Industrial Crops & Products 131 (2019) 182–193

Contents lists available at ScienceDirect

Industrial Crops & Products

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

Comparative transcriptome analysis reveals key genes in the regulation of squalene and β-sitosterol biosynthesis in *Torreya grandis*

Jinwei Suo[1], Ke Tong[1], Jiasheng Wu*, Mingzhu Ding, Wenchao Chen, Yi Yang, Heqiang Lou, Yuanyuan Hu, Weiwu Yu, Lili Song*

State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou, Zhejiang 311300, China

ARTICLE INFO

Keywords:
Torreya grandis
Cultivars
Transcriptome
Squalene
β-Sitosterol
Biosynthesis

ABSTRACT

Torreya grandis (*T. grandis,* Taxaceae) is both economically and medicinally valuable species being rich in various bioactive compounds (e.g., squalene and β-sitosterol). However, the contents of these compounds are various and cultivar specific, and the complicated regulatory mechanisms of their biosynthesis in *T. grandis* are still unknown. To uncover the underlying molecular mechanisms that control the differences in the accumulation of squalene and β-sitosterol, a comprehensive transcriptome was constructed from nine different *T. grandis* cultivars. A total of more than 60,372 unigenes were obtained, of which over 60% were successfully annotated. Identification and expression analysis of the differentially-expressed genes (DEGs) showed that 39 candidate genes were involved in squalene and β-sitosterol biosynthesis in *T. grandis* seeds. In particular, the expression patterns of genes related to the mevalonate (MVA) and methylerythritol phosphate (MEP) pathways indicates that both pathways promote the upstream biosynthesis of isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) in different *T. grandis* cultivars. Moreover, several key regulatory steps controlling the differential accumulation of squalene and β-sitosterol between *T. grandis* cultivars were also discussed.

1. Introduction

Torreya is a genus in the yew family (Taxaceae) that is composed of six species and two varieties, which are distributed in localized areas of China, the United States, Japan, and Korea (Kang and Tang, 1995). *Torreya grandis* (*T. grandis*), a large and ancient evergreen tree species with linear, bristle-pointed leaves, dioecious flowers, and drupe-like fruits with nut-seeds, is native to China (He et al., 2016; Yu et al., 2016). *T. grandis* seed is not only one of the world's most rare nut, but is also an important traditional medicine that is widely used to cure coughs, excess phlegm, diarrhea, to expel parasites, and prevent malnutrition (Ni and Shi, 2014; Ni et al., 2015; Wu et al., 2018). The aforementioned medicinal quality of this plant is largely a result of the various bioactive compounds, most notably squalene and phytosterol found in the seeds (Chen et al., 2006; Ni et al., 2015; He et al., 2016).

Squalene and phytosterols are the compounds present in the unsaponifiable lipid fraction of kernel oils and beneficial to health (Weihrauch and Gardner, 1978; Delgado-Zamarreno et al., 2009). Squalene is a 30-carbon isoprenoid with six double bonds which acts as the key intermediate for the biosynthesis of phytosterols (Liao et al.,

Abbreviations: CAS, cycloartenol synthase; CMK, 4-(cytidine-5-diphospho)-2-C-methy-D-erythritol kinase; COG, clusters of orthologous groups of proteins; DEGs, differentially expressed genes; DMAPP, dimethylallyl diphosphate; DOXP, 1-deoxy-D-xylulose-5-phosphate; DXR, 1-deoxy-D-xylulose 5-phosphate reductoisomerase; DXS, 1-deoxy-D-xylulose-5-phosphate synthase; FPKM, fragments per kilobase per million fragments; FPP, farnesyl diphosphate; FPS, farnesyl diphosphate synthase; GAP, glyceraldehyde-3-phosphate; GFP, green fluorescent protein; GGPP, geranylgeranyl diphosphate; GGPS, geranyl diphosphate synthase; GPS, geranyl diphosphate synthase; GO, gene ontology; GPP, geranyl diphosphate; HDR, 1-hydroxy-2-methyl-2-(*E*)-butenyl-4-diphosphate reductase; HMBPP, (*E*)-4-hydroxy-3-methylbut-2-enyl diphosphate; HMGR, 3-hydroxy-3-methyl glutaryl coenzyme A reductase; HMGS, 3-hydroxy-3-methyl glutaryl coenzyme A synthase; IDSs, isoprenyl diphosphate synthases; IPI, isopentenyl diphosphate isomerase; IPP, isopentenyl diphosphate; KEGG, kyoto encyclopedia of genes and genomes; MCS, 2-C-methyl-D-erythritol-2,4-cyclodiphosphate synthase; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; MDC, mevalonate-5-pyrophosphate decarboxylase; MEP, methylerythritol phosphate; MK, mevalonate kinase; MVA, mevalonate; NCBI, national center for biotechnology information; Nr, NCBI non-redundant protein sequences; Nt, NCBI non-redundant nucleotide sequences; ORF, open reading frames; PMK, phosphomeralonate kinase; RT-qPCR, quantitative real-time PCR; SMTs, Δ24-sterol methyl transferases; SQE, squalene epoxidase; SQS, squalene synthase; SwissProt, SwissProt protein sequence database

* Corresponding authors.
E-mail addresses: wujs@zafu.edu.cn (J. Wu), lilisong@zafu.edu.cn (L. Song).
[1] These authors have contributed equally to this work.

https://doi.org/10.1016/j.indcrop.2019.01.035
Received 17 August 2018; Received in revised form 15 January 2019; Accepted 20 January 2019
Available online 01 February 2019

2016). In plants, the dominant phytosterols include β-sitosterol, campesterol, and stigmasterol, the first one being the most abundant. Squalene and β-sitosterol are considered pharmacologically significant in antimicrobial, anti-inflammatory, anti-oxidative, anticancer, and immunomodulating effects (Ye and Chang, 2010; Ambavade et al., 2014). Also, squalene is an important constituent of skin-care products, oxidation-resistant industrial lubricants, and numerous vaccines (Fox, 2009). These compounds' highly coveted pharmacological and biological characteristics have given rise to increasing demand for squalene and β-sitosterol, and their biosynthetic pathway has attracted a substantial amount of attention (Paramasivan et al., 2018; Qiao et al., 2018; Ramadan et al., 2019).

Thus far, the biosynthetic pathways of squalene and phytosterols have been well documented (Valitova et al., 2016), as both are derived from a common 5-carbon precursor, isopentenyl diphosphate (IPP) and its isomer dimethylallyl diphosphate (DMAPP), which are formed via either the cytosolic mevalonate (MVA) pathway or the plastidial methylerythritol phosphate (MEP) pathway (Nagegowda, 2010). The MVA pathway begins with 3-acetyl-CoA giving rise to IPP, while the MEP pathway synthesizes IPP and DMAPP from pyruvate and glyceraldehyde-3-phosphate (GAP). Among them, 3-hydroxy-3-methyl-glutaryl-coenzyme A synthase (HMGS) and 3-hydroxy-3-methyl-glutaryl-coenzyme A reductase (HMGR) catalyze the first and second committed steps in the MVA pathway (Bach, 1995). The overexpression of both leads to a higher sterol content in numerous plants, such as *Arabidopsis*, *Nicotiana tabacum*, and *Ginkgo biloba* (Harker et al., 2003; Wang et al., 2012; Li et al., 2014). Furthermore, it has been reported that HMGR activity in *Arabidopsis* has been observed to be regulated not only at the transcriptional level but also at the post-translational level (Hemmerlin et al., 2013). In addition, 1-deoxy-D-xylulose-5-phosphate synthase (DXS), 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR), and 1-hydroxy-2-methyl-2-(*E*)-butenyl-4-diphosphate reductase (HDR) are also reported as the key enzymes in controlling carbon flux in the MEP pathway (Rodríguez-Concepción, 2006). The overexpression of *WsDXR2* from *Withania somnifera* in transgenic tobacco (*Nicotiana tabacum* L.) exhibited a 2.5- to 3-fold increase in enzyme activity and an increased accumulation of sterols (Singh et al., 2014). In addition, a positive correlation between HDR transcript level and plastid isoprenoid accumulation was observed in tomato (*Lycopersicon esculentum* L.) fruit and *Arabidopsis* seedlings (Botella-Pavía et al., 2004).

In the downstream pathway of phytosterol biosynthesis, IPP and DMAPP precursors are further converted to C_{15} farnesyl diphosphate (FPP) by farnesyl diphosphate synthase (FPS), then squalene synthase (SQS) catalyzes the condensation of two molecules of FPP to produce squalene (Pandit et al., 2000). Subsequently, squalene epoxidase (SQE) converts squalene to 2,3-oxidosqualene, and cycloartenol synthase (CAS) catalyzes 2,3-oxidosqualene to cycloartenol, which is finally converted into end-product phytosterols after alkylation (catalyzed by two distinct sterol C24-methyltransferases [SMTs]). In recent years, these key genes (i.e., *SQS*, *SQE*, *SMT1*, and *SMT2*) have been cloned and widely studied in transgenic plants (Holmberg et al., 2002; Busquets et al., 2008; Neelakandan et al., 2010; Laranjeira et al., 2015). The overexpression of membrane-bound SQS results in elevated levels of phytosterols, while the expression of different copies of SQS exhibited tissue- and organ-specific regulation in many plants (Busquets et al., 2008; Kim et al., 2011). Both *AtSQE1* and *AtSQE3* seem to be functionally important in the biosynthesis of sterols, where *AtSQE3* is more likely to be involved in root tissues (Laranjeira et al., 2015). The quantity and composition of sterols in leaf and seed tissues were altered in response to *SMT1* overexpression in tobacco (Holmberg et al., 2002). These results suggest that the regulation of sterol biosynthesis is plant and tissue specific. Thus, a comprehensive understanding of squalene and phytosterol biosynthesis is necessary to uncover the key genes and critical regulatory steps for further investigation into increasing squalene and phytosterol levels in plants in a beneficial manner.

In the present study, a comparative transcriptomic analysis of *T. grandis* seeds from nine different cultivars were carried out to detect candidate genes involved in squalene and phytosterols biosynthesis, and analyze the expression of these candidate genes. Subsequently, the relative expression of the candidate genes was validated via quantitative real-time PCR (RT-qPCR). Combined with the physiological analysis of squalene and β-sitosterol contents in different *T. grandis* cultivars, a more complete picture of the squalene and β-sitosterol biosynthetic pathway and their underlying regulatory mechanisms in *T. grandis* seeds were provided. This study will not only improve our understanding of this important pathway, provide a valuable genomic resource for the further exploration of the expression and regulation mechanisms underlying squalene and β-sitosterol biosynthesis, but it will also be helpful for further genetic engineering and molecular breeding of productive *T. grandis* cultivars with high squalene and phytosterol levels.

