Identification, Classification and Differential Expression of Oleosin Genes in Tung Tree (Vernicia fordii)

Heping Cao¹*, Lin Zhang², Xiaofeng Tan², Hongxu Long², Jay M. Shockey¹

¹ U.S. Department of Agriculture, Agricultural Research Service, Southern Regional Research Center, Commodity Utilization Research Unit, New Orleans, Louisiana, United States of America, ² Key Laboratory of Cultivation and Protection for Non-Wood Forest Trees, Ministry of Education, Central South University of Forestry and Technology, Changsha, Hunan Province, People’s Republic of China

Abstract

Triacylglycerols (TAG) are the major molecules of energy storage in eukaryotes. TAG are packed in subcellular structures called oil bodies or lipid droplets. Oleosins (OLE) are the major proteins in plant oil bodies. Multiple isoforms of OLE are present in plants such as tung tree (Vernicia fordii), whose seeds are rich in novel TAG with a wide range of industrial applications. The objectives of this study were to identify OLE genes, classify OLE proteins and analyze OLE gene expression in tung trees. We identified five tung tree OLE genes coding for small hydrophobic proteins. Genome-wide phylogenetic analysis and multiple sequence alignment demonstrated that the five tung OLE genes represented the five OLE subfamilies and all contained the “proline knot” motif (PX5SPX3P) shared among 65 OLE from 19 tree species, including the sequenced genomes of Prunus persica (peach), Populus trichocarpa (poplar), Ricinus communis (castor bean), Theobroma cacao (cacao) and Vitis vinifera (grapevine). Tung OLE1, OLE2 and OLE3 belong to the S type and OLE4 and OLE5 belong to the SM type of Arabidopsis OLE. TaqMan and SYBR Green qPCR methods were used to study the differential expression of OLE genes in tung tree tissues. Expression results demonstrated that 1) All five OLE genes were expressed in developing tung seeds, leaves and flowers; 2) OLE mRNA levels were much higher in seeds than leaves or flowers; 3) OLE1, OLE2 and OLE3 genes were expressed in tung seeds at much higher levels than OLE4 and OLE5 genes; 4) OLE mRNA levels rapidly increased during seed development; and 5) OLE gene expression was well-coordinated with tung oil accumulation in the seeds. These results suggest that OLE genes 1–3 probably play major roles in tung oil accumulation and/or oil body development. Therefore, they might be preferred targets for tung oil engineering in transgenic plants.

Introduction

Tung tree (Vernicia fordii) is an economically important tree with a very limited growing area in the United States [1,2]. Tung seeds contain approximately 50–60% oil (dry weight basis) with about 80 mole % α-eleostearic acid (9cis, 11trans, 13trans octadecatrienoic acid) [3]. Tung oil is readily oxidized because of the three unique conjugated double bonds in eleostearic acid. Dried tung oil is impervious to heat, moisture, dust and many chemical challenges. Tung oil, unlike other drying oils, does not darken with age. These properties of tung oil make it a widely used drying ingredient in paints, varnishes, coatings and finishes [4,5]. Recently, tung oil has been explored as a raw material to produce biodiesel [6–8], polyurethane and wood flour composites [9], thermosetting polymer [10] and repairing agent for self-healing epoxy coatings [11].

Our project focuses on alternative ways of producing tung oil-like fatty acids and other high-value industrial oils by engineering tung oil biosynthetic genes into oilseed crops. Many tung oil biosynthetic genes have been identified in our laboratories, including those coding for diacylglycerol acyltransferases (DGAT) [12,13], delta-12 oleic acid desaturase (FAD2) and delta-12 fatty acid conjugase (FADX) [14], omega-3 fatty acid desaturase (FAD3) [15], α-1-CoA binding proteins [16], cytochrome b5 [17], cytochrome b5 reductase [18], glycerol-3-phosphate acyltransferase (GPAT) [19], plastid-type omega-3 fatty acid desaturase (TnDES2) [20], aquaporin [21] and glutaredoxin [21]. However, selection of target genes for genetic engineering of plant oils is difficult because oil is biosynthesized by at least 10 enzymatic steps and each step is catalyzed by multiple isozymes [22–24]. Furthermore, it has been difficult to study tung oil biosynthesis at the protein level because these enzymes are mostly hydrophobic and membrane-localized proteins [25,26].

Triacylglycerols (TAG) such as tung oil accumulate in discrete subcellular structures called oil bodies in plants, similar to oil droplets in animals. Plant oil bodies mainly consist of TAG surrounded by a monolayer of phospholipids, with the hydrophobic acyl moieties of the phospholipids interacting with TAG and the hydrophilic head groups facing the cytosol [27,28]. In addition, plant oil bodies contain a number of proteins including oleosins (OLE) and caleosins [27,29,30].

OLE are a group of hydrophobic proteins localized on the surfaces of plant oil bodies formed primarily in the seeds and pollen. The precise functions of OLE are unknown. They may...
function to stabilize oil bodies at low water potential and/or regulate the sizes of oil bodies [31]. Recent research suggests that OLE may be bifunctional enzymes with both monoacylglycerol acyltransferase and phospholipase activities regulated by serine/threonine/tyrosine protein kinases [32,33]. The tung genome is known to contain multiple oleosin genes [34,35] but the oleosin protein makeup of tung seed oil bodies is unknown. A number of OLE EST sequences from tung tree have been deposited in the GenBank database [36]. Two tung tree OLE cDNA clones were described in a preliminary report published by a Chinese journal [36]. However, no data is available on the expression of any OLE gene in tung tree [35,36].

The objectives of this study were to identify OLE genes, classify OLE proteins and analyze OLE gene expression in tung trees. We identified five OLE genes in tung tree. We performed genome-wide phylogenetic analysis and multiple sequence alignment and classified the five tung OLE genes based on 65 OLE from 19 tree species including the sequenced genomes of *Prunus persica* (peach) [37], *Populus trichocarpa* (poplar) [38], *Ricinus communis* (castor bean) [39], *Theobroma cacao* (cacao) [40] and *Vitis vinifera* (grapevine) [41]. We also classified tung OLE according to the well-known 17 subfamilies of *Arabidopsis* OLE. Finally, we used TaqMan and SYBR Green qPCR assays to evaluate the relative abundance and tissue distribution of the five OLE mRNA in the seeds, leaves and flowers of tung trees.

**Materials and Methods**

**Plant Materials**

Tung trees were grown in the American Tung Oil Corporation orchard in Lambert, Mississippi. John Corley, the Company officer, granted permission for this field study. Tung fruits were collected weekly for 11 weeks beginning June 23, 2006 (week 1). The developmental stage of week 1 seeds corresponded to approximately 9 weeks after full bloom and 1 month before the initiation of storage oil synthesis. Tung tree seeds were removed from the trees and kernels and immediately frozen in liquid N2 and stored at −80°C. The oil and fatty acid profiles of these tung seeds were reported previously by Cao et al, 2013 [13].

