Industrial Crops & Products 120 (2018) 47–60

Contents lists available at ScienceDirect

Industrial Crops & Products

journal homepage: www.elsevier.com/locate/indcrop

De novo transcriptome sequencing of *Torreya grandis* reveals gene regulation in sciadonic acid biosynthesis pathway

Jiasheng Wu[a,*], Jiandiao Huang[a], Yiwei Hong[a], Haizhen Zhang[b], Mingzhu Ding[a], Heqiang Lou[a], Yuanyuan Hu[a], Weiwu Yu[a], Lili Song[a,*]

[a] State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou, Zhejiang 311300, China
[b] Hubin Administration of Hangzhou West Lake Scenic Area, Hangzhou Municipal Bureau of Landscape and Cultural Relic, Hangzhou 310001, China

ARTICLE INFO

Keywords:
Torreya grandis
Transcriptome
Sciadonic acid
Δ^6-elongase
Δ^5-desaturase

ABSTRACT

Torreya grandis (*T. grandis*), a conifer of the Cephalotaxaceae family, yields rare and unique dried fruits of great nutritional value that contains a high level of novel UFA, sciadonic acid. The composition of sciadonic acid in different tissues and cultivars, and the molecular mechanisms underlying the biosynthesis of sciadonic acid in *T. grandis*, both of which remain unknown however, were investigated. Two cultivars, *T. grandis cv*. Yuanfei and Xifei were used to reveal in comparison the fatty acid composition of roots, stems, leaves, arils and kernels. It was found that the Yuanfei kernels had the highest levels of sciadonic acid with 17% in oil. In transcriptome analysis, a total of 80,235 and 82,454 unigenes were generated from five libraries of Yuanfei and Xifei cultivars, respectively. KEGG orthology enrichment of differentially expressed genes showed that lipid metabolism pathways were the most highly represented. In particular, six candidate unigenes (Δ^6-elongase and Δ^5-desaturase) encoding sciadonic acid elongase and desaturases with high expression levels in Yuanfei and Xifei kernels were identified. This study provides the first comprehensive genomic analysis of sciadonic acid biosynthesis and contributes to the understanding of the molecular mechanisms responsible for *de novo* fatty acid biosynthesis and high sciadonic acid accumulation in gymnosperm species.

1. Introduction

Unsaturated fatty acids (UFA) are a major focus in the fields of nutrition and medicine due to their preventative effects on angiosclerosis, obesity, certain types of cancers, and neurological disorders. The formation of UFA has evolved during *de novo* fatty acid synthesis and fatty acid elongation and desaturation pathways (Beisson et al., 2003; Harwood, 1996; Ohlrogge and Browse, 1995), subsequently linking three carbons of the glycerol backbone, accumulating as triacylglycerols (TAGs) in plant seeds (Brown et al., 2012). Currently, investigation into fatty acid synthesis in plants has focused almost entirely on oil seeds in angiosperm species, such as Arabidopsis, Pecan and Tree peony (Huang et al., 2017; Li et al., 2015; Li-Beisson et al., 2010). UFA biosynthesis in gymnosperm species, which are recommended as an important oil seed resource, is difficult to study because they have very large effective population sizes and their genomes are highly heterozygous (Nystedt et al., 2013). Sciadonic acid, also known as Cis-5,11,14 eicosatrienoic acid, is a newly identified type of non-methylene-interrupted polyunsaturated fatty acid (NMI-PUFA) with a Δ^{5cis}-ethylenic bond, and is considered to be characteristic of gymnosperms (Wolff, 1999). Sciadonic acid was initially discovered by Itabashi and Takagi (1982) in the seeds, arils, and leaf lipids of *Taxus cuspidate* (Japanese yew) (Itabashi and Takagi, 1982). Sciadonic acid mainly occurs in seed oils of various gymnosperm (conifer) species, such as *Podocarpus andinus*, *Sciadopitys verticillata*, and *Podocarpus nagi* (Berger et al., 2002; Wolff et al., 1999a,b), although the content and profile vary depending on the family, genus, or even the species. Sciadonic acid as a taxonomic marker has likely been around for over 250 million years and accompanied the emergence, radiation, and expansion of gymnosperms, whereas almost all angiosperm species have lost the capability to introduce supplementary Δ^5-desaturation in unsaturated fatty acids in C20 acids dating back to the Devonian period (Wolff, 1999; Aitzetmuller, 1995). Recently, considerable research on increasing sciadonic acid content has garnered interest because of its association with human health and nutrition, including inflammatory responses, blood pressure regulation, blood clotting, and cell signaling (Berger et al., 2002; Niu et al., 2011; Tanaka et al., 2001).

Although sciadonic acid biosynthesis has not been fully elucidated, it is known that the synthesis of sciadonic acid requires the existence of a Δ^5-desaturase that would use 9,12-C18:2 and their elongation

* Corresponding authors.
E-mail addresses: wujs@zafu.edu.cn (J. Wu), lilisong@zafu.edu.cn (L. Song).

https://doi.org/10.1016/j.indcrop.2018.04.041
Received 16 November 2017; Received in revised form 10 April 2018; Accepted 15 April 2018
Available online 07 May 2018
0926-6690/

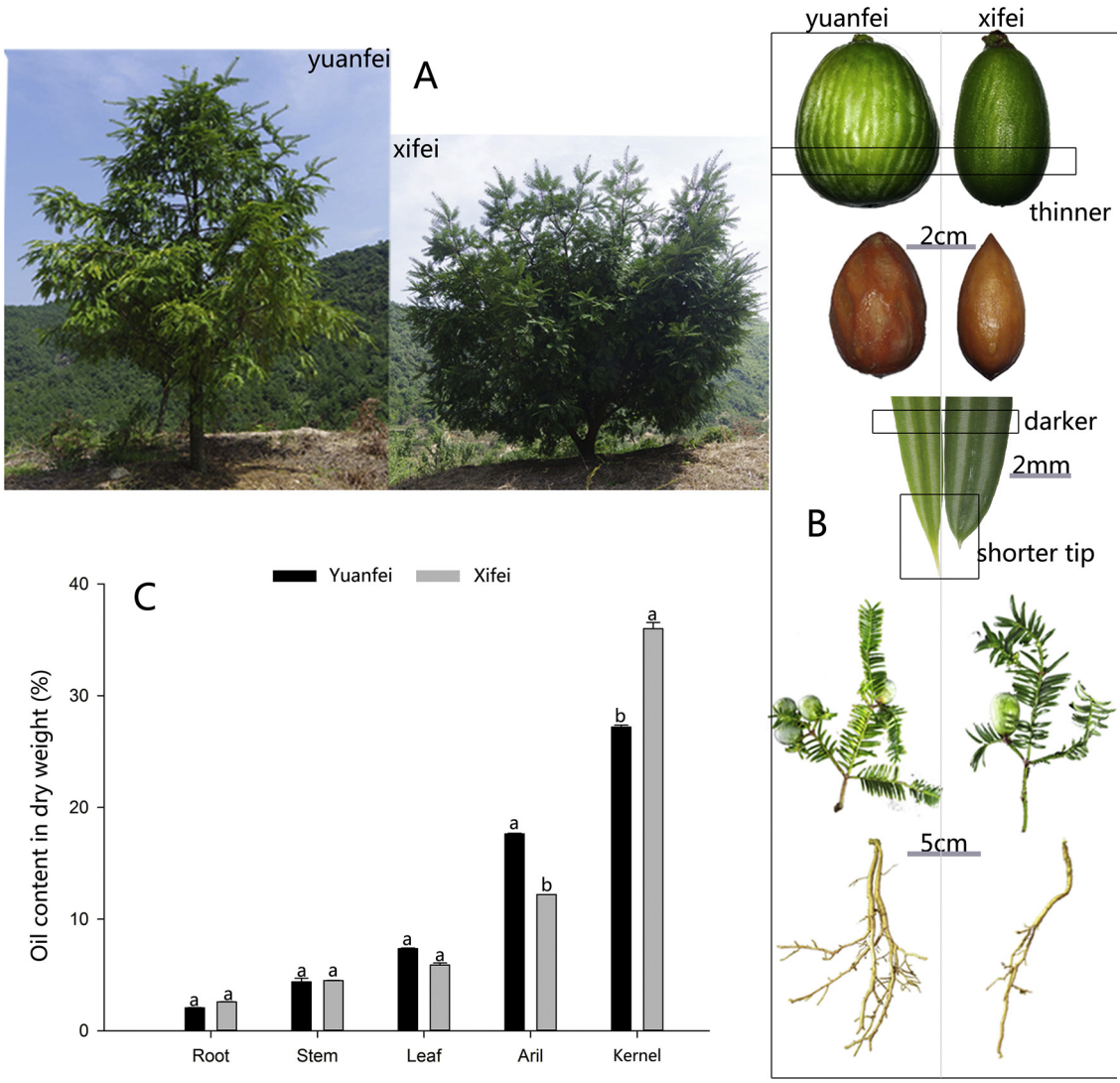


Fig. 1. The oil contents, appearance and different tissues of *T. grandis* cv. Yuanfei and Xifei. (A) The plants. (B) Root, stem, leaf, matured seed with or without a fleshy covering. (C) Oil contents of root, stem, leaf, aril and kernel tissue in 2015.

products, 11,14-C20:2 acids, as substrates (Wolff et al., 1999a,b). Furthermore, two candidate cDNAs for C20 Δ^5-desaturase from the developing seeds of *Anemone leveillei* were functionally characterized in transgenic Arabidopsis plants (Sayanova et al., 2007), in which the Δ^5-desaturase was biochemically demonstrated to use acyl-CoA as a substrate. The molecular mechanism of sciadonic acid biosynthesis in gymnosperms species, particularly with respect to the full set of genes involved in this pathway, remains unexplored.

Torreya, a large, evergreen coniferous tree, is a primitive member of the gymnospermous yew family (Taxaceae). *Torreya grandis* (*T. grandis*) is a common ornamental plant in China and has been listed as a national key protected wild plant (second group) (Saeed et al., 2007; Shen et al., 2014). Due to the influence of the environment and cultivation management, *T. grandis* is morphologically variable in China and exhibits different seed shapes and qualities, including Yuanfei, Xifei, Xiangyafei, Zhimafei, *etc*. Yuanfei (also named as Mufei) is the seedling of *T. grandis* and Xifei is a commercially important grafted variety that is widely cultivated for nut production (Li et al., 2005). In practice, the two cultivars are significantly different in external appearance (Fig. 1A and B). Furthermore, both Yuanfei and Xifei seed oils have rich unsaturated fatty acids with healthy and nutritional values, preventing angiosclerosis and coronary heart disease. Interestingly, siadonic acid has been found in abundance in *T. grandis* seeds including Yuanfei and Xifei, representing 8.3–16.4% of the oil, while Yuanfei has been established to be higher in sciadonic acid content in the seed oil than Xifei (Wang et al., 2016). Other gymnosperm species such as *Pinus* spp. seeds have sciadonic fatty acid content less than 1.24% of the oil (Wolff et al., 2001), although the presence of sciadonic acid in seed lipids of all coniferophyte families is a common biochemical feature. Thus, *T. grandis* species that contain high levels of this fatty acid constitute a useful and reliable model for identifying and determining sciadonic acid biosynthesis. For the past years, considerable research has been focused on the fatty acid composition, distribution and extraction in the seeds over *T. grandis* cultivars (Niu et al., 2011). Unfortunately, little information is available on the genomic analysis of *T. grandis* regarding the molecular mechanisms of fatty acid biosynthesis, especially associated with sciadonic acid biosynthesis.

