Tree Physiology 37, 1659–1671
doi:10.1093/treephys/tpx091

Research paper

Identification of potential genes that contributed to the variation in the taxoid contents between two *Taxus* species (*Taxus media* and *Taxus mairei*)

Chunna Yu[1,2], Hong Guo[1,2], Yangyang Zhang[1], Yaobin Song[1,3], Erxu Pi[1], Chenliang Yu[4], Lei Zhang[5], Ming Dong[1,3], Bingsong Zheng[6], Huizhong Wang[1,2] and Chenjia Shen[1,2,7]

[1]College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou 310036, China; [2]Zhejiang Provincial Key Laboratory for Genetic Improvement and Quality Control of Medicinal Plants, Hangzhou Normal University, Hangzhou 310036, China; [3]Key Laboratory of Hangzhou City for Ecosystem Protection and Restoration, College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou 310036, China; [4]Vegetable Research Institute, Zhejiang Academy of Agricultural Sciences, Hangzhou 310021, China; [5]Department of Plant Pathology, Washington State University, Pullman, WA 99164-6430, USA; [6]State Key Laboratory of Subtropical Silviculture, Zhejiang A & F University, Linan, Hangzhou 311300, China; [7]Corresponding author (shencj@hznu.edu.cn)

Received February 22, 2017; accepted June 22, 2017; published online July 17, 2017; Handling Editor Chung-Jui Tsai

Taxol is currently a valuable anticancer drug; however, the accumulated mixture of taxoids can vary greatly among *Taxus* species. So far, there is very little genomic information for the genus *Taxus*, except for *Taxus baccata*. Transcriptome analysis is a powerful approach to explore the different regulatory mechanisms underlying the taxoid biosynthesis pathway in *Taxus* species. First, we quantified the variation in the taxoid contents between *Taxus media* and *Taxus mairei*. The contents of paclitaxel and 10-deacetylpaclitaxel in *T. media* are higher than that in *T. mairei*. Then, the transcriptome profiles of *T. media* and *T. mairei* were analyzed to investigate the altered expressions. A total of 20,704 significantly differentially expressed genes (DEGs), including 9865 unigenes predominantly expressed in *T. media* and 10,839 unigenes predominantly expressed in *T. mairei*, were identified. In total, 120 jasmonic acid-related DEGs were analyzed, suggesting variations in 'response to JA stimulus' and 'JA biosynthetic process' pathways between *T. media* and *T. mairei*. Furthermore, a number of genes related to the precursor supply, taxane skeleton formation and hydroxylation, and C13-side chain assembly were also identified. The differential expression of the candidate genes involved in taxoid biosynthetic pathways may explain the variation in the taxoid contents between *T. media* and *T. mairei*.

Keywords: HPLC–MS/MS, taxoid biosynthetic pathway, Taxol, *Taxus*, transcriptome.

Introduction

The natural product Taxol (generic name paclitaxel), which was first extracted from *Taxus brevifolia*, is an effective antitumor drug worldwide (Ketchum et al. 1999). Since the 1990s, Taxol's efficacy has been well documented, and it was approved by the US FDA for the treatment of several cancers, such as ovarian, breast, lung, and head and neck cancers (Rowinsky et al. 1993, Kohler and Goldspiel 1994). As the cancer incidence increases, so does the commercial value of Taxol. However, due to the limited *Taxus* spp. resources, the demand is far greater than Taxol production (Li et al. 2012).

Species in the genus *Taxus* contain a large number of taxanes, a class of diterpenoid with a taxane skeleton (Wang et al. 2011). The fully functionalized diterpenoid Taxol is one of the hundreds of identified taxoid secondary metabolites (Xiao et al. 2004, Wang et al. 2011). Synthesis is a costly approach to Taxol production (Kaspera and Croteau 2006). Therefore, a deep understanding of Taxol biosynthesis, enzymes and the related genes is essential for the genetic improvement of Taxol production.

Taxol biosynthesis is an intricate metabolic pathway that involves 19 steps (Howat et al. 2014, Jennewein et al. 2004*b*).

In detail, the diterpenoid taxane core is derived by three units of C5 isoprenoid precursor isopentenyl diphosphate (IPP) and one dimethylallyl diphosphate (DMAPP) derived from the 2-C-methyl-D-erythritol phosphate (MEP) pathway (Eisenreich et al. 1996). The conversion of IPP to DMAPP is catalyzed by a plastidial IPP isomerase (Rohdich et al. 2002, Jennewein et al. 2004*b*). The cyclization of the geranylgeranyl skeleton to form the taxane skeleton is catalyzed by taxadiene synthase (TS) (Harrison et al. 1966). The cyclization by TS is a slow step in the pathway, but not a rate-limiting step (Hezari et al. 1995, Williams et al. 2000). An important enzyme, 10-deacetylbaccatin III-10-O-acetyltransferase (DBAT), takes part in the formation of baccatin III (Walker and Croteau 2000). A phenylalanine-derived C_{13}-side chain is appended to baccatin III by baccatin III-3-amino, 3-phenylpropanoyltransferase (BAPT) enzyme, leading to the formation of 3′-*N*-debenzoyl-2′-deoxytaxol (Nasiri et al. 2016). The hydroxylation of the taxane core involve a series of cytochrome P450 taxoid hydroxylases, including the 2α-, 5α-, 7β-, 9α-, 10β-, 13α- and 14β-hydroxylases (Chau and Croteau 2004, Chau et al. 2004). Several more enzymatic reactions, such as an oxidation at C9 site and the formation of the oxetane, occur late in the baccatin III formation pathway (Giner and Faraldos 2003). Through P450-related hydroxylation of the side chain, 3′-*N*-debenzoyl-2′-deoxytaxol is converted into paclitaxel by 3′-*N*-debenzoyl-2′-deoxytaxol-*N*-benzoyl transferase (DBTNBT) (Lenka et al. 2015). Lastly, the assembly of C13-side chain attached to baccatin III is thought to be the final step of the Taxol biosynthesis pathway (Kingston 1995, Jiménez-Barbero et al. 2002).

A whole-genome sequencing of *Taxus baccata* (English Yew) to construct a draft genome assembly has been published (Nystedt et al. 2013). Recently, RNA sequencing (RNA-Seq) has been applied to uncover various aspects of Taxol biosynthesis in different *Taxus* species. In 2010, the first transcriptome of the *Taxus* genus was published using cultured cambial meristematic cells from *Taxus cuspidata* (Lee et al. 2010). Then, the tissue-specific transcriptome of *Taxus mairei* was sequenced using Illumina paired-end sequencing technology (Hao et al. 2011). Later in the same year, the transcriptome of *T. cuspidata* needles was sequenced using 454 pyro-sequencing (Wu et al. 2011). Applications of methyl jasmonate (MeJA) significantly promote the contents of baccatin III and paclitaxel in *Taxus* spp. suspension cultures (Yukimune et al. 1996), and a transcriptome analysis indicated the involvement of MeJA in the activation of the terpenoid biosynthesis branch of Taxol synthesis (Li et al. 2012). A deep sequencing of *Taxus media* cells found that the expression of 18 genes associated with terpenoid backbone and paclitaxel biosynthesis increased following induction by MeJA (Sun et al. 2013). A comparison between the transcriptomes of mock and MeJA-elicited cultures also suggested a role of MeJA in cell division by inhibiting the expression of genes participating in cell cycle progression (Patil et al. 2014).

Most of the previous RNA-Seq data about *Taxus* spp. were from only one *Taxus* species or one *Taxus* species under different treatments. Thus, it is hard to compare two transcriptomes from different *Taxus* species based on the published data. *Taxus media*, a hybrid *Taxus* species with a high content of Taxol, is a major source for industrial Taxol production (Kai et al. 2005, Shen et al. 2016*a*). *Taxus mairei*, another member of the Taxaceae family, is widely distributed and broadly planted in southeastern China. Our study compared two different species using data generated by the same platform, and it gave us an opportunity to analyze the differences between *T. media* and *T. mairei* at the transcriptional level.

Materials and methods

Plant materials and RNA isolation

Fresh twig samples were collected from 3-year-old cultivated *T. media* and *T. mairei* plants in March 2015 (Figure 1A). The plants were grown in a greenhouse of Hangzhou Normal University, Hangzhou, China (East longitude 118° 21′-120° 30′; North latitude 29° 11′-30° 33′), at a temperature of 25 ± 1 °C with a light/dark cycle of 12/12 h and 60–70% relative humidity. Total RNAs were isolated by a TRIzol reagent (Invitrogen,

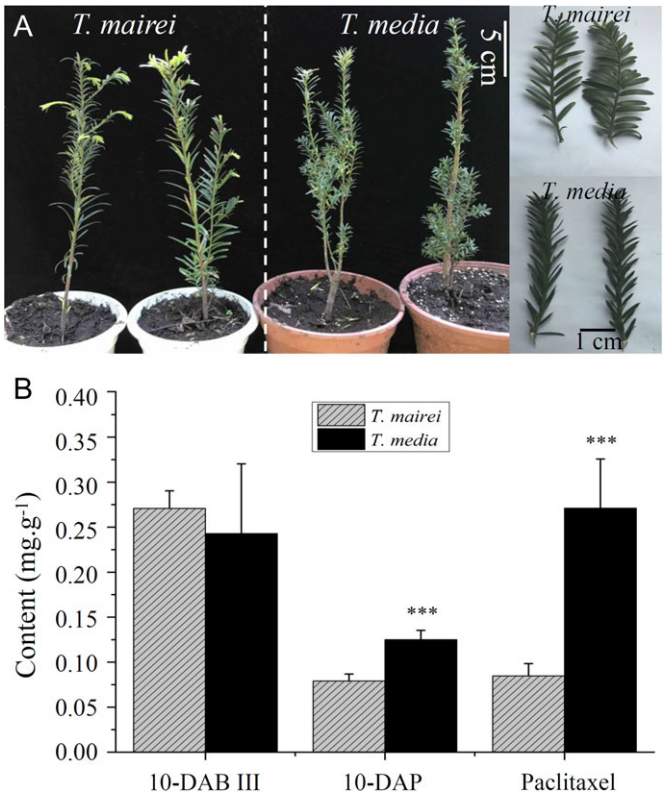

Figure 1. Variation of the taxoid contents between *T. media* and *T. mairei*. (A) A picture of *T. media* and *T. mairei* under greenhouse condition. (B) The contents of paclitaxel and two intermediates, including 10-DAB III and 10-DAP, were quantified by HPLC–MS/MS method. The significant variations ($P < 0.001$) in paclitaxel and 10-DAP contents are indicated by ***. Error bars represent mean + SD ($n = 6$).

