Comparative analysis of the transcriptomes and primary metabolite profiles of adventitious roots of five *Panax ginseng* cultivars

Yun Sun Lee 1,*, Hyun-Seung Park 1,*, Dong-Kyu Lee 2,*, Murukarthick Jayakodi 1, Nam-Hoon Kim 1, Sang-Choon Lee 1, Atreyee Kundu 1,3, Dong-Yup Lee 4,5, Young Chang Kim 6, Jun Gyo In 7, Sung Won Kwon 2,*, Tae-Jin Yang 1,8,*

1 Department of Plant Science, Plant Genomics and Breeding Institute, Research Institute of Agriculture and Life Sciences, College of Agriculture and Life Sciences, Seoul National University, Seoul, Korea
2 College of Pharmacy and Research Institute of Pharmaceutical Sciences, Seoul National University, Seoul, Korea
3 Department of Biological Sciences, Presidency University, Kolkata, West Bengal, India
4 Bioprocessing Technology Institute, A*STAR (Agency for Science, Technology and Research), Centros, Singapore
5 Department of Chemical and Biomolecular Engineering, Synthetic Biology Research Consortium, National University of Singapore, Singapore
6 Ginseng Research Division, National Institute of Horticultural and Herbal Science, Rural Development Administration, Eumseong, Korea
7 Ginseng Resources Research Laboratory, R&D Headquarters, Korea Ginseng Corporation, Daejeon, Korea
8 Crop Biotechnology Institute/GreenBio Science and Technology, Seoul National University, Pyeongchang, Korea

A R T I C L E   I N F O

Article history:
Received 26 June 2015
Received in Revised form 15 December 2015
Accepted 31 December 2015
Available online 9 January 2016

Keywords:
adventitious root
metabolome
*Panax ginseng*
primary metabolic pathways
transcriptome

A B S T R A C T

Background: Various *Panax ginseng* cultivars exhibit a range of diversity for morphological and physiological traits. However, there are few studies on diversity of metabolic profiles and genetic background to understand the complex metabolic pathway in ginseng.

Methods: To understand the complex metabolic pathway and related genes in ginseng, we tried to conduct integrated analysis of primary metabolite profiles and related gene expression using five ginseng cultivars showing different morphology. We investigated primary metabolite profiles via gas chromatography–mass spectrometry (GC-MS) and analyzed transcriptomes by Illumina sequencing using adventitious roots grown under the same conditions to elucidate the differences in metabolism underlying such genetic diversity.

Results: GC-MS analysis revealed that primary metabolite profiling allowed us to classify the five cultivars into three independent groups and the grouping was also explained by eight major primary metabolites as biomarkers. We selected three cultivars (Chumpoon, Cheongsun, and Sunhyang) to represent each group and analyzed their transcriptomes. We inspected 100 unigenes involved in seven primary metabolite biosynthesis pathways and found that 21 unigenes encoding 15 enzymes were differentially expressed among the three cultivars. Integrated analysis of transcriptomes and metabolomes revealed that the ginseng cultivars differ in primary metabolites as well as in the putative genes involved in the complex process of primary metabolic pathways.

Conclusion: Our data derived from this integrated analysis provide insights into the underlying complexity of genes and metabolites that co-regulate flux through these pathways in ginseng.

Copyright © 2016, The Korean Society of Ginseng, Published by Elsevier. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

*Panax ginseng* Meyer, is a traditional herb grown in Korea and China that is widely used as a source of herbal medicines due to its tonic, stimulant, restorative, and anti-aging properties [1]. Recent pharmacological studies have contributed to our understanding of the medicinal efficacy of ginseng, revealing its pain-relieving, anti-carcinogen, anti-diabetes, antihypertension, and immunization-
enhancing effects [2]. P. ginseng (commonly known as Korean ginseng) is widely cultivated and represents one of most valuable medicinal crops in Korea. Korean ginseng includes a number of cultivars derived from inbred lines developed by pedigree selection from three local landraces, Jalgyung, Chungkyung, and Hwangsook [3].

Different ginseng cultivars have different phenotypes [4], genetic polymorphisms [3,5,6], and ginsenoside contents [7–9]. Many previous studies have revealed genetic and metabolic diversities based on DNA polymorphisms and metabolite profiles. Genetic diversity among cultivars has been revealed by simple sequence repeat [3], single nucleotide polymorphism [5], and sequence characterized amplified region marker [6] analyses based on DNA polymorphisms. The metabolic diversity of ginseng cultivars has been investigated via HPLC [7], Fourier transform infrared [8], and nuclear magnetic resonance [9] analyses. However, to date, only a few studies have been performed on the diversity of primary metabolites in ginseng cultivars [9].

