Transcriptome Analysis of Cultivated and Wild Ginseng in Different Growth Years Revealed the Regulatory Mechanism of Ginsenosides

Xiaoxue Fang

Northeast Normal University

Manqi Wang

Northeast Normal University

Xinteng Zhou

Northeast Normal University

Huan Wang

Northeast Normal University

Huaying Wang (✉ wanghy609@nenu.edu.cn)

Northeast Normal University

Hongxing Xiao

Northeast Normal University

Research Article

Keywords: ginsenoside, growth years, Panax ginseng, transcriptome, transcription factors

Posted Date: January 11th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1091473/v2>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Background: As a famous Chinese medicine, ginseng has been used in the world for nearly 5,000 years. Wild ginseng is endangered due to environmental damage. Thus, cultivated ginseng is developed to replace wild ginseng. The morphological and physiological characteristics of both wild ginseng and cultivated ginseng change during growth, and the mechanism of this change is not yet understood.

Results: This study performed transcriptome sequencing on the roots, stems and leaves of cultivated ginseng and wild ginseng with different growth years, exploring the effect of growth years on gene expression in ginseng. The number of DEGs in cultivated ginseng is more than that in wild ginseng. Based on the weighted gene co-expression network analysis, we found that the growth years significantly affected the gene expression of MAPK signaling pathway - plant and terpenoid backbone biosynthesis pathway in cultivated ginseng, but had no effects in wild ginseng. Furthermore, the growth years had significant effects on the genes related to ginsenoside synthesis in cultivated ginseng, and the effects were different in the roots, stems and leaves. However, it had little influence on the expression of genes related to ginsenoside synthesis in wild ginseng and no effect on leaves. These results showed wild ginseng was less affected by growth years than cultivated ginseng. Furthermore, HMGR, SS, DXS, DS, IspF, AACT, CYP450 and UGTs were related with MYB, NAC, AP2/ERF, bHLH and WRKY transcription factors. Growth years may regulate genes for ginsenoside synthesis by influencing these transcription factors, thereby affecting the content of ginsenosides.

Conclusions: This study complemented the gaps in the genetic information of wild ginseng in different growth periods and different tissues and provided a new insight into the mechanism of ginsenoside regulation.

Background

Ginseng (*Panax ginseng* C.A. Meyer) belongs to the genus *Panax* (Araliaceae family) and has been widely used as a significant source of natural medicine for thousands of years in East Asia, particularly in China, Korea and Japan [1]. Ginseng has been highly valued and vigorously promoted due to its important medicinal properties. The pharmacological active substances in ginseng include ginsenosides, flavonoids and polysaccharides, among which ginsenosides have been proved to be the main bioactive compounds in ginseng [2]. Studies have demonstrated that ginsenosides has various biological activities, such as immune modulation, anti-inflammation, anti-tumor, anti-amnestic and anti-aging activities [3–6]. In addition, ginsenosides also have a defensive effect on pathogenic microorganisms and herbivorous insects [7, 8].

Ginsenosides, triterpenoid saponins with a four-ring skeleton structure, are unique to ginseng genus. To date, more than 200 triterpene saponins have been isolated and characterized from the root, stem, leaf, flower and fruit of *P. ginseng* [9]. There are two main classes of ginsenosides based on the skeletons of their aglycones, namely, the dammarane type and the oleanane type. Dammarane type ginsenosides are

divided into two groups based on their structure, the protopanaxadiol (PPD) group, including Ra1, Ra2, Rb1, Rb2, Rb3, Rc, Rd, Rg3, Rh2 and others and protopanaxatriol (PPT) group included Rg1, Rg2, Re, Rf, Rh1 and so on, while oleanane-group ginsenoside only has one saponin, Ro [10, 11]. Based on previous studies, the biosynthetic pathways of ginsenosides are as follows. Firstly, *isopentenyl pyrophosphate* (IPP) and *dimethylallyl pyrophosphate* (DMAPP) are synthesized by the *mevalonic acid* (MVA) and *2-C-methyl-D-erythritol-4-phosphate* (MEP) pathways, respectively. The MVA pathway is located in cytosol while the MEP pathway is in chloroplast. Then, 2,3-oxidosqualene is produced by IPP and DMAPP via cyclization reactions catalyzed by *geranyl pyrophosphate synthase* (GPS), *farnesyl pyrophosphate synthase* (FPS), *squalene synthase* (SS) and *squalene epoxidase* (SE) [12]. The 2, 3-oxidosqualene is cyclized by *oxidosqualene cyclases* (OSCs), generating the two triterpenoid backbones, *dammarenediol-II* and β -*amyrin*. The OSC enzyme family include *dammarenediol synthase* (DS) and β -*Amyrin synthase* (β -AS). Finally, triterpenoids are modified by some specific cytochrome P450-dependent monooxygenases (CYP450s) and UDP-dependent glycosyltransferases (UGTs), resulting in a number of various ginsenosides [13].

The content of ginsenosides exhibited various among different tissues and growing years for ginsengs. Liu et al (2017) suggested the content of ginsenoside was different in various tissues, and it was significantly higher in roots than that in the stems and leaves of ginseng [14]. It is generally believed that the longer the ginseng grows, the higher the content of ginsenosides, which has been reported by many previous studies [15]. For example, Zhang et al (2014) found that the content of ginsenoside risen with the increase of cultivation years by comparing the content of ginsenosides from one to thirteen years of ginsengs [16]. So far, transcriptome analysis has been widely used to study the complex biosynthetic pathways of ginsenosides and expression characteristics of related genes. Jayakodi et al (2015) found 38 encoding enzymes involved in ginsenoside biosynthesis pathway by transcriptome sequencing of one-year-old and six-year-old ginseng roots [17]. In different tissues of five-year-old ginseng, the expression level of the MEP pathway genes were similar to those of the MVA pathway gens in roots, but higher in leaves [18]. However, the effects of different tissues and growth years of ginseng on the expression levels of genes related to ginsenoside biosynthesis have not been studied in detail.

Moreover, due to endangered resource of wild ginseng resulted from excessive consumption, ginsengs are mainly obtained through cultivation via modern agricultural technologies [19]. Cultivated ginseng is planted in farmland that was once forested. Wild ginseng grows naturally in the forest, and the growth rate of wild ginseng is much slower than that of cultivated ginseng. Generally, cultivated ginseng faces great pressure from various plant diseases and does not usually survive beyond sixth year. In contrast, wild ginseng could grow up to hundreds of years under natural conditions [20–22]. These differences in growth environment and growth time may be related to differences in pharmacological activity in wild ginseng and cultivated ginseng. However, the difference in these mechanisms is currently unclear. Hence, in this study, high-throughput transcriptome sequencing was performed on roots, stems and leaves of wild ginseng and cultivated ginseng with different growing years to explore the effect of growing years and tissues on gene expression, especially the DEGs involved in ginsenoside biosynthesis and regulation mechanism of ginsenoside biosynthesis. It may provide information resources for improving

transcriptome data of ginseng and also has profound significance for further research on ginseng cultivation breeding and related functional candidate genes.