Shanghai, China) according to the manufacturer's protocol. Total RNAs were quantified using a Bioanalyzer 2100 and the quality was confirmed using an RNA 6000 Nano LabChip Kit (Agilent, Santa Clara, CA, USA) with an RNA integrity number >7.0.

Chemicals and reagents

Paclitaxel (≥99%; CAS No. 33069-62-4) and 10-deacetylbaccatin III (10-DAB III; ≥98%; CAS No. 32981-86-5) were purchased from Aladdin Biochemical Technology (Shanghai, China). 10-Deacetylpaclitaxel (10-DAP; 98%; CAS No. 78432-77-6) was obtained from J & K Scientific (Beijing, China). HPLC-grade methanol and formic acid were purchased from Merck (Kenilworth, NJ, USA).

Preparation of crude extracts from Taxus species

Fresh twigs were cut from *T. media* and *T. mairei*. They were dried at 40 °C and powdered. The powder was passed through a filter (0.42 mm pore size). To prepare crude extracts, a previously published method was applied with some modifications (Li et al. 2009). In brief, 2.0 g fine powder was mixed with 30 ml of 100% methanol and then the mixture was ultrasonicated for 60 min. After centrifugation at 5000 rpm for 5 min, the supernatant was transferred to a new tube. Then, the crude extracts of *T. media* and *T. mairei* were diluted at ratios of 1:10 and 1:2, respectively. Before HPLC–MS/MS analysis, the samples were filtered through 0.22-μm membrane filters.

HPLC–MS/MS analysis

The quantifications of paclitaxel, 10-DAB III and 10-DAP were carried out by using Thermo Dionex UltiMate 3000 series HPLC system equipped with a Finnigan TSQ Quantum Discovery triple quadrupole MS (Thermo Fisher Scientific, Waltham, MA, USA). The separation of 10-DAB III, 10-DAP and paclitaxel was achieved on a Phenomenex Kinetex C18 column with 100 × 4.6 mm and 2.6-μm particle size (Phenomenex, Torrance, CA, USA). The mobile phase consisted of 35% (v/v) 2 mM ammonium formate–0.1% formic acid aqueous solution and 65% (v/v) methanol. The column oven was kept at 30 °C. The flow rate was set at 0.2 ml min^{-1}, and the injection volume was 5 μl. The positive electrospray ionization mode was employed. The capillary temperature was set at 270 °C, and the ion spray voltage was set at 3000 V. Nitrogen was used as both the sheath gas and auxiliary gas. High purity argon was used as collision gas. Multiple reaction monitoring was applied for the determination. The transition of *m/z* 567.2 → 445.3 was used for 10-DAB III quantification and the transitions of *m/z* 567.2 → 385.2 and 567.2 → 427.3 were utilized for confirmation. The transitions of *m/z* 829.4 → 286.1 and 829.4 → 122.0 were measured for 10-DAP quantification and confirmation, respectively. The transition of *m/z* 876.4 → 308.1 was chosen for paclitaxel quantification and the transitions of *m/z* 876.4 → 531.2 and 876.4 → 591.4 were utilized for confirmation. The Xcalibur 2.2 software (Thermo Fisher Scientific) was used for data acquisition and processing.

cDNA library preparation and sequencing

An amount of 10 μg RNA for each sample was used for cDNA library preparation. The RNA samples were collected for three biological replications RNAs were subjected to poly(T) oligo-attached magnetic beads (Thermo Fisher Scientific) and then were fragmented with divalent cations at a relatively high temperature. An mRNA-Seq sample preparation kit (Illumina, San Diego, CA, USA) was used to reverse-transcribe the fragmented RNAs to the final cDNA library used in sequencing following the manufacturer's procedure. The paired-end sequencing was performed on an Illumina Hiseq2500 platform following the manufacturer's protocol, and 100/50-bp paired/single-end reads were produced.

Quality checking, de novo assembly and read annotation

Raw reads were processed using perl scripts provided by LC-Bio, Hangzhou, China. The low quality reads and adapter- or poly-N-containing reads were removed. To evaluate the quality of clean reads, several parameters, such as GC-content, Q20, Q30 and sequence duplication level, were calculated. A de novo strategy was employed to assemble the transcriptomes of *T. media* and *T. mairei* based on 25.75 and 17.85 Gb of data, respectively, using the Trinity assembly program with default parameters (Grabherr et al. 2011). The unigene annotation was performed by a locally installed BLASTall algorithm-based program, with an *E*-value threshold of 10^{-10}, to several protein databases, including NCBI non-redundant (Nr) (http://www.ncbi.nlm.nih.gov/protein/), Swiss-Prot protein (http://www.uniprot.org/) and Kyoto Encyclopedia of Genes and Genomes (KEGG) (http://www.genome.jp/kegg/). A classification of the metabolic pathway was produced by KEGG, and Gene Ontology (GO) classifications were carried out by Blast2GO (https://www.blast2go.com/) (Shen et al. 2016*b*).

Calculating transcript's expression levels and a DEG analysis

The clean reads were mapped to the assembled transcript by Bowtie (version 0.12.7) using the 'single-end' method. Normalized expression values of each gene were calculated by the reads per kilobase per million reads method (Mortazavi et al. 2008). Data sets of three distinct biological samples from the same *Taxus* spp. were treated as a group, and the differential expression between two groups was analyzed using the DESeq R package v1.10.1. Here, the false discovery rate + was used to calculate the *P*-value in the significance tests, and then the resulting *P*-values were adjusted using the Benjamini and Hochberg's method. $P < 0.05$ was used as a threshold to determine a significant difference in the comparison between two groups. The absolute value of the log2(Ratio) of differential

expression between different samples from *T. media* and *T. mairei* was calculated.

Phylogenetic tree building

The predicated full-length sequences of P450 proteins were used for multiple sequence alignments. The alignments were performed by ClustalW with default parameters and were visualized by the software GeneDoc (http://www.nrbsc.org/gfx/genedoc/). An unrooted phylogenetic tree was built using MEGA6.1 (http://www.megasoftware.net/) employing the neighbor joining method (Saitou and Nei 1987).

Real-time PCR validation

Total RNA from different samples was isolated by a Plant RNeasy Mini kit (Qiagen, Hilden, Germany) according to our previous work (Shen et al. 2015). Genomic DNA contamination was cleaned by DNase I. Then, the cDNA was synthesized by ReverAid First Strand cDNA Synthesis Kit (Thermo Scientific, Shanghai, China).

QRT-PCR was performed using the SYBR Premix Ex Taq Kit (TaKaRa, Dalian, China) and a DNA Sequence Detection System (ABI PRISM 7700, Applied Biosystems, Shanghai, China). Independent cDNA samples from *T. media* and *T. mairei* were used for real-time PCR experiments. Inhibition of reverse-transcription activity and PCR has been checked by dilution of the sample. No inhibition has been observed. An *ACTIN* sequence was used as the internal standard gene to calculate relative fold differences by the values of comparative cycle threshold ($2^{-\Delta\Delta Ct}$). Sterile ddH_2O was used as no-template control. All the primer sequences are listed in Table S1 available as Supplementary Data at *Tree Physiology* Online. In detail, 1 μl cDNA was added to 5 μl of 2 × SYBR® Green buffer, 0.1 μM of each primer and ddH_2O to a working volume of 10 μl. The PCR program was 95 °C for 10 min, followed by 40 cycles of 95 °C for 10 s and 60 °C for 60 s, 72 °C for 15 s at last. Five technical repetitions were performed in qRT-PCR experiments. Software LERTPA-V1.0 was utilized for calculation of gene expression data.

The detection limit and amplification efficiency of qRT-PCR were determined using 10-fold serial dilution of cDNA isolated from one sample to create the standard curve. The PCR efficiency of each primer pair was calculated by the slopes and correlation coefficients of its standard curve. The value of PCR efficiency and linear dynamic range for each primer pair are also listed in Table S1 available as Supplementary Data at *Tree Physiology* Online. The expression analysis was performed for three biological replications, and the figures show the average values.