Given that most gymnosperm species have a large genome, high proportions of repetitive elements and numerous pseudogenes, genetic information currently available for these species is limited. In this study, we generated the first tree transcriptome using high-throughput Illumina sequencing technology to uncover genes related to sciadonic acid biosynthesis, to observe gene expression patterns and characterize candidate genes/proteins responsible for sciadonic acid synthesis in *T. grandis*. More efforts has been done to complement and enhance the study, mainly by analyzing fatty acid composition of different tissues

including roots, stems, leaves, arils, and kernels in two cultivars, *T. grandis* cv. Yuanfei and Xifei, thus to gain insight into the characterization in sciadonic acid biosynthesis. These findings may not only provide a basis for dissecting the metabolic pathways involved in sciadonic acid biosynthesis, but also present potentially genetic engineering approaches to increase the sciadonic acid content of *T. grandis*, as well as in the breeding of favorable cultivars.

2. Materials and methods

2.1. Plant materials

One-hundred-year old Yuanfei and Xifei plants were randomly collected from the *T. grandis* commercial planting base in Fuyang County, Zhejiang province, China in early September 2015 and 2016. The daily temperature varied between 24 °C and 30 °C, and the relative humidity was 70%. The field soil was sandy loam with a pH of 5.5–7.0. Roots, stems, leaves and seeds of Yuanfei and Xifei plants were selected from the same sampled tree, respectively, and three biological replicates per tissue type were prepared. Roots, stems and leaves were sampled from one-year old healthy growing roots and branches, whereas stems and leaves were from the current bearing shoots. After the sarcotesta (arils) were removed and sampled by hand, the remaining hard seeds were used for kernel sampling after their shell and testa (seed coat) were removed by hand. Additionally, the center of the seeds (entirely the endosperm) was selected. After the samples were collected and sectioned separately, they were flash-frozen in liquid nitrogen and stored at −80 °C to be used for lipid analysis and RNA extraction, respectively.

2.2. Lipid analysis

Oils from the roots, stems, leaves, arils, and kernels of *T. grandis* cv. Yuanfei and Xifei (freeze-dried powder) plants were extracted with petroleum ether at 50 °C for 8 h. The fatty acid composition of the oil was analyzed by gas chromatography (Waters H-class UPLC, Agilent 7890A, Agilent Technologies, USA) according to the fatty acid methyl esters (FAME) method (the ISO, method 5509). Altogether 37 kinds of methyl-esterified fatty acids were mixed as the standard for the fatty acid composition (97.8%–99.9%, Supelco USA). Additionally, 5(Z),11(Z),14(Z)-eicosatrienoic acid (> 98%, Larodan USA) were methyl-esterified and used as the standard for assay of sciadonic acid content. For each tissue, lipid analysis was repeated in triplicate.

2.3. Total RNA extraction, cDNA library construction, and high throughput sequencing

Total RNA was isolated from the roots, stems, leaves, arils, and kernels of *T. grandis* cv. Yuanfei and Xifei plants, respectively, with a TIANGEN RNA Prep Pure Plant kit (Tiangen Biotech Co. Ltd, Beijing, China), and purified with a Dynabeads® Oligo (dT)25 kit (Life Technologies, CA, USA). RNA concentration and integrity was evaluated with a 1.5% agarose gel stained with Goldview, and detected by the RNA Nano 6000 Assay Kit on the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA). Since there was no significant difference in oil content among three biological replicates per tissue type, including roots, stems, leaves, arils, and kernels tissue of both cultivars in *T. grandis* in the same year (Additional file 1), one sample per tissue was chosen. The tissues in Yuanfei and Xifei were designated g1, j1, y1, q1, h1 and G1, J1, Y1, Q1, and H1 in 2015, and g2, j2, y2, q2, h2, and G2, J2, Y2, Q2, and H2 in 2016, respectively, and chosen for RNA-Seq analysis (Additional file 2). A total of 3 μg RNA per sample was used as input material for the RNA sample preparation. Sequencing libraries were generated using the NEBNext®Ultra™ RNA Library Prep Kit and library quality was assessed with the Agilent Bioanalyzer 2100 system. Sequencing was performed using an Illumina HiSeq™ 4000 paired-end sequencing system.

2.4. Sequence analysis and assembly

Raw data (raw reads) in FASTQ format was first processed using in-house perl scripts. In this step, reads containing adapters and poly-Ns, as well as low quality reads from the raw data, were removed. Simultaneously, Q20, Q30, GC-content, and sequence duplication levels of the clean data were calculated, serving as the base of all the downstream analyses. Cured reads were assembled using the program Trinity (Grabherr et al., 2011). Gene function was annotated based on the following databases: Nr (NCBI non-redundant protein sequences); Nt (NCBI non-redundant nucleotide sequences); Pfam (Protein family); KOG/COG (Clusters of Orthologous Groups of proteins); Swiss-Prot (a manually annotated and reviewed protein sequence database); KEGG (Kyoto Encyclopedia of Genes and Genomes pathway database); and GO (Gene Ontology).

2.5. Differential expression of unigenes

Differential expression analysis of the two cultivars was performed using the DEGseq (2010) R package (Zheng et al., 2015). *P*-values were adjusted using q values (Storey, 2002). FDR < 0.001 and |log2-Ratio| ≥ 2 found by DESeq was set as the threshold for significant differential expression.

2.6. GO and KEGG orthology enrichment analyses

GO enrichment analysis of the differentially expressed genes (DEGs) was performed using GOseq R packages based on the Wallenius non-central hyper-geometric distribution (Young et al., 2010), which can adjust for gene length bias in DEGs. KEGG is a database resource for understanding high-level functions and utilities of the biological system, such as the cell, the organism, and the ecosystem, from molecular-level information, for example, large-scale molecular datasets generated by genome sequencing and other high-throughput experimental technologies (Kanehisa et al., 2008) (http://www.genome.jp/kegg/). KOBAS software was used to test the statistical enrichment of DEGs in the KEGG pathways (Mao et al., 2005). To facilitate access and analysis of transcriptome sequencing data of *T. grandis*, the Transcriptome Assembly Sequence Database were uploaded to NCBI with accession numbers SRR6187237.

2.7. Validation of RNA-Seq data by quantitative real-time RT-PCR (qRT-PCR)

Unigenes related to fatty acid metabolic pathway gene synthesis were selected for validation by qRT-PCR. Specific primers used in qRT-PCR were obtained using Primer Quest (http://www.ncbi.nlm.nih.gov/tools/primer-blast/) according to the open reading frames of the target genes (Additional file 3). qRT-PCR was performed using SYBR Premix ExT aq™II (Takara) in 20 μL volumes on the Roche Light Cycler480 system (Roche Diagnostics, CA, USA). Each reaction was performed as follows: 95 °C for 30 s, followed by 40 cycles of 95 °C for 5 s, and 60 °C for 34 s. All samples were tested in triplicate, and experiments were performed on three biological replicates to ensure reproducibility and reliability.

2.8. Statistical analysis

All experiments in this study were conducted employing completely randomized designs. The data was tested by analysis of variance using SPSS Version 11.0. All measurements were tested by a one-way ANOVA for all the studied tissues of *T. grandis* cv. Yuanfei and Xifei. Least significance differences (LSDs) were calculated at $P < 0.05$.

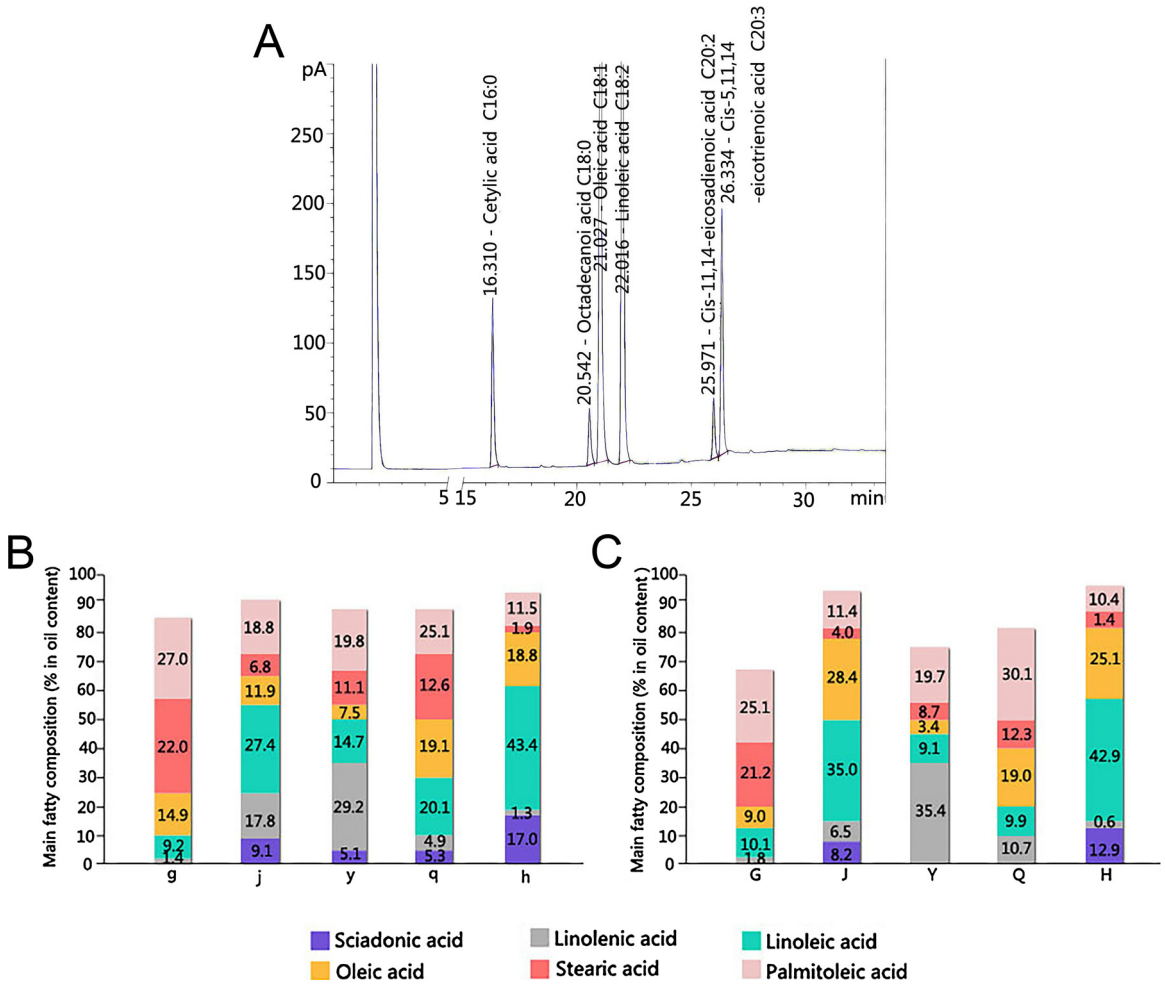


Fig. 2. Main fatty acid compositions of *T. grandis* cv. Yuanfei (A) and Xifei (B).

3. Results

3.1. Changes in fatty acid composition in different tissues of T. grandis cv. Yuanfei and Xifei

The oil content and fatty acid composition in *T. grandis* vary depending on the tissues and cultivars. As shown in Fig. 1C, compared with the root, stem and leaf tissue, the aril and kernel had significantly high oil content in both cultivars cv. Yuanfei and Xifei ($P < 0.05$). Interestingly, no significant difference was found in oil content in the root, stem and leaf tissue between Yuanfei and Xifei ($P > 0.05$).

Six dominant fatty acids exist in *T. grandis*, which are palmitic acid (C16:0), stearic acid (C18:0), oleic acid (C18:1 Δ^9), linoleic acid (C18:2 $\Delta^{9,12}$), linolenic acid (C18:3$\Delta^{9,12,15}$), and sciadonic acid (C20:3 $\Delta^{5,11,14}$) (Fig. 2A). Accumulation of major fatty acids varies in the different tissues, among which kernels had the highest ratio of unsaturated fatty acids to total oil content in both cultivars (Fig. 2B and C). Furthermore, the three dominant fatty acids in the roots, stems, and arils are palmitic acid, stearic acid, and oleic acid, whereas oleic, lenoleic and sciadonic acids are abundant in kernel tissue in both cultivars (Fig. 2B and C). Compared with the four other tissues, kernels exhibited the highest levels of sciadonic acid, with 17.0% and 12.9% in Yuanfei and Xifei, respectively (Fig. 2B and C). Interestingly, the dominant fatty acid component of the leaves of Yuanfei and Xifei was linolenic acid with 29.29% and 35.40%, respectively.

