the plant journal

The Plant Journal (2020) doi: 10.1111/tpj.14710

Tissue-specific study across the stem of *Taxus media* identifies a phloem-specific TmMYB3 involved in the transcriptional regulation of paclitaxel biosynthesis

Chunna Yu[1,2], Xiujun Luo[1,2], Chengchao Zhang[1,2], Xinyun Xu[1,2], Jiefang Huang[1,2], Yueyue Chen[1,2], Shangguo Feng[1,2], Xiaori Zhan[1,2], Lei Zhang[3], Huwei Yuan[4,5], Bingsong Zheng[4,5], Huizhong Wang[1,2] and Chenjia Shen[1,2,*]

[1]*College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou 311121, China,*

[2]*Zhejiang Provincial Key Laboratory for Genetic Improvement and Quality Control of Medicinal Plants, Hangzhou Normal University, Hangzhou 311121, China,*

[3]*Department of Plant Pathology, Washington State University, Pullman, WA 99164-6430, USA,*

[4]*State Key Laboratory of Subtropical Silviculture, Zhejiang A & F University, Hangzhou 311300, China, and*

[5]*Center for Cultivation of Subtropical Forest Resources (CCSFR), Zhejiang A & F University, Hangzhou 311300, China*

Received 10 September 2019; revised 6 January 2020; accepted 22 January 2020.
*For correspondence (e-mail shencj@hznu.edu.cn).

SUMMARY

Taxus **stem barks can be used for extraction of paclitaxel. However, the composition of taxoids across the whole stem and the stem tissue-specificity of paclitaxel biosynthesis-related enzymes remain largely unknown. We used cultivated** ***Taxus media*** **trees for analyses of the chemical composition and protein of major stem tissues by an integrated metabolomic and proteomic approach, and the role of TmMYB3 in paclitaxel biosynthesis was investigated. The metabolomic landscape analysis showed differences in stem tissue-specific accumulation of metabolites. Phytochemical analysis revealed that there is high accumulation of paclitaxel in the phloem. Ten key enzymes involved in paclitaxel biosynthesis were identified, most of which are predominantly produced in the phloem. The full-length sequence of** ***TmMYB3*** **and partial promoter sequences of five paclitaxel biosynthesis-related genes were isolated. Several MYB recognition elements were found in the promoters of** ***TBT***, ***DBTNBT*** **and** ***TS***. **Further** ***in vitro*** **and** ***in vivo*** **investigations indicated that TmMYB3 is involved in paclitaxel biosynthesis by activating the expression of** ***TBT*** **and** ***TS***. **Differences in the taxoid composition of different stem tissues suggest that the whole stem of** ***T. media*** **has potential for biotechnological applications. Phloem-specific TmMYB3 plays a role in the transcriptional regulation of paclitaxel biosynthesis, and may explain the phloem-specific accumulation of paclitaxel.**

Keywords: metabolome, phloem, proteome, R2R3-MYB, paclitaxel biosynthesis, ***Taxus media*****.**

INTRODUCTION

Paclitaxel is one of the most important chemotherapeutic agents, and its efficacy against ovarian, breast and head and neck cancers has been widely documented (Bernabeu *et al.*, 2017). The importance of *Taxus* trees as natural sources of paclitaxel and its derivatives has been highlighted in numerous biological, phytochemical and omics studies (Hao *et al.*, 2012). The commercial value of paclitaxel has driven considerable research interest in *Taxus* trees in recent decades (Sanchez-Munoz *et al.*, 2018).

The biosynthesis of paclitaxel, a structurally complex representative of approximately 400 identified taxoids isolated from *Taxus* trees, can be conceptually divided into 19 steps, from a diterpenoid progenitor to the final product (Croteau *et al.*, 2006). First, the 2-*C*-methyl-D-erythritol 4-phosphate (MEP) pathway yields isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP) to synthesize the universal diterpenoid precursor geranylgeranyl diphosphate (GGPP) (Eisenreich *et al.*, 1996). Subsequently, formation of the taxane skeleton involves cyclization of GGPP into taxa-4(5),11(12)-diene, which is catalyzed by a slow-starter enzyme, taxadiene synthase (TS) (Koepp *et al.*, 1995). To obtain a key intermediate 10-deacetylbaccatin III (10-DAB), a number of functional groups are added to the taxane core by a series of cytochrome P450-mediated hydroxylations, CoA-dependent acyl transfers and a C9 site oxidation (Kaspera and Croteau, 2006). One step of these reactions is conducted by a taxane 2α-*O*-benzoyltransferase (TBT), which transforms a 2-debenzoyl taxoid-type intermediate to 10-DAB (Walker and Croteau, 2000).

Next, a C13 side chain is appended to baccatin III (BAC), yielding 3′-*N*-debenzoyl-2′-deoxytaxol (Kaspera and Croteau, 2006). Finally, the formation of the functional paclitaxel molecule involves an important enzyme, 3′-*N*-debenzoyl-2′-deoxytaxol-*N*-benzoyltransferase (DBTNBT) (Onrubia *et al.*, 2011).

Omics studies on *Taxus* trees have identified a large number of metabolites and proteins involved in the paclitaxel biosynthesis pathway (Yu *et al.*, 2017). For example, six taxoids were isolated from cultured *Taxus* seedlings using a metabolomic approach (Tanaka *et al.*, 2011). Comparative metabolomics revealed that most of the intermediates of paclitaxel biosynthesis varied hugely in distinct *Taxus* species, providing a possible explanation for their differential accumulation of paclitaxel (Yu *et al.*, 2018). Comparative proteomic analyses identified four MEP pathway-related enzymes and three cytochrome P450 taxoid oxygenases that play important roles in determining interspecific differences in taxoid accumulation (Hao *et al.*, 2017). An integrated metabolomic and proteomic analysis indicated that 1-deoxy-D-xylulose 5-phosphate reductoisomerase (DXR) and 1-hydroxy-2-methyl-2-(*E*)-butenyl 4-diphosphate reductase (HDR), two key enzymes in the MEP pathway, are responsible for the stimulation of paclitaxel production (Zheng *et al.*, 2016).

Recently, numerous transcription factors (TFs) have been reported to play crucial roles in the transcriptional regulation of different paclitaxel biosynthesis-related genes (Kuang *et al.*, 2019). Among these TFs, MYC family members are believed to be involved in the regulation of jasmonic acid (JA) signal transduction, which is involved in paclitaxel biosynthesis (Cui *et al.*, 2019). Three basic helix–loop–helix (bHLH) TFs, TcJAMYC1, TcJAMYC2 and TcJAMYC4, negatively regulate the expression of paclitaxel biosynthetic genes in *Taxus cuspidata* (Lenka *et al.*, 2015). Another study showed that TcMYC2a may regulate the expression of *TS*, *TAT*, *DBTNBT*, *T13OH* and *T5OH* via ethylene responsive factor (ERF) regulators that depend on the JA signaling pathway in *Taxus chinensis* (Zhang *et al.*, 2018b). Also in *T. chinensis*, ERF12 and ERF15, act as repressor and activator of paclitaxel biosynthesis, respectively, by binding to the GCC-box in the JA-responsive element of the *TS* promoter (Zhang *et al.*, 2015). In *T. chinensis*, a WRKY transcription factor, TcWRKY1, participates in the transcriptional activation of the promoter of 10-deacetylbaccatin-III-10-β-*O*-acetyltransferase (*DBAT*) (Li *et al.*, 2013). Another two WRKY TFs, TcWRKY8 and TcWRKY47, significantly increased the expression levels of several paclitaxel biosynthesis-related genes (Zhang *et al.*, 2018a).

The plant MYB family consists of four subfamilies, 1R-, R2R3-, R1R2R3- and 4R-MYB, which have different numbers of MYB domain repeats (Dubos *et al.*, 2010). In particular, R2R3-MYB is a large subfamily, members of which are involved in secondary metabolism in plants (Stracke *et al.*, 2001). In poplar, MYB165 and MYB194 are broad repressors of the flavonoid and phenylpropanoid biosynthetic pathways (Ma *et al.*, 2018). In peach, activator-type R2R3-MYB genes balance the accumulation anthocyanin and pro-anthocyanidin by inducing a repressor-type R2R3-MYB gene *PpMYB18* (Zhou *et al.*, 2019a). However, the role of R2R3-MYB TFs in the regulation of paclitaxel biosynthesis is largely unknown.

Stem bark tissues are major sources of medicinal compounds, such as lignanoids from the bark of *Eucommia ulmoides* and taxoids from the bark of *Taxus* species (Uniyal, 2013; Wang *et al.*, 2019). First isolated from the bark of the Pacific yew *Taxus brevifolia*, hundreds of taxoids have since been identified in the stem bark of various *Taxus* species (Chan *et al.*, 1994; Shen *et al.*, 2001; Nadeem *et al.*, 2002). Although the whole *Taxus* stem is rich in secondary metabolites, the composition of taxoids across distinct stem tissues remains poorly understood. Studies of the chemical composition of the major tissues of the whole stem will give us a good opportunity to explore the stem tissue-specificity of taxoid accumulation. Herein, molecular fingerprints were determined for four major tissue types in the stem of *T. media*, namely the cortex, phloem, xylem and pith. The results reveal the metabolite and protein profiles of different stem fractions based on integrated metabolomic and proteomic analyses. Furthermore, a phloem-specific R2R3-MYB TF, TmMYB3, was identified, and its function in the regulation of paclitaxel biosynthesis was investigated. Our findings may help to elucidate the regulatory mechanism underlying paclitaxel biosynthesis and accelerate the breeding of species with a high paclitaxel yield.

