Shared differentially expressed genes of Chinese fir (*Cunninghamia lanceolata* (Lamb.) Hook) seedlings under low nitrogen and phosphorus stress

Jian-Hui Li [1], Gui-Fang Ma [1], Ding-Wei Luo [1], Zai-Kang Tong [1], Jin-Liang Xu [2], Yong-Quan Lu[1]
[1]State Key Laboratory of Subtropical Silviculture, Zhejiang A & F University, Lin'an, Zhejiang, China
[2]Kaihua Woodland ,Kaihuan, Quzhou,Zhejiang province,China
Corresponding Author:
Yong-Quan Lu[1]
Email address: luyongquan@zafu.edu.cn

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2916v1 | CC BY 4.0 Open Access | rec: 8 Apr 2017, publ: 8 Apr 2017

Abstract:

Chinese fir (*Cunninghamia lanceolata* (Lamb.) Hook) is an excellent fast-growing timber species and has significant value in the forestry industry. In order to increase the nitrogen and phosphorus absorption and utilization in Chinese fir, shared differentially expressed genes under low nitrogen and phosphorus stress were screened in this study.

Seedling of Chinese fir clone X6 was cultivated in aeroponic system with 3 treatments, namely, low nitrogen (LN), low phosphorus (LP) and a control check (with nitrogen and phosphorus sufficient, CK). After 4 months of treatment, the roots from the LN, LP and CK groups were collected and transcriptome sequencing was done by LC Sciences (USA) using an Illumina platform.

When comparing the LN stress group with the CK group, 977 SDGEs were detected, 264 of which had KEGG annotations; 931 SDGEs were detected when comparing the LP stress group with the CK group, of which 189 had KEGG annotations; 297 SDGEs were detected in both the LN stress and LP stress groups, 78 of which had KEGG annotations representing 98 metabolic pathways. Among the 78 selected SDEGs that were differentially expressed under both LN and LP stress conditions, Twenty-one SDEGs were selected based on the metabolic pathways that nitrogen and phosphorus are involved in, these genes are *PNR*, *PSBA*, *EGLC*, *GLC*, *END*, *BGLU*, *AMY*, *A1E*, *PAL*, *GOGAT*, *NIR1*, *NIR2*, *C4M*, *PAL*, *PRDX6*, *POX*, *CCR*, *CCoAOMT*, *FDH*, *CHS* and *ANR*. These genes can potentially be used in breeding to improve both nitrogen and phosphorus utilization efficiency in Chinese fir.

Key words: Shared differentially expressed genes, Chinese fir, low nitrogen stress, low phosphorus stress, nitrogen and phosphorus utilization efficiency

INTRODUCTION

Chinese fir (*Cunninghamia lanceolata* (Lamb.) Hook) is a gymnosperm in the cypress family Taxodiaceae. It is an excellent fast-growing timber species and has significant value in the forestry industry (Ma et al. 2016). It has been widely cultivated in southern China and accounts for 20–30% of the total commercial timber production in present-day China(Orwa et al. 2009).

Nitrogen and phosphorus are essential nutrients for plant growth and development (Bhattacharyya et al. 2008; Gonzalez-Dugo et al. 2010; Moseley & Grossman 2009). With increasing application rates of nitrogen and phosphorus fertilizers in agriculture, substantial nitrogen fertilizers are released into environment, causing damage not only to ecosystems, but also compromising food and drinking water safety (Sekhon 1995; Spiertz 2010). Released phosphorus fertilizers are adsorbed, fixed and mineralized in the soil, and consequently, plant phosphorus absorption and utilization efficiency are reduced dramatically (Vance et al. 2003). Therefore, increasing the nitrogen and phosphorus absorption and utilization efficiency of plants is of great significance.

The selection of a high nitrogen and phosphorus efficient genotype is a necessary means of improving plant nitrogen and phosphorus absorption and utilization efficiency. Many genotypes with high nitrogen and phosphorus efficiency have been selected (HUANG et al. 2015; Silva et al. 2016; Singh et al. 2015; Xiangqing et al. 2002), and several studies have revealed that regulating key genes involved in nitrogen and phosphorus metabolic pathways can enhance nitrogen and phosphorus absorption and utilization efficiency (Coelho et al. 2010; Hayes et al. 2000; Hu et al. 2015; Lapis-Gaza et al. 2014; Lu et al. 2011; Sharma et al. 2016). However, the molecular mechanisms of high nitrogen and phosphorus absorption and utilization efficiency in Chinese fir are still unknown.

Most current studies focus on ways of improving the absorption and utilization efficiency of

single nutritive elements in plants. However, studies have also shown that there are interactions between nitrogen and phosphorus absorption and utilization pathways (Ierna et al. 2012; Schlüter et al. 2013). Plants need substantial nitrogen and phosphorus to fulfill their biosynthetic and metabolic processes (Ågren et al. 2012; Rivas-Ubach et al. 2012), many of which need both nitrogen and phosphorous, indicating that some genes are involved in both nitrogen and phosphorus metabolism. Selecting genes that are involved in both nitrogen and phosphorous metabolism is the first step towards unraveling these mechanisms. Subsequently, these selected genes can potentially be used to improve both nitrogen and phosphorus absorption efficiency and utilization efficiency. In this study, we used next generation sequencing technology (RNA-seq) to sequence and compare the transcriptomes of Chinese fir under a control check condition, low nitrogen stress condition, and low phosphorus stress condition. Genes that were differentially expressed in the root tissues under both LN and LP stress conditions were selected. This study will inform the breeding of Chinese firs with high nitrogen and phosphorus efficiency.

MATERIALS AND METHODS

In October 2013, stem cuttings of 1 highly adaptive Chinese fir clone, named X6, was collected from a garden in the Kaihua Woodland in Zhejiang province and propagated in the Germplasm Center of the Zhejiang Agriculture and Forestry University. In the spring of 2014, disease-free seedlings with complete root systems that appeared to be similar in structure and displayed uniform growth were selected as basic experimental materials.

An aeroponic system was designed for our research, and the nutrition solutions are shown in Table 1. Hoagland's solution was used for the major elements except for nitrogen and phosphorus, while Amon's solution was used for trace elements. According to the concentration of nitrogen and phosphorus, the test set a total of 3 treatments, namely, low nitrogen (LN), low

phosphorus (LP) and a control check (with nitrogen and phosphorus sufficient, CK). HCl and NaOH were used to adjust the pH value of the nutrient solutions to 5.8; KNO_3 and $Ca(NO_3)_2$ were used as nitrogen sources; KH_2PO_4 was used as a phosphorus source; and HCl and $CaCl_2$ were used to maintain a stable ion concentration when the ion concentration changed due to changes in NO_3^- or $(PO_4)^{3-}$. Each group contained 5 plants, and each experiment was triplicated. The selected root cuttings from the 4 clones were planted in an aeroponic culture apparatus, of which the root chamber volume is 50cm ×50cm×50cm, and the space of seedlings is 10 cm × 10 cm each (Figure 1).

After 4 months of treatment, the roots from the LN, LP and CK groups were collected and snap frozen in liquid nitrogen, and then stored at −80 ℃ after being kept in liquid nitrogen for 3 hours. CTAB+Trizol reagent (Invitrogen, CA, USA) was used to extract RNA. Transcriptome sequencing was done by LC Sciences (USA) using an Illumina platform.

The unigene obtained by splicing was compared with the protein sequences in 5 open databases (NR, Swiss-Prot, KEGG, KOG and Pfam) and evaluated using a BLAST (Blast Local Alignment Search Tool) algorithm. The threshold was set as evalue $\leq 1e^{-10}$ to obtain similar sequences, and then the functions were annotated.

In this study, the results of the 3 transcriptions were divided into 2 groups: one for the low-nitrogen group (LN treatment and CK), and the other for the low-phosphorus group (LP treatment and CK). RPKM (reads per kilobase of exon model per million mapped reads) was used to measure the abundance of gene expression(Mortazavi et al. 2008); the expression of unigenes in each group was calculated by the log2fold_change (logarithm of the expression ratio of LN or LP to CK using base 2), and the absolute value of log2fold_change≥1 and false discovery rate (FDR) <0.001 were considered the thresholds for significantly differentially expressed genes (SDEG) between the treatment and the CK in the 2 groups(Jiang et al. 2012).

Transcriptome sequencing (mRNA-seq)

After combining equal amounts of RNA from the 3 samples, the sequencing was performed on an Illumina HiSeq 2000/2500 platform. After removing the adapters, low-quality reads and reads in which the N-containing ratio exceeding 5%, a total of 5.20 G bases from 41,615,988 reads were obtained for the CK group with an Q20 value of 91.72% and a GC content value of 42.37%. Likewise, a total of 4.31 G bases from 34,498,198 reads were obtained for the LN stress group with an Q20 value of 91.91% and a GC content value of 41.92%; a total of 6.99 G bases from 55,955,276 reads were obtained for the LP stress group with an Q20 value of 95.15% and a GC content value of 42.42% (Table 2). These high quality reads were assembled using the Trinity package, and a total of 33,712,721 bases from 33,070 unigenes were obtained, with an average assembly length of 1,019 bp, an N50 value of 1,762 nt and 54,188 transcripts in total (Table 3). The assembled unigenes had a wide length distribution, approximately 44.81% (14,818) of the unigenes were between 200 and 500 bp, and 14.54% (4,802) of the unigenes were more than 2,000 bp. We detected 54,188 transcripts from the obtained unigenes, with a total length of 60,691,480 bp, ranging from 201 to 12,214 bp. The average length of all obtained transcripts was 1,120 bp with an N50 value of 1,809 bp. Approximately 58.87% (31,889) of the total transcripts were between 200 and 1,000 bp, and 16.52% (8,954) of the total transcripts were more than 2,000 bp. According to the length distribution of the unigenes and transcripts obtained, the number of unigenes and transcripts decreased as the length of unigenes and transcripts increased between 200 and 1,900 bp (Figure 2 and 3).