Results

Overview of RNA-seq data

A total of 33 individuals were sequenced with an average amount of clean bases per individual was 6.5 Gb. After filtering low-quality reads, the clean reads were mapped to the ginseng reference genome with an average mapping rate of 87.44%. We first evaluated the variation in gene expression among biological replicates for each sample. The correlation values were high between them (Additional file 1: Fig. S1, $r^2 > 0.80$), except for between the two replicates in TSBT_R (Additional file 1: Fig. S1, $r^2 = 0.56$). However, considering that only two samples in the TSBT_R and the $r^2 > 0.50$. Therefore, all samples were used for subsequent analysis.

Differential expression analysis in three groups

The all ginsengs were divided into three comparison groups, including YNJY_vs_JYSH, TSBT_vs_PHQH and YONE_vs_YSWD. There were 18,923, 11,009 and 1,241 DEGs in YNJY_vs_JYSH, TSBT_vs_PHQH and YONE_vs_YSWD, respectively (Fig. 1). The number of DEGs in cultivated ginseng is higher than that in wild ginseng. Furthermore, the number of DEGs in YNJY_vs_JYSH was larger than that in TSBT_vs_PHQH. In addition, the numbers of DEGs in different tissues of three groups were also different. In YNJY_vs_JYSH, the number of DEGs in roots were almost the same as that in leaves, while DEGs in stems were the least (6,392) (Fig. 1A). In TSBT_vs_PHQH, the number of DEGs in leaves was the highest (7,198), while DEGs in roots were the least (2,243) (Fig. 1B). In YONE_vs_YSWD, the number of DEGs in root was the highest (696), then DEGs in leaves were the least (181) (Fig. 1C).

Construction of Gene Co-Expression Network

A total of 39,086, 38,823 and 39,780 genes were construct WGCNA analysis for YNJY_vs_JYSH, TSBT_vs_PHQH and YONE_vs_YSWD, and he soft thresholding of three groups were set at 14, 10 and 8, respectively (Additional file 1: Fig. S2). While the scale-free topology fit index reached 0.85, indicating approximate scale-free topology (Additional file 1: Fig. S2). Finally, we identified 34, 57 and 69 distinct modules in YNJY_vs_JYSH, TSBT_vs_PHQH, YONE_vs_YSWD, respectively (Additional file 1: Fig. S3).

Co-expression modules related to growing years of ginseng

In WGCNA, the module eigengene (ME) is representative of the corresponding module's gene expression profile correlated with a defined trait (growing years). Several modules were significantly correlated with

the growing years ($p < 0.05$). In YNJY_vs_JYSH, blue, royalblue and red modules were significantly correlated with growing years, and the correlation coefficient were 0.91, 0.9 and -0.96, respectively ($p < 0.05$) (Additional file 1: Fig. S3). In blue module, the enrichment pathways included Citrate cycle (TCA cycle) (ko00020), C5-Branched dibasic acid metabolism (ko00660) and Carbon metabolism (ko01200), and the most genes in these pathways were highly expressed in YNJY (Fig. 2). In red module, genes were enriched in protein processing in endoplasmic reticulum (ko04141) (Fig. 3A). Differential clustering of genes involved in this pathway showed that most of the genes were expressed at higher levels in YNJY than in JYSH (Fig. 3B). In royalblue module, the enriched pathway was MAPK signaling pathway-plant (ko04016) (Fig. 3C), in which genes were up-regulated in JYSH (Fig. 3D).

In TSBT_vs_PHQH, three modules were significantly correlated with growing years, namely white, grey60 and skyblue3 modules ($p < 0.05$). However, only Alpha-linolenic acid metabolism (ko00592) and terpenoid backbone biosynthesis (ko00900) were enriched in grey60 module (Fig. 4A), and the expression levels of most genes in these pathways were lower in PHQH than that in TSBT (Fig. 4B, C). Nevertheless, there were no enriched pathways in the significant related modules (orangered4 and salmon4 modules) for YONE_vs_YSWD (Additional file 1: Fig. S3).

In addition, a large number of genes for ginsenoside biosynthesis were found in four modules related to ginseng growth years, such as farnesyl diphosphate synthase (FPS), diphosphomevalonate decarboxylase (MVD), phosphomevalonate kinase (PMK), UGTs, SS, SE and CYP450 were found in the blue and royalblue modules from YNJY_vs_JYSH. In the grey60 module from TSBT_vs_PHQH, we found 10 enzymes involved in the synthesis of ginsenosides, including hydroxymethylglutaryl-CoA reductase (HMGR), (E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase (ISPG), 4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase (ISPH), SS, 1-deoxy-D-xylulose-5-phosphate synthase (DXS), DS, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (ISPF), acetyl-CoA C-acetyltransferase (AACT) and Hydroxymethylglutaryl-CoA synthase (HMGS). There was only 1 enzyme involved in ginsenoside biosynthesis in the orangered4 module from YONE_vs_YSWD, which was 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase (ISPE) (Table 1).

Table 1
Genes involved in the ginsenoside
synthetic pathway in the module

Module	Gene_id	Gene family
Blue	Pg_S0304.36	FPS
	Pg_S3074.4	MVD
	Pg_S3321.6	PMK
	Pg_S6708.3	UGTs
royalblue	Pg_S0992.8	SS
	Pg_S2390.5	UGTs
	Pg_S3064.5	SE
	Pg_S3293.6	CYP450
	Pg_S4157.4	UGTs
	Pg_S4174.7	UGTs
	Pg_S4493.1	UGTs
	Pg_S4733.5	CYP450
	Pg_S6308.10	SE
	Pg_S0126.10	HMGR
	Pg_S0247.51	IspG
	Pg_S0913.16	HMGR
grey60	Pg_S1005.15	IspH
	Pg_S1678.33	SS
	Pg_S1908.21	DXS
	Pg_S3318.3	DS
	Pg_S4604.8	IspF
	Pg_S6240.3	AACT
	Pg_S6896.2	HMGS
orangered4	Pg_S2198.2	IspE

In order to find the key regulatory transcription factors (TFs) related to ginsenoside biosynthesis from these modules, we constructed a gene correlation network for each module. In the blue module, 12 TFs

were identified, B3-ARF (Pg_S0055.1), E2F (Pg_S0029.9, Pg_S1183.4, Pg_S0517.8 and Pg_S0284.10), C2C2-LSD (Pg_S1482.2 and Pg_S5995.1), NAC (Pg_S2339.2, Pg_S2698.4 and Pg_S4475.1), GARP-G2-like (Pg_S2430.1 and Pg_S4768.11), MYB (Pg_S0682.20 and Pg_S5278.7), AP2/ERF (Pg_S4268.4 and Pg_S4859.1), LOB (Pg_S1960.28 and Pg_S0146.8), HB-PHD (Pg_S5577.18), FAR1 (Pg_S7424.8) and TCP (Pg_S5350.8) were highly positively related to MVD and PMK. However, Tify (Pg_S3560.17) was highly positively correlated with FPS and UGTs (Fig. 5A, 6A). A total of 5 genes encoding 3 TFs were found in the royalblue module, including AP2/ERF (Pg_S0575.7 and Pg_S0315.1), bHLH (Pg_S1163.1 and Pg_S3713.24) and WRKY (Pg_S5167.2) were highly related to UGTs, SS, SE and CYP450 (Fig. 5B, 6B). In the grey60 module, NAC (Pg_S1059.27), MYB (Pg_S1913.1) and bHLH (Pg_S4358.2, Pg_S6447.1 and Pg_S3268.1) were highly positively correlated with IspG, HMGR, IspH, SS, DXS, DS, IspF, AACT and HMGS (Fig. 5C, 6C). In the orangered4 module, bZIP (Pg_S0602.25) was highly positively related to IspE (Fig. 6D). These results indicated that TFs affected genes for ginsenoside synthesis.