2. Materials and methods

2.1. Plant materials

Seed samples of nine *T. grandis* cultivars (named Z06, A19, R18, R26, C03, Z03, Y11, Y37, and X039) used in the study were collected from five cities of Zhejiang province (namely Zhuji, Hangzhou, Shengzhou, Shaoxing and Fuyang) and one city of Anhui province (Huangshan) (Additional file 1). The trees were about fifty-year-old, and 800 g (90 seeds) were collected at four direction branches from three trees of each cultivar. After collection, the sarcotesta (arils) and the testa (seed coat) were removed, and the remaining hard seeds were used for kernel sampling. Subsequently, the kernel samples were sectioned separately, immediately frozen in liquid nitrogen, and stored at −80 °C until further analysis.

2.2. Squalene and β-sitosterol analysis

Frozen *T. grandis* kernel samples from each cultivar were ground into powder, and crude oil was extracted in hexane, filtered through a 0.2 μm cellulose acetate filter, and stored at −20℃ in the dark for squalene and β-sitosterol isolation (Fernández-Cuesta et al., 2013). The extraction of squalene and β-sitosterol from the nine cultivars of *T. grandis* kernel oil was determined according to the method of Giacometti (2001) with slight modifications. Approximately 100 mg crude oil was mixed thoroughly with 2 ml of 2% KOH (w/v) ethanol solution. Samples were vortexed briefly and then incubated at 80 °C for 15 min in a water bath. The unsaponifiable lipid fraction was extracted with 2 ml hexane and centrifuged at 340 × *g* for 10 min. Next, the upper hexane layer was transferred into a new tube and maintained in oven at 37.5 °C overnight. Finally, dried pellets were redissolved in 100 μl of Sylon BTZ (Supelco, Bellefonte, PA, USA) and conserved at −20 °C until further analysis.

Gas chromatography of squalene and β-sitosterol were performed using an Agilent 7890 gas chromatograph (Agilent Technologies, Santa Clara, CA, USA) according to the methods of Górnaś et al. (2016) with slight modifications. A 0.5 μl aliquot of each sample was injected in splitless mode onto a HP-5 capillary column (id = 0.32 mm, length = 30 m, film thickness = 0.25 μm). The injector and detector temperatures were 300 °C. The initial column temperature was set to 160 °C for 2 min, then programmed to 280 °C at a rate of 15 °C/min, held for 7 min, increased at 5 °C/min to 300 °C, and held for 5 min. Standard stock solutions of squalene (1000 ng/μl) and β-sitosterol (1000 ng/μl) were used as the external standards for quantifications, and the concentrations of squalene and β-sitosterol were calculated according to the standard curves.

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

Total RNA was isolated from the kernel of nine *T. grandis* cultivars,

respectively, using the RNAprep Pure Plant Kit (TIANGEN, Beijing, China). RNA quality was assessed by 1% agarose gel electrophoresis. All RNA samples were quantified and examined (RNA purity and integrity) using a NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies Inc., Wilmington, DE, USA). A total of 3 μg of *T. grandis* kernel RNA per sample was used as input material, and cDNA libraries were constructed according to the manufacturer's instructions. Subsequently, library preparations were assessed using the Agilent Bioanalyzer 2100 system (Agilent Technologies, CA, USA) and sequenced using the Illumina HiSeq™ 4000 platform (Illumina Inc., San Diego, CA, USA).

2.4. De novo assembly and functional annotation

After the raw data was filtered to remove low-quality reads and adaptor sequences, the high-quality sequence reads were further assembled to obtain unigenes using Trinity software (version 6.0). The functional annotation of the unigenes was obtained by searching against the following six public databases: NCBI non-redundant protein sequences (Nr), NCBI non-redundant nucleotide sequences (Nt), SwissProt protein sequence database (SwissProt), Clusters of Orthologous Groups of proteins (COG), the Kyoto Encyclopedia of Genes and Genomes (KEGG), and Gene Ontology (GO).

2.5. Differentially-expressed genes (DEGs) analysis and function enrichment

The expression levels of the unigenes from nine *T. grandis* cultivars were calculated using the fragments per kilobase per million fragments (FPKM) method (Mortazavi et al., 2008), and the analysis of the differential expression of DEGs was carried out using the "DESeq" R package (version 1.10.1) with a significant threshold of $|\log_2$ (fold change)$| \geq 2$, a false discovery rate (FDR) of < 0.001, and a *p*-value of < 0.05 (Anders and Huber, 2010). GO and KEGG enrichment analysis were performed to investigate the functions of the DEGs. The "GOseq" R package was used for the GO enrichment analysis based on Wallenius non-central hyper-geometric distribution (Young et al., 2010), while the KEGG statistical enrichment analysis was tested using KOBAS software (version 2.0) (Kanehisa et al., 2012).

2.6. Validation of gene expression profiles by RT-qPCR

In this study, RT-qPCR was performed to validate the expression patterns of the candidate genes— determined by RNA-seq analysis—involved in squalene and phytosterol biosynthesis. Specific primer pairs (Additional file 2) used in RT-qPCR were designed by the online software program Primer3 (version 0.4.0) (http://bioinfo.ut.ee/primer3-0.4.0/). RT-qPCR amplification was performed on a CFX96 Touch Real-Time PCR System (Bio-Rad, California, USA) using the SYBR® Green Real-Time PCR Master Mix (Toyobo, Osaka, Japan) according to the following program: an initial denaturation step of 95 °C for 10 min, followed by 45 cycles of 95 °C for 10 s, 57 °C for 10 s, and 72 °C for 20 s. The relative expression of each gene was normalized to actin and calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001). All samples were analyzed in at least three biological and technical replicates.

2.7. Subcellular localization of TgFPS

To investigate the subcellular localization of key proteins in squalene and β-sitosterol biosynthesis pathway, green fluorescent protein (GFP) fusions were constructed. Full-length open reading frames (ORF) of *TgFPS* (without the stop codon) were amplified by PCR using appropriate primers (Additional file 3). Then the PCR products were cloned into the binary vector 35S::GFP (modified from pCAMBIA1300) to produce plasmids. GFP-protein fusion construct (*35S::TgFPS::GFP*) was transiently expressed by agroinfiltration (*Agrobacterium tumefaciens* strain GV3101) of tobacco (*N. benthamiana*) leaves (Sparkes et al., 2006). Images of leaf epidermal cells were examined using confocal laser-scanning microscopy (LSM710, Karl Zeiss).

2.8. Statistical analyses

All results in the present study are displayed as the mean ± standard deviation of at least three replicates. Statistical analyses were performed using a one-way ANOVA in SPSS 18.0 (SPSS Inc., Chicago, IL, USA). The least significant difference (LSD) was used to test the means and a *p*-value of < 0.05 was considered statistically significant.

3. Results

3.1. Squalene and β-sitosterol contents in the seeds of various T. grandis cultivars

Squalene and β-sitosterol levels were examined in *T. grandis* seeds from nine cultivars by gas chromatography, and the compounds were identified by comparing retention time with standard records of squalene (retention time of 11.0 min) and β-sitosterol (retention time of 16.7 min) (Fig. 1). According to the results, the contents of squalene and β-sitosterol varied depending on the cultivar (Fig. 2a and b). Squalene was accumulated to a much greater degree in the cultivars Y11, Y37, and X039, with the concentration of 61.87 mg/kg, 66.20 mg/kg, and 71.40 mg/kg, respectively. In contrast, β-sitosterol was present in the largest quantities in cultivars X039, Z03, and Y37, with the content of 3409.13 mg/kg, 3608.33 mg/kg, and 4082.40 mg/kg, respectively. Both compounds exhibited the minimum contents in the cultivars Z06 (14.92 ± 1.61 mg/kg) and A19 (1650.81 ± 76.97 mg/kg), respectively (Fig. 2a and b). In addition, moderate-strong positive correlations (correlation coefficient of 0.6239) were found between the contents of squalene and β-sitosterol (Fig. 2c). In addition, in investigating oil yield in the different *T. grandis* cultivars, oil content was found fluctuated between 20% to 50%, and was the highest in the cultivars Z06 and Y11 (Fig. 2d). A negative correlation was observed between oil yield and β-sitosterol content (Fig. 2e).

3.2. Transcriptome assembly, annotation, and the classification of unigenes in T. grandis cultivars

After filtering out low-quality reads and adaptor sequences, more than 58.29 million clean reads were obtained from nine different *T. grandis* cultivars, where the total sequence length for each cultivar was greater than 8,743,986,300 nt and the average GC content was 46.20% (Table 1). The Q30 and Q20 values were greater than 96.24% and 90.69% (Table 1). Subsequently, the *de novo* assembly of these high-quality reads resulted in a total of 60,372 to 66,948 unigenes from the nine *T. grandis* cultivars (Table 2). A functional annotation of the unigenes was performed by BLASTing the sequences against the Nr, Nt, SwissProt, COG, KEGG, and GO databases using an E-value threshold of 10^{-5}. As a result, total of 36,357 to 40,271 unigenes (over 60.22%) were successfully aligned to known genes in at least one of the six public databases, and 2,958 to 3,181 unigenes were annotated in all seven databases (Table 2).

Based on the GO functional annotations, a total of 59,071 unigenes were categorized into three main categories: biological process, cellular component, and molecular function (Additional file 4). There were a total of 23 classifications in the biological process category, where cellular process (33,868 unigenes) and metabolic process (34,147 unigenes) represented two of the largest groups. In the cellular component category, cell (30,414 unigenes) and cell part (30,226 unigenes) represented the two most abundant GO terms of the 20 different groups. In terms of the molecular function category, unigenes were clustered into 15 classifications, where binding (23,472 unigenes) and catalytic