Recently, integrated transcriptome and metabolome analysis has been introduced as a new approach to systematically elucidate the mechanisms that regulate metabolic networks in plants [10–12]. First, genes regulating metabolic pathways and relevant metabolic correlations are extensively analyzed by high-throughput sequencing and non-targeted metabolome analysis, respectively. The resulting profiles are then comparatively analyzed to elucidate gene–metabolite correlations and to identify the functions of unknown genes. The coordinated regulation of genes and metabolites involved in phenylpropanoid metabolism has been explored using this approach in Medicago truncatula [13], Populus x canescens [14], and Arabidopsis thaliana [15]. Moreover, to analyze alkaloid and triterpene metabolism, comparative analyses of transcriptomes and metabolomes have been carried out in Papaver somniferum [16], M. truncatula [17], and Catharanthus roseus [18]. A similar approach has also been applied to Solanum tuberosum using gas chromatography–mass spectrometry (GC-MS); a major platform for primary metabolite analysis [19].

In this study, we performed integrated analysis of the transcriptomes and metabolomes via GC-MS and Illumina platform next-generation sequencing technology in adventitious roots from several ginseng cultivars. We compared the data derived from the same tissues to obtain an integrated understanding of primary metabolic pathways and related genes. The results of this study increase our understanding of the complexity of metabolite biosynthesis and will enhance efforts to identify functional genes in each step of primary metabolite biosynthesis in ginseng.

2. Materials and methods

2.1. Chemicals and reagents

Growth hormones and media required for adventitious root culture were purchased from Duchefa (Haarlem, The Netherlands). All authentic compounds, alkaline mixtures, and solvents for GC-MS analysis were obtained from Sigma–Aldrich (St. Louis, MO, USA) and J.T. Baker (Phillipsburg, NJ, USA). All reference chemicals were dissolved in 100% methanol.

2.2. Plant materials

Adventitious roots from five P. ginseng cultivars [Chunpoong (CP), Cheongsun (CS), Sunhyang (SH), Goppong (GO), and Sunun (SU)] were used in this study. Adventitious roots were induced and cultivated in bioreactors as previously described [20]. Adventitious roots maintained in bioreactors were transferred into 250-mL flasks containing 100 mL Schenk and Hildebrandt liquid medium [21] supplemented with 3 mg/L indole-3-butyric acid (IBA) and 5% sucrose and cultured at 25°C on a rotary shaker under dark conditions. After 1 month, the adventitious roots were collected and used for GC-MS and transcriptome analysis.

2.3. Sample preparation for GC-MS analysis

Analysis of primary metabolites using GC-MS was carried out as follows. Freeze-dried adventitious roots from five ginseng cultivars were pulverized and extracted with 100% methanol under sonication for 30 min at room temperature (10 mg/mL methanol). The extracts were centrifuged at 13,500 g for 5 min at room temperature, filtered through a 0.5-µm polytetrafluoroethylene syringe filter (Toyo Roshi Kaisha, Tokyo, Japan), and mixed with 40 µL 1,000 ppm C17:0 fatty acid as an internal standard. The extracts were completely dried under N2 purge. The residues were resuspended in 100 µL 20 mg/mL methoxyamine hydrochloride in pyridine at 37°C for 90 min and then trimethylsilylated with 100 µL N,O-bis (trimethylsilyl) trifluoroacetamide at 37°C for 30 min. The prepared solutions were transferred to vials, which were capped immediately. All samples collected under the same conditions were analyzed with five individual replicates.

2.4. GC-MS conditions

GC-MS analysis was conducted on a GCMS 2010 plus (Shimadzu, Tokyo, Japan). A DB-5 capillary column (30 m × 0.25 mm, 0.25 µm thickness; Agilent, Santa Clara, CA, USA) was used with helium at a constant flow rate of 1 mL/min. The oven temperature conditions were as follows: 100°C for 2 min, ramping up to 300°C at a rate of 5°C/min, and holding at 300°C for 10 min. The sample (1 µL) was injected in split mode (50: 1). The ionization energy was 70 eV in electron impact mode. The transfer line and ion source temperatures were set at 300°C and 250°C, respectively. Mass spectra were obtained at 3.06 scans/s with a mass range of 50–500 m/z.