**Identification of Oleosin Genes in Tung Tree**

The tung seed cDNA library used for EST analysis was constructed previously [14] using the TriplEx system (Clontech, Mountain View, CA, USA). Initial gene discovery was derived from random sequencing of a plasmid-based tung seed cDNA library, as described previously [14]. Additional gene discovery was enabled through 454 pyrosequencing of cDNA samples from developing tung seeds, created from reverse transcription of RNA samples extracted using the method of Wan and Wikins, as described previously [12,42]. Normalized and non-normalized 454 samples were prepared as described previously [16].

**Computational Methods**

OLE sequences from other organisms were obtained from database searches using the keyword “oleosin” and BlastP searches [43,44] against the National Center for Biotechnology Information (NCBI)’s non-redundant protein sequence databases (http://blast.ncbi.nlm.nih.gov/Blast.cgi) using tung tree OLE sequences. The properties and amino acid compositions of OLE were analyzed using Vector NTI software (Life Technologies, Carlsbad, CA) [45]. Statistics were performed using Microsoft Excel. Phylogenetic analysis for studying the presumed evolutionary relationships among the OLE proteins was performed using the Vector NTI software based on the neighbor-joining method of Saitou and Nei [46]. Multiple sequence alignment was performed using the ClustalW algorithm [47,48] of the AlignX program of the Vector NTI software. This method is based on algorithms that assign scores to aligned residues and detect sequence similarities. Identical amino acid residues in alignment have higher scores than those not identical and less similar residues.

**RNA Isolation**

Total RNAs from tung seeds, leaves and flowers were isolated as described by Spectrum Plant Total RNA Kit (Sigma) [49] and the hot borate method [42]. RNA concentrations and integrity were determined using RNA 6000 Nano Assay Kit and the Bioanalyzer 2100 (Agilent Technologies) with RNA 6000 Ladder as the standards [50]. The RNAs isolated from tung tissues were high quality because the RNA preparations have high rRNA ratio (28S/18S = 1.9) and the RNA integrity number (RIN = 8.7) [49].

**cDNA Synthesis**

The cDNAs were synthesized from total RNAs using SuperScript II reverse transcriptase as described [49]. The cDNA synthesis mixture in 20 μl contained 5 μg total RNA, 2.4 μg oligo(dT)12-18 primer, 0.1 μg random primers, 500 μM dNTPs, 10 mM DTT, 40 μM RnasOUT, and 200 μ SuperScript II reverse transcriptase in 1X first-strand synthesis buffer (Life Technologies, Carlsbad, CA). The cDNA synthesis reaction was kept at 42°C for 30 min. The cDNAs were stored in −80°C freezer before qPCR analyses.

**qPCR Primers and Probes**

PCR primers and TaqMan probes were designed using Primer Express software (Applied Biosystems, Foster City, CA, USA). The *T*~m~ for the probes were approximately 10°C higher than the corresponding primers. They were synthesized by Biosearch Technologies, Inc (Navato, CA, USA). The amplicon sizes and the nucleotide sequences (5’ to 3’) of the forward primers, TaqMan probes (TET–BHQ1) and reverse primers of OLE are described in Table 1. The reference genes coding for tung 60 s ribosome protein L19 (Rpl19b), ubiquitin protein ligase (Ubl) and glyceraldehyde 3 phosphate dehydrogenase (Gapdh) were described [51]. Tung Rpl19b and Ubl are preferable to Gapdh as the best reference genes for both TaqMan and SYBR Green qPCR assays for quantitative gene expression analysis in tung trees [52].

**qPCR Assays**

The optimized qPCR reaction mixtures contained variable amounts of total RNA-derived cDNA (2.5, 5, 12.5, and 25 ng), 200 nM each of the forward primer, reverse primer, and TaqMan probe and 1× Absolute qPCR Mix (ABgene House, Epsom, Surrey, UK) (TaqMan qPCR) or 1× SYBR Green Supermix (Bio-Rad Laboratories) without the TaqMan probes (SYBR Green qPCR) [50]. The reactions were performed in 96-well clear plates sealed by adhesives with a CFX96 real-time system-C1000 Thermal Cycler (Bio-Rad Laboratories). The thermal cycle conditions for TaqMan assay were as follows: 2 min at 50°C and 15 min at 95°C (This step is required for the activation of Thermo-Start DNA polymerase), followed by 50 cycles at 95°C for 15 s and 60°C for 60 s. The thermal cycle conditions for SYBR Green assay were as follows: 3 min at 95°C, followed by 40 cycles at 95°C for 10 s, 65°C for 30 s and 72°C for 30 s. Agarose gel electrophoresis was used to confirm the specificity of qPCR amplification using 3% agarose gel for separating qPCR products at 100 V for 30 min as shown previously [13].
Data Analysis

The ∆ΔCT method of relative quantification was used to determine the fold change in expression [53]. This was done by first normalizing the threshold cycle (CT) values of the target mRNAs to the CT values of the internal reference mRNA Rpl19b, Gapdh or Ubl in the same samples (first normalizing the threshold cycle (CT) values of the target determine the fold change in expression [53]). This was done by (DDCT was further normalized with a calibrator, the sample control (ΔΔCT). The fold change in expression was then obtained (2−ΔΔCT)). The amplification efficiency of qPCR assay was estimated on the basis of the equation E=(10−1/ΔΔCT−1)×100 [54]. The means and standard deviations presented in the tables and figures were determined from 4–6 assays for each mRNA.

Results

Identification of OLE Genes in Tung Tree

Anonymous tung seed cDNA sequences from week 6 were generated and analyzed by pyrosequencing (‘‘454’’) technology as described previously [16]. Multiple forms of OLE genes were found in these U.S. samples. Four of them corresponded to the sequences identified in the Chinese samples, which were deposited in the Genbank database (Table 1). The other tung oleosin in U.S. samples (named OLE4) shared significant conservation with the common tung OLE at the protein and nucleotide levels; 16 amino acid residues out of 169 and 123 nucleotides out of 820 were completely conserved among the five tung OLE at the protein and DNA levels, respectively (Figure 1 and Figure S1). The signature sequences of OLE called ‘‘proline knot’’ (PX5SPX3P) were completely conserved among the five tung tree OLE (Figure 1). These five forms of OLE were all small proteins with an average of 154 amino acid residues and calculated molecular mass of 16.5 kDa (Table 2). These proteins possessed high isoelectric point (9.35) and high percentage of hydrophobic resides (40.45%) (Table 2).

