Pak. J. Bot., 48(3): 885-898, 2016.

DE NOVO TRANSCRIPTOME SEQUENCEING AND COMPARATIVE ANALYSIS OF DIFFERENTIALLY EXPRESSED GENES IN *DRYOPERIS FRAGRANS* UNDER TEMPERATURE STRESS

WEN-ZHONG WANG[1,2√], WEI-SHUANG TONG[1√], YAN LI[1], RUI GAO[1], LIGUO ZHANG[1,3] AND YING CHANG[1]

[1]*College of Life Sciences, Northeast Agricultural University, Harbin 150030, China*
[2] *Virus-free Seedling Research Institute, Heilongjiang Academy of Agricultural Sciences, Harbin 150086, China*
[3] *Institute of Industrial Crops, Heilongjiang Academy of Agricultural Sciences, Harbin 150086, China*
[*]*Corresponding author's e-mail: changying@neau.edu.cn; Tel: +86 0451 55190410*
[√]*Authors contributed equally to this work.*

Abstract

Dryopteris fragrans is a species of fern and contains flavonoids compounds with medicinal value. This study explain the temperature stress impact flavonoids synthesis in *D. fragrans* tissue culture seedlings under the low temperature at 4ºC, high temperature at 35ºC and moderate temperature at 25ºC. By using Illumina HiSeq 2000 sequencing, 80.9 million raw sequence reads were *de novo* assembled into 66,716 non-redundant unigenes. 38,486 unigenes (57.7%) were annotated for their function. 13,973 unigenes and 29,598 unigenes were allocated to gene ontology (GO) and clusters of orthologous group (COG), respectively. 18,989 sequences mapped to 118 Kyoto Encyclopedia of Genes and Genomes Pathway database (KEGG), 204 genes were involved in flavonoid biosynthesis, regulation and transport. 25,292 and 16,817 unigenes exhibited marked differential expression in response to temperature shifts of 25ºC to 4ºC and 25ºC to 35ºC, respectively. *4CL* and *CHS* genes involved in flavonoid biosynthesis were tested and suggested that they were responsible for biosynthesis of flavonoids. This study provides the first published data to describe the *D. fragrans* transcriptome and should accelerate understanding of flavonoids biosynthesis, regulation and transport mechanisms. Since most unigenes described here were successfully annotated, these results should facilitate future functional genomic understanding and research of *D. fragrans.*

Key words: Transcriptome; *Dryopteris fragrans*; Flavonoids biosynthesis; DEG (differentially expressed genes).

Introduction

Dryopteris fragrans (L.) Schott, a species of fern belonging to the *Dryopteridaceae* family, is abundant in Wudalianchi, Heilongjiang Province, China (Wang, 1989). *D. fragrans* typically lives in lava rock in the temperature range of 30°C to 60°C and can endure temperatures below −20°C (Ao & Li, 1990). Such an environment has probably played a role in development of this species' unique pharmacological activities. *D. fragrans* has been used as a Chinese medicinal crop for thousands of years and contains various bioactive components with distinct, but overlapping, activities for treatment of arthritis and various skin diseases (Li *et al.*, 2005). Phloroglucinol, flavonoids, terpenoids and other plant secondary substances with known medical utility have been isolated and characterized from this fern (Banerjee, 1980; Widen *et al.*, 2001; Shen *et al.*, 2006). The major bioactive compounds in *D. fragrans* are flavonoids, which have been reported to have a variety of medical applications including prevention and treatment of cancer (Zhang *et al.*, 2003), cell activation leading to cancer cell death (Erhart *et al.*, 2005) and treatment of cardiovascular disease (Graf *et al.*, 2005).

Ferns occupy an important evolutionary niche and exhibit independent gametophyte and sporophyte life stages. *D. fragrans* is a unique kind of fern that grows in extremely frigid zones. The study of the relationship between *D. fragrans* and the environment is an important research field, given that expression of diverse flavonoids varies with temperature stress and that temperature is known to influence flavonoids composition, quality and yield. Because the molecular mechanisms of bioflavonoid biosynthesis may be diverse and complex, modern genetic tools are needed to better understand flavonoids biosynthesis and transport in *D. fragrans*.

Because of their economic importance, the extraction and purification of bioactive components from *D. fragrans* has attracted much research interest (Li *et al.*, 2012; Sun *et al.*, 2013). However, only recently a few research papers have been published focusing on the molecular biology of this organism, including work describing cloning of 1 *DfMADS1* gene (Huang *et al.*, 2014), 1 *psb*A gene (Gao *et al.*, 2014), 4 *Df4CL* genes (Li *et al.*, 2015a), 1 *DfCHS* gene (Sun *et al.*, 2014), 3 *DfPAL* genes and 1 *DfC4H* gene (Li *et al.,* 2015b). While several molecular genomic and gene expression studies have already focused on the genus *Dryopteris*, newer methods, such as next-generation sequencing (NGS), have not yet been widely used. NGS methods hold great promise for advancing this field due to their advantages of high speed, high throughput, generation of large amounts of data, amenability to a high degree of automation, low cost, among others. Many plants and crops have been sequenced using this technology (Goff *et al.*, 2002; Paterson *et al.*, 2009; AI-dous *et al.*, 2011). However, until 2011 only the transcriptome analysis of the bracken fern (*Pteridium aquilinum*) had been completed by Roche using the 454 GS-FLX sequencing method (Der *et al.*, 2011). In order to perform more extensive transcriptome analyses, NGS methods have replaced the older 454 pyrosequencing technology. Because of much greater output, Illumina HiSeq™ 2000 whole transcriptome sequencing has been used for such transcriptome projects successfully (Qi *et al.* 2011).

Because the study of differentially expressed gene (DEG) transcripts has shed light on our understanding of many vital pathways and their regulation in other organisms, this methodology was used here to decipher molecular mechanisms underlying *D. fragrans* flavonoids biosynthesis and transport in response to temperature stress. Due to the large amount of *D. fragrans* transcriptome data generated in this study using Illumina HiSeq™ 2000, an extensive repository of differentially expressed fern genes will be available to inform future gene expression and functional genomic studies. It will further our understanding of the evolutionary and functional genomics of ferns.

Materials and Methods

Sporophyte culture and temperature treatments: The plant and spores of *D. fragrans* (plant specimen number: WDLC-1) were collected from a single individual plant which was assayed over the course of one year in Wudalianchi, China (126° 07'07" N, 48° 42'38" E). No specific permissions were required before studying these samples, because these samples were studied within a laboratory that allows researchers to legally conduct such research.

D. fragrans spores were cultivated as described previously (Sun *et al.*, 2014). When sporophytes appeared, we performed secondary culture of the sporophytes on 1/2 MS culture media at 25°C then cultured them under various temperature conditions (4°C or 35°C) or under the moderate temperature condition (25°C) as a control. Treated plants were labeled T1 (4°C), T2 (35°C) and the control was labeled CK (25°C). 48 h later, the differentially treated plant samples were flash frozen in liquid nitrogen and preserved at −80°C.

Total flavonoids measurements: The samples were prepared from sporophytes, gametophytes, sporophyte leaves, petioles and roots of *D. fragrans* which were all pooled and cultured seedlings grown under the three temperature stress conditions. Total flavonoids content was determined using a colorimetric method (Müller *et al.*, 2005; Zhang *et al.*, 2012) and total flavonoids levels were determined using calibration curves.

RNA isolation, cDNA library preparation and sequencing: Total RNA (100 mg) was isolated and purified using RNAplant Plus Reagent and an RNAprep Pure Kit (for Plant) (TIANGEN, Beijing, China), respectively. The RNA samples were pooled from sporophytes, gametophytes, sporophyte leaves, petioles and roots of *D. fragrans* tissue cultures seedlings grown under the three temperature stress conditions. For cDNA library preparation, 20 μg of RNA (of a 400 ng/μL stock) was reverse-transcribed using the Haigene Reverse Transcription Kit and sequenced by the Beijing Genomics Institute (BGI) using Illumina HiSeq™ 2000.

Sequencing and analysis of the *D. fragrans* transcriptome: Raw RNA-sequencing reads were generated using high-throughput sequencing and included dirty reads, adapter sequences and sequences containing unknown base calls or low quality sequences. Because suboptimal sequence data would negatively affect subsequent bioinformatics analysis, such reads were discarded. After data screening, the SOAPdenovo program was used for transcriptome analysis (Li *et al.*, 2010) and unigenes were assembled sequentially. A final set of unigenes from the samples for each temperature shift condition were assembled and redundant unigenes were removed using sequencing clustering software to generate a final set of 66,716 non-redundant unigenes.

For functional annotation and metabolic pathway analyses of non-redundant unigenes sequences, multiple databases were used, including the NCBI non-redundant protein (Nr) database, Swiss-Prot database, the Kyoto Encyclopedia of Genes and Genomes (KEGG) and the Cluster of Orthologous Groups (COGs). The unigenes sequences were compared to the databases using BLASTX (E-value $<10^{-5}$) and sequence matches were captured using the best alignment results. Gene Ontology (GO) annotations were performed as previously described (Conesa *et al.*, 2005; Ye *et al.*, 2006).

In order to pinpoint genes exhibiting different expression levels under different temperature conditions, a previously described method was used (Audic & Claverie, 1997). GO functional analysis provided GO functional specification annotation for DEGs to direct assignment of GO and KEGG pathway annotations. After analyses were complete for each experimental condition under study, integration of data across all compared experimental conditions was performed using cluster analysis of expression patterns.

Quantitative RT-PCR validation and analysis: In preparation for RNA sequencing, RNA was converted to cDNA using the RT-PCR kit V3.0 (TaKara). The primers for *Df4CL* validation were designed from Unigene18514_All, Unigene29145_All, Unigene15437_All, Unigene19002_All, Unigene25561_All and known *4CL* genes cDNA sequences. The *DfCHS* primers were designed from Unigene26015_All, Unigene31220_All, Unigene33005_All, Unigene36123_All and known *CHS* genes cDNA sequences (Table 1).