2.5. GC-MS data processing and statistical analysis

All raw data analyzed by GC-MS were converted into cdf format and exported to MZmine software version 2.10 (http://mzmine.sourceforge.net/). The raw data were deconvoluted using a Savitzky–Golay filter and aligned through the RANdom SAmple Consensus algorithm. The aligned data were exported to a CSV-formatted bucket table, followed by normalization based on the area of the internal standard. Statistical analysis of the data was carried out using SIMCA-P+ version 12.0 (Umetrics, Umeå, Sweden) to compare metabolite variations among the five ginseng cultivars. The measured abundances of all variables were scaled to unit variance (UV scaling), to calculated base weight as 1/standard deviation, for a selection of all samples. The resulting data were subjected to supervised partial least square discriminant analysis (PLS-DA), and a permutation test was carried out 200 times to validate the PLS-DA model. Metabolites that were significantly different (p < 0.05) and had high variable importance for projection values (VIP value > 1.0) were regarded as biomarkers for determining metabolic differences among ginseng cultivars.

2.6. RNA isolation and Illumina sequencing

Total RNA from the adventitious roots of two ginseng cultivars (CS and SH) was isolated using a Plant RNAeasy Mini Kit (Qiagen, Hilden, Germany). The concentration and quality of the extracted RNA were checked by ND-1000 (NanoDrop Technologies Inc.,
aggravous roots of CS and SH were used to construct cDNA libraries. Paired-end sequencing of fragments 150 bp in size was performed using the Illumina sequencing platform (NextSeq 500; Lab Genomics Co., Pankyo, Korea). The raw RNA-seq data from CS and SH were deposited in the NCBI Sequence Read Archive (SRA, http://www.ncbi.nlm.nih.gov/Traces/sra) under accession numbers SRR1688723 and SRR1688724, respectively. The raw RNA-seq data from CP were obtained from our previous report [20] (Accession number: SRR619718). Adapter sequences and low-quality bases were filtered using NGS QC Toolkit [22]. Adventitious roots of CS and SH collected under the same conditions were analyzed with three individual replicates.

2.7. Identification of candidate genes involved in primary metabolism and digital expression analysis

The sequences of genes involved in primary metabolite biosynthesis pathways in ginseng were retrieved from the CP root transcriptome through comparisons with homologous sequences in the NCBI and TAIR databases using TBLASTN [23]. Genes with e values < 10^{-5} and coverage above 80 were selected as candidate genes. The expression levels of the candidate genes in adventitious roots of CP, CS and SH were determined using fragments per kilobase of exon per million fragments (FPKM) values calculated from RNA-seq data via Expectation Maximization [24]. Analysis of variance (p < 0.05) and false discovery rate (q < 0.05) tests were performed using the R program (version 3.1.0, www.R-project.org) to select genes exhibiting significantly different expression patterns among the three cultivars. Mean values of FPKM for each gene were transformed to log₂-fold changes and visualized using the R (version 3.1.0) heatmap package.

3. Results and discussion

3.1. Primary metabolite profiles of the adventitious roots of five ginseng cultivars

We performed primary metabolite profiling using adventitious roots derived from five officially registered inbred ginseng cultivars that were selected among 10 cultivars based on normal adventitious root growth in our preliminary survey (Fig. 1). Previous studies demonstrated that ginseng cultivars have different phenotypes, ginsenoside types, and contents [7]. However, there have been no in-depth studies on the variation in primary metabolism responsible for the diversities among the cultivars. Therefore, we aimed to investigate metabolic diversities among the cultivars and to identify the genes corresponding with the identified metabolic differences using integrative analysis of transcriptome and metabolome, focused on primary metabolism. Peaks obtained by GC-MS were subjected to multivariate analysis to examine the differences in metabolite composition among ginseng cultivars (Fig. 2). Based on PLS-DA analysis, which was validated by 200 random permutation tests (R² = 0.727 and Q² = −0.589), the five ginseng cultivars were separated into three groups: Group 1 consisted of CS; Group 2 consisted of GO, SH and SU; and Group 3 consisted of CP (Fig. 3). These results indicate that the metabolite compositions of GO, SH, and SU are similar, while those of CS and CP are distinct. SH, SU, and GO were bred from Jakyung, while CS and CP were bred from Chungkyung, through pure-line selection [3]. This breeding pedigree reflects the grouping results of our PLS-DA analysis, and the phylogenetic relationships of nine Korea ginseng cultivars revealed by analysis with Simple sequence repeats (SSR)-based markers [3] is also consistent with the clustering based on PLS-DA analysis. We identified some metabolites from the GC-MS analysis by comparing their MS spectra with those available in the National Institute of Standards and Technology (NIST) library (www.nist.gov) or with authentic reference standards. A total of 20 compounds were identified in the five ginseng cultivars (Table 1). Among these, eight compounds were selected as biomarker compounds, as they enabled discrimination among each of the five ginseng cultivars; these markers include L-alanine, malonate, L-valine, phosphoric acid, propanoic acid, L-serine, DL-malate, and L-glutamine.