Identification of DEG related to ginsenoside biosynthesis

The expression level of genes in the ginsenoside biosynthesis pathway were further inspected among different tissues and sample. These genes were significantly various among cultivated ginseng with different growth years. In YNJY_vs_JYSH, most DEGs involved in the ginsenoside biosynthetic pathway showed higher expression levels in JYSH than in YNJY. Only genes encoded MVD was highly expressed in three tissues of YNJY. The numbers of DEGs involved in the ginsenoside biosynthesis pathway were more in roots and leaves than in stems (Fig. 7). Fourteen enzymes, including AACT, HMGS, Mevalonate Kinase (MVK), DXS, 1-deoxy-D-xylulose-5-phosphate reductoisomerase (DXR), ISPE, ISPF, ISPG, FPS, SS, SE, DS, CYP450 and UGTs were highly expressed in the roots of JYSH. In the stems, level of gene expression encoding the six enzymes (MVD, DXS, ISPH, SE, β -AS and UGTs) were significantly different, moreover, three of them (SE, β -AS and UGTs) were highly expressed in JYSH. In addition, the twelve enzymes included HMGR, DXR, ISPE, ISPF, ISPG, ISPH, FPS, SS, SE, DS, CYP450 and UGTs were highly expressed in leaves of JYSH (Fig. 7).

In TSBT_vs_PHQH, the numbers of DEGs associated with ginsenoside biosynthesis were significant more in leaves than that in roots and stems. In roots and stems, the expression level of six enzymes (ISPE, ISPH, SE, β -As, CYP450 and UGTs) were different, only SE and CYP450 showed higher expression levels in PHTS than TSBT, while the other DEGs were highly expressed in TSBT, including ISPE, ISPH, β -As and UGTs. In the leaves, sixteen enzymes (AACT, HMGS, HMGR, DXS, ISPD, ISPE, ISPF, ISPG, ISPH, FPS, SS, SE, β -As, DS, CYP450 and UGTs) were highly expressed in TSBT than PHQH (Fig. 7).

In contrast, there were fewer DEGs for ginsenoside biosynthesis in the roots, stems and leaves of wild ginseng group (YONE_vs_YSWD) compared with cultivated ginseng comparison groups. Five enzymes, HMGS, DXS, SE, CYP450 and UGTs were highly expressed in the root of YSWD, however, only ISPH was highly expressed in the stem of YONE (Fig. 7). The result indicated that the growth years had little effect on ginsenoside synthesis in wild ginseng.

Validation of DEGs by qRT-PCR

To ensure the reliability and accuracy of the results, we randomly selected fifteen DEGs of different ginseng tissue samples to verify the relative expression levels by qRT-PCR analysis. The validation results showed that the expression trends of examined genes were consistent with that derived from RNA-seq (Additional file 1: Fig. S4). The correlation between RNA-seq results and qRT-PCR experimental results was significant and the correlation coefficient was 0.864 by spearman test ($p < 0.01$).

Discussion

Expression profiles of the DEGs in different growth years of cultivated ginseng

Ginseng is a kind of perennial herbaceous plant, its morphology and physiology changes greatly during the growth process. Here we describe dynamic gene expression profiles of cultivated and wild panax ginseng with different growth years. In the WGCNA analysis of YNJY_vs_JYSH, the genes in blue module related to growth years were enriched in citrate cycle (TCA cycle) (ko00020), C5-Branched dibasic acid metabolism (ko00660) and Carbon metabolism (ko01200), and the most genes in these pathways are highly expressed in YNJY. These pathways are common primary metabolic pathways in plants. This might be that one-year-old cultivated ginseng (YNJY) was in a stage of rapid growth and development and requires more energy, so the primary metabolic pathway was more active. The morphological and physiological components of cultivated ginseng grown to six years have been basically matured, so the energy required for growth is decreased.

The genes in royalblue module related to growth years were enriched in MAPK signaling pathway – plant in YNJY_vs_JYSH. This pathway is involved in plant disease resistance and abiotic stress response [23–25]. In *Arabidopsis thaliana*, overexpression of MKK2 gene increased MPK4 and MPK6 activities, and increased plant frost resistance and salt tolerance [26]. Moreover, the most genes of this pathway were expressed a higher level in JYSH, indicating that six-years-old ginseng (JYSH) may have higher stress resistance. When cultivated ginseng increases with planting years, it faces various plant diseases and great environmental pressure [19]. Hence, the expression of genes related to stress resistance pathways is continuously increased to resist the bad environment [27].

For ginseng that has been grown for more than six years, we found the genes in the growth age-related modules were mainly related to the terpenoid backbone biosynthesis in the WGCNA analysis of TSBT_vs_PHQH. And the expression level of most genes in terpenoid backbone biosynthesis was higher in TSBT than in PHQH. Terpenoid backbone biosynthesis is a key step in the synthesis of ginsenosides for ginseng [28]. Previous studies have reported 15 years represents a turning point in the maturation of ginseng growth and the accumulation of these bioactive components in ginseng, moreover, most ginsenosides accumulated during years 5–15 and decreased after 15 years [29]. This might be related to

the low expression of terpenoid backbone biosynthesis genes in ginseng that had been grown for more than 15 years (PHQH), resulting low ginsenoside content.

Effects of growth years on the expression levels of genes related to ginsenosides synthesis in cultivated ginseng

Ginsenosides have been used in the treatment of various diseases [13]. In our study, we explored the effect of growth years on genes related to ginsenoside biosynthesis in cultivated and wild ginseng. In YNJY_vs_JYSH, most genes involved in the ginsenoside biosynthetic pathway showed higher expression levels in JYSH than in YNJY. Previous studies have suggested that the contents of ginsenosides increased with cultivation ages in cultivated ginseng [30, 31]. The high expression level of these DEGs might contribute to the accumulation of ginsenosides in six-year-old ginseng (JYSH). Moreover, the amounts of secondary metabolites produced by plants for defense purposes increases under environmental stress conditions such as nutrient deficiency [32, 33]. Currently, ginsenosides, as secondary metabolites, have been shown to play an important physiological role to protect plants from pathogens [34]. When ginseng grows to six years old, it faces greater environmental pressure. Therefore, it is more necessary to synthesize ginsenoside to cope with many biotic or abiotic stresses [35, 36].