Statistical analyses

Statistical analyses were carried out using SPSS software version 19.0 (SPSS Inc., Chicago, IL, USA), and a one-way ANOVA was applied to compare taxoid content differences between the two *Taxus* species. A *P*-value < 0.05 was considered to be statistically significant and indicated by *; $P < 0.01$ was indicated by **; $P < 0.001$ was indicated by ***.

Results

Variation in the taxoid contents between T. media and T. mairei

The dried twigs from *T. media* contained 0.17–0.32 mg g^{-1} paclitaxel while *T. mairei* contained 0.07–0.11 mg g^{-1} paclitaxel. The average content of paclitaxel in *T. media* was 3.2-fold higher than that in *T. mairei*. Additionally, *T. media* contained significantly more 10-DAP than *T. mairei*. However, the 10-DAB III content was not significantly different between the two species (Figure 1B and Figure S1 available as Supplementary Data at *Tree Physiology* Online).

Transcriptome sequencing of T. media and T. mairei twig samples

To investigate the differentially expressed genes (DEGs) between *T. media* and *T. mairei*, the intricate gene expression profiles were analyzed. More than 96.55% of the clean reads had quality scores at the level of Q20, and more than 92.77% of the clean reads had quality scores at the level of Q30 (Table S2 available as Supplementary Data at *Tree Physiology* Online). All reads obtained from the two *Taxus* species were assembled using Trinity software, resulting in 136,635 transcripts (N50: 1,426), with a mean length of 817 bp, and 88,194 unigenes (N50: 1543), with a mean length of 715 bp. There were 72,621 unigenes (82.33%) with lengths <1000 bp, 8943 unigenes (10.15%) in the range of 1000–2000 bp, and 6630 (7.52%) with lengths >2000 bp in the *T. mairei* library (Figure 2A). The GC content frequency distribution was shown in Figure 2B. There were 34,523 unigenes (39.14%) matching the protein sequences in the Nr database, 22,539 (25.56%) in the SwissProt database, 12,361 (14.02%) in the KEGG database, 26,814 (30.40%) in the KOG database, 25,912 (29.38%) in the Pfam database and 20,434 (23.17%) in the GO database (Figure 2C and Table S3 available as Supplementary Data at *Tree Physiology* Online). A number of *Taxus* unigenes showed high similarities to the genes in other plant species. The species distribution of the annotated unigenes is shown in Figure 2D. The largest number of *Taxus* homologous genes was identified in *Picea sitchensis*, a woody plant.

A total of 18,848 unigenes (19.8%) were grouped into 25 KOG functional categories. Interestingly, 1137 unigenes (~5%) belonged to the KOG category 'secondary metabolites biosynthesis', which contained most of the Taxol biosynthesis-related genes. In detail, the CYP2 subfamily (KOG0156), CYP4/19/26 subfamilies (KOG0157) and iron/ascorbate family oxidoreductase (KOG0143) were identified as the top three KOG terms in the 'secondary metabolites biosynthesis' category (Figure 2E).

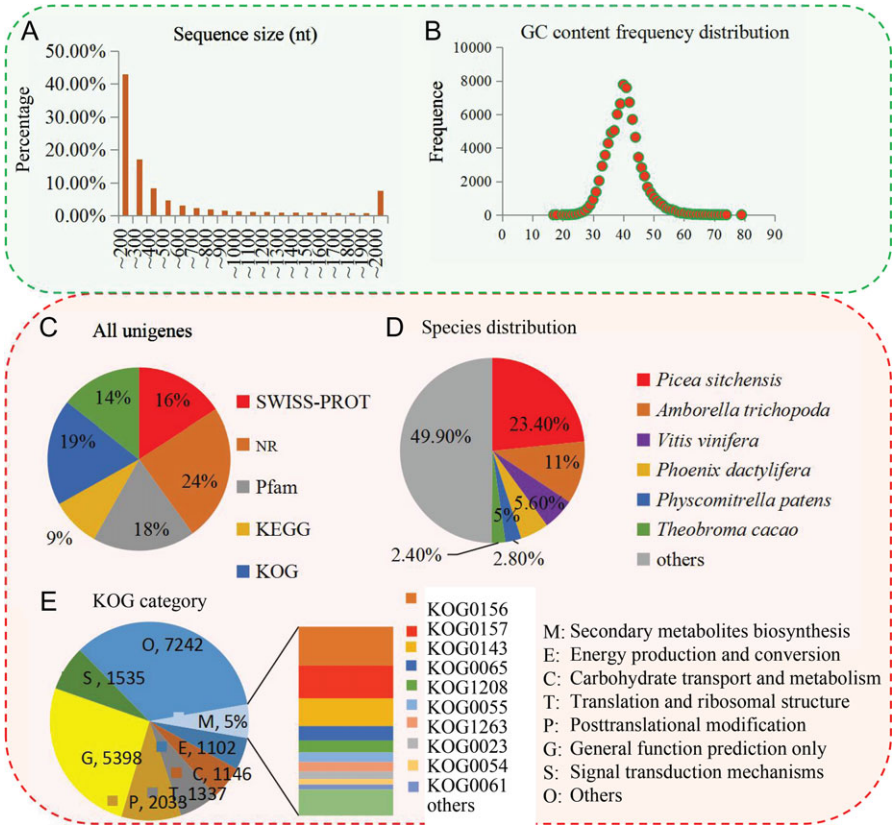

Figure 2. Illumina sequencing of *T. media* and *T. mairei*. Classification of raw reads generated by Illumina sequencing. (A) The length distribution of all assembled unigenes. (B) The GC content frequency distribution of all assembled unigenes. (C) The number of unigenes annotated by different databases, including Nr, Swissprot, KOG, KEGG, GO and Pfam. (D) Species distribution of all annotated unigenes. (E) KOG category classification of all unigenes.

Classifications of enriched GO and KEGG terms

A total of 20,434 unigenes (21.4%) could be assigned to at least one GO term. Within the biological process category, the most highly represented terms were 'regulation of transcription' and 'transcription, DNA-dependent'. Within the cellular component category, 'integral to membrane' and 'nucleus' were the two most abundant terms. The most enriched terms within the molecular function category were 'ATP binding', 'binding' and 'protein serine/threonine kinase' (Figure 3A).

A total of 7458 unigenes (7.8%) were classified into various KEGG signaling and metabolic pathways, including pathways related to organismal systems, metabolism, cellular process, environmental information processing and genetic information processing. Interestingly, the most enriched KEGG pathways included those related to metabolic pathways, such as 'carbohydrate metabolism' (1168 unigenes), 'amino acid metabolism' (1036 unigenes), 'energy metabolism' (819 unigenes) and 'biosynthesis of other secondary metabolites' (727 unigenes). For organismal systems, the largest number of unigenes (444 unigenes) was assigned into 'immune system'. For genetic information processing, the 'translation' term contained the largest number of unigenes (973 unigenes). For environmental information processing, the 'signal transduction' term contained the largest number of unigenes (439 unigenes) (Figure 3B).

Identification of DEGs between T. media *and* T. mairei

A total of 20,704 significantly DEGs, including 9865 *T. media* predominantly expressed unigenes and 10,839 *T. mairei* predominantly expressed unigenes, were identified and analyzed using criteria of twofold differences and $P_{adj} < 0.05$ (Figure 4A).

Within the DEGs, 204 GO terms, such as 'sequence-specific DNA binding', 'electron carrier activity', 'defense response', 'heme binding' and 'ethylene-mediated signaling pathway', were significantly enriched (Figure 4B). We mapped the DEGs to the canonical reference pathways in KEGG database to identify the differential metabolic pathways between *T. media* and *T. mairei*. All of the DEGs could be assigned to 247 predicted metabolic pathways, of which 27 pathways were significantly enriched ($P < 0.05$). A large number of *T. media* predominantly expressed unigenes were enriched in six KEGG pathways, ribosome (5%), phenylpropanoid biosynthesis (5%), starch and sucrose metabolism (5%), purine metabolism (5%), limonene and pinene degradation (4%), and pyrimidine metabolism