Unigene function annotation

To predict the putative functions of the assembled unigenes, all unigenes were compared with 3 public protein databases, namely NCBI non-redundant protein (NR), Swiss-Prot and Pfam

using the BLAST algorithm with an E-value threshold of $1e^{-10}$. Among all 33,070 unigenes, 60.16% (19,896), 40.11% (13,265) and 48.65% were identified and showed significant similarity to known proteins in NR, Swiss-Prot and Pfam database, respectively (Table 4). However, 39.84% of assembled unigenes did not show similarity to any known proteins due to the lack of *Cunninghamia lanceolata* genome and EST data. Comparing all assembled unigenes against the NR plant protein database, we found that for 28.2% of the unigenes the most similar proteins sequences were from *Picea sitchensis*, followed by *Amborella trichopoda* (13.8%), *Phoenix dactylifera* (5.9%), *Vitis vinifera* (5.6%), *Physcomitrella patens* (3.1%) and *Theobroma cacao* (2.5%). The remaining 41% of the unigenes were most similar to protein sequences from other plant species (Figure 4). These results indicate that the Illumina sequencing produced a substantial fraction of the *Cunninghamia lanceolata* genes.

Gene ontology analysis

To functionally categorize the unigenes obtained, Gene Ontology analysis was conducted. We assigned 11,772 unigenes to GO classes with 60,288 terms. The assigned terms were summarized into 3 main GO categories (biological process, molecular function and cellular component) and 4,591 sub-categories. Molecular function comprised 23,659 (39.24%) GO annotations followed by biological process (19,664, 32.62%) and cellular component (16,965, 28.14%). The information of the top 10 sub-categories from each category are shown in Figure 5. In the molecular function category, 10,608 unigenes were assigned to 1,669 sub-categories, with ATP binding (2,889) representing the most. In the biological process category, 8,348 unigenes were assigned to 2,349 sub-categories, with transcription and DNA-dependent (686) representing the most. In the cellular component category, 9,750 unigenes were assigned to 573 sub-categories with integral to membrane (2,852) representing the most. In the molecular function category,

10,608 unigenes were assigned to 1,669 sub-categories, with ATP binding (2,889) representing the most.

To further predict gene function and evaluate the completeness of the transcriptome library, all assembled unigenes were searched against the euKaryotic Ortholog Groups (KOG) database. Overall, 16,497 unigenes (49.83%) were assigned with KOG classification (Table 4). Altogether, 12,295 KOG functional annotations were produced and the SOG-annotated putative proteins were classified into 25 functional categories (Figure 6). Among all the categories, the general function prediction only represented the largest group with 2,024 unigenes, followed by posttranslational modification, protein turnover, chaperones (1289), signal transduction mechanisms (1,065), carbohydrate transport and metabolism (680). The smallest group was cell motility with only 4 unigenes (Figure 4).

KEGG pathway mapping

To understand the biological pathways in which the assembled unigenes are involved, the unigenes were compared against the KEGG database. The results showed that 9,424 unigenes (28.50%) had significant matches and were assigned with KEGG annotations. At the top level, the unigenes were divided into 5 categories, with metabolism (5,782) represented the largest category, followed by Genetic Information Processing (1,287), environmental information processing (807), cellular processes (1464) and organismal systems (1,750). At the second level, the unigenes were further divided into 30 sub-categories, including lipid metabolism, metabolism of cofactors and vitamins, energy metabolism, nucleotide metabolism, biosynthesis of other secondary metabolites and amino acid metabolism. Among all the sub-categories, carbohydrate metabolism represented the largest group with 844 unigenes, followed by amino acid metabolism with 662 unigenes. The signaling molecules and interaction represented the smallest group with

only 10 unigenes (Figure 7).

Identification of differentially expressed genes

When comparing the LN stress group with the CK group, 977 SDGEs were detected, 264 of which had KEGG annotations; 931 SDGEs were detected when comparing the LP stress group with the CK group, of which 189 had KEGG annotations; 297 SDGEs were detected in both the LN stress and LP stress groups, 78 of which had KEGG annotations representing 98 metabolic pathways. Apart from the SDGEs that were detected in both the LN stress and LP stress groups, 680 and 634 SDGEs were specifically detected in the LN stress and LP stress group, with 174 and 130 KEGG annotations representing 160 and 138 metabolic pathways, respectively.

SGEGs that were specifically detected in the LN stress group could be classified into different metabolism pathways, including the phenylpropanoid biosynthesis pathway with 22 SDEGs; phenylalanine metabolism pathway with 17 SDEGs; starch and sucrose metabolism pathway with 15 SDEGs; methane metabolism pathway with 14 SDEGs; and cysteine and methionine metabolism pathway with 12 SDEGs (Figure 8).

SGEGs that were specifically detected in the LP stress group could be classified into different metabolism pathways, including the starch and sucrose metabolism pathway with 17 SDEGs; phenylpropanoid biosynthesis pathway with 14 SDEGs; phenylalanine metabolism pathway with 12 SDEGs; methane metabolism pathway with 8 SDEGs and nitrogen metabolism pathway with 8 SDEGs (Figure 9).

Among the 78 selected SDEGs (with KEGG annotation) that were differentially expressed under both LN and LP stress conditions, 6 of them showed different expression patterns (No. 1–5 were upregulated under LP stress and downregulated under LN stress; No. 6 was upregulated under LN conditions and downregulated under LP conditions), and 72 of them showed similar

expression patterns under different conditions. Sorting the 72 SDGEs into descending order based on expression value differences under LN stress, 34 (No. 7–40) of them were upregulated and 38 (No. 41–78) were downregulated under both conditions. The top 7 (No. 7–13) ranking SDGEs which showed infinite upregulation encode proteins for H+-transporting ATPase, polyneuridine-aldehyde esterase, L-ascorbate oxidase, glycerol-3-phosphate dehydrogenase (NAD+), glutathione S-transferase, nucleolin and stearoyl-CoA desaturase (delta-9 desaturase). The following 6 SDGEs (No. 14–19) which showed significant upregulation (with an increase of log2fold_change more than 5) under LN and LP stress conditions encode proteins for cytochrome P450, family 1, subfamily A, polypeptide 1, H+-transporting ATPase, phospholipase A2, cytochrome P450, family 26, subfamily A, aminocyclopropanecarboxylate oxidase and tubulin alpha. Among the 38 SDEGs that were downregulated, the last 8 SDGEs showed a decrease in the log2fold_change value (more than 5) under LN stress conditions; the value of the log2fold_change value also decreased under LP stress conditions, but generally lower than that under LN stress conditions. The 8 SDGEs encode proteins for peroxiredoxin 6, 1-Cys peroxiredoxin, aminocyclopropanecarboxylate oxidase, glucan endo-1,3-beta-D-glucosidase, aminocyclopropanecarboxylate oxidase, peroxidase, beta-glucosidase, flavonol synthase and N1-acetylpolyamine oxidase.

Identification of shared differentially expressed genes

To better identify key genes that function under both LN and LP stress, we further analyzed the abovementioned shared SDEGs. Twenty-one SDEGs were further selected based on the metabolic pathways that nitrogen and phosphorus are involved in.

Photosynthesis, starch and sucrose metabolism, and glycolysis/ gluconeogenesis are major metabolic pathways in which nitrogen and phosphorus are involved. Two of the detected genes

were involved in photosynthesis, including *PNR* and *PSBA*; 5 genes were involved in starch and sucrose metabolism, inducing *EGLC*, *GLC*, *END*, *BGLU* and *AMY*; and only 1 detected gene was involved in glycolysis/gluconeogenesis, namely *A1E*.

Nitrogen metabolism is an important biological pathway in which nitrogen is involved. In this pathway, 4 genes were detected, including *PAL*, *GOGAT*, *NIR1* and *NIR2*.

Phosphorus plays important roles in phenylpropanoid biosynthesis, phenylalanine metabolism and methane metabolism. Moreover, there are also interactions between these 3 pathways. Seven detected genes were related with phenylpropanoid biosynthesis, including *C4M*, *PAL*, *PRDX6*, *POX*, *CCR*, *CCoAOMT* and *BGLU*. Meanwhile, *C4M*, *PAL*, *PRDX6* and *CCoAOMT* were involved in phenylalanine metabolism; *PRDX6* and *POX* were involved in methane metabolism; and *PAL* and *BGLU* were also involved in nitrogen and starch metabolism.

Flavonoids are important secondary metabolites in plants and enhance the adaptation capacity of plants to complex changing environments. In our study, 5 genes were detected in relation to flavonoid biosynthesis, including *C4M*, *FLS*, *CHS*, *CCoAOMT* and *ANR*. The *C4M* gene was also involved in phenylpropanoid and phenylalanine metabolism.