Interestingly, we only found the genes encoding MVD was expressed higher in YNJY than in JYSH. MVD are key enzymes in the MVA pathway, and the low expression levels of MVD may be influenced by the downstream products of MVA pathway and secondary metabolites [37, 38]. Moreover, the number of DEGs involved in the ginsenoside biosynthetic pathway was more in roots and leaves than in stems for YNJY_vs_JYSH, suggesting the influence of growth years on the genes related to ginsenoside synthesis was greater in roots and leaves than in stems. Schramek et al (2014) proposed that ginsenosides were transported from leaves to roots [39]. More recently, Xue et al (2019) showed that genes related to ginsenoside synthesis pathway were expressed in living cells, but not in vascular and xylem cells of ginseng [18]. Therefore, we consider that the content of ginsenoside synthesis in roots and leaves is more than in stems, and roots and leaves are greatly affected by grown years.

For TSBT_vs_PHQH, the most DEGs related to ginsenoside biosynthesis were highly expressed in TSBT than in PHQH, such as ISPE, ISPF and UGTs et al. When the ginseng grows more than six years old, the plant becomes lignified [40]. Thus, the activity of genes in related metabolic pathways decreases, which leads to a decrease in the expression of certain genes. Furthermore, the number of DEGs related to ginsenoside biosynthesis was more in leaves than in stems and roots. The different expression profiles of genes related to ginsenoside synthesis in ginseng roots, stems and leaves suggested that there might be significant differences in ginsenosides or other triterpenoid precursors synthesized in different tissues of ginseng. This differentiation probably led to the accumulation of different types of ginsenosides in the roots, stems and leaves, thus explaining tissue-specific ginsenosides production. Kim et al (2015) indicated the main ginsenosides in root and leaf were Rb1, Rb2, Rc, Rd (PPD-type) and Re, Rg1, Rf (PPT-

type), respectively [34]. Furthermore, Kim et al (2018) also showed that the content of most ginsenosides, Rb2, Rg1, Rf, F1, F2, Rf and Rh1, were obviously different in leaves of 1, 4 and 6-years old ginsengs [41].

Effects of growth years on the expression levels of genes related to ginsenosides synthesis in wild ginseng

The germplasm resources of wild ginseng are scarce, so there are few studies on gene expression level of wild ginseng. In our study, we found the numbers of DEGs in YNJY_vs_JYSH and TSBT_vs_PHQH was more than that in YONE_vs_YSWD, suggesting gene expression levels were less affected by growth years in wild ginseng compared with cultivated ginseng. Wild ginseng grows more slowly than cultivated ginseng and can be grown in nature for hundreds of years, so it may be less affected by a few short decades. Furthermore, ginsenoside biosynthesis related genes, including HMGS, DXS and SE were more highly expressed in the roots in YSWD than in YONE. HMGS is a key catalytic enzyme in MVA pathway, and HMGS overexpression increased triterpenoids in *Ganoderma lucidum* (Ang, 2013). Ye et al (2018) suggesting the enhanced expression level of SE genes may cause increased production of triterpenoids (Ye et al., 2018). In addition, DXS is important in terpenoid biosynthesis, and DXS is the first rate-limiting enzyme of the MEP pathway (Wei et al., 2019). CYP450 and UGTs were modified in the synthesis of triterpenoid saponins, also were highly expressed in YSWD. However, there was no DEGs involved in ginsenoside biosynthesis in the leaves of YONE_vs_YSWD. The results indicated that the effect of growth years on ginsenoside biosynthesis in wild ginseng was mainly in roots, but had little effect on leaves. Wild ginseng increases the number of leaves every 3-5 years, eventually producing 4-6 compound leaves. These leaves are then all shed and grow again from the first leaf (Kim et al., 2018). The leaf state of twenty-year-old wild ginsengs was similar to that of one-year-old wild ginsengs, thus, the number of DEGs in the leaves was the least and there were few DEGs related to ginsenoside biosynthesis.

A regulated gene network for ginsenoside biosynthesis

Biosynthesis of particular compounds is influenced by a variety of factors by influencing transcription factors (TFs), which regulate the expression of relevant genes [42, 43]. In this study, we used WGCNA to screen the growth age-related modules, and found genes related to ginsenoside biosynthesis and TFs in four modules (blue, royalblue, grey60 and orangered4). In these modules, MYB (Pg_S0682.20, Pg_S5278.7, Pg_S0682.20, Pg_S5278.7 and Pg_S1913.1), NAC (Pg_S2339.2, Pg_S2698.4, Pg_S1059.27 and Pg_S4475.1), AP2/ERF (Pg_S4268.4, Pg_S4859.1, Pg_S0575.7 and Pg_S0315.1), bHLH (Pg_S1163.1, Pg_S4358.2, Pg_S6447.1, Pg_S3268.1, Pg_S3713.24, Pg_S4358.2, Pg_S6447.1 and Pg_S3268.1), bZIP (Pg_S0602.25) and WRKY (Pg_S5167.2) were highly correlated with genes related to ginsenoside biosynthesis.

MYB family is the most widely distributed and powerful transcription factor. The MYB domain typically consists of one to four incomplete repeats, each containing approximately 52 amino acid residues [44]. A

ginseng PgMYB2, was confirmed to positively regulate the expression of DS in ginseng [45]. In our study, MYB was also positively correlated with DS. In addition, NAC and AP2/ERF highly correlated with genes related to ginsenoside biosynthesis, such as HMGR, IspG, IspH, SS, DXS, DS, IspF, AACT and HMGS et al. Xu et al. (2017) identified a number of transcription factors, including NAC, that might be involved in the synthesis of triterpenoids in *Cyclocarya paliurus* (Bata1) Iljinskaja [46]. Deng et al. (2017) also found that the expression levels of DS and SS and total saponins content in PnERF1 transgenic *Panax notoginseng* cell lines were higher than those in non-transgenic cell lines [47]. Moreover, WRKY was highly related to SS, SE, CYP450 and UGTs. WRKY transcription factors mainly exist in plants and play important roles in many biological processes [48]. A PqWRKY1 transcription factor isolated by transcriptomic sequencing from *Panax quinquefolius* enhances the content of saponins by regulating the transcription of some genes related to biosynthesis of saponins [49]. There were 8 genes encoded bHLH, which were significant positive correlation with genes related to ginsenoside biosynthesis, such as UGTs, CYP450, SS, DS, IspF, AACT and SE et al. In *Panax notoginseng*, the overexpression of transcription factor PnbHLH caused the increase of saponins [50]. In YONE_vs_YSWD, only bZIP was highly positively related to IspE in orangered4 module. However, there are few studies on the effect of bzip on the synthesis of ginsenosides in *Panax* genus, which required subsequent yeast one-hybrid or two-hybrid to verify its function. These result suggested growth years might regulate the expression of ginsenoside synthesis genes by regulating TFs (MYB, NAC, AP2/ERF, bHLH, bZIP and WRKY), thereby affecting the content of ginsenosides.