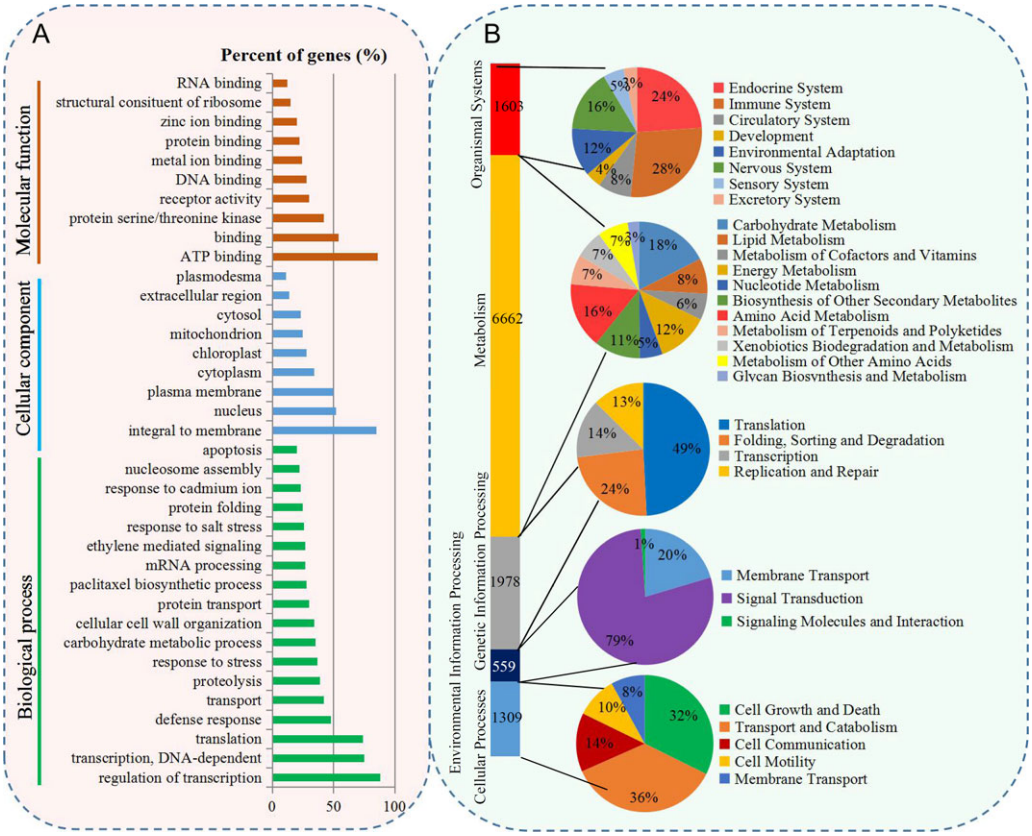

Figure 3. Classification of enriched GO and KEGG terms. (A) A total of 20,434 unigenes were assigned to different GO terms. (B) A total of 7458 unigenes were assigned to different KEGG terms.

(4%). On the other hand, many *T. mairei* predominantly expressed unigenes were enriched in six pathways, ribosome (17%), cysteine and methionine metabolism (5%), oxidative phosphorylation (5%), phenylpropanoid biosynthesis (4%), flavonoid biosynthesis (4%) and pyridine alkaloid biosynthesis (4%) KEGG pathways (Figure 4C).

Differential expression of JA metabolism- and signaling pathway-related genes

Increasing evidence showed that exogenous MeJA could significantly promote the production of paclitaxel and baccatin III in various *Taxus* species (Yukimune et al. 1996, Sun et al. 2013, Patil et al. 2014). In our study, we compared the expression of the JA metabolism- and signaling pathway-related genes between *T. media* and *T. mairei* based on their GO terms. In total, 120 DEGs were grouped into 10 GO terms, including GO:0009753, GO:0009695, GO:2000022, GO:0071395, GO:0009861, GO:0009864, GO:0009867, GO:0009694, GO:0080141 and GO:0009871, which were identified as JA-related unigenes (Figure 5A). Among these GOs, 'response to JA stimulus' (GO:0009753, $P = 0.0066$) and 'JA biosynthetic process' (GO:0009695, $P = 0.0168$) showed significant differences between *T. media* and *T. mairei* (Figure 5B). Furthermore, GO:0009753 contained the largest number of JA-related DEGs, including 27 DEGs predominantly expressed in *T. media* and 36 DEGs predominantly expressed in *T. mairei*. In addition, GO:0009695 contained the second largest number of JA-related DEGs, including 19 *T. media* predominantly expressed DEGs and five *T. mairei* predominantly expressed DEGs (Figure 5C).

DEGs related to precursor supply in Taxol biosynthesis

Our transcriptome data revealed at least one unigene encoding each of the key steps in the MEP and mevalonate (MVA) pathways upstream of Taxol biosynthesis (Croteau et al. 2006). In the MEP pathway, three unigenes encoding 1-deoxy-D-xylulose 5-phosphate synthase (DXS) were annotated, and two of them, DXS1: DN36084 and DXS2: DN36007, were substantially more abundant in *T. media* than in *T. mairei*. 1-Deoxy-D-xylulose 5-phosphate reductoisomerase (DXR), 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase (MCT), 4-(cytidine 5′-diphospho)-2-C-methyl-D-erythritol kinase (CMK), 4-hydroxy-3-methylbut-2-enyl-diphosphate synthase (HDS) and 4-hydroxy-3-methylbut-2-enyl-diphosphate reductase (HDR)-encoding genes were each represented by one unigene. Among these five unigenes, three (DN31655 for DXR, DN34746 for MCT and DN39532 for HDR) were predominantly expressed in *T. media*. In addition, three unigenes, DN12291, DN24680 and DN32294,

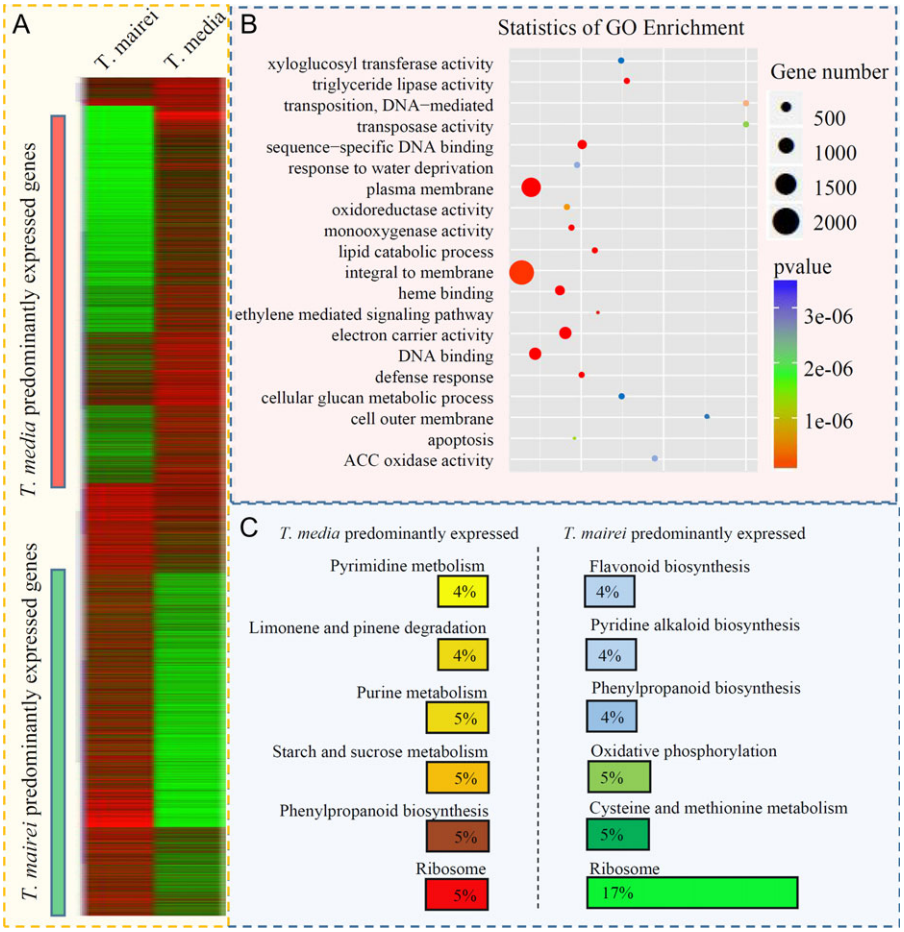

Figure 4. Transcriptional variation between *T. media* and *T. mairei*. (A) Expression profiles of the DEGs between *T. media* and *T. mairei* are shown by a heatmap. The red block indicates the *T. media* predominantly expressed genes and the green block indicates the *T. mairei* predominantly expressed genes. (B) The GO term enrichment analysis of DEGs between *T. media* and *T. mairei*. The top 20 enriched GO terms are shown. (C) The KEGG analysis of DEGs between *T. media* and *T. mairei*. The top six enriched KEGG terms of *T. media* predominantly expressed and *T. mairei* predominantly expressed DEGs are shown.

encoding 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase (MDS) were also identified. Interestingly, the three MDS genes were more predominantly expressed in *T. media* than in *T. mairei*. In the MVA pathway, two unigenes encoding acetyl-CoA C-acetyltransferase (AACT) were annotated, and one of them (AACT 1: DN43039) was predominantly expressed in *T. media*. Unigenes encoding 3-hydroxy-3-methylglutaryl-CoA reductase and diphospho-MVA decarboxylase were substantially more abundant in *T. media* than in *T. mairei*. On the contrary, unigenes encoding 3-hydroxy-3-methylglutaryl-CoA synthase and MVA kinase were substantially more abundant in *T. mairei* than in *T. media* (Figure 6).

In addition to these key steps, a unigene (DN31318) encoding IPP isomerase, which controls the reversible conversion of IPP to DMAPP in the cytosol and mitochondria, respectively, was also identified (Berthelot et al. 2012). This unigene (DN31318) was more predominantly expressed in *T. mairei* than in *T. media*. Three units of IPP and one unit of DMAPP are required to form GGPP, a universal, acyclic precursor for the formation of the taxane skeleton, catalyzed by GPP synthase (GPPS) (Eisenreich et al. 1996). Two full-length unigenes encoding GPPS were identified. One, DN35175, was predominantly expressed in *T. media*, while the other, DN34725, showed similar expression levels in *T. media* and *T. mairei*. Another unigene, DN31147, which was predominantly expressed in *T. media*, was identified as GGPP synthase (Figure 6).

DEGs related to the formation of the taxane skeleton and C13-side chain assembly

The formation of the taxane skeleton was first suggested by Lythgoe's group many years ago (Harrison et al. 1966). The cyclization of GGPP skeleton to taxa-4(5),11(12)-diene is catalyzed by TS. In our study, only one unigene (DN34420) encoding TS was identified, and its expression was very low in *T. mairei* (Figure 7A and B).