RESULTS

Localization of taxoid accumulation

The stem of *Taxus* species, especially the stem bark, is widely used for paclitaxel extraction (Shen *et al.*, 2001; Nguyen *et al.*, 2003). Four major tissues across the stem of *T. media*, the cortex, phloem, xylem and pith, were separated by simply peeling. A cross-section of the stem is shown in Figure 1(a). Quantification analysis was performed to explore the variation of taxoid content among the four major stem tissues. Four important taxoids, including paclitaxel, 10-deacetylpaclitaxel (DAP), BAC and 10-DAB, were quantified by ultra-performance liquid chromatography – tandem mass spectrometry (UPLC-MS/MS) and their total ion chromatography (TIC) chromatogram is shown in Figure 1(b). Most taxoids accumulated to varying levels in these four stem tissues. Specifically, both paclitaxel and 10-DAB accumulated to the highest levels in the phloem, while BAC predominantly accumulated in the pith and DAP greatly accumulated in the cortex and phloem

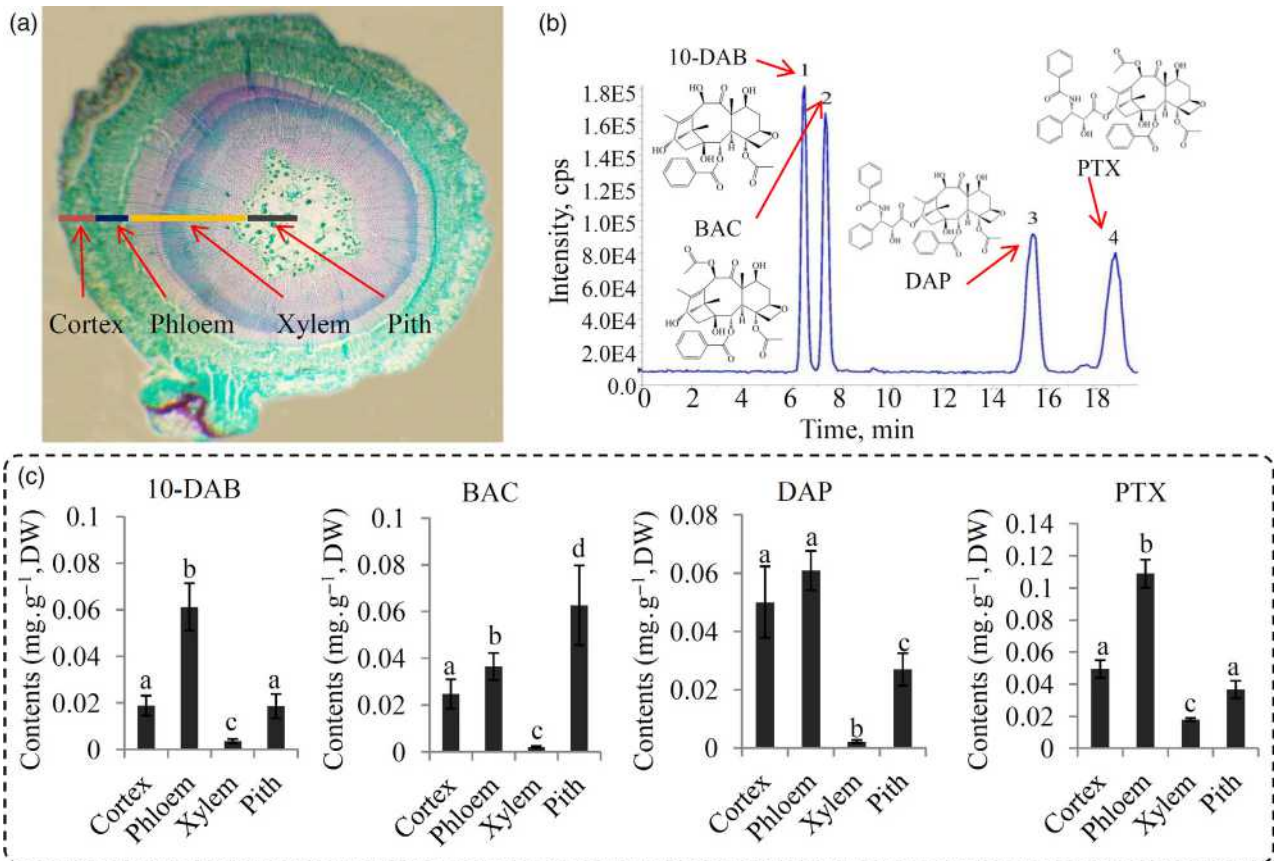

Figure 1. Microstructure of the stem and differential accumulation of four representative taxoids across the stem tissues of *Taxus media*.
(a) Paraffin section of the stem. Four major stem tissues, including cortex, phloem, xylem and pith, were clearly identified.
(b) A representative total ion current chromatogram of four taxoids: peak 1, 10-deacetylbaccatin III (10-DAB); peak 2, baccatin III (BAC); peak 3, 10-deacetylpaclitaxel (DAP); peak 4, paclitaxel (PTX). cps, counts per second.
(c) Corresponding histograms indicated the differences in the four representative taxoids among different stem tissues. Different letters indicated significant differences in the contents of four representative taxoids from different stem tissues ($P < 0.05$). DW, dry weight.

tissues. The xylem contained the lowest levels of these four taxoids (Figure 1c).

Overview of the metabolomes of different stem tissues

To explore the variation in phytochemical composition of different stem tissues, an untargeted metabolomic analysis ($n = 10$) was performed, and 9512 ion features corresponding to 5162 potential metabolites with annotations were identified (Table S2 in the online Supporting Information). Analysis of five checking parameters, including TIC, average *m/z* distribution, retention-time width, metabolite intensity distribution and coefficient of variation distribution indicated effective sample preparation and high-quality raw data (Figures 2a and S1). Principal component analysis (PCA) showed that PC1 and PC2 explained 39.3% and 15.51% of variation, respectively. Metabolite profiling of the four stem tissues separated their metabolomes into three groups (Figure 2b). Metabolomes from the cortex and phloem were grouped into the same cluster, indicating high similarity between these tissues (Figure 2c). A total of 5077 annotated metabolites were mapped to different known metabolic pathways (Table S3). The largest number of metabolites belonged to the category 'metabolism of terpenoids and polyketides' (275 metabolites), followed by 'amino acid metabolism' (239 metabolites) and 'lipid metabolism' (224 metabolites).

Analysis of differentially accumulated metabolites (DAMs) in different stem tissues

To analyze stem tissue-specific accumulated metabolites, DAMs were clustered into four major clusters (I–IV). Metabolites belonging to Cluster I predominantly accumulated in the cortex; metabolites grouped into Cluster II were highly accumulated in the phloem; metabolites in Cluster III were abundant in the pith; and metabolites belonging to Cluster IV predominantly accumulated in both the cortex and phloem (Figure 2d). In Cluster I, the major DAMs were sugars, flavonoids, amino acids and alkaloids; in Cluster II, the largest number of DAMs belonged to amino acids, flavonoids and terpenoids; in Cluster III,