DISCUSSION

Chinese fir is an important timber wood experiencing rapid increasing demand, and therefore fast growing trees are needed to meet the production requirements of Chinese fir. Nitrogen and phosphate fertilizer is generally applied to satisfy the growth demands of Chinese fir. Nitrogen and phosphorus metabolism constitute one of the basic physiological processes of plants. Previous research on Chinese fir has focused mostly on the screening of nitrogen or phosphorus efficient genotypes, while few have reported on the genes related to nitrogen and phosphorus metabolism.

In this study, the transcriptome of the root systems of Chinese fir seedlings were studied under LN stress, LP stress and CK, respectively, using the Illumina platform. Seventy-eight SDEGs that were differentially expressed under both LN and LP stress were obtained, and twenty-one shared differentially expressed genes were obtained. These genes can potentially be used in breeding to improve both nitrogen and phosphorus utilization efficiency in Chinese fir.

Among the selected SDEGs, PNR and PSBA were identified as important genes involved in photosynthesis. Photosynthesis is an important source of energy for many organisms. Under most abiotic stress conditions, plant growth is inhibited and biomass production experiences a reduced demand for photosynthetic products. Consequently, carbohydrate accumulation leads to the feedback inhibition of photosynthetic gene expression (Paul & Pellny 2003). In our study, both PNR and PSBA were downregulated and fell into this expression pattern (Table 5 and 6). One possible explanation is that under LN stress, plants downregulate chlorophyll and protein synthesis to save limited nitrogen resources (Congming & Jianhua 2000). However, under LP stress, due to the lack of phosphate, ATP synthesis is hindered and consequently, photosynthesis related proteins are downregulated(Fan et al. 2016; Jacob & Lawlor 1993).

Starch and sucrose metabolism is vitally important for energy conversion in living organisms. Under LN and LP stress, 5 SDEGs including *EGLC, GLC, END, BGLU* and *AMY* were downregulated. Since starch and sucrose are synthesized from glyceraldehyde-3P via a series of enzyme reactions during photosynthesis, starch and sucrose metabolism is dependent on substrates derived from photosynthesis. Therefore, the expression of genes involved in starch and sucrose metabolism might be limited by the speed of photosynthesis, and a reduced photosynthetic rate may lead to the downregulation of these genes(Zhang et al. 2014). However, we also found that SDEGs encoding END and BGLU were upregulated under LN and LP stress. This needs further investigation.

Glycolysis/gluconeogenesis is another important energy conversion pathway in living organisms. In this study, *A1E* was downregulated under LN and LP stress. A1E (EC:5.1.3.3) catalyzes the interconversion between α-D-Glucose and β-D-Glucose (Li et al. 2013). As a result, the interconversion between α-D-Glucose and β-D-Glucose was downregulated.

Nitrogen metabolism is another important metabolic pathway in plants. Under LN and LP stress, 3 SDEGs related to nitrogen metabolism were detected and 2 of them were downregulated (*GOGAT* and *NIR1*). Similar results have also been shown in maize: under LN and LP conditions, the expression of genes involved in nitrate assimilation was inhibited (Schlüter et al. 2013), as nitrate assimilation requires a lot of energy and reducing power which is not available under stress conditions (Hu et al. 2011). However, the upregulation of NIR2 led to a higher conversion rate of nitrite into nitricoxide, indicating under LN and LP stress conditions, Chinese firs might convert more nitrogen into nitricoxide.

There are large overlaps between phenylpropanoid biosynthesis, phenylalanine metabolism, flavonoid biosynthesis and methane metabolism pathways. Many enzymes are involved in more than 2 of abovementioned pathways, rendering all the pathways closely connected. In this study, PAL was detected as a SDEG and downregulated under LN and LP stress. Moreover, *PLA* was also involved in phenylpropanoid biosynthesis and phenylalanine metabolism pathways. PAL (EC:4.3.1.24) is a key regulating enzyme in phenylpropanoid biosynthesis, which converts phenylalanine into cinnamic acid (Vogt 2010). The phenylpropanoid biosynthesis pathway is also inhibited under low LN and LP stress due to the downregulation of PLA. Similarly, in the phenylalanine metabolism pathway, PAL can convert phenylalanine into trans-cinnamate, and downstream products synthesis might be inhibited due to decreased levels of PAL under LN and LP stress. In this research both *POX and C4M* were involved in phenylpropanoid biosynthesis, phenylalanine metabolism and methane metabolism. Many SDEGs encoding POX were

downregulated. POX can be divided into 2 functional categories: one is related to plant morphogenesis, for example phenylpropanoid biosynthesis is related to lignin synthesis; and the other is related to plant resistance capacity by functioning as important defense enzymes(Apel & Hirt 2004; Nounjan et al. 2012; Park et al. 2003). As a result, LN and LP stress conditions do not only inhibit plant morphogenesis, but also reduce plant resistance capacity. *C4M* transcripts were upregulated under both LN and LP stress, indicating that stress conditions enhance the expression of this gene. Another selected SDEG that was upregulated in both conditions encodes CCoAOMT. CCoAOMT is related to lignin synthesis in zinnias(Meyermans et al. 2000). CCoAOMT (EC:1.1.1.219) catalyses the conversion of caffeoyl-CoA into feruloyl-CoA and when *CCoAOMT* expression is downregulated, the amount of lignin is significantly decreased in maize. *CCR* is only involved in phenylpropanoid biosynthesis pathways and has multiple substrates, for example, CCR catalyses cinnamoyl-CoA, p-coumaroyl-CoA and caffeoyl-CoA into the corresponding aldehyde. In this study, 2 SDEGs encoded CCR: 1 was upregulated under both LN and LP stress, while the other was downregulated in both conditions. *CCR* is a multigene family, therefore the 2 identified SDGEs need further characterization. *PAL*, *POX*, *C4M*, *CCoAOMT* and *CCR* all are involved in lignin synthesis (Huang et al. 2012; Liu et al. 2008). As an important timber wood in southern China, the selection of candidate genes with high efficiency under LN and LP stress conditions (involved in lignin synthesis) is one means of increasing timber wood production.

FDH (EC:1.2.1.2) is involved in the methane metabolism pathway, where it catalyzes the oxidation reaction of aromate into CO2, and NAD^+ is reduced into NADH. In potato, FDH expression levels were found to rapidly increase after 1–3 days of placing plants in abiotic stress conditions, including drought and low temperature(Hourton-Cabassa et al. 1998). However, in our study, FDH expression was significantly downregulated under LN and LP stress conditions.

This might be attributed to the duration of the stress treatment, which led to very low levels of nitrogen and phosphorus, resulting in insufficient materials for FDH synthesis. FLS, CHS and ANR all are involved in the flavonoid biosynthesis pathway. Flavonoids have a broad range of pharmacological functions, including antiviral functions, antineoplastic functions and prevention of cardio-cerebrovascular diseases (Middleton et al. 2000). Flavonoids also play an important role in the adaptation of plants to different stress conditions (Treutter 2006). FLS (EC:1.14.11.23) is an important enzyme that catalyses the conversion of dihydrokaempferol and dihydroquercetin into the corresponding flavones and flavonols. In this study, 4 selected SDEGs encode FLS; 2 of them were upregulated under both LN and LP stress conditions; 1 was downregulated under both LN and LP stress conditions; and 1 was upregulated under LP stress conditions and downregulated under LN stress conditions. Six FLS family members have been identified in *Arabidopsis*, therefore, future studies are needed to further characterize the 4 selected SDEGs. CHS (EC:2.3.1.74) is the first specific enzyme that functions in flavonoid biosynthesis pathways and catalyses p-cinnamoyl-COA, caffeoyl-CoA and feruloyl-CoA conversion into chalcone. Chalcone constitutes the backbone for the synthesis of flavonoids and locates upstream of the flavonoid biosynthesis pathway. Therefore, the CHS step is an important rate-limiting step in flavonoid biosynthesis(Shih et al. 2008; Wang et al. 2010). Interestingly, the SDEH encoding CHS was upregulated under LP stress conditions and downregulated under LN stress conditions. ANR (EC:1.3.1.77) is a key enzyme that is involved in anthocyanin synthesis. In the flavonoid biosynthesis pathway, ANR catalyzes pelargonidin, cyanidin and delphinidin conversion into (-)-Epiafzelechin, (-)-Epicatechin and (-)-Epigallocatechin. ANR exhibits similar expression patterns as CHS, and is upregulated under LP stress and downregulated under LN stress. Future studies are needed to explain the similarity.

References

Ågren GI, Wetterstedt J, and Billberger MF. 2012. Nutrient limitation on terrestrial plant growth–modeling the interaction between nitrogen and phosphorus. *New phytologist* 194:953-960.

Apel K, and Hirt H. 2004. Reactive oxygen species: metabolism, oxidative stress, and signal transduction. *Annu Rev Plant Biol* 55:373-399.

Bhattacharyya R, Kundu S, Prakash V, and Gupta H. 2008. Sustainability under combined application of mineral and organic fertilizers in a rainfed soybean–wheat system of the Indian Himalayas. *European Journal of Agronomy* 28:33-46.