Classification of Tung Tree OLE

GenBank database search identified 75 unique OLE from 22 trees. These OLE include 1 from Coffea arabica (coffee) [55], 5 from Coffea canephora (coffee) [55], 2 from Cocos nucifera (coconut palm) [57], 5 from Camellia oleifera (tea oil), 5 from Carya illinoinensis (pencil pine) [58], 1 from Citrus sinensis (orange) [59], 1 from Elaeis guineensis (oil palm) [60], 2 from Ficus carica (climbing fig) [61], 5 from Jatropha curcas (barbados nut) [62–64], 1 from Juglans regia (walnut), 1 from Olea europaea (olive) [65], 1 from Persica americana (avocado) [66], 1 from Prunus armeniaca (apricot) [67], 1 from Prunus dulcis (almond) [68], 6 from Prunus persica (peach) [37], 2 from Pinus taeda (loblolly pine) [69], 9 from Populus trichocarpa (poplar) [38], 5 from Ricinus communis (castor bean) [39], 6 from Thelohania cacao (cacao) [40] and 9 from Vitis vinifera (grapevine) [41]. Sixty-five of these sequences from 19 tree species including five identified in this report contained the perfect ‘‘proline knot’’ motif (PX5SPX3P) (Figure S2). Phylogenetic analysis showed that tung OLE were most closely related to castor bean OLE (Figure 2A). The five tung OLE represented the five subfamilies of OLE from diverse tree species (Figure 2A). Tung OLE contained all four invariant residues in the ‘‘proline knot’’ motif (PX5SPX3P) (Figure S2). In addition, the amino and carbonyl termini of tung OLE were properly aligned with those of the other plant OLE (Figure S2). These sequence analyses clearly support our conclusion that the identified OLE from tung tree are full-length.

Arabidopsis contains 17 subfamilies of OLE including 5 forms of seed-specific OLE (S type), 3 forms of seed-and-microspore-specific OLE (SM type) and 9 forms of tapetum-specific OLE (T type) [70–72]. Twenty-three reference genes code for OLE in Arabidopsis genome [73–79]. Tung OLE aligned with S and SM types but none of them aligned with T type of Arabidopsis OLE (Figure 2B). Tung OLE1 closely aligned with Arabidopsis S5 and S5 OLE, whereas tung OLE2 and OLE3 aligned with Arabidopsis S1, S2 and S4 OLE (Figure 2B). Tung OLE4 aligned with Arabidopsis SM1 and SM2 OLE and tung OLE5 aligned with Arabidopsis SM3 OLE (Figure 2B).

Optimization, Specificity and Efficiency of qPCR Assay for OLE Gene Expression

TaqMan and SYBR Green qPCR assays are widely used for quantitative analysis of gene expression. Figure 3A shows an example of primer and probe optimization for OLE1 mRNA quantification. Two hundred nM primer pair and probe concentrations saturated the TaqMan qPCR reactions. We thus used 200 nM for each of the primers and probes in the following qPCR assays. SYBR Green qPCR may generate false positive signals if nonspecific products or primer-dimers are present in the assay because the dye binds to all double-stranded DNA. Melt curve analysis shows that each OLE gene-specific qPCR resulted in a single peak of PCR amplification signal using total cDNA from seeds, leaves and flowers (Figure 3B and Figure S3). Agarose gel electrophoresis shows that amplification from each OLE gene resulted in a single DNA fragment matching the predicted size of the amplicon (Figure 3B and Figure S3). Both analyses indicate that SYBR Green qPCR assays were reliable for evaluating OLE family gene expression. Both qPCR assays generated similar slopes and correlation co-efficiencies but the TaqMan assays generated higher y-intercepts (Figure 3C and Figure S4). The TaqMan qPCR assay was selected for further analysis of qPCR efficiency using RNA from other stages of tung seeds and leaves and flowers. TaqMan qPCR generated high correlation co-efficiency ($r^2>0.99$).
Variations of OLE Gene Expression among Tung Trees

The development of tung trees varies significantly in terms of flowering time, which affects gene expression levels, seed development and oil accumulation in later stages. These variations affect data calculation using multiple tung trees. We therefore examined the variations of OLE gene expression among the tung trees using cDNA from multiple stages of developing tung seeds under optimized primer pair and probe concentrations. TaqMan qPCR shows that OLE1 mRNA levels in the seeds had the least variations among the trees with a mean and standard deviation of 0.90±0.09, whereas those of OLE2, OLE3 and OLE4 mRNA were 0.81±0.24, 0.80±0.35 and 0.75±0.24, respectively (Table 3).

The variation of OLE5 gene expression was similar to that of OLE4 (data not shown). Similar results were obtained by SYBR Green qPCR assays (Table S2). These variations of OLE gene expression among the trees, in agreement with previous data on DGAT family gene expression [13], illustrate the difficulties of averaging data among the three trees and therefore data from tree 1 are presented in the following experiments.

OLE Gene Expression among Tung Tissues

TaqMan and SYBR Green qPCR was used to analyze OLE gene expression among tung tissues using cDNA from tung seeds, leaves and flowers with optimized primer pair and probe concentrations. Seeds had the highest expression levels for all five OLE genes by either qPCR assay (Table 4 and data not shown). TaqMan qPCR shows that OLE1, OLE2 and OLE3 mRNA levels were hundreds of folds higher in the seeds than those in the leaves and flowers, respectively, whereas OLE4 mRNA levels in the seeds were tens of folds higher than those in the leaves and several-fold higher than those in the flowers (Table 4). The mean CT values of the TaqMan qPCR amplification were included in the table to illustrate the existence of OLE mRNA in the leaves and flowers. OLE5 gene expression was similar to that of OLE4 in tung tissues (data not shown). SYBR Green qPCR, which utilizes different detection chemistry, was used to confirm the results from TaqMan qPCR assays. SYBR Green qPCR assays generated similar expression patterns in the three tissues (Table 4). Although the expression levels of OLE genes in tung leaves and flowers were low, melt curve analysis and agarose gel electrophoresis showed that these SYBR Green qPCR generated DNA amplicons matched the predicted sizes (Figure 3B and Figure S3 vs. Table 1).

Table 2. The properties and amino acid composition of the tung Ole proteins.

<table>
<thead>
<tr>
<th>Properties and amino acid composition (% by frequency)</th>
<th>OLE1</th>
<th>OLE2</th>
<th>OLE3</th>
<th>OLE4</th>
<th>OLE5</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Length (amino acid residue)</td>
<td>137</td>
<td>154</td>
<td>169</td>
<td>169</td>
<td>142</td>
<td>154±11</td>
</tr>
<tr>
<td>Molecular weight (Da)</td>
<td>14409</td>
<td>16788</td>
<td>17964</td>
<td>18289</td>
<td>15106</td>
<td>16511±1269</td>
</tr>
<tr>
<td>Isoelectric point (pI)</td>
<td>9.87</td>
<td>9.93</td>
<td>8.25</td>
<td>9.17</td>
<td>9.52</td>
<td>9.35±0.62</td>
</tr>
<tr>
<td>Polar (NCOSTY) (%)</td>
<td>26.28</td>
<td>24.03</td>
<td>28.40</td>
<td>27.22</td>
<td>23.94</td>
<td>25.97±1.96</td>
</tr>
<tr>
<td>Hydrophobic (AILFWV) (%)</td>
<td>42.34</td>
<td>38.96</td>
<td>34.32</td>
<td>44.38</td>
<td>42.25</td>
<td>40.45±3.80</td>
</tr>
</tbody>
</table>