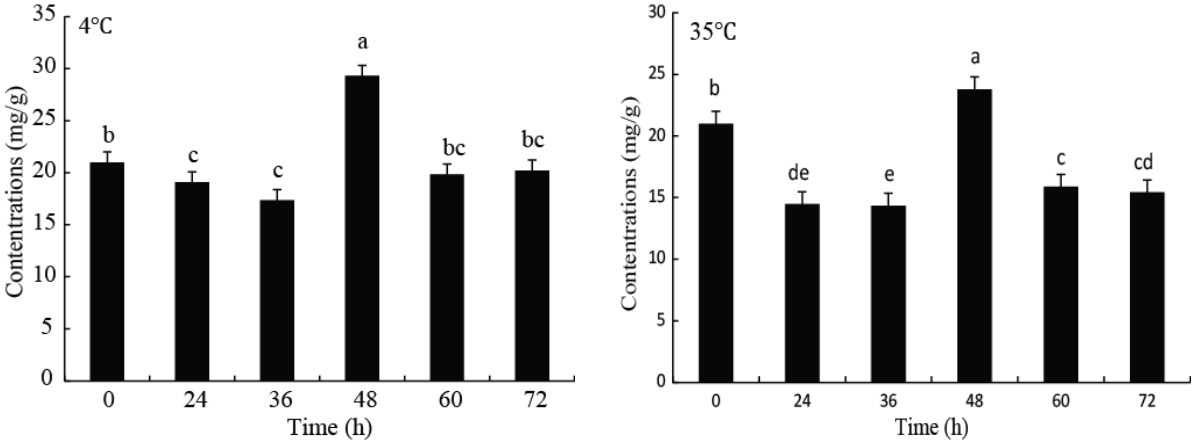
A real time PCR system was used to quantify cDNA levels (Applied Biosystems 7500, USA) using SYBR Green staining and the following reaction system (20 μL reaction including 1 μL template, 0.5 μL forward primer, 0.5 μL reverse primer, 10 μL SYBR Green QPCR Master Mix and 0.4 μL ROX) (TaKaRa, Japan). The 18S rRNA gene from *D. fragrans* was used as an internal reference. The PCR protocol was as follows: 95°C for 10 s, 40 cycles of 95°C for 5 s and 60°C for 35 s. The experiments were each repeated three times. Quantitative changes were analyzed using the Ct-value and calculated using a standard curve.

Results

Total flavonoids content of *D. fragrans* cultured seedling samples: For analyses of the total flavonoids content under different temperature conditions, we used the *D. fragrans* cultured seedling samples at 25ºC (0 h) as a control. According to calculations using the standard curve, the flavonoid content of the control (25ºC) was 21.00±0.76 mg/mL ($p<0.05$), while flavonoids contents for 48 h were 29.31±1.06 mg/mL at 4ºC (Fig. 1A, $p<0.05$) and 23.80±0.61 mg/mL at 35ºC (Fig. 1B, $p<0.05$), respectively. These results indicated that flavonoids content exhibited significant variations under different temperature treatments, and higher flavonoid levels relative to the control was found at lower and higher temperature samples.

Table 1. Sequences of the primers used in this study.

Primer	Length	Sequence (5' to 3')
Df4CL-F	29	CTCATCGAGCACAGCAACACTGAATTCAA
Df4CL-R	29	CTGGTACAACAGGCTTACCCAAAGGAGTC
DfCHS-F	26	GGAAACATGTCGAGTGCGTGCGTGAT
DfCHS-R	25	GACCCCACTCAAACCCTTCCCCTGT
18s-F	23	GCTTTCGCAGTAGTTCGTCTTTC
18s-R	24	TGGTCCTATTATGTTGGTCTTCGG



A= Low temperature (4ºC) **B=** High temperature (35ºC)

Fig. 1. Total flavonoid content of *Dryopteris fragran*s (L.) Schott under temperature stress.

Sequencing and *de novo* assembly results: To acquire the *D. fragrans* transcriptome involved in flavonoids biosynthesis, regulation and transport under different temperature treatments, RNA was extracted from the pools of whole sporophytes. Three cDNA libraries were built for three temperature shifts from 25°C to the following: low-temperature condition of 4°C (T1), high-temperature condition of 35°C (T2) or moderate-temperature condition of 25°C (CK). After rigorous data screening, approximately 80.9 million RNA-Seq reads were generated from all samples in the three libraries and included 27,039,850 reads (T1), 27,047,380 reads (T2) and 26,775,874 reads (CK). From the T1, T2 and CK libraries, the Q20 bases (sequencing error rate <1%) were 91.31%, 92.47% and 92.36% and GC percentages were 48.68%, 49.88% and 49.57%, respectively. All of these results demonstrate a high degree of accuracy for the sequencing and assembly results.

For T1, the contigs were assembled into 52,561 unigenes with a median length of 333 bp and an N50 size of 356 bp. For T2, the contigs were assembled into 64,292 unigenes with a median length of 387 bp and an N50 size of 449 bp. For CK, the contigs were assembled into 56,095 unigenes with a median length of 401 bp and an N50 size of 477 bp. In the final step, these three sets of unigenes were merged using the TIGR Gene Indices Clustering Tool (TGICL), resulting in a total final collection of 66,716 non-redundant unigenes with a median length of 524 bp and an N50 size of 652 bp (Table 2). 66.92% (44,646 unigenes) of all unigenes lengths were between 100 and 500 bp, 21.89% (14,606 unigenes) were between 500 and 1,000 bp, 6.96% (4,644 unigenes) were between 1,000 and 1,500 bp, 2.77% (1,851 unigenes) were between 1,500 and 2,000 bp, 1.45% (969 unigenes) were longer than 2,000 bp (Fig. 2). BLAST searches identified 39,174 potential coding sequences (CDSs); of these, the lengths of 18,141 unigenes were greater than 300 bp, while 38,262 unigenes with positive match results had no gaps in their sequence alignments.

Functional annotation and classification by GO and COG: The result indicated that 35,458 (53.1%), 28,194 (42.3%), 18,989 (28.5%), 17, 327 (26.0%) and 13, 973 (21.0%) unigenes demonstrated significant similarity to known proteins in Nr, Swiss-Prot, KEGG, COG and GO databases, respectively (Table 3). Furthermore, 24.88% of the mapped sequences exhibited significant homology to known proteins (<1.0E-45) (Fig. 3A). These results implied that the unigenes involved in flavonoids biosynthesis, regulation and transport were intimately tied to responses to temperature stresses. Interestingly, 29.98% of the unigenes showed significant homology to sequences of *Arabidopsis thaliana*, 16.32% and 7.33% of the mapped sequences showed significant homology to the sequences of *Oryza sativa* Japonica Group and *Zea mays,* respectively (Fig. 3B). In summary, annotation results suggested that a huge number of assembled transcripts of *D. fragrans* were successfully characterized using HiSeq™ 2000 sequencing.

Table 2. Overview of the sequencing and assembly.

	T1	T2	CK	
Total reads	27,039,850	27,039,850	26,775,874	
Total clean nucleotides (bp)	2,433,586,500	2,434,264,200	2,409,828,660	
Q20 percentage	91.31%	92.47%	92.36%	
GC percentage	48.68%	49.88%	49.57%	
Total number of contigs	153,608	156,386	146,105	
Length of all contig (bp)	25,670,172	32,188,795	29,456,543	
Average length of contigs (bp)	167	206	202	
Contig N50 (bp)	196	255	260	
Total number of unigenes	52,561	64,292	56,095	66,716
Length of all unigene (bp)	17,480,556	24,874,416	22,481,690	34,970,672
Average length of unigenes (bp)	333	387	401	524
Unigene N50 (bp)	356	449	477	652

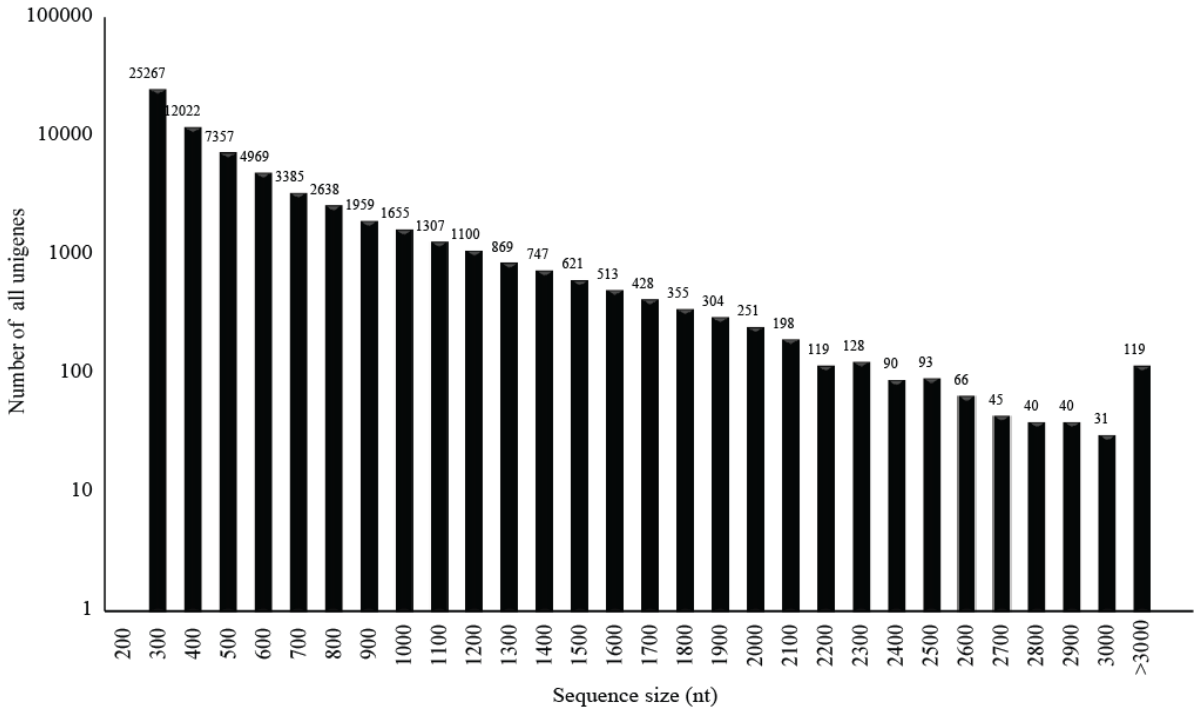
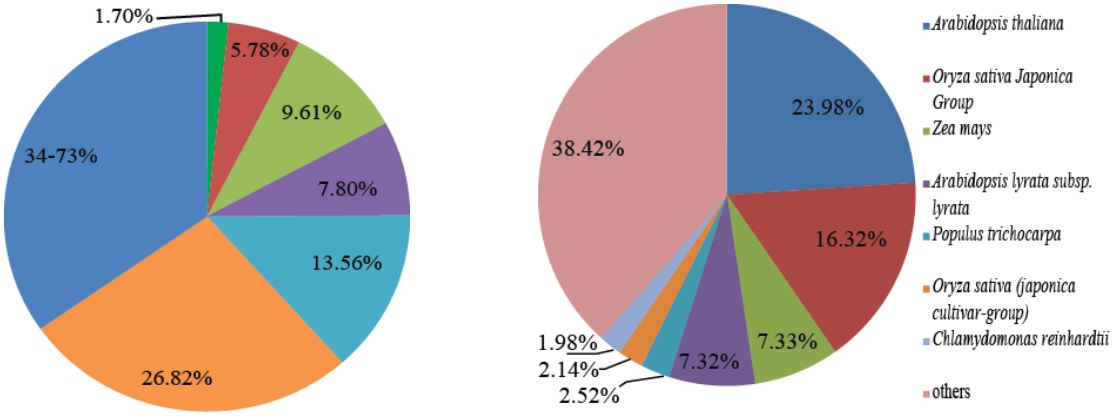


Fig. 2. The length distribution of all unigenes.