Metabolite profiling through GC-MS analysis has been widely used as strategy for detecting polar (organic acids, amino acids, sugars, and sugar alcohols) and nonpolar (fatty acids and sterols) compounds representing primary metabolites [25]. GC-MS analysis, with the high accuracy of MS and excellent reproducibility of retention time, has been used in many studies to discriminate among species [26,27] or cultivars [28,29]. GC-MS analysis has also been used to detect differences in volatile compounds among P. ginseng, Panax notoginseng and Panax quinquefolius [30]. Although metabolite accumulation patterns are diverse among accessions in same species [28,29], no previous report has described the intra-species diversity of metabolites within P. ginseng. Our data demonstrate that there is much diversity in the contents and composition of primary metabolites in P. ginseng (Table 2). The levels of compounds in the amino acid and carboxylic acid biosynthesis pathways were higher in CS and SH than in CP. The levels of phosphoric acid (in the inorganic acid category) as well as glucose and fructose (monosaccharides) were lower in CS and SH than in CP.

Fig. 1. Adventitious roots from five Panax ginseng cultivars maintained in bioreactors (CS, Cheogsun; SU, Sunun; SH, Sunhyang; CP, Chunpoong; GO, Gopoong). Adventitious roots from each cultivar were cultured in Schenk and Hildebrandt medium including 5 mg/L IBA for 1 year with subculturing every 4 weeks. The adventitious roots, which were maintained under dark conditions, were transferred to flasks and collected after 1 month for GC-MS and transcriptome analysis. IBA, indole-3-butyric acid; GC-MS, gas chromatography–mass spectrometry.
3.2. Transcriptome analysis and isolation of candidate genes involved in primary metabolite biosynthesis pathways

Ginsenosides, which are representative secondary metabolites in *P. ginseng*, are biosynthesized via the triterpenoid biosynthesis pathway. Acetyl coenzyme A (acetyl-CoA), which is one of the precursors for entry into the triterpenoid biosynthesis pathway, undergoes multiple chemical changes, resulting in the production of diverse ginsenosides. Previous studies have investigated some of the putative genes associated with the ginsenoside biosynthesis pathway via transcriptome analysis [20,31] and have characterized their functions [32–34]. However, genes involved in primary metabolite biosynthesis, which is an upstream reaction of ginsenoside biosynthesis, have not been reported in ginseng.

Our PLS-DA analysis demonstrated differential accumulation of primary metabolites among ginseng cultivars (Fig. 3). Seven pathways were identified in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database (http://www.genome.jp/kegg/), including the glycolysis pathway (ko00010), valine, leucine, and isoleucine biosynthesis pathway (ko00290), glycine, serine, and threonine biosynthesis pathway (ko00260), citrate cycle [tricarboxylic acid (TCA) cycle] (ko00020), fructose and mannose biosynthesis pathway (ko00051), and arginine and proline pathway (ko00330) (Table 3). These pathways are responsible for the biosynthesis of primary metabolites. We retrieved information about 30 genes involved in the seven primary metabolite biosynthesis pathways in other species from public databases (NCBI and TAIR database) and identified 100 unigenes in the ginseng unigene sets that were assembled from 12 root transcriptomes of the CP cultivar (Table 3) [23].

To identify genes involved in the seven primary metabolite biosynthesis pathways, we selected three representative cultivars from each group (Group 1, CS; Group 2, SH; and Group 3, CP) and performed transcriptome analysis using the Illumina platform. In
3.3. Integrated analysis of gene expression and metabolome data

Integrated analysis of transcriptome and metabolome data is a useful technique for discovering genes that are potentially involved in metabolic networks. We therefore examined the correlation of metabolite pathways using the transcriptome data from the three ginseng cultivars. Digital gene expression profiling based on FPKM values revealed 21 unigenes (encoding 15 enzymes) with statistically significant differential gene expression between two cultivars, CS and SH (Table 5).

3.4. Fructose and mannose metabolism and the glycolysis pathway

Sugars are important carbon sources that serve as energy sources for plant growth and development [35]. Among sugars, we detected glucose and fructose in three ginseng cultivars. The table 2 Relative abundance of primary metabolites in CS, SH and CP

| Class                  | Fold change (compared to CP) | CS          | SH          | CP          |
|------------------------|-----------------------------|-------------|-------------|-------------|
| Amino acids            |                             | 1.32*       | 1.39*       | 1*          |
| Carboxylic acids       |                             | 2.09*       | 0.96*       | 1*          |
| Fatty acids            |                             | 1.07        | 0.99        | 1           |
| Inorganic acids        |                             | 0.64*       | 0.55*       | 1*          |
| Monosaccharides        |                             | 0.91*       | 0.50*       | 1*          |

Different letters in each column indicate significant differences based on least significant difference test (p < 0.05).