The β-phenylalanoyl-type side chain is a constituent of taxoids. α-Phenylalanine is converted to β-phenylalanine by a phenylalanine aminomutase (Croteau et al. 2006). Four full-length unigenes (DN40261, DN40642, DN41966 and DN33648) were identified as phenylalanine aminomutase-

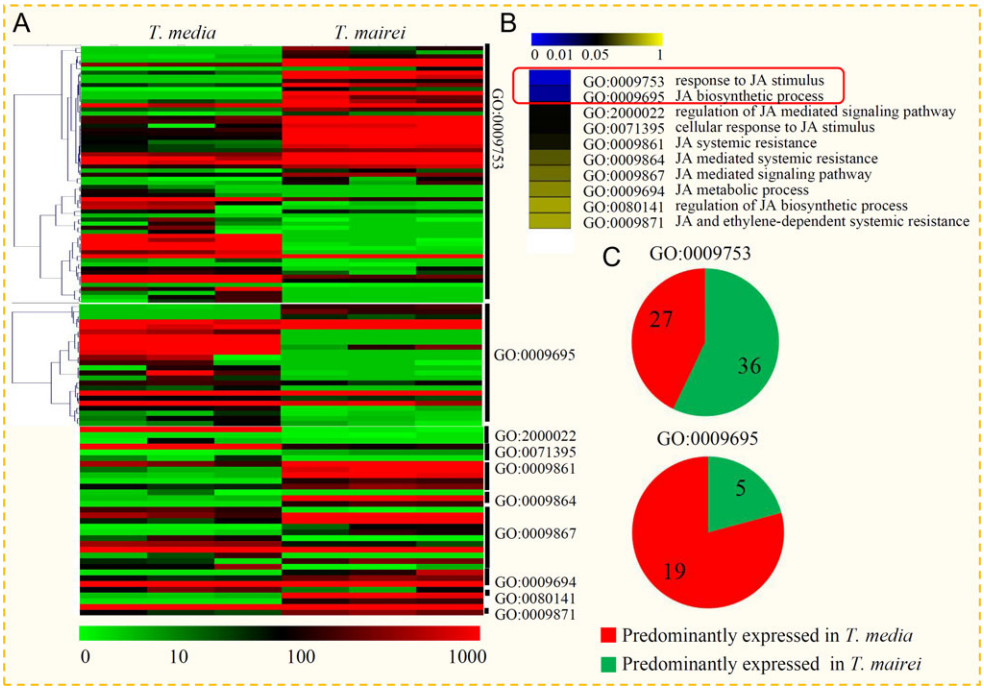

Figure 5. Differential expression of the genes related to jasmonic acid metabolism and signaling pathway. (A) Expression profiles of the genes related to JA metabolism and signaling pathway are shown by a heatmap. All the JA-related genes were assigned into 10 GO terms. (B) Significance analysis of 10 JA-related GO terms between *T. media* and *T. mairei*. The red box indicates significant GO terms. (C) The number of up-regulated genes and down-regulated genes that grouped into GO:0009753 and GO:0009695. Red indicates up-regulated genes and green indicates down-regulated genes.

encoding genes. Based on our transcriptome data, the expression of DN40261 and DN33648 were higher in *T. media* than in *T. mairei*, and the expression of DN40642 was more abundant in *T. mairei* than in *T. media*. However, the expression of DN41966 was almost the same in *T. media* and *T. mairei* (Figure 7A and C).

The transfer of the taxane core involves the benzoylation at the C2α-position of the substrate 2-debenzoyl-7,13-diacetylbaccatin III and is catalyzed by taxoid 2α-O-benzoyl transferase (Walker and Croteau 2000). Two unigenes were annotated in our study. One unigene, DN34347, was predominantly expressed in *T. mairei* and another one, DN12834, was predominantly expressed in *T. media* (Figure 7D and E)

In total, eight unigenes encoding four important cytochrome P450 taxoid oxygenases were identified. The unigenes encoding the taxadiene-5α-hydroxylase and the taxoid-14β-hydroxylase were more highly expressed in *T. media* than in *T. mairei*. Three unigenes, DN40916, DN29240 and DN36663, encoding taxoid-10β-hydroxylase were identified, and all three were more predominantly expressed in *T. media* than in *T. mairei*. In addition, three unigenes, DN39262, DN30809 and DN34169, encoding taxoid-13α-hydroxylase (T13OH) were also identified (Figure 8B and C). Interestingly, all eight cytochrome P450 taxoid oxygenase-encoding genes were grouped into Clan 85, suggesting a common origin for these P450 genes.

Differential expression of the cytochrome P450 taxoid oxygenase-encoding genes

Cytochrome P450 monooxygenases play essential roles in Taxol biosynthesis (Kaspera and Croteau 2006). In our study, the full-length cDNAs of 90 cytochrome P450 candidate genes were used to construct a phylogenetic tree. These P450 genes could be grouped into seven major groups, Clans 74, 97, 86, 72, 71, 710 and 85, according to the nomenclature (Rasool and Mohamed 2016). The largest group was Clan 71, which contained 33 P450 members, followed by Clan 85. Clan 710 was the smallest, with only one identified P450 member (Figure 8A).

Differential expression of other genes alongside production of associated taxanes

BAPT, *DBAT* and *DBTNBT* genes were reported to contributed to the taxol biosynthetic pathway (Nasiri et al. 2016). However, no BAPT-, DBAT- or DBTNBT-encoding unigenes were identified in our transcriptomes. In this study, the variations in transcriptional levels of *DBAT*, *BAPT* and *DBTNBT* genes were analyzed by qRT-PCR. *DBAT* and *BAPT* were more predominantly expressed in *T. media* than in *T. mairei*. The expression level of *DBTNBT* was almost the same in *T. media* and *T. mairei* (Figure S2 available as Supplementary Data at *Tree Physiology* Online).

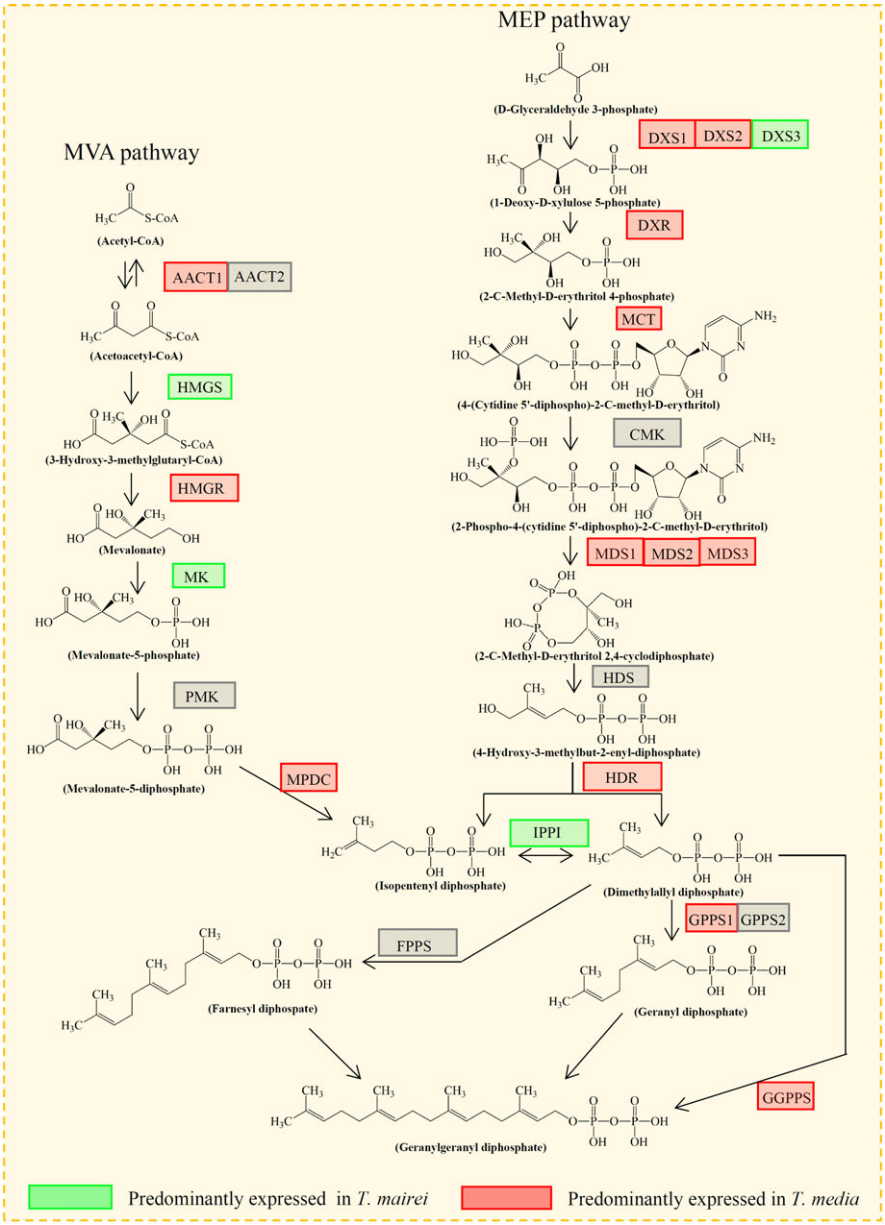

Figure 6. Differential expression of the genes related to precursor supply of Taxol biosynthesis. Overview of the MVA and MEP pathways in *Taxus* species. The red block indicates the *T. media* predominantly expressed genes and the green block indicates the *T. mairei* predominantly expressed genes. Enzyme abbreviations are: AACT, acetyl-CoA C-acetyltransferase; HMGS, 3-hydroxy-3-methylglutaryl-CoA synthase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; MK, MVA kinase; PMK, phospho-MVA kinase; MPDC, diphospho-MVA decarboxylase; DXS, 1-deoxy-D-xylulose 5-phosphate synthase; DXR, 1-deoxy-D-xylulose 5-phosphate reductoisomerase; MCT, 2-C-methyl-D-erythritol 4-phosphate cytidylyltransferase; CMK, 4-(cytidine 5-diphospho)-2-C-methyl-D-erythritol kinase; MDS, 2-C-methyl-D-erythritol 2,4-cyclodiphosphate synthase; HDS, 4-hydroxy-3-methylbut-2-enyl-diphosphate synthase; HDR, 4-hydroxy-3-methylbut-2-enyl diphosphate reductase.