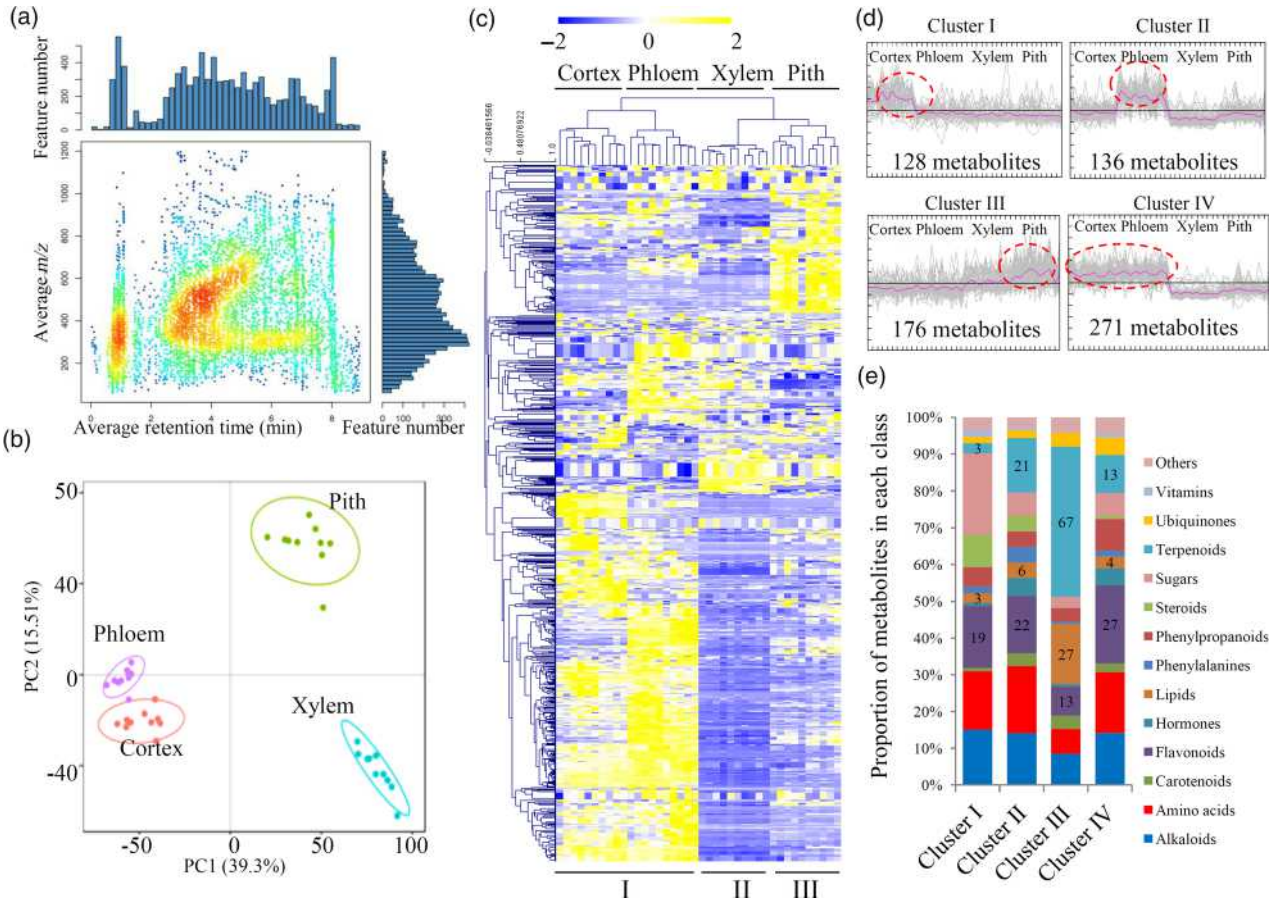

Figure 2. Untargeted metabolomes identified differentially accumulated metabolites among the four stem tissues.
(a) Quality control parameters, including *m/z* widths and retention-time widths of the metabolomes.
(b) A principal component (PC) analysis of the metabolomes from four stem tissues of *Taxus media*.
(c) A heatmap of the abundance of metabolites in the four stem tissues (n = 10). The heatmap scale ranges from −2 to +2 on a $\log_2$ scale.
(d) MeV cluster analysis of tissue-specific accumulated metabolites from the untargeted metabolomic profiles. The red ovals indicated stem tissue-specific accumulated metabolites. Cluster I indicates the cortex-specific accumulated metabolites, Cluster II indicates the phloem-specific accumulated metabolites, Cluster III indicates the pith-specific accumulated metabolites and Cluster IV indicates the metabolites predominantly accumulated in both the cortex and phloem.
(e) Proportions of metabolites in each class.

terpenoids, lipids and alkaloids were the main metabolites; and in Cluster IV, the predominant metabolites were flavonoids, amino acids and alkaloids (Figure 2e).

Overview of the proteomes of different stem tissues

To explore stem tissue-specific proteins, protein samples (n = 3) from the four stem tissues were isolated for the quantification of dynamic changes in their proteomes (Figure S2). The distribution of mass errors and peptide lengths indicated a high mass accuracy for the mass spectrometry data. After searching the database, the number of effectively matched spectra was 132 730, and the utilization rate of spectra was 17.42%. Based on the matched spectra, 56 902 peptides were identified, including 47 396 unique peptides (Figure S3). After filtering, 5604 proteins were annotated, among which 4461 were quantified (Table S4). To assess statistical consistency, pairwise Pearson's correlation coefficients, PCA and relative standard deviation (RSD) were investigated using all quantified proteins, suggesting that three samples from each group displayed good repeatability (Figure S4).

Gene Ontology (GO) annotation, Eukaryotic Orthologous Groups (KOG) classification and subcellular localization prediction were performed on all identified proteins. The most representative GO terms were 'catalytic activity' (2273 proteins), 'metabolic process' (2194 proteins), 'binding' (2031 proteins), 'cellular process' (1433 proteins) and 'single-organism process' (1344 proteins); the most abundant KOG terms were 'general function prediction only' (600 proteins), 'post-translational modification' (599 proteins) and 'energy production and conversion' (378 proteins); the dominant subcellular localization categories were 'cytoplasm' (1879 proteins), 'chloroplast' (1825 proteins) and 'nucleus' (999 proteins) (Table S4).

Analysis of differentially produced proteins (DPPs) in different stem tissues

The numbers of up- and downregulated proteins from different comparisons (cortex/xylem, cortex/phloem, cortex/pith, xylem/pith, phloem/xylem and phloem/pith) are shown in Figure S5. Furthermore, proteomic profiling of the four stem tissues clustered tissue-specific proteins into four groups (I–IV; Figure 3a). In total, 329 cortex-, 307 phloem-, 94 xylem- and 90 pith-specific proteins were identified (Figure 3b). Further studies were carried out on phloem-specific produced proteins (PPPs) to investigate their relationship with phloem-specific accumulation of paclitaxel (Table S5). Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis classified 307 PPPs into five significantly enriched KEGG pathways. Interestingly, the 'terpenoid backbone biosynthesis' pathway, which provides precursors for paclitaxel biosynthesis, was the most significantly enriched KEGG pathway (Figure 3c). The GO enrichment analysis of DPPs showed that 'oxidoreductase activity' was the dominant term in the molecular function category, and 'isoprenoid biosynthetic process' was the major biological process subcategory (Figure 3d).

Proteomic analysis identified nine classical TFs, namely MYB3, NAC25, GATA9-like, TCP2, TFIIB, bHLH82, WRKY31,