3.2. De novo assembly sequencing

To study the transcriptome of *T. grandis*, roots, stems, leaves, arils, and kernels of both cultivars were collected. Total RNA was extracted from each tissue and mRNA was isolated and sequenced with the Illumina HiSeq™ 4000 platform. Adapter sequences and reads that were shorter than 50 bp were filtered out and ultimately, 8.6, 9.7, 9.7, 9.3, and 9.9 Gbp of high-quality (HQ) reads were generated for the roots, stems, leaves, arils, and kernels in Yuanfei, and 9.4, 9.7, 9.5, 9.5, and 9.4 Gbp of HQ reads were generated for the roots, stems, leaves, arils, and kernels in Xifei (Table 1). The *de novo* transcriptome assembly was performed using Trinity with default parameters and a total of 80,235 and 82,454 non-redundant unigenes were obtained in *T. grandis* cv. Yuanfei and Xifei, respectively. The mean length was 902 nt in Yuanfei and 841 nt in Xifei, and contig N50 of Yuanfei was 1733 nt and 1662 nt in Xifei (Table 2).

3.3. Annotation and tissue-specific differential expression of transcripts

To annotate genes, we performed a homology search in BLASTX (E-value cut-off 1.0E-5), which showed that 45,277 (56.43%), 37,495 (46.73%), 32,442 (40.43%), 28,591 (35.63%) and 53,731 (66.97%) of the 80,235 contigs in Yuanfei, and 47,823 (57.80%), 39,278 (47.64%), 34,004 (41.24%), 29,669 (35.98%) and 45,379 (55.04%) of the 82,454 contigs in Xifei, respectively, had significant matches with NCBI sequences in the Nr, Nt, SwissProt, KEGG, and trEMBLE databases. These genes were tentatively annotated according to known sequences that showed the best match.

Table 1
Summary of transcriptome data sequenced by Illumina platform in tissues of *T. grandis* cv. Yuanfei and Xifei, which roots, stems, leaves, arils and kernels in Yuanfei, named as g, j, y, q and h, and roots, stems, leaves, arils and kernels in Xifei, named as G, J, Y, Q and H.

Cultivar	Tissue	Total reads (bp)	Clean reads (bp)	Clean data (bp)	GC (%)
Yuanfei	g	61,195,804	57,578,768	8,636,815,200	43.19
	j	70,666,208	64,659,020	9,698,853,000	43.72
	y	70,974,558	64,830,834	9,724,625,100	43.49
	q	67,711,310	62,412,890	9,361,933,500	44.19
	h	70,299,562	65,939,142	9,890,871,300	44.83
Xifei	G	67,439,340	62,401,966	9,360,294,900	43.86
	J	69,005,292	64,404,488	9,660,673,200	43.72
	Y	68,121,008	63,103,932	9,465,589,800	43.75
	Q	68,432,748	63,504,794	9,525,719,100	44.12
	H	66,558,372	62,591,592	9,388,738,800	44.95

Table 2
De novo assembly and annotation of *T. grandis* cv. Yuanfei and Xifei.

Cultivar	Mean length (nt)	N50 (bp)	Number of unigenes	Number of annotated unigenes
Yuanfei	902	1733	80,235	55,821
Xifei	841	1662	82,454	57,773

The best-matched species for the gene in Yuanfei included a 32.7% match with that of *Picea sitchensis*, 9.1% with *Amborella trichopoda*, 5.0% with *Indian lotus*, and 2.3% with *Physcomitrella patens* (Fig. 3B). In Xifei, there was a 33.6% match with *Picea sitchensis*, 8.9% with *Amborella trichopoda*, 4.9% with *Indian lotus* and 2.1% with *Physcomitrella patens* (Fig. 3C).

In the Yuanfei cultivar, unigenes from root, stem, leaf, and aril tissue were compared to those from kernel tissue, respectively. A total of 13,607 DEGs, constituting the highest number, were identified in leaf tissue, with 5120 being up-regulated and 8487 being down-regulated. The arils possessed the smallest number of DEGs, including 3552 up-regulated unigenes and 5814 down-regulated unigenes (Fig. 4A). In Xifei, the root tissue exhibited the highest number of DEGs with a total of 16,167 unigenes, of which 8575 were up-regulated and 7612 were down-regulated. As observed in Yuanfei, aril tissue possessed the least DEGs, of which 4691 were up-regulated and 4912 were down-regulated

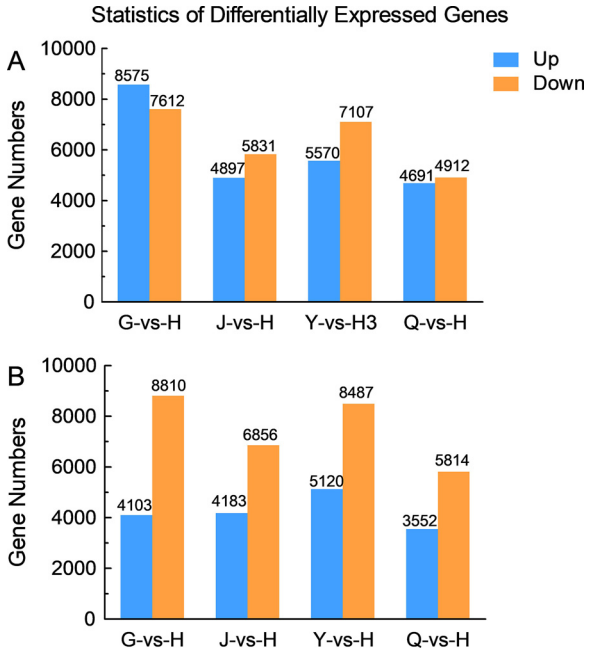


Fig. 4. Different expression genes and venn diagram for different tissues in *T. grandis* cv. Yuanfei and Xifei.

(Fig. 4B; Additional file 4).

3.4. Functional classification of differential genes in T. grandis by GO and KEGG

Functional annotation with GO terms was carried out using the BLASTp algorithm against the Swiss Prot and TrEMBL databases by the GoPipe program according to gene2go software (E-value cut-off, 1.0E-5). All DEGs were categorized into 46 functional groups under three main divisions (biological processes, cellular components, and molecular functions). In the biological process category, metabolic process and single-organism process were the predominant groups identified in both cultivars (Additional file 5).

Pathway-based analysis can further enhance our understanding of the biological functions and interactions of genes. A total of 3138 DEGs

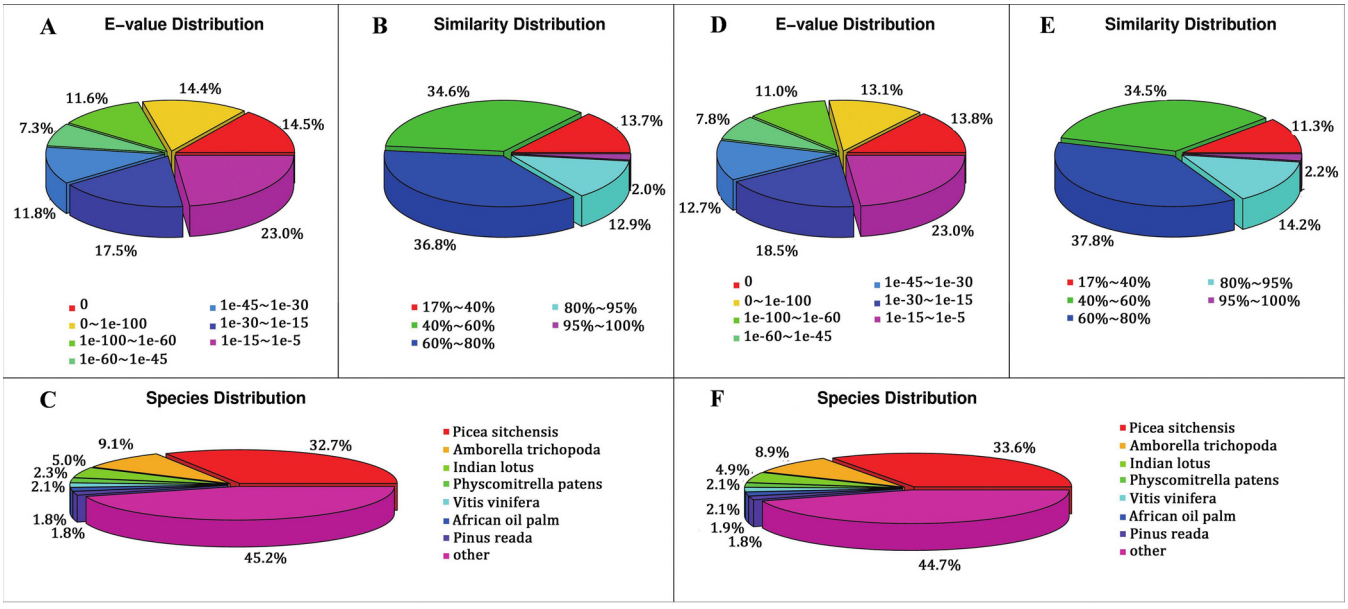


Fig. 3. Species distribution of top BLAST hits of sequences *T. grandis* var. Yuanfei (A) and Xifei (B) with other plant species.

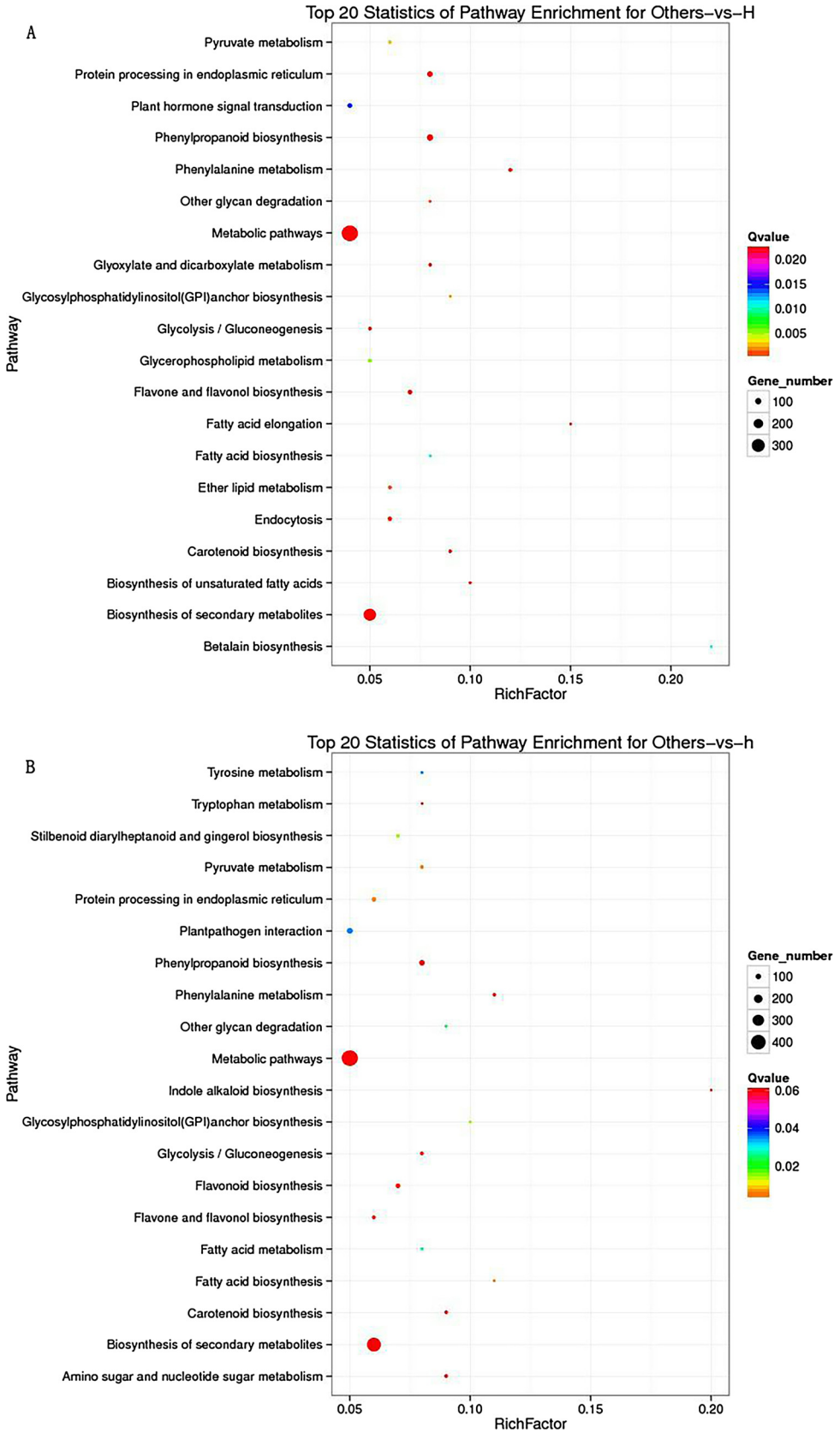


Fig. 5. Different expression genes KEGG enrichment in *T. grandis* cv. Yuanfei and Xifei.