Discussion

Large variations in the contents of paclitaxel and related taxoids among different *Taxus* species have been previously documented (Bala et al. 1999, van Rozendaal et al. 2000, Onrubia et al. 2011, Sabater-Jara et al. 2014, Nasiri et al. 2016). Our data confirmed that there are great differences in the contents of 10-DAP and paclitaxel between *T. media* and *T. mairei* (Figure 1B). However, the mechanisms responsible for the interspecific difference in taxoid accumulations between different *Taxus* species has not been well-studied. Thus, determining the DEGs provides the opportunity to elucidate the molecular mechanisms underlying the paclitaxel biosynthetic pathway.

Our results indicated that there are a large number of transcriptional alterations between *T. media* and *T. mairei*. In recent years, the availability of transcriptome data for *Taxus* species has increased. For example, data sets from three different tissues of cultivated *T. mairei* having 23,515 unigenes were generated by

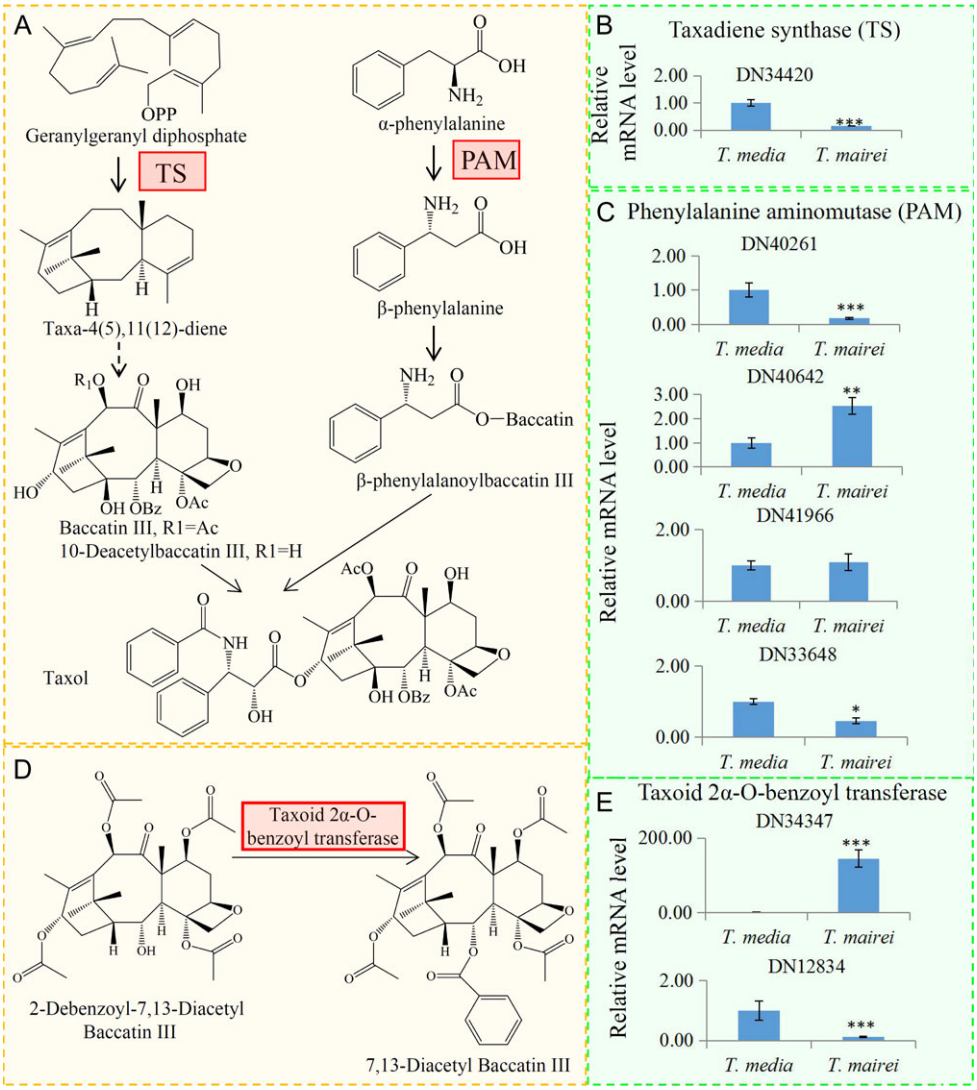

Figure 7. Differential expression of the genes related to the formation of the taxane skeleton and C13-side chain assembly. (A) Outline of the Taxol biosynthetic pathway. (B) The transfer to the taxane core involves benzoylation at the C2α-position of the surrogate substrate 2-debenzoyl-7,13-diacetylbaccatin III and is catalyzed by the taxoid 2α-O-benzoyl transferase. (C) Differential expression of one TS-encoding gene, four PAM-encoding genes and two taxoid 2α-O-benzoyl transferase-encoding genes. The abbreviations are: TS, taxadiene synthase and PAM, phenylalanine aminomutase. The significant variations ($P < 0.05$) are indicated by *; the significant variations ($P < 0.01$) are indicated by **; the significant variations ($P < 0.001$) are indicated by ***. Error bars represent mean ± SD ($n = 6$).

Yang's group (Hao et al. 2011); two samples having 46,581 unigenes were published by Yu's group (Li et al. 2012); and 40,348 unigenes were published by Qiu's group (Sun et al. 2013). According to previous studies, any change in taxane production occurs simultaneously with a reprogramming of cell metabolism, particularly secondary metabolism (Cusido et al. 2014). In our study, KEGG analysis revealed a large number of metabolism-related genes, including 727 'biosynthesis of other secondary metabolites'-related genes.

Applications of exogenous MeJA to the medium at the early culture stage increase the production of taxane by increasing the expression of genes in the terpenoid biosynthesis pathway (Yukimune et al. 1996, Bentebibel et al. 2005, Li et al. 2012). In our study, a large number of JA-related DEGs were identified, suggesting variations in the JA metabolism and signaling pathways between *T. media* and *T. mairei*. Interestingly, one of the most significantly differentially expressed GO terms was GO:0009695 'JA biosynthetic process'. For example, 19 GO:0009695-related unigenes, including *DAD1*, *CLL5*, *ORP7* and *PED1*, were predominantly expressed in *T. media*, while only five unigenes, such as *AIM1*, *ORP3*, *AOC2* and *CYP74A*, were predominantly expressed in *T. mairei* (Figure 5C). Thus, the activated JA biosynthetic process may lead to a higher JA level in *T. media* compared with in *T. mairei*.

For the taxoid biosynthetic pathways, the diterpenoid taxane core is derived through the MEP pathway (Eisenreich et al.