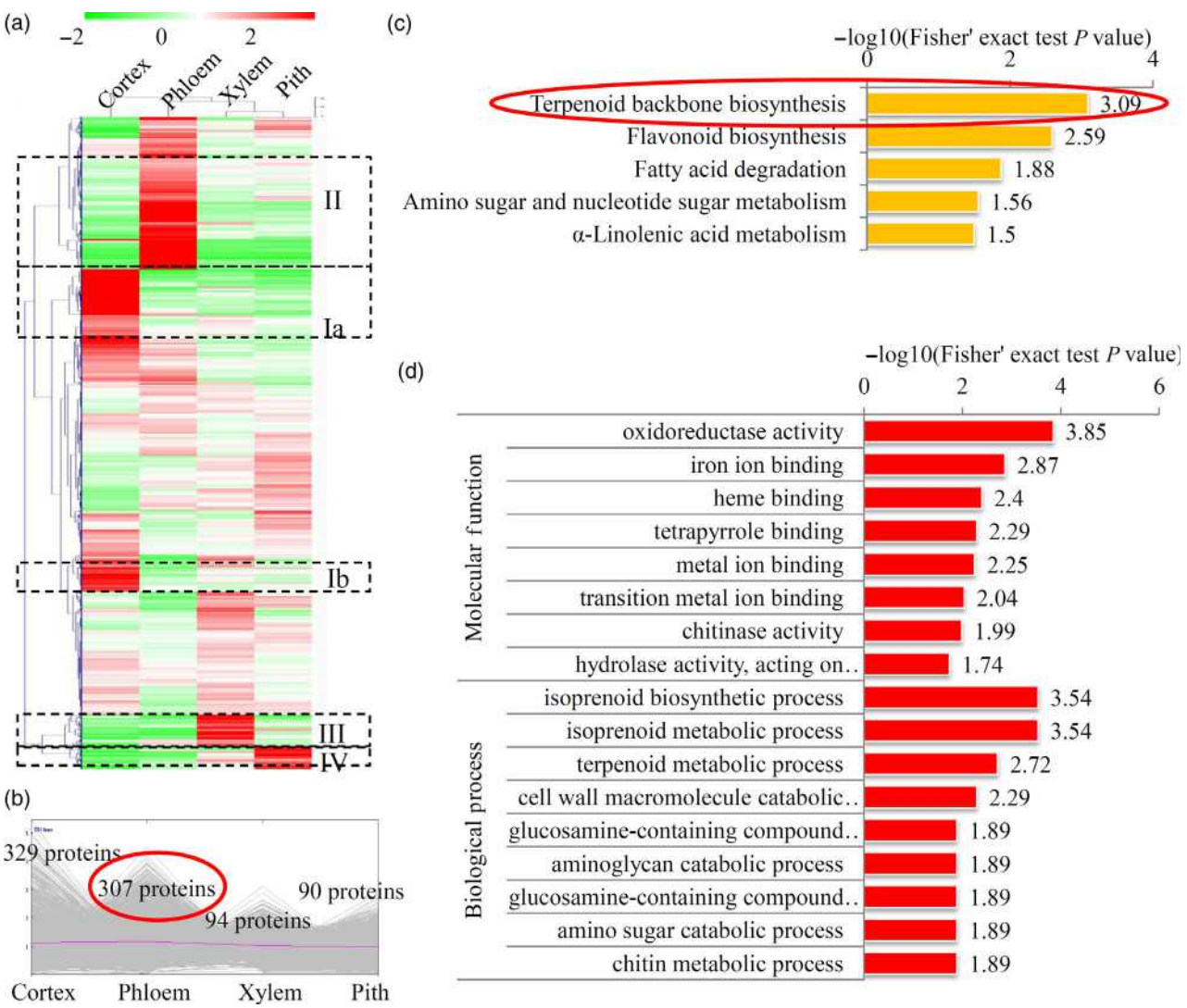

Figure 3. Analysis of differentially produced proteins (DPPs) among the four stem tissues.
(a) A heatmap of the abundance of proteins in the four stem tissues ($n = 3$). The heatmap scale ranges from −2 to +2 on a $\log_2$ scale. The black dotted line frames indicated stem tissue-specific produced proteins.
(b) MeV cluster analysis showing the numbers of tissue-specific produced proteins. The red oval indicates the number of phloem-specific produced proteins (PPPs).
(c) Kyoto Encyclopedia of Genes and Genomes enrichment analysis of the PPPs.
(d) Gene Ontology enrichment analysis of the PPPs.

DIVARICATA and TFIID (Table S6). Among these TFs, MYB3, NAC25 and GATA9-like were predominantly produced in the phloem.

Integrated metabolomic and proteomic analysis of taxoid metabolic pathways in *T. media*

Paclitaxel biosynthesis is a sophisticated metabolic pathway that involves a series of intermediate metabolites and enzymes (Croteau *et al.*, 2006). In our present study, 12 intermediate metabolites involved in paclitaxel biosynthesis were identified, among which taxa-4(20),11(12)-dien-5α-yl acetate, 10-DAB, 3′-*N*-debenzoyl-2′-deoxytaxol, 3′-*N*-debenzoyltaxol and paclitaxel predominantly accumulated in the phloem, while taxa-4(5),11(12)-diene and taxa-4(20),11(12)-dien-5α,13α-diol accumulated at high levels in the xylem, and taxa-4(20),11(12)-dien-5α-ol, 10β,14β-dihydroxytaxa-4(20),11(12)-dien-5α-yl acetate and BAC significantly accumulated in the pith (Figure 4a). Additionally, 10 paclitaxel biosynthesis-related enzymes were identified. The precise positions of the identified intermediate metabolites and key enzymes involved in paclitaxel biosynthesis are shown in Figure 4(b). Interestingly, most of the paclitaxel-related enzymes, such as geranylgeranyl diphosphate synthase (GGPPS), TS, taxadiene 5α-hydroxylase (T5αOH), taxadien-5α-ol O-acetyltransferase. (TAT), taxane 10β-hydroxylase (T10βOH), TBT, DBAT, baccatin III:3-amino-3-phenylpropanoyltransferase (BAPT) and DBTNBT, were predominantly produced in the phloem (Figure 4c).

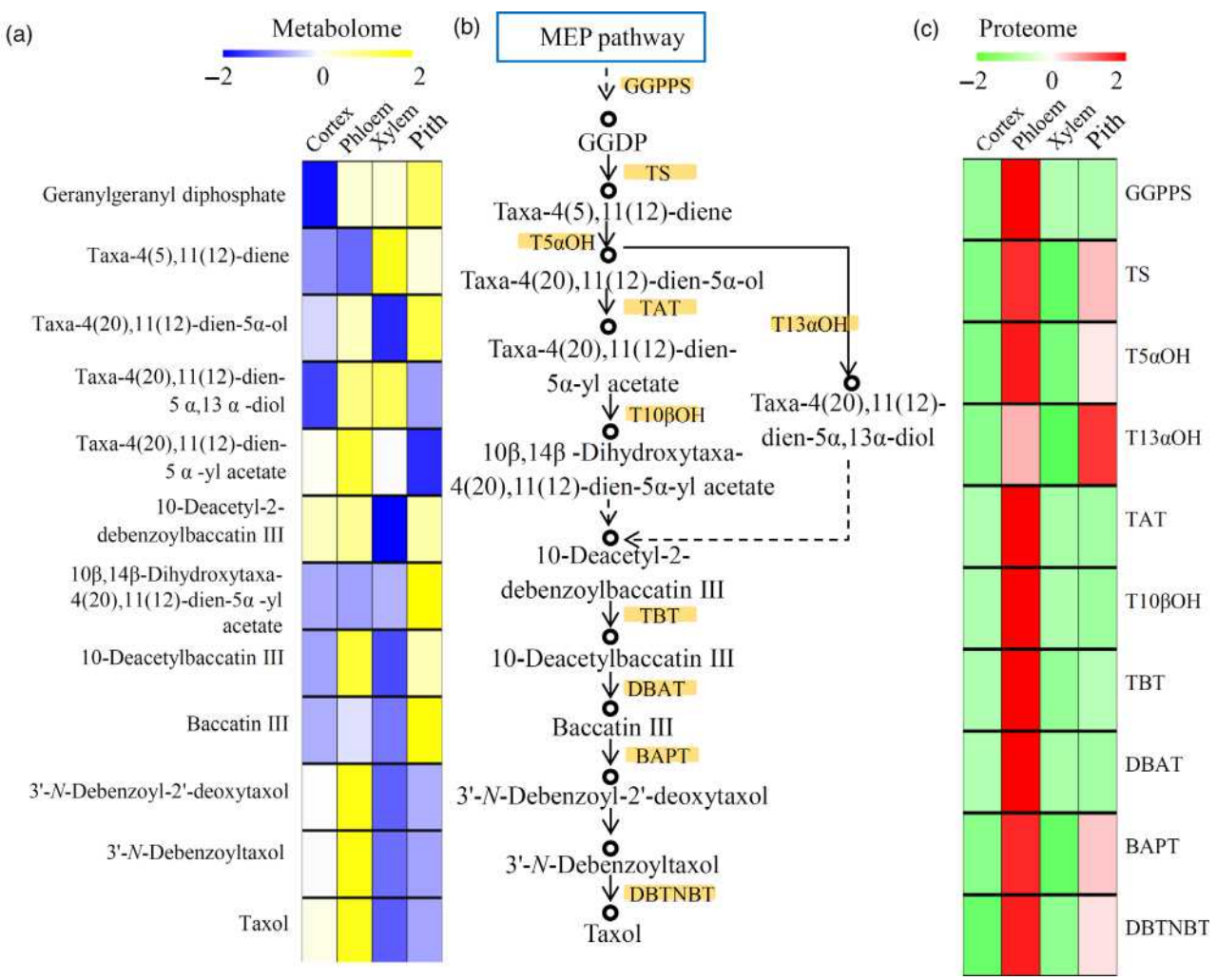

Figure 4. Integrated metabolomic and proteomic analysis of the paclitaxel biosynthesis pathway.
(a) Differential accumulation of 12 intermediate metabolites involved in paclitaxel biosynthesis from the four major stem tissues.
(b) Overview of the paclitaxel biosynthesis pathway. The orange background indicates the enzymes identified by the proteome. BAPT, baccatin III:3-amino-3-phenylpropanoyltransferase; DBAT, 10-deacetylbaccatin-III-10-β-O-acetyltransferase; DBTNBT, 3′-*N*-debenzoyltaxol *N*-benzoyltransferase; GGPPS, geranylgeranyl diphosphate synthase; MEP, 2-C-methyl-d-erythritol 4-phosphate; T10βOH, taxane 10β-hydroxylase; T13αOH, taxane 13α-hydroxylase; T5αOH, taxadiene-5α-hydroxylase; TAT, taxadien-5α-ol O-acetyltransferase; TBT, taxane 2α-O-benzoyltransferase; TS, taxadiene synthase.
(c) Differential production of 10 key enzymes involved in paclitaxel biosynthesis from the four major stem tissues. The heatmap scale ranges from −2 to +2 on a log2 scale.

There are a number of dead-end metabolites and 14-hydroxylated taxoids in *Taxus* trees; however, none of these intermediate metabolites leads to paclitaxel biosynthesis (Wang *et al.*, 2011). Herein, four taxusin-like metabolites, (+)-taxusin, 2α-hydroxytaxusin, 7β-hydroxytaxusin and 2α,7β-dihydroxytaxusin, and one 14-hydroxylated taxoid, taxuyunnanin C, were identified (Figure 5a). All four taxusin-like metabolites accumulated in the pith, and taxuyunnanin C accumulated in both the cortex and phloem (Figure 5b). Two important enzymes, taxadiene-2α-hydroxylase (T2αOH) and taxane-7β-hydroxylase (T7βOH), involved in the metabolic pathway of taxusin-like metabolites were produced at high levels in the pith. Taxoid 14β-hydroxylase (T14βOH), catalyzing an intermediate step to synthesize taxuyunnanin C, was produced in both the cortex and phloem (Figure 5c).