Conclusions

In this study, through transcriptome analysis of wild and cultivated ginseng with different growth years, we found the DEGs were much more in cultivated ginsengs than that in wild ginseng. WGCNA analysis identified modules obviously affected by growth years in cultivated ginseng, and genes in these modules were significantly enriched in MAPK signaling pathway - plant and terpenoid backbone biosynthesis pathway. However, there was no enrichment pathway for genes in modules significantly related to growth years in wild ginseng. Furthermore, the growth years had a significant effect on the gene expression of ginsenoside synthesis in cultivated ginseng, and the degree of influence varies in different tissues. However, it had little effect on genes related to ginsenoside synthesis in wild ginseng and no effect on leaves. These results indicated that the influence of growth years on cultivated ginseng was greater than that of wild ginseng. In addition, HMGR, SS, DXS, DS, IspF, AACT and HMGS were highly correlated with MYB, NAC, AP2/ERF, bHLH, bZIP and WRKY. Growth years might regulate genes for ginsenoside synthesis by influencing these TFs, then it affected the content of ginsenosides. This study provides comprehensive transcriptional information for wild and cultivated ginseng at different growth stages, and also supplies new insights for understanding the metabolic regulation of ginsenosides.

Methods

Plant materials and preparation

One- and six-year-old cultivated ginsengs were collected from Jingyu county, Jilin province of China (marked as YNJY and JYSH, respectively), and the one- and twenty-year-old wild ginsengs were collected from Korean Autonomous County of Changbai, Jilin province (marked as YONE and YSWD, respectively). In addition, we sampled six-year-old cultivated ginsengs from Taishang town of Jilin province (marked as TSBT), meanwhile, we also collected three samples of more than fifteen-year-old cultivated ginsengs from the same location (marked as PHQH) (Additional file 1: Table S1).

All cultivated ginseng and wild ginseng samples were collected in August 2018. After cleaning with distilled water on the spot, the roots, stems and leaves of all ginseng were separately cut into small pieces quickly and packaged in aluminum foils. Then these samples are immediately frozen in the liquid nitrogen and stored at -80°C. According to sampling location, these ginsengs were divided into three comparison groups, namely YNJY_vs_JYSH, TSBT_vs_PHQH and YONE_vs_YSWD.

RNA extraction, cDNA library construction and sequencing

Total RNA was extracted from all tissues of each sample using TRIzol (Invitrogen, USA) and digested with DNase I (Waryong, China). Finally, the RNA was dissolved by adding 50 µl DEPC-treated water. After confirmed the integrity and quality, the total RNA was immediately stored at -80°C for sequencing. The RNA samples with high purity ($28S/18S \geq 1.4$) and high integrity ($RIN \geq 7.0$) were employed for cDNA library construction.

Then, library construction and sequencing were performed by the Beijing Genomics Institute (BGI, China). Briefly, Oligo(dT)-attached magnetic beads were used to isolate mRNA. Purified mRNA was broken into short fragments by mixing with fragmentation buffer at appropriate temperature. The short fragments were purified and resolved with the Elution buffer for end repair and single nucleotide A addition, and then connected with adapters. Target fragments were selected as templates for PCR amplification to construct the cDNA sequencing libraries. Each cDNA library was sequenced by an Illumina HiSeq X-ten platform with paired end (PE) reads of 150 bp. The clean data were deposited in the NCBI (National Centre for Biotechnology Information) under the accession numbers PRJNA762437.

Data filtering, mapping and gene expression profiling

To obtain high-quality clean reads, the adaptor sequences and the reads with unknown nucleotides larger than 5% or with low-quality nucleotides (≤ 10) larger than 20% were filtered out. Then the clean reads were mapped to the ginseng reference genome v1 (Ginseng Genome Database) using Bowtie2 (v2.2.5) [51], followed by gene expression levels of each sample were calculated using RSEM with default parameters, and the FPKM (Fragments Per Kilobase of transcript per Million mapped reads) value was used to quantify gene expression levels [52]. Pearson correlation coefficients (r^2) of expression levels were calculated between each pair of ginseng samples using R package (v4.2.0), and samples with $r^2 < 0.5$ across samples were removed. Differential expression analysis of each comparison group was

performed according to the method described by DESeq2 (v1.16.1) [53]. The genes with a threshold of $|\log_2(\text{fold change})| \geq 1$ and $p \leq 0.05$ were identified as differentially expressed genes (DEGs).

Weighted gene co-expression network construction

Weighted gene co-expression network analysis (WGCNA) was performed to construct co-expression networks of genes that were differentially expressed in different growth years of ginseng using the R package (v4.2.0) [54]. After background correction and quantile normalization, the all genes with top 50% variant in the variance analysis for WGCNA analysis. Firstly, the hclust function was used to conduct cluster analysis on the expression data of the samples. Average was selected as the clustering method to remove the outlier samples. Then the function pickSoftThreshold was used to filter Soft Threshold (power). The similarity between the two genes was calculated by topological overlap (TOM), and the module gene tree was obtained by hierarchical clustering. The genes with the same expression patterns were grouped into the same module by dynamicTreeCut method, and the minimum number of genes was set as 30 in the module (minModuleSize = 30). The first principal component was used to calculate module eigengenes (ME) with mergeCutHeight = 0.25, and similar modules were merged through module characteristic genes. The cor (MEs, datTraits) function was used to estimate the module-trait relationship, then the Pearson correlation coefficient between the characteristic gene of the module and the growing years. The module with a correlation coefficient $\geq |0.8|$ and $p \leq 0.05$ was regarded as significant growing years-related modules. Finally, in order to further investigate these genes in significant related module, the enrichment analysis of these genes were analyzed based on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway by clusterProfiler (v3.16.1)[55]. Cytoscape (v3.7.0) was used to visualize the most significantly correlated genes with a WGCNA edge weight > 0.25 , then we considered the top 20% of the connected genes as hub genes in the module [56]. The correlation between genes in the ginsenoside synthesis pathway and hub genes in the module was analyzed using R package (v4.2.0).

Identification of differential expression genes (DEGs) associated with ginsenoside biosynthesis

Ginsenoside is the most important medicinal component in the ginseng. Therefore, the expression levels of genes involved in the biosynthesis of ginsenoside were analyzed. Then, DEGs associated with ginsenoside biosynthesis in each comparison group was also performed by DESeq2 (v1.16.1), which was consistent with the above (Love et al., 2014).

qRT-PCR validation

Fifteen DEGs were randomly selected for qRT-PCR to validate the RNA-seq data. Total RNAs as described for RNA-seq were used for qRT-PCR. Total RNA from each sample was reverse transcribed into cDNA using PrimeScriptTM RT reagent Kit with gDNA Eraser (TaKaRa, Japan) following manufacturer's protocol.

Gene-specific primers were designed by Primer Premier 6 software [57]. The qRT-PCR reactions were performed in a final volume of 10µl using 2x SYBR Green qPCR Master Mix (Vazyme, China) on the Applied Biosystems® QuantStudio® 3 (Thermo Fisher Scientific, USA). PCR amplification was conducted under the following conditions: 95°C for 5 min, 40 cycles of 95°C for 30 s and 60°C for 30 s, and with a dissociation stage of 95°C for 15 s, 60°C for 60 s and 95°C for 15 s. Each biological replicate was technically replicated three times. The relative expression levels of the selected genes were calculated with the $2^{-\Delta\Delta CT}$ method using GAPDH as the internal reference gene [58].

