Zhao L *et al.* (2022)
Notulae Botanicae Horti Agrobotanici Cluj-Napoca
Volume 50, Issue 2, Article number 12721
DOI:10.15835/nbha50212721
Research Article

Deciphering the stem variations in ginseng plant using RNA-Seq

Lu ZHAO[1], Yan-Shuang YU[1], Xin-Fang ZHOU[1], Huxitaer REHEMAN[1], Fu-Hui WEI[1], Da-pu ZHO[1], Ping FANG[2], Jin-Zhuang GONG[2], Yong-Hua XU[1,2*]

[1]*Jilin Agricultural University, College of Chinese Medicinal Materials, Department of Traditional Chinese Medicine Resources and Development, Changchun City, Jilin Province, China; zhaol035@jlau.edu.cn; shuang429429@163.com; zhouxinfang0127@163.com; hux199624@163.com; weifuhui163@163.com; zhou877561749@163.com; xuyonghua777@yeah.net (*corresponding author)*

[2]*Jilin Agriculture University, State Local Joint Engineering Research Center of Ginseng Breeding and Application, Changchun City, Jilin Province, China; 2502891316@qq.com; 1428106161@qq.com; xuyonghua777@yeah.net (*corresponding author)*

Abstract

Ginseng is an important herb widely grown in East Asia that has medicinal and nutritional uses. Multi-stem ginseng plants undergo rapid growth, are of good quality, and have a high main-root yield. The multi-stem trait is important in ginseng breeding. To understand the molecular mechanisms responsible for the multi-stem formation, the physiological changes before and after overwintering bud formation, we analysed the transcriptomes of multi- and single-stem ginseng plants. RNA sequencing of overwintering buds from multi- and single-stem ginseng plants was performed using high-throughput second-generation sequencing. We obtained 47.66 million high quality reads at a sequencing efficiency of greater than 99% from the multi- and single-stem transcriptome. An analysis of significantly enriched gene ontology functions and comparisons with Kyoto Encyclopedia of Genes and Genomes pathways revealed expression level changes in genes associated with plant hormones, photosynthesis, steroids biosynthesis, and sugar metabolism. Plant hormones are involved in multi-stem formation in ginseng. Auxin, cytokinin, brassinolide, and strigolactone have positive effects on multi-stem formation, but further research is needed to elucidate their mechanisms. Our results have important implications in ginseng cultivation and breeding.

Keywords: ginseng; plant hormone; plant morphology; RNA-Seq Data; terpenoids

Introduction

Ginseng (*Panax ginseng* C. A. Meyer) is an important medicinal plant of the family Araliaceae. In China, Korea, and Southeast Asia, ginseng has a long history of use as a medicinal material and tonic (Kim *et al.*, 2013; Lee *et al.*, 2015), presently, ginseng is also used in food and cosmetic products (Zhai *et al.*, 2013). China and South Korea are major producers of ginseng and supply a substantial number of products to the Asian market each year. In ginseng production, the two greatest concerns of producers are yield and quality. The yield and quality of ginseng are closely associated with the cultivar, cultivation techniques, and environmental factors

Received: 03 Apr 2022. Received in revised form: 03 May 2022. Accepted: 10 May 2022. Published online: 24 May 2022.
From Volume 49, Issue 1, 2021, Notulae Botanicae Horti Agrobotanici Cluj-Napoca journal uses article numbers in place of the traditional method of continuous pagination through the volume. The journal will continue to appear quarterly, as before, with four annual numbers.

(light, water, and soil nutrients). Genotype, cultivation practices, physiology, and environmental factors influence the growth of ginseng. The F1 progeny of reciprocal crosses between two given individuals show heterosis, compared with the non-hybrid parental lines, as well as increased root weights and saponin contents (Kim *et al.*, 2017). Thus, reciprocal crossings may be useful in increasing ginseng's yield and saponin content. The leaves of 4- and 6-year-old ginseng plants are rich in osmoprotectants and oxidizing chemicals. In the latter, the saponin content decreases in the leaves, but increases in the roots by 1.2- to six-fold compared with the former (Kim *et al.*, 2018). This indicates that the long-distance transport of saponins from leaves to roots increases as ginseng plants attain maturity. Liu *et al.* (year) studied the metabolic changes in the lateral ginseng roots using 30 root samples at five developmental stages (1, 2, 3, 4, and 5 years of age) and analysed the Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways of genes expressed in the same tissues at the different growth stages. Significant changes in carbohydrate metabolism, pentose phosphate metabolism, and other metabolic pathways were observed (Liu *et al.*, 2019). Environmental factors also affect ginseng growth, with plants having the ability to adapt to soil cadmium concentrations ≤ 1.0 mg/kg, while at soil cadmium concentrations > 1.0 mg/kg their biomass decreases (Guo *et al.,* 2016). For ginseng saponin biosynthesis, a genomic analysis revealed the presence of 225 genes involved in UDP glycosyltransferases (Xu *et al.*, 2017). Liu *et al.* (2017) analysed the transcriptomes of ginseng stem and leaf tissues, and screened for genes involved in regulating growth, stress response, disease resistance, and chlorophyll metabolism, as well as those encoding enzymes associated with saponin biosynthesis (Liu *et al.*, 2017). MYB transcription factors play important roles in plant development and defense processes. Ginseng PgMYB1's expression is up-regulated by abscisic acid, salicylic acid, NaCl, and low-temperature treatment while is down-regulated by methyl jasmonate, which indicates that PgMYB1 might be involved in responses to environmental stresses and hormonal signals (Afrin *et al.*, 2015).

Studies from rice and maize have confirmed that plant biological and economic yields are closely associated with plant morphology (Sarasketa *et al.*, 2014; Tuhina-Khatun *et al.*, 2016). In general, ginseng develops a succulent straight root system, with 2-5 roots produced at the end of the main root, and few lateral roots; the aboveground stem is erect, unbranched, with verticillate compound leaves at the top of the stem (Ingram *et al.*, 2011). In a ginseng plantation in Ji'an City, Jilin Province, China, we observed that among plants growing on the same farmland, the growth form may differ markedly, even between adjacent plants. The majority of plants produced a single stem, whereas approximately 2% of the plants developed 2, 3, or 4 stems. The multi-stem growth form generally appeared in more than 3 years old plants. The multi-stem growth form resulted in an enhanced leaf area, strong growth potential, and a rapidly growing main root. The average weight of a 4-year-old multi-stem root was 15%-20% greater than that of a single-stem root, and the yield was significantly higher than that of a single-stem plant. Whether the multi-stem growth form is determined by germplasm, cultivation techniques, or environmental factors is presently unknown

Among plant hormones, auxin and cytokinin are associated with maintaining and releasing apical dominance in plants (Trainotti *et al.*, 2007; Kobayashi *et al.*, 2012; Schaller *et al.*, 2015). The contents and the concentration ratio of these two hormones in plants control the dominance of the terminal root and terminal stem over the lateral roots and lateral buds, respectively (Haver *et al.*, 2002; Woodworth *et al.*, 2007). Hence, we aim to understand is multi-stem formation in ginseng regulated by these hormones? Under identical environmental conditions, what causes differences in hormone contents between individuals? Is multi-stem development a reflection of differential gene expression in a hybrid? To answer these questions, we analysed the transcriptomes and phytohormone contents of single- and multi-stem ginseng plants grown in the same environment.