Figure 1. Amino acid sequence alignment of the five tung tree OLE proteins. Multiple sequence alignment was performed using the ClustalW algorithm of the AlignX program of the Vector NTI software. OLE name is on the left of alignment followed by the start of the amino acid sequence of each OLE protein. The numbers at the top of the alignment are the positions of the multiple sequence alignment. The letters at the bottom of the alignment are the consensus amino acid residues. Residues in red on yellow represent those conserved in all five OLE sequences at a given position, whereas those in black on blue represent residues conserved in majority of the sequences at a given position.

doi:10.1371/journal.pone.0088409.g001
Figure 2. Phylogenetic analysis of 65 OLE from 19 tree species and 23 reference OLE from *Arabidopsis*. (A) Phylogenetic analysis of OLE from tung tree and other tree species. (B) Phylogenetic analysis of OLE from tung tree and *Arabidopsis*. Tung OLE are highlighted in red. The names of 17 subfamilies from 23 *Arabidopsis* OLE are highlighted in green. S, seed-specific OLE; SM, seed-microspore-specific OLE; T, tapetum-specific OLE. The abbreviations of the organisms are: Ath, *Arabidopsis thaliana*; Car, *Coffea arabica* (coffee); Cca, *Coffea canephora* (coffee); Cav, *Corylus avellana* (hazelnut); Col, *Camellia oleifera* (tea oil); Csi, *Citrus sinensis* (orange); Egu, *Elaeis guineensis* (oil palm); Fpu, *Ficus pumila* (climbing fig); Jcu, *Jatropha curcas* (barbados nut); Jre, *Juglans regia* (walnut); Oeu, *Olea europaea* (olive); Pam, *Persea American* (avocado); Pdu, *Prunus dulcis* (almond); Ppe, *Prunus persica* (peach); Ptr, *Prunus trichocarpa* (poplar); Vfo, *Vernicia fordii* (tung tree); Vvi, *Vitis vinifera* (grapevine).

doi:10.1371/journal.pone.0088409.g002
Figure 3. qPCR optimization, specificity and efficiency for OLE assay. (A) TaqMan qPCR optimization. TaqMan qPCR reactions contained 5 ng RNA-equivalent cDNA from tung seeds, various concentrations of the primers and TaqMan probe. Ole1 assay optimization is presented. (B) Specificity of SYBR Green qPCR by melt curve analysis and gel electrophoresis of amplification products. The qPCR reactions contained 5 ng RNA-equivalent cDNA from tung tree seeds. The qPCR products were separated by agarose gel electrophoresis. Lane 100 bp represents DNA ladders with 100 bp as the smallest band, increasing upward in 100 bp increments. The results using RNA isolated from leaves and flowers are presented in Figure S3. (C) qPCR efficiency for OLE assay. TaqMan and SYBR Green qPCR reaction mixtures contained variable concentrations of RNA-equivalent cDNA from tung seeds, the optimized concentrations of each primer and probe (200 nM), and Absolute QPCR Mix (TaqMan qPCR) or each primer and 1 x iQ SYBR Green Supermix (SYBR Green qPCR). The results using RNA isolated from stage 4 seeds of tree 1 are shown in the figure. The results for Ole2, Ole3 and Ole4 assays are presented in Figure S4. The results using RNA from other stages of tung seeds, leaves and flowers are presented in Table S1. doi:10.1371/journal.pone.0088409.g003
Table 3. Variation of Ole gene expression among tung trees.

<table>
<thead>
<tr>
<th>mRNA</th>
<th>Tree 1 (fold)</th>
<th>Tree 2 (fold)</th>
<th>Tree 3 (fold)</th>
<th>Mean ± SD (fold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ole1</td>
<td>1</td>
<td>0.89±0.27</td>
<td>0.81±0.40</td>
<td>0.90±0.09</td>
</tr>
<tr>
<td>Ole2</td>
<td>1</td>
<td>0.89±0.27</td>
<td>0.55±0.27</td>
<td>0.81±0.24</td>
</tr>
<tr>
<td>Ole3</td>
<td>1</td>
<td>1.01±0.30</td>
<td>0.40±0.20</td>
<td>0.80±0.35</td>
</tr>
<tr>
<td>Ole4</td>
<td>1</td>
<td>0.72±0.22</td>
<td>0.53±0.26</td>
<td>0.75±0.24</td>
</tr>
</tbody>
</table>

TaqMan qPCR reaction mixtures (25 μL) contained 25 ng of RNA-equivalent cDNA from various stages of tung seeds, the optimized concentrations of each primer and probe (200 nM) and QPCR Mix. The expression levels under each tree represent the means and standard deviations of the expression fold calculated using three reference mRNA (Rpl19b, Gapdh and Ubl) from 11 stages of seeds with 2–4 assays for each stage. Ole gene expression in tree 1 seeds was used as the calibrator for the calculation of Ole gene expression in tree 2 and tree 3 seeds.

doi:10.1371/journal.pone.0088409.t003

Table 4. Ole gene expression among tung tissues.

<table>
<thead>
<tr>
<th>qPCR method</th>
<th>mRNA</th>
<th>Seed</th>
<th>Leaf</th>
<th>Flower</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TaqMan</td>
<td>Ole1</td>
<td>1</td>
<td>0.0002±0.0001 (33.58)</td>
<td>0.0003±0.0001 (36.18)</td>
</tr>
<tr>
<td></td>
<td>Ole2</td>
<td>1</td>
<td>0.0006±0.0004 (32.52)</td>
<td>0.0003±0.0001 (33.43)</td>
</tr>
<tr>
<td></td>
<td>Ole3</td>
<td>1</td>
<td>0.0007±0.0004 (32.28)</td>
<td>0.0056±0.0023 (28.87)</td>
</tr>
<tr>
<td></td>
<td>Ole4</td>
<td>1</td>
<td>0.0206±0.0116 (32.62)</td>
<td>0.1247±0.0514 (29.67)</td>
</tr>
<tr>
<td>SYBR Green</td>
<td>Ole1</td>
<td>1</td>
<td>0.0004±0.0002 (28.14)</td>
<td>0.0002±0.0001 (29.74)</td>
</tr>
<tr>
<td></td>
<td>Ole2</td>
<td>1</td>
<td>0.0010±0.0005 (27.34)</td>
<td>0.0005±0.0002 (28.55)</td>
</tr>
<tr>
<td></td>
<td>Ole3</td>
<td>1</td>
<td>0.0007±0.0003 (28.36)</td>
<td>0.0061±0.0031 (25.39)</td>
</tr>
<tr>
<td></td>
<td>Ole4</td>
<td>1</td>
<td>0.0473±0.0221 (25.84)</td>
<td>0.4323±0.2201 (22.73)</td>
</tr>
</tbody>
</table>