A= E-value distribution of BLAST hits for each unigene sequence.

B= Species distribution of the all unigenes sequences matching the top BLAST hits in Nr databases.

Fig. 3. Outcome of homology search of unigenes to Nr databases.

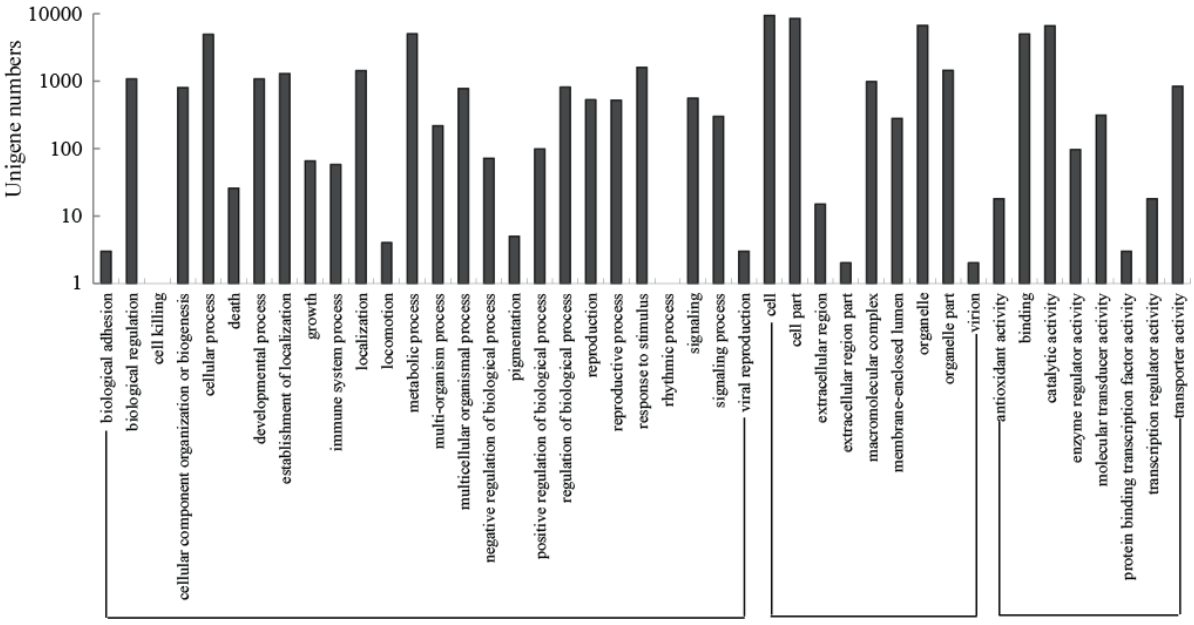


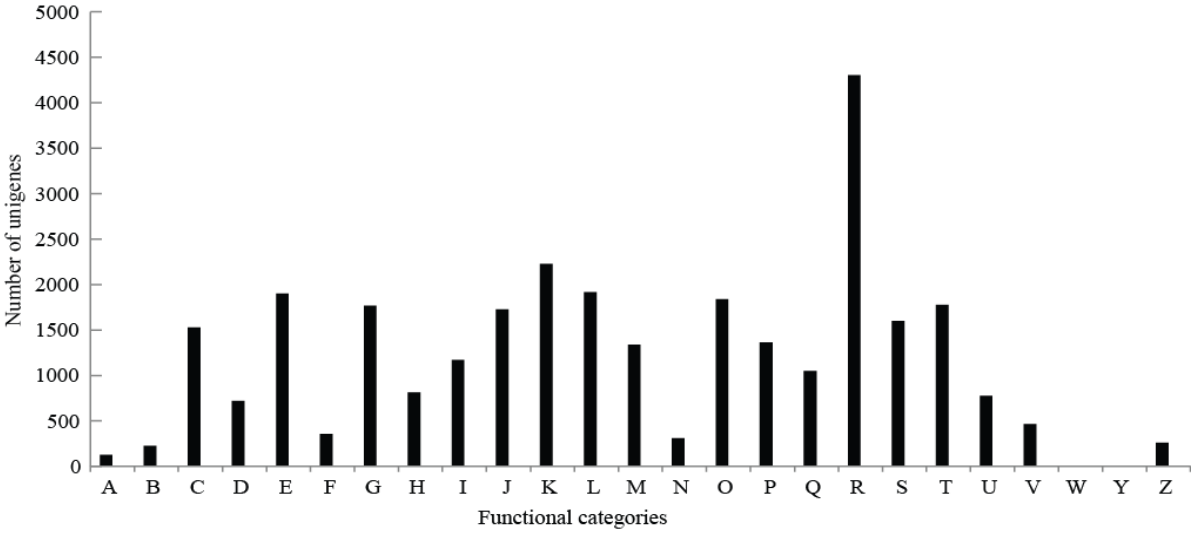
Fig. 4. The histogram of Gene Ontology classifications and analysis of transcripts. The unigenes were summarized in three main classifications: biological process, cellular component and molecular function.

In total, 13,973 unigenes (20.94%) were allocated to GO classes and described by 61,997 functional terms. Of these, assignments to the cellular component represented the majority (27,529 unigenes, 44.40%), followed by biological processes (21,471 unigenes, 34.63%) and molecular function (12,997 unigenes, 20.96%) (Table 3 and Fig. 4). Cell (9,492 unigenes, 67.93%), metabolic process (5,101 unigenes, 36.51%) and catalytic activity (6,681 unigenes, 47.81%) were the dominant subclasses in the cellular component, biological process and molecular function categories, respectively. Moreover, cell part (8,530 unigenes, 61.05%), organelle (6,759 unigenes, 48.37%), binding (5,019 unigenes, 35.92%) and cellular process (4,953 unigenes, 35.45%) were also well represented, suggesting that many novel genes extensively involved in metabolic activities could be play important roles in adaptation of ferns to temperature stress. However, few genes were assigned to the categories biological adhesion (3 unigenes), protein binding transcription factor activity (3 unigenes), extracellular region part (2 unigenes), virion (2 unigenes), cell killing (1 unigene) and rhythmic process (1 unigene).

After accounting for COG databases, 29,598 unigenes (44.36%) were allocated to 25 COG classifications (Fig. 5). The cluster of general function prediction only represented the largest group (4,305 unigenes, 14.54%), followed by transcription (2,229 unigenes, 7.53%), replication, recombination and repair (1,918 unigenes, 6.48%), amino acid transport and metabolism (1,901 unigenes, 6.42%) and posttranslational modification, protein turnover and chaperones (1,841 unigenes, 6.22%). However, only a few unigenes were allocated to RNA processing and modification, nuclear structure and extracellular structures. It was noteworthy that 1,051 unigenes (3.55%) were assigned to the group of secondary metabolites biosynthesis, transport and catabolism, these results should stimulate future research to elucidate the interrelationships and functions of these genes in *D. fragrans*.

Metabolic pathway analysis by KEGG: The activities and interactions of proteins at the molecular level within particular functional or structural networks have been uncovered using pathway databases. In this research, 18,989 sequences (28.46%) of the 66,716 non-redundant unigenes were allocated to 118 KEGG pathways (Table 3 and Additional file 1) and other sequences were assigned to 37 pathways, each with over 200 unigenes assigned. Most unigenes (5,012, 26.39%) were allocated to metabolic pathways, while the second largest group was assigned to biosynthesis of secondary metabolites (2,788, 14.68%). Remaining sequences received assignments such as spliceosome (1,384, 7.29%), plant-pathogen interaction (1,128 5.94%), protein processing in the endoplasmic reticulum (573, 3.02%), purine metabolism (483, 2.54%), phenylpropanoid biosynthesis (462, 2.43%), starch and sucrose metabolism (461, 2.43%), ribosome (451, 2.38%) and pyrimidine metabolism (406, 2.14%). However, no more than 10 unigenes were assigned to glycosphingolipid biosynthesis-ganglio series, anthocyanin biosynthesis and betalain biosynthesis. These results suggested that under temperature stress, secondary metabolic processes were active pathways in *D. fragrans*.

The candidate genes involved in flavonoids biosynthesis and transport: Among many categories of secondary metabolites in *D. fragrans,* flavonoids are one of the most vital medicinal groups. Moreover, it is known that flavonoid content can significantly influence the efficacy of fern-derived medicinal treatments. Therefore, due to great research interest, the synthetic pathway of flavonoids has been clearly elucidated in some crops and plants such as *Zea mays* and *Arabidopsis thaliana.* In these plants, the major classes of flavonoids synthesis pathways produce flavones, anthocyanins, flavonols, proanthocyanidins, phlohaphenes and aurones. However, the overall molecular mechanisms of flavonoids biosynthesis and their accumulation in *D. fragrans* are not fully understood.