Coelho GT, Carneiro NP, Karthikeyan AS, Raghothama KG, Schaffert RE, Brandão RL, Paiva LV, Souza IR, Alves VM, and Imolesi A. 2010. A phosphate transporter promoter from Arabidopsis thaliana AtPHT1; 4 gene drives preferential gene expression in transgenic maize roots under phosphorus starvation. *Plant Molecular Biology Reporter* 28:717-723.

Congming L, and Jianhua Z. 2000. Photosynthetic CO2 assimilation, chlorophyll fluorescence and photoinhibition as affected by nitrogen deficiency in maize plants. *Plant Science (Limerick)* 151:135-143.

Fan F, Ding G, and Wen X. 2016. Proteomic analyses provide new insights into the responses of Pinus massoniana seedlings to phosphorus deficiency. *Proteomics* 16:504-515.

Gonzalez-Dugo V, Durand J-L, and Gastal F. 2010. Water deficit and nitrogen nutrition of crops. A review. *Agronomy for Sustainable Development* 30:529-544.

Hayes J, Richardson A, and Simpson R. 2000. Components of organic phosphorus in soil extracts that are hydrolysed by phytase and acid phosphatase. *Biology and Fertility of Soils* 32:279-286.

Hourton-Cabassa C, Ambard-Bretteville F, Moreau F, de Virville JD, Rémy R, and des Francs-Small CC. 1998. Stress induction of mitochondrial formate dehydrogenase in potato leaves. *Plant Physiology* 116:627-635.

Hu B, Wang W, Ou S, Tang J, Li H, Che R, Zhang Z, Chai X, Wang H, and Wang Y. 2015. Variation in NRT1. 1B contributes to nitrate-use divergence between rice subspecies. *Nature Genetics* 47:834-838.

Hu B, Zhu C, Li F, Tang J, Wang Y, Lin A, Liu L, Che R, and Chu C. 2011. LEAF TIP NECROSIS1 plays a pivotal role in the regulation of multiple phosphate starvation responses in rice. *Plant Physiology* 156:1101-1115.

Huang H-H, Xu L-L, Tong Z-K, Lin E-P, Liu Q-P, Cheng L-J, and Zhu M-Y. 2012. De novo characterization of the Chinese fir (Cunninghamia lanceolata) transcriptome and analysis of candidate genes involved in cellulose and lignin biosynthesis. *BMC Genomics* 13:648.

HUANG Y-L, LI M-M, LU M, MO J-L, HOU G-G, LONG Q-Z, WANG H-M, and TANG X-Y. 2015. Selection of Rice Germplasm with High Nitrogen Utilization Efficiency and Its Analysis of the Related Characters. *Journal of Plant Genetic Resources* 16:87-93.

Ierna A, Mauro RP, and Mauromicale G. 2012. Improved yield and nutrient efficiency in two globe artichoke genotypes by balancing nitrogen and phosphorus supply. *Agronomy for Sustainable Development* 32:773-780.

Jacob J, and Lawlor D. 1993. In vivo photosynthetic electron transport does not limit photosynthetic capacity in phosphate‐deficient sunflower and maize leaves. *Plant, Cell & Environment* 16:785-795.

Jiang H, Peng S, Zhang S, Li X, Korpelainen H, and Li C. 2012. Transcriptional profiling analysis in Populus yunnanensis provides insights into molecular mechanisms of sexual differences in salinity tolerance. *Journal of Experimental Botany* 63:3709-3726.

Lapis-Gaza HR, Jost R, and Finnegan PM. 2014. Arabidopsis PHOSPHATE TRANSPORTER1 genes PHT1; 8 and

PHT1; 9 are involved in root-to-shoot translocation of orthophosphate. *Bmc Plant Biology* 14:334.

Li S, Ha S-J, Kim HJ, Galazka JM, Cate JH, Jin Y-S, and Zhao H. 2013. Investigation of the functional role of aldose 1-epimerase in engineered cellobiose utilization. *Journal of Biotechnology* 168:1-6.

Liu X, Deng Z, Gao S, Sun X, and Tang K. 2008. A new gene coding for p-coumarate 3-hydroxylase from Ginkgo biloba. *Russian Journal of Plant Physiology* 55:82-92.

Lu Y, Luo F, Yang M, Li X, and Lian X. 2011. Suppression of glutamate synthase genes significantly affects carbon and nitrogen metabolism in rice (Oryza sativa L.). *Science China Life Sciences* 54:651-663.

Ma Z, Huang B, Xu S, Chen Y, Cao G, Ding G, and Lin S. 2016. Ion Flux in Roots of Chinese Fir (Cunninghamia lanceolata (Lamb.) Hook) under Aluminum Stress. *Plos One* 11:e0156832.

Meyermans H, Morreel K, Lapierre C, Pollet B, De Bruyn A, Busson R, Herdewijn P, Devreese B, Van Beeumen J, and Marita JM. 2000. Modifications in lignin and accumulation of phenolic glucosides in poplar xylem upon down-regulation of caffeoyl-coenzyme A O-methyltransferase, an enzyme involved in lignin biosynthesis. *Journal of Biological Chemistry* 275:36899-36909.

Middleton E, Kandaswami C, and Theoharides TC. 2000. The effects of plant flavonoids on mammalian cells: implications for inflammation, heart disease, and cancer. *Pharmacological reviews* 52:673-751.

Mortazavi A, Williams BA, McCue K, Schaeffer L, and Wold B. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods* 5:621-628.

Moseley J, and Grossman AR. 2009. Phosphate metabolism and responses to phosphorus deficiency. *The Chlamydomonas sourcebook* 2:189-216.

Nounjan N, Nghia PT, and Theerakulpisut P. 2012. Exogenous proline and trehalose promote recovery of rice seedlings from salt-stress and differentially modulate antioxidant enzymes and expression of related genes. *Journal of Plant Physiology* 169:596-604.

Orwa C, Mutua A, Kindt R, Jamnadass R, and Simons A. 2009. Agroforestree database: a tree species reference and selection guide version 4.0. *World Agroforestry Centre ICRAF, Nairobi, KE*.

Park S, Ryu S, Kwon S, Lee H, Kim J, and Kwak S. 2003. Differential expression of six novel peroxidase cDNAs from cell cultures of sweetpotato in response to stress. *Molecular Genetics and Genomics* 269:542-552.

Paul MJ, and Pellny TK. 2003. Carbon metabolite feedback regulation of leaf photosynthesis and development. *Journal of Experimental Botany* 54:539-547.

Rivas-Ubach A, Sardans J, Pérez-Trujillo M, Estiarte M, and Peñuelas J. 2012. Strong relationship between elemental stoichiometry and metabolome in plants. *Proceedings of the National Academy of Sciences* 109:4181-4186.

Schlüter U, Colmsee C, Scholz U, Bräutigam A, Weber AP, Zellerhoff N, Bucher M, Fahnenstich H, and Sonnewald U. 2013. Adaptation of maize source leaf metabolism to stress related disturbances in carbon, nitrogen and phosphorus balance. *BMC Genomics* 14:442.

Sekhon G. 1995. Fertilizer-N use efficiency and nitrate pollution of groundwater in developing countries. *Journal of Contaminant Hydrology* 20:167-184.

Sharma V, Kumar A, Archana G, and Kumar GN. 2016. Ensifermeliloti overexpressing Escherichia coli phytase gene (appA) improves phosphorus (P) acquisition in maize plants. *The Science of Nature* 103:76.

Shih CH, Chu H, Tang LK, Sakamoto W, Maekawa M, Chu IK, Wang M, and Lo C. 2008. Functional characterization of key structural genes in rice flavonoid biosynthesis. *Planta* 228:1043-1054.

Silva DAd, Esteves JAdF, Gonçalves JGR, Azevedo CVG, Ribeiro T, Chiorato AF, and Carbonell SAM. 2016.

Evaluation of common bean genotypes for phosphorus use efficiency in Eutrophic Oxisol. *Bragantia*:0-0.

Singh A, Chaudhari V, and Ajay B. 2015. Screening of groundnut genotypes for phosphorus efficiency under field conditions. *Indian Journal of Genetics and Plant Breeding (The)* 75:363-371.

Spiertz J. 2010. Nitrogen, sustainable agriculture and food security. A review. *Agronomy for Sustainable Development* 30:43-55.

Treutter D. 2006. Significance of flavonoids in plant resistance: a review. *Environmental Chemistry Letters* 4:147-157.

Vance CP, Uhde ‐ Stone C, and Allan DL. 2003. Phosphorus acquisition and use: critical adaptations by plants for securing a nonrenewable resource. *New phytologist* 157:423-447.

Vogt T. 2010. Phenylpropanoid biosynthesis. *Molecular Plant* 3:2-20.

Wang Y, Li J, and Xia R. 2010. Expression of chalcone synthase and chalcone isomerase genes and accumulation of corresponding flavonoids during fruit maturation of Guoqing No. 4 satsuma mandarin (Citrus unshiu Marcow). *Scientia Horticulturae* 125:110-116.

Xiangqing M, Aiqin L, Baolong H, and Youli C. 2002. Study on selection of high-nitrogen-efficiency-genotypes of Chinese fir clones. *Scientia Silvae Sinicae* 38:53-57.

Zhang K, Liu H, Tao P, and Chen H. 2014. Comparative proteomic analyses provide new insights into low phosphorus stress responses in maize leaves. *Plos One* 9:e98215.