Ginseng is a heterotetraploid plant, 2n = 48, with an estimated genome size of approximately 3 Gb (Choi *et al.*, 2014). Genomic sequencing and analysis are expensive and labor-intensive. Although Korean researchers have completed a draft genome sequence for the Korean ginseng cultivar 'Chunpoong' (Jayakodi *et al.*, 2018; Boopathi *et al.*, 2020). However, we cannot sequence and compare the genomes of the two forms of

ginseng in a short period of time. Therefore, here, we analysed and compared the transcriptomes of the two growth forms, which were grown in a uniform environment from a single-seed source. We considered the multi-stem form as the experimental plant group and the single-stem form as the control plant group. An Illumina (San Diego, CA, USA) sequencing platform at Shanghai Ouyi Biomedical Technology Co., Ltd. (Shanghai, China) was used to sequence the samples' transcriptomes. Phytohormone contents in the samples were detected using high-performance liquid chromatography-tandem mass spectrometry (HPLC-MS/MS).

Materials and Methods

Plant material

At the onset of overwintering bud development in late July 2015, samples were collected from a ginseng plantation in Ji'an City, Jilin Province, China. "Two-maya" type 4-year-old ginseng roots (Figure 1) were sampled from multi- and single-stem plants growing in the same farmland under the same environmental conditions. Three individual plants (representing three biological replicates) were collected for each growth form. The single-stem type was the control plant group, and the multi-stem type was the experimental plant group. After the plants were frozen in liquid nitrogen, RNA was extracted from the overwintering buds using a mirVana™ miRNA Isolation Kit (Ambion, Austin, TX, USA) and a TruSeq Stranded mRNA LT Sample Prep Kit (Illumina).

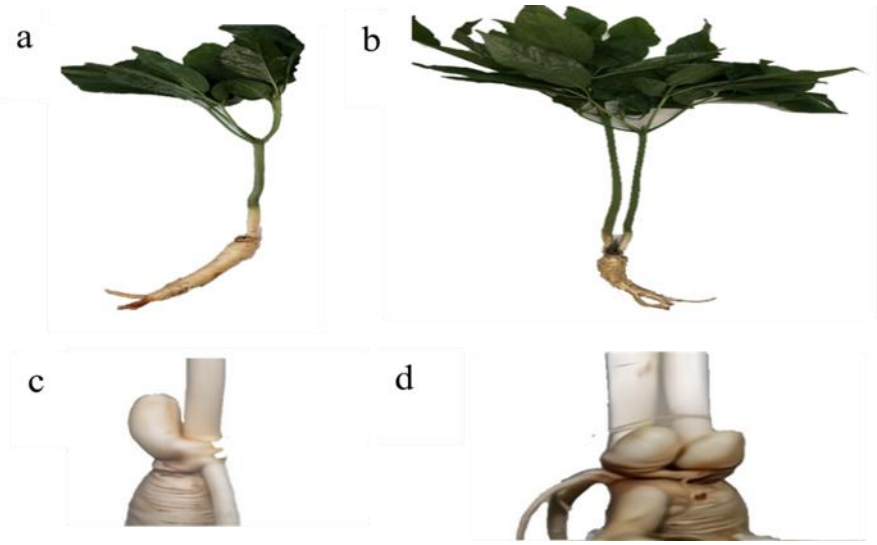


Figure 1. Morphology and quantity of ginseng stems and buds. (a) Single-stem ginseng plant, one bud in the root; (b) ginseng plant with two stems; (c) one spore of a single-stem ginseng root; (d) two buds in the root

Methods

Illumina sequencing and quality assessment

The quality of the nucleic acid sequence raw reads obtained with the Illumina HiSeq™ 2500 platform was evaluated (Figure 2) using the following procedure. Low-quality reads were filtered using a quality threshold of 20 and a length threshold of 70%. Low-quality bases were removed from the 3′ end of the sequence using a quality threshold of 20. The N-containing sequences of reads were removed at a length threshold of 35 bp. To detect contaminating sequences, 250,000 pairs (500,000) were randomly extracted from the sequence dataset and compared with the nt library (ftp://ftp.ncbi.nih.gov/blast/db), with the criteria e < 1e-10 and 80% coverage of the optimal nucleic acid sequence. The quality control procedures were performed using FastQC software (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The data were preprocessed using

NGS QC Toolkit v2.3.3 software (Patel *et al.*, 2012) to remove low-quality fragments, and the resulting data constituted the transcriptome dataset (http://59.163.192.90:8080/ngsqctoolkit/). This data set has been uploaded to the public database (No. SUB4753701).

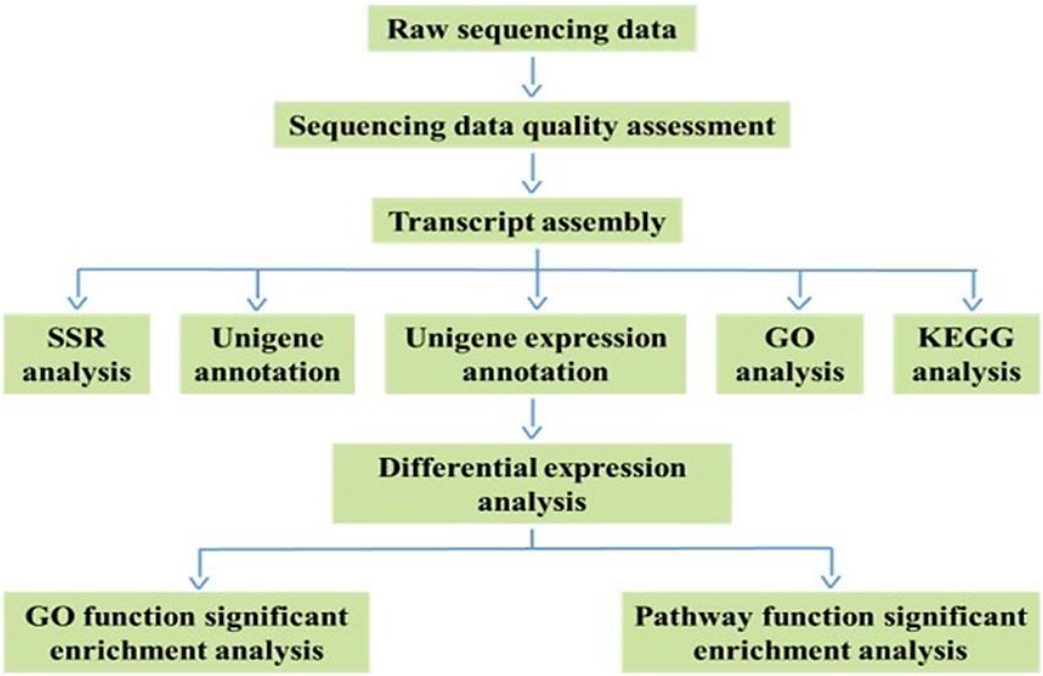


Figure 2. procedure used for transcriptome data analysis

De novo assembly

Reference genome-based spliced sequences were assembled using the paired-end splicing method with Trinity software (Grabherr *et al.*, 2011), clustered using TGICL software (Pertea *et al.*, 2003) and redundant sequences were removed to generate the final unigene sequences.

Unigene annotation

Using the BLAST algorithm (Altschul *et al.*, 1990), the unigene sequences were compared with the nr, Swiss-Prot, and EuKaryotic Orthologous Groups databases. The screening condition was e < 1e-5 and proteins with the highest sequence and functional similarities were obtained. After alignment, screening, and annotation, all the unigene data were combined into the transcriptome dataset.