CP = Chunpoong; CS = Cheongsun; SH = Sunhyang.

3.5. Alanine, aspartate, and glutamate metabolism

Pyruvate, which is produced via the glycolysis pathway, serves as a carbon substrate for the biosynthesis of a wide range of compounds [37]. Alanine, valine, and serine (derived from pyruvate) were also detected in our GC-MS analysis. The relative concentration of alanine was higher in CS and SH than in CP. The transcript levels of Pg_Root127493_c1_seq3 (No. 5), encoding alanine–glutamate aminotransferase (EC 2.6.1.44), were highest in the CS cultivar. A previous study demonstrated that overexpression of the gene encoding alanine–glutamate aminotransferase increases the nitrogen uptake ability of Oryza sativa [38]. Therefore, the rapid growth rates of CS adventitious roots [20] might result from increased expression of alanine–glutamate aminotransferase. However, the relatively high concentration of alanine in SH did not correlate with the expression pattern of alanine–glutamate aminotransferase, suggesting that other factors led to an increase in alanine content in SH.

Four unigenes encoding two enzymes in the glutamate biosynthesis pathway, that is, succinate–semialdehyde dehydrogenase (EC 1.2.1.24, which converts succinate to semialdehyde) and glutamine synthetase (EC 6.3.1.2, which directly converts i-glutamate to i-glutamine), were differentially expressed among the three ginseng cultivars. Glutamine levels were highest in CS, followed by SH and CP. However, the gene expression patterns were not well correlated with the accumulation patterns of glutamine. Two unigenes, Pg_Root11075_c0_seq7 (No. 7) and Pg_Root110835_c0_seq14 (No. 8), both encoding succinate–semialdehyde dehydrogenase (EC 1.2.1.24), were identified in the CP root transcriptome. Pg_Root11075_c0_seq7 (No. 7) was highly expressed in CP, whereas Pg_Root110835_c0_seq14 (No. 8) was highly expressed in SH and CP. Moreover, Pg_Root126266_c2_seq17 (No. 10), encoding glutamine synthetase (EC 6.3.1.2), were highly expressed in SH and CP, respectively. This result suggested that various enzymes might be associated with glutamine synthesis.
3.6. Valine, leucine, and isoleucine biosynthesis

Unigenes encoding acetylacetate synthase (EC 2.2.1.6) and ketol-acid reductoisomerase (EC 1.1.1.86), which function in the valine biosynthesis pathway, were isolated in the CP root transcriptome. Among these, Pg_Root123419_c1_seq1 (No. 11), encoding ketol-acid reductoisomerase (EC 1.1.1.86), which catalyzes the conversion of serine to glycine (via the formation of threonine) exhibited similar expression patterns among cultivars. The valine levels were higher in CS and SH than in CP, but the expression of Pg_Root123419_c1_seq1 (No. 11) was highest in SH. The accumulation patterns of valine in the cultivars did not correlate with the expression of the unigenes involved in its biosynthesis, indicating that expression of other genes involved in valine biosynthesis might be more increased in CS than other cultivars.