Table 1. Nutrient formulas for the mist

	CK (mg/L)	LN (mg/L)	LP (mg/L)
KNO_3	255	0.158	225
$Ca(NO_3)_2$	410	0	410
$MgSO_4.7H_2O$	490	490	490
KH_2PO_4	136	136	0.278
KCl	188	376	262.5
$CaCl_2$	193	470.5	193
H_3BO_3	2.86	2.86	2.86
$CuSO_4.5H2O$	0.08	0.08	0.08
$ZnSO_4.7H2O$	0.22	0.22	0.22
$MnCl_2.4H2O$	1.81	1.81	1.81
$H_2MoO.4H2O$	0.09	0.09	0.09
Fe_2EDTA	20	20	20

Table 2. RNA sequence pretreated results

sample	Raw Data Read	Valid Data Read	Base	Valid%	Q20%	Q30%	GC%
LN	36141806	34498198	4.31G	95.45	91.91	84.73	41.92
LP	56623784	55955276	6.99G	98.82	95.15	90.45	42.42
CK	43558660	41615988	5.20G	95.54	91.72	84.43	42.37

Table 3. Blast annotation results

	All	Mean GC%	Mean Length	Total Assembled bases	N50
unigene	33070	40.61	1019	33712721	1762
transcript	54188	40.69	1120	60691480	1809

Table 4. Blast annotation results

gene_number	swiss-prot	nr	Pfam	KEGG	KOG	GO
33070	13265	19896	16090	9424	16479	11772
100%	40.11 %	60.16 %	48.65 %	28.50 %	49.83 %	35.60 %

Table 5. Expression analysis of 78 SDEGs that were differentially expressed under both LN and LP stresses.

No.	Gene	gene_ID	Annotation	log2fold_change LN/CK	log2fold_change LP/CK
1	-	comp8974_c0	unknown	−1.516487272	1.195394176
2	MDL3	comp27665_c0	mandelonitrile lyase	−1.378643262	1.765767021
3	FLS	comp26906_c0	flavonol synthase	−1.602412955	1.295111419
4	CHS	comp26905_c0	chalcone synthase	−1.541431858	1.030163712
5	PD1	comp15924_c0	prephenate dehydratase	−1.782315176	1.407269327
6	MEE	comp24335_c0	DNA (cytosine-5-)-methyltransferase	1.770421146	−1.777607579
7	ATP6	comp30027_c0	H+-transporting ATPase	Inf	Inf
8	PAE	comp25227_c0	polyneuridine-aldehyde esterase	Inf	Inf
9	LOC	comp23944_c0	L-ascorbate oxidase	Inf	Inf
10	GPDH	comp22791_c0	glycerol-3-phosphate dehydrogenase (NAD+)	Inf	Inf
11	GSTF	comp22459_c0	glutathione S-transferase	Inf	Inf
12	NCL	comp18015_c0	nucleolin	Inf	Inf

13	SCD	comp10090_c0	stearoyl-CoA desaturase (delta-9 desaturase)	Inf	Inf
14	CYP2A1	comp26769_c0	cytochrome P450, family 1, subfamily A, polypeptide 1	9.07881795	9.260919534
15	ATP6	comp30691_c0	H+-transporting ATPase	8.434975656	7.95806932
16	PLA2	comp23451_c0	phospholipase A2	8.220048481	8.408136712
17	CYP26A	comp19769_c0	cytochrome P450, family 26, subfamily A	6.043721377	5.453872093
18	ACO	comp14611_c0	aminocyclopropanecarboxylate oxidase	5.742897447	5.465646202
19	TUA	comp25356_c0	tubulin alpha	5.31567212	5.606191644
20	C4M	comp20761_c0	trans-cinnamate 4-monooxygenase	4.737467711	3.846220345
21	CYP90D2	comp21003_c0	cytochrome P450, family 90, subfamily D, polypeptide 2 (steroid 3-oxidase)	4.450402653	4.528473263
22	FLS	comp8151_c0	flavonol synthase	3.869385758	3.981635233
23	KAO	comp26628_c0	ent-kaurenoic acid	3.337734191	3.07126943

			hydroxylase		
24	CCR	comp19489_c0	cinnamoyl-CoA reductase	3.223595419	2.887430397
25	EKO	comp25073_c0	ent-kaurene oxidase	3.0997515	2.152844636
26	FTH	comp23566_c0	ferritin heavy chain	2.800582421	2.097772574
27	FLS	comp19811_c0	flavonol synthase	2.496902224	1.513693427
28	CSD	comp23102_c0	Cu/Zn superoxide dismutase	2.462752617	2.439249289
29	END	comp5610_c0	endoglucanase	2.274480932	4.187408059
30	NIR2	comp22661_c0	nitrite reductase (NO-forming)	2.215205881	3.20026961
31	PLC	comp26823_c0	phospholipase C	2.154108863	3.299712243
32	ACP	comp16483_c0	acid phosphatase	2.105112412	4.566104345
33	BCK1	comp16113_c0	mitogen-activated protein kinase kinase kinase	2.008028799	2.293431018
34	CSD	comp10352_c0	Cu/Zn superoxide dismutase	1.662156377	1.330518926
35	CCoAOMT	comp27096_c0	caffeoyl-CoA O-methyltransferase	1.653593622	2.596690051
36	ESTA	comp26812_c0	esterase / lipase	1.529892865	1.165978227
37	UBC	comp23207_c0	ubiquitin C	1.459907237	1.085286718
38	BGLU	comp27001_c0	beta-glucosidase	1.417604976	1.488274508

39	MTK	comp26720_c0	5-methylthioribose kinase	1.313433715	1.957357635
40	AVP	comp23810_c0	inorganic pyrophosphatase	1.051120994	2.182704343
41	MDL3	comp22267_c0	mandelonitrile lyase	−1.028912326	−1.546383801
42	CCR	comp27038_c0	cinnamoyl-CoA reductase	−1.189221443	−1.003163156
43	PNR	comp23465_c0	ferredoxin--NADP+ reductase	−1.234942099	−1.39438315
44	CARB	comp25382_c0	carbamoyl-phosphate synthase large subunit	−1.306665634	−1.385071979
45	ACO	comp26762_c0	aminocyclopropanecarboxylate oxidase	−1.386226588	−2.22235954
46	GOGAT	comp21774_c0	glutamate synthase (NADPH/NADH)	−1.424752654	−1.391107616
47	POX	comp21916_c0	peroxidase	−1.439595303	−1.457382788
48	ANR	comp14096_c0	anthocyanidin reductase	−1.451046997	1.047309522
49	CYP26A	comp25986_c0	cytochrome P450, family 26, subfamily A	−1.509892306	−1.642910197
50	RPB1	comp24144_c0	DNA-directed RNA polymerase II subunit	−1.545553702	−1.723572386

			A		
51	TPS-CIN	comp22756_c0	1,8-cineole synthase	−1.697546348	−1.589005389
52	MEK1	comp26881_c0	mitogen-activated protein kinase kinase 1	−1.778545327	−2.795249817
53	AGAL	comp24611_c0	alpha-galactosidase	−1.815998835	−1.986456294
54	PHOB	comp23219_c0	phosphatidylinositol-bisphosphatase	−1.856042615	−3.791825989
55	PAL	comp26210_c0	phenylalanine ammonia-lyase	−1.908962334	−1.571166988
56	PSBA	comp24767_c0	photosystem II PsbA protein	−2.374644429	−3.143063306
57	CHI	comp26078_c0	chitinase	−2.578014078	−1.946984816
58	COBA	comp24511_c0	uroporphyrin-III C-methyltransferase	−2.608584431	−2.196979126
59	FDH	comp25550_c0	formate dehydrogenase	−2.639091743	−2.366238221
60	NIR1	comp14490_c0	ferredoxin-nitrite reductase	−2.97950818	−1.435176293
61	AMY	comp22510_c0	alpha-amylase	−3.05932428	−2.002585476
62	A1E	comp27133_c0	aldose 1-epimerase	−3.258306998	−1.729975945
63	CHI	comp23644_c0	chitinase	−3.634234447	−2.416812452

64	POX	comp18870_c0	peroxidase	−3.730929125	−1.698092365
65	AOS	comp26923_c0	hydroperoxide dehydratase	−3.862004198	−1.921469537
66	LOX	comp26770_c0	lipoxygenase	−3.985137611	−2.291296314
67	POX	comp15460_c0	peroxidase	−4.427606173	−3.188819313
68	POX	comp25585_c0	peroxidase	−4.58382786	−3.066988081
69	EGLC	comp25401_c0	glucan endo-1,3-beta-D-glucosidase	−4.604169973	−1.159739518
70	GLC	comp23680_c0	glucan 1,3-beta-glucosidase	−4.624490865	−3.440251267
71	PRDX6	comp6035_c0	peroxiredoxin 6, 1-Cys peroxiredoxin	−5.009123526	−5.449696118
72	ACO	comp21662_c0	aminocyclopropanecarboxylate oxidase	−5.337654855	−4.722220002
73	EGLC	comp26227_c0	glucan endo-1,3-beta-D-glucosidase	−5.772142607	−2.61431182
74	ACO	comp25697_c0	aminocyclopropanecarboxylate oxidase	−6.46881556	−4.952101361
75	POX	comp20256_c0	peroxidase	−6.533978572	−3.689433091
76	BGLU	comp18357_c0	beta-glucosidase	−7.009465786	−1.834672174
77	FLS	comp31858_c0	flavonol synthase	−8.85997842	−4.448164821
78	PAOX	comp10040_c0	N1-acetylpolyamine oxidase	−9.814048332	−6.249263713