Unigene expression abundance

Transcript abundances were quantified as indicators of gene expression levels. In the unigene database, the expression abundance of each unigene in each sample was determined by a sequence similarity alignment using the software Bowtie2 (Langmead *et al.*, 2012) (*http://bowtie-bio.sourceforge.net*/bowtie2/manual.shtml) and eXpress (Roberts *et al.*, 2013) (http://www.rna-seqblog.com/express-a-tool-for-quantification-of-rna-seq-data/). The unigene expression level was calculated using the fragments per kilobase per million reads method.

Differential expression analysis

Differential expression analysis was performed to identify unigenes that were differentially expressed between the samples. After screening a comparison between samples, the differentially expressed unigenes were analyzed for significantly enriched gene ontology (GO) functions and KEGG pathways. A unigene's differential expression was determined in accordance with the negative binomial distribution test in the DESeq

software package (*http://bioconductor.org*/packages/release/bioc/html/DESeq.html). The screening criteria for unigene differential expression were 0.5 > fold change (FC) > 2 and P ≤ 0.05 (Xiao *et al.*, 2015).

Verification of differentially expressed genes by quantitative RT-PCR

Using the above-mentioned screening criteria, 10 genes involved in the plant hormone biosynthesis and metabolic pathways were selected and their differential expression was verified. Total RNA was extracted using a mirVana miRNA Isolation Kit (Ambion) and stored at −80 °C. Reverse transcription was performed using the PrimeScript™ RT Reagent Kit (TaKaRa, Shiga, Japan) on an ABI 9700 PCR thermal cycler. The reaction mixture comprised 3 μL total RNA, 2 μL PrimeScript RT Master Mix, and 5 μL nuclease-free H_2O. The reaction condition was 37 °C for 15 min. After reaction termination, 90 μL nuclease-free H_2O was added to the reaction solution. Quantitative PCR was performed using the LightCycler® 480 SYBR Green I Master kit (Roche, Mannheim, Germany) on a Roche LightCycler 480 Instrument II. The reaction mixture consisted of 5 μL 2× LightCycler 480 SYBR Green I Master, 0.2 μL 10-μM forward primer, 0.2 μL 10-μM reverse primer, 3.6 μL nuclease-free H2O, and 1 μL cDNA. The reaction protocol was 40 cycles of 95 °C for 10 min, 95 °C for 10 s, and 60 °C for 30 s. The relative gene expression level was calculated using the 2−ΔΔCT method.

HPLC-MS/MS analysis of plant hormone contents in roots

Roots of the study material were collected in the early (May), mid (July), and late (September) stages of overwintering bud development. Five individual plants of each growth form were sampled. Phytohormones were extracted from the rhizomes with 80% methanol + 2% β-mercaptoethanol, concentrated to the aqueous phase, dissolved in 1 mol/L formic acid, and centrifuged for 20 min at 10,000 ×g, to obtain the supernatant. The hormone extract was purified using MCX and MAX columns to obtain basic and acidic hormone samples, respectively (Cai *et al.*, 2018). The hormone content was determined using HPLC-MS/MS (Agilent Technologies 6420, Santa Clara, CA, USA) on a ZORBAX SB-C18 column. For positive-ion-mode chromatography, mobile phase A was 5 mmol/L ammonium formate aqueous solution, mobile phase B was methanol, the flow rate was 0.15 mL/min, and the injection volume was 2 μL, and the column temperature was 40 °C. For negative-ion-mode chromatography, mobile phase A was 5 mmol/L ammonium acetate aqueous solution, mobile phase B was methanol, the flow rate was 0.15 mL/min, and the injection volume was 2 μL, and the column temperature was 40 °C. The standard products indole acetic acid (IAA), trans-zeatin (tZT), isopentenyl adenine (iP), and gibberellin A (GA) 3 (GA3) were purchased from Shanghai Yuanye Biotechnology Co., Ltd. (Shanghai, China), and the standards gibberellin A4 (GA4) and gibberellin A1 (GA1) were purchased from OlChemIm (Olomouc, Czech Republic).

Results

Sequencing data quality assessment and de novo assembly

RNA sequencing of the multi- and single-stem ginseng samples generated ~48,000,000 raw reads, from which ~47,660,000 clean reads were obtained after screening (Table 1). The Q30 proportion in each sample was >90%, The reads and Unigene comparison rate of each sample was greater than 75%, and the only comparison rate of Reads and Unigene was greater than 53%. The GC content was 43%~44%, indicating that the base sequence was stable and that the genome composition of each sample was similar. A cluster analysis showed that for each growth form (Sample_dan1.2.3 represents three sample plants of the single-stem form, and Sample_duo1.2.3 represents three sample plants of the multi-stem form.) the replicate samples showed high similarity levels, which indicated that the data were representative of the growth forms and that the biological replicates were consistent (Figure 3). The resulting high-quality reads were subjected to de novo assembly to obtain the final unigene as a reference sequence for subsequent analyses (Table 2). A total of 94,063

assembled unigenes were obtained. Among the unigenes, 21,680 had sequences ≥ 1,000 bp in length, and the average sequence length was 836.99 bp. The N50 sequence was 1,148 bp, indicating that the sequencing quality was excellent.

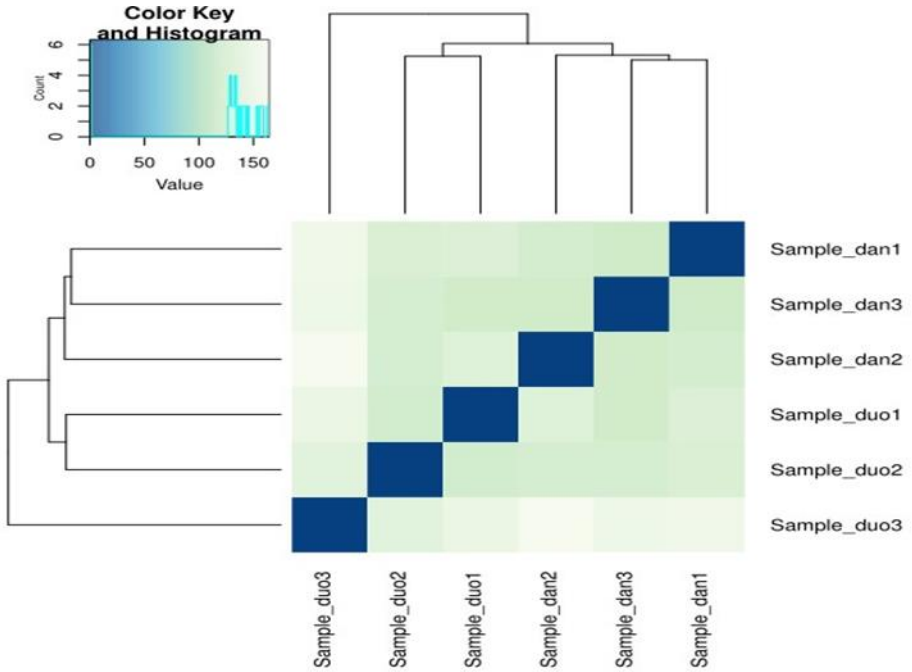


Figure 3. Cluster analysis of transcriptome data characteristics for single-stem and multi-stem ginseng samples

Table 1. Transcriptome data quality assessment

Sample ID	Clean reads	Clean bases	Q30	GC content	Mapped reads	Uniquely mapped
Sample_dan1	47665248	5958156000	91.79%	44.00%	76.78%	53.07%
Sample_dan2	47664752	5958094000	91.89%	44.00%	76.13%	54.87%
Sample_dan3	47662364	5957795500	91.90%	43.00%	75.98%	55.43%
Sample_duo1	47663862	5957982750	91.93%	43.00%	75.33%	55.89%
Sample_duo2	47663320	5957915000	92.02%	43.00%	75.64%	55.57%
Sample_duo3	47665116	5958139500	91.13%	43.00%	75.32%	54.24%

Table 2. Statistics for de novo assembly of transcriptome data

	All	>=500 bp	>=1000 bp	N50	Total Length	Max Length	Min Length	Average Length
Unigene	94063	49158	21680	1148	78729752	15906	301	836.99

Unigene expression abundance

The abundance of unigenes were determined by using gene expression levels. The average number of each sample was between 8.55 and 9.35, the median was between 0.63 and 0.83. The average value of each biological duplicate sample of single-stem-shaped ginseng is more similar, showing that the sample has good repeatability and consistent characteristics. The multi-stem morphology of the ginseng samples 2 and 3 are closer to each other, which is different from the average value of sample 1, indicating that samples 2 and 3 have better repeatability (Table 3).