A= RNA processing and modification (127); **B=** Chromatin structure and dynamics (227); **C=** Energy production and conversion (1529); **D=** Cell cycle control, cell division, chromosome partitioning (720); **E=** Amino acid transport and metabolism (1901); **F=**Nucleotide transport and metabolism (356); **G=** Carbohydrate transport and metabolism (1768); **H=** Coenzyme transport and metabolism (815); **I=** Lipid transport and metabolism (1172); **J=** Translation, ribosomal structure and biogenesis (1729); **K=** Transcription (2229); **L=** Replication, recombination and repair (1918); **M=** Cell wall/membrane/envelope biogenesis (1341); **N=** Cell motility (311); **O=** Posttranslational modification, protein turnover, chaperones (1841); **P=** Inorganic ion transport and metabolism (1365); **Q=** Secondary metabolites biosynthesis, transport and catabolism (1051); **R=** General function prediction only (4305); **S=** Function unknown (1602); **T=** Signal transduction mechanisms (1776); **U=** Intracellular trafficking, secretion, and vesicular transport (778); **V=** Defense mechanisms (468); **W=** Extracellular structures (4); **Y=** Nuclear structure (3); **Z=** Cytoskeleton (262).

Fig. 5. Histogram presentation of orthologous groups (COG) classification.

Table 3. Statistics of all Unigene annotations.

Database	Annotated (n)	Percentage (%)
Nr	35,458	53.1
Swiss-Port	28,194	42.3
KEGG	18,989	28.7
COG	17,327	26.0
GO	13,973	21.0
Total	38,486	57.7

Table 4. The numbers of candidate genes involved in flavonoid biosynthesis and transport in *D. fragrans* transcriptome.

Family	Number of annotated sequences
Phenylalanine ammonia lyase (PAL)	21
Cinnamate 4-hydroxylase (C4H)	10
4-coumarate CoA ligase (4CL)	31
Chalcone synthase (CHS)	24
Chalcone isomerase (CHI)	3
Dihydroflavonol 4-reductase (DFR)	6
Flavanone 3-hydroxylase (F3H)	1
Flavonoid 3′-hydroxylase (F3′H)	4
Flavonoid 3′, 5′-hydroxylase (F3′5′H)	7
O-methyltransferase (OMT)	3
Leucoanthocyanidin dioxygenase (LDOX)	1
Anthocyanidin reductase (ANR)	2
MYB	25
basic helix–loop–helix (bHLH)	23
WD40	43

As shown in Table 4, which displayed the annotated *D. fragrans* transcriptome, most known enzymes involved in flavonoid synthesis were found. First, phenylalanines are converted to chalcones via the phenylpropanoid pathway by the enzymes phenylalanine ammonia lyase (PAL, 21 unigenes), cinnamate 4-hydroxylase (C4H, 10 unigenes), 4-coumarate CoA ligase (4CL, 31 unigenes) and chalcone synthase (CHS, 24 unigenes). Subsequently, chalcone synthase catalyzes conversion of 4-coumaroyl CoA to naringenin chalcone or to aurones. Chalcone isomerase (CHI, 3 unigenes) catalyzes conversion of naringenin chalcone to naringenin or to flavones. Dihydroflavonol 4-reductase (DFR, 6 unigenes) catalyzes conversion of naringenin to phlohaphenes; naringenin is converted to dihydroflavonols by flavanone 3-hydroxylase (F3H, 1 unigene), flavonoid 3′-hydroxylase (F3′H, 4 unigenes) and flavonoid 3′, 5′-hydroxylase (F3′5′H, 7 unigenes). O-methyltransferase (OMT, 3 unigenes) catalyzes conversion of dihydroflavonols to flavonol. Dihydroflavonols are converted to leucoanthocyanidins by DFR. Leucoanthocyanidin dioxygenase (LDOX, 1 unigene) catalyzes conversion of leucoanthocyanidins to pelargonidin. Anthocyanidin reductase (ANR, 2 unigenes) catalyzes conversion of pelargonidin to proanthocyanins, while OMT catalyzes conversion of pelargonidin to anthocyanins (Fig. 6). We didn't identify the genes encoding flavonol synthase (*FLS*), anthocyanidin synthase (*ANS*) and leucopelargonidin-4-reductase (*LAR*), probably because they were not expressed at the start of transcription or not present in *D. fragrans*.

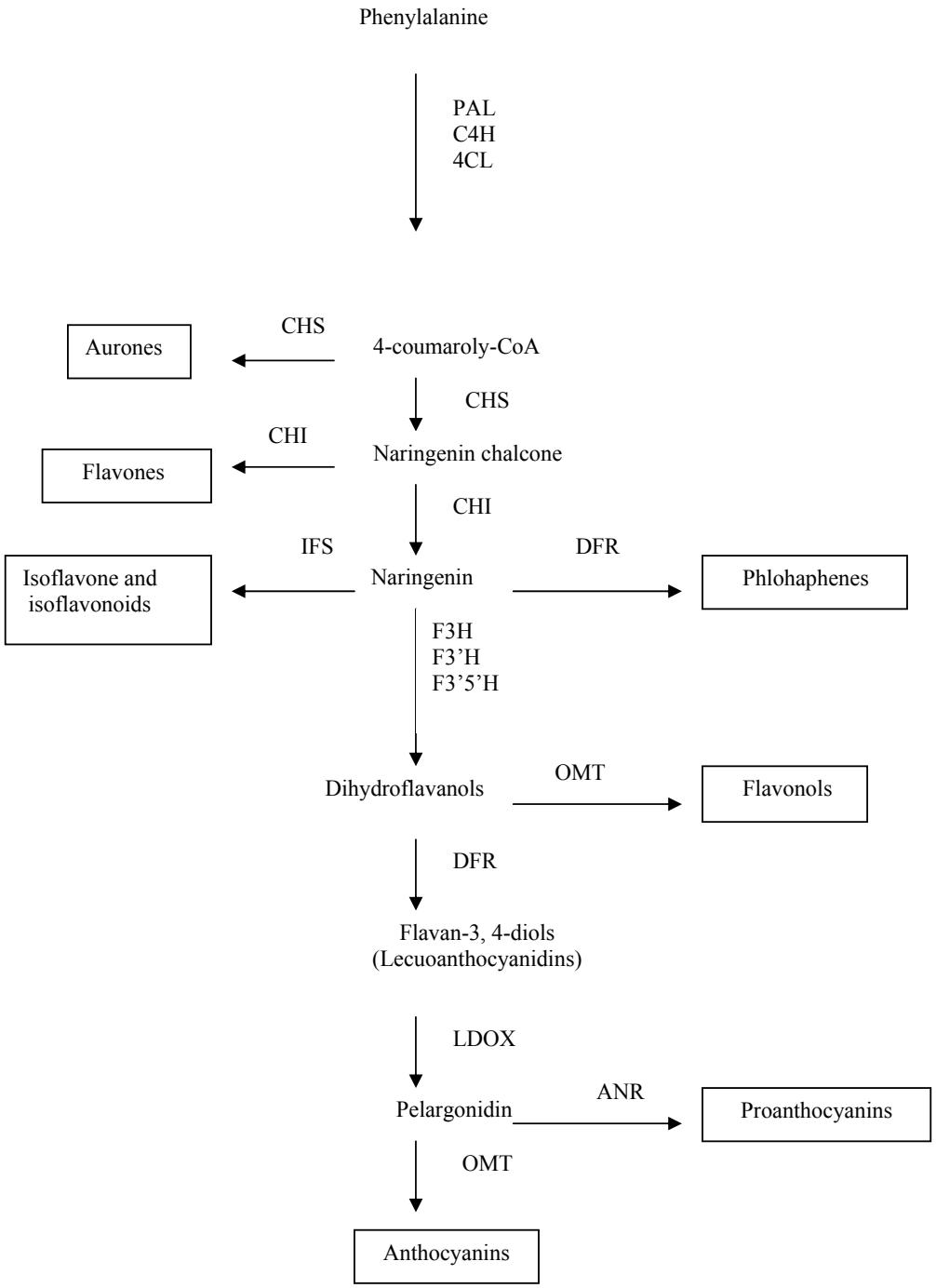


Fig. 6. The flavonoid biosynthetic pathway diagram for *Dryopteris fragran*s (L.) Schott.

The modulation of expression of plant flavonoid biosynthesis is controlled by two types of genes. Transcription factors comprise one type of control gene and include the WD-repeat-containing proteins, MYB proteins and basic helix–loop–helix (bHLH) transcription factors (TFs). WD-repeat proteins have four or more copies of the WD (tryptophan-aspartate) repeat and they also comprise a large gene family which coordinates protein interactions (De Vetten *et al.*, 1997). MYB TFs and bHLH TFs play a part in controlling plant phenylpropanoid secondary metabolism and each belongs to multiple gene families (Stracke *et al.*, 2001; Jones, 2004; Czemmel *et al.*, 2012). These transcription factors could control and regulate the final steps of the flavonoid biosynthetic pathway through formation of ternary complexes of MYB-bHLH-WD40 (MBW) (Ramsay & Glover, 2005). Based on our annotated transcription data, a total of 91 unigenes were identified as MYB (25 unigenes), bHLH (23 unigenes) and WD40 (43 unigenes). MYB proteins are characterized by four domain types. Of

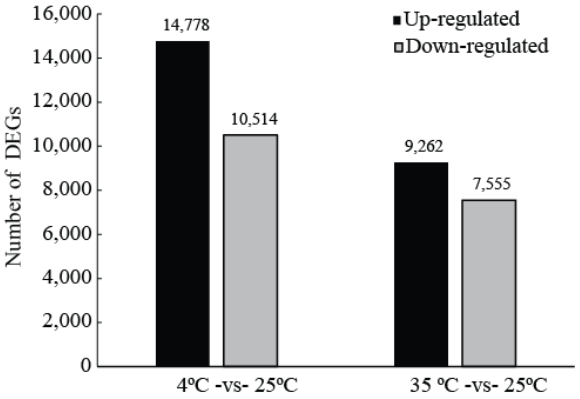
particular note, we identified 10 genes annotated as *R2R3-type MYB* and *R2R3-type MYB* genes, which are known to play a role in plant secondary metabolism**.** These results suggested that under temperature stress, genes which encode such secondary metabolite processes played important roles in *D. fragrans*.

Differential gene expression between low and high temperature stress: Statistically marked expression differences of DEGs under temperature stress might be illustrated by comparing the expression profiles for 4°C and 35°C samples *vs.* the 25°C control. The false discovery rates (FDR) control serves as a threshold to evaluate the statistical significance of DEGs and we chose a cutoff of an FDR $\leq 10^{-3}$ and analyzed DEGs with an expression ratio larger than 2. From our data sets, 42,109 genes were significantly variably expressed in the pair-wise comparisons between T1 (25°C to 4°C) and T2 (25°C to 35°C) and the CK library. 25,292 DEGs (14,778 up-regulated/10,514 down-regulated) were observed between T1 and CK libraries. 16,817 DEGs (9,262 up-regulated/7,555 down-regulated) were observed between T2 and CK libraries (Fig. 7A). More genes were highly expressed in T1 library, suggesting that more genes were involved in formation of complex metabolites during the low temperature stress of *D. fragrans*. Moreover, of all the DEGs, 7,191 genes were significantly differentially expressed in both T1 and T2 libraries (Fig. 7B).