The qPCR reaction mixtures contained 5 ng of RNA-equivalent cDNA from various stages of tung tree 1 seeds, leaves and flowers, the optimized concentrations of each primer and probe (200 nM) and QPCR Mix. The expression levels under "seed" represent the means of the expression fold of 4 stages of seeds (weeks 2, 4, 6 and 10) (TaqMan qPCR) or 6 stages of seeds (weeks 2, 4, 5, 6, 8 and 10) (SYBR Green qPCR) calculated using three reference mRNA (Rpl19b, Gapdh and Ubl) with 2–4 assays for each stage. The expression levels under "leaf" and "flower" represent the means and standard deviations of the expression fold calculated using three reference mRNA (Rpl19b, Gapdh and Ubl) with 2–4 assays and Ole gene expression in seeds as the calibrator. The mean Ct values generated using RNA from leaves and flowers are included after expression fold as a proof of low but reliable levels of Ole mRNA detection in these tissues.

doi:10.1371/journal.pone.0088409.t004

SYBR Green qPCR Analysis of OLE Gene Expression in Developing Tung Seeds

TaqMan qPCR assays were performed under optimal concentrations of primers and probe for evaluating the expression profiles of the five OLE genes in developing tung seeds using reference genes Rpl19b, Gapdh and Ubl. Figure 4A shows that the expression of OLE1 and OLE2 genes dramatically increased following the initiation of seed development, which were detectable at week 4 seeds and increased significantly at week 5 seeds. OLE3 mRNA levels were detected in week 5 seeds with lower levels than those of OLE1 and OLE2 mRNA (Figure 4A). While OLE4 mRNA levels were relatively low, OLE5 gene expression in tung seeds was minimal (Figure 4A). Similar profiles of OLE gene expression in tung seeds were observed using the three reference genes Rpl19b (Figure 4A), Gapdh (Figure S5A) and Ubl (Figure S5B).

Among the individual OLE genes, OLE1 mRNA levels were increased 11-, 40- and 97-fold in stage 4, 6 and 10 seeds compared to those in stage 2 seeds (Table S3). Similar increases of OLE2 mRNA levels were observed in the seeds. OLE3 mRNA levels were upregulated even more dramatically, with 7-, 41- and 413-fold increases in stage 4, 6 and 10 seeds compared to stage 2 seeds (Table S3). In contrast, OLE4 mRNA levels were only increased 1.2-, 3.1- and 6.7-fold in stage 4, 6 and 10 seeds compared to stage 2 seeds (Table S3). Interestingly, OLE5 mRNA levels were increased 16- and 18-fold in stage 4 and 6 seeds but declined to 0.23-fold in stage 10 seeds compared to stage 2 seeds (Table S3).

Discussion

Tung tree is a tropical plant with a very limited growing area in the United States [1,2]. Tung orchards in the southern United States were largely destroyed by hurricanes including Hurricanes Betsy in 1965, Camille in 1969, and Katrina and Rita in 2005. The losses of tung orchards due to hurricanes spurred interest in trying to preserve a reliable domestic source of tung oil in the US by engineering tung oil biosynthetic pathway in traditional, temperate oilseeds. To this end, multiple isoforms of many gene families from the tung oil biosynthetic pathway have been cloned by engineering tung oil biosynthetic pathway in traditional, temperate oilseeds. To this end, multiple isoforms of many gene families from the tung oil biosynthetic pathway have been cloned in our laboratories in recent years.

In this study, we identified five tung OLE genes by EST sequencing and pyrosequencing (“454”) technology using anonymous tung cDNA sequences from week 6 seeds. All five OLE shared significant conservation at the protein and nucleotide
The five OLE genes coded for small proteins with an average of 154 amino acid residues and an average of 16.5 kDa. They possessed high isoelectric point and high percentage of hydrophobic residues. These properties of OLE proteins were similar to those of the three DGAT in tung tree [13] and DGAT in other plants, animals and fungi [24,80]. OLE1 and OLE2 were identical to those isolated from cDNA library of tung seeds (OleI-GU245884 and OleII-GU245885) and EST (vf3-GR217899 and vf2-GR217906) [35,36]. OLE3 and OLE5 are identical to those isolated from cDNA library of tung seed EST (vf1-GR217754 and vf5-GR218198) [35]. The reported vf4 has an N-terminal extension and only six varied amino acid residues at the C-terminus of the protein compared to vf3 and clustered together in the phylogenetic tree [35], suggesting that these two genes constitute a paralogous pair, a common occurrence in eudicot species such as Arabidopsis [71]. It might be preferable to name the reported vf3 and vf4 as OLE3a and OLE3b, respectively. Therefore, we named the new OLE as OLE4 in this report since only five major groups of OLE were present in tree species.

The numbers of OLE genes in plants are widely different. We performed a genome-wide search for OLE genes in the sequenced tree genomes including Prunus persica (peach) [37], Populus trichocarpa (poplar) [38], Ricinus communis (castor bean) [39], Theobroma cacao (cacao) [40] and Vitis vinifera (grapevine) [41]. The completed genome sequences from the five trees provide a great opportunity for classifying tree OLE proteins. Phylogenetic analysis of all OLE clearly grouped these proteins into five subfamilies. Each of the five tung OLE genes grouped with each of the five OLE subfamilies. These results suggest that tung tree might only contain five subfamilies of OLE and this report might represent the complete OLE gene family in tung tree. Interestingly, the five tung OLE only aligned with the S and SM type but not the T type of Arabidopsis OLE. Tung OLE1, OLE2 and OLE3 belong to the S type and OLE4 and OLE5 belong to the SM type of Arabidopsis OLE. The more abundant expression of OLE1, OLE2 and OLE3 than OLE4 and OLE5 in tung tree seed support the classification of tung OLE based on Arabidopsis OLE.

It is important to determine which isoform is expressed in developing seeds to understand the genetic control of oil biosynthesis and to guide rational design of successful transgenic plants. In this report, we analyzed expression of OLE gene family quantitatively in 3 tung trees, 3 tung tissues (seeds, leaves and flowers).
flowers) and 11 stages of seed development. Under optimized assay conditions, qPCR exhibited similar amplification efficiencies between the OLE genes (Figure 4C, Figure S4 and Table S1) and the three reference genes reported previously [51], which allowed for comparison of the relative expression levels in different tissues and developmental stages. This study provides complete expression profiles of the OLE genes in tung tissues. Both qPCR methods show that OLE genes were mainly expressed in developing tung seeds and that low levels of OLE mRNA were detected in tung leaves and flowers. OLE1, OLE2 and OLE3 were the major OLE mRNAs in the seeds whose mRNA levels were rapidly increased in developing seeds when tung oil begins to accumulate and their expression levels are sustained in all subsequent stages [13]. These developmental expression profiles of OLE genes in tung tree seed trees are very similar to those in the seeds of Coffea canephora [55] and Brassica napus [81].