Analysis of differentially expressed genes (DEGs)

Comparative analysis of differentially expressed Unigene in two forms of ginseng plants, screening conditions were P≤ 0.05. A total of 1,389 unigenes were up-regulated, 1,368 unigenes were down-regulated, and 2,757 unigenes were differentially expressed (Table 4). Plant morphogenesis is closely associated with the endogenous plant hormones. Bud differentiation and development are regulated by auxins, cytokinins, and sterols (Lakshmanan *et al.*, 1997; Sanyal *et al.*, 1998)The multi-stem form results from the differentiation and development of multiple lateral buds.

Table 3. Unigene expression abundance in single-stem and multi-stem ginseng plants

Sample	Min.	1st Qu.	Median	Mean	3rd Qu.	Max.	Sd.	Sum.
Sample_dan1	0.00	0.00	0.63	9.32	2.58	20174.84	124.72	876276.80
Sample_dan2	0.00	0.00	0.75	9.35	2.75	18416.36	124.05	879490.80
Sample_dan3	0.00	0.15	0.83	9.13	2.98	17940.57	108.55	858620.80
Sample_duo1	0.00	0.00	0.81	8.55	2.89	24798.99	110.50	804146.80
Sample_duo2	0.00	0.00	0.71	8.96	2.72	16088.04	103.45	843260.50
Sample_duo3	0.00	0.14	0.83	9.12	2.96	22038.92	123.96	857479.60

Based on the results of DEGs, GO and KEGG enrichment analysis, Unigene related to phytohormone biosynthesis and metabolism were screened. The initial screening detected unigenes associated with auxin (568), cytokinin (322), gibberellin (162), brassinosteroid (117), ethylene (47), strigolactone (8), abscisic acid (647), jasmonates (279), and salicylic acid (189) were enriched. Using the screening criteria P < 0.05 and FC> 3 or FC < 0.1, the largest numbers of differentially expressed unigenes were associated with auxin, cytokinin, and abscisic acid. Among these unigenes, comp59494_c0_seq1: ABC transporter B family member 15 is involved in auxin efflux transmembrane transporter activity; comp59582_c0_seq2 participates in responses to auxin, abscisic acid, brassinosteroids, cytokinin, and gibberellin; and comp74106_c0_seq1: Transcription factor MYB12 participates in responses to auxin and ethylene. Two unigenes were associated with cytokinin, comp29964_c0_seq2: Two-component response regulator ARR2, which is involved in the cytokinin- and ethylene-activated signaling pathways, and comp40100_c0_seq1: cytokinin trans-hydroxylase, which is involved in t-ZT biosynthesis. The unigene comp32988_c1_seq1: steroid 5-alpha-reductase DET2 is primarily involved in brassinosteroid biosynthesis (Figure 2). Ten Unigenes were selected (Table 5) based on DEGs and their relative expression levels were quantified by qRT-PCR to verify the accuracy of the RNA-seq data. The results are shown in Figure 3. The qRT-PCR results showed that the changes in the relative expression of these Unigenes were consistent with the changes in RNA-seq.

GO functional annotation and enrichment analysis

The results of the GO enrichment analysis of the unigenes were shown in Figure 4. The highest proportion of annotations in the Biological Process ontology was cellular processes, comprising 19,665 unigenes with an annotated ratio of 68.67%. Metabolic processes comprised 17,049 unigenes with an annotated ratio of 59.54%. The lowest percentage of annotations was for cell death, with 33 unigenes and an annotated ratio of 0.12%, and biological adhesion, with 87 unigenes and an annotation ratio of 0.30%. The highest percentage of annotations in the Cellular Component ontology was cell, with 22,185 unigenes and an annotation ratio of 77.48%, followed by cell part, with 22,121 unigenes and an annotation ratio of 77.25%. The lowest percentage of annotations was for extracellular matrix, with 14 unigene annotations and an annotation ratio of 0.05%, The highest proportion of annotations in the Molecular Function ontology was binding, with 16,377 unigenes and an annotation ratio of 57.19%, followed by catalytic activity, with 14,587 unigenes and an annotation ratio of 50.94%. The least frequently annotated terms were receptor regulator activity and protein tag with one and eight unigenes, respectively, and annotation ratios of 0% and 0.03%, respectively. The large number of high-level unigene annotations in the Biological Process and Cellular

Component ontology terms indicated that the tested samples were from a period of vigorous physiological activity and that genes associated with physiological processes were abundantly expressed. The GO enrichment analysis revealed that for 1,010 up-regulated unigenes, 613 terms were enriched at $P < 0.05$ and 358 terms were enriched at a FDR < 0.05. For 999 down-regulated unigenes, 568 terms were enriched at $P < 0.05$ and 387 terms were enriched at a FDR < 0.05.

Table 4. Partially expressed Unigenes in FoldChange ≥ 10 or FoldChange ≤ 0.1 in the ginseng transcriptome

Gene ID	NR. Des	Other description	Fold Change	Pathway
comp100216_ c0_seq1	Disease resistance protein family	Apoptotic ATPase (KOG)	46.84	--
comp100658_ c0_seq1	Unnamed protein product [Vitis vinifera]	Beta-xylosidase/ arabinofuranosidase (SWI)	23.64	ko00500; ko00520
comp81902_ c0_seq2	ethylene-responsive transcription factor ERF026-like	Ethylene-responsive transcription factor ERF027(SWI)	19.35	GO:0005634; GO:0003677; GO:0003700
comp90562_ c0_seq1	calcium-transporting ATPase 13	Calcium transporting ATPase (KOG)	12.12	GO:0005887; GO:0043231; GO:0005524
comp70881_ c0_seq1	zinc finger protein ZAT12-like	Zinc finger protein ZAT12 (SWI); Transcription factor activity (GO)	11.87	GO:0005634; GO:0046872; GO:0003700
comp65136_ c0_seq2	ABC transporter G family member 31 isoform X2	Pleiotropic drug resistance proteins (PDR1-15) (KOG)	0.09	GO:0016021; GO:0005524; GO:0016887
comp80502_ c0_seq1	ribulose bisphosphate carboxylase small chain	Ribulose bisphosphate carboxylase small chain (SWI)	0.09	ko00630; ko00710; ko01200
comp29964_ c0_seq2	two-component response regulator ARR14-like isoform X1	Two-component response regulator ARR2 (SWI); GATA-4/5/6transcription factors (KOG)	0.06	ko04075 GO:0005634; GO:0003677
	Cytochrome-C oxidase subunit 3	Cytochrome oxidase subunit III and related proteins (KOG) Cytochrome-C oxidase, subunit II,		

comp76169_ c0_seq1	Cytochrome-C oxidase subunit 2 (mitochondrion)	and related proteins (KOG)	0.05	ko00190; ko04260 GO:0016021; GO:0005743
comp69352_ c0_seq1			0.04	ko00190; ko04260 GO:0016021; GO:0005751