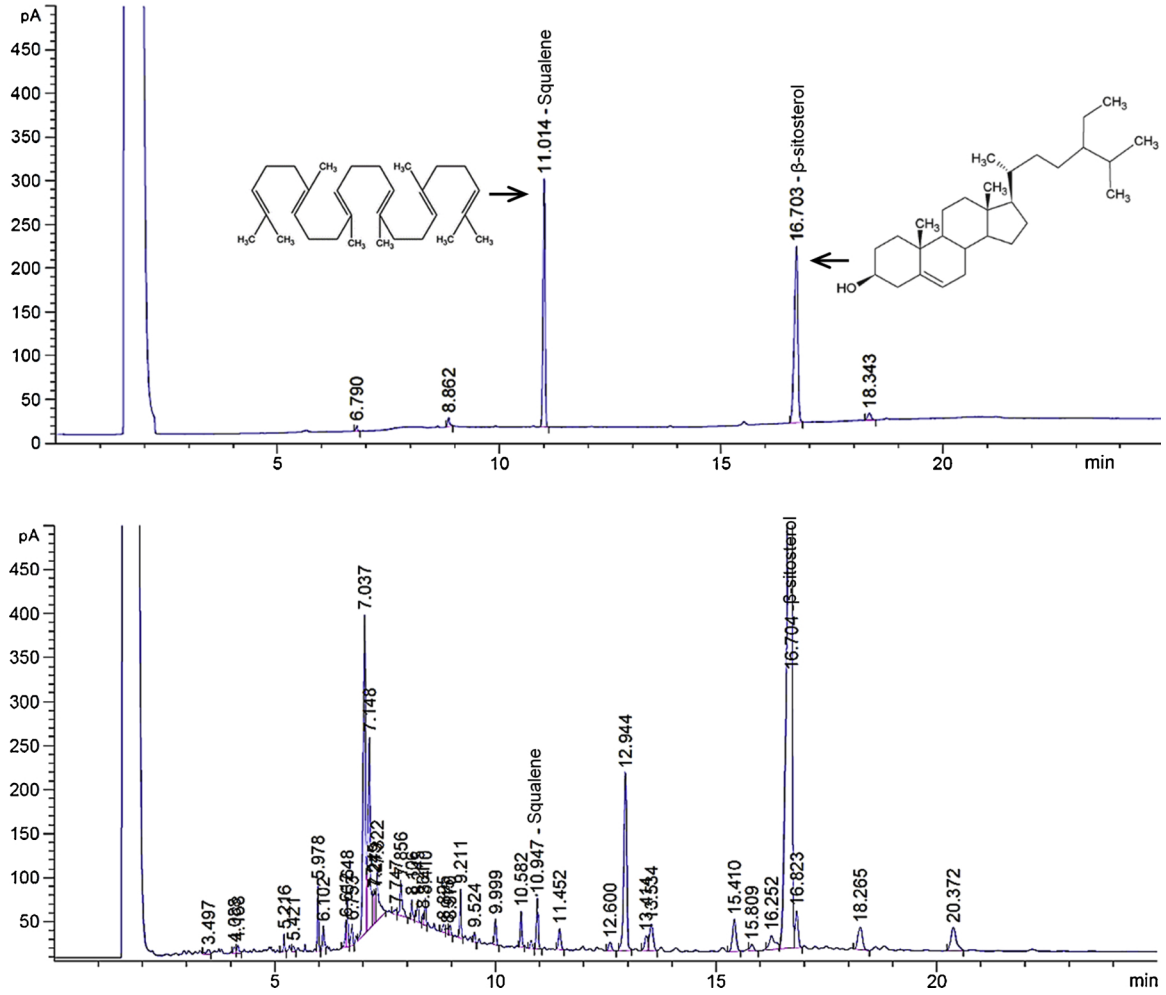

Fig. 1. Gas chromatogram of squalene and β-sitosterol of standard solution (a) and *Torreya grandis* seed sample (b). The retention times were approximately 11.0 min and 16.7 min for squalene and β-sitosterol, respectively.

activity (27,041 unigenes) were observed to be the two dominant subcategories (Additional file 4).

3.3. Analysis of DEGs from different T. grandis cultivars

In order to investigate DEGs in the nine *T. grandis* cultivars, expression levels of the unigenes were compared with cultivar Z06. It was observed that more than 3,521 upregulated and more than 7,492 downregulated unigenes were identified (Fig. 3). In order to further explore the functions of DEGs, a KEGG pathway enrichment analysis was performed. In total, 20 significantly enriched pathways in the up- and downregulated unigenes were found (Fig. 4). Metabolic pathways and the biosynthesis of secondary metabolites represented two of the dominant pathways associated with most of upregulated unigenes in *T. grandis* cultivars (Fig. 4a). Most notably, genes involved in sesquiterpenoid and triterpenoid biosynthesis and steroid biosynthesis exhibited high relative expression levels (Additional file 5). Of the downregulated unigenes, the enrichment analysis yielded five significantly pathways (each pathway contains more than 200 unigenes), including plant hormone signal transduction, starch and sucrose metabolism, RNA transport, mRNA surveillance, and spliceosome (Fig. 4b).

3.4. Identification and expression of candidate genes involved in squalene and β-sitosterol biosynthesis in T. grandis cultivars

Based on transcriptome results, a total of 39 candidate unigenes were identified to be involved in squalene and β-sitosterol biosynthetic pathways (Fig. 5; Additional file 6). Of these, nine transcripts were found to encode five enzymes in the MVA pathway: HMGS, HMGR, mevalonate kinase (MK), phosphomevalonate kinase (PMK), and mevalonate-5-pyrophosphate decarboxylase (MDC). Gene expression levels of the five enzymes were induced in most of the *T. grandis* cultivars in comparison with Z06 (Fig. 5; Additional file 6). In addition, 15 of the unigenes representing six enzymes in the plastid MEP pathway were also identified: DXS, DXR, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT), 4-(cytidine-5-diphospho)-2-C-methy-D-erythritol kinase (CMK), 2-C-methyl-D-erythritol-2,4-cyclodiphosphate synthase (MCS), and HDR. DXS catalyzes the first step in the MEP pathway to form 1-deoxy-D-xylulose-5-phosphate (DOXP), whereupon DOXP is further converted to MEP by DXR, and MEP is finally transformed into IPP by the consecutive enzymatic actions catalyzed by MCT, CMK, MCS, and HDR (Valitova et al., 2016). From the transcriptome results, DXS, MCT, CMK, and MCS were found to be upregulated in most *T. grandis* cultivars, while the expression of DXR and HDR were only increased in the *T. grandis* cultivars A19 and R18 in comparison with Z06 (Fig. 5; Additional file 6). Therefore, the gene expression levels of the majority of genes involved in the MVA and MEP pathways were found to be upregulated in most *T. grandis* cultivars, where the increased accumulation of these upstream components of the MVA and MEP pathways could contribute to the subsequent biosynthesis of squalene and β-sitosterol.

In terms of the downstream components of the squalene and β-sitosterol biosynthetic pathways, the reversible conversion between IPP

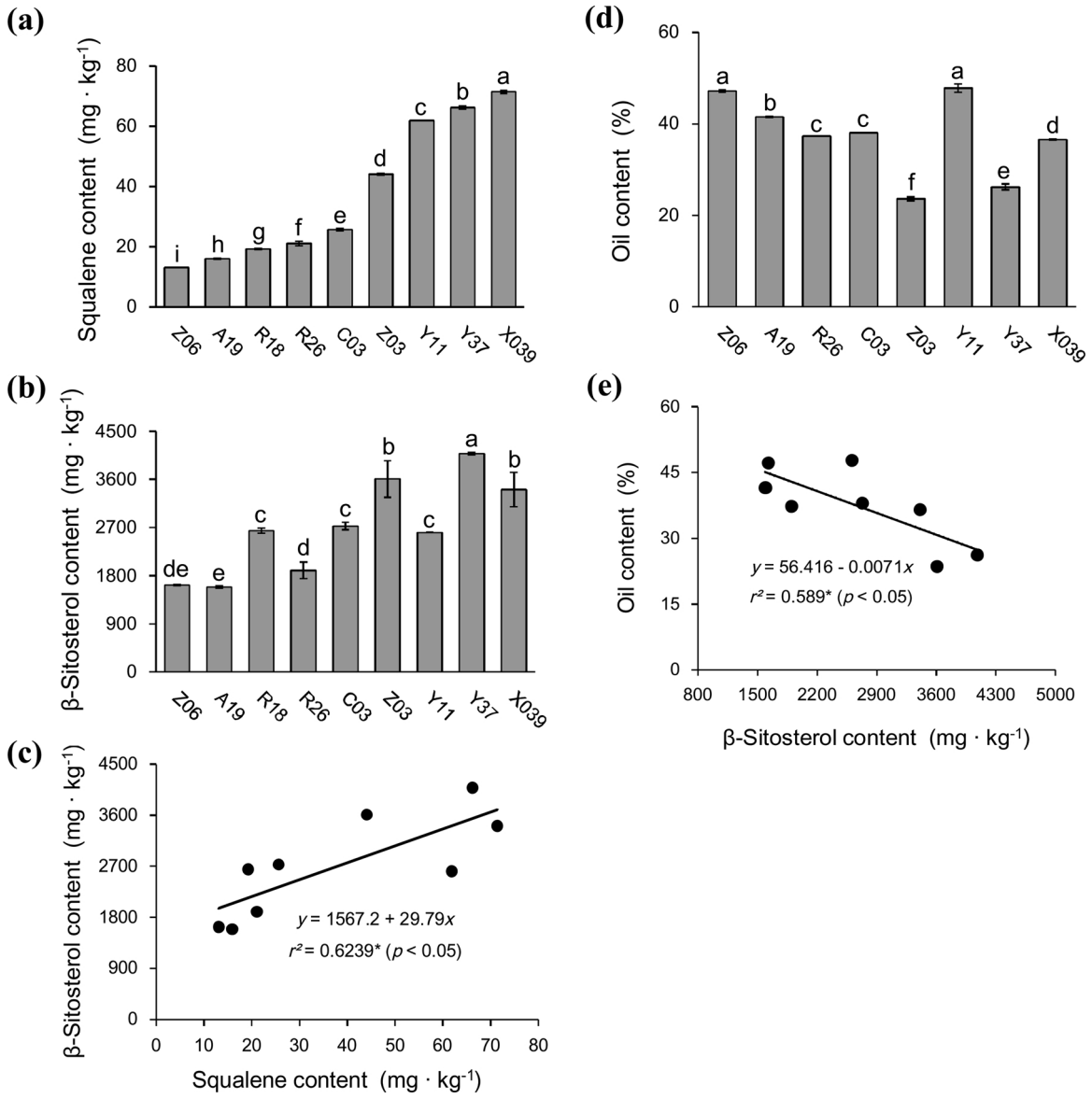

Fig. 2. Squalene and β-sitosterol content (a–b), correlation between squalene and β-sitosterol (c), oil content (d), and correlation between oil and β-sitosterol content (e) in seeds of different *Torreya grandis* cultivars.

Table 1
Summary of sequences analysis of *Torreya grandis* transcriptomes.