It is noteworthy that OLE transcripts were detected in leaves and flowers of tung tree, although their mRNA levels were extremely low compared to those in the seeds. The expression of OLE genes in these non-seed tissues was verified by melt curve and gel electrophoresis analyses of the qPCR products with predicted sizes of the amplicons. It was generally accepted and gel electrophoresis analyses of the qPCR products with OLE genes in these non-seed tissues was verified by melt curve extremely low compared to those in the seeds. The expression of and flowers of tung tree, although their mRNA levels were extremely low compared to those in the seeds, as tobacco leaf [88] and olive fruit [89]. These results demonstrate oil body biogenesis is present in tissues other than seeds such as tobacco leaf [88] and olive fruit [89]. These results demonstrate that OLE genes are expressed in tissues other than the seeds, pollen and tapetum, and suggest that they may play a role in these tissues.

Conclusions and Future Research

We identified five OLE genes in tung tree. Genome-wide phylogenetic analysis and multiple sequence alignment demonstrated that the five tung OLE represented the five OLE subfamilies and contained the “proline knot” motif (PX8PSPX3P) shared among 65 OLE from 19 tree species. Tung OLE1, OLE2 and OLE3 belong to the S type and OLE4 and OLE5 belong to the SM type of Arabidopsis OLE. Expression results demonstrated that all five OLE genes were expressed in developing tung seeds, leaves and flowers but their expression levels were much higher in the seeds than leaves or flowers. In addition, OLE1, OLE2 and OLE3 genes were expressed in tung seeds at much higher levels than OLE4 and OLE5 genes. Finally, the amounts of OLE mRNAs were rapidly increased in developing seeds and their expression levels were well-coordinated with tung oil accumulation. The information on OLE expression profiles suggests that OLE1, OLE2 and OLE3 genes may play major roles in tung oil biosynthesis and/or tung oil body development. Therefore, they might be preferred targets for tung oil engineering in transgenic plants.

There is no information about the genes coding for other oil-body proteins including caleosins and steroleosins in tung tree at this time. Identification and characterization of additional oil body proteins and their corresponding genes at the genomics and proteomics levels would enhance the understanding of genetic and mechanistic control of tung oil biosynthesis. This information is essential for improving important oils such as tung oil in the future.

Supporting Information

Figure S1 Nucleotide sequence alignment of the five OLE genes in tung tree. Multiple sequence alignment was performed using the ClustalW algorithm of the AlignX program of the Vector NTI software. Ole sequence name is on the left of alignment followed by the genusBank accession number and the start of the nucleotide sequence of each Ole gene. The numbers at the top of the alignment are the positions of the multiple sequence alignment. The letters at the bottom of the alignment are the consensus nucleotides. Nucleotides in red on yellow represent those conserved in all five Ole sequences at a given position, whereas those in black on blue represent nucleotides conserved in majority of the sequences at a given position. The underlined nucleotides represent the forward primers, TaqMan probes and the complementary sequences of the reverse primers used in qPCR assays.

Figure S2 Multiple sequence alignment for the identification of amino acid residues and sequence motifs conserved in OLE. Multiple sequence alignment was performed using the ClustalW algorithm of the AlignX program of the Vector NTI software. Each OLE sequence name is on the left of the alignment followed by the position of amino acid residue of OLE protein sequence in the alignment. The letters at the bottom of the alignment are the consensus residues. Color codes for amino acid residues are as follows: 1) red on yellow: consensus residue derived from a completely conserved residue at a given position; 2) black on green: consensus residues derived from the occurrence of greater than 50% of a single residue at a given position; 3) blue on cyan: consensus residue derived from a block of similar residues at a given position; 4) green on white: residue weakly similar to consensus residue at a given position; 5) black on white: non-similar residues. The abbreviations of the organisms are: Car, Coffea arabica (coffee); Cca, Coffea canephora (coffee); Cav, Corylus avellana (hazelnut); Col, Canellia oleifera (tea oil); Citrus sinensis (orange); Egu, Elaeis guineensis (oil palm); Fpu, Ficus pumila (climbing fig); Jcu, Jatropha curcas (barbados nut); Jre, Juglans regia (walnut); Oeu, Olea europaea (olive); Pam, Persea americana (avocado); Pcl, Prunus dulcis (almond); Ppe, Prunus persica (peach); Pta, Pinus taeda (lobolly pine); Ptr, Populus trichocarpa (poplar); Rco, Ricinus communis (castor bean); Tca, Theobroma cacao (cacao); Vfo, Vitis vinifera (grapevine).

Figure S3 Specificity of SYBR Green qPCR Assay. The qPCR reactions contained 5 ng RNA-equivalent cDNA from tung tree leaves and flowers. The qPCR products were separated by agarose gel electrophoresis. Lane 100 bp represents DNA ladders with 100 bp as the smallest band, increasing upward in 100 bp increments. The results using tung tree seeds are shown in Figure 3B. (A) Melt curve analysis, (B) Gel electrophoresis.

Figure S4 qPCR efficiency for OLE assay. TaqMan and SYBR Green qPCR reaction mixtures contained variable concentrations of RNA-equivalent cDNA from tung tree leaves and flowers. The qPCR products were separated by agarose gel electrophoresis. Lane 100 bp represents DNA ladders with 100 bp as the smallest band, increasing upward in 100 bp increments. The results using tung tree seeds are shown in Figure 3B. (A) Melt curve analysis, (B) Gel electrophoresis.
SYBR Green Supersmix (SYBR Green qPCR). The results using RNA isolated from stage 4 seed of tree 1 are shown. The qPCR efficiency for Ole1 mRNA detection is presented in Figure 3C. The results using RNA from other stages of tung seeds, leaves and flowers are presented in Tables S1, A) qPCR efficiency for Ole1 mRNA detection, B) qPCR efficiency for Ole2 mRNA detection, C) qPCR efficiency for Ole3 mRNA detection, D) qPCR efficiency for Ole4 mRNA detection.

Figure S5 Relative levels of OLE gene expression in developing tung seeds by TaqMan qPCR. The qPCR reaction mixtures contained 25 ng of RNA-equivalent cDNA from various stages of tung seed, leaves and flowers and 200 nM of each primer. The means of mRNA expression levels calculated from two qPCR assays in each seed stage are presented. (A) Gapdh as the reference mRNA, (B) Ubl as the reference mRNA.

Figure S6 Relative levels of OLE gene expression in tung tissues by SYBR Green qPCR. The qPCR reaction mixtures contained 5 ng of RNA-equivalent cDNA from various stages of tung seed, leaves and flowers and 200 nM of each primer. The means of mRNA expression levels calculated from two qPCR assays in each stage using Rp119b as the reference mRNA are presented in Figure 4B. The means of mRNA expression levels calculated from two qPCR assays in each stage are presented.

Table S1 TaqMan qPCR efficiency for quantifying Ole mRNA in tung tree tissues.

Table S2 Variation of Ole gene expression among tung trees.

Table S3 Ole gene expression among different stages of tung seeds.

Acknowledgments

The authors thank Drs. PingKuang Chang, Zhongyi He, Steve Bové and Kanniah Rajasekaran for helpful comments on the manuscript. Mention of trade names or commercial products in this publication is solely for the purpose of providing specific information and does not imply recommendation or endorsement by the U.S. Department of Agriculture. USDA is an equal opportunity provider and employer.