Table 5. Expression changes of Unigenes in RNA-seq data in experimental and control groups

Gene	Gene ID	Description	Fold Change	Pathway
XYL4	comp100658_c0_seq1	beta-D-xylosidase 4	23.64	ko00500; ko00520
CKX	comp85049_c0_seq1	cytokinin dehydrogenase	1.90	ko00908
BZR1/2	comp92613_c0_seq6	brassinosteroid resistant 1/2	1.49	ko04075
EIN3	comp90339_c0_seq1	ethylene-insensitive protein 3	1.57	ko04075
FTH1	comp29719_c0_seq1	ferritin heavy chain	39.24	ko04978
S30e	comp73943_c0_seq2	small subunit ribosomal protein S30e	0.69	ko03010
734A1	comp95119_c1_seq1	PHYB activation tagged suppressor 1	0.61	ko00905
CYP735A	comp90352_c0_seq1	cytokinin trans-hydroxylase	0.16	ko00908
90B/724B	comp74204_c0_seq1	steroid 22-alpha-hydroxylase	0.56	ko00905
L38e	comp70664_c1_seq1	large subunit ribosomal protein L38e	0.69	ko03010

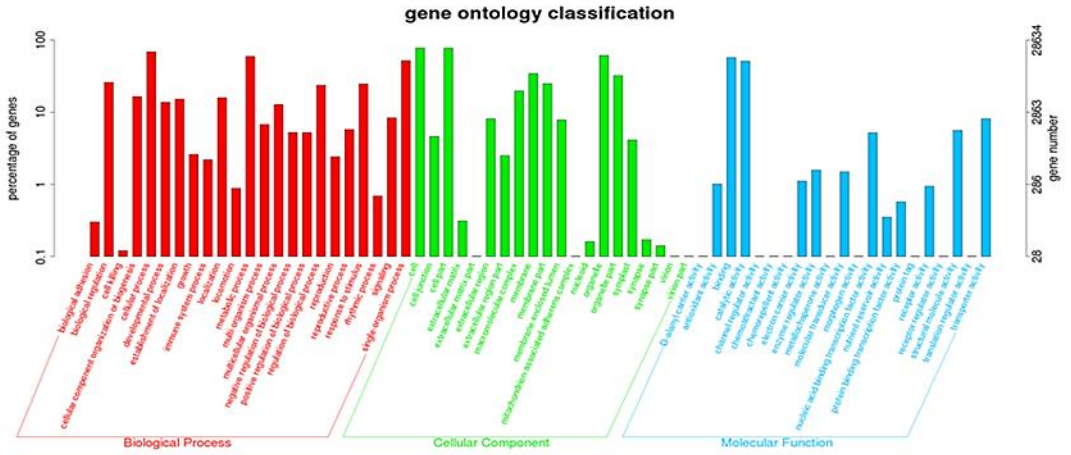


Figure 4. GO annotation statistics under level 2 level (A total of 64 types of functions were annotated, of which biological processes accounted for 23 types, cellular processes accounted for 21 types, and molecular Functions accounted for 20 types.)

KEGG functional annotation and enrichment analysis

A total of 11,445 unigenes were mapped to KEGG pathways, accounting for 12.17% of the total number of unigenes. The KEGG pathway analysis was performed on the differentially expressed unigenes and the enrichment significances were calculated. The analysis yielded 208 Tested term. At P < 0.05, 51 terms were enriched, and at a FDR < 0.05, 12 terms were enriched. Among the significantly up-regulated Tested term at P < 0.05, 47 terms were enriched, and at a FDR < 0.05, 15 terms were enriched; among significantly down-regulated Tested term at P < 0.05, 35 terms were enriched, and at a FDR < 0.05, 9 terms were enriched (Table 6). The most significantly enriched among the top 20 pathways (Figure 5) were photosynthesis–antenna proteins (P = 2.16E-07), followed by photosynthesis (P = 2.67E-07), sesquiterpenoid and triterpenoid biosynthesis (P = 3.03E-05), and porphyrin and chlorophyll metabolism (P = 5.22E-05) (Figure 5). In the KEGG enrichment analysis, the brassinosteroid biosynthesis pathway (P = 0.01633) was significantly enriched among differentially expressed unigenes. For the photosynthesis and sesquiterpenoid and triterpenoid biosynthesis pathways, under the screening criteria P ≤ 0.05 or 0.5 > FC > 2, 49 and 15 unigenes were detected, respectively. Under the screening criteria P ≤ 0.05 or 0.1 > FC > 3, one unigene (comp80502_c0_seq1, encoding ribulose-bisphosphate carboxylase large chain [EC:4.1.1.39]) and two unigenes (com54918_c0_seq1 and comp54918_c1_seq1, both encoding ent-kaurenoic acid hydroxylase [EC:1.14.13.79]), respectively, were detected. It was found in the KEGG enrichment analysis that the most enriched pathways for differentially expressed genes were Ribosome (31 Unigenes), Starch and sucrose metabolism (19 Unigenes), Photosynthesis (17 Unigenes), Plant-pathogen interaction (17 Unigenes), Plant hormone signal transduction (17 Unigenes), Carbon metabolism (17 Unigenes), Phenylpropanoid biosynthesis (15 Unigenes), Amino sugar and nucleotide sugar metabolism (15 Unigenes), Endocytosis (15 Unigenes), Biosynthesis of amino acids (14 Unigenes) etc.Here, we also pay special attention to the differences in the expression of plant hormone synthesis and metabolism and signalling pathway-related genes in the two forms of ginseng plants (Figure 6). Among them, the Auxin input vector AUX1 was down-regulated in multi-stem ginseng, which reduced the inhibition of AUX / IAA by TIR1. Down-regulation of SAUR expression indicates that lateral transport of auxin is reduced, and contralateral bud development may be beneficial. Cytokinin receptor A-ARR is directly induced by cytokinin expression, A-ARR negative feedback inhibits cytokinin receptor B-ARR expression, B-ARR expression is down-regulated, indicating that receptor A-ARR and cytokinin are in a relative High concentration, which facilitates cell division and initial development of buds (Figure 6a). In the brassinosteroid signalling pathway, the expression of the transcription factor BZR1 / 2 gene is up-regulated, which inhibits

downstream BR gene transcription and positively regulates the brassinosteroid signalling pathway, indicating that brassinosteroid transport is active in multi-stem ginseng (Figure 6b).