Cultivars	Clean reads (bp)	Clean nucleotide (nt)	Q20 [a] (%)	Q30 [b] (%)	GC content (%)
Z06	64,368,136	9,655,220,400	97.08%	92.25%	45.96%
A19	62,790,404	9,418,560,600	97.10%	92.52%	46.95%
R18	62,018,746	9,302,811,900	96.82%	91.90%	46.38%
R26	58,293,242	8,743,986,300	96.24%	90.69%	46.38%
C03	61,873,548	9,281,032,200	97.65%	93.61%	45.88%
Z03	59,683,260	8,952,489,000	97.20%	92.46%	45.57%
Y11	62,375,848	9,356,377,200	98.33%	95.62%	46.00%
Y37	64,010,630	9,601,594,500	98.37%	95.74%	46.79%
X039	65,105,340	9,765,801,000	98.34%	95.53%	45.87%

[a] Q20, the base quality score (Q score) was no less than 20.
[b] Q30, the base quality score was no less than 30.

and DMAPP is catalyzed by isopentenyl diphosphate isomerase (IPI), whereupon IPP and DMAPP are transformed to geranyl diphosphate (GPP, C_{10}), farnesyl diphosphate (FPP, C_{15}), and geranylgeranyl diphosphate (GGPP, C_{20}) by GPP synthase (GPS), FPP synthase (FPS), and GGPP synthase (GGPS), respectively. Subsequently, SQS catalyzes the condensation of FPPs to produce squalene (Pandit et al., 2000). A total of five transcripts—encoding IPI, GGPS, FPS, and SQS—were identified, three out of the five transcripts (GGPS, FPS, and SQS) were found to be upregulated in *T. grandis* cultivars, most notably in Y37 and X039 (Fig. 5; Additional file 6). Sequence analyses showed that the deduced amino acid sequence of *TgSQS* encodes 409 amino acids with three conserved domains (A–C) and one transmembrane region in the C-terminal region that is composed of 18–20 amino acid residues (Additional file 7). TgSQS exhibited a high degree of similarity with AtSQS1 and AtSQS2 sequences, and several conserved residues were also found to be present in their corresponding domains, especially the highly conserved Phe^{287} residue (Phe^{285} in TgSQS), which is considered to be important for the SQS activity of AtSQS1 (Additional file 7). Moreover, a correlation analysis found that the expression patterns of *TgSQS*—resulting from the *T. grandis* transcriptome analysis (FPKM-fold change)—exhibited a moderate positive correlation ($r^2 = 0.4705$; $p < 0.05$) with squalene content in different *T. grandis* cultivars (Fig. 6a). In addition, several enzymes (SQE, CAS, and SMTs) involved in the conversion of squalene to β-sitosterol were also identified (Additional file

Table 2
Statistics of annotations for assembled unigenes in the nine cultivars of *Torreya grandis* transcriptomes.

Cultivars	Nr[a]	Nt[b]	SwissProt[c]	COG[d]	KEGG[e]	GO[f]	Annotated in all database	Annotated in at least one database	Total unigenes
Z06	40,271	7,860	32,249	21,497	27,595	26,135	3,181	40,271	66,948
A19	39,128	7,670	31,410	21,017	26,865	25,568	3,126	39,135	64,947
R18	39,038	7,670	31,305	20,984	26,860	25,573	3,151	39,046	64,691
R26	39,052	7,646	31,323	21,045	26,874	25,590	3,131	39,093	64,546
C03	38,682	7,607	31,106	20,861	26,644	25,368	3,130	38,730	64,396
Z03	38,696	7,579	31,067	20,854	26,639	25,382	3,110	38,741	64,238
Y11	38,188	7,515	30,601	20,611	26,307	25,064	3,105	38,232	63,397
Y37	36,316	7,187	29,075	19,656	25,054	23,861	2,958	36,357	60,372
X039	39,716	7,778	31,915	21,303	27,332	25,878	3,165	39,757	66,160

[a] NCBI non-redundant protein sequences (Nr).
[b] NCBI non-redundant nucleotide sequences (Nt).
[c] SwissProt protein sequence database (SwissProt).
[d] Clusters of Orthologous Groups of proteins (COG).
[e] Kyoto Encyclopedia of Genes and Genomes (KEGG).
[f] Gene Ontology (GO).

6). SQE and CAS catalyze the two-step cyclization of squalene into cycloartenol, and SMT1 and SMT2 subsequently catalyze the first and second alkylations of cycloartenol (Holmberg et al., 2002; Valitova et al., 2016). Transcriptome analysis found that SQE, CAS, and SMT2 transcripts exhibited various patterns of expression in different *T. grandis* cultivars, while the level of *SMT1* expression was significantly increased, especially in cultivars X039, Z03, and Y37 (Additional file 6). Furthermore, a correlation analysis showed that the expression of *TgSMT1* was strongly correlated with β-sitosterol ($r^2 = 0.8065$; $p < 0.01$) (Fig. 6b).

3.5. RT-qPCR confirmation of RNA-seq data

To validate the accuracy and reliability of the gene expression patterns from RNA-Seq results, 16 candidate genes involved in squalene and β-sitosterol biosynthesis were selected from the DEGs for further RT-qPCR analysis. As expected, the patterns of differential expression for these genes in the nine *T. grandis* cultivars exhibited great similarity to those calculated by the FPKM method from the transcriptome data (Fig. 7a). Furthermore, a Pearson correlation analysis—used to test the correlation between the RT-qPCR and RNA-Seq results—yielded a significant positive correlation (correlation coefficient of 0.7637) between the two methods (Fig. 7b).

3.6. Subcellular localization of GFP-tagged proteins involved in squalene and β-sitosterol biosynthesis

To experimentally determine the subcellular localization of the key proteins in the squalene and β-sitosterol biosynthetic pathways, the ORFs of relative genes were cloned and fused to GFP under the control of the CaMV35S promoter. Finally, TgFPS::GFP was successfully constructed and expressed in the epidermis cells of tobacco leaves. FPS belongs to the short-chain prenyltransferase subfamily of proteins, which are distributed in the cytosol, mitochondria, chloroplast, and peroxisome (Thabet et al., 2011). While TgFPS::GFP in the present study displayed multi-localizations, strong signals were observed in the plasma membrane, while additional fluorescent signals were detected in the nucleus (Fig. 8a–d).

4. Discussion

4.1. Accumulation patterns of squalene and β-sitosterol in T. grandis seeds varies among cultivars

Owing to the valuable biological properties of squalene and β-sitosterol, more and more studies have focused on quantifying and investigating the biosynthesis of these compounds (Nagegowda, 2010). In the present study, the contents of squalene and β-sitosterol in the seed of nine *T. grandis* cultivars varied from 13 to 72 mg/kg and from 1500 to 4100 mg/kg, respectively (Fig. 2a and b). These results were higher than those reported by He et al. (2016) and Shi et al. (2018), where squalene contents ranged from 13 to 38 mg/kg, and β-sitosterol ranged from 900 to 1300 mg/kg and 1200 to 2500 mg/kg, respectively. However, it is important to note that the observed differences in the concentrations of squalene and β-sitosterol were cultivar-dependent, both in the present study as well as in previous reports (He et al., 2016; Shi et al., 2018). As we know, squalene is the necessary intermediate in the β-sitosterol biosynthesis pathway (Liao et al., 2016), and thus the

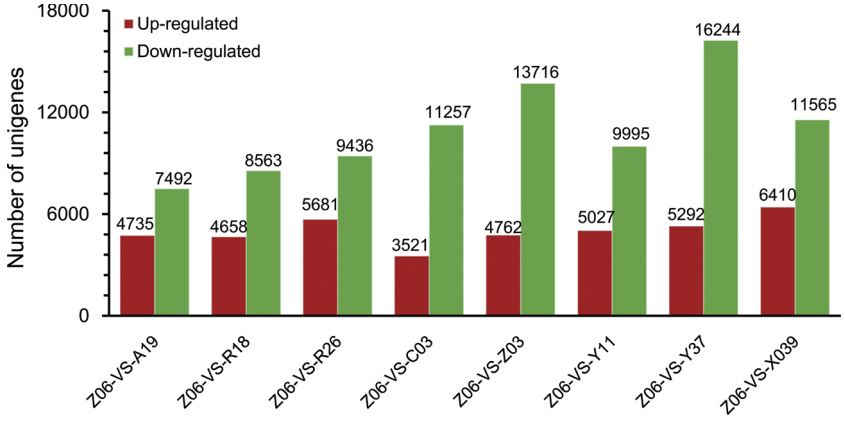

Fig. 3. Analysis of differentially expressed genes (DEGs) in different *Torreya grandis* cultivars.

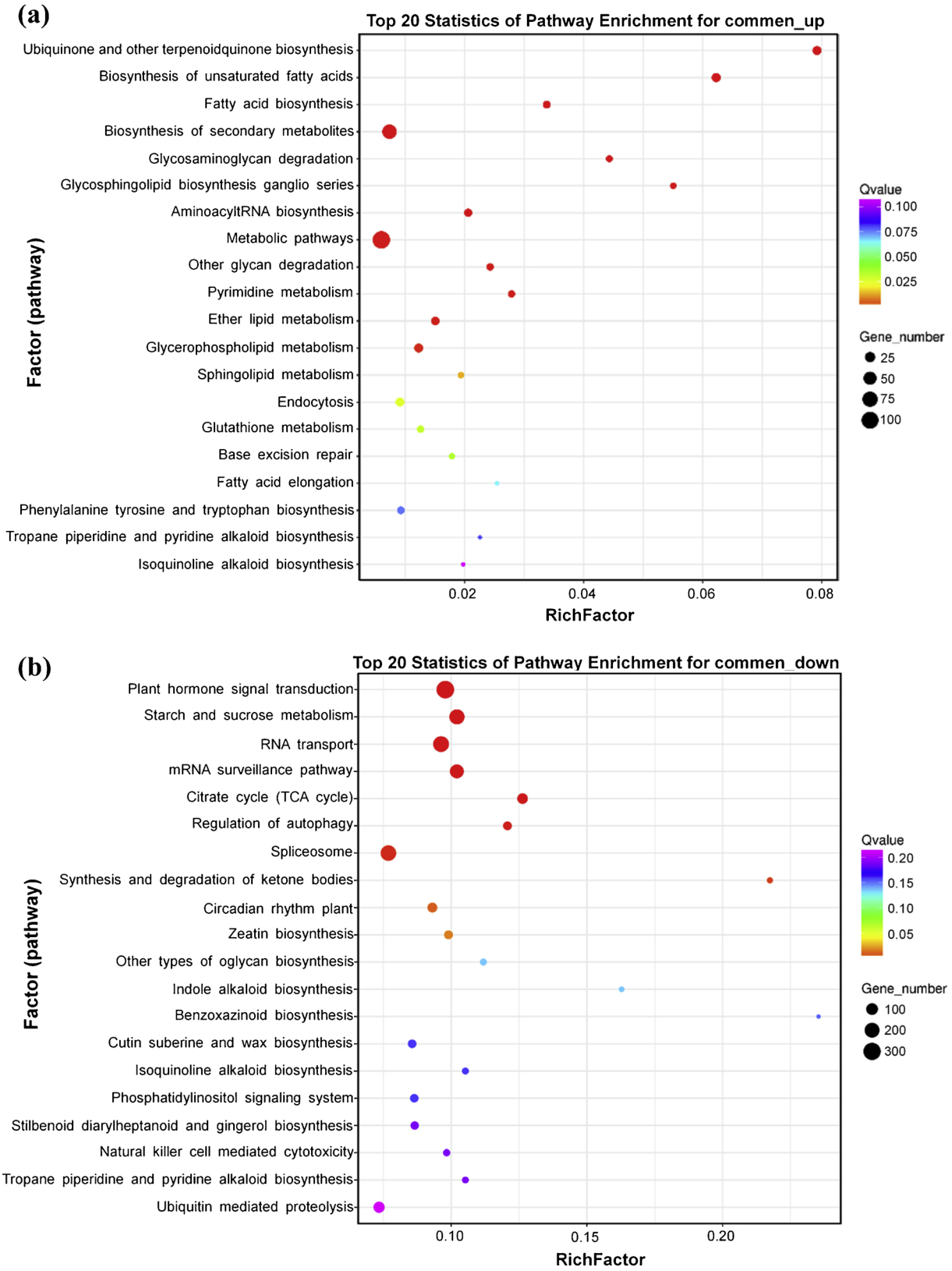

Fig. 4. KEGG enrichment of up-regulated (a) and down-regulated (b) genes among different *Torreya grandis* cultivars.