were assigned to 116 pathways in the KEGG database in Yuanfei, and 2792 were assigned to 108 pathways in the KEGG database in Xifei (Additional file 6). The most represented pathways in *T. grandis* included "metabolic pathways", "biosynthesis of secondary metabolites" and "plant-pathogen interaction", consisting of 453 and 388, 380 and 279, and 116 and 101 unigenes in the three pathways in Yuanfei

Table 3
Differential Expression of lipid synthesis in *T. grandis* cv. Yuanfei and Xifei.

Cultivar	Pathway	Sequence ID	Sequence description	Expression level (RPKM)				
				Root	Stem	Leaf	Aril	Kernel
Yuanfei	Fatty acid biosynthesis	TR92797|c0_g1	3-oxoacyl-[acyl-carrier-protein] synthase I	28.80	21.39	17.66	28.15	85.63
		TR94949|c1_g1	Acetyl-coenzyme A carboxylase carboxyl transferase subunitalpha, chloroplastic	28.18	3.58	3.17	10.03	90.19
		TR96418|c0_g1	Acyl-[acyl-carrier-protein] desaturase	51.08	71.26	31.58	47.61	319.19
		TR110170|c0_g1	Enoyl-[acyl-carrier protein] reductase I	0.00	0.00	0.00	1.71	20.06
		TR118991|c2_g12	Oleoyl-acyl carrier protein thioesterase, chloroplastic	1.51	2.37	0.46	20.98	91.66
		TR127800|c1_g2	3-Oxoacyl-[acyl-carrier-protein] reductase, chloroplastic	0.87	4.02	2.86	1.03	67.47
		TR127800|c1_g3	Short-chain type dehydrogenase/reductase	0.43	5.16	7.04	1.59	94.04
		TR127800|c1_g9	Tropine dehydrogenase	1.42	3.24	3.85	1.30	34.33
		TR156286|c0_g1	NAD(P)-binding Rossmann-fold superfamily protein	19.11	16.69	17.43	20.88	101.26
		TR30875|c0_g1	Biotin carboxyl carrier protein of acetyl-CoA carboxylase, chloroplastic	29.01	13.01	17.39	29.89	102.60
		TR47870|c0_	Acyl-[acyl-carrier-protein] desaturase, chloroplastic	88.30	119.13	74.20	58.38	422.22
		TR50173|c0_g1	Oleoyl-ACP thioesterase	1.76	4.15	2.85	37.30	87.18
		TR50173|c0_g3	Oleoyl-ACP thioesterase	4.46	5.28	3.65	44.10	123.22
		TR55271|c0_g1	11-Beta-hydroxysteroid dehydrogenase-like protein	0.00	0.07	0.02	0.06	30.50
	Unsaturated fatty acidsbiosynthesis	TR96418|c0_g1	Acyl-[acyl-carrier-protein] desaturase	51.08	71.26	31.58	47.61	319.19
		TR127800|c1_g2	3-Oxoacyl-[acyl-carrier-protein] reductase, chloroplastic	0.87	4.02	2.86	1.03	67.47
		TR127800|c1_g3	Short-chain type dehydrogenase/reductase	0.43	5.16	7.04	1.59	94.04
		TR127800|c1_g9	Tropine dehydrogenase	1.42	3.24	3.85	1.30	34.33
		TR141933|c0_g1	Omega-3 fatty acid desaturase, chloroplastic	1.49	14.67	4.39	30.09	90.71
		TR156286|c0_g1	NAD(P)-binding Rossmann-fold superfamily protein	19.11	16.69	17.43	20.88	101.26
		TR158563|c0_g1	Polyunsaturated fatty acid specific elongation enzyme 1	76.61	52.34	74.43	8.67	420.94
		TR168891|c0_g1	Very-long-chain (3*R*)-3-hydroxyacyl-[acyl-carrier protein] dehydratase	0.33	0.00	0.32	1.06	6.18
		TR168891|c0_g2	Very-long-chain (3*R*)-3-hydroxyacyl-[acyl-carrier protein] dehydratase	8.38	7.12	10.68	18.47	41.19
		TR47870|c0_g1	Acyl-[acyl-carrier-protein] desaturase, chloroplastic	88.30	119.13	74.20	58.38	422.22
		TR55271|c0_g1	11-Beta-hydroxysteroid dehydrogenase-like protein	0.00	0.07	0.02	0.06	30.50
	Fatty acid elongation	TR158563|c0_g1	Polyunsaturated fatty acid specific elongation enzyme 1	76.61	52.34	74.43	8.67	420.94
		TR163010|c0_g3	1,4-Dihydroxy-2-naphthoyl-CoA synthase	1.32	1.88	2.52	3.21	20.81
		TR164584|c0_g1	3-Ketoacyl-CoA synthase 1	0.00	0.05	0.00	0.04	2.45
		TR168891|c0_g1	Very-long-chain (3*R*)-3-hydroxyacyl-[acyl-carrier protein] dehydratase	0.33	0.00	0.32	1.06	6.18
	Glycerolipid metabolism	TR168891|c0_g2	Very-long-chain (3*R*)-3-hydroxyacyl-[acyl-carrier protein] dehydratase	8.38	7.12	10.68	18.47	41.19
		TR61455|c0_g1	NADPH-dependent codeinone reductase 2	10.96	4.91	14.37	3.05	0.10
		TR137605|c0_g2	1,2-Dihydroxy-3-keto-5-methylthiopentene dioxygenase 2	1.29	1.26	0.43	2.95	14.16
		TR142408|c1_g3	Aldo-keto reductase family 4 member C9	0.81	0.29	0.42	0.66	13.85
		TR165726|c4_g27	Alpha-galactosidase	2.05	0.47	0.29	0.07	7.90
		TR179013|c11_g2	1-Acyl-sn-glycerol-3-phosphate acyltransferase 5	6.92	29.15	1.83	0.00	45.55
		TR179398|c0_g2	Diacylglycerol acyltransferase 1-like	1.30	0.94	0.78	1.40	13.95
		TR189854|c4_g22	NADP-dependent D-sorbitol-6-phosphate dehydrogenase	0.22	2.03	1.96	2.19	21.95
		TR45755|c0_g2	Alpha-galactosidase	0.03	0.09	0.02	0.05	53.05
Xifei	Fatty acid biosynthesis	TR119967|c1_g1	Delta 9 steroyl ACP desaturase	0.00	0.00	0.00	0.08	18.69
		TR121450|c3_g2	3-Oxoacyl-[acyl-carrier-protein] reductase, chloroplastic	0.96	6.75	3.16	0.44	124.90
		TR121450|c3_g3	Short-chain type dehydrogenase/reductase	0.89	0.48	0.51	0.11	78.30
		TR121450|c3_g4	3-Oxoacyl-[acyl-carrier-protein] reductase, chloroplastic	0.00	1.13	1.54	0.83	39.97
		TR122370|c1_g2	2,4-Dienoyl-CoA reductase	20.82	26.24	28.99	52.30	157.26
		TR173964|c1_g3	Nelumbo nucifera acyl-[acyl-carrier-protein] desaturase, chloroplastic-like	0.00	0.00	0.00	2.73	509.63
		TR173964|c1_g6	Stearoyl-ACP desaturase	0.00	0.00	0.00	0.91	367.38
		TR173964|c1_g10	Stearoyl-acyl carrier protein desaturase	0.00	0.00	0.00	1.51	411.98
		TR173964|c1_g11	Stearoyl acyl carrier protein desaturase	0.00	0.00	0.00	0.00	212.80
		TR173964|c1_g12	Stearoyl-[acyl-carrier-protein] 9-desaturase 2, chloroplastic	0.00	0.00	0.00	0.50	276.82
		TR178531|c5_g15	Short-chain type dehydrogenase/reductase	29.79	31.02	25.91	3.70	122.61
		TR194026|c0_g1	Stearoyl-ACP desaturase	0.08	0.00	0.00	0.86	208.02
		TR64948|c0_g1	Acyl-[acyl-carrier-protein] desaturase, chloroplastic	0.00	0.00	0.00	0.42	62.34
	Unsaturated fatty acids biosynthesis	TR77988|c0_g1	Fatty acid desaturase	0.00	0.00	0.00	0.00	10.65
		TR119967|c1_g1	Delta 9 steroyl ACP desaturase	0.00	0.00	0.00	0.08	18.69
		TR121450|c3_g2	3-Oxoacyl-[acyl-carrier-protein] reductase, chloroplastic	0.96	6.75	3.16	0.44	124.90
		TR121450|c3_g3	Short-chain type dehydrogenase/reductase	0.89	0.48	0.51	0.11	78.30
		TR121450|c3_g4	3-Oxoacyl-[acyl-carrier-protein] reductase, chloroplastic "	0.00	1.13	1.54	0.83	39.97
		TR122370|c1_g2	2,4-Dienoyl-CoA reductase	20.82	26.24	28.99	52.30	157.26
		TR128343|c0_g1	Very-long-chain (3*R*)-3-hydroxyacyl-[acyl-carrier protein] dehydratase	0.19	0.99	2.31	2.72	10.81
		TR150287|c0_g1	Polyunsaturated fatty acid specific elongation enzyme 1	72.68	51.48	71.05	4.89	379.12
		TR173964|c1_g3	Nelumbo nucifera acyl-[acyl-carrier-protein] desaturase, chloroplastic-like	0.00	0.00	0.00	2.73	509.63
		TR173964|c1_g6	Stearoyl-ACP desaturase	0.00	0.00	0.00	0.91	367.38

(*continued on next page*)

Table 3 (*continued*)