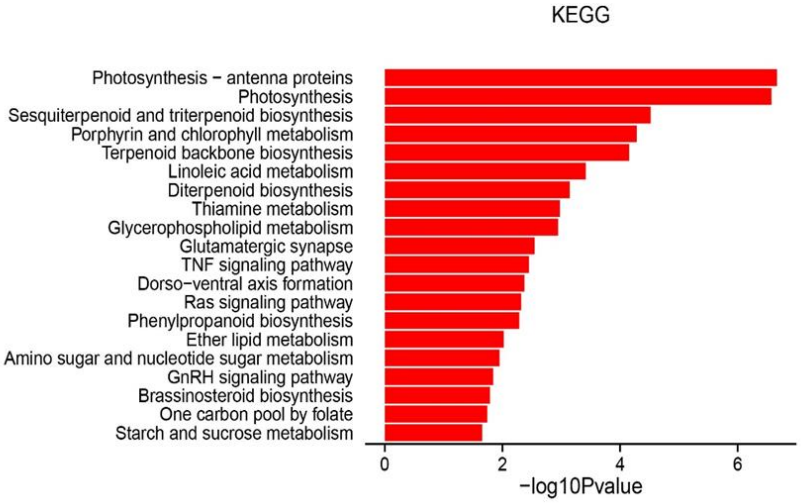
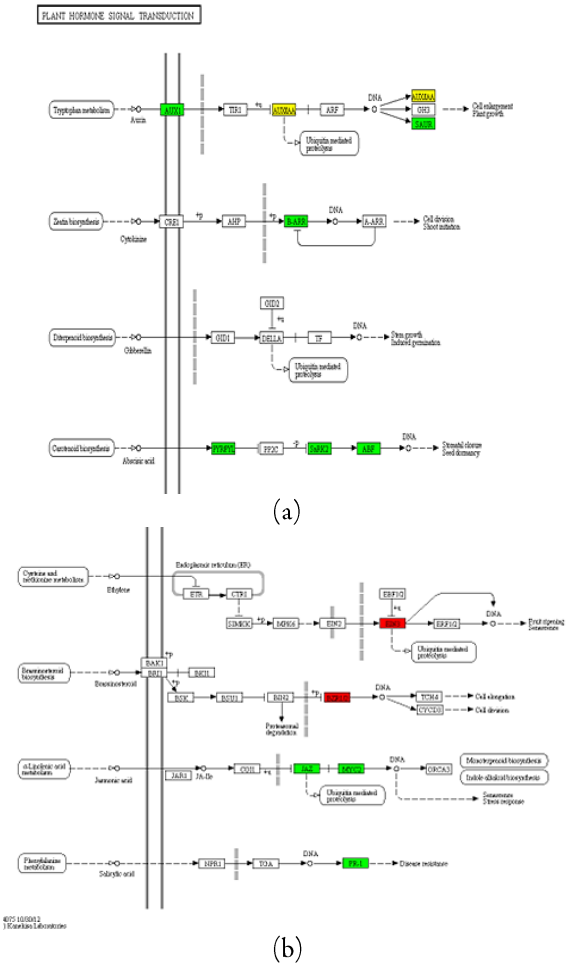


Figure 5. KEGG enrichment analysis of unigenes



(a)

(b)

Figure 6. Plant hormone signal transduction pathway

(a) Auxin, Cytokinine, Gibberellin, Abscisic acid signal transduction pathway; (b) Ethylene, Brassinosteroid, Jasmonic acid, Salicylic acid signal transduction pathway. Down- and up-regulations are marked with green and red colors.

Table 6. KEGG enrichment analysis statistics

Groups	Tested term	P Value <0.05	P Value <0.01	FDR <0.05	FDR <0.01
Sample_dan-Sample_duo up	171	47	22	15	9
Sample_dan-Sample_duo Total	208	51	18	12	5
Sample_dan-Sample_duo down	168	35	15	9	7

Analysis of the main plant hormones in ginseng roots

The contents of the phytohormones IAA, t-ZT, iP, GA1, GA3, and GA4 in the roots at early, mid, and late stages of overwintering bud development were determined (Figure 7). The IAA contents in the roots of single-stem plants were consistently higher than those in the roots of multi-stem plants at the same developmental stage, especially in the early and late stages, when they were 1.89-fold and 2.85-fold greater, respectively. The main active cytokinins detected in the roots were t-ZT and iP. The t-ZT contents at the early and mid-stages of multi-stem plants were 1.58- and 1.1-fold those of single-stem plants, whereas at the late stage, the t-ZT content of single-stem plants was 1.37-fold that of the multi-stem plants. The iP contents were 1.4- and 3.17-fold higher at the early and mid-stages, respectively, in single-stem plants compared with those of the multi-stem plants, whereas the iP content in the late-stage multi-stem plants was 1.13-fold that of the single-stem plants.

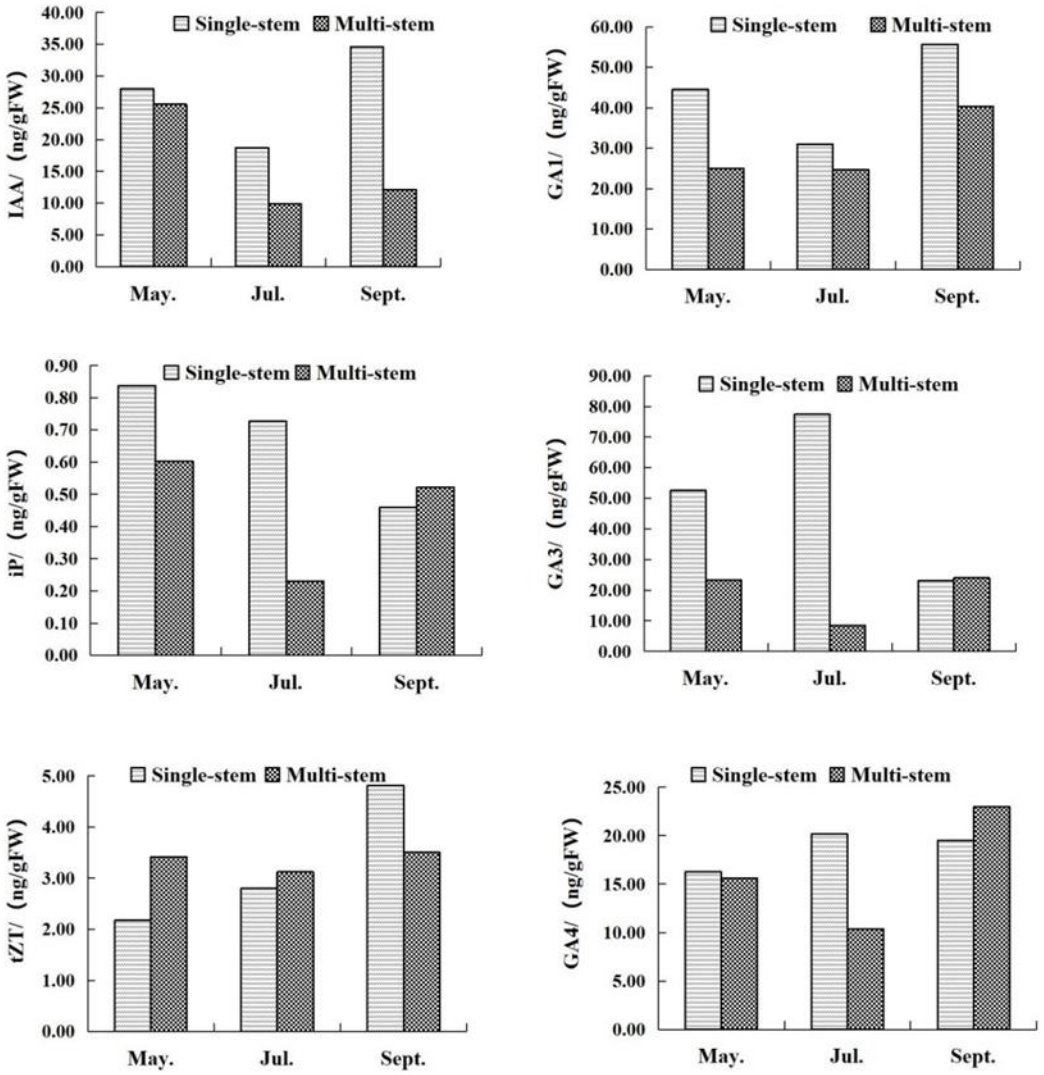


Figure 7. Contents of auxin, cytokinins, and gibberellins in the root of ginseng plants

The physiologically active GA1, GA3, and GA4 contents differed between single- and multi-stem plants. Their contents at the early and mid-stages were higher in single-stem compared with in multi-stem plants. In the late stage of overwintering bud formation, the GA1 content in single-stem plants remained higher than that in multi-stem plants, whereas the GA3 and GA4 contents were similar or lower than those in multi-stem plants. In the early and mid-stages of overwintering bud formation, the IAA, GA1, GA3, and GA4 contents in plants of the two growth forms were consistent.