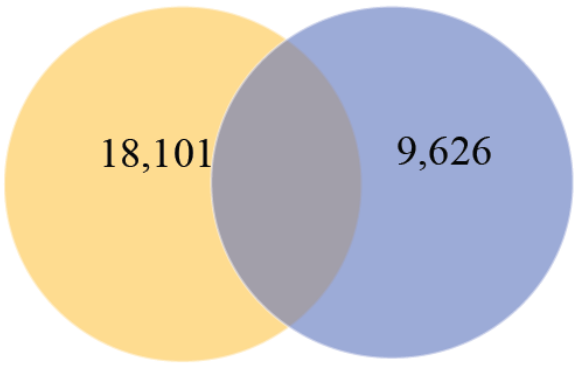
Comparison of DEGs for 4°C *vs.* 25°C generated a total of 20,134 DEGs that were assigned to 36 significantly enriched GO terms, of which the cellular component term predominated (8,812, 43.77%), followed by biological process (6,913, 34.33%) and molecular function (4,409, 21.90%) (Fig. 8A). The most abundant GO terms were cell (3,161), metabolic process (1,677) and catalytic activity (2,384) in the cellular component, biological process and molecular function categories, respectively. Comparison of DEGs for 35°C *vs.* 25°C generated a total of 15,306 DEGs that were assigned to 38 significantly enriched GO terms, of which the cellular component predominated (6,814, 44.52%), followed by biological process (5,190, 33.91%) and molecular function (3,302, 21.57%) (Fig. 8B). The most abundant GO terms were cell (2,380), metabolic process (1,236) and catalytic activity (1,730) in the cellular component, biological process and molecular function categories, respectively. Moreover, some interesting terms associated with flavonoids biosynthesis and transport, such as metabolic process, regulation of biological process and transporter activity, were observed.

Of 6,826 DEGs with pathway annotations in 4°C *vs.* 25°C, 40 pathways were enriched significantly in DEGs (Additional file 2). Of these, metabolic pathways made up the most significant representation (2,151, 31.51%), followed by biosynthesis of secondary metabolites (1,233, 18.06%) and valine, leucine and isoleucine degradation (168, 2.46%). Phenylalanine metabolism (137, 2.01%), flavonoid biosynthesis (120, 1.76%) and flavone and flavonol biosynthesis (49, 0.72%) were also enriched. Of 5,182 DEGs with pathway annotations in 35°C *vs.* 25°C, 29 pathways were enriched significantly in DEGs (Additional file 3). Of these, phenylpropanoid biosynthesis made up the most significant portion (215, 4.15%), followed by flavonoid biosynthesis (149, 2.88%) and biosynthesis of secondary metabolites (944, 18.22%). Flavone and flavonol biosynthesis (68, 1.31%) and phenylalanine metabolism (106, 2.01%) were also enriched significantly. These patterns could reflect a higher specialization in partitioning of secondary metabolic intermediates towards creation of different classes of end derivatives which accumulate in the same organ during different kinds of temperature stress during flavonoids biosynthesis and transport in *D. fragrans.*

Regulation of genes involved in flavonoid biosynthesis: The transcriptome dataset contained several flavonoid biosynthesis genes, suggestive of high coverage. Many of these genes appeared to have several family members, implying that one or more rounds of genome duplication might have occurred during *D. fragrans* genome evolution. Some genes involved in flavonoid biosynthesis identified here have high similarity to known flavonoid biosynthesis genes in *D. fragrans*, including *PAL*, *C4H*, *CHS* and *4CL* (Li *et al.*, 2015b; Sun *et al.*, 2014; Li *et al.*, 2015a). Because some genes identified here showed low similarity to known genes, this result suggested that there were novel putative genes contributing to this pathway in this transcriptome.

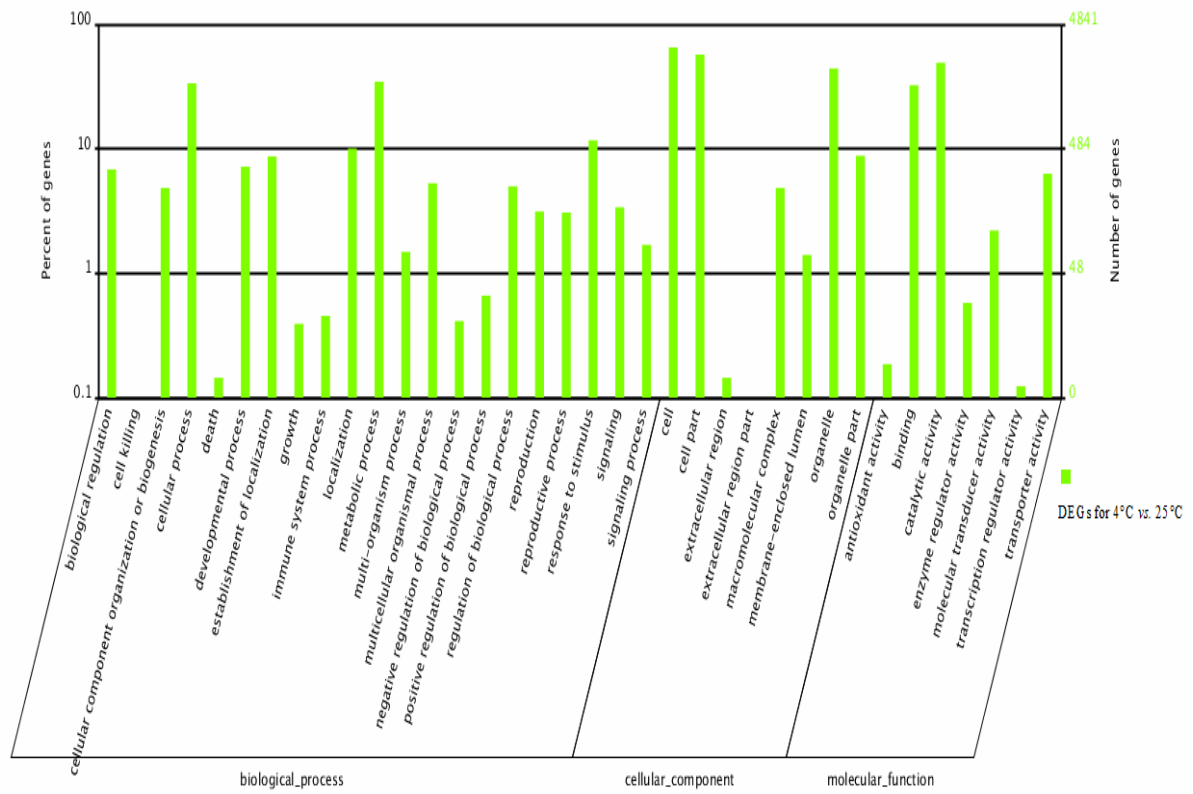


A= The numbers of up- and down-regulated genes in comparisons of 4ºC *vs.* 25ºC and 35 ºC *vs.* 25ºC two temperature stress libraries.

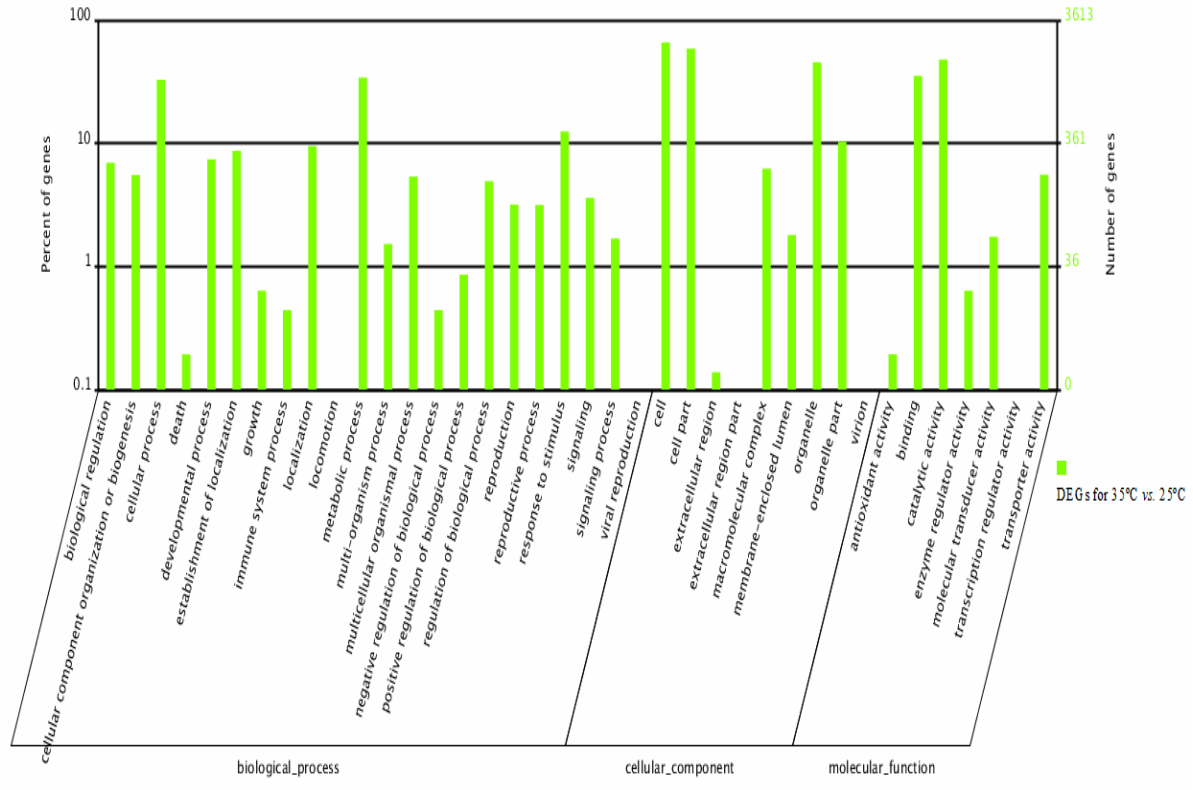


B= Venn diagram showing the comparison of DEGs between two temperature stress libraries of 4ºC *vs.* 25ºC and 35 ºC *vs.* 25ºC in *Dryopteris fragran*s (L.) Schott.
http://bioinfogp.cnb.csic.es/tools/venny_old/index.html

Fig. 7. The analysis of DEGs under different temperature treatments in *Dryopteris fragran*s (L.) Schott.