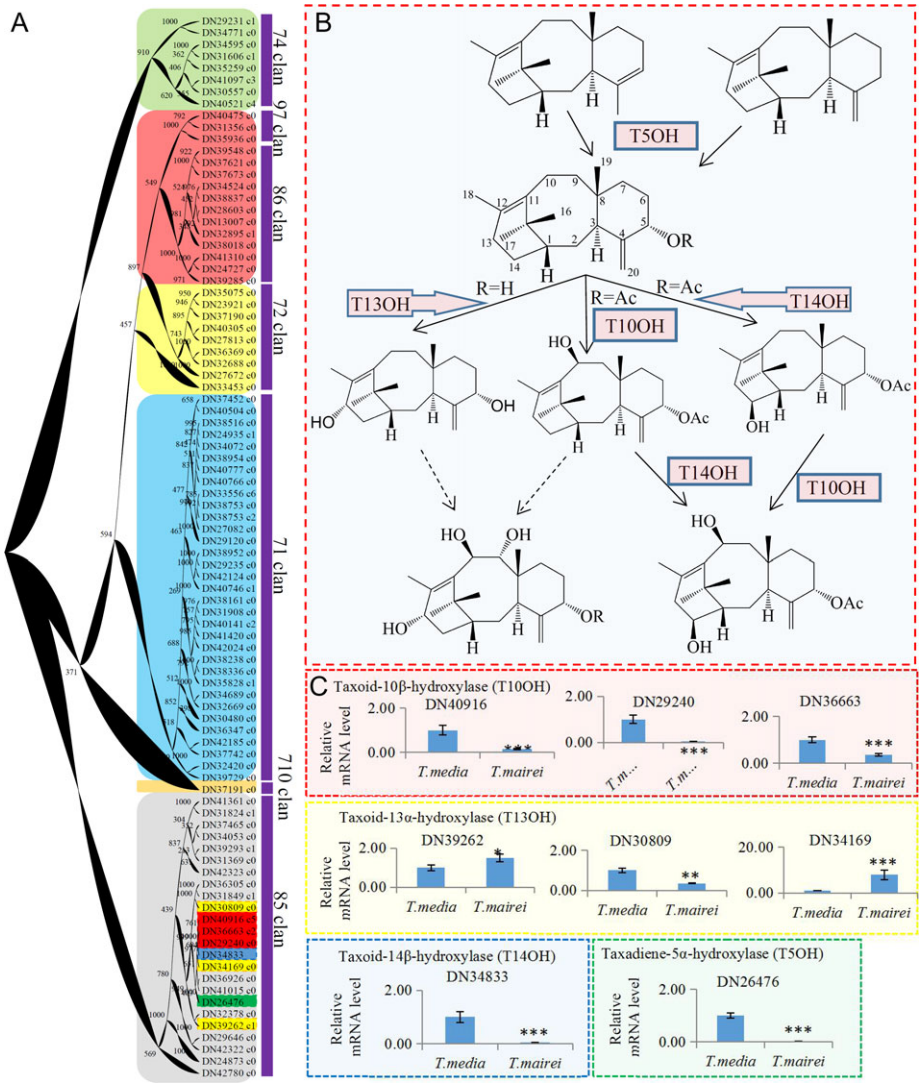

Figure 8. Differential expression of the cytochrome P450 taxoid oxygenase-encoding genes. (A) A phylogenetic tree of 90 P450 candidate genes with full-length sequences. Various background colors indicate different P450 clans. (B) The intermediate oxygenation steps of the Taxol biosynthetic pathway. (C) Differential expression of three T10OH-encoding genes, three T13OH-encoding genes, one T14OH-encoding gene and one T5OH-encoding gene. The abbreviations are: T5OH, taxadiene 5α-hydroxylase; T10OH, taxoid 10β-hydroxylase; T13OH, taxoid 13α-hydroxylases; and T14OH, taxoid 14β-hydroxylase. The significant variations ($P < 0.05$) are indicated by *; the significant variations ($P < 0.01$) are indicated by **; the significant variations ($P < 0.001$) are indicated by ***. Error bars represent mean ± SD ($n = 6$).

1996). The MEP pathway starts with the condensation of thiamin, derived from pyruvate, and D-glyceraldehyde 3-phosphate, which is catalyzed by DXS. The latter is encoded by multiple gene paralogs in most plant species (Vranova et al. 2013). Similar to the model plant *Arabidopsis*, three *DXS*-like genes were identified and named as *DXS1/2/3* in our study (Estevez et al. 2000). Unlike DXS, the other enzymes involved in the MEP pathway in *Arabidopsis* are encoded by single copy genes (Vranova et al. 2013). However, three distinct MDS-encoding genes were identified in our study. DXR has been indicated as controlling a slow step in the synthesis of plastid-derived terpenoids (Mahmoud and Croteau 2001). Among the genes encoding MEP pathway enzymes, most genes, including the DXR-encoding gene, were predominantly expressed in *T. media*, suggesting a more active MEP pathway in *T. media* than in *T. mairei*.

Taxadiene synthase (TS) catalyzes the formation of the taxane nucleus from the branch-point intermediate GGPP (Hezari et al. 1995). The corresponding cDNA sequence of the TS-encoding gene was isolated by a homology-based cloning strategy from the stem of *T. brevifolia* (Wildung and Croteau 1996). The TS reaction is a slow step in the whole pathway (Williams et al. 2000). In our study, only one TS-encoding gene was identified in both *T. media* and *T. mairei*. The expression level of the TS-encoding gene in *T. media* was significantly higher than that in *T. mairei* (Figure 7B). Our data suggested that the relatively high expression level of the TS-encoding gene may be the major cause of the high taxoid accumulation in *T. media*.

After the formation of the taxane skeleton, a series of hydroxylation steps, most of which are mediated by cytochrome P450 in the endoplasmic reticulum, are required on the taxane core (Chau and Croteau 2004, Croteau et al. 2006). Functional studies on these genes led to the identification of taxoid 2α-, 5α-, 13α-, 7β-, 10β- and 14β-hydroxylases (Jennewein et al. 2001, 2003, 2004*a*). An unrooted phylogenetic tree was built based on the complete sequences of 90 cytochrome P450 candidate genes. In our study, three taxoid 10β-hydroxylase-encoding genes and three T13OH-encoding genes were identified, indicating a potential catalytic redundancy in the hydroxylation step of the pathway. An expression analysis showed that most of the P450 oxygenase-encoding genes involved in Taxol biosynthesis were more highly expressed in *T. media*, except for T13OH genes (Figure 8C). A number of predominantly expressed candidate genes in *T. media* may provide a possible explanation for the variation in the taxoid contents of different *Taxus* species.

The expression of several key genes, such as *TS*, *BAPT*, *DBAT* and *DBTNBT*, related to the Taxol biosynthetic pathway was affected by some environmental cues, such as day length, temperature, sunlight and relative humidity (Onrubia et al. 2011, Nasiri et al. 2016). Predominantly expressed *DBAT* and *BAPT* genes in *T. media* may suggest a potential role of environmental cues in the variation in the taxoid contents of two distinct species.

Conclusions

Variations between the taxoid contents of *T. media* and *T. mairei* were determined. The contents of paclitaxel and 10-DAP in *T. media* are significantly higher than those in *T. mairei*. Screening and classification of DEGs showed a significant transcriptional difference between *T. media* and *T. mairei*. Several JA-related DEGs were identified, suggesting a variation in JA metabolism and signaling pathway between *T. media* and *T. mairei*. Furthermore, a large number of genes related to the precursor supply, taxane skeleton formation and hydroxylation, and C13-side chain assembly were also identified. The differential expression of candidate genes involved in taxoid biosynthetic pathways may provide a possible explanation for the variation in the taxoid contents between *T. media* and *T. mairei*.

Supplementary Data

Supplementary Data for this article are available at *Tree Physiology* Online.

Acknowledgments

We are grateful to Kaidong Liu (Life Science and Technology School, Lingnan Normal University) for reading this manuscript. We are also grateful to LC Sciences company for technical support.

Conflict of interest

None declared.

Funding

Our work was funded by the National Key Research and Development Project (2016YFC0503100), Zhejiang Provincial Natural Science Foundation of China under Grant No. LQ14C060001 and National Natural Science Foundation of China (31401935).

References

Bala S, Uniyal GC, Chattopadhyay SK et al. (1999) Analysis of taxol and major taxoids in Himalayan yew, *Taxus wallichiana*. J Chromatogr A 858:239–244.

Bentebibel S, Moyano E, Palazon J, Cusido RM, Bonfill M, Eibl R, Pinol MT (2005) Effects of immobilization by entrapment in alginate and scale-up on paclitaxel and baccatin III production in cell suspension cultures of *Taxus baccata*. Biotechnol Bioeng 89:647–655.

Berthelot K, Estevez Y, Deffieux A, Peruch F (2012) Isopentenyl diphosphate isomerase: a checkpoint to isoprenoid biosynthesis. Biochimie 94:1621–1634.

Chau M, Croteau R (2004) Molecular cloning and characterization of a cytochrome P450 taxoid 2alpha-hydroxylase involved in Taxol biosynthesis. Arch Biochem Biophys 427:48–57.

Chau M, Jennewein S, Walker K, Croteau R (2004) Taxol biosynthesis: molecular cloning and characterization of a cytochrome P450 taxoid 7 beta-hydroxylase. Chem Biol 11:663–672.

Croteau R, Ketchum RE, Long RM, Kaspera R, Wildung MR (2006) Taxol biosynthesis and molecular genetics. Phytochem Rev 5:75–97.

Cusido RM, Onrubia M, Sabater-Jara AB, Moyano E, Bonfill M, Goossens A, Angeles Pedreno M, Palazon J (2014) A rational approach to improving the biotechnological production of taxanes in plant cell cultures of *Taxus* spp. Biotechnol Adv 32:1157–1167.

Eisenreich W, Menhard B, Hylands PJ, Zenk MH, Bacher A (1996) Studies on the biosynthesis of taxol: the taxane carbon skeleton is not of mevalonoid origin. Proc Natl Acad Sci USA 93:6431–6436.

Estevez JM, Cantero A, Romero C, Kawaide H, Jimenez LF, Kuzuyama T, Seto H, Kamiya Y, Leon P (2000) Analysis of the expression of CLA1, a gene that encodes the 1-deoxyxylulose 5-phosphate synthase of the 2-C-methyl-D-erythritol-4-phosphate pathway in *Arabidopsis*. Plant Physiol 124:95–104.

Giner JL, Faraldos JA (2003) Facile orthoester formation in a model Compound of the Taxol oxetane: are biologically active epoxy esters, orthoesters, and oxetanyl esters latent electrophiles? Helv Chim Acta 86:3613–3622.

Grabherr MG, Haas BJ, Yassour M et al. (2011) Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat Biotechnol 29:644–652.