Cultivar	Pathway	Sequence ID	Sequence description	Expression level (RPKM)				
				Root	Stem	Leaf	Aril	Kernel
		TR173964|c1_g10	Stearoyl-acyl carrier protein desaturase	0.00	0.00	0.00	1.51	411.98
		TR173964|c1_g11	Stearoyl acyl carrier protein desaturase	0.00	0.00	0.00	0.00	212.80
		TR173964|c1_g12	Stearoyl-[acyl-carrier-protein] 9-desaturase 2, chloroplastic	0.00	0.00	0.00	0.50	276.82
		TR178531|c5_g15	Short-chain type dehydrogenase/reductase	29.79	31.02	25.91	3.70	122.61
		TR194026|c0_g1	Stearoyl-ACP desaturase	0.08	0.00	0.00	0.86	208.02
		TR196757|c2_g11	Acyl-coenzyme A oxidase 2	8.65	11.18	7.16	13.80	45.19
		TR201154|c0_g1	Elongation of very long chain fatty acids protein 5	0.00	13.68	19.29	0.91	109.50
		TR64948|c0_g1	Acyl-[acyl-carrier-protein] desaturase, chloroplastic	0.00	0.00	0.00	0.42	62.34
		TR120163|c0_g1	3-Ketoacyl-CoA synthase 11	0.00	0.00	0.00	0.66	12.53
		TR120163|c0_g2	3-Ketoacyl-CoA synthase 11	0.00	0.43	0.00	6.80	20.58
	Fatty acid elongation	TR120163|c1_g1	3-Ketoacyl-CoA synthase 10	0.00	0.96	0.00	15.49	77.25
		TR120163|c1_g2	3-Ketoacyl-CoA synthase 10	0.02	0.82	0.02	11.68	37.04
		TR122309|c1_g2	3-Ketoacyl-CoA synthase 6	0.03	4.87	1.79	1.41	57.70
		TR128343|c0_g1	Very-long-chain (3*R*)-3-hydroxyacyl-[acyl-carrier protein] dehydratase	0.19	0.99	2.31	2.72	10.81
		TR131852|c3_g1	3-Ketoacyl-CoA synthase	0.00	0.00	0.00	0.03	1.83
		TR150287|c0_g1	Polyunsaturated fatty acid specific elongation enzyme 1	72.68	51.48	71.05	4.89	379.12
		TR167400|c1_g3	1,4-Dihydroxy-2-naphthoyl-CoA synthase	0.91	2.44	1.60	2.96	11.81
		TR201154|c0_g1	Elongation of very long chain fatty acids protein 5	0.00	13.68	19.29	0.91	109.50
	Glycerolipid metabolism	TR135441|c0_g1	Diacylglycerol acyltransferase 1-like	1.06	2.08	1.91	1.29	14.85
		TR136723|c0_g1	Alpha-galactosidase	0.03	0.05	0.05	0.06	4.89
		TR164059|c2_g11	Alpha-galactosidase	0.00	0.00	0.00	0.00	20.70

and Xifei, respectively (Fig. 5). Notably, some pathways closely linked to changes in oil content and composition were enriched in both cultivars, such as “fatty acid biosynthesis”, “biosynthesis of unsaturated fatty acids”, “fatty acid elongation”, and “glycerolipid metabolism” (Fig. 5). These provided critical clues for identifying and characterizing key functional genes involved in UFA biosynthesis (Additional file 7).

3.5. Identification of unigenes in lipid synthesis in different tissues of *T. grandis* cv. Yuanfei and Xifei

The DEGs involved in lipid synthesis were compared between Yuanfei and Xifei (Table 3). Compared to the root, stem, leaf and aril tissue, most unigenes associated with the putative “fatty acid biosynthesis”, “biosynthesis of unsaturated fatty acids”, “fatty acid elongation” and “glycerolipid metabolism” showed higher expression levels in kernel tissues in both cultivars, although the number of unigenes involved in lipid synthesis were different between Yuanfei and Xifei (Table 3). Interestingly, the biosynthesis of UFA related enzymes, such as stearoyl-ACP desaturase and stearoyl-[acyl-carrier-protein] 9-desaturase II had almost no expression in root, stem, leaf and aril tissue, but were rapidly up-regulated in Xifei kernels.

3.6. Characterization of genes involved in sciadonic acid biosynthesis

All the reported genes involved in putative sciadonic acid biosynthesis could be detected from the tissue-based transcriptomes of both cultivars (Additional file 8), namely two subunits (biotin carboxylase (BC) and biotin carboxyl carrier protein (BCCP)) of acetyl-CoA carboxylase (ACCase), enoyl-ACP reductase (EAR), ketoacyl-ACP synthase I, II (KAS I, II), acyl-ACP thioesterase A/B (FATA/B), stearoyl-ACP desaturase (SAD), oleate desaturase (FAD6), linoleate desaturase (FAD8 in Yuanfei and FAD7 in Xifei), elongase, and Δ^5-desaturase. In many cases, more than one unique sequence was annotated as encoding the same enzyme, probably because they represent fragments of a single transcript, or different members of a gene family, or both (Hyun et al., 2012). ACCase, which catalyzes the first reaction to generate an intermediate malonyl-CoA, is a regulatory enzyme that controls, at least in part, the rate of fatty acid synthesis (Fig. 6). Unigenes TR62704 and TR158628 encoding BC were highly expressed in all studied tissues in Yuanfei. Subsequently, repeated condensations of malonyl-ACP with a growing acyl-ACP chain were primed by fatty acid synthase (FAS) subunits, consecutively adding two carbon units to form 16:0-ACP (Fig. 6). In this process, seven, four, five, three, and two unigenes encoding BC, BCCP, EAR, KAS I, and KAS II, respectively, were identified in Yuanfei, whereas only five, three, one, two, and one unigenes encoding BC, BCCP, EAR, KAS I, and KAS II, respectively, were identified in Xifei.

During fatty acid desaturation, ten and fourteen unigenes encoding unsaturated fatty acids were detected, including two and five unigenes for stearoyl-ACP desaturase (SAD), two and five unigenes for oleate desaturase (FAD6), and two and one unigene for omega-3 fatty acid desaturase (FAD7/8) in Yuanfei and Xifei, respectively. In Yuanfei, the relative abundance of TR96418 encoding SAD was up-regulated in kernel samples, whereas TR29422 and TR64485 encoding FAD6 were up-regulated in leaf tissue. In Xifei, the relative abundance of TR173964 encoding SAD was up-regulated in the aril, and unigene TR80792 encoding FAD6 was up-regulated in the stem and leaf. Conversely, the relative abundance of unigenes encoding FAD7/8 were differentially expressed in Yuanfei and Xifei, in which two unigenes encoding FAD 7 were up-regulated in the stem and kernel tissues in Yuanfei, but the unigene encoding FAD8 was lower in Xifei (Fig. 6). According to the putative pathway, specific elongase and Δ^5-desaturase genes were involved in the formation of sciadonic acid, where one and two unigenes encoding elongase and two and one unigenes encoding Δ^5-desaturase were detected in Yuanfei and Xifei, respectively (Fig. 6).

Subsequent sequence analysis showed that both elongase and Δ^5-desaturase constituted the part-length homologue in *T. grandis* (Fig. 7). The alignment of the predicted protein sequences of Δ^5-desaturase and elongase from *T. grandis* cv. Yuanfei and Xifei showed that one and two elongase genes in Yuanfei and Xifei shared 99.7% identity, and two and one Δ^5- desaturase genes shared 100% identity to each other, respectively, indicating that all elongase genes in Yuanfei and Xifei are considered to represent the same genes named *TgELO1* and *TgELO2*, respectively. The Δ^5-desaturase genes are named *TgΔ5-1* and *TgΔ5-2*, respectively. BLAST showed that *TgELO* in *T. grandis* shared 53.3% and 43.2% identity with the Δ^6-elongase gene of *Marchantia polymorpha* and *Pyramimonas cordata*, respectively, in which two characteristic motifs (L–H–X–X–H–H and M–Y–X–Y–Y, boxed in Fig. 7C) are conserved in the ELO-like protein sequences (Kajikawa et al., 2006a,b). It should be noted that from a phylogenetic perspective, *TgELO1* and 2 are closer to *OtELO1* in algae, which is a putative C18Δ^6-elongase (Fig. 7C). Additionally, Δ^5-desaturase genes in *T. grandis* shared 61% and 53%

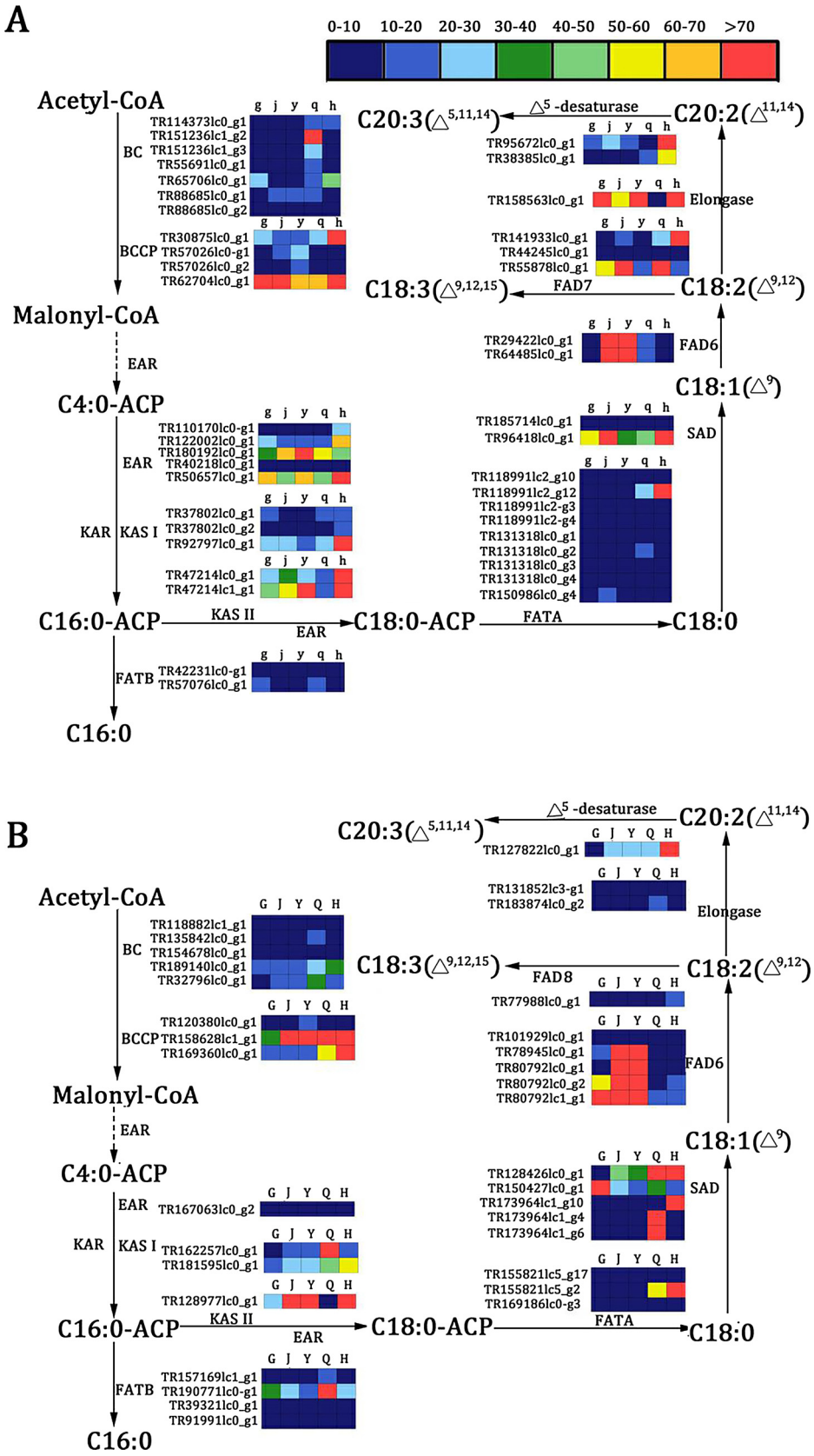
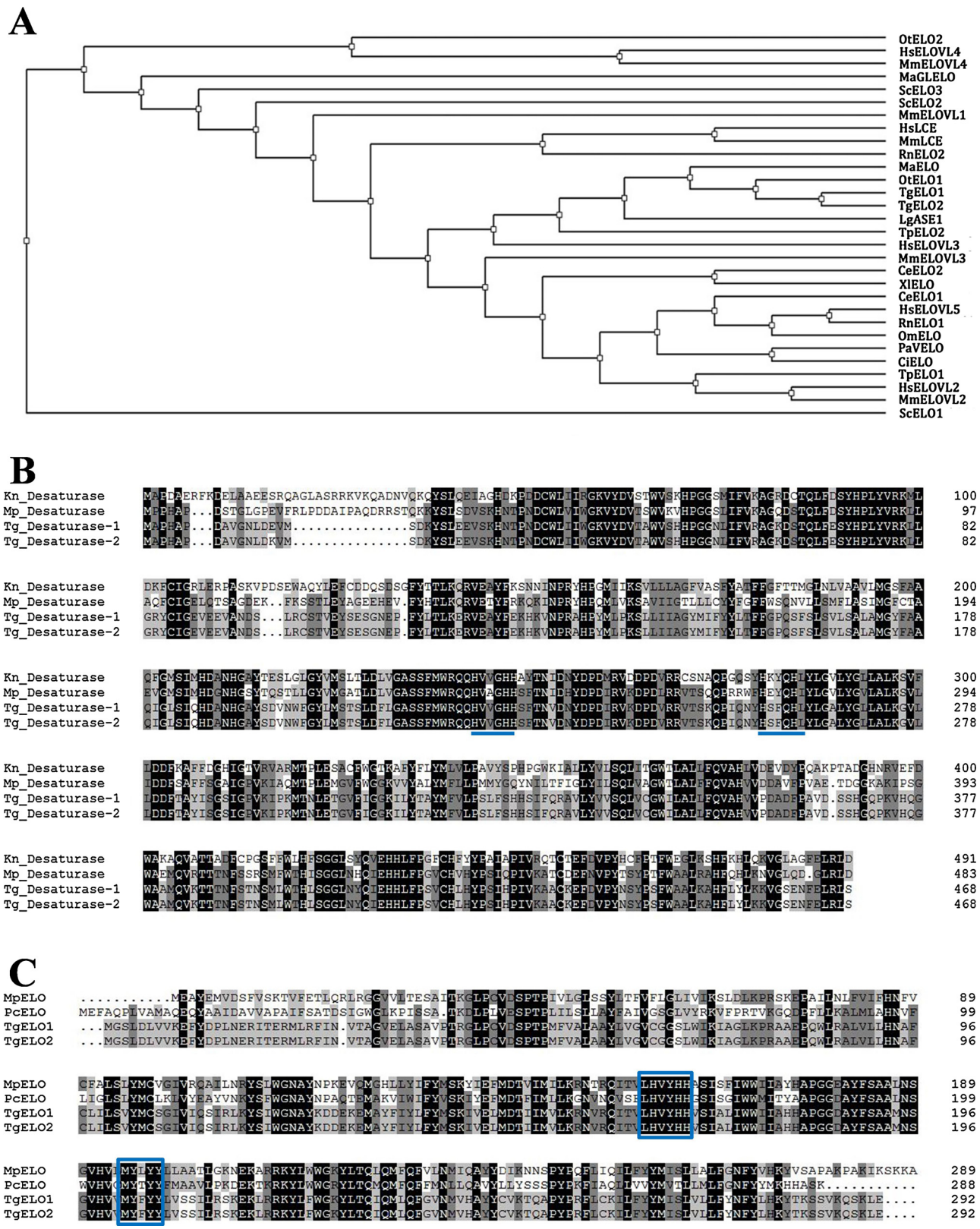