Discussion

The physiological basis and molecular mechanisms responsible for the morphological differences in ginseng growth forms were investigated using transcriptome sequencing and changes in the main morphogenesis-regulating phytohormone contents during overwintering bud formation. This is the first report on the morphological differences in ginseng growth forms. During the formation of overwintering buds, the IAA, t-ZT, iP, and GA contents are dynamically changing, as is the auxin-to-cytokinin concentration ratio.

Significant differences have been detected in the expression levels of genes and transcription factors involved in hormone biosynthesis, metabolism, and signal transduction, such as IAA, t-ZT, iP, GA, brassinolide, and strigolactone. Our results provide insights into the roles of plant hormones in ginseng development. A plant with a single overwintering bud has high IAA, GA1, GA3, and GA4 contents, which maintains apical dominance and inhibits lateral bud development on the roots. The roots of plants that develop multiple overwintering buds have higher t-ZT and iP contents, suggesting that apical dominance is weakened. The number of overwintering buds determines the number of stems in the second year of ginseng growth. Is multi-stem formation in ginseng a result of environmental factors or is it determined by the genetic constitution? Previous studies have indicated that nitrogen fertilizer applications to soil, resulting in high soil nitrogen contents, promote the synthesis of endogenous cytokinins. High cytokinin contents are beneficial in relieving apical dominance, and they promote lateral bud and root formation, leading to multi-bud development. The plant materials used in the present study were continuously grown on the same farmland for 4 years, and the growth environment and cultivation techniques were consistent. The proportion of multi-stem plants in the population was ~20%. Multi-stem morphological traits were observed in plants more than 3 years old and were rarely observed in 1- and 2-year-old plants. Additional research is required to determine whether the development of multiple overwintering buds is influenced by environmental factors and cultivation techniques.

Ginseng is a heterotetraploid plant, and the floral structure shows that it requires cross-pollination to set seeds. The morphology of the aboveground plant parts differs markedly with age. This difference is determined by the developmental characteristics of the species and the genetic constitution. Therefore, we speculate that the growth of the multi-stem form is determined by a difference in individual genetic backgrounds. Korean researchers have assembled a complete draft genome sequence of the cultivar 'Chunpoong' and constructed a database based on data provided by international scholars, but it is not yet possible to analyse the complete genomes of the multi- and single-stem forms, nor analyse gene alignments; therefore, additional evidence is needed to determine the molecular mechanisms underlying the difference in growth form.

At present, our research is focused on the following topics. First, whether plant traits that contribute to high-yield and high-quality breeding objectives can be stably and continuously expressed in offspring and, if so, whether such plants can serve as ginseng breeding materials. The second is when planting ginseng, can nitrogen levels be increased by increasing the application of nitrogen fertilizer or rotation with leguminous plants to promote the biosynthesis and accumulation of cytokinin hormones in ginseng plants and promote the development of multiple overwintering buds. Stem, increase photosynthesis area, increase yield, these questions need to be answered by further research.

As can be seen from the sample-to-sample comparison in Figure 3, the three samples of the single-stem (Sample_dan1/2/3) cluster in a group. In particular, the characteristics of Sample_dan1 and Sample_dan3 are very similar. Meanwhile, in the three samples of multi-stem, Sample_duo1 and Sample_duo2 are very close. Although there are differences in characteristics among Sample_duo3, Sample_dan1/2/3 cluster group and Sample_duo1/2 cluster group, the main features of the samples are consistent and they finally cluster into a group. The results may be related to the difference in plant morphology. Sample_duo3 has 3 stems and 3 overwintering buds, and sample_duo1/2 has 2 stems and 2 overwintering buds, respectively. In addition, there was no significant difference between the median values (Median) of the two groups in table 3. These results suggest that for such heterozygosis as ginseng, an experimental sample consisting of 3-5 plants or even more and repeating at least 3 times can effectively avoid the effect of individual differences of the plant. Then the experimental data which is more characteristic with better repeatability can be obtained.

Conclusions

The number of ginseng overwintering buds determines the number of stems and the morphology of ginseng plants. An analysis of significantly enriched gene ontology functions of the transcriptomes of multi- and single-stem ginseng revealed expression level changes in genes associated with plant hormones, photosynthesis, steroids biosynthesis, and sugar metabolism. Plant hormones are involved in multi-stem formation in ginseng. Auxin, cytokinin, brassinolide, and strigolactone have positive effects on multi-stem formation.

Authors' Contributions

Conceptualization, Lu Zhao and Yong-Hua Xu; Formal analysis, Xin-Fang Zhou, Yan-Shuang Yu, Ping Fang, Jin-Zhuang Gong; Methodology, Lu Zhao, Da-pu Zhou, Huxitaer Reheman and Fu-hui Wei; Project administration, Yonghua Xu; Writing original draft, Xin-Fang Zhou; Writing review & editing, Lu Zhao.

All authors read and approved the final manuscript.

Ethical approval (for researches involving animals or humans)

Not applicable.

Acknowledgements

This work was supported by the Project of Jilin Province Science and Technology Development Project, grant number 20190304023YY; and funded by the National Key R & D Projects of China, grant number 2017YFC1702101.

Conflict of Interests

The authors declare that there are no conflicts of interest related to this article.

References

Afrin S, Zhu J, Cao H, Huang J, Xiu H, Luo T, Luo Z (2015). Molecular cloning and expression profile of an abiotic stress and hormone responsive MYB transcription factor gene from *Panax ginseng*. Acta Biochimica et Biophysica Sinica 47(4):267-277. *https://doi.org/10.1093/abbs/gmv01221.23*

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990). Basic local alignment search tool. Journal of Molecular Biology 215(3):403-410. *https://doi.org/10.1016/s0022-2836(05)80360-2*

Boopathi V, Subramaniyam S, Mathiyalagan R, Yang DC (2020). Till 2018: a survey of biomolecular sequences in genus *Panax*. Journal of Ginseng Research 44(1):33-43. *https://doi.org/10.1016/j.jgr.2019.06.004*

Cai T, Meng X, Liu X, Liu T, Wang H, Jia Z, ... Ren X (2018). Exogenous hormonal application regulates the occurrence of wheat tillers by changing endogenous hormones. Frontiers in Plant Science 9:1886. *https://doi.org/10.3389/fpls.2018.01886*

Choi HI, Waminal NE, Park HM, Kim NH, Choi BS, Park M, ... Yang TJ (2014). Major repeat components covering one-third of the ginseng (*Panax ginseng* CA Meyer) genome and evidence for allotetraploidy. The Plant Journal 77(6): 906-916. *https://doi.org/10.1111/tpj.12441*