Table 3

| Pathway | EC number | Enzyme | Queries | Number of genes from Panax ginseng |
|---------|-----------|--------|---------|-----------------------------------|
| Glycolysis pathway ko00010 | EC 3.1.3.11 | Fructose-1,6-bisphosphatase | AT3G45050.1 | 3 |
|        | EC 2.7.1.11 | 6-Phosphofructokinase | AT2G22480.1 | 4 |
|        | EC 4.1.2.13 | Fructose-bisphosphate aldolase | AT2G22480.1 | 3 |
|        | EC 1.2.1.12 | Glyceraldehyde-3-phosphate dehydrogenase | AT2G3763.1 | 1 |
|        | EC 2.7.2.3 | Phosphoglycerate kinase | AT1G56190.1 | 1 |
|        | EC 4.2.1.11 | Phosphophruvate hydratase | AT2G29560.1 | 4 |
|        | EC 2.7.1.40 | Pyruvate kinase | AT5G5290.1 | 6 |
|        | EC 1.2.1.24 | Succinate—semialdehyde dehydrogenase | AT1G79440.1 | 11 |
|        | EC 2.6.1.19 | 4-Aminobutyrate-2-oxoglutarate transaminase | AT3G22200.1 | 4 |
|        | EC 6.3.1.12 | Glutamine synthase | AT5G16570.1 | 2 |
|        | EC 2.2.1.6 | Acetolactate synthase | AT3G48560.1 | 3 |
|        | EC 1.1.1.86 | Ketol-acid reductoisomerase | XM_002525069.1 | 1 |
| Glycine, serine and threonine metabolism ko00260 | EC 4.3.1.17 | l-Serine ammonia-lyase | AT1G21540.1 | 6 |
|        | EC 2.1.2.1 | Serine hydroxymethyltransferase | AT4G13890.1 | 1 |
|        | EC 4.1.2.5 | l-Threonine aldolase | AT3G08870.1 | 1 |
|        | EC 2.7.1.39 | Homoserine kinase | XM_002517652.1 | 2 |
|        | EC 4.2.3.1 | Threonine synthase | XM_002521317.1 | 1 |
| Citrate cycle (TCA cycle) ko00020 | EC 1.2.4.1 | Pyruvate dehydrogenase | XM_002523354.1 | 2 |
|        | EC 1.8.1.4 | Dihydrolipoamide dehydrogenase | NM_001247841.1 | 2 |
|        | EC 1.1.1.37 | Cytosolic malate dehydrogenase | NM_001247223.1 | 1 |
| Fructose and mannose metabolism ko00051 | EC 4.2.1.11 | Phosphopyruvate hydratase | AT2G29560.1 | 1 |
|        | EC 1.1.1.14 | Sorbitol dehydrogenase | XM_002510809.1 | 6 |
|        | EC 1.1.1.21 | Aldo-keto reductase | XM_002529821.1 | 16 |
| Arginine and proline metabolism ko00330 | EC 6.3.4.5 | Arginosuccinate synthase | AT4G24830.1 | 1 |
|        | EC 3.5.3.1 | Arginase | AT4G08870.1 | 1 |
|        | EC 2.6.1.13 | Ornithine aminotransferase | AT5G08820.1 | 4 |
|        | EC 1.5.1.2 | Pyrroline-5-carboxylate reductase | AT4G16570.1 | 1 |

Table 4

| Advenitious root sample | Raw data | QC filtered data |
|-------------------------|----------|-----------------|
|                        | Total no. of reads | Length (bp) | Total no. of reads | Length (bp) |
| CS, replicate 1         | 16,731,664 | 2,321,621,040 | 15,291,398 | 2,115,447,628 |
| CS, replicate 2         | 14,306,820 | 1,977,938,763 | 12,831,592 | 1,766,225,247 |
| CS, replicate 3         | 14,801,822 | 2,056,937,756 | 13,407,302 | 1,857,154,376 |
| SH, replicate 1         | 17,526,160 | 2,446,652,196 | 15,710,392 | 2,185,318,113 |
| SH, replicate 2         | 16,947,000 | 2,354,285,233 | 15,409,700 | 2,134,138,490 |
| SH, replicate 3         | 16,426,236 | 2,281,803,413 | 14,810,282 | 2,050,279,385 |
| CP                      | 90,242,024 | 9,114,444,424 | 85,355,736 | 8,441,707,472 |

Total | 186,981,726 | 22,553,682,825 | 172,796,402 | 20,550,270,711 |

CP = Chunpoong; CS = Cheongsun; SH = Sunhyang.

* RNA-seq data from CP were obtained from a previous study [20].

3.7. Glycine, serine, and threonine metabolism

Serine, glycine, threonine, and aspartate were monitored in adventitious roots from the three ginseng cultivars through GC-MS analysis. The contents of serine, threonine, and aspartate were significantly different among the three cultivars. CP had the lowest levels of serine. However, unigenes encoding l-serine ammonia-lyase (EC 4.3.1.17), which converts pyruvate into serine, and serine hydroxymethyltransferase (EC 2.1.2.1) catalyze the conversion of serine to glycine, exhibited similar expression patterns among cultivars.

According to the KEGG database, five enzymes regulate the formation of threonine from aspartate. Aspartate kinase (EC 2.7.2.4), aspartate—semialdehyde dehydrogenase (EC 1.2.1.11), and homoserine dehydrogenase (EC 1.1.1.3) catalyze the conversion of aspartate to homoserine, followed by the formation of threonine by homoserinekinase (EC 2.7.1.39) and threonine synthase (EC 4.2.3.1). l-Threonine aldolase (EC 4.1.2.5) is required for the conversion of threonine to glycine. We identified unigenes putatively encoding three of these enzymes, that is, l-threonine aldolase (EC 4.1.2.5), homoserinekinase (EC 2.7.1.39), and threonine synthase (EC 4.2.3.1), in the CP root transcriptome. High levels of threonine and aspartate were detected in SH, while the levels of homoserinekinase (EC 2.7.1.39) and threonine synthase (EC 4.2.3.1), which
regulate the final step in threonine biosynthesis, were similar among the three cultivars. However, the transcript levels of Pg_Root118760_c1_seq1 encoding cytosolic malate dehydrogenase (EC 1.1.1.37), were highest in SH, as were threonine levels, indicating that threonine aldolase (EC 4.1.2.5) might partially account for the increased threonine biosynthesis observed in SH.