Fig. 6. The sequences associated with de novo fatty acid and sciadonic acid biosynthetic pathway in *T. grandis* cv. Yuanfei (A) and Xifei (B). Enzyme names, unigene ids and expression patterns are indicated at each step. Grids with 8 different colors from blue to red show the RPKM values. 0–10, 10–20, 20–30, 30–40, 40–50, 50–60, 60–70 and over 70 represented by colors 1–8, respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)



(caption on next page)

identity with that of *M. polymorpha* and *P. cordata*, respectively.

3.7. Validation of RNA-Seq data by qRT-PCR analysis

We conducted comparative analysis of the expression data obtained by RNA-Seq with qRT-PCR in root, stem and leaf tissues based on FPKM values to confirm the reliability of the RNA-Seq data. In the present study, twelve genes involved in fatty acid biosynthesis, fatty acid elongation and unsaturated fatty acid synthesis were selected and their expression patterns were detected by qRT-PCR. High consistency was

Fig. 7. Phylogenetic tree of ELO-like elongase family (A) (based on Kajikawa et al., 2006a,b; Meyer et al., 2004), and Alignment of the amino acid sequences of elongase (B) and Δ^5-desaturases (C) of *Torreya grandis* cv. Yuanfei and Xifei. GenBank accession numbers of the sequences are NP_012339 (ScELO1; *Saccharomyces cerevisiae*, *S. cerevisiae*), NP_009963 (ScELO2; *S. cerevisiae*), NP_013476 (ScELO3; *S. cerevisiae*), AAF71789 (MaELO; *Mortierella alpina*, *M. musculus*), NP_031729 (MmELOVL3; *Mus musculus*), AAG17875 (HsELOVL3; *Homo sapiens*, *H. sapiens*), NP_076995 (HsLCE; *H. sapiens*), AAL14239 (MmLCE; *M. musculus*), BAB69888 (RnELO2; *Rattus norvegicus*, *R. norvegicus*), NP_503114 (CeELO2; *Caenorhabditis elegans*), AAF70462 (CeELO1; *C. elegans*), AAV33630 (PavELO; *Pavlova* sp.), AAV67800 (TpELO2; *Thalassiosira pseudonana*, *T. pseudonana*), AAV67798 (OtELO2; *Ostreococcus tauri*, *O. tauri*), AAL37626 (IgASE1; *Isochrysis galbana*, *I. galbana*), AAV67797 (OtELO1; *O. tauri*), AAV67799 (TpELO1; *T. pseudonana*), AAF70417 (MaGLELO; *M. alpina*), XP_002040 (HsELOVL1; *H. sapiens*), NP_062295 (MmELOVL1; *M. musculus*), AAV67801 (XlELO; *Xenopus laevis*, *X. laevis*), NP_073563 (HsELOVL4; *H. sapiens*), AAG47667 (MmELOVL4; *M. musculus*), AAV67802 (CiELO; *Ciona intestinalis*, *C. intestinalis*), NP_060240 (HsELOVL2; *H. sapiens*), NP_062296 (MmELOVL2; *M. musculus*), NP_068586 (HsELOVL5; *H. sapiens*), NP_599209 (RnELO1; *R. norvegicus*), and AAV67803 (OmELO; *Oncorhynchus mykiss*, *O. mykiss*). The conserved His boxes are underlined and the characteristic motifs are boxed.

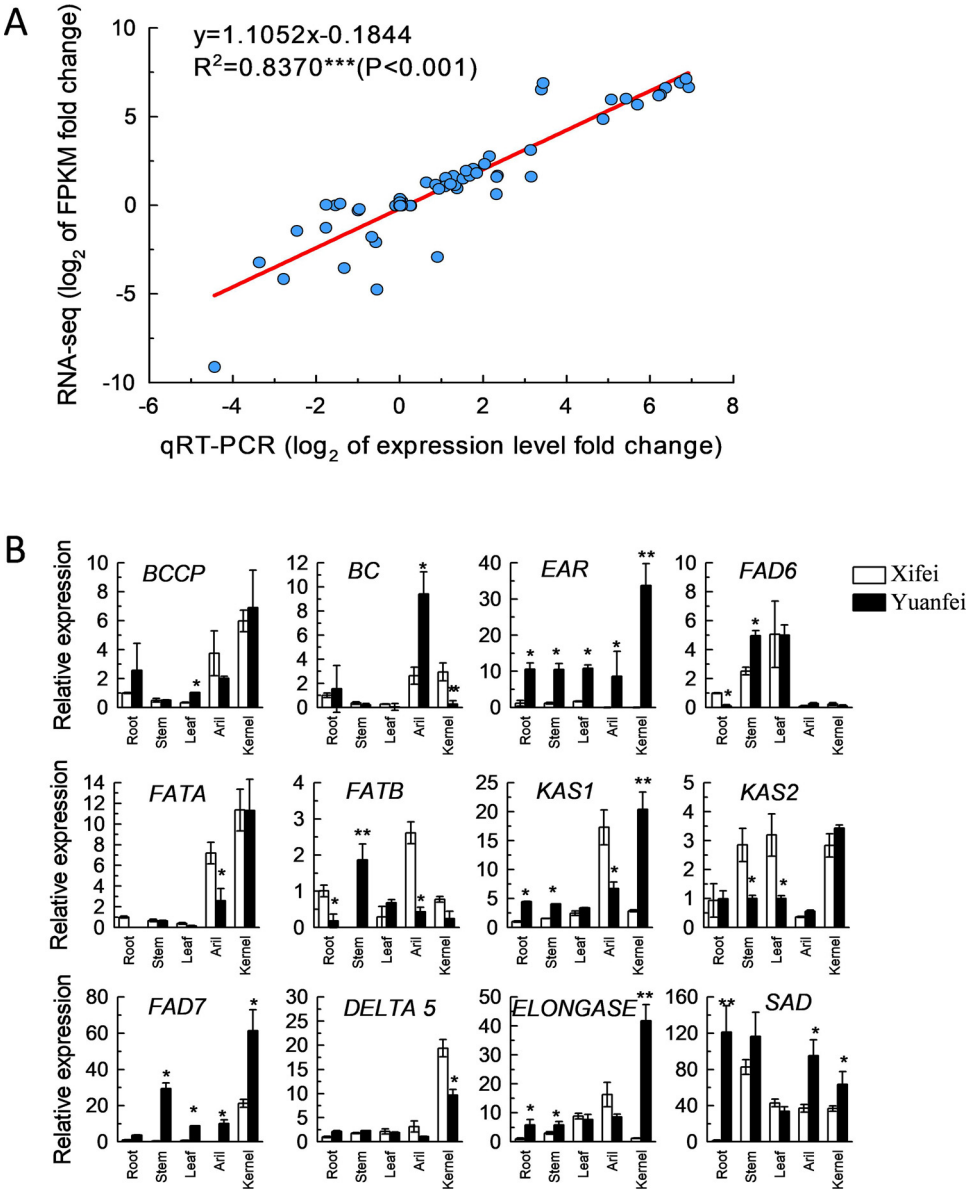


Fig. 8. RT-qPCR validation of RNA-seq relative expression estimation. (A) Correlation plot between the RT-qPCR fold-change (FC) compared to FC calculated from RPKM values. (B) RT-qPCR analysis showing the fold of expression level in the root, stem, leaf, aril and kernel tissue in *T. grandis* cv. Yuanfei and Xifei. Error bars show the confidence intervals calculated from 1SE of ΔΔCt.

verified between RNA-Seq and qRT-PCR results for the twelve candidate unigenes ($r^2 = 0.837$, $p < 0.001$), indicating that the expression data obtained by RNA-Seq was reliable (Fig. 8A).

Expression analysis using qRT-PCR was also performed to compare the relative transcript levels of the unigenes in arils and kernels of *T. grandis* (Fig. 8B). Genes encoding BCCP, KAS I and II exhibited significantly high expression levels in all of the studied tissues for both cultivars. BCCP, the subunit of ACCase, is partly a regulatory enzyme controlling the rate of fatty acid synthesis, and KAS I and II are related to elongation of palmitoyl-ACP to stearoyl-ACP. Interestingly, genes encoding SAD, FAD6, and FAD7/8 exhibited different expression patterns in all tissues of both cultivars, among which the expression levels of genes encoding SAD, and FAD7/8 were higher in Yuanfei than in Xifei, whereas genes encoding FAD6 exhibited higher expression levels

in the stem and leaf tissue than in other tissues in both cultivars. Additionally, kernels showed the highest expression levels of genes encoding Δ^5-desaturase among all five studied tissues in both cultivars. In kernels, Δ^5-desaturase of Xifei was up-regulated about one-fold compared with Yuanfei, whereas elongase in Yuanfei displayed much higher expression than in Xifei.