A= Comparison of DEGs for 4°C *vs.* 25°C



B= Comparison of DEGs for 35ºC *vs.* 25ºC

Fig. 8. GO function enrichment analysis of DEGs.

Expression difference analysis of unigenes showed numerous genes exhibiting differential expression profiles after a temperature shift from 25°C to 4°C (CK *vs.* T1). Although *C4H* genes were neither up- nor down-regulated, many other genes did show expression changes: a total of 15 *PAL* genes (10 up-regulated/5 down-regulated), 20 *4CL* genes (17 up-regulated/3 down-regulated) and 15 *CHS* genes (12 up-regulated/3 down-regulated). Expression changes were also noted in a total of 6 *PAL* genes (0 up-regulated/6 down-regulated), 2 *C4H* genes (1 up-regulated/1 down-regulated), 12 *4CL* genes (8 up-regulated/4 down-regulated) and 18 *CHS* genes (0 up-regulated/18 down-regulated) after a temperature shift from 25°C to 35°C (CK *vs.* T2). After this temperature shift, *PAL* exhibited a slight change in expression level, *C4H* exhibited a very slight change in expression level and *CHS* and *4CL* exhibited large expression level changes. Notably, PAL catalyzes the initial step of the phenylpropanoid pathway and is a rate-limiting enzyme step in the pathway (Bate *et al.*, 1994), while CHS and C4H catalyze the first and second committed step of flavonoids biosynthesis in the phenylpropanoid metabolic pathway (Hahlbrock & Scheel, 1989). Moreover, 4CL catalyzes the conversion of hydroxycinnamates into corresponding CoA esters for biosynthesis of flavonoids and lignins. Recently, research focusing on *4CL*, *PAL* and *C4H* genes produced similar results as observed in our study (Li *et al.*, 2015a; Li *et al.*, 2015b). While the increased expression of *CHS* genes in this study did not correlate precisely with temperature effects observed in other studies (Sun *et al.*, 2014), this could be due to the fact that *CHS* genes belong to a multigene family (Sommer & Saedler, 1986) and some of these genes are very conserved.

DFR, LDOX and ANR are enzymes that act late in the flavonoid synthetic pathway. DFR catalyzes a rate-limiting enzymatic step in anthocyanin biosynthesis and also plays a vital role in rutin, tannin and biosynthesis of other secondary metabolites. LDOX/ANS catalyzes conversion of leucocyanidin into anthocyanin. LAR and ANR are active in two different metabolic biosynthetic pathways which produce proanthocyanidins. We identified *D. fragrans* genes encoding these enzymes and unigenes expression difference analysis showed that *ANR* genes were neither up- nor down-regulated, while a total of 3 *DFR* genes (1 up-regulated/2 down-regulated) and 1 *LDOX* gene (1 up-regulated) exhibited expression changes when the temperature was shifted from 25°C to 4°C (CK *vs.* T1). However, the *LDOX* gene was not up- or down-regulated and a change in expression of 1 *DFR* gene (1 down-regulated) and 1 *ANR* gene (1 down-regulated) were detected when the temperature was shifted from 25°C to 35°C (CK *vs*. T2), although these expression level changes were modest.

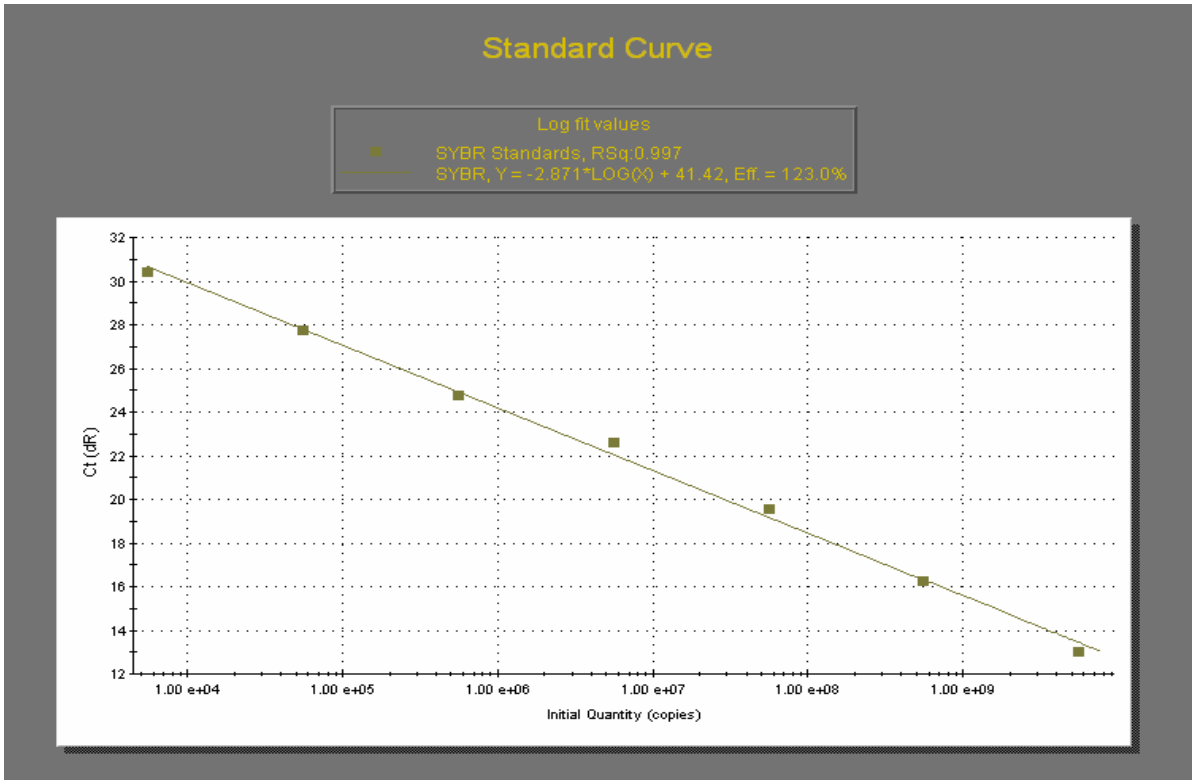
Three kinds of regulatory genes were identified that all take part in flavonoid biosynthesis and transport. Unigenes expression difference analysis showed that a total of 5 *WD40* gene profiles (1 up-regulated/4 down-regulated), 8 *MYB* (2 up-regulated/6 down-regulated) and 9 *bHLH* genes (1 up-regulated/8 down-regulated) were detected with differences in expression at 25°C *vs.* 4°C (CK *vs*. T1). A total of 5 *WD40* gene profiles (3 up-regulated/1 down-regulated), 7 *MYB* (1 up-regulated/6 down-regulated) and 9 *bHLH* genes (2 up-regulated/7 down-regulated) were observed at 25°C *vs.* 35°C (CK *vs.* T2), respectively. Therefore, after temperature treatments, these three kinds of regulatory genes exhibited down-regulated expression levels from 25°C to 4°C (CK *vs*. T1) and different expression levels from 25°C to 35°C (CK *vs*. T2).

Quantitative RT-PCR analysis of *4CL* and *CHS* genes expression patterns: Production of flavonoids, vital natural secondary metabolites, is a pivotal area of investigation needed before performing more intensive exploration of their medicinal value. Based on comparison of sequence data to public databases, a total of 66,716 non-redundant unigenes were assembled and 38,486 unigenes were annotated using matched database genes. By searching the 38,486 annotated unigenes for key flavonoids biosynthesis enzymes, *4CL* and *CHS* genes were represented at the highest frequency. Among these 31 *4CL* genes and 24 *CHS* genes, 5 *4CL* genes and 4 *CHS* genes have homology to known *4CL* and *CHS* genes, respectively. To validate the data, qRT-PCR assays were performed on *Df4CL* and *DfCHS* genes involved in *D. fragrans* cultured seedlings under different temperature treatments.

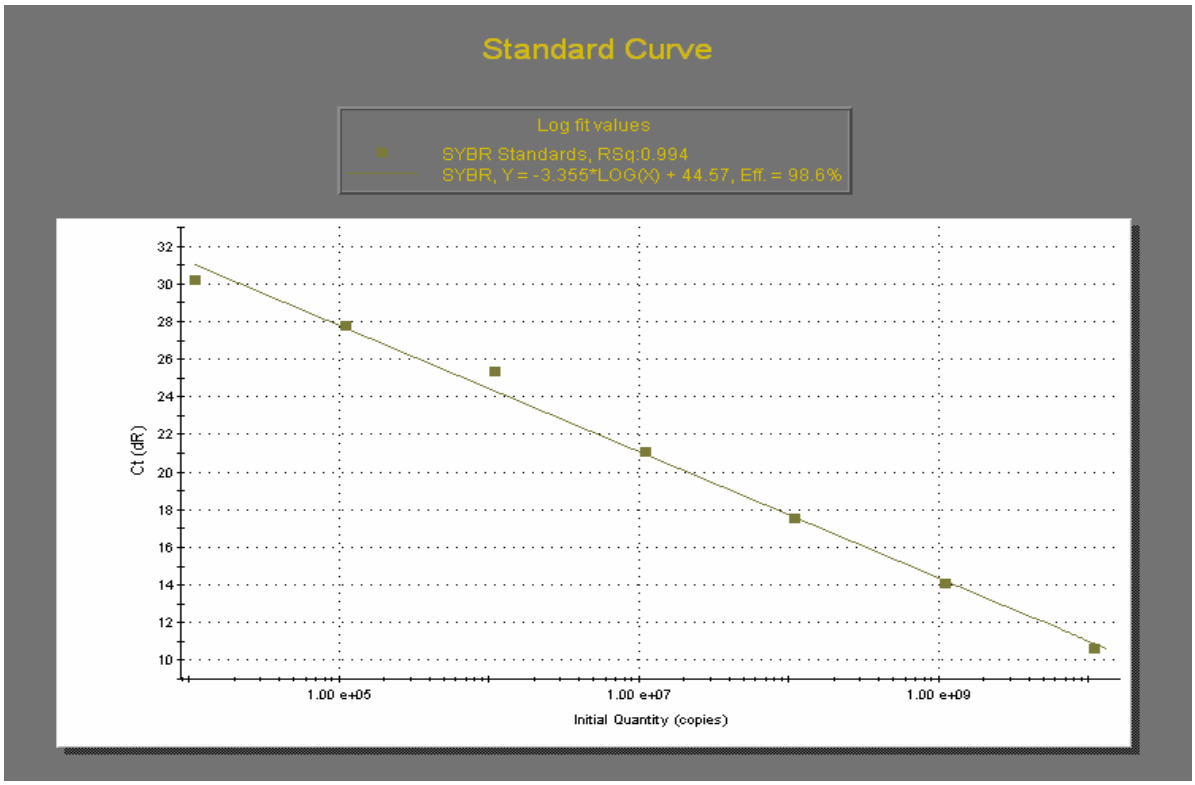
E-values were used to evaluate amplification efficiency. The standard curve of standard gene *Df4CL* was C_T= -2.871*LOG(X) +41.42, R=0.997. The standard curve of standard gene *DfCHS* was C_T= -3.553*LOG(X) +44.57, R=0.994. Thus, the standard curves of PCR products created using *Df4CL* and *DfCHS* primers, which were designed from the *de novo* transcriptome sequencing data, exhibited good linear relationships, and exhibited E-values between 0.9 and 1.1 (Fig. 9).