concentration of squalene may affect the accumulation of β-sitosterol, which was consistent with the observations that squalene and β-sitosterol contents were found significantly correlated, and both of them accumulated relatively high levels in the cultivars Y37 and X039 (Fig. 2a–c). Similarly, the substantial influence of cultivar on squalene and β-sitosterol contents has been reported in many other plant species (Manzi et al., 1998), where squalene content in different cultivars of olive (*Canarium album* L.) oil typically ranges from 200 to 7500 mg/kg (Fernández-Cuesta et al., 2013) and 100 to 170 mg/kg in different cultivars of grape (*Vitis vinifera* L.) seed oils (Wen et al., 2016). The

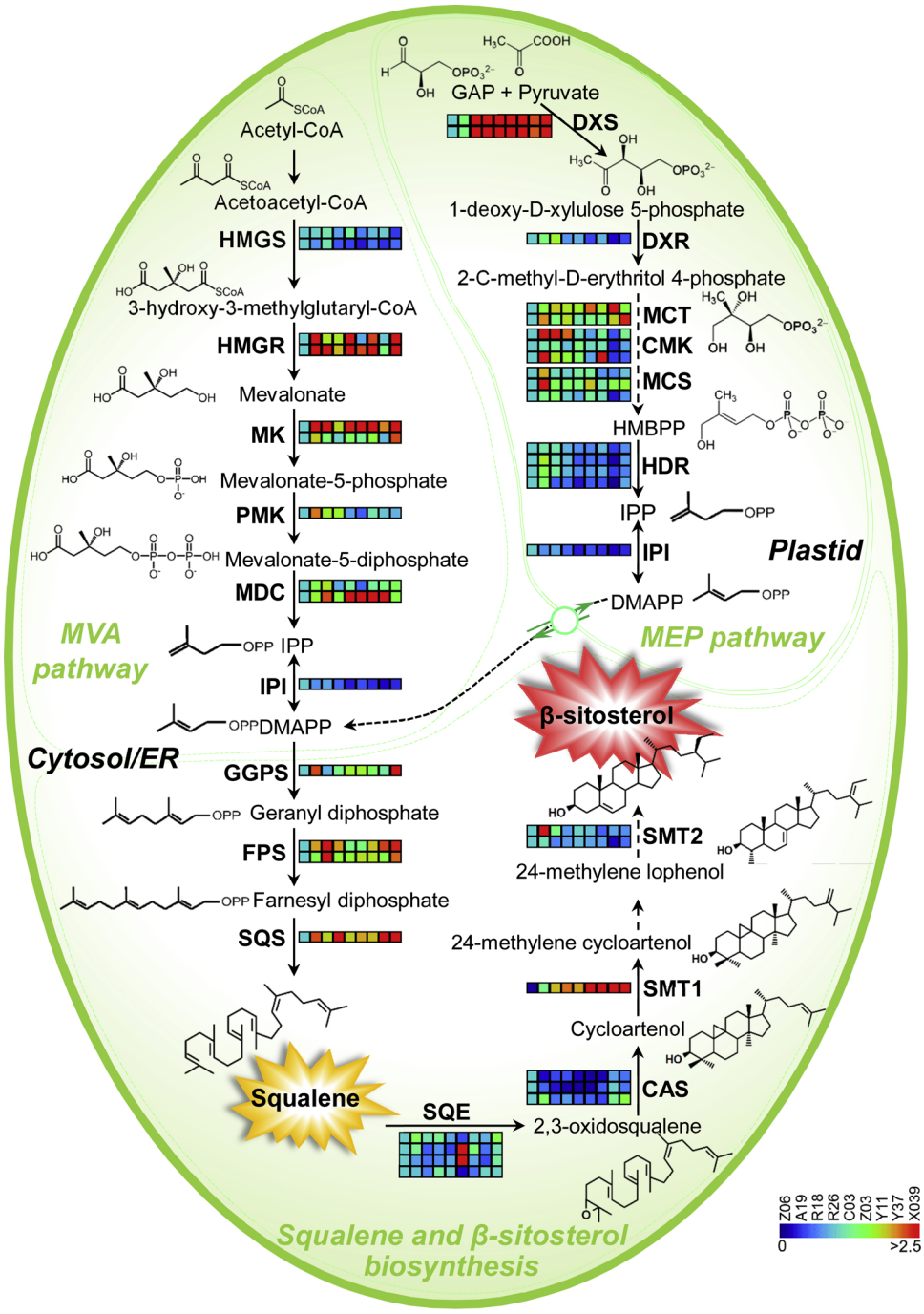

Fig. 5. Schematic presentation of squalene and β-sitosterol biosynthetic pathways in *Torreya grandis* seeds. Enzymes and their relative gene expression patterns are shown in the pathway. The scale bar indicates FPKM ratios of different *T. grandis* cultivars (Z06, A19, R18, R26, C03, Z03, Y11, Y37, and X039), while different colors from blue to red indicate the relative gene expression levels of *T. grandis* cultivars compared with Z06. Abbreviations: CAS, cycloartenol synthase; CMK, 4-(cytidine-5-diphospho)-2-C-methy-D-erythritol kinase; DMAPP, dimethylallyl diphosphate; DXR, 1-deoxy-D-xylulose 5-phosphate reductoisomerase; DXS, 1-deoxy-D-xylulose-5-phosphate synthase; FPS, farnesyl diphosphate synthase; GAP, glyceraldehydes^{-3}-phosphate; GGPS, geranyl diphosphate synthase; HDR, 1-hydroxy-2-methyl-2-(*E*)-butenyl-4-diphosphate reductase; HMBPP, (*E*)-4-hydroxy-3-methylbut-2-enyl diphosphate; HMGR, 3-hydroxy-3-methyl glutaryl coenzyme A reductase; HMGS, 3-hydroxy-3-methyl glutaryl coenzyme A synthase; IPI, isopentenyl diphosphate isomerase; IPP, isopentenyl diphosphate; MCS, 2-C-methyl-D-erythritol-2,4-cyclodiphosphate synthase; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; MDC, mevalonate-5-pyrophosphate decarboxylase; MEP, methylerythritol phosphate; MK, mevalonate kinase; MVA, mevalonate; PMK, phosphomeralonate kinase; SMTs, Δ24-sterol methyl transferases; SQE, squalene epoxidase; SQS, squalene synthase.

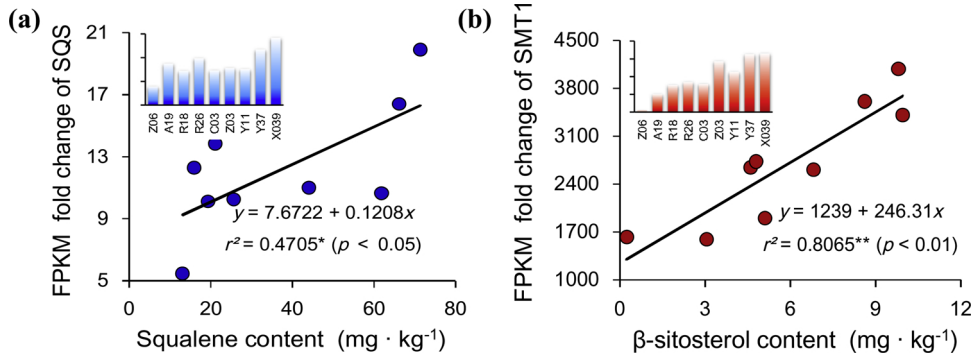

Fig. 6. Correlation analysis between *TgSQS* FPKM fold change and squalene content (a), *TgSMT1* FPKM fold change and β-sitosterol content in different *Torreya grandis* cultivars (b).