Most of the paclitaxel biosynthesis-related enzymes are found mainly in the phloem, while the intermediate metabolites involved in paclitaxel biosynthesis are distributed in different stem tissues. For example, GGPP, taxa-4(20),11(12)-dien-5α-ol, 10β,14β-dihydroxytaxa-4(20),11(12)-dien-5α-yl acetate and BAC might be produced in the phloem and then transfer to the pith; taxa-4(5),11(12)-diene might be produced in the phloem and pith and transfer to the xylem; taxa-4(20),11(12)-dien-5α,13α-diol might be produced in the pith and then migrate to the xylem and phloem; and 10-deacetyl-2-debenzoylbaccatin III might be produced in the phloem and then transfer to the cortex and pith (Figure S6). Our data indicated possible migration routes of the intermediate metabolites involved in paclitaxel biosynthesis.

Isolation of the promoters of paclitaxel synthesis-related genes

Due to the unavailability of information on the *T. media* genome, promoters of most paclitaxel biosynthesis-related genes have not yet been isolated. Using the chromosome walking approach, partial promoter sequences of five paclitaxel biosynthesis-related genes, *TS*, *T7OH*, *T13OH*, *TBT* and *DBTNBT*, were successfully cloned in the present work (Figure S7). To investigate their potential regulators, all five promoter sequences were scanned for known TF motifs. The results showed that the promoter sequences of *TS*, *TBT* and *DBTNBT* contained at least one classical MYB-binding element (MBE; Figure S7). Interestingly, TmMYB3 was the most significant phloem-specific TF identified by proteomic analysis, hence we focused on TmMYB3 and its regulatory role in paclitaxel biosynthesis in subsequent studies.

Cloning and basic analysis of TmMYB3

Based on the transcriptomes of *T. media*, the partial sequence of *TmMYB3* was assembled, and using the assembled fragment as a specific template the full-length coding sequence (CDS) of *TmMYB3* was cloned using 5′/3′-rapid amplification of cDNA ends (RACE). Sequence

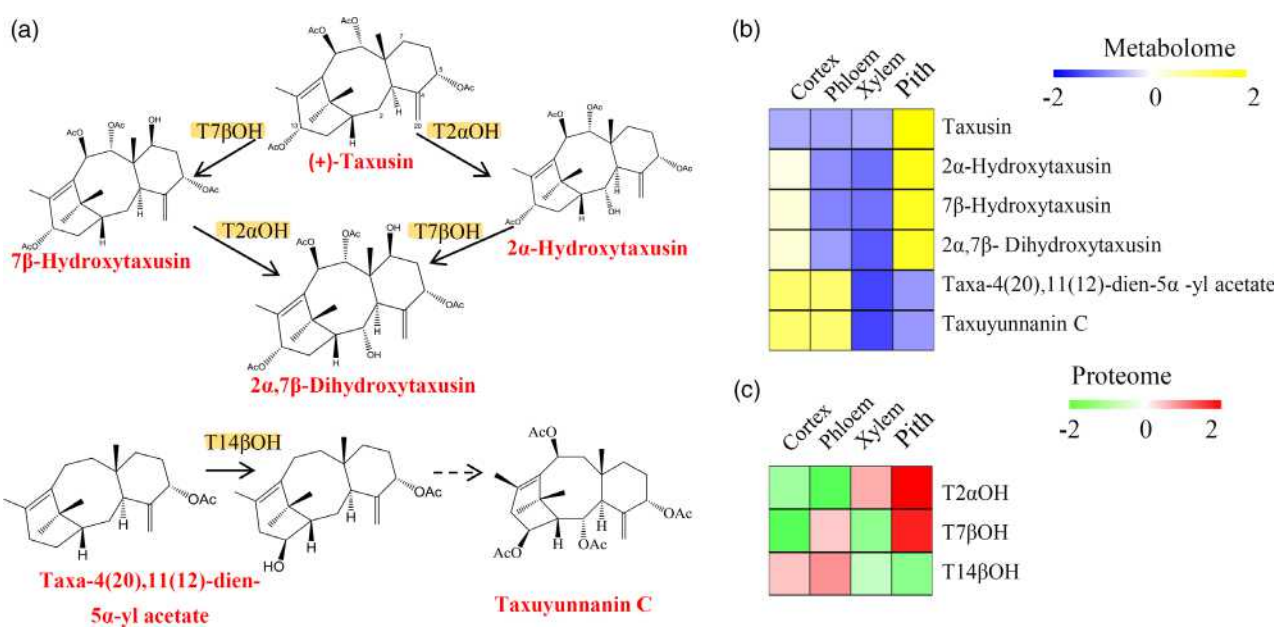

Figure 5. Integrated metabolomic and proteomic analysis of the biosynthetic pathways of other taxoids.
(a) Overview of the metabolic pathway of taxusin-like metabolites and taxuyunnanin C. T2αOH, taxadiene-2α-hydroxylase; T7βOH, taxane-7β-hydroxylase; T14βOH, Taxoid 14β-hydroxylase.
(b) Differential accumulation of taxusin-like metabolites and taxuyunnanin C.
(c) Differential production of three key enzymes involved in taxusin-like metabolites and taxuyunnanin C biosynthesis pathway. The heatmap scale ranges from −2 to +2 on a $\log_2$ scale.

analysis indicated that *TmMYB3* encodes a putative protein of 257 amino acid residues with a calculated molecular weight of 28.0 kDa and a predicted isoelectric point of 7.26. Multiple sequence alignment revealed that TmMYB3 possesses a complete R2R3 domain at the N-terminus. No common repressor domain, such as those containing EAR or TLLLFR motifs, was identified at the C-terminus of TmMYB3 (Figure 6a). BLAST hits and phylogenetic analysis indicated great similarity between TmMYB3 and *Gynura bicolor* MYB1, *Picea glauca* MYB10 and *Ginkgo biloba* MYB3 proteins (Figure S8).

To determine its subcellular localization, enhanced GFP-tagged TmMYB3 was transiently produced in tobacco epidermal cells. As shown in Figure 6(b), TmMYB3 was localized to the cell nucleus, consistent with the prediction that TmMYB3 acts as a TF. Based on the proteomic analysis, the abundance of TmMYB3 suggested a significant stem tissue-specific pattern, with the highest abundance in the phloem (Figure 6c). To confirm the phloem-specific expression of TmMYB3, Western blotting analysis using an anti-TmMYB3 antibody was performed. The results showed that the TmMYB3 protein was abundant in the