### 3.8. Arginine and proline metabolism

We detected ornithine and proline in the adventitious roots of the three ginseng cultivars. According to the KEGG database, three enzymes regulate the biosynthesis pathway from citrulline to ornithine. Among these, Pg_Root123415_c0_seq10 (No. 20), encoding argininosuccinate synthase (EC 6.3.4.5), were highest in SH, as were threonine levels, indicating that threonine aldolase (EC 4.1.2.5) might partially account for the increased threonine biosynthesis observed in SH.

#### 3.9. TCA cycle

We identified three enzymes involved in the citrate cycle in the Panax ginseng cultivars, including cytosolic malate dehydrogenase (EC 1.1.1.37), which converts oxaloacetate to malate as well as pyruvate dehydrogenase (EC 1.2.4.1) and dihydrolipoamide dehydrogenase (EC 1.8.1.4), where the activities of both enzymes together convert pyruvate to acetyl-CoA. For acetyl-CoA to enter the ginsenoside biosynthesis pathway, two acetyl-CoAs must be condensed to acetoacetyl-CoA by acetyl-CoA acetyltransferase (EC 2.3.1.9). Pg_Root120502_c0_seq16 (No. 14), encoding pyruvate dehydrogenase (EC 1.2.4.1), were expressed at low levels in SH and CS, respectively, suggesting that the different expression levels of these genes might lead to distinct ginsenoside accumulation patterns among cultivars. Malate levels were highest in CS, and the expression of Pg_Root118760_c1_seq1 encoding cytosolic malate dehydrogenase (EC 1.1.1.37), was also highest in CS. The levels of malonate were higher in CS and SH than in CP; however, we failed to detect genes encoding enzymes involved in this pathway in the Panax ginseng transcriptomes.

More than 20 ginseng cultivars have been officially registered with the Korea Seed and Variety Service and cultivated in Korea. Each cultivar has unique genetic and morphological diversity. In this study, we tried to understand the diversity of primary metabolites and the expression levels of genes related to the biosynthesis of these metabolites among Korean ginseng cultivars. The results demonstrated that the ginseng cultivars could be discriminated based on the variation of primary metabolite profiles and classified with eight primary metabolites as biomarkers. We also report expression level differences for candidate genes in the related biosynthetic pathways among ginseng cultivars. Our data demonstrate the value of an integrated approach of transcriptome and metabolome analysis to promote our understanding of complex metabolic pathways of plants.

### Conflicts of interest

The authors confirm that they have no conflicts of interest.
Acknowledgments

This research was carried out with the support of “Next-Generation BioGreen21 Program for Agriculture & Technology Development (Project No. PJ01103001)” Rural Development Administration, Republic of Korea.