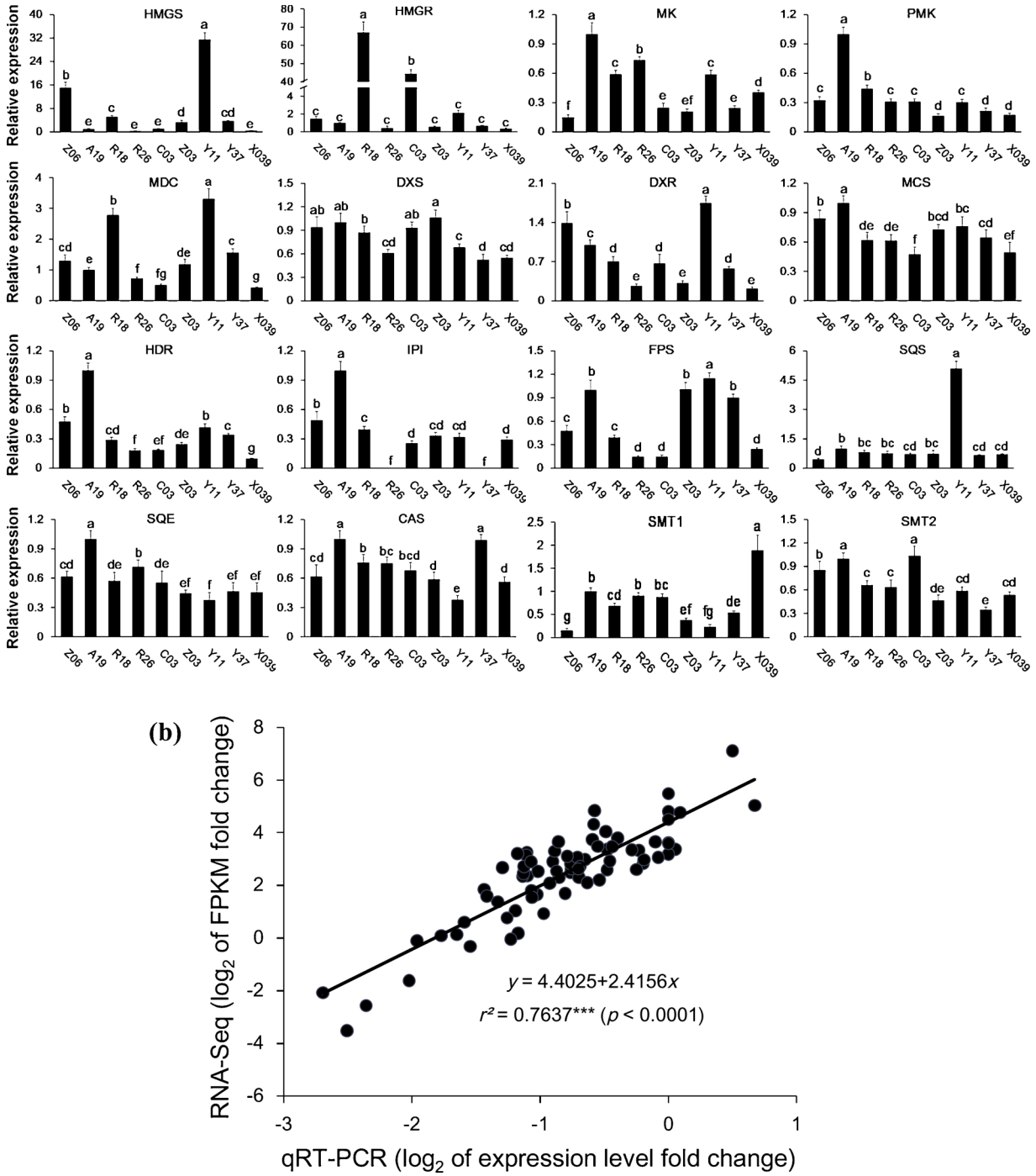

Fig. 7. RT-qPCR validation of RNA-seq relative expression estimation. (a) RT-qPCR analysis of gene expression patterns of 16 selected candidate genes involved in squalene and phytosterol biosynthesis among different *Torreya grandis* cultivars. (b) Correlation of the expression levels of the selected genes measured by RT-qPCR and RNA-seq.

content of β-sitosterol has also been shown to be cultivar-dependent in *Cocos nucifera* (50–380 mg/kg), *Trachycarpus fortunei* oil (60–400 mg/kg), and in *Arachis hypogaea* (120–1200 mg/kg) (Phillips et al., 2002; de Jong et al., 2003), and was generally lower in other plant species than in *T. grandis* cultivars X039, Z03, and Y37. Altogether, these results suggest that these three *T. grandis* cultivars could be a potential natural source of β-sitosterol, and that the biosynthesis of β-sitosterol might be different regulated among *T. grandis* cultivars. In addition, the significant negative correlation between oil content and β-sitosterol content found in the present study was consistent with previous results obtained for apple (*Malus domestica* L.) seeds (Górnaś et al., 2014). These results may provide valuable information that could be used for the preliminary estimates of β-sitosterol content merely based on assessments of oil yields in different plant cultivars. However, further studies should be performed to fully understand and explain this phenomenon.

4.2. Expression pattern of genes involved in MVA and MEP pathways in *T. grandis* cultivars

The *de novo* biosynthesis of squalene and β-sitosterol has been studied in many plants (Nes, 2011). Some studies indicated that

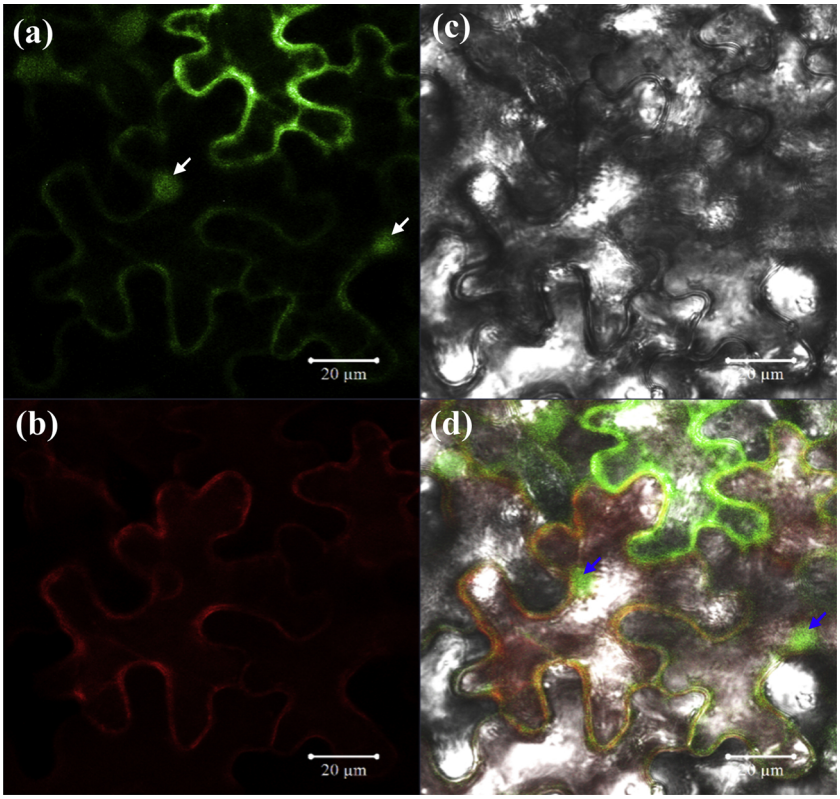

Fig. 8. Subcellular localization of TgFPS. Confocal images of TgFPS::GFP (a) and plasma membrane marker (b), corresponding transmitted light images (c), and their merge images (d). White and blue arrows indicate nuclei, Bar = 20 μm.

phytosterols are mainly synthesized via the MVA pathway rather than the MEP pathway (Rosenwasser et al., 2014), while it also showed that metabolite exchanges or cross-regulation seem to occur between the two biosynthetic pathways in some plants (Laule et al., 2003). In the present study, genes (including *HMGR*, *MK*, and *MDC*) involved in the MVA pathway were upregulated in most *T. grandis* cultivars in comparison with Z06 (Fig. 5; Additional file 6). Among them, HMGR catalyzes the rate limiting step in the MVA pathway (Bach, 1995). The overexpression of rubber (*Hevea brasiliensis* L.) *HMGR* in transgenic tobacco increased phytosterol levels in its seeds (Harker et al., 2003). Thus the elevated expression of *TgHMGR* (BRD_TGR92 and BRD_TGR9470) suggested a positive role for the MVA pathway in the accumulation of IPP and DMAPP in most *T. grandis* cultivars. Furthermore, the higher relative expression of *TgMK* and *TgMDC* in *T. grandis* seeds may also contribute to the accumulation of these initial precursors. This is consistent with the overexpression of *PgMDC* in transgenic lines of *Panax ginseng*, which exhibited 4.4 folds higher phytosterol content than that of the wild-type control (Kim et al., 2014). In addition, six genes involved in the MEP pathway were identified in transcriptome, where the expression of *DXS* and *MCT* was also upregulated in most *T. grandis* cultivars in comparison with Z06 (Fig. 5; Additional file 6). Both of these enzymes have been shown to play important roles in the MEP pathway. The overexpression of the *DXS* gene triggered a remarkable increase in isoprenoid precursors in numerous plants, such as *Lavandula latifolia* (Miñoz-Bertomeu et al., 2006), *Catharanthus roseus* (Peebles et al., 2011), and *Salvia miltiorrhiza* (Kai et al., 2011). Furthermore, the downregulation of *AtMCT* resulted in a dramatic reduction in IPP and DMAPP biosynthesis in transgenic *Arabidopsis* (Okada et al., 2002). In this case, the upregulation of these key genes suggests that both the MVA and MEP pathways may be active in different *T. grandis* cultivars during the early stage of steroid biosynthesis (Fig. 5; Additional file 6). However, the upregulated genes from both pathways did not exhibit significant differences between *T. grandis* cultivars or display consistent trends in the contents of squalene and β-sitosterol, which varied depending on the cultivar (Fig. 2; Additional file 6). These results suggest that it is likely that the upstream isoprenoid biosynthetic pathway does not determine the biosynthetic abilities of squalene and β-sitosterol in *T. grandis* seeds. The C_5 precursors (IPP and DMAPP) not only contribute to the accumulation of squalene and β-sitosterol, but also are the universal terpene precursors for the biosynthesis of a wide range of steroids. Therefore, the MVA and MEP pathways may be important for the biosynthesis of IPP and DMAPP in *T. grandis* cultivars, while the accumulation of squalene and β-sitosterol probably are regulated by related key genes in the downstream pathway.

4.3. The key regulatory steps controlling squalene and β-sitosterol biosynthesis in different T. grandis cultivars

In the downstream steroid biosynthesis pathway, the short-chain isoprenyl diphosphate synthases (IDSs; including GPS, FPS, and GGPS) have been proposed as the first key branch-point enzymes in the catalyzation of IPP and/or DMAPP to form the different precursors (GPP, FPP, and GGPP), thereby controlling the biosynthesis of different classes of steroids (Schmidt and Gershenzon, 2008). Previous studies have suggested that each specific IDS generally corresponds to only one single product (Schmidt and Gershenzon, 2008). However, a novel bifunctional PaIDS1 was found to possess the ability to form 90% GPP and 10% GGPP in the gymnosperm of *Picea abies* (Schmidt et al., 2010). In the present study, the upregulated unigene BRD_TGR72198 was identified with the annotation of putative TgGGPS, while a sequence analysis indicated that it was slightly more similar to PaIDS1 (61.5%) than to GPS (PaIDS2, 55.9%) and GGPS (PaIDS5, 58.5%) (Additional files 6 and 8). These results indicate that the role of TgGGPS may be similar to that of PaIDS1, and may provide not only GGPP, but also a substantial portion of the GPP required for squalene and β-sitosterol

biosynthesis. Furthermore, the membrane-bound squalene synthase (SQS) is another key enzyme in the squalene and sterol biosynthesis pathway (Radisky and Poulter, 2000; Huang et al., 2007). There are two members of SQS-annotated sequences in *Arabidopsis* (AtSQS1, At4g34640 and AtSQS2, At4g34650), while only AtSQS1 was found to be a functional SQS (Busquets et al., 2008). In this study, one SQS (*TgSQS*; BRD_TGR604) was identified. Sequence analysis showed that the deduced amino acid sequence of TgSQS shared high similarities with AtSQS1 and AtSQS2 (Additional file 7). Moreover, the highly conserved residue Phe^{287} (domain C) was also found in TgSQS, as this specific site is considered to be important for the SQS activity of AtSQS1 (Additional file 7) (Busquets et al., 2008). These data indicate the catalytic/functional activity of TgSQS in *T. grandis*. Furthermore, a positive correlation between *TgSQS* transcript levels and squalene content was observed (Fig. 6a). This suggested that the expression patterns of *TgSQS* may be important to regulate the accumulation of squalene in different cultivars, indicating that TgSQS is the key control point in the squalene biosynthesis pathway in *T. grandis*.