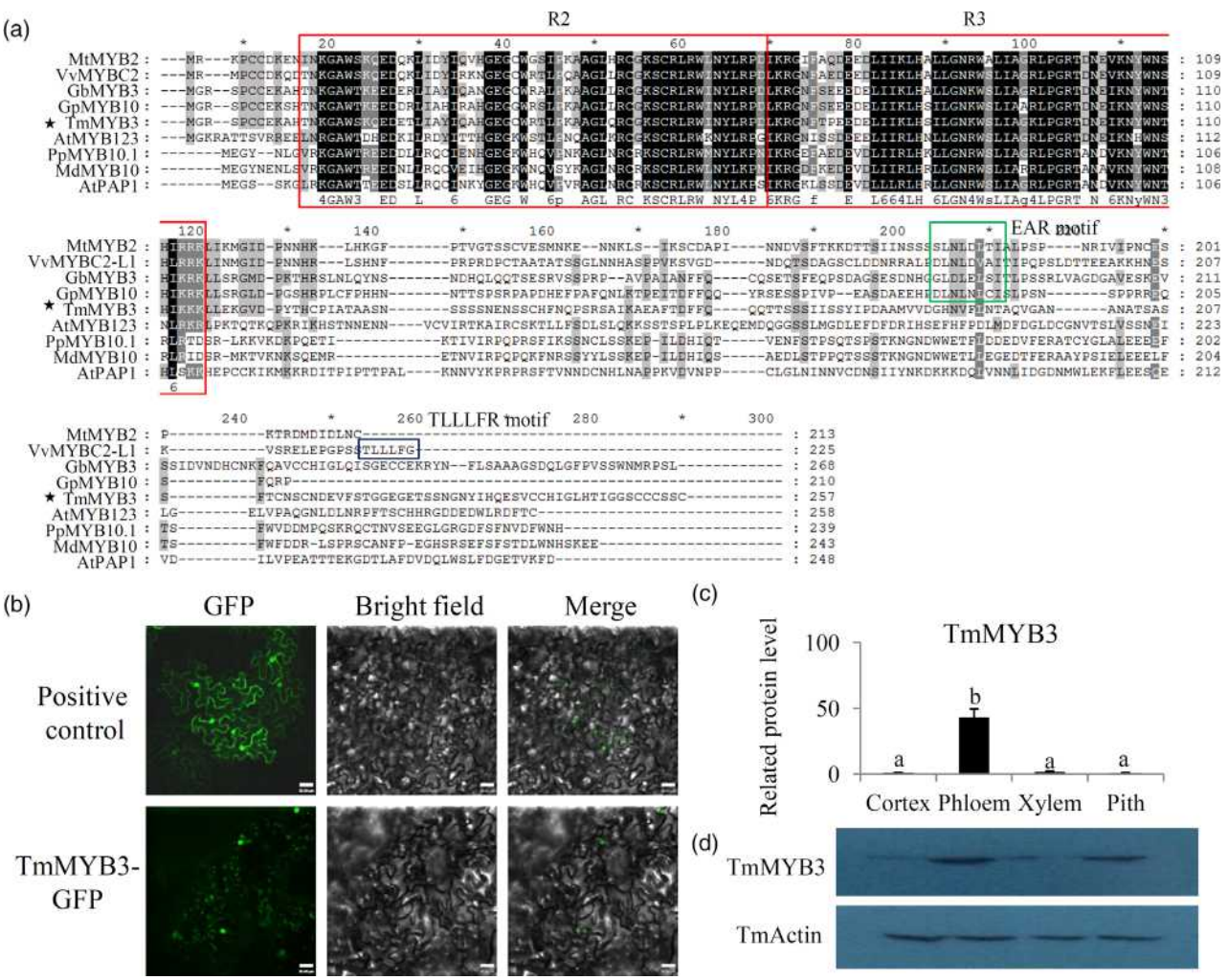

Figure 6. Cloning and basic analysis of TmMYB3.

(a) Amino acid sequence alignment of TmMYB3 and other known R2R3-MYBs in plants. Conserved residues are highlighted in black and partial conservation is shown in gray. The R2 and R3 domains are indicated with red boxes, whereas conserved motifs in the C-terminus are boxed. The following GenBank or Arabidopsis TAIR accession numbers were used: *Arabidopsis thaliana* AtPAP1 (AT1G56650), AtMYB123 (AT5G35550); *Medicago truncatula* MtMYB2 (XM_003616340); *Vitis vinifera* VvMYBC2-L1 (JX050227); *Prunus persica* PpMYB10.1 (XM_007216468); *Malus domestica* MdMYB10 (DQ267897); *Ginkgo biloba* GbMYB3 (ASR18088.1), GpMYB10 (ABQ51226.1).

(b) Subcellular localization of TmMYB3.

(c) Differential expression analysis of TmMYB3 in the four major stem tissues. The histogram shows expression levels of TmMYB3 protein based on the proteomic data. Different letters indicate significant differences in the expression levels of TmMYB3 in different stem tissues ($P < 0.05$).

(d) Western blot confirmed that TmMYB3 was highly produced in the phloem. TmActin was used as a control.

phloem (Figure 6c,d), which was consistent with the proteomic results.

TmMYB3 binds to the promoters of *TmTS*, *TmTBT* and *TmDBTNBT*

To assess the role of TmMYB3 in the transcription of paclitaxel synthesis-related genes, an *in vitro* electrophoretic mobility shift assay (EMSA) experiment was performed to investigate whether some paclitaxel biosynthesis-related genes are direct targets of TmMYB3. First, the promoter sequences of *TmTBT*, *TmDBTNBT* and *TmTS* were screened for TF motifs and all of them were found to bear several putative MYB recognition elements (MREs). Specifically, one MRE in the promoter of *TmTBT* (position −52 to −58), one MRE in the promoter of *DBTNBT* (position −609 to −603) and two MREs in the promoter of *TmTS* [positions −179 to −173 (TS1) and −790 to −784 (TS2)] were identified (Figure 7a). The TmMYB3-His fusion protein and the promoter fragments of *TmTBT*, *TmDBTNBT* and *TmTS* were prepared, and the EMSA results showed that TmMYB3 bound directly to the MREs in the promoters of *TmTBT*, *TmDBTNBT* and *TmTS*. Addition of excess cold competitive probe effectively weakened the binding of the His-TmMYB3 fusion protein to the DNAs, but addition of excess mutated probe did not affect their binding (Figure 7b–e).

Binding of TmMYB3 to these three promoters *in vivo* was further verified by chromatin immunoprecipitation (ChIP)-quantitative (q)PCR. As expected, the promoter regions containing MREs gave a strong signal using the TmMYB3 antibody compared with the negative control IgG (Figure 7f,g). Both *in vitro* and *in vivo* results suggested

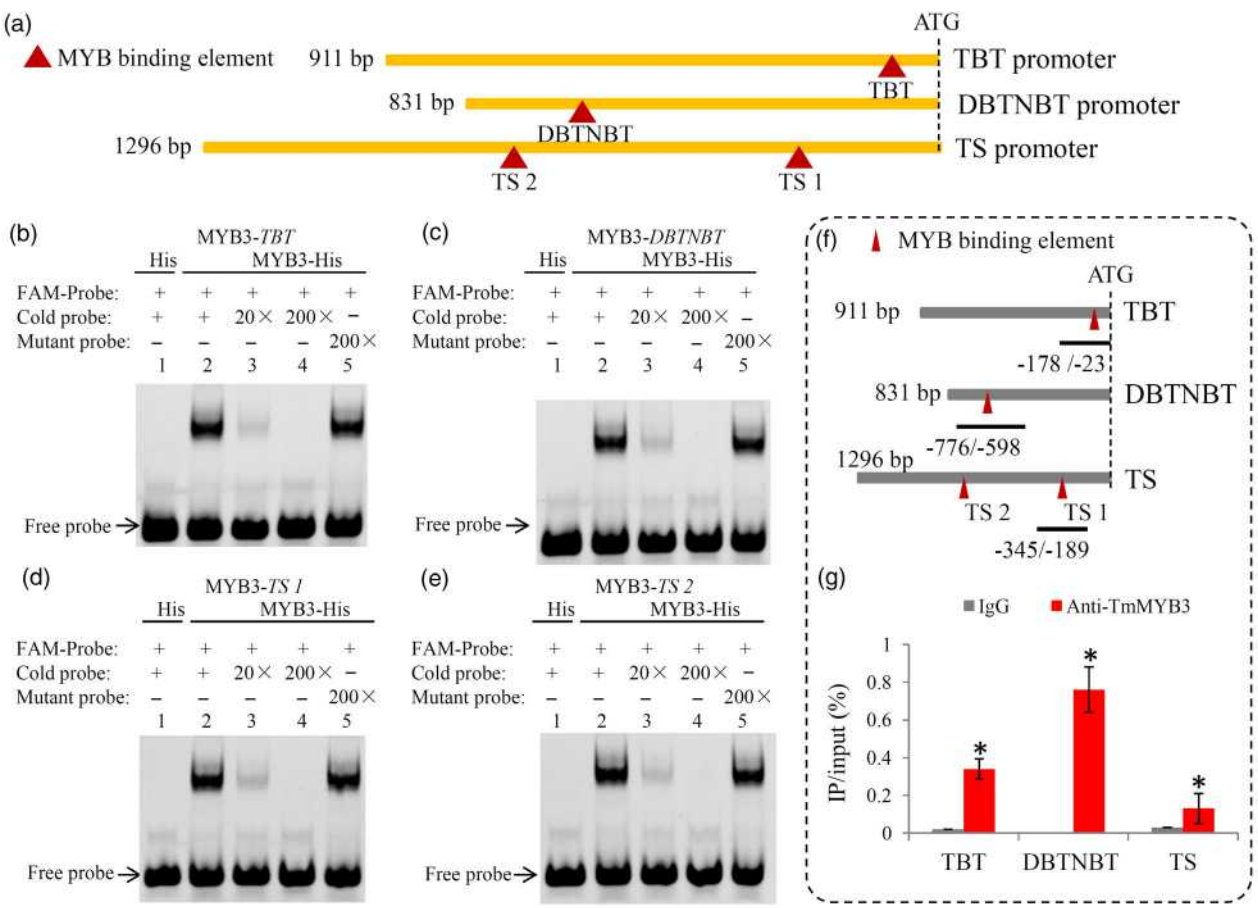

Figure 7. Electrophoretic mobility shift assay of the binding of TmMYB3 to the promoters of *TBT*, *DBTNBT* and *TS* genes.
(a)–(e) Screening of MYB-binding elements (MBEs) in the promoter regions of *TBT*, *DBTNBT* and *TS* genes. Crimson triangles indicated the MBEs. The His only or TmMYB3-His fusion protein was incubated with probes containing the MBEs derived from the promoters of *TBT* (b), *DBTNBT* (c) and *TS* (d,e) genes. − and + represent absence and presence, respectively, and 20× or 200× show increasing amounts of probes for competition. DBTNBT, 3′-*N*-debenzoyltaxol *N*-benzoyltransferase; TBT, taxane 2α-*O*-benzoyltransferase; TS, taxadiene synthase.
(f) Schematic representation of promoter structure of *TBT*, *DBTNBT* and *TS* genes. Crimson triangles indicate the MBEs. Black lines indicated the DNA fragments that were used for chromatin immunoprecipitation-quantitative PCR analyses.
Values are the percentage of DNA fragments that were co-immunoprecipitated with TmMYB3 antibody or anti-IgG relative to the input DNA. Error bars represent the SE of three independent repeats. Asterisks indicate statistically significant differences as determined by Student's *t*-test ($P < 0.05$).

that *TmTBT*, *TmDBTNBT* and *TmTS* might be downstream target genes of TmMYB3.