Hao DC, Ge G, Xiao P, Zhang Y, Yang L (2011) The first insight into the tissue specific *Taxus* transcriptome via illumina second generation Sequencing. PLoS One 6:e21220.

Harrison JW, Scrowston RM, Lythgoe B, Harrison JW, Scrowston RM, Lythgoe B (1966) Taxine. Part IV. The constitution of taxine-I. J Chem Soc C, 1933–1945.

Hezari M, Lewis NG, Croteau R (1995) Purification and characterization of taxa-4(5),11(12)-diene synthase from Pacific yew (*Taxus brevifolia*) that catalyzes the first committed step of taxol biosynthesis. Arch Biochem Biophys 322:437–444.

Howat S, Park B, Oh IS, Jin YW, Lee EK, Loake GJ (2014) Paclitaxel: biosynthesis, production and future prospects. Nat Biotechnol 31: 242–245.

Jennewein S, Rithner CD, Williams RM, Croteau RB (2001) Taxol biosynthesis: taxane 13α-hydroxylase is a cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci USA 98:13595–13600.

Jennewein S, Rithner CD, Williams RM, Croteau R (2003) Taxoid metabolism: taxoid 14beta-hydroxylase is a cytochrome P450-dependent monooxygenase. Arch Biochem Biophys 413:262–270.

Jennewein S, Long RM, Williams RM, Croteau R (2004*a*) Cytochrome p450 taxadiene 5alpha-hydroxylase, a mechanistically unusual monooxygenase catalyzing the first oxygenation step of taxol biosynthesis. Chem Biol 11:379–387.

Jennewein S, Wildung MR, Chau M, Walker K, Croteau R (2004*b*) Random sequencing of an induced *Taxus* cell cDNA library for identification of clones involved in Taxol biosynthesis. Proc Natl Acad Sci USA 101:9149–9154.

Jiménez-Barbero J, Amat-Guerri F, Snyder JP (2002) The solid state, solution and tubulin-bound conformations of agents that promote microtubule stabilization. Curr Med Chem Anticancer Agents 2:91–122.

Kai G, Zhao L, Zhang L, Li Z, Guo B, Zhao D, Sun X, Miao Z, Tang K (2005) Characterization and expression profile analysis of a new cDNA encoding taxadiene synthase from *Taxus media*. J Biochem Mol Biol 38:668–675.

Kaspera R, Croteau R (2006) Cytochrome P450 oxygenases of Taxol biosynthesis. Phytochem Rev 5:433–444.

Ketchum REB, Luong JV, Gibson DM (1999) Efficient extraction of paclitaxel and related taxoids from leaf tissue of *taxus* using a potable solvent system. J Liq Chromatogr Relat Technol 22:1715–1732.

Kingston DGI (1995) Recent advances in the chemistry and structure-activity relationships of paclitaxel. In: ACS Symposium Series (USA). Oxford University Press, Oxford, pp 203–216.

Kohler DR, Goldspiel BR (1994) Paclitaxel (taxol). Pharmacotherapy 14: 3–34.

Lee E-K, Jin Y-W, Park JH et al. (2010) Cultured cambial meristematic cells as a source of plant natural products. Nat Biotech 28:1213–1217.

Lenka SK, Nims NE, Vongpaseuth K, Boshar RA, Roberts SC, Walker EL (2015) Jasmonate-responsive expression of paclitaxel biosynthesis genes in *Taxus cuspidata* cultured cells is negatively regulated by the bHLH transcription factors TcJAMYC1, TcJAMYC2, and TcJAMYC4. Front Plant Sci 6:115.

Li S, Fu Y, Zu Y, Zu B, Wang Y, Efferth T (2009) Determination of paclitaxel and its analogues in the needles of *Taxus* species by using negative pressure cavitation extraction followed by HPLC-MS-MS. J Sep Sci 32:3958–3966.

Li ST, Zhang P, Zhang M, Fu CH, Zhao CF, Dong YS, Guo AY, Yu LJ (2012) Transcriptional profile of *Taxus chinensis* cells in response to methyl jasmonate. BMC Genomics 13:295.

Mahmoud SS, Croteau RB (2001) Metabolic engineering of essential oil yield and composition in mint by altering expression of deoxyxylulose phosphate reductoisomerase and menthofuran synthase. Proc Natl Acad Sci USA 98:8915–8920.

Mortazavi A, Williams BA, Mccue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods 5:621.

Nasiri J, Naghavi MR, Alizadeh H, Moghadam MR (2016) Seasonal-based temporal changes fluctuate expression patterns of *TXS*, *DBAT*, *BAPT* and *DBTNBT* genes alongside production of associated taxanes in *Taxus baccata*. Plant Cell Rep 35:1103–1119.

Nystedt B, Street NR, Wetterbom A et al. (2013) The Norway spruce genome sequence and conifer genome evolution. Nature 497:579–584.

Onrubia M, Moyano E, Bonfill M, Palazon J, Goossens A, Cusido RM (2011) The relationship between *TXS*, *DBAT*, *BAPT* and *DBTNBT* gene expression and taxane production during the development of *Taxus baccata* plantlets. Plant Sci 181:282–287.

Patil RA, Lenka SK, Normanly J, Walker EL, Roberts SC (2014) Methyl jasmonate represses growth and affects cell cycle progression in cultured *Taxus* cells. Plant Cell Rep 33:1479–1492.

Rasool S, Mohamed R (2016) Plant cytochrome P450s: nomenclature and involvement in natural product biosynthesis. Protoplasma 253: 1197–1209.

Rohdich F, Hecht S, Gartner K, Adam P, Krieger C, Amslinger S, Arigoni D, Bacher A, Eisenreich W (2002) Studies on the nonmevalonate terpene biosynthetic pathway: metabolic role of IspH (LytB) protein. Proc Natl Acad Sci USA 99:1158–1163.

Rowinsky EK, Eisenhauer EA, Chaudhry V, Arbuck SG, Donehower RC (1993) Clinical toxicities encountered with paclitaxel (Taxol). Semin Oncol 20:1–15.

Sabater-Jara AB, Onrubia M, Moyano E, Bonfill M, Palazon J, Pedreno MA, Cusido RM (2014) Synergistic effect of cyclodextrins and methyl jasmonate on taxane production in *Taxus x media* cell cultures. Plant Biotechnol J 12:1075–1084.

Saitou N, Nei M (1987) The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol Biol Evol 4:406–425.

Shen C, Yue R, Sun T, Zhang L, Xu L, Tie S, Wang H, Yang Y (2015) Genome-wide identification and expression analysis of auxin response factor gene family in *Medicago truncatula*. Front Plant Sci 6:73.

Shen C, Xue J, Sun T, Guo H, Zhang L, Meng Y, Wang H (2016*a*) Succinyl-proteome profiling of a high taxol containing hybrid *Taxus* species (*Taxus x media*) revealed involvement of succinylation in multiple metabolic pathways. Sci Rep 6:21764.

Shen C, Yang Y, Liu K, Zhang L, Guo H, Sun T, Wang H (2016*b*) Involvement of endogenous salicylic acid in iron-deficiency responses in *Arabidopsis*. J Exp Bot 67:4179–4193.

Sun G, Yang Y, Xie F, Wen J-F, Wu J, Wilson IW, Tang Q, Liu H, Qiu D (2013) Deep sequencing reveals transcriptome re-Programming of *Taxus* × *media* cells to the elicitation with methyl jasmonate. PLoS One 8:e62865.

van Rozendaal EL, Lelyveld GP, van Beek TA (2000) Screening of the needles of different yew species and cultivars for paclitaxel and related taxoids. Phytochemistry 53:383–389.

Vranova E, Coman D, Gruissem W (2013) Network analysis of the MVA and MEP pathways for isoprenoid synthesis. Annu Rev Plant Biol 64: 665–700.

Walker K, Croteau R (2000) Taxol biosynthesis: molecular cloning of a benzoyl-CoA:taxane 2alpha-O-benzoyltransferase cDNA from *Taxus* and functional expression in *Escherichia coli*. Proc Natl Acad Sci USA 97:13591–13596.

Wang YF, Shi QW, Dong M, Kiyota H, Gu YC, Cong B (2011) Natural taxanes: developments since 1828. Chem Rev 111:7652–7709.

Wildung MR, Croteau R (1996) A cDNA clone for taxadiene synthase, the diterpene cyclase that catalyzes the committed step of taxol biosynthesis. J Biol Chem 271:9201–9204.

Williams DC, Wildung MR, Jin AQ, Dalal D, Oliver JS, Coates RM, Croteau R (2000) Heterologous expression and characterization of a 'Pseudomature' form of taxadiene synthase involved in paclitaxel (Taxol) biosynthesis and evaluation of a potential intermediate and inhibitors of the multistep diterpene cyclization reaction. Arch Biochem Biophys 379:137–146.

Wu Q, Sun C, Luo H, Li Y, Niu Y, Sun Y, Lu A, Chen S (2011) Transcriptome analysis of *Taxus cuspidata* needles based on 454 pyrosequencing. Planta Med 77:394–400.

Xiao Z, Itokawa H, Lee KH (2004) Total synthesis of taxoids. ChemInform 35. doi:10.1002/chin.200415239.

Yukimune Y, Tabata H, Higashi Y, Hara Y (1996) Methyl jasmonate-induced overproduction of paclitaxel and baccatin III in *Taxus* cell suspension cultures. Nat Biotechnol 14:1129–1132.