In addition, as a necessary intermediate for β-sitosterol biosynthesis, squalene is known to be transformed to cycloartenol by a two-step cyclization (Valitova et al., 2016). Cycloartenol is further converted into β-sitosterol after alkylation events that are catalyzed by two distinct SMTs (Holmberg et al., 2002). SMT1 is a critical enzyme that catalyzes the first committed step in regulating the carbon flux toward sterol biosynthesis (Holmberg et al., 2002). It has been proposed that SMT1 could act as a regulator of seed sterol content, where overexpressing *SMT1* was shown to increase the quantity of β-sitosterol by approximately 30%–50% in tobacco seeds (Holmberg et al., 2002). Consistent with this, transcriptional levels of *TgSMT1* were upregulated in *T. grandis* seeds in the present study, especially in the cultivars X039, Z03, and Y37 (Fig. 5; Additional file 6). The expression of *TgSMT1* was also positively correlated with β-sitosterol content in the seeds of *T. grandis* (Fig. 6b). The fact that the expression of *TgSMT2* was clearly downregulated in most *T. grandis* cultivars in comparison with Z06 (Fig. 5; Additional file 6) suggests that post-transcriptional regulation may be important for the enzymatic activity of TgSMT2. These findings confirm the strong correlation of squalene and β-sitosterol contents with the transcript levels of biosynthetic genes in different cultivars of *T. grandis*. Moreover, several key regulatory steps have been proposed to play important roles during squalene and β-sitosterol biosynthesis in different *T. grandis* cultivars. However, more research is clearly required to further elucidate the molecular mechanisms that regulate squalene and β-sitosterol biosynthesis in different cultivars of *T. grandis*.

5. Conclusion

In the present study, the concentrations of squalene and β-sitosterol were investigated in different cultivars of *T. grandis*. Results noted clear cultivar-specific patterns in the accumulation of squalene and β-sitosterol, which suggests that the biosynthetic regulation of these bioactive compounds might be different among *T. grandis* cultivars. In addition, the identification and analysis of DEGs from the transcriptome analysis of *T. grandis* cultivars indicates roles for both MVA and MEP pathways, suggesting that the biosynthesis of squalene and β-sitosterol occur via both pathways. Moreover, transcriptional profiling also provides new insights into our understanding of the molecular mechanisms and key regulatory steps underlying squalene and β-sitosterol biosynthesis.

Conflict of interest

The authors declare that there are no conflicts of interest.

Acknowledgments

This work was supported by the Special Fund for Forest Scientific Research in the Public Welfare (201504708); the National Natural Science Foundation of China (31670687); the Key Technical Integration and Demonstration Promotion of the Efficient Ecological Cultivation of Torreya Young Forest ([2015]No.TS03), the Selective Breeding of New Cultivars in *Torreya grandis* (2016C02052-12); the Zhejiang Provincial Natural Science Foundation of China (LQ19C160008); and the Launching Funds for Zhejiang A&F University (2018FR021).

Appendix A. Supplementary data

Supplementary material related to this article can be found, in the online version, at doi:https://doi.org/10.1016/j.indcrop.2019.01.035.

References

Ambavade, S.D., Misar, A.V., Ambavade, P.D., 2014. Pharmacological, nutritional, and analytical aspects of β-sitosterol: a review. Orient. Pharm. Exp. Med. 14, 193–211. https://doi.org/10.1007/s13596-014-0151-9.

Anders, S., Huber, W., 2010. Differential expression analysis for sequence count data. Genome Biol. 11, R106. https://doi.org/10.1186/gb-2010-11-10-r106.

Bach, T.J., 1995. Some new aspects of isoprenoid biosynthesis in plants—A review. Lipids 30, 191–202. https://doi.org/10.1007/BF02537822.

Botella-Pavía, P., Besumbes, O., Phillips, M.A., Carretero-Paulet, L., Boronat, A., Rodríguez-Concepción, M., 2004. Regulation of carotenoid biosynthesis in plants: evidence for a key role of hydroxymethylbutenyl diphosphate reductase in controlling the supply of plastidial isoprenoid precursors. Plant J. 40, 188–199. https://doi.org/10.1111/j.1365-313X.2004.02198.x.

Busquets, A., Keim, V., Closa, M., del Arco, A., Boronat, A., Arro, M., Ferrer, A., 2008. *Arabidopsis thaliana* contains a single gene encoding squalene synthase. Plant Mol. Biol. 67, 25–36. https://doi.org/10.1007/s11103-008-9299-3.

Chen, B.Q., Cui, X.Y., Zhao, X., Zhang, Y.H., Piao, H.S., Kim, J.H., Lee, B.C., Pyo, H.B., Yun, Y.P., 2006. Antioxidative and acute antiinflammatory effects of *Torreya grandis*. Fitoterapia 77, 262–267. https://doi.org/10.1016/j.fitote.2006.03.019.

de Jong, A., Plat, J., Mensink, R.P., 2003. Metabolic effect of plant sterols and stanols. J. Nutr. Biochem. 4, 362–369. https://doi.org/10.1016/S0955-2863(03)00002-0.

Delgado-Zamarreno, M.M., Bustamante-Rangel, M., Martinez-Pelarda, D., Carabias-Martinez, R., 2009. Analysis of β-sitosterol in seed and nuts pressurized liquid extraction and liquid chromatography. Anal. Sci. 25, 765–768. https://doi.org/10.2116/analsci.25.765.

Fernández-Cuesta, A., León, L., Velasco, L., De la Rosa, R., 2013. Changes in squalene and sterols associated with olive maturation. Food Res. Int. 54, 1885–1889. https://doi.org/10.1016/j.foodres.2013.07.049.

Fox, C.B., 2009. Squalene emulsions for parenteral vaccine and drug delivery. Molecules 14, 3286–3312. https://doi.org/10.3390/molecules14093286.

Giacometti, J., 2001. Determination of aliphatic alcohols, squalene, alpha-tocopherol and sterols in olive oils: direct method involving gas chromatography of the unsaponifiable fraction following silylation. Analyst 126, 472–475. https://doi.org/10.1039/b007090o.

Górnaś, P., Rudzińska, M., Segliņa, D., 2014. Lipophilic composition of eleven apple seed oils: a promising source of unconventional oil from industry by-products. Ind. Crop. Prod. 60, 86–91. https://doi.org/10.1016/j.indcrop.2014.06.003.

Górnaś, P., Rudzińska, Raczyk, M., Mišina, I., Soliven, A., Segliņa, D., 2016. Composition of bioactive compounds in kernel oils recovered from sour cherry (*Prunus cerasus* L.) by-products: impact of the cultivar on potential applications. J. Crop. Prod. Process. 82, 44–50. https://doi.org/10.1016/j.indcrop.2015.12.010.

Harker, M., Holmberg, N., Clayton, J.C., Gibbard, C.L., Wallace, A.D., Rawlins, S., Hellyer, S.A., Lanot, A., Safford, R., 2003. Enhancement of seed phytosterol levels by expression of an N-terminal truncated *Hevea brasiliensis* (rubber tree) 3-hydroxy-3-methylglutaryl-CoA reductase. Plant Biotechnol. J. 1, 13–121. https://doi.org/10.1046/j.1467-7652.2003.00011.x.

He, Z.Y., Zhu, H.D., Li, W.L., Zeng, M.M., Wu, S.F., Chen, S.W., Qin, F., Chen, J., 2016. Chemical components of cold pressed kernel oils from different *Torreya grandis* cultivars. Food Chem. 209, 196–202. https://doi.org/10.1016/j.foodchem.2016.04.053.

Hemmerlin, A., 2013. Post-translational events and modifications regulating plant enzymes involved in isoprenoid precursor biosynthesis. Plant Sci. 203–204, 41–54. https://doi.org/10.1016/j.plantsci.2012.12.008.

Holmberg, N., Harker, M., Gibbard, C.L., Wallace, A.D., Clayton, J.C., Rawlins, S., Hellyer, A., Safford, R., 2002. Sterol C-24 methyltransferase type 1 controls the flux of carbon into sterol biosynthesis in tobacco seed. Plant Physiol. 130, 303–311. https://doi.org/10.1104/pp.004226.

Huang, Z., Jiang, K., Pi, Y., Hou, R., Liao, Z., Cao, Y., Han, X., Wang, Q., Sun, S., Tang, K., 2007. Molecular cloning and characterization of the yew gene encoding squalene synthase from *Taxus cuspidate*. J. Biochem. Mol. Biol. 40, 625–635. https://doi.org/10.5483/BMBRep.2007.40.5.625.

Kai, G., Xu, H., Zhou, C., Liao, P., Xiao, J., Luo, X., You, L., Zhang, L., 2011. Metabolic engineering tanshinone biosynthetic pathway in *Salvia miltiorrhiza* hairy root cultures. Metab. Eng. 13, 319–327. https://doi.org/10.1016/j.ymben.2011.02.003.

Kanehisa, M., Goto, S., Sato, Y., Furumichi, M., Tanabe, M., 2012. KEGG for integration and interpretation of large-scale molecular data sets. Nucleic Acids Res. 40, 109–114. https://doi.org/10.1093/nar/gkr988.

Kang, N., Tang, Z.X., 1995. Studies on the taxonomy of the genus Torreya. Bull. Bot. Res. 15, 349–362.

Kim, T.D., Han, J.Y., Huh, G.H., Choi, Y.E., 2011. Expression and functional characterization of three squalene synthase genes associated with saponin biosynthesis in *Panax ginseng*. Plant Cell Physiol. 52, 125–137. https://doi.org/10.1093/pcp/pcq179.

Kim, Y.K., Kim, Y.B., Uddin, M.R., Lee, S., Kim, S.U., Park, S.U., 2014. Enhanced triterpene accumulation in *Panax ginseng* hairy roots overexpressing mevalonate-5-pyrophosphate decarboxylase and farnesyl pyrophosphate synthase. ACS Synth. Biol. 3, 773–779. https://doi.org/10.1021/sb400194g.

Laranjeira, S., Amorim-Silva, V., Esteban, A., Arró, M., Ferrer, A., Tavares, R.M., Botella, M.A., Rosado, A., Azevedo, H., 2015. Arabidopsis squalene epoxidase 3 (SQE3) complements SQE1 and is important for embryo development and bulk squalene epoxidase activity. Mol. Plant 8, 1090–1102. https://doi.org/10.1016/j.molp.2015.02.007.

Laule, O., Fürholz, A., Chang, H.S., Zhu, T., Wang, X., Heifetz, P.B., Gruissem, W., Lange, M., 2003. Crosstalk between cytosolic and plastidial pathways of isoprenoid biosynthesis in *Arabidopsis thaliana*. Proc. Natl. Acad. Sci. U. S. A. 100, 6866–6871. https://doi.org/10.1073/pnas.1031755100.