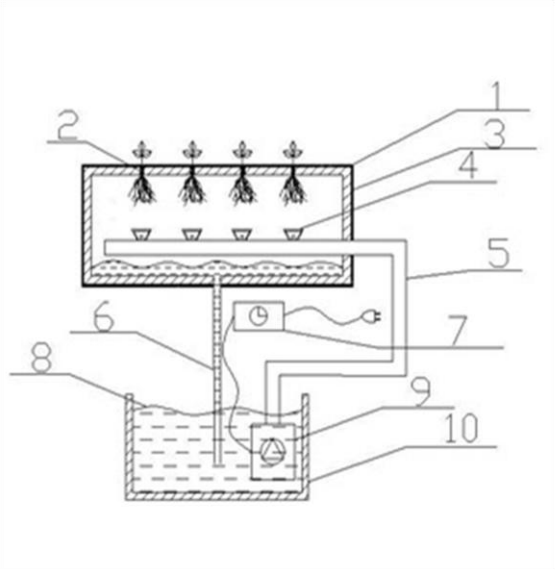
PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2916v1 | CC BY 4.0 Open Access | rec: 8 Apr 2017, publ: 8 Apr 2017

Table 6. Shared differentially expressed genes under LN and LP stresses.

No.	Gene	Annotation	KEGG Pathway	gene_ID
1	PNR	ferredoxin--NADP+ reductase	Photosynthesis	comp23465_c0
2	PSBA	photosystem II PsbA protein	Photosynthesis	comp24767_c0
3	EGLC	glucan endo-1,3-beta-D-glucosidase	Starch and sucrose metabolism	comp26227_c0; comp25401_c0
4	GLC	glucan 1,3-beta-glucosidase	Starch and sucrose metabolism	comp23680_c0
5	END	endoglucanase	Starch and sucrose metabolism	comp5610_c0
6	BGLU	beta-glucosidase	Starch and sucrose metabolism & Phenylpropanoid biosynthesis	comp27001_c0; comp18357_c0
7	AMY	alpha-amylase	Starch and sucrose metabolism	comp22510_c0
8	A1E	aldose 1-epimerase	Glycolysis / Gluconeogenesis	comp27133_c0
9	PAL	phenylalanine ammonia-lyase	Nitrogen metabolism & Phenylpropanoid biosynthesis & Phenylalanine metabolism	comp26210_c0
10	GOGAT	glutamate synthase	Nitrogen metabolism	comp21774_c0

		(NADPH/NADH)		
11	NIR1	ferredoxin-nitrite reductase	Nitrogen metabolism	comp14490_c0
12	NIR2	nitrite reductase (NO-forming)	Nitrogen metabolism	comp22661_c0
13	C4M	trans-cinnamate 4-monooxygenase	Phenylpropanoid biosynthesis & Flavonoid biosynthesis & Phenylalanine metabolism	comp20761_c0
14	PRDX6	peroxiredoxin 6, 1-Cys peroxiredoxin	Phenylpropanoid biosynthesis & Phenylalanine metabolism & Methane metabolism	comp6035_c0
15	POX	peroxidase	Phenylpropanoid biosynthesis & Phenylalanine metabolism & Methane metabolism	comp25585_c0; comp21916_c0; comp20256_c0; comp18870_c0; comp15460_c0
16	CCR	cinnamoyl-CoA reductase	Phenylpropanoid biosynthesis	comp27038_c0; comp19489_c0
17	CCoAOMT	caffeoyl-CoA O-methyltransferase	Phenylpropanoid biosynthesis & Flavonoid biosynthesis & Phenylalanine metabolism	comp27096_c0
18	FDH	formate	Methane metabolism	comp25550_c0

		dehydrogenase		
19	FLS	flavonol synthase	Flavonoid biosynthesis	comp8151_c0; comp31858_c0; comp26906_c0; comp19811_c0
20	CHS	chalcone synthase	Flavonoid biosynthesis	comp26905_c0
21	ANR	anthocyanidin reductase	Flavonoid biosynthesis	comp14096_c0



1. Black opaque cover
2. Planting device
3. Incubator
4. Nutrient solution outlet nozzles
5. Nutrient solution transport tubes
6. Nutrient solution recovery tubes
7. Timer
8. Nutrient solution
9. Frequency conversion submersible pump
10. Nutrient solution storage vessel.

Figure 1

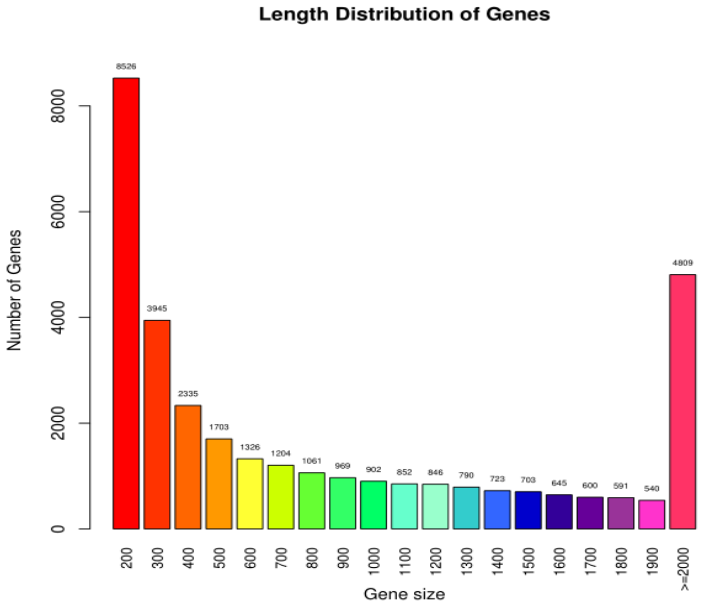

Length Distribution of Genes
Number of Genes
0
2000
4000
6000
8000
8526
3945
2335
1703
1326
1204
1061
969
902
852
846
790
723
703
645
600
591
540
4809
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
>=2000
Gene size

Figure 2

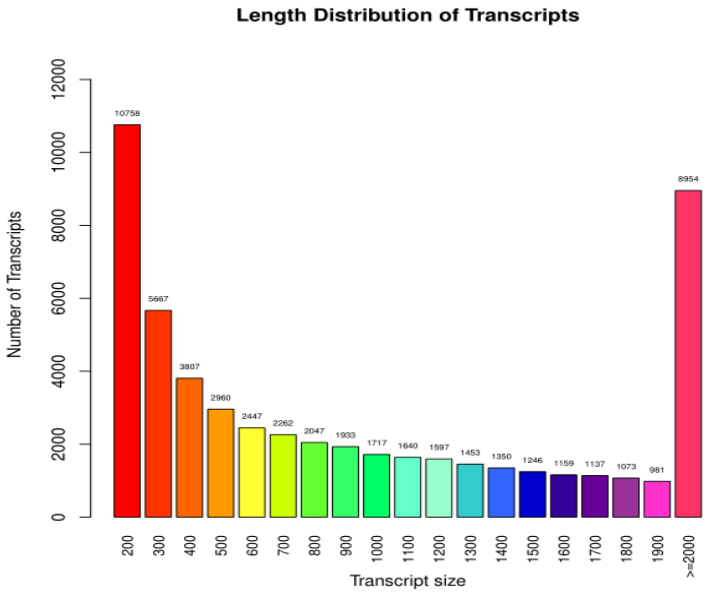

Length Distribution of Transcripts
Number of Transcripts
0
2000
4000
6000
8000
10000
12000
10758
5667
3807
2960
2447
2262
2047
1933
1717
1640
1597
1453
1350
1246
1159
1137
1073
981
8954
200
300
400
500
600
700
800
900
1000
1100
1200
1300
1400
1500
1600
1700
1800
1900
>=2000
Transcript size