TmMYB3 enhances the expression of *TmTBT* and *TmTS*

To determine the transcriptional activities of TmMYB3 *in vivo*, dual-luciferase reporter assays were performed in *Nicotiana benthamiana* leaves (Figures 8a and S9). Dual-luciferase analysis showed that TmMYB3 effectively activated the expression of *TmTBT* and *TmTS* pro-LUC reporter genes compared with the empty construct, indicating that TmMYB3 functions as a transcriptional activator by binding directly to the promoters of *TmTBT* and *TmTS* genes. However, addition of TmMYB3 did not activate the expression of the *TmDBTNBT* gene (Figure 8b,c).

DISCUSSION

The stem bark of *Taxus* trees is a major source for paclitaxel extraction (Shen *et al.*, 2001; Uniyal, 2013). Increasing market demand for paclitaxel has led to vast amounts of illegal logging, destroying thousands of rare *Taxus* trees (Li *et al.*, 2012b). Despite its importance, current knowledge about the metabolism and proteome dynamics of *Taxus* species is very limited.

Recently, a number of taxoids have been isolated from *Taxus* trees, and their distribution and levels vary largely among different tissues (Ge *et al.*, 2010). A previous study showed great differences in the content of paclitaxel and 10-DAB among needle, stem and root tissues (Mubeen *et al.*, 2018). The distribution and amounts of paclitaxel vary greatly in different shoot parts of *T. cuspidata* (Fett Neto and DiCosmo, 1992). However, the composition of taxoids across the whole stem of *T. media* has remained poorly understood due to a lack of stem-specific studies. Herein, molecular fingerprints of four major stem tissue types (cortex, phloem, xylem and pith) were explored using an untargeted metabolomic method. Studies on all major tissues in the context of the whole stem could reveal the stem tissue specificity of medicinal compound accumulation.

The phloem is an important stem tissue for the biosynthesis and transport of secondary metabolites (Choudhary

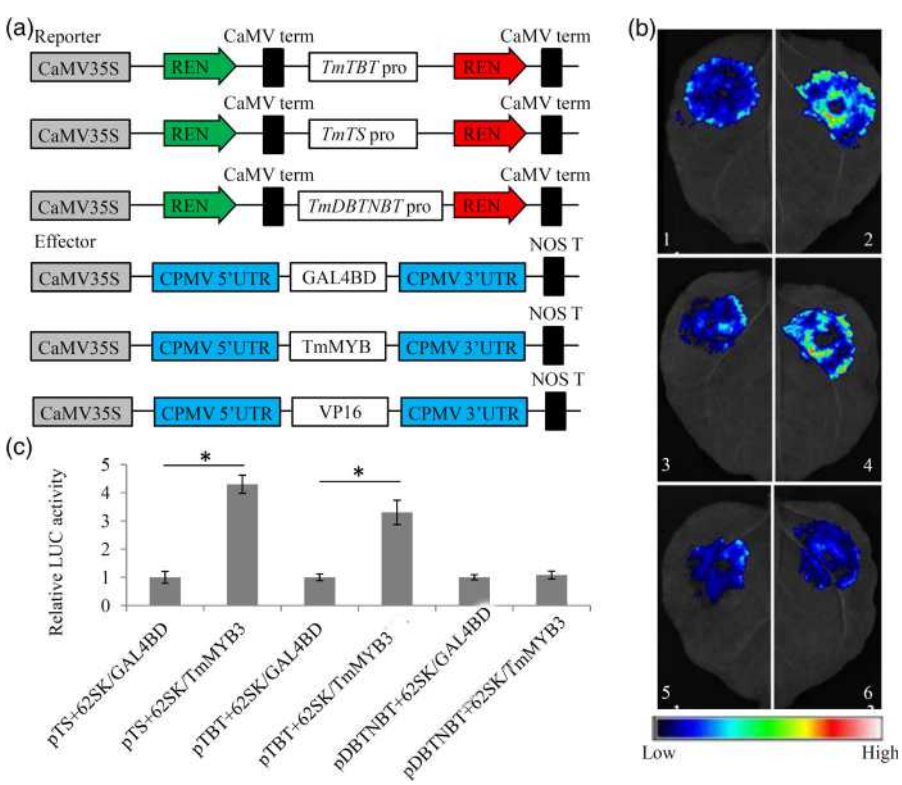

Figure 8. Transcriptional activation ability of TmMYB3.
(a) Schematic view of the constructs used for transient expression analysis and detailed information about plasmid combinations of dual REN/LUC reporters and effectors. The promoter fragments of *TmTS*, *TmDBTNBT* and *TmTBT* were cloned into pGreenII 0800-LUC vector to generate the reporter constructs. The effectors were generated by recombining the *TmMYB3* gene into the pGreenII 62SK vector. CPMV, cowpea mosaic virus; LUC, firefly luciferase; REN, *Renilla* luciferase; UTR, untranslated regions.
(b) The dual luciferase assays in tobacco leaves showed that co-transformation of TmMYB3 activates both *TmTS* and *TmTBT* promoters. Representative pictures were taken: 1, pTS+62SK/GAL4BD; 2, pTS+62SK/TmMYB3; 3, pTBT+62SK/GAL4BD; 4, pTBT+62SK/TmMYB3; 5, pDBTNBT+62SK/GAL4BD; 6, pDBTNBT+62SK/TmMYB3.
(c) The transcriptional activation ability of TmMYB3 is indicated by the LUC/REN ratio. The pBD that produced the GAL4 DNA-binding domain alone was used as a negative control. Each value is the mean ± SE of three biological repeats. The asterisk (*) represents significant differences ($P < 0.05$).

et al., 2016). In Norway spruce, phenolic stilbene glucosides, including astringin, isorhapontin and piceid, are mainly accumulated in the phloem (Li *et al.*, 2012a; Jyske *et al.*, 2016). In *T. media*, paclitaxel, as well as 10-DAB and DAP, accumulate at high levels in the phloem (Figure 1c), suggesting that this tissue may be the most significant source for paclitaxel extraction. Although the stem bark is a major source for isolation of taxoids, the pith contains high levels of BAC (Lin *et al.*, 2018). This suggests that the barkless trunk of *Taxus* trees might be a valuable material for extraction of BAC, and could be a cost-effective source of the key paclitaxel precursor (Lin *et al.*, 2018).

A broad range of taxusin-like metabolites are considered dead-end metabolites rather than intermediates leading to paclitaxel formation (Croteau *et al.*, 2006). A synthetic study demonstrated that a number of putative paclitaxel biosynthesis metabolites could be prepared by Barton deoxygenation of the C_9 and C_{10} hydroxyl groups of protected derivatives of taxusin (Li *et al.*, 2008). For many years, taxusin-like metabolites have been thought to predominantly accumulate in yew heartwood (Banskota *et al.*, 2003; Yu *et al.*, 2018). Our metabolomic results showed that taxusin-like metabolites predominantly accumulated in the pith rather than xylem of *T. media*. Knowledge of the precise distribution of taxusin and its analogues in the stem tissues will be useful for exploring the utilization of prominent side-route metabolites in *Taxus* trees.

The phloem is a key stem tissue that mediates stress tolerance by affecting plant–environment interactions (Savage *et al.*, 2016). A number of secondary metabolites, such as flavonoids, play essential roles in the responses to abiotic and biotic stresses (Brunetti *et al.*, 2018). Flavonoids accumulate highly in the phloem of *T. media*, suggesting that activation of the biosynthesis of flavonoids might play a role in phloem-mediated resistance to environmental stresses (Agati *et al.*, 2012). Analysis of the chemical composition showed that both the phloem and cortex were similar in terms of primary and secondary metabolites, whereas the numbers of pith-specific accumulated lipids (27 metabolites) and terpenoids (67 metabolites) were larger than in other stem tissues. Due to the differential metabolite composition of different stem tissues, the whole stem has broad potential for biotechnological applications.

Interestingly, the localizations of enzymes and the intermediate metabolites involved in paclitaxel biosynthesis are inconsistent in the *T. media* stem tissues. For example, TS is produced in the phloem and its downstream product taxa-4(5),11(12)-diene basically accumulates in the xylem. Another important intermediate, 10-DAB, accumulates in the phloem, but its downstream product BAC accumulates in the pith (Figures 4 and S6). Our data indicate the movements of the intermediates from one tissue to another. Similar to the taxoids, gibberellins (GAs) belong to another major class of diterpenoids, and the movements of GAs have been well studied in recent years (Binenbaum *et al.*, 2018). Several previous studies have demonstrated the presence of GAs in the phloem sap and the movement of GAs through the phloem (Hoad and Bowen, 1968). GA12 moved through the xylem in a root-to-shoot direction but through the phloem in a shoot-to-root manner (Regnault *et al.*, 2015). Due to the ability to move out of cells, the existence of GA efflux transporters is predicted for GA to effectively move locally, at both tissue and cellular levels (Kramer, 2006). However, taxoid efflux transporters have not yet been identified. In addition to active transport, the ability of taxoids to diffuse through membranes may be another explanation for the movements of taxoids from one tissue to another.