Li, W., Liu, W., Wei, H., He, Q., Chen, J., Zhang, B., Zhu, S., 2014. Species-specific expansion and molecular evolution of the 3-hydroxy-3-methylglutaryl coenzyme A reductase (HMGR) gene family in plants. PLoS One 9, e94172. https://doi.org/10.1371/journal.pone.0094172.

Liao, P., Hemmerlin, A., Bach, T.J., Chye, M.L., 2016. The potential of the mevalonate pathway for enhanced isoprenoid production. Biotechnol. Adv. 34, 697–713. https://doi.org/10.1016/j.biotechadv.2016.03.005. Epub 2016 Mar 16.

Livak, K.J., Schmittgen, T.D., 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. Methods 25, 402–408. https://doi.org/10.1006/meth.2001.1262.

Manzi, P., Panfili, G., Esti, M., Pizzoferrato, L., 1998. Natural antioxidants in the unsaponifiable fraction of virgin olive oils from different cultivars. J. Sci. Food Agric. 77, 115–120. https://doi.org/10.1002/(SICI)1097-0010(199805)77:1<115::AID-JSFA13>3.0.CO;2-N.

Miñoz-Bertomeu, J., Arrillaga, I., Ros, R., Segura, J., 2006. Up-regulation of 1-deoxy-D-xylulose-5-phosphate synthase enhances production of essential oils in transgenic spike lavender. Plant Physiol. 142, 890–900. https://doi.org/10.1104/pp.106.086355.

Mortazavi, A., Williams, B.A., Mccue, K., Schaeffer, L., Wold, B., 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat. Methods 5, 621–628. https://doi.org/10.1038/nmeth.1226.

Nagegowda, D.A., 2010. Plant volatile terpenoid metabolism: biosynthetic genes, transcriptional regulation and subcellular compartmentation. FEBS Lett. 584, 2965–2973. https://doi.org/10.1016/j.febslet.2010.05.045.

Neelakandan, A.K., Nguyen, H.T., Kumar, R., Tran, L.S., Guttikonda, S.K., Quach, T.N., Aldrich, D.L., Nes, W.D., Nguyen, H.T., 2010. Molecular characterization and functional analysis of *Glycine max* sterol methyl transferase 2 genes involved in plant membrane sterol biosynthesis. Plant Mol. Biol. 74, 503–518. https://doi.org/10.1007/s11103-010-9692-6.

Nes, W.D., 2011. Biosynthesis of cholesterol and other sterols. Chem. Rev. 11, 6423–6451. https://doi.org/10.1021/cr200021m.

Ni, L., Shi, W.Y., 2014. Composition and free radical scavenging activity of kernel oil from *Torreya grandis*, *Carya cathayensis* and *Myrica rubra*. Iran. J. Pharm. Res. 13, 221–226.

Ni, Q.X., Gao, Q.X., Yu, W.W., Liu, X.Q., Xu, G.Z., Zhang, Y.Z., 2015. Supercritical carbon dioxide extraction of oils from two *Torreya grandis* varieties seeds and their physicochemical and antioxidant properties. LWT-Food Sci. Technol. 60, 1226–1234. https://doi.org/10.1016/j.lwt.2014.09.007.

Okada, K., Kawaide, H., Kuzuyama, T., Seto, H., Curtis, I.S., Kamiya, Y., 2002. Antisense and chemical suppression of the nonmevalonate pathway affects ent-kaurene biosynthesis in Arabidopsis. Planta 215, 339–344. https://doi.org/10.1007/s00425-002-0762-0.

Pandit, J., Danley, D.E., Schulte, G.K., Mazzalupo, S., Pauly, T.A., Hayward, C.M., Hamanaka, E.S., Thompson, J.F., Harwood, H.J., 2000. Crystal structure of human squalene synthase. A key enzyme in cholesterol biosynthesis. J. Biol. Chem. 275, 30610–30617. https://doi.org/10.1074/jbc.M004132200.

Paramasivan, K., Rajagopal, K., Mutturi, S., 2018. Studies on squalene biosynthesis and the standardization of its extraction methodology from *Saccharomyces cerevisiae*. Appl. Biochem. Biotechnol. 1–17. https://doi.org/10.1007/s12010-018-2845-9.

Peebles, C.A., Sander, G.W., Hughes, E.H., Peacok, R., Shanks, J.V., San, K.Y., 2011. The expression of 1-deoxy-D-xylulose synthase and geraniol-10-hydroxylase or anthranilate synthase increases terpenoid indole alkaloid accumulation in *Catharanthus roseus* hairy roots. Metab. Eng. 13, 234–240. https://doi.org/10.1016/j.ymben.2010.11.005.

Phillips, K.M., Ruggio, D.M., Toivo, J.I., Swank, M.A., Simpkins, A.H., 2002. Free and esterified sterol composition of edible oils and fats. J. Food Compos. Anal. 15, 123–142. https://doi.org/10.1006/jfca.2001.1044.

Qiao, W., Li, C., Mosongo, I., Liang, Q., Liu, M., Wang, X., 2018. Comparative transcriptome analysis identifies putative genes involved in steroid biosynthesis in *Euphorbia tirucalli*. Genes 9, E38. https://doi.org/10.3390/genes9010038.

Radisky, E.S., Poulter, C.D., 2000. Squalene synthase: steady-state, pre-steady-state, and isotope-trapping studies. Biochemistry 39, 1748–1760. https://doi.org/10.1021/bi9915014.

Ramadan, A.M., Azeiz, A.A., Baabad, S., Hassanein, S., Gadalla, N.O., Hassan, S., Algandaby, M., Bakr, S., Khan, T., Abouseadaa, H.H., Ali, H.M., Al-Ghamdi, A., Osman, G., Edris, S., Eissa, H., Bahieldin, A., 2019. Control of β-sitosterol biosynthesis under light and watering in desert plant *Calotropis procera*. Steroids 141, 1–8. https://doi.org/10.1016/j.steroids.2018.11.003.

Rodríguez-Concepción, M., 2006. Early steps in isoprenoid biosynthesis: multilevel regulation of the supply of common precursors in plant cells. Phytochem. Rev. 5, 1–15. https://doi.org/10.1007/s11101-005-3130-4.

Rosenwasser, S., Mausz, M.A., Schatz, D., Sheyn, U., Malitsky, S., Aharoni, A., Weinstock, E., Tzfadia, O., Ben-Dor, S., Feldmesser, E., Pohnert, G., Vardi, A., 2014. Rewiring host lipid metabolism by large viruses determines the fate of *Emiliania huxleyi*, a bloom-forming alga in the ocean. Plant Cell 26, 2689–2707. https://doi.org/10.1105/tpc.114.125641.

Schmidt, A., Gershenzon, J., 2008. Cloning and characterization of two different types of geranyl diphosphate synthases from Norway spruce (*Picea abies*). Phytochemistry 69, 49–57. https://doi.org/10.1016/j.phytochem.2007.06.022.

Schmidt, A., Wächtler, B., Temp, U., Krekling, T., Séguin, A., Gershenzon, J., 2010. A bifunctional geranyl and geranylgeranyl diphosphate synthase is involved in terpene oleoresin formation in *Picea abies*. Plant Physiol. 152, 639–655. https://doi.org/10.1104/pp.109.144691.

Shi, L.K., Mao, J.H., Zheng, L., Zhao, C.W., Jin, Q.Z., Wang, X.G., 2018. Chemical characterization and free radical scavenging capacity of oils obtained from *Torreya grandis* Fort. ex. Lindl. and *Torreya grandis* Fort. var. Merrillii: A comparative study using chemometrics. Ind. Crop. Prod. 115, 250–260. https://doi.org/10.1016/j.indcrop.2018.02.037.

Singh, S., Pal, S., Shanker, K., Chanotiya, C.S., Gupta, M.M., Dwivedi, U.N., Shasany, A.K., 2014. Sterol partitioning by HMGR and DXR for routing intermediates toward withanolide biosynthesis. Physiol. Plant. 152, 617–633. https://doi.org/10.1111/ppl.12213.

Sparkes, I.A., Runions, J., Kearns, A., Hawes, C., 2006. Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nat. Protoc. 1, 2019–2025. https://doi.org/10.1038/nprot.2006.286.

Thabet, I., Guirimand, G., Courdavault, V., Papon, N., Godet, S., Dutilleul, C., Bouzid, S., Giglioli-Guivarc'h, N., Clastre, M., Simkin, A.J., 2011. The subcellular localization of periwinkle farnesyl diphosphate synthase provides insight into the role of peroxisome in isoprenoid biosynthesis. J. Plant Physiol. 168, 2110–2116. https://doi.org/10.1016/j.jplph.2011.06.017.

Valitova, J.N., Sulkarnayeva, A.G., Minibayeva, F.V., 2016. Plant sterols: diversity, biosynthesis, and physiological functions. Biochemistry (Mosc) 81, 1050–1068. https://doi.org/10.1134/S0006297916080046.

Wang, H., Nagegowda, D.A., Rawat, R., Bouvier-Nave, P., Guo, D., Bach, T.J., Chye, M.L., 2012. Overexpression of *Brassica juncea* wild-type and mutant HMG-CoA synthase 1 in Arabidopsis up-regulates genes in sterol biosynthesis and enhances sterol production and stress tolerance. Plant Biotechnol. J. 10, 31–42. https://doi.org/10.1111/j.1467-7652.2011.00631.x.

Weihrauch, J.L., Gardner, J.M., 1978. Sterol content of foods of plant origin. J. Am. Diet. Assoc. 73, 39–44.

Wen, X., Zhu, M., Hu, R., Zhao, J., Chen, Z., Li, J., Ni, Y., 2016. Characterisation of seed oils from different grape cultivars grown in China. J. Food Sci. Technol. 53, 3129–3136. https://doi.org/10.1007/s13197-016-2286-9.

Wu, J., Huang, J., Hong, Y., Zhang, H., Ding, M., Lou, H., Hu, Y., Yu, W., Song, L., 2018. De novo transcriptome sequencing of *Torreya grandis* reveals gene regulation in sciadonic acid biosynthesis pathway. Ind. Crop. Prod. 120, 47–60. https://doi.org/10.1016/j.indcrop.2018.04.041.

Ye, J.C., Chang, W.C., Hsieh, D.J.Y., Hsiao, M.W., 2010. Extraction and analysis of β-sitosterol in herbal medicines. J. Med. Plants Res. 4, 522–527. https://doi.org/10.5897/JMPR10.337.

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, R14. https://doi.org/10.1186/gb-2010-11-2-r14.

Yu, Y.J., Ni, S., Wu, F., Sang, W.G., 2016. Chemical composition and antioxidant activity of essential oil from *Torreya grandis* cv. merrillii Arils. J. Essent. Oil Bear. Plant 19, 1170–1180. https://doi.org/10.1080/0972060X.2014.989183.