Authors Contributions

Conceived and designed the experiments: HC. Performed the experiments: HC. Analyzed the data: HC. Contributed reagents/materials/analysis tools: HC LZ XT HL JMS. Wrote the paper: HC. Revised the manuscript: HC LZ XT HJ. MJS.

References


**Figure S1. Nucleotide sequence alignment of the five Ole genes in tung tree.** Multiple sequence alignment was performed using the ClustalW algorithm of the AlignX program of the Vector NTI software. Ole sequence name is on the left of alignment followed by the GenBank accession number and the start of the nucleotide sequence of each Ole gene. The numbers at the top of the alignment are the positions of the multiple sequence alignment. The letters at the bottom of the alignment are the consensus nucleotides. Nucleotides in red on yellow represent those conserved in all five Ole sequences at a given position, whereas those in black on blue represent nucleotides conserved in majority of the sequences at a given position. The underlined nucleotides represent the forward primers, TaqMan probes and the complementary sequences of the reverse primers used in qPCR assays.
Figure S2. Multiple sequence alignment for the identification of amino acid residues and sequence motifs conserved in OLE. Multiple sequence alignment was performed using the ClustalW algorithm of the AlignX program of the Vector NTI software. Each OLE sequence name is on the left of the alignment followed by the position of amino acid residue of OLE protein sequence in the alignment. The letters at the bottom of the alignment are the consensus residues. Color codes for amino acid residues are as follows: 1) red on yellow: consensus residue derived from a completely conserved residue at a given position; 2) black on green: consensus residue derived from the occurrence of greater than 50% of a single residue at a given position; 3) blue on cyan: consensus residue derived from a block of similar residues at a given position; 4) green on white: residue weakly similar to consensus residue at a given position; 5) black on white: non-similar residues. The abbreviations of the organisms are: Car, Coffea arabica (coffee); Cca, Coffea canephora (coffee); Cav, Corylus avellana (hazelnut); Col, Camellia oleifera (tea oil); Citrus sinensis (orange); Egu, Elaeis guineensis (oil palm); Fpu, Ficus pumila (climbing fig); Jcu, Jatropha curcas (barbados nut); Jre, Juglans regia (walnut); Oeu, Olea europaea (olive); Pam, Persea americana (avocado); Pdu, Prunus dulcis (almond); Ppe, Prunus persica (peach); Pta, Pinus taeda (lobolly pine); Ptr, Populus trichocarpa (poplar); Rco, Ricinus communis (castor bean); Tca, Theobroma cacao (cacao); Vfo, Vernicia fordii (tung tree); Vvi, Vitis vinifera (grapevine).
Figure S3. Specificity of SYBR Green qPCR Assay. The qPCR reactions contained 5 ng RNA-equivalent cDNA from tung tree leaves and flowers. The qPCR products were separated by agarose gel electrophoresis. Lane 100bp represents DNA ladders with 100 bp as the smallest band, increasing upward in 100 bp increments. The results using tung tree seeds are shown in Figure 4B. (A) Melt curve analysis, (B) Gel electrophoresis.
Figure S4. qPCR efficiency for OLE assay. TaqMan and SYBR Green qPCR reaction mixtures contained variable concentrations of RNA-equivalent cDNA from tung seeds, the optimized concentrations of each primer and probe (200 nM), and Absolute QPCR Mix (TaqMan qPCR) or each primer and 1 x iQ SYBR Green Supermix (SYBR Green qPCR). The results using RNA isolated from stage 4 seed of tree 1 are shown. The qPCR efficiency for Ole1 mRNA detection are presented in Figure 4C. The results using RNA from other stages of tung seeds, leaves and flowers are presented in Table S1. (A) qPCR efficiency for Ole2 mRNA detection, (B) qPCR efficiency for Ole3 mRNA detection, (C) qPCR efficiency for Ole4 mRNA detection.
Figure S5. Relative levels of OLE gene expression in developing tung seeds by TaqMan qPCR. The qPCR reaction mixtures contained 25 ng of RNA-equivalent cDNA from tung seeds and 200 nM of each primer and probe. The means of mRNA expression levels calculated from two qPCR assays in each seed stage using Rp19b as the reference mRNA are presented in Figure 5A. The means of mRNA expression levels calculated from two qPCR assays in each seed stage are presented. (A) Gapdh as the reference mRNA, (B) Ubl as the reference mRNA.
Figure S6. Relative levels of OLE gene expression in tung tissues by SYBR Green qPCR. The qPCR reaction mixtures contained 5 ng of RNA-equivalent cDNA from various stages of tung seed, leaves and flowers and 200 nM of each primer. The means of mRNA expression levels calculated from two qPCR assays in each seed stage using Rpl19b as the reference mRNA are presented in Figure 5A. The means of mRNA expression levels calculated from two qPCR assays in each seed stage are presented. (A) Gapdh as the reference mRNA, (B) Ubl as the reference mRNA.
Table S1. TaqMan qPCR efficiency for quantifying Ole mRNA in tung tree tissues.