Figure 3

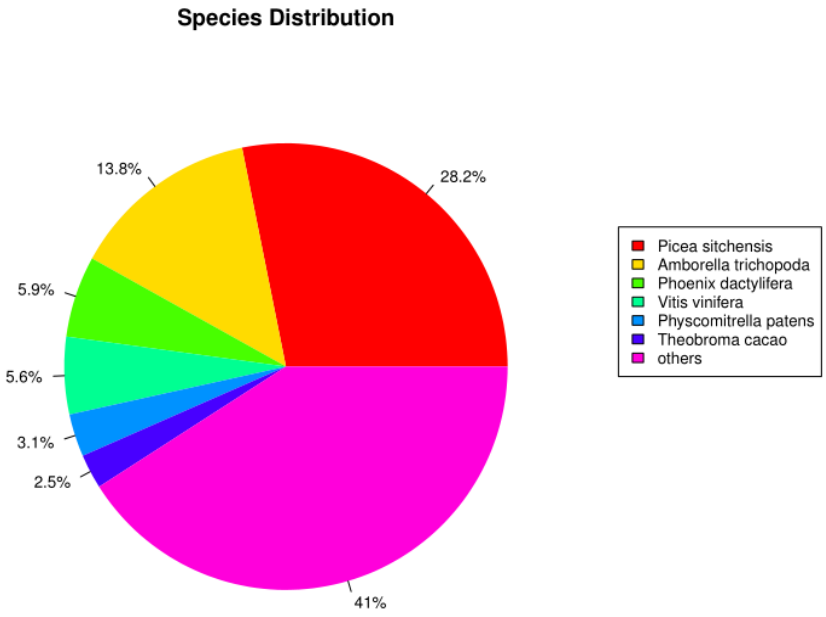


Figure 4

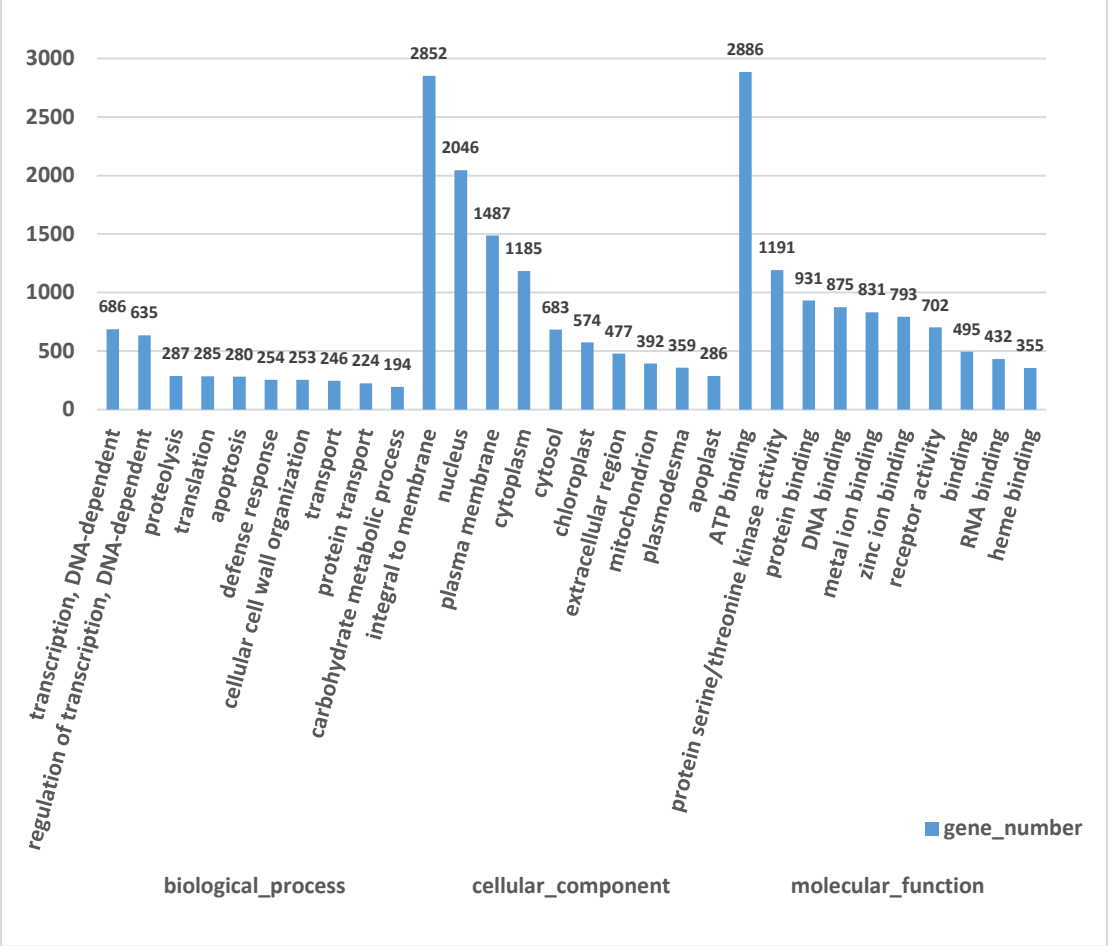

3000
2500
2000
1500
1000
500
0
686 635 287 285 280 254 253 246 224 194 2852 2046 1487 1185 683 574 477 392 359 286 2886 1191 931 875 831 793 702 495 432 355
transcription, DNA-dependent
regulation of transcription, DNA-dependent
proteolysis
translation
apoptosis
defense response
cellular cell wall organization
transport
protein transport
carbohydrate metabolic process
integral to membrane
nucleus
plasma membrane
cytoplasm
cytosol
chloroplast
extracellular region
mitochondrion
plasmodesma
apoplast
ATP binding
protein serine/threonine kinase activity
protein binding
DNA binding
metal ion binding
zinc ion binding
receptor activity
binding
RNA binding
heme binding
gene_number
biological_process
cellular_component
molecular_function

Figure 5

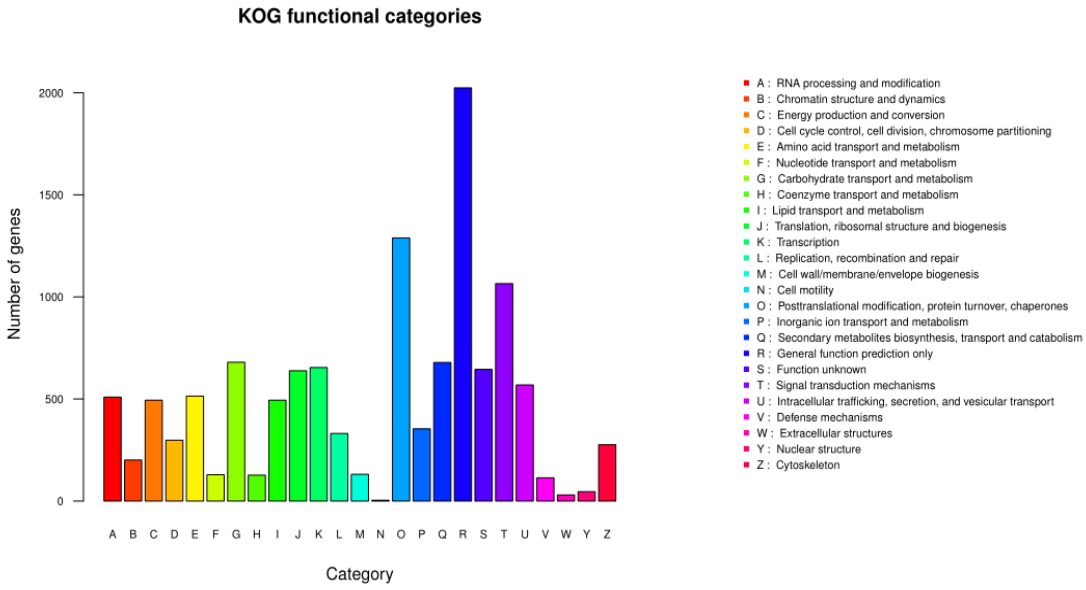

KOG functional categories
Number of genes
2000
1500
1000
500
0
A B C D E F G H I J K L M N O P Q R S T U V W Y Z
Category
A : RNA processing and modification
B : Chromatin structure and dynamics
C : Energy production and conversion
D : Cell cycle control, cell division, chromosome partitioning
E : Amino acid transport and metabolism
F : Nucleotide transport and metabolism
G : Carbohydrate transport and metabolism
H : Coenzyme transport and metabolism
I : Lipid transport and metabolism
J : Translation, ribosomal structure and biogenesis
K : Transcription
L : Replication, recombination and repair
M : Cell wall/membrane/envelope biogenesis
N : Cell motility
O : Posttranslational modification, protein turnover, chaperones
P : Inorganic ion transport and metabolism
Q : Secondary metabolites biosynthesis, transport and catabolism
R : General function prediction only
S : Function unknown
T : Signal transduction mechanisms
U : Intracellular trafficking, secretion, and vesicular transport
V : Defense mechanisms
W : Extracellular structures
Y : Nuclear structure
Z : Cytoskeleton