Abbreviations

PPD: protopanaxadiol; PPT: protopanaxatriol; IPP: derivescisopentenyl pyrophosphate; DMAPP: dimethylallyl pyrophosphate; MVA: mevalonic acid; MEP: 2-C-methyl-D-erythritol-4-phosphate; GPS: geranyl pyrophosphate synthase; FPS: farnesyl pyrophosphate synthase; SS: squalene synthase; SE: squalene epoxidase; OSCs: oxidosqualene cyclases; DS: dammarenediol synthase; β -AS: β -Amyrin synthase; CYP450s: cytochrome P450-dependent monooxygenases; UGTs: UDP-dependent glycosyltransferases; ME: module eigengene; WGCNA: Weighted gene co-expression network analysis; FPS: farnesyl diphosphate synthase; MVD: diphosphomevalonate decarboxylase; PMK: phosphomevalonate kinase; HMGR: hydroxymethylglutaryl-CoA reductase; ISPG: (E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase; ISPH: 4-hydroxy-3-methylbut-2-en-1-yl diphosphate reductase; DXS: 1-deoxy-D-xylulose-5-phosphate synthase; ISPF: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; AACT: acetyl-CoA C-acetyltransferase; HMGS: Hydroxymethylglutaryl-CoA synthase; ISPE: 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; MVK: Mevalonate Kinase; DXR: 1-deoxy-D-xylulose-5-phosphate reductoisomerase; KEGG: Kyoto Encyclopedia of Genes and Genomes; FPKM: Fragments Per Kilobase of transcript per Million mapped reads.

Declarations

Acknowledgments

We thank reviewers and editors for their constructive comments that greatly improved our manuscript. We also thank Wei Zhang for helping to collected the samples in this study.

Author Contributions

H. W. and X. F. designed the experiments. X. F. and M. W. performed most of experiments and analyzed the data. Other authors assisted in experiments and discussed the results. X. F. and H. X. wrote the manuscript.

Funding

Availability of data and materials

The datasets generated and analyzed in the current study are available from the corresponding author on reasonable request. All data generated or analyzed during this study are included in this published article and its Supplementary information files. The clean RNA-seq data are freely available at: www.ncbi.nlm.nih.gov/bioproject/PRJNA762437.

Ethics approval and consent to participate

The wild ginsengs were collected from Korean Autonomous County of Changbai, Jilin Province in this study, and identified by Prof. Hongxing Xiao from the College of Life Sciences, Northeast Normal University. The voucher specimens were stored in the Herbarium of Northeast Normal University (NENU403274). The research conducted in this study neither required approval from an ethics committee, nor involved any human or animal subjects. No specific permits were required for the described field studies. The location is not privately-owned or protected in any way. We complied with the IUCN Policy Statement on Research Involving Species at Risk of Extinction and the Convention on the Trade in Endangered Species of Wild Fauna and Flora.

Consent for publication

Not applicable.

Conflict of Interest

The authors declare that they have no competing interests.

References

1. Yun T-K: **Panax ginseng—a non-organ-specific cancer preventive?** *The Lancet Oncology* 2001, **2**(1):49–55.
2. Kim SK, Park JH: **Trends in ginseng research in 2010.** *Journal of Ginseng Research* 2011, **35**(4):389–398.
3. Buriana K, Hristo N, Christina H, Petkov VD: **Immunomodulating Activity of Ginsenoside Rg1 from Panax Ginseng.** *Japanese Journal of Pharmacology* 1990, **54**(4):447–454.
4. Matsuda H, Samukawa KI, Kubo M: **Anty-inflammaotory activity of ginsenoside ro1.** *Planta Medica* 1990, **56**(1):19–23.

5. Mochizuki M, Yoo YC, Matsuzawa K, Sato K, Tono-Oka S, Azuma I, Saiki I, Samukawa KI: **Inhibitory Effect of Tumor Metastasis in Mice by Saponins, Ginsenoside-RB2, 20(R)- and 20(S)-Ginsenoside-Rg3, of Red Ginseng.** *Biological and Pharmaceutical Bulletin* 1995, **18**(9):1197–1202.
6. Cheng Y, Shen LH, Zhang JT: **Anti-amnestic and anti-aging effects of ginsenoside Rg1 and Rb1 and its mechanism of action.** *Acta Pharmacologica Sinica* 2005, **26**(2):143–149.
7. Zhang AH, Lei FJ, Xu YH, Zhang LX: **Allelopathic effects of ginsenosides on main soil-borne diseases of ginseng.** In: *The fourth Symposium on allelopathy of Plants in China: 2009.* 1.
8. Tan SQ, Zhang AH, Xie JY, Lei FJ, Zhang LX: **Anti-feeding effect of total ginsenoside from Panax ginseng on Heliothis dipsacea larvae** *China journal of chinese materia medica* 2013, **38**(01):37–39.
9. Chen W, Balan P, Popovich DG: **Chapter 6 - Comparison of the ginsenoside composition of Asian ginseng (Panax ginseng) and American ginseng (Panax quinquefolius L.) and their transformation pathways.** In: *Studies in Natural Products Chemistry.* Edited by Atta ur R, vol. 63: Elsevier; 2019: 161-195.
10. Chen CF, Chiou WF, Zhang JT: **Comparison of the pharmacological effects of Panax ginseng and Panax quinquefolium.** *Acta Pharmacologica Sinica* 2008, **29**(9):1103–1108.
11. Kim S-J, Murthy HN, Hahn E-J, Lee HL, Paek K-Y: **Parameters affecting the extraction of ginsenosides from the adventitious roots of ginseng (Panax ginseng C.A. Meyer).** *Separation and Purification Technology* 2007, **56**(3):401–406.
12. Thimmappa R, Geisler K, Louveau T, O'Maille P, Osbourn A: **Triterpene biosynthesis in plants.** In: *Annual Review of Plant Biology.* vol. 65; 2014: 225-257.
13. Jayakodi M, Choi BS, Lee SC, Kim NH, Park JY, Jang W, Lakshmanan M, Mohan SVG, Lee DY, Yang TJ: **Ginseng Genome Database: An open-access platform for genomics of Panax ginseng.** *BMC Plant Biology* 2018, **18**(1).
14. Liu J, Liu Y, Wang Y, Abozeid A, Zu YG, Zhang XN, Tang ZH, Berger S: **GC-MS metabolomic analysis to reveal the metabolites and biological pathways involved in the developmental stages and tissue response of panax ginseng.** *Molecules* 2017, **22**(3).
15. Shi W, Wang Y, Li J, Zhang H, Ding L: **Investigation of ginsenosides in different parts and ages of Panax ginseng.** *Food Chemistry* 2007, **102**(3):664–668.
16. Zhang YC, Li G, Jiang C, Yang B, Yang HJ, Xu HY, Huang LQ: **Tissue-specific distribution of ginsenosides in different aged ginseng and antioxidant activity of ginseng leaf.** *Molecules* 2014, **19**(11):17381–17399.
17. Jayakodi M, Lee SC, Lee YS, Park HS, Kim NH, Jang W, Lee HO, Joh HJ, Yang TJ: **Comprehensive analysis of Panax ginseng root transcriptomes.** *BMC plant biology* 2015, **15**:138.
18. Xue L, He Z, Bi X, Xu W, Wei T, Wu S, Hu S: **Transcriptomic profiling reveals MEP pathway contributing to ginsenoside biosynthesis in Panax ginseng.** *BMC Genomics* 2019, **20**(1).
19. Fan H, Li K, Yao F, Sun L, Liu Y: **Comparative transcriptome analyses on terpenoids metabolism in field- and mountain-cultivated ginseng roots.** *BMC Plant Biology* 2019, **19**(1).