The expression levels of *Df4CL* by qRT-PCR analysis showed that after 48 h under the 4°C and 35°C conditions, *Df4CL* was up- and down-regulated after the treatments, respectively (Fig. 10A). Comparison with the measurement of total flavonoid content and *Df4CL* unigenes expression difference analysis, they all exhibited up-regulated expression level under the 4°C condition and these suggested that *Df4CL* was sensitive to low temperature treatment. Under the 35°C condition, the total flavonoid content and the unigenes expression difference analysis exhibited up-regulated expression level, but the expression levels of *Df4CL* was opposite. These suggested that *Df4CL* could have an adaptive processes accumulate total flavonoid content to the high temperature treatment (Fig. 10A).

The expression levels of *DfCHS* by qRT-PCR analysis showed that after 48 h under the 4°C and 35°C conditions, *DfCHS* was down-regulated after treatment (Fig. 10B). Comparison with the measurement of total flavonoid content and *DfCHS* unigenes expression difference analysis, the total flavonoid content and unigenes expression difference analysis exhibited up-regulated expression level, but the expression levels of *DfCHS* was down-regulated under the 4°C condition. Under the 35°C condition, they all exhibited down-regulated during the whole process from 0 h to 48 h. These suggested that could be a complex relationship between *DfCHS* and total flavonoid content under temperature treatment, As we known, CHS is the first key enzyme in flavonoids biosynthesis and the regulation could not obvious (Winkel-Shirley, 2002). When the total flavonoid content was highest, and the expression level of *DfCHS* was down-regulated and this is also same as previous research (Sun *et al.*, 2014).

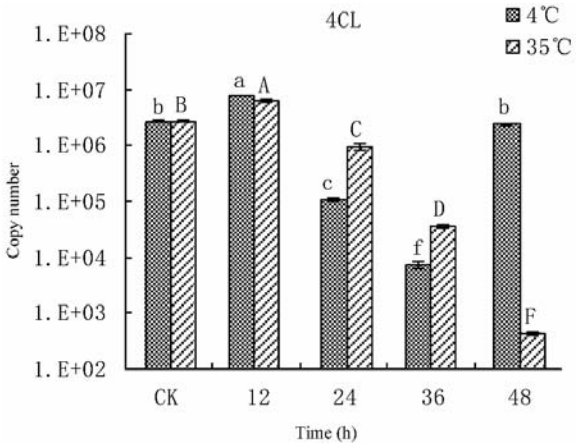


A= The standard curves of *Df4CL*. C_T= -2.871*LOG(X) +41.42，R=0.997.

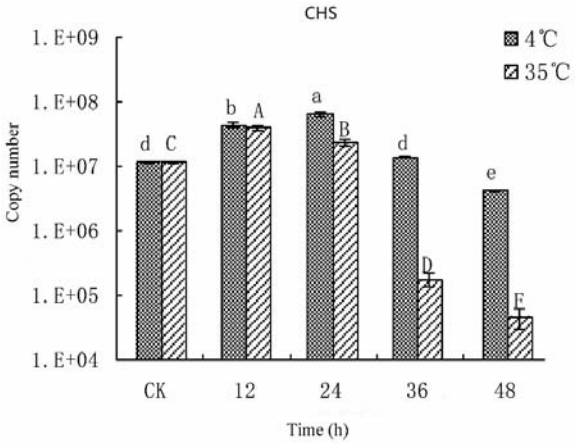


B= The standard curves of *DfCHS*. C_T= -3.553*LOG(X) +44.57，R=0.994

Fig. 9. The standard curves.



A= The level of transcripts of *Df4CL* under different temperature stress



B= The level of transcripts of *DfCHS* under different temperature stress

Fig. 10. Quantitative real-time polymerase chain reaction analyses of *Df4CL* and *DfCHS* in *Dryopteris fragran*s (L.) Schott under different temperature stress.
Different letters indicate significant differences in gene expression as calculated by Duncan statistical analysis ($p \leq 0.05$).

Discussion

Characterization of the *D. fragrans* transcriptome: Now no genomic research about *D. fragrans* and only a few focused on extraction, purification and molecular biology. Thus more molecular genomic data will be useful for researches studying. High transcriptome sequencing technology could provide more information about genes and help us understand functional genomic regulations and important bioprocess molecular mechanisms in model and non-model plants. In this study, a *de novo* transcriptome of the different temperature stress of *D. fragrans* tissues culture seedlings were assembled based. In total, 66,716 genes were newly discovered with a median length of 524 bp and similar with *Pteridium aquilinum* with a median length of 547.2 bp (Der *et al.*, 2011). Approximately 57.7% of the unigenes were annotated using sequence comparisons to public databanks. 21.0%, 26.0% and 28.5% of the unigenes were classified by GO, COG and KEGG database, respectively. All these will improve the understanding of gene functions for important economic traits, exploring the flavonoids biosynthesis regulation and transport mechanisms in *D. fragrans*. At the same time all these indicated that 40% of the unigenes have no apparent homologs and they probably have new functions.

Potential candidate involved in flavonoids biosynthesis: The RNA-Seq analysis revealed that 42,109 DEGs were response to temperature stress. 7,191 genes were significantly differentially expressed both under the high and low temperature stress conditions, this indicated that the unigenes involved in flavonoids biosynthesis and transport might be specifically expressed in similar ways. Genes related with flavonoids biosynthesis and regulation have been extensively studied in many plants. However, only 3 *PAL*, 1 *C4H*, 1 *CHS* and 4 *4CL* genes have been identified based on the **R**ACE technology (Li *et al.*, 2015b; Sun *et al.*, 2014; Li *et al.*, 2015a) in *D. fragrans.* In this study, with the help of transcriptome analysis, 204 genes were involved in flavonoid biosynthesis, regulation and transport, including 21 *PAL* genes, 10 *C4H* genes, 31 *4CL* genes, 24 *CHS* genes, 3 *CHI* genes, 6 *DFR* genes, 1 *F3H* gene, 4 *F3'H* genes, 7 *F3'5'H* genes, 3 *OMT* genes, 1 *LDOX* gene, 2 *ANR* genes, 25 *MYB* genes, 23 *bHLH* genes and 43 *WD40* genes. Most of them were first identified in *D. fragrans* and indicated that they were possible due to the complicate progress of flavonoids biosynthesis and the further functions of them need to study. DEGs analysis showed under the low temperature treatment, *PAL*, *4CL*, *CHS* exhibited large expression level changes and under the high temperature treatments, *PAL* and *C4H* exhibited a very slight change, *CHS* and *4CL* exhibited large expression level changes. All these indicated that the different temperature stress don't have the same impact on flavonoids biosynthesis.

Notably, PAL catalyzes the initial step of the phenylpropanoid pathway and is a rate-limiting enzyme step in the pathway (Bate *et al.*, 1994), while CHS and C4H catalyze the first and second committed step of flavonoids biosynthesis in the phenylpropanoid metabolic pathway (Hahlbrock & Scheel, 1989). Moreover, 4CL catalyzes the conversion of hydroxycinnamates into corresponding CoA esters for biosynthesis of flavonoids and lignins. Recently, research focusing on *4CL*, *PAL* and *C4H* genes produced similar results as observed in our study (Li *et al.*, 2015a; Li *et al.*, 2015b). While the increased expression of *CHS* genes in this study did not correlate precisely with temperature effects observed in other studies (Sun *et al.*, 2014), this could be due to the fact that *CHS* genes belong to a multigene family (Sommer & Saedler, 1986) and some of these genes are very conserved. Among these genes, 5 *4CL* genes and 4 *CHS* genes have homology to known *4CL* and *CHS* genes, respectively. Their expression levels showed sensitive to temperature treatment, this indicated that these genes are responsible for the flavonoids biosynthesis and the relationship are very complex.

Therefore, the identification of these genes will be accelerate the further understanding and the mechanisms of flavonoids biosynthesis and regulation.

Conclusions

In our study, the transcriptome data of *D. fragrans* was first characterized by *de novo* sequencing using the Illumina platform without relying on guidance from a reference genome. After subjecting the organism to low and high temperature stresses, the cellular component annotation was assigned to the majority of differentially expressed genes (43.77%), followed by biological process (34.33%) and molecular function genes (21.57%). Further analysis of the transcriptome provided information regarding the molecular mechanisms of flavonoid biosynthesis and transport underlying adaptation to temperature stress. The results demonstrated that most known structural and regulatory genes involved in flavonoid biosynthesis and transport were found to be regulated by temperature. Moreover, the genes encoding enzymes FLS, ANS and LAR were not found, indicating that flavonoid biosynthesis and transport were distinct in *D. fragrans*. In addition, several genes were identified here for the first time and possess low similarity to known genes; these could be as yet unknown genes contributing to the flavonoids biosynthetic pathway in this transcriptome. Our study provides the largest number of unigenes to date and lays the groundwork for in-depth transcriptomic profiling of *D. fragrans***.** Furthermore, this research provides useful data for future functional genomic research of this fern species and will likely facilitate future research of other closely related species toward development of improved methods for isolating medicinal flavonoids from ferns, including *D. fragrans*.