References

[1] Vogler BK, Pittler MH, Erns TE. The efficacy of ginseng. A systematic review of randomised clinical trials. Eur J Clin Pharmacol 1999;55:567–75.
[2] Choi KT. Botanical characteristics, pharmacological effects and medicinal components of Korean Panax ginseng C A Meyer. Acta Pharmacol Sin 2008;29:1109–18.
[3] Kim NH, Choi H, Ahn IO, Yang TJ. EST-SSR marker sets for practical authentication of all nine registered ginseng cultivars in Korea. J Ginseng Res 2012;36:298–307.
[4] Kwon WS, Lee MG, Lee JH. Characteristics of flowering and fruiting in new varieties and lines of Panax ginseng CA Meyer. J Ginseng Res 2001;25:41–4.
[5] Sun H, Wang HT, Kwon WS, Kim YJ, In JG, Yang DC. A simple and rapid technique for the authentication of the ginseng cultivar, Yumpong, using an SNP marker in a large sample of ginseng leaves. Gene 2011;478:75–9.
[6] Lee JW, Kim YC, Jo BH, Seo AY, Lee JH, Kim GT, Hyno DV, Cha SW, Bang KH, Cho JH. Development of an ISSR-derived SCAR marker in Korean ginseng cultivars (Panax ginseng CA Meyer). J Ginseng Res 2011;35:52–9.
[7] Ahn JO, Lee SS, Lee JH, Lee MJ, Jo BG. Comparison of ginsenoside contents and pattern similarity between root parts of new cultivars in Panax ginseng CA Meyer. J Ginseng Res 2008;32:15–8.
[8] Kwon YK, Ahn MS, Park JS, Liu JR, In DS, Min BW, Kim SW. Discrimination of cultivation ages and cultivars of ginseng leaves using Fourier transform infrared spectroscopy combined with multivariate analysis. J Ginseng Res 2014;38:52–8.
[9] Lee EJ, Shakhurbudinov R, Weljie AM, Vogel HJ, Facchini PJ, Park SU, Kim YK, Yang TJ. Quality assessment of ginseng by 1H NMR metabolite fingerprinting and profiling analysis. J Agri Food Chem 2009;57:7513–22.
[10] Goossens A, Rischer H. Implementation of functional genomics for gene discovery in alkaloid producing plants. Phytochem Rev 2007;6:35–49.
[11] Öksman-Calderón KM, Saito K. Integrating genomics and metabolomics for engineering plant metabolic pathways. Curr Opin Biotech 2005;16:174–9.
[12] Fukushima A, Kanaya S, Nishida K. Integrated network analysis and effective tools in plant systems biology. Front Plant Sci 2014;5:1–9.
Jayakodi M, Lee SC, Lee YS, Park HS, Jang W, Lee HO, Joh HJ, Yang TJ, Patel RK, Jain M. NGS QC Toolkit: a toolkit for quality control of next generation sequencing data. PLoS ONE 2012;7:e30619.

Jayakodi M, Lee SC, Lee YS, Park HS, Kim NH, Jang W, Lee HO, Joh HJ, Yang TJ. Comprehensive transcriptome analysis of Panax ginseng root tissues identifies tissue-specific and year dependent genes. BMC Plant Biol 2015;15:1–12.

Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. BMC Bioinformatics 2011;12:1–16.

Lee DK, Yoon MH, Kang YP, Yu J, Park JH, Lee J, Sung WK. Comparison of primary and secondary metabolites for suitability to discriminate the origins of Schisandra chinensis by GC/MS and LC/MS. Food Chem 2013;141:3931–7.

Lisko JG, Stanfill SB, Duncan BW, Watson CH. Application of GC-MS/MS for the analysis of tobacco alkaloids in cigarette filler and various tobacco species. Anal Chem 2008;80:168–75.

Francisco CS, Messina GB, Lopes LM, Tininis AG, de Oliveira JE, Capellari L. Classification of Aristolochia species based on GC–MS and chemometric analyses of essential oils. Phytochemistry 2008;79:168–75.

Enhanced triterpene and phytosterol biosynthesis in Panax ginseng L. Lett 2006;580:5143–8.

Cho IH, Lee HJ, Kim YS. Differences in the volatile compositions of ginseng species (Panax sp.). J Agric Food Chem 2012;60:7616–22.

Chen S, Luo H, Li Y, Sun Y, Wu Q, Niu Y, Song J, Lu A, Zhu Y, Sun C. 454 EST analysis detects genes putatively involved in ginsenoside biosynthesis in Panax ginseng. Plant Cell Rep 2011;30:1593–601.

Tasakul P, Shibuya M, Kushiro T, Ebizuka Y. Dammaradiol-II synthase, the first dedicated enzyme for ginsenoside biosynthesis in Panax ginseng. FEBS Lett 2006;580:5143–9.

Lee MH, Jeong JH, Seo JW, Shin CG, Kim YS, In JG, Yang DC, Yi JS, Choi YE. Enhanced triterpene and phytosterol biosynthesis in Panax ginseng over-expressing squalene synthase gene. Plant Cell Physiol 2004;45:976–84.

Han JY, Hwang HS, Choi SW, Kim HJ, Choi YE. Cytochrome P450 CYP716A53v2 catalyzes the formation of protopanaxatriol from protopanaxadiol during ginsenoside biosynthesis in Panax ginseng. Plant Cell Physiol 2012;53:1535–45.

Huston M, Smith T. Plant succession: life history and competition. Am Nat 1987;130:168–98.

Causer T, Shanmugam K, Yomano L, Ingram L. Engineering Escherichia coli for efficient conversion of glucose to pyruvate. Proc Natl Acad Sci U S A 2004;101:2235–40.

Shawat AK, Carroll RT, DePauw M, Taylor GJ. Genetic engineering of improved nitrogen use efficiency in rice by the tissue-specific expression of alanine aminotransferase. Plant Biotechnol J 2008;6:722–32.