Figure 6

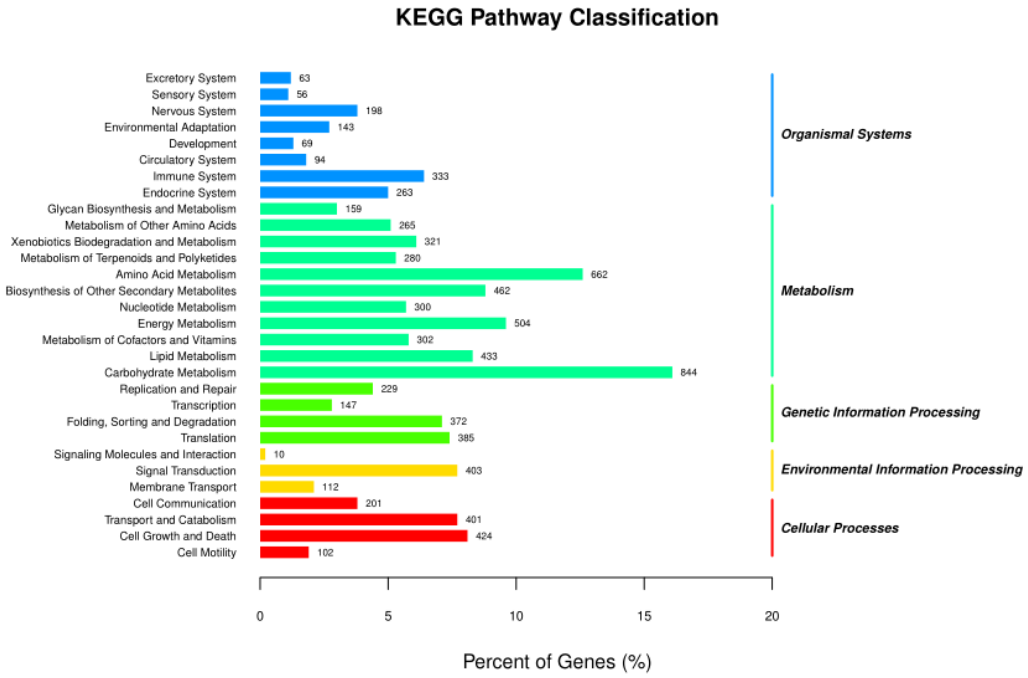

KEGG Pathway Classification
Excretory System 63
Sensory System 56
Nervous System 198
Environmental Adaptation 143
Development 69
Circulatory System 94
Immune System 333
Endocrine System 263
Glycan Biosynthesis and Metabolism 159
Metabolism of Other Amino Acids 265
Xenobiotics Biodegradation and Metabolism 321
Metabolism of Terpenoids and Polyketides 280
Amino Acid Metabolism 662
Biosynthesis of Other Secondary Metabolites 462
Nucleotide Metabolism 300
Energy Metabolism 504
Metabolism of Cofactors and Vitamins 302
Lipid Metabolism 433
Carbohydrate Metabolism 844
Replication and Repair 229
Transcription 147
Folding, Sorting and Degradation 372
Translation 385
Signaling Molecules and Interaction 10
Signal Transduction 403
Membrane Transport 112
Cell Communication 201
Transport and Catabolism 401
Cell Growth and Death 424
Cell Motility 102
Organismal Systems
Metabolism
Genetic Information Processing
Environmental Information Processing
Cellular Processes
0
5
10
15
20
Percent of Genes (%)

Figure 7

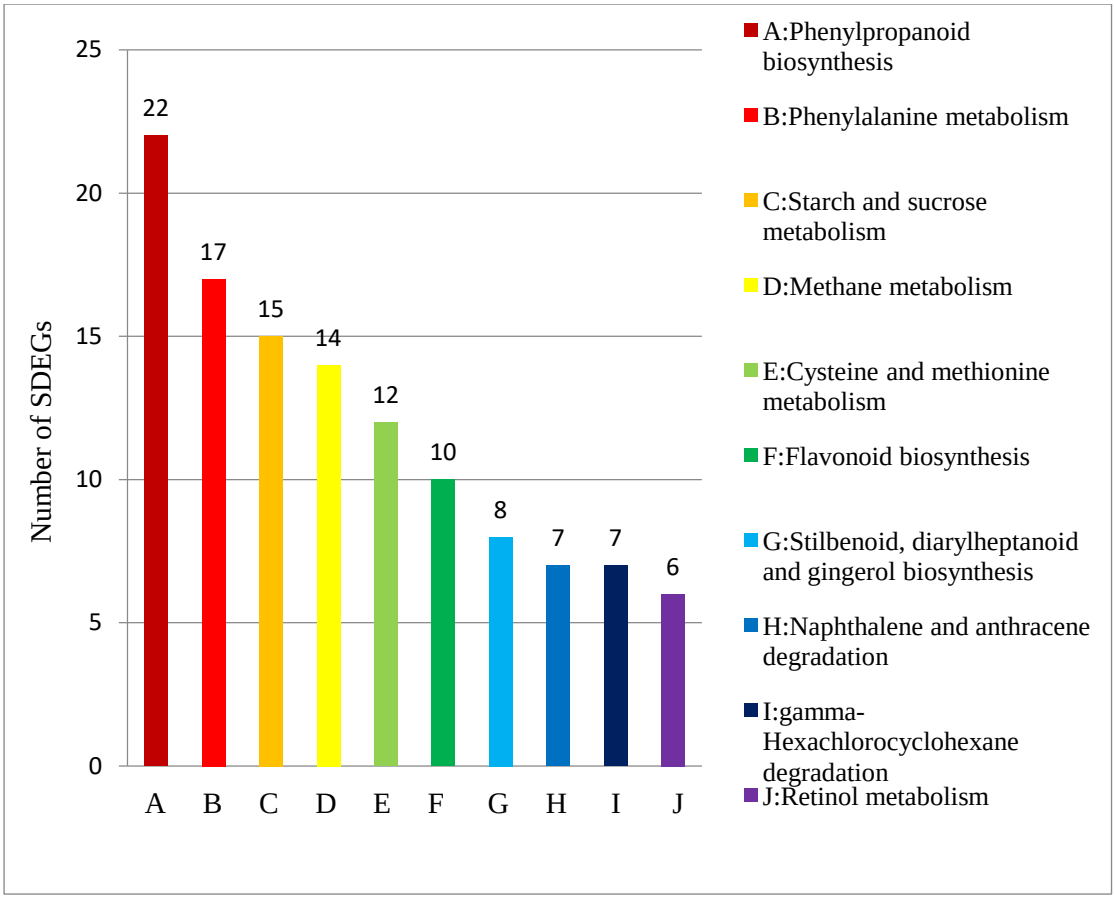

Number of SDEGs
25
20
15
10
5
0
22
17
15
14
12
10
8
7
7
6
A
B
C
D
E
F
G
H
I
J
A:Phenylpropanoid biosynthesis
B:Phenylalanine metabolism
C:Starch and sucrose metabolism
D:Methane metabolism
E:Cysteine and methionine metabolism
F:Flavonoid biosynthesis
G:Stilbenoid, diarylheptanoid and gingerol biosynthesis
H:Naphthalene and anthracene degradation
I:gamma-Hexachlorocyclohexane degradation
J:Retinol metabolism

Figure 8

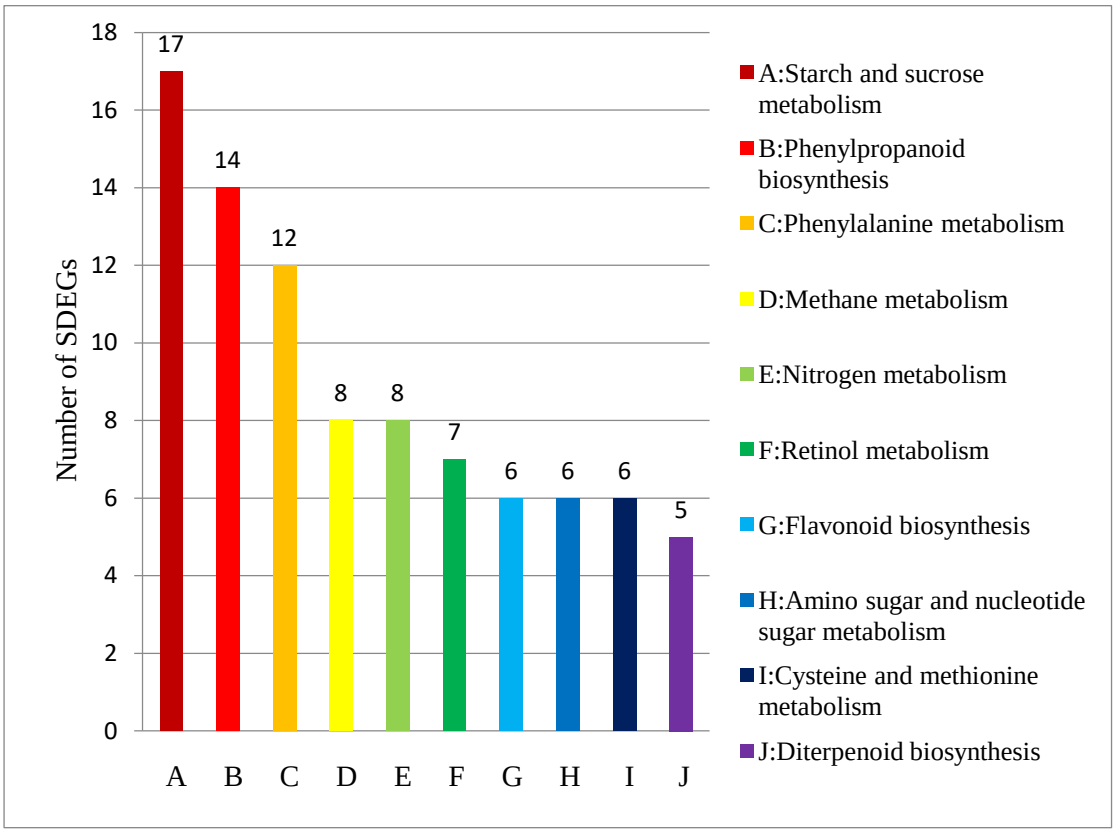

Number of SDEGs
18
16
14
12
10
8
6
4
2
0
17
14
12
8
8
7
6
6
6
5
A
B
C
D
E
F
G
H
I
J
A:Starch and sucrose metabolism
B:Phenylpropanoid biosynthesis
C:Phenylalanine metabolism
D:Methane metabolism
E:Nitrogen metabolism
F:Retinol metabolism
G:Flavonoid biosynthesis
H:Amino sugar and nucleotide sugar metabolism
I:Cysteine and methionine metabolism
J:Diterpenoid biosynthesis

Figure 9