76.*

Matthaus, B.; Ozcak, M. M.; Al-Juhaimi, F. Fatty Acid, Tocopherol, and Mineral Contents of Onopordum acanthum seed and oil. *Chem. Nat. Compd.* 2014, 50, 1092−1093.

Harrabi, S.; Curtis, S.; Hayet, F.; Mayer, P. M. Changes in the sterol compositions of milk thistle oil (Silybum marianum L.) during seed maturation. *Grasas Aceites 2016, 67, 123.*

Harrabi, S.; Ferchichi, A.; Bacheli, A.; Fella, H. Policosanol composition, antioxidant and anti-arthritic activities of milk thistle (Silybum marianum L.) oil at different seed maturity stages. *Lipids Health Dis.* 2018, 17, 82.

Grabberr, M. G.; Haas, B. J.; Yassour, M.; Levin, J. Z.; Thompson, D. A.; Amit, I.; Adiconis, X.; Fan, L.; Raychowdhury, R.; Zeng, Q. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 2011, 29, 654.

Fu, L.; Niu, B.; Zhou, Z.; Wu, S.; Li, W. CD-HIT: accelerated for clustering the next-generation sequencing data. *Bioinformatics* 2012, 28, 3150−3152.

Ye, J.; Fang, L.; Zheng, H.; Zhang, Y.; Chen, J.; Zhang, Z.; Wang, J.; Li, S.; Li, R.; Bolund, L.; Wang, J. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res.* 2006, 34, W293−W297.

Whelan, J. A.; Russell, N. B.; Whelan, M. A. A method for the absolute quantification of cDNA using real-time PCR. *J. Immunol. Methods* 2003, 278, 261−269.

Kim, J. K.; Choi, S. R.; Lee, J.; Park, S.-Y.; Song, S. Y.; Na, J.; Kim, S. W.; Kim, S.-J.; Nou, I.-S.; Lee, Y. H.; Park, S. U.; Kim, H. R. Metabolic differentiation of diamondback moth (*Plutella xylostella* (L.)) resistance in cabbage (*Brassica oleracea* L. ssp. capitata). *J. Agric. Food Chem.* 2013, 61, 11222−11230.

Kim, T. J.; Lee, K. B.; Baek, S.-A.; Choi, J.; Ha, S.-H.; Lim, S.-H.; Park, S.-Y.; Yeo, Y.; Park, S. U.; Kim, J. K. Determination of lipophilic metabolites for species discrimination and quality assessment of nine leafy vegetables. *J Korean Soc Appl Biol Chem 2015, 58, 909−918.*

Kim, Y. K.; Li, X.; Xu, H.; Park, N. I.; Uddin, M. R.; Pyon, J. Y.; Park, S. U. Production of phenolic compounds in hairy root culture of *tartary buckwheat* (*Fagopyrum tataricum* Gaertn.). *J. Crop Sci. Biotech.* 2009, 12, 53−57.