Ingram P, Dettmer J, Helariutta Y, Malamy JE (2011). Arabidopsis lateral root development 3 is essential for early phloem development and function, and hence for normal root system development. Plant Journal 68(3):455-67. *https://doi.org/10.1111/j.1365-313X.2011.04700.x*

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, ... Regev A (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nature Biotechnology 29(7):644-652. *https://doi.org/10.1038/nbt.1883*

Guo H, Hong C, Chen X, Xu Y, Liu Y, Jiang D, Zheng B (2016). Different growth and physiological responses to cadmium of the three *Miscanthus* species. PloS One 11(4):e0153475. *https://doi.org/10.1371/journal.pone.0153475*

Haver DL, Schuch UK, Lovatt CJ (2002). Exposure of *Petunia* seedlings to ethylene decreased apical dominance by reducing the ratio of auxin to cytokinin. Journal of Plant Growth Regulation 21(4):459-468. https://doi.org/10.1007/s00344-002-0022-3

Jayakodi M, Choi BS, Lee SC, Kim NH, Park JY, Jang W, ... Yang TJ (2018). Ginseng Genome Database: an open-access platform for genomics of *Panax ginseng*. BMC Plant Biology 18(1):1-7. *https://doi.org/10.1186/s12870-018-1282-9*

Kim H, Kim JH, Lee PY, Bae KH, Cho S, Park BC, ... Park SG (2013). Ginsenoside Rb1 is transformed into Rd and Rh2 by *Microbacterium trichothecenolyticum*. Journal of Microbiology and Biotechnology 23(12):1802-1805. *https://doi.org/10.4014/jmb.1307.07049*

Kim YJ, Silva J, Zhang D, Shi J, Joo SC, Jang MG, ... Yang DC (2016). Development of interspecies hybrids to increase ginseng biomass and ginsenoside yield. Plant Cell Reports 35(4):779-790. *https://doi.org/10.1007/s00299-015-1920-8*

Kim YJ, Joo SC, Shi J, Hu C, Quan S, Hu J, ... Zhang D (2018). Metabolic dynamics and physiological adaptation of Panax ginseng during development. Plant Cell Reports 37(3):393-410. *https://doi.org/10.1007/s00299-017-2236-*

Kobayashi K, Baba S, Obayashi T, Sato M, Toyooka K, Keränen M, ... Masuda T (2012). Regulation of root greening by light and auxin/cytokinin signaling in *Arabidopsis*. The Plant Cell 24(3):1081-1095. *https://doi.org/10.1105/tpc.111.092254*

Lakshmanan P, Ng SK, Loh CS, Goh CJ (1997). Auxin, cytokinin and ethylene differentially regulate specific developmental states associated with shoot bud morphogenesis in leaf tissues of mangosteen (*Garcinia mangostana* L.) cultured *in vitro*. Plant Cell Physiology 38(1):59-64. *https://doi.org/10.1093/oxfordjournals.pcp.a029085*

Langmead B, Salzberg SL (2012). Fast gapped-read alignment with Bowtie 2. Nature Methods 9(4):357-359. *https://doi.org/10.1038/nmeth.1923*

Lee SM, Bae BS, Park HW, Ahn NG, Cho BG, Cho YL, Kwak YS (2015). Characterization of Korean red ginseng (*Panax ginseng* Meyer): History, preparation method, and chemical composition. Journal of Ginseng Research 39(4):384-391. *https://doi.org/10.1016/j.jgr.2015.04.009*

Liu J, Liu Y, Wang Y, Abozeid A, Zu YG, Zhang XN, Tang ZH (2017). GC-MS metabolomic analysis to reveal the metabolites and biological pathways involved in the developmental stages and tissue response of *Panax ginseng*. Molecules 22(3):496. *https://doi.org/10.3390/molecules22030496*

Liu S, Liu M, Wang S, Lin Y, Zhang H, Wang Q, Zhao Y (2017). Analysis of the *Panax ginseng* stem/leaf transcriptome and gene expression during the leaf expansion period. Molecular Medicine Reports 16(5):6396-6404. *https://doi.og/10.3892/mmr.2017.7377*

Patel RK, Jain M (2012). NGS QC Toolkit: A toolkit for quality control of next generation sequencing data. PLoS One 7(2):e30619. *https://doi.org/10.1371/journal.pone.0030619*

Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, ... Quackenbush J (2003). TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics 19(5):651-652. *https://doi.org/10.1093/bioinformatics/btg034*

Roberts A, Pachter L (2013). Streaming fragment assignment for real-time analysis of sequencing experiments. Nature Methods 10(1):71-73. *https://doi.org/10.1038/nmeth.2251*

Sanyal D, Bangerth F (1998). Stress induced ethylene evolution and its possible relationship to auxin-transport, cytokinin levels, and flower bud induction in shoots of apple seedlings and bearing apple trees. Plant Growth Regulation 24(2):127-134. *https://doi.org/10.1023/A:1005948918382*

Sarasketa A, González-Moro MB, González-Murua C, Marino D (2014). Exploring ammonium tolerance in a large panel of *Arabidopsis thaliana* natural accessions. Journal of Experimental Botany 65(20):6023-6033. *https://doi.org/10.1093/jxb/eru342*

Schaller GE, Bishopp A, Kieber JJ (2015). The Yin-Yang of hormones: cytokinin and auxin interactions in plant development. The Plant Cell 27(1):44. *https://doi.org/10.1105/tpc.114.133595*

Trainotti L, Tadiello A, Casadoro G (2007). The involvement of auxin in the ripening of climacteric fruits comes of age: the hormone plays a role of its own and has an intense interplay with ethylene in ripening peaches. Journal of Experimental Botany 58(12):3299-3308. *https://doi.org/10.1093/jxb/erm178*

Tuhina-Khatun M, Hanafi MM, Rafii MY, Wong MY, Aktar-Uz-Zaman M (2016). Characterisation of globally diverse blast-resistant upland rice (*Oryza sativa* L.) germplasms based on morpho-physiological and yield attributes. Acta Agriculturae Scandinavica, Section B—Soil & Plant Science 66(5):417-431. *https://doi.org/10.1080/09064710.2016.1155640*

Woodworth JC, Tokach MD, Nelssen JL, Goodband RD, Dritz SS, Koo SI, ... Owen KQ (2007). Influence of dietary L-carnitine and chromium picolinate on blood hormones and metabolites of gestating sows fed one meal per day. Journal of Animal Science 85(10):2524-2537. *https://doi.org/10.2527/jas.2006-284*

Xiao X, Ma J, Sun Y, Yao Y (2015). A method for the further assembly of targeted unigenes in a transcriptome after assembly by Trinity. Frontiers in Plant Science 6:843. *https://doi.org/10.3389/fpls.2015.00843*

Xu J, Chu Y, Liao B, Xiao S, Yin Q, Bai R, ... Chen S (2017). Panax ginseng genome examination for ginsenoside biosynthesis. Gigascience 6(11):gix093. *https://doi.org/10.1093/gigascience/gix093*

Zhai J, Wang Y, Sun C, Jiang S, Wang K, Zhang Y, ... Zhang M (2013). A plant-transformation-competent BIBAC library of ginseng (*Panax ginseng* CA Meyer) for functional genomics research and characterization of genes involved in ginsenoside biosynthesis. Molecular Breeding 31(3):685-692. *https://doi.org/10.1007/s11032-012-9826-4*