4. Discussion

The high levels of unsaturated fatty acids in oil seeds has been linked to preventative cardiovascular disease, cancer, inflammation, and autoimmune disease (Simopoulos, 2002, 2006). An expansion of the available genomic and transcriptomic resources for nonmodel species constitutes one means of facilitating comparative analyses of fatty acid metabolism across multiple species. In contrast to angiosperms and more basal plant groups, for which 30 genomes have already been sequenced, entire sequenced genomes of gymnosperm species are limited due to their size, though the Norway spruce genome has been sequenced (Nystedt et al., 2013; Zimin et al., 2017). This is especially true for *T. grandis,* which is characterized by wind-dispersed pollen, high heterozygosity, and a long life span of decades to centuries (Zeng et al., 2014). RNA-Seq is a highly effective and powerful technology for generating comprehensive transcriptome profiles, especially in the context of the large and unknown genomic information of species (Dai et al., 2014; Yi and Bao, 2016). In the present study, we first generated the tree transcriptome using high-throughput Illumina sequencing to identify features specific to high oil production, uncovered genes related to sciadonic acid biosynthesis, and provided insights into seed lipid metabolism in gymnosperms.

4.1. Unsaturated fatty acid levels, especially sciadonic acid content, varies depending on the tissue and cultivar

In current study, *T. grandis* is characterized by abundant unsaturated fatty acids (UFA), whereas the content of UFA varies depending on the tissues and cultivars (Figs. 1C and 2B, C). Compared with the vegetative tissues of roots, stems and leaves, the seeds had significantly higher levels of UFA, among which kernels exhibited the highest ratio of UFA to total oil content in both cultivars (Fig. 2B and C). Among the UFA, sciadonic acids were a major group that mainly exist in kernel tissue in both cultivars (Fig. 2B and C). This result was also reported by Wolff et al. (1997a,b,c) who showed that sciadonic acid mainly occurs in the seed oils of various gymnosperm (conifer) species, such as *Podocarpus andinus*, *Sciadopitys verticillata,* and *Podocarpus nagi*. Furthermore, compared with Xifei, a superior *T. grandis* cultivar, the wild cultivar Yuanfei contained higher levels of sciadonic acid in the seed kernels (Fig. 2), indicating a new oleaginous source for sciadonic acid accumulation in Yuanfei. Wang et al. (2016) also found sciadonic acid content to be higher in Yuanfei seeds than in Xifei.

Linolenic acid is another essential dietary unsaturated fatty acid not synthesized by the human body. Interestingly, the dominant fatty acid compound in the leaves of Yuanfei and Xifei was linolenic acid (29.29% and 35.40%, respectively), of which smaller amounts were found in the kernel (1.62% and 0.68%, respectively) (Fig. 2). This result was consistent with what was previously reported by Hernández et al. (2016), who found that young leaves had the highest levels of linolenic acid in both *Picual* and *Arbequina* cultivars.

4.2. De novo fatty acid biosynthesis were identified in T. grandis cv. Yuanfei and Xifei

In the present study, genes involved in putative *de novo* fatty acid biosynthesis were identified in *T. grandis* cv. Yuanfei and Xifei transcriptomes, in which unigenes TR62704 and TR158628 encoding BC were highly expressed in all studied tissues in Yuanfei. *De novo* fatty acid elongation was performed, manifested as identified unigenes encoding BC, BCCP, EAR, KAS I and KAS II (Fig. 6), and the verification of selected genes with qRT-PCR confirmed the changes (Fig. 8B). This result indicated that the acyl-CoA-dependent pathway, which is the major pathway for TAG assembly in plants (Ohlrogge and Browse, 1995), was conserved in fatty acid biosynthesis in *T. grandis*, consistent with genes in tree peony seeds (Li et al., 2015).

Desaturation is the key step that results in unsaturated fatty acids, when fatty acyl chains are desaturated by fatty acid desaturases family through introducing double bonds into fatty acid chains. In the plastids, soluble stearoyl-ACP desaturases (SAD) catalyse the initial desaturation of stearoyl-ACP to oleoyl-ACP and form oleic acid. Subsequently, oleic acid can be desaturated to linoleic acid and further linolenic acid by FAD2, FAD3, FAD4, FAD5, FAD6, FAD7 and FAD8 in the endoplasmic reticulum (ER)/plastids (Ohlrogge and Browse, 1995). FAD2 and FAD6 are oleate desaturases and located in the endoplasmic reticulum and plastids, respectively (Chi et al., 2011), whereas FAD3, FAD7 and FAD8 encode ER-localized omega-3 fatty acid desaturase and plastidial isozymes of this enzyme, respectively (Ohlrogge and Browse, 1995; Matsuda et al., 2005). In this study, genes encoding SAD, FAD6 and FAD7/8 were identified in *T. grandis* transcriptomes, consistent with the unsaturated fatty acids contents in kernel tissue in both cultivars (Fig. 2B and C), indicating the role of these genes in the formation of unsaturated fatty acids (Ohlrogge and Browse, 1995). Also, Li et al. (2015) reported that the formation of unsaturated fatty acids was associated with the high expression of FAD8 in tree peony seeds. Interestingly, the accumulation of linolenic acid in leaf tissue in both *T. grandis* cultivars was not accompanied by transcriptional expression level of *FAD7/8* gene, indicating the involvement of post-translational regulation mechanism (Matsuda et al., 2005). However, it is necessary to further examine substrate specificity and regulation of enzymes related to mega-3 fatty acid desaturase in *T. grandis*.

4.3. Genes related to sciadonic acid biosynthesis were identified and characterized in T. grandis cv. Yuanfei and Xifei

It has been established that the synthesis of sciadonic acid requires the existence of a Δ^5-desaturase that would use the product catalyzed from 9-C18:1 by the enzyme of FAD6, 9,12-C18:2 and their elongation products, 11,14-C20:2 acids, as substrates (Wolff et al., 1999a,b). In addition to Δ^5-desaturase, Δ^9-elongating activity was identified and characterized from the aquatic microbe *Isochrysis galbana*, and is assumed to be involved in the synthesis of methylene-interrupted PUFAs (Qi et al., 2004). Sayanova et al. (2007) further showed that the co-expression of the Δ^5-desaturase gene with a C18Δ^9-elongase in transgenic *Arabidopsis* plants resulted in the production of sciadonic acid. In the present study, two and one unigenes encoding Δ^5- desaturase, and one and two unigenes encoding elongase were detected in Yuanfei and Xifei, respectively (Fig. 6), which indicated that Δ^5- desaturase and elongase were responsible for the synthesis of sciadonic acid in *T. grandis*. Differential transcription of these related genes in both cultivars is described in detail in Figs. 6 and 8B. Unigenes encoding Δ^5-desaturase and elongase in the kernel tissue in Yuanfei were up-regulated relative to other tissues, whereas in Xifei, Δ^5-desaturase unigenes were up-regulated but unigenes encoding elongase were down-regulated relative to other tissues. This result was consistent with expression levels of genes for Δ^5-desaturase and elongase in both cultivars, where elongase in Yuanfei kernel tissue had higher expression than in Xifei, but the reverse was observed for Δ^5-desaturase. Thus, we deduced that the high content of sciadonic acid in Yuanfei seeds was attributed to the activity and abundance of elongase enzymes. Further investigation of the relationship between sciadonic acid biosynthesis genes and artificial regulation should provide novel insights into increasing sciadonic acid content in *T. grandis* using genetic engineering or molecular breeding.

Furthermore, in our study, elongase genes in Yuanfei and Xifei are believed to represent *TgELO1* and *TgELO2*, which are close to *OtELO1* (a putative C18 Δ^6-elongase) in algae, and possessed two characteristic

motifs (L–H–X–X–H–H and M–Y–X–Y–Y, boxed in Fig. 7A) conserved in the ELO-like protein sequences (Kajikawa et al., 2006a,b). Sequence identity comparisons showed 53.3% identity to the previously annotated Δ^6-elongase sequence, indicating that elongating activity in the biosynthesis of sciadonic acid is governed by Δ^6-elongase (Fig. 7). Further investigation is necessary to examine the role of elongase involved in sciadonic acid biosynthesis *via* genetic engineering or artificial regulation.

5. Conclusion

Fatty acid composition in the root, stem, leaf, aril, and kernel tissue of *T. grandis* cv. Yuanfei and Xifei was analyzed in this study. Our results indicate that *T. grandis* seed oil could be beneficial to human health owing to its high levels of sciadonic acid. Additionally, we report here the first transcriptome dataset derived from high-throughput sequencing technology (Illumina) for the root, stem, leaf, aril, and kernel tissue of *T. grandis* cv. Yuanfei and Xifei. Transcriptome analyses from *T. grandis* resulted in 80,235 and 82,454 unigenes in Yuanfei and Xifei, respectively. Genes involved in the putative *de novo* fatty acid and sciadonic acid biosynthesis pathways were identified in both transcriptomes. In particular, candidate genes such as Δ^6-elongase and Δ^5-desaturase that are associated with sciadonic acid biosynthesis should provide insight into the molecular mechanisms underlying the high levels of sciadonic acid biosynthesis in *T. grandis* seeds.

Conflicts of interest

The authors declare that there are no conflicts of interest.

Acknowledgments

We thank LetPub (www.letpub.com) for its linguistic assistance during the preparation of this manuscript. This work was financially supported by the Special Fund for Forest Scientific Research in the Public Welfare (201504708), the National Natural Science Foundation of China (31670687) and the Open fund of Zhejiang Key Laboratory for Modern Silvicultural Technology (201607).

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.indcrop.2018.04.041.