20. Zhen G, Zhang L, Du YN, Yu RB, Liu XM, Cao FR, Chang Q, Deng XW, Xia M, He H: **De novo assembly and comparative analysis of root transcriptomes from different varieties of *Panax ginseng* C. A. Meyer grown in different environments.** *Science China Life Sciences* 2015, **58**(11):1099–1110.
21. Yang BW, Hahm YT: **Transcriptome analysis using de novo RNA-seq to compare ginseng roots cultivated in different environments.** *Plant Growth Regulation* 2018, **84**(1):149–157.
22. Xu HP: **A brief discussion on the classification and identification of wild ginseng.** *Chinese rural medicine* 2021, **28**(01):13–14.
23. Ichimura K, Shinozaki K, Tena G, Sheen J, Henry Y, Champion A, Kreis M, Zhang S, Hirt H, Wilson C *et al*: **Mitogen-activated protein kinase cascades in plants: A new nomenclature.** *Trends in Plant Science* 2002, **7**(7):301–308.
24. Mao G, Meng X, Liu Y, Zheng Z, Chen Z, Zhang S: **Phosphorylation of a WRKY Transcription Factor by Two Pathogen-Responsive MAPKs Drives Phytoalexin Biosynthesis in Arabidopsis.** *The Plant cell* 2011, **23**:1639–1653.
25. Kim SH, Woo DH, Kim JM, Lee SY, Chung WS, Moon YH: **Arabidopsis MKK4 mediates osmotic-stress response via its regulation of MPK3 activity.** *Biochemical Biophysical Research Communications* 2011, **412**(1):150–154.
26. Teige M, Scheikl E, Eulgem T, Doczi R, Ichimura K, Shinozaki K, Dangel J, Hirt H: **The MKK2 Pathway Mediates Cold and Salt Stress Signaling in Arabidopsis.** *Molecular cell* 2004, **15**:141–152.
27. Ying YX, Ding WL, Li Y: **Characterization of soil bacterial communities in rhizospheric and nonrhizospheric soil of *panax ginseng*.** *Biochemical Genetics* 2012, **50**(11-12):848–859.
28. Wang K, Jiang S, Sun C, Lin Y, Yin R, Wang Y, Zhang M: **The Spatial and Temporal Transcriptomic Landscapes of Ginseng, *Panax ginseng* C. A. Meyer.** *Scientific Reports* 2015, **5**(1):18283.
29. Guo N, Yang Y, Yang X, Guan Y, Yang J, Quan J, Yan H, Hou W, Zhang G: **Growth age of mountain cultivated ginseng affects its chemical composition.** *Industrial Crops and Products* 2021, **167**.
30. Xiu Y, Li X, Sun X, Xiao D, Miao R, Zhao H, Liu S: **Simultaneous determination and difference evaluation of 14 ginsenosides in *Panax ginseng* roots cultivated in different areas and ages by high-performance liquid chromatography coupled with triple quadrupole mass spectrometer in the multiple reaction–monitoring mode combined with multivariate statistical analysis.** *Journal of Ginseng Research* 2019, **43**(4):508–516.
31. Liu Z, Wang CZ, Zhu XY, Wan JY, Zhang J, Li W, Ruan CC, Yuan CS: **Dynamic changes in neutral and acidic ginsenosides with different cultivation ages and harvest seasons: Identification of chemical characteristics for *panax ginseng* quality control.** *Molecules* 2017, **22**(5).
32. Coley PD, Bryant JP, Chapin Iii FS: **Resource availability and plant antiherbivore defense.** *Science* 1985, **230**(4728):895–899.
33. Pollastrini M, Di Stefano V, Ferretti M, Agati G, Grifoni D, Zipoli G, Orlandini S, Bussotti F: **Influence of different light intensity regimes on leaf features of *Vitis vinifera* L. in ultraviolet radiation filtered condition.** *Environmental and Experimental Botany* 2011, **73**(1):108–115.

34. Kim YJ, Zhang D, Yang DC: **Biosynthesis and biotechnological production of ginsenosides.** *Biotechnology Advances* 2015, **33**(6):717–735.
35. Oh JY, Kim YJ, Jang MG, Joo SC, Kwon WS, Kim SY, Jung SK, Yang DC: **Investigation of ginsenosides in different tissues after elicitor treatment in *Panax ginseng*.** *Journal of Ginseng Research* 2014, **38**(4):270–277.
36. 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 physiology* 2012, **53**(9):1535–1545.
37. Mazein A, Watterson S, Hsieh WY, Griffiths WJ, Ghazal P: **A comprehensive machine-readable view of the mammalian cholesterol biosynthesis pathway.** *Biochemical Pharmacology* 2013, **86**(1):56–66.
38. Liu ZC, Sun TT, Wang SX, Ma YS, Wang XT, Sun J, Zou L: **The cloning and expression analysis of mevalonate pyrophosphate decarboxylase gene c DNA sequence from *Sanguangporus baumii*.** *Journal of Nanjing Forestry University* 2020, **44**(04):79–85.
39. Schramek N, Huber C, Schmidt S, Dvorski SE, Kmispel N, Ostrozhenkova E: **Biosynthesis of ginsenosides in field-grown *Panax ginseng*.** *JSM Biotechnol Bioeng* 2014, **2**.
40. Wu W, Sun L, Zhang Z, Guo Y, Liu S: **Profiling and multivariate statistical analysis of *Panax ginseng* based on ultra-high-performance liquid chromatography coupled with quadrupole-time-of-flight mass spectrometry.** *Journal of Pharmaceutical and Biomedical Analysis* 2015, **107**:141–150.
41. Kim YJ, Joo SC, Shi J, Hu C, Quan S, Hu J, Sukweenadhi J, Mohanan P, Yang DC, Zhang D: **Metabolic dynamics and physiological adaptation of *Panax ginseng* during development.** *Plant Cell Reports* 2018, **37**(3):393–410.
42. Tai Y, Liu C, Yu S, Yang H, Sun J, Guo C, Huang B, Liu Z, Yuan Y, Xia E *et al*: **Gene co-expression network analysis reveals coordinated regulation of three characteristic secondary biosynthetic pathways in tea plant (*Camellia sinensis*).** *BMC Genomics* 2018, **19**(1).
43. Guo Y, Gao C, Wang M, Fu FF, El-Kassaby YA, Wang T, Wang G: **Metabolome and transcriptome analyses reveal flavonoids biosynthesis differences in *Ginkgo biloba* associated with environmental conditions.** *Industrial Crops and Products* 2020, **158**.
44. Dubos C, Stracke R, Grotewold E, Weisshaar B, Martin C, Lepiniec L: **MYB transcription factors in *Arabidopsis*.** *Trends in Plant Science* 2010, **15**(10):573–581.
45. Liu T, Luo T, Guo X, Zou X, Zhou D, Afrin S, Li G, Zhang Y, Zhang R, Luo Z: **PgMYB2, a MeJA-responsive transcription factor, positively regulates the dammarenediol synthase gene expression in *Panax Ginseng*.** *International Journal of Molecular Sciences* 2019, **20**(9).
46. Xu X, X.: **Transcriptome sequencing and differentially expressed triterpenoid synthesis gene analysis and cloning of *Cyclocarya paliurus* suspension cells induced by *Aspergillus niger* elicitor.** Jiangxi Agricultural University; 2017.
47. Deng B, Huang Z, Ge F, Liu D, Lu R, Chen C: **An AP2/ERF Family Transcription Factor PnERF1 Raised the Biosynthesis of Saponins in *Panax notoginseng*.** *Journal of Plant Growth Regulation* 2017, **36**(3):691–701.