<table>
<thead>
<tr>
<th>Tung tissue</th>
<th>mRNA</th>
<th>Slope</th>
<th>Y-Intercept</th>
<th>Correlation Co-efficiency</th>
<th>Efficiency $E=\left[10(1/-S)-1\right]\times100$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Seed (2 weeks)</td>
<td>Ole1</td>
<td>-3.3962</td>
<td>26.861</td>
<td>0.9998</td>
<td>97</td>
</tr>
<tr>
<td></td>
<td>Ole2</td>
<td>-3.6044</td>
<td>27.547</td>
<td>0.9986</td>
<td>89</td>
</tr>
<tr>
<td></td>
<td>Ole3</td>
<td>-3.8020</td>
<td>27.912</td>
<td>0.9978</td>
<td>83</td>
</tr>
<tr>
<td></td>
<td>Ole4</td>
<td>-3.4817</td>
<td>29.681</td>
<td>0.9977</td>
<td>94</td>
</tr>
<tr>
<td>Seed (4 weeks)</td>
<td>Ole1</td>
<td>-3.4709</td>
<td>23.897</td>
<td>0.9997</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Ole2</td>
<td>-3.5082</td>
<td>24.502</td>
<td>0.9997</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>Ole3</td>
<td>-3.6924</td>
<td>25.358</td>
<td>0.9984</td>
<td>87</td>
</tr>
<tr>
<td></td>
<td>Ole4</td>
<td>-3.3893</td>
<td>29.928</td>
<td>0.9929</td>
<td>90</td>
</tr>
<tr>
<td>Seed (6 weeks)</td>
<td>Ole1</td>
<td>-3.4313</td>
<td>23.272</td>
<td>0.9994</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Ole2</td>
<td>-3.6393</td>
<td>23.364</td>
<td>0.9994</td>
<td>88</td>
</tr>
<tr>
<td></td>
<td>Ole3</td>
<td>-3.5800</td>
<td>23.707</td>
<td>0.9983</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td>Ole4</td>
<td>-3.6150</td>
<td>30.064</td>
<td>0.9898</td>
<td>89</td>
</tr>
<tr>
<td>Seed (10 weeks)</td>
<td>Ole1</td>
<td>-3.3300</td>
<td>22.220</td>
<td>0.9992</td>
<td>100</td>
</tr>
<tr>
<td></td>
<td>Ole2</td>
<td>-3.4564</td>
<td>22.545</td>
<td>0.9991</td>
<td>95</td>
</tr>
<tr>
<td></td>
<td>Ole3</td>
<td>-3.5145</td>
<td>20.878</td>
<td>0.9990</td>
<td>93</td>
</tr>
<tr>
<td></td>
<td>Ole4</td>
<td>-3.4344</td>
<td>29.110</td>
<td>0.9967</td>
<td>95</td>
</tr>
<tr>
<td>Leaf</td>
<td>Ole1</td>
<td>-3.4126</td>
<td>35.856</td>
<td>0.9958</td>
<td>96</td>
</tr>
<tr>
<td></td>
<td>Ole2</td>
<td>-3.1539</td>
<td>34.505</td>
<td>0.9959</td>
<td>107</td>
</tr>
<tr>
<td></td>
<td>Ole3</td>
<td>-3.4722</td>
<td>34.761</td>
<td>0.9841</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Ole4</td>
<td>-3.6923</td>
<td>34.791</td>
<td>0.9866</td>
<td>87</td>
</tr>
<tr>
<td>Flower</td>
<td>Ole1</td>
<td>-3.0942</td>
<td>37.921</td>
<td>0.9831</td>
<td>110</td>
</tr>
<tr>
<td></td>
<td>Ole2</td>
<td>-3.5341</td>
<td>35.983</td>
<td>0.9982</td>
<td>92</td>
</tr>
<tr>
<td></td>
<td>Ole3</td>
<td>-3.4672</td>
<td>31.241</td>
<td>0.9997</td>
<td>94</td>
</tr>
<tr>
<td></td>
<td>Ole4</td>
<td>-3.4919</td>
<td>31.982</td>
<td>0.9970</td>
<td>93</td>
</tr>
</tbody>
</table>

qPCR reaction mixtures contained variable amounts of RNA-equivalent cDNA from tung seed (0.05, 0.5, 2.5, 5, 12.5 and 25 ng), the optimized concentrations of each primer (200 nM) and probe (200 nM) and qPCR Mix.
Table S2. Variation of Ole gene expression among tung trees by SYBR Green qPCR assay.

<table>
<thead>
<tr>
<th>mRNA</th>
<th>Tree 1 (fold)</th>
<th>Tree 2 (fold)</th>
<th>Tree 3 (fold)</th>
<th>Mean ± SD (fold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ole1</td>
<td>1</td>
<td>0.81 ± 0.17</td>
<td>0.88 ± 0.17</td>
<td>0.90 ± 0.10</td>
</tr>
<tr>
<td>Ole2</td>
<td>1</td>
<td>0.68 ± 0.15</td>
<td>0.52 ± 0.10</td>
<td>0.73 ± 0.24</td>
</tr>
<tr>
<td>Ole3</td>
<td>1</td>
<td>0.74 ± 0.16</td>
<td>0.47 ± 0.09</td>
<td>0.74 ± 0.26</td>
</tr>
<tr>
<td>Ole4</td>
<td>1</td>
<td>0.64 ± 0.14</td>
<td>0.60 ± 0.11</td>
<td>0.75 ± 0.22</td>
</tr>
</tbody>
</table>

The SYBR Green qPCR reaction mixtures (12.5 µl) contained 5 ng of RNA-equivalent cDNA from various stages of tung seeds, the optimized concentrations of each primer (200 nM) and QPCR Mix. The expression levels under each tree represent the means and standard deviations of the expression fold calculated using three reference mRNA (Rpl19b, Gapdh and Ubl) from 6 stages of seeds (weeks 2, 4, 5, 6, 8 and 10) with 2-4 assays for each stage. Ole gene expression in tree 1 seeds was used as the calibrator for the calculation of Ole gene expression in tree 2 and tree 3 seeds.
Table S3. Ole gene expression among different stages of tung seeds.

<table>
<thead>
<tr>
<th>qPCR method</th>
<th>mRNA</th>
<th>2 (fold)</th>
<th>4 (fold)</th>
<th>6 (fold)</th>
<th>10 (fold)</th>
</tr>
</thead>
<tbody>
<tr>
<td>TaqMan</td>
<td>Ole1</td>
<td>1</td>
<td>11.05 ± 2.49</td>
<td>39.58 ± 18.26</td>
<td>97.29 ± 74.92</td>
</tr>
<tr>
<td></td>
<td>Ole2</td>
<td>1</td>
<td>9.40 ± 2.12</td>
<td>48.38 ± 22.32</td>
<td>109.67 ± 84.45</td>
</tr>
<tr>
<td></td>
<td>Ole3</td>
<td>1</td>
<td>6.77 ± 1.33</td>
<td>40.91 ± 2.19</td>
<td>413.10 ± 2.61</td>
</tr>
<tr>
<td></td>
<td>Ole4</td>
<td>1</td>
<td>1.15 ± 0.26</td>
<td>3.11 ± 1.43</td>
<td>6.74 ± 5.198</td>
</tr>
<tr>
<td></td>
<td>Ole5</td>
<td>1</td>
<td>15.63 ± 3.52</td>
<td>17.98 ± 8.30</td>
<td>0.23 ± 0.18</td>
</tr>
<tr>
<td>SYBR Green</td>
<td>Ole1</td>
<td>1</td>
<td>14.81 ± 2.66</td>
<td>80.13 ± 63.97</td>
<td>107.53 ± 103.69</td>
</tr>
<tr>
<td></td>
<td>Ole2</td>
<td>1</td>
<td>8.01 ± 1.44</td>
<td>110.56 ± 88.25</td>
<td>99.30 ± 95.75</td>
</tr>
<tr>
<td></td>
<td>Ole3</td>
<td>1</td>
<td>6.17 ± 0.28</td>
<td>148.21 ± 1.45</td>
<td>306.52 ± 1.60</td>
</tr>
<tr>
<td></td>
<td>Ole4</td>
<td>1</td>
<td>1.38 ± 0.25</td>
<td>4.63 ± 3.70</td>
<td>5.58 ± 5.38</td>
</tr>
</tbody>
</table>

The qPCR reaction mixtures contained 25 ng in 25 µl (TaqMan qPCR) or 5 ng in 12.5 µl (SYBR Green qPCR) of RNA-equivalent cDNA from various stages of tung tree 1 seeds, the optimized concentrations of each primer and probe (200 nM) and QPCR Mix. The expression levels under each seed stage represent the means and standard deviations of the expression fold calculated using three reference mRNA (Rpl19b, Gapdh and Ubl) each with 2-4 assays. Ole gene expression in stage 2 seeds was used as the calibrator for the calculation of Ole gene expression in other stage seeds.