References

Aitzetmuller, K., 1995. Fatty acid patterns of Ranunculaceae seed oils: phylogenetic relationships. Plant Syst. Evol. 9, 229–240.
Beisson, F., Koo, A.J.K., Ruuska, S., Schwender, J., Pollard, M., Thelan, J.J., Paddock, T., Salas, J.J., Savage, L., Milcamps, A., Mhaske, V.B., Cho, Y., Ohlrogge, J.B., 2003. Arabidopsis genes involved in acyl lipid metabolism. A 2003 census of the candidates a study of the distribution of expressed sequence tags in organs, and a web-based database. Plant Physiol. 132, 681–697.
Berger, A., Baur, M., Charbonnet, C., Safonova, I., Jomard, A., 2002. Epidermal anti-inflammatory properties of 20:3$\Delta^{5,11,14}$: effects on mouse ear edema, PGE2 levels in cultured keratinocytes, and PPAR activation. Lipids Health Dis. 1, 1–12.
Brown, A.P., Kroon, J.T.M., Swarbreck, D., Febrer, M., Larson, T.R., Graham, I.A., Caccamo, M., Slabas, A.R., 2012. Tissue-specific whole transcriptome sequencing in castor, directed at understanding triacylglycerol lipid biosynthetic pathways. PLoS One 7 (2), e30100.
Chi, X.Y., Yang, Q.L., Lu, Y.D., Wang, J.Y., Zhang, Q.F., Pan, L.J., Chen, M.N., He, Y.N., Yu, S.L., 2011. Genome-wide analysis of fatty acid desaturases in soybean (Glycine max). Plant Mol. Biol. Rep. 29 (4), 769–783.
Dai, F., Chen, Z.H., Wang, X., Li, Z., Jin, G., Wu, D., Cai, S., Wang, N., Wu, F., Nevo, E., Zhang, G., 2014. Transcriptome profiling reveals mosaic genomic origins of modern cultivated barley. Proc. Natl. Acad. Sci. U. S. A. 111 (37), 13403–13408.
Grabherr, M.G., Haas, B.J., Yassour, M., Levin, J.Z., Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q.D., Chen, Z.H., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., Palma, F.D., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., Regev, A., 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. Nat. Biotechnol. 29, 644–652.
Harwood, J.L., 1996. Recent advances in the biosynthesis of plant fatty acids. Biochim. Biophys. Acta 1301, 7–56.
Hernández, M.L., Sicardo, M.D., Martínez-Rivas, J.M., 2016. Differential contribution of endoplasmic reticulum and chloroplast ω-3 fatty acid desaturase genes to the linolenic acid content of Olive (*Olea europaea*) Fruit. Plant Cell Physiol. 57 (1), 138–151.
Huang, R.M., Huang, Y.J., Sun, Z.C., Huang, J.Q., Wang, Z.J., 2017. Transcriptome analysis of genes involved in lipid biosynthesis in the developing embryo of Pecan (*Carya illinoinensis*). J. Agric. Food Chem. 65, 4223–4236.
Hyun, K.S., Song, J.Y., Jeong, J.E., Young, L.K., Young, S.S., Han, L.Y., Min, A.Y., Sik, K.Y., 2012. Genetic association of thee gr2gene with bipolar disorder in Korea. Exp. Mol. Med. 44 (2), 121–129.
Itabashi, Y., Takagi, T., 1982. Cis-5-olefinic nonmethylene-interrupted fatty acids in lipids of seeds, arils and leaves of Japanese Yew. J. Jpn. Oil Chem. Soc. 31 (9), 574–579.
Kajikawa, M., Yamato, K.T., Kohzu, Y., Shoji, S., Matsui, K., Tanaka, Y., Sakai, Y., Fukuzawa, H.A., 2006a. Front-end desaturase from Chlamydomonas reinhardtii produces pinolenic and coniferonic acids by omega-13 desaturation in methylotrophic yeast and tobacco. Plant Cell Physiol. 47, 64–73.
Kajikawa, M., Yamato, K.T., Sakai, Y., Fukuzawa, H., Ohyama, K., Kohchi, T., 2006b. Isolation and functional characterization of fatty acid Δ^5-elongase gene from the liverwort *Marchantia polymorpha* L. FEBS Lett. 580, 149–154.
Kanehisa, M., Araki, M., Goto, S., Hattori, M., Hirakawa, M., Itoh, M., Katayama, T., Kawashima, S., Okuda, S., Tokimatsu, T., Yamanishi, Y., 2008. KEGG for linking genomes to life and the environment. Nucleic Acids Res. 36, 480–484.
Li, Z.J., Luo, C.F., Chen, X.J., Feng, X.J., Yu, W.W., 2005. Component analysis and nutrition evaluation of seeds of Torreya grandis 'Merrillii'. J. Zhejiang For. Coll. 22 (5), 540–544 (in Chinese with English abstract).
Li, S.S., Wang, L.S., Shu, Q.Y., Wu, J., Chen, L.G., Shao, S., Yin, D.D., 2015. Fatty acid composition of developing tree peony (*Paeonia* section *Moutan* DC.) seeds and transcriptome analysis during seed development. BMC Genom. 16 (1), 208.
Li-Beisson, Y., Shorrosh, B., Beisson, F., Andersson, M.X., Arondel, V., Bates, P.D., Baud, S., Bird, D., Debono, A., Durrett, T.P., Franke, R.B., Graham, I.A., Katayama, K., Kelly, A.A., Larson, T., Markham, J.E., Miquel, M., Molina, I., Nishida, I., Rowland, O., Samuels, L., Schmid, K.M., Wada, H., Welti, R., Xu, C., Zallot, R., Ohlrogge, J., 2010. Acyl-lipid metabolism. Arabidopsis Book 8 (8), e0133.
Mao, X., Cai, T., Olyarchuk, J.G., Wei, L., 2005. Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. Bioinformatics 21, 3787–3793.
Matsuda, O., Sakamoto, H., Hashimoto, T., Iba, K., 2005. A temperature-sensitive mechanism that regulates post-translational stability of a plastidial omega-3 fatty acid desaturase (FAD8) in *Arabidopsis* leaf tissues. J. Biol. Chem. 280 (5), 3597–3604.
Meyer, A., Kirsch, H., Domergue, F., Abbadi, A., Sperling, P., Bauer, J., Cirpus, P., Zank, T.K., Moreau, H., Roscoe, T.J., Zähringer, U., Heinz, E., 2004. Novel fatty acid elongases and their use for the reconstitution of docosahexaenoic acid biosynthesis. J. Lipid Res. 45, 1899–1909.
Niu, L.Y., Wu, X.Q., Zhang, Y., 2011. Analysis on fatty acid and unsaponifiable matter in *Torreya grandis* var. Merriii seed oil. J. Chin. Cereals Oils Assoc. 26 (6), 52–55 (in Chinese with English abstract).
Nystedt, B., Street, N.R., Wetterbom, A., Zuccolo, A., Lin, Y.C., Scofield, D.G., Vezzi, F., Delhomme, N., Giacomello, S., Alexeyenko, A., Vicedomini, R., Sahlin, K., Sherwood, E., Elfstrand, M., Gramzow, L., Holmberg, K., Hällman, J., Keech, O., Klasson, L., Koriabine, M., Kucukoglu, M., Käller, M., Luthman, J., Lysholm, F., Niittylä, T., Olson, A., Rilakovic, N., Ritland, C., Rosselló, J.A., Sena, J., Svensson, T., Talavera-López, C., Theißen, G., Tuominen, H., Vanneste, K., Wu, Z.Q., Zhang, B., Zerbe, P., Arvestad, L., Bhalerao, R., Bohlmann, J., Bousquet, J., Gil, R.G., Hvidsten, T.R., Jong, P., MacKay, J., Morgante, M., Ritland, K., Sundberg, B., Thompson, S.L., Peer, Y.V., Andersson, B., Nilsson, O., Ingvarsson, P., Lundeberg, J., Jansson, S., 2013. The Norway spruce genome sequence and conifer genome evolution. Nature 497 (7451), 579–584.
Ohlrogge, J., Browse, J., 1995. Lipid biosynthesis. Plant Cell 7, 957–970.
Qi, B.X., Fraser, T., Mugford, S., Dobson, G., Sayanova, O., Butler, J., Napier, J.A., Stobart, A.K., Lazarus, C.M., 2004. Production of very long chain polyunsaturated omega-3 and omega-6 fatty acids in plants. Nat. Biotechnol. 22, 739–745.
Saeed, M.K., Deng, Y., Parveen, Z., Dai, R., Ahmad, W., Yu, Y., 2007. Studies on the chemical constituents of *Torreya grandis* fort. ex lindl. J. Appl. Sci. 7 (2), 269–273.
Sayanova, O., Haslam, R., Caleron, M.V., Napier, J.A., 2007. Cloning and characterization of unusual fatty acid desaturases from *Anemone leveillei*: identification of an acyl-coenzyme A C20Δ^5-desaturase responsible for the synthesis of sciadonic acid. Plant Physiol. 144, 455–467.
Shen, C.H., Hu, Y.Y., Du, X.H., Li, T.T., Tang, H., Wu, J.S., 2014. Salicylic acid induces physiological and biochemical changes in *Torreya grandis* cv. Merrillii seedlings under drought stress. Trees 28, 961–970.
Simopoulos, A.P., 2002. The importance of the ratio of omega-6/omega-3 essential fatty acids. Biomed. Pharmacother. 56 (8), 365–379.
Simopoulos, A.P., 2006. Evolutionary aspects of diet, the omega-6/omega-3 ratio and genetic variation: nutritional implications for chronic diseases. Biomed. Pharmacother. 60 (9), 502–507.
Storey, J.D., 2002. A direct approach to false discovery rates. J. R. Stat. Soc. 64, 479–498.
Tanaka, T., Morishige, J., Takimoto, T., Takai, Y., Satouchi, K., 2001. Metabolic characterization of sciadonic acid (5c,11c,14c-eicosatrienoic acid) as an effective substitute for arachidonate of phosphatidylinositol. Eur. J. Biochem. 268, 4928–4939.
Wang, Y.B., Liu, B.T., Qin, Y.C., Wang, L.L., Tong, X.Q., Tong, P.Z., Qian, H., 2016. Compositions of fatty acid and flavor compounds in different cultivars of *Torrey grandis* seed oil. Chin. Oil Fat 41 (2), 101–105 (in Chinese with English abstract).
Wolff, R.L., Comps, B., Deluc, L.G., Marpeau, A.M., 1997a. Fatty acids of the seeds from pine species of the *Ponderosa Banksiana* and *Halepensis* sections. The peculiar taxonomic position of *Pinus pinaster*. J. Am. Oil Chem. Soc. 75, 45–50.

Wolff, R.L., Comps, B., Marpeau, A.M., Deluc, L.G., 1997b. Taxonomy of pinus species based on the seed oil fatty acid compositions. Trees 12 (2), 113–118.
Wolff, R.L., Deluc, L.G., Marpeau, A.M., Comps, B., 1997c. Chemotaxonomic differentiation of conifer families and genera based on the seed oil fatty acid compositions: multivariate analyses. Trees 12, 57–65.
Wolff, R.L., Christie, W.W., Pedrono, F., Marpeau, A.M., Tsevegsuren, N., Aitzetmuller, K., Gunstone, F.D., 1999a. Δ5-olefinic acids in the seed lipids from four *Ephedra* species and their distribution between the alpha and beta positions of triacylglycerols. characteristics common to coniferophytes and cycadophytes. Lipids 34, 855–864.
Wolff, R.L., Pédrono, F., Marpeau, A.M., Gunstone, F.D., 1999b. The seed fatty acid composition and the distribution of Δ^5-olefinic acids in the triacylglycerols of some *Taxares* (*Cephalotaxus* and *Podocarpus*). J. Am. Oil Chem. Soc. 76, 469–473.
Wolff, R.L., Lavialle, O., Pedrono, F., Pasquier, E., Deluc, L.G., Marpeau, A.M., Aitzetmuller, K., 2001. Fatty acid composition of *Pinaceae* as taxonomic markers. Lipids 36, 439–451.
Wolff, R.L., 1999. The phylogenetic significance of sciadonic (all-cis 5, 11, 14–20: 3) acid in gymnosperms and its quantitative significance in land plants. J. Am. Oil Chem. Soc. 76 (12), 1515–1516.
Yi, G.M., Bao, Y.C., 2016. Sequencing and bioinformatic analysis for transcriptome of *Torreya grandis* Fort. ex Lindl. cv. merrillii. J. Shandong Agric. Univ. 47 (1), 19–24 (in Chinese with English abstract).
Young, M.D., Wakefield, M.J., Smyth, G.K., Oshlack, A., 2010. Gene ontology analysis for rna-seq: accounting for selection bias. Genome Biol. 11 (2), R14.
Zeng, Y.R., Ye, S.Y., Yu, W.W., Wu, S., Hou, W., Wu, R.L., Dai, W.S., Chang, J., 2014. Genetic linkage map construction and QTL identification of juvenile growth traits in *Torreya grandis*. BMC Genet. 15 (1), S2.
Zheng, C., Zhao, L., Wang, Y., Shen, J.Z., Zhang, Y.F., Jia, S.S., Li, Y.S., Ding, Z.T., 2015. Integrated RNA-Seq and sRNA-seq analysis identifies chilling and freezing responsive key molecular players and pathways in tea plant (*Camellia sinensis*). PLoS One 10 (4), e0125031.
Zimin, A.V., Stevens, K.A., Crepeau, M.W., Puiu, D., Wegrzyn, J.L., Yorke, J.A., Langley, C.H., Neale, D.B., Salzberg, S.L., 2017. An improved assembly of the loblolly pine mega-genome using long-read single-molecule sequencing. GigaScience 6 (1), 1.