48. Rushton PJ, Somssich IE, Ringler P, Shen QJ: **WRKY transcription factors**. *Trends in Plant Science* 2010, **15**(5):247–258.
49. Sun Y, Niu Y, Xu J, Li Y, Luo H, Zhu Y, Liu M, Wu Q, Song J, Sun C *et al*: **Discovery of WRKY transcription factors through transcriptome analysis and characterization of a novel methyl jasmonate-inducible PqWRKY1 gene from Panax quinquefolius**. *Plant Cell, Tissue and Organ Culture (PCTOC)* 2013, **114**(2):269–277.
50. Zhang X, Ge F, Deng B, Shah T, Huang Z, Liu D, Chen C: **Molecular Cloning and Characterization of PnbHLH1 Transcription Factor in Panax notoginseng**. 2017, **22**(8):1268.
51. Langmead B, Salzberg SL: **Fast gapped-read alignment with Bowtie 2**. *Nature Methods* 2012, **9**(4):357–359.
52. Li B, Dewey CN: **RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome**. *BMC Bioinformatics* 2011, **12**(1):323.
53. Love MI, Huber W, Anders S: **Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2**. *Genome Biology* 2014, **15**(12).
54. Langfelder P, Horvath S: **WGCNA: An R package for weighted correlation network analysis**. *BMC Bioinformatics* 2008, **9**.
55. Yu G, Wang LG, Han Y, He QY: **ClusterProfiler: An R package for comparing biological themes among gene clusters**. *OMICS A Journal of Integrative Biology* 2012, **16**(5):284–287.
56. Shannon, Paul, Markel, Andrew, Ozier, Owen, Baliga, Nitin, S., Research WJG: **Cytoscape: A Software Environment for Integrated Models of Biomolecular Interaction Networks**. 2003, **13**(11):2498–2504.
57. Singh VK, Mangalam AK, Dwivedi S, Naik S: **Primer premier: Program for design of degenerate primers from a protein sequence**. *BioTechniques* 1998, **24**(2):318–319.
58. Livak KJ, Schmittgen TD: **Analysis of relative gene expression data using real-time quantitative PCR and the 2- $\Delta\Delta$ CT method**. *Methods* 2001, **25**(4):402–408.

Figures

Figure 1

Venn diagram of differentially expressed genes (DEGs) in different compared groups. The overlapping portions of the different circles represent the number of DEGs common to these groups. (A) YNJY_vs_JYSH; (B) TSBT_vs_PHQH; (C) YONE_vs_YSWD. R, root; L, leaf; S, stem.

Figure 2

Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis and expression profile associated with the blue module in YNJY_vs_JYSH group. (A), KEGG pathway enrichment of blue module; (B), expression of genes related to Citrate cycle (TCA cycle); (C), expression of genes related to C5-Branched dibasic acid metabolism; (D), expression of genes related to Carbon metabolism. R, root; L, leaf; S, stem. Various color blocks represent FPKM (Fragments Per Kilobase of transcript per Million mapped reads) normalized expression values. Red represents high gene expression, and blue represents low gene expression.

Figure 3

Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis and expression profile associated with the red module and royalblue module in YNJY_vs_JYSH group. (A), KEGG pathway enrichment of red module; (B), expression of genes related to Protein processing in endoplasmic reticulum; (C), KEGG pathway enrichment of royalblue module; (D), expression of genes related to MAPK signaling pathway-plant. R, root; L, leaf; S, stem. Various color blocks represent FPKM (Fragments Per Kilobase of transcript per Million mapped reads) normalized expression values. Red represents high gene expression, and blue represents low gene expression.

Figure 4

Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment and differential clustering analysis of the grey60 module in TSBT_vs_PHQH group. (A), KEGG pathway enrichment of grey60 module; (B), expression of genes related to alpha-Linolenic acid; (C), expression of genes related to terpenoid backbone biosynthesis. R, root; L, leaf; S, stem. Various color blocks represent FPKM (Fragments Per Kilobase of transcript per Million mapped reads) normalized expression values. Red represents high gene expression, and blue represents low gene expression.

Figure 5

Gene co-expression subnetwork in the (A) blue, (B) royalblue, (C) grey60 modules.

Figure 6

Correlation analysis of transcription factors (TFs) and polysaccharide synthesis genes in the (A) blue, (B) royalblue, (C) grey60 and (D) orangered4 modules. Network was reconstructed by edge weight cutoff = 0.25 and visualized by Cytoscape.

Figure 7

Heatmaps of DEG related to ginsenoside biosynthesis in YNJY_vs_JYSH, TSBT_vs_PHQH and YONE_vs_YSWD groups. * represent DEGs in each comparison group. Each row corresponds to a gene related to ginsenoside biosynthesis. R, root; L, leaf; S, stem. Various color blocks represent FPKM (Fragments Per Kilobase of transcript per Million mapped reads) normalized expression values. Red represents high gene expression, and blue represents low gene expression. DMAPP: dimethylallyl pyrophosphate; MVA: mevalonic acid; MEP: 2-C-methyl-D-erythritol-4-phosphate; GPS: geranyl pyrophosphate synthase; FPS: farnesyl pyrophosphate synthase; SS: squalene synthase; SE: squalene epoxidase; OSCs: oxidosqualene cyclases; DS: dammarenediol synthase; β -AS: β -Amyrin synthase; CYP450s: cytochrome P450-dependent monooxygenases; UGTs: UDP-dependent glycosyltransferases; ME: module eigengene; WGCNA: Weighted gene co-expression network analysis; FPS: farnesyl diphosphate synthase; MVD: diphosphomevalonate decarboxylase; PMK: phosphomevalonate kinase; HMGR: hydroxymethylglutaryl-CoA reductase; ISPG: (E)-4-hydroxy-3-methylbut-2-enyl-diphosphate synthase; ISPH: 4-hydroxy-3-methylbut-2-en-1-yl

diphosphate reductase; DXS: 1-deoxy-D-xylulose-5-phosphate synthase; ISPF: 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; AACT: acetyl-CoA C-acetyltransferase; HMGS: Hydroxymethylglutaryl-CoA synthase; ISPE: 4-diphosphocytidyl-2-C-methyl-D-erythritol kinase; MVK: Mevalonate Kinase; DXR: 1-deoxy-D-xylulose-5-phosphate reductoisomerase;

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Additionalfile1.docx](#)