Acknowledgements

This research was supported by China National Natural Science Foundation (Project 31070291 and J1210069).

References

AI-Dous, E.K., B. George, M.E. AI-Mahmoud, M.Y. AI-Jaber, H. Wang, Y.M. Salameh, E.K. AI-Azwani, S. Chaluvadi, A.C. Pontaroli, J. DeBarry, V. Arondel, J. Ohlrogge, I.J. Saie, K.M.S. Elmeer, J.L. Bennetzen, R.R. Kruegger and J.A. Malek. 2011. De novo genome sequencing and comparative genomics of date palm (*Phoenix dactylifera*). *Nat. Biotechnol.*, 29(6): 521-528.

Ao, Z.W. and G.F. Li. 1990. *Ferns of Heilongjiang Province*. Northeast Forestry University Press, Harbin.

Audic, S. and J. Claverie. 1997. The significance of digital gene expression profiles. *Genome Res*., 7(10): 986-95.

Banerjee, R. D. 1980. Antibiotic activity of *Pteridophyte*. *Econ. Bot.,* 34(3): 284-293.

Bate, N.J., J. Orr, W. Ni, A. Meromi, T.N. Hassar, P.W. Doermer, R.A. Dixon, C.J. Lamb and Y. Elkind. 1994. Quantitative relationship between phenylalanine ammonia-lyase levels and phenylpropanoid accumulation in transgenic tobacco identifies a rate-determining step in natural product synthesis. *Proc. Natl. Acad. Sci.,* 91: 7608-7612.

Conesa, A., S. Götz, J.M.G. Gómez, J. Terol, M. Talón and M. Robles. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, 21(18): 3674-3676.

Czemmel, S., S.C. Heppel and J. Bogs. 2012. R2R3 MYB transcription factors: Key regulators of the flavonoid biosynthetic pathway in grapevine. *Protoplasma*, 249: 109-118.

De Vetten, N., F. Quattrocchio, J. Mol and R. Koes. 1997. The an11 locus controlling flower pigmentation in petunia encodes a novel WD-repeat protein conserved in yeast, plants, and animals. *Gen Dev.,* 11: 1422-1434.

Der, J.P., M.S. Barker, N.J. Wickett, C.W. de Pamphilis and P.G. Wolf. 2011. De novo characterization of the gametophyte transcriptome in bracken fern, *Pteridium aquilinum*. *BMC Genomics*, 12: 99.

Erhart, L.M., B.L. Buttgereit, H. Schmidt, U. Wenzel, H. Daniel and R. Göke. 2005. Flavone initiates a hierarchical activation of the capase-cascade in colon cancer cells. *Apoptosis*, 10(3): 611-617.

Gao. R., Y. T. Liang and Y. Chang. 2014. Gene cloning and sequence analysis of psbA in *Dryopteris fragrans* (L.) Schott. *North. Hort.*, 19: 98-102.

Goff, S.A., D. Ricke, T.H. Lan, G. Presting, R.L. Wang, M. Dunn, J. Glazebrook, A. Sessions, P. Oeller, H. Varma, D. Hadley, D. Hutchison, C. Martin, F. Katagiri, B.M. Lange, T. Moughamer, Y. Xia, P. Budworth, J.P. Zhong, T. Miguel, U. Paszkowski, S. Zhang, M. Colbert, W. L. Sun, L.L. Chen, B. Cooper, S. Park, T.C. Wood, L. Mao, P. Quail, R. Wing, R. Dean, Y. Yu, A. Zharkikh, R. Shen, S. Sahasrabudhe, A. Thomas, R. Cannings, A. Gutin, D. Pruss, J. Reid, S. Tavtigian, J. Mitchell, G. Eldrege, T. Scholl, R.M. Miller, S. Bhatnagar, N. Adey, T. Rubano, N. Tusneem, R. Robinson, J. Feldhaus, T. Macalma, A. Oliphant and S. Briggs. 2002. A Draft Sequence of the Rice Genome (*Oryza sativa* L. ssp. *japonica*). *Sci*., 296(5): 92-100.

Graf, B.A., P.E. Milbury and J. B. Blumberg. 2005. Flavonols, flavones, flavanones and human health: epidemiological evidence. *J. Med. Food.*, 8(3): 281-290.

Hahlbrock, K. and D. Scheel. 1989. Physiology and molecular biology of phenylpropanoid metabolism. *Annu. Rev. Plant Physiol. Plant Mol. Biol.,* 40: 347-369.

Huang, Q.Y., W.H. Li, R.F. Fan and Y. Chang. 2014. New MADS-box gene in fern: cloning and expression analysis of DfMADS1 from *Dryopteris fragrans*. *PLoS One*, 9: e86349.

Jones, S. 2004. An overview of the basic helix-loop-helix proteins. *Genome Biol*., 5(6): 226.

Li, H.Z., Z.B. Shen, Y. Zhao and L.D. Zhang. 2005. Therapeutic effect of *Dryopteris Fragrans* ointment to rats with experimental tinea corporis. *J. Chinese Med. Materials.*, 28(10): 637-640.

Li, R., H. Zhu, J. Ruan, W. Qian, X. Fang, Z. Shi, Y. Li, G. Shan, K. Kristiansen, S. Li, H. Yang, J. Wang and J. Wang. 2010. De novo assembly of human genomes with massively parallel short read sequencing. *Genome Res*., 20: 265-272.

Li, S.S., Y. Li, L.L. Sun, B.Z. Hu and Y. Chang. 2015a. Identification and expression analysis of 4-Coumarate: Coenzyme A ligase gene family in *Dryopteris fragrans*. *Cell. Mol. Biol*., 61(4): 25-33.

Li, X.J., W. Wang, M. Luo, C.Y. Li, Y.G. Zu, P.S. Mu and Y.J. Fu. 2012. Solvent-free microwave extraction of essential oil from *Dryopteris fragrans* and evaluation of antioxidant activity. *Food Chem.*, 133: 437-444.

Li, Y. L.L. Sun, H.M. Wang, R. Gao, J.Z. Zhang, B.Z. Hu and Y. Chang. 2015b. Cloning and Expression Analysis of Phenylalanine Ammonia-Lyase (PAL) Gene Family and Cinnamate 4-Hydroxylase (C4H) From *Dryopteris Fragrans*. *Biologia*, 70(5): 606-614.

Müller, S.D., S.B. Vasconcelos, M. Coelho and M.W. Biavatti. 2005. LC and UV determination of flavonoids from *Passiflora alata* medicinal extracts and leaves. *J. Pharm. Biomed. Anal.*, 37(2): 399-403.

Paterson, A.H., J.E. Bowers, R. Bruggmann, I. Dubchak, J. Grimwood, H. Gundlach, G. Haberer, U. Hellsten, T. Mitros, A. Poliakov, J. Schmutz, M. Spannaql, H. Tang, X. Wang, T. Wicker, A.K. Bharti, J. Chapman, F.A. Feltus, U. Gowik, I. V. Grigorev, E. Lyons, C.A. Maher, M. Martis, A. Narechania, R.P. Otillar, B.W. Penning, A.A. Salamov, Y. Wang, L. Zhang, N. C. Carpita, M. Freeling, A.R. Gingle, C.T. Hash, B. Keller, P. Klein, S. Kresovich, M.C. McCann, R. Ming, D.G. Peterson, Mehboob-ur-Rahman, D. Ware, P. Westhoff, K.F. Mayer, J. Messing and D.S. Rokhsar. 2009. The Sorghum bicolor genome and the diversification of grasses. *Nat.*, 457(7229): 551-556.

Qi, Y.X., Y.B. Liu and W.H. Rong. 2011. RNA-Seq and its applications: a new technology for transcriptomics. *HEREDITAS* (*Beijing*), 33(11): 1191-1201.

Ramsay, N.A. and B.J. Glover. 2005. MYB-bHLH-WD40 protein complex and the evolution of cellular diversity. *Trends Plant Sci*., 10(2): 63-70.

Shen, Z.B., W.Y. Lou, Y.S. Yan and J.F. Zhu. 2006. Study on terpene of *Dryopteris fragrans* (L.). *Chinese herbal medicines*, 30(5): 448-449.

Sommer, H. and H. Saedler. 1986. Structure of the chalcone synthase gene of Antirrhinum majus. *Mol. Gen. Genet.*, 202: 429-434.

Stracke, R., M. Werber and B. Weisshaar. 2001. The R2R3-MYB gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol*., 4: 447-456.

Sun, L.L., Y. Li, S.S. Li, X.J. Wu, B.Z. Hu and Y. Chang. 2014. Identification and characterisation of DfCHS, a chalcone synthase gene regulated by temperature and ultraviolet in *Dryopteris fragrans*. *Cell Mol. Biol.*, 60(6): 1-7.

Sun, Y., F. Mu, C. Li, W. Wang, M. Luo, Y. Fu and Y. Zu. 2013. Aspidin BB, a phloroglucinol derivative, induces cell cycle arrest and apoptosis in human ovarian HO-8910 cells. *Chemico-biological interactions*, 204: 88-97.

Wang, J.Z. 1989. A taxonomic study on Pteridophyte in N.E. China. *Bull. Bot. Res.,* 9(4): 73-82.

Widen, C.J., C.R. Fraser-Jenkins, T. Reichstein and J. Sarvela. 2001. A survey of phenolic compounds in *Dryopteris* and related fern genera. *Ann. Bot. Fenn.,* 38(2): 99-138.

Winkel-Shirley, B. 2002. Biosynthesis of flavonoids and effects of stress. *Curr. Opin. Plant Biol.,* 5: 218-223.

Ye, J., L. Fang, H. Zheng, Y. Zhang, J. Chen, Z. Zhang, J. Wang, S. Li, R. Li, L. Bolund and J. Wang. 2006. WEGO: a web tool for plotting GO annotations. *Nucleic Acids Res*., 34: W293-297.

Zhang, C., J. Kan and Z. Chen. 2003. Research progress of bioflavonoid antitumous effect. *China Food Additives*, 83(3): 17-19.

Zhang, M., J. Cao, X. Dai, X. Chen and Q. Wang. 2012. Flavonoids contents and free radical scavenging activity of extracts from leaves, stems, rachis and roots of *Dryopteris erythrosora*. *Iranian J. Pharm. Res.*, 11: 991-997.

(Received for publication 17 May 2015)