Metabolic engineering has the potential to achieve higher paclitaxel output (Wilson and Roberts, 2014). Improving the production yield of paclitaxel depends critically upon a deep understanding of the regulation of the paclitaxel biosynthesis pathway (Li *et al.*, 2013). Transcription factors can increase the accumulation of bioactive ingredients in medicinal plants by regulating genes in a specific biosynthetic pathway (Zhou *et al.*, 2019b). As a cultivated *Taxus* species with a high paclitaxel content, *T. media* is a major source for industrial production of paclitaxel (Shen *et al.*, 2016). Our previous studies identified most genes in the paclitaxel biosynthesis pathway of *T. media*, providing a solid foundation for further studies on the transcriptional regulation of paclitaxel biosynthesis (Hao *et al.*, 2017; Yu *et al.*, 2017).

In other *Taxus* species, several TFs, including MYCs in *T. cuspidata* and *T. chinensis*, WRKYs in *T. chinensis* and ERFs in *T. chinensis*, were deemed to be involved in the paclitaxel biosynthesis pathway (Li *et al.*, 2013; Lenka *et al.*, 2015; Zhang *et al.*, 2015; 2018a; Cui *et al.*, 2019). To date, no MYB TFs have been functionally identified in *Taxus* species. Herein, a R2R3-MYB TF, TmMYB3, was identified in *T. media*. In plants, secondary metabolism is regulated by a complex regulatory network in which MYBs are important components (Chezem and Clay, 2016). Previous studies showed that R2R3-MYB TFs mediate terpenoid biosynthesis by regulating the expression of MEP pathway genes such as *DXS*, *DXR* and *GGPPS* (Zhang *et al.*, 2017). In *Stevia rebaudiana* Bert, MYB TFs are thought to control the downstream metabolic flux of MEP and mevalonate pathways (Singh *et al.*, 2017). Paclitaxel is a bioactive substance built around a diterpenoid taxane core, and the expression of key genes involved in paclitaxel biosynthesis might be regulated by MYB TFs (Croteau *et al.*, 2006).

Sequence-specific DNA-binding sites of R2R3-MYB have been demonstrated in model plants (Romero *et al.*, 1998). Due to the unavailability of the *Taxus* genome, cloning of promoter sequences is needed to explore the relationships between TFs and their target genes. Using *T. cuspidata*

suspension cells, promoters of seven known paclitaxel pathway genes have been isolated (Lenka *et al.*, 2015). In *T. chinensis*, a 550 bp 5′-flanking sequence of the *TS* gene and a 1740 bp 5′-flanking sequence of the *DBAT* gene have been reported (Li *et al.*, 2015). Additionally, the role of a Y-patch promoter region of the *BAPT* gene has been revealed using *Taxus* cell cultures (Sanchez-Munoz *et al.*, 2018). To verify the target genes of TmMYB3, we isolated the promoter sequences of five important paclitaxel biosynthesis-related genes. After scanning the MREs, three potential target genes of TmMYB3, *TS*, *TBT* and *DBTNBT*, were identified in *T. media* (Figure 7a). Taxadiene synthase catalyzes a committal step in paclitaxel biosynthesis by producing the taxane core from the intermediate GGPP (Edgar *et al.*, 2017), TBT catalyzes the conversion of 10-deacetyl-2-debenzoylbaccatin III to 10-DAB (Walker and Croteau, 2000) and DBTNBT catalyzes the conversion of 3′-*N*-debenzoyltaxol into the target compound, paclitaxel (Nasiri *et al.*, 2016). All three enzymes play essential roles in paclitaxel biosynthesis (Croteau *et al.*, 2006). Herein, at least one potential MBE was identified in the promoter sequences of *TS*, *TBT* and *DBTNBT* genes, and EMSA assays confirmed the binding of TmMYB3 to these three promoters (Figure 7). Although physical binding occurs, exactly how TmMYB3 regulates downstream target genes by activation or repression still needs to be addressed. For example, TcMYC2 in *T. chinensis* is a positive regulator of *TS*, *TAT*, *DBTNBT*, *T13OH* and *T5OH* genes (Zhang *et al.*, 2018b), and two WRKYs in *T. chinensis*, TcWRKY8 and TcWRKY47, function as transcriptional activators in paclitaxel biosynthesis (Zhang *et al.*, 2018a). Two JA-responsive factors, TcERF12 and TcERF15, act as a repressor and an activator of the *TS* gene, respectively (Zhang *et al.*, 2015). In our present study, dual-luciferase transient expression analysis demonstrated the transcriptional activation activity of TmMYB3 against the promoters of the *TS* and *TBT* genes, but not the *DBTNBT* gene (Figure 8). Although most of the intermediate metabolites involved in paclitaxel biosynthesis might be produced in the phloem, several intermediate metabolites accumulated in different stem tissues. Interestingly, the downstream product of TS, taxa-4 (5),11(12)-diene, was transferred to the xylem, and the upstream product of TBT, 10-deacetyl-2-debenzoylbaccatin III, migrated to the pith and cortex. Taken together, our results revealed that TmMYB3 is a positive regulator involved in paclitaxel biosynthesis, possibly by activating the expression of *TmTBT* and *TmTS*. Moreover, possible movements of TS- and TBT-related products among different stem tissues are predicted (Figure 9).

In conclusion, the integration of proteomic and metabolomic profiling of *Taxus* stem tissues provides a molecular framework for revealing the mechanism underlying tissue-specific accumulation of medicinal compounds such as taxoids. The differential metabolite compositions of different stem tissues highlight the broad potential for biotechnological applications. Moreover, as a phloem-specific TF, TmMYB3 has a significant effect on paclitaxel biosynthesis by activating the expression of the *TS* and *TBT* genes, providing a potential explanation for the phloem-specific accumulation of paclitaxel.

EXPERIMENTAL PROCEDURES

Plant materials and sampling

Five-year-old *T. media* (cultivar '*Zike*') plants were cultivated in an experimental field within the campus of Hangzhou Normal University, Hangzhou, China. In our study, four different stem tissues, including cortex, phloem, xylem and pith, were tested. Independent samples from the different stem tissues were used for metabolomic analysis (25 mg each, $n = 10$) and proteomic analysis (50 mg each, $n = 3$). After the isolation of stem tissues, cross sections were observed under a microscope to localize the anatomical position. Young stems of 5-year-old *T. media* trees were collected and prepared for microsection observation according to a previously published paper (Chen *et al.*, 2014). All slides were mounted with synthetic resin and observed under a LSM510 laser scanning system (Carl Zeiss, https://www.zeiss.com/).

Quantitative UPLC-MS/MS and untargeted metabolomic analyses

Fresh stems were harvested from *T. media* plants ($n = 6$), and four major tissues of each stem, including cortex, phloem, xylem and pith, were prepared by peeling and scraping. The tissue samples were dried at 40°C and ground into a fine powder. Our previously published method was used for the preparation of crude extracts and UPLC-MS/MS quantitative analysis (Yu *et al.*, 2018). For the metabolomic analysis, metabolite extraction, UPLC-MS/MS analysis and bioinformatic analysis were performed according to our previous work (Yu *et al.*, 2018).

Proteomic analysis

For proteomic analysis, protein extraction, trypsin digestion, Tandem Mass Tag (TMT) labeling, peptide fractionation, LC-MS/MS analysis, protein annotation and functional enrichment were performed according to our previous work (Hao *et al.*, 2017). The MS/MS data generated by MaxQuant with integrated Andromeda search engine v.1.5.2.8 were searched against the transcriptome of *T. media* (NCBI accession number GSM3438662).

Promoter isolation by chromosome walking

High-quality *T. media* DNA was isolated from young twigs using the classical cetyl trimethyl ammonium bromide method. The upstream promoter regions of *TBT*, *TS*, *DBTNBT*, *T7OH* and *T13OH* genes were amplified using a Universal Genome Walker Kit (Clontech, https://www.takarabio.com/). In brief, genomic DNA was digested with four blunt-end restriction enzymes, *Eco*RV, *Pvu*II, *Dra*I and *Stu*I, and the digested products were ligated with Genome Walker adaptors to produce templates for first-round nested PCR. The products of first-round PCR were used as templates in secondary PCR with nested primers. The PCR products were ligated with the pMD18-T vector (TaKaRa, https://www.takarabio.com/) and sequences were analyzed using the PlantCARE program (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). The sequences of primers used for the chromosome walking are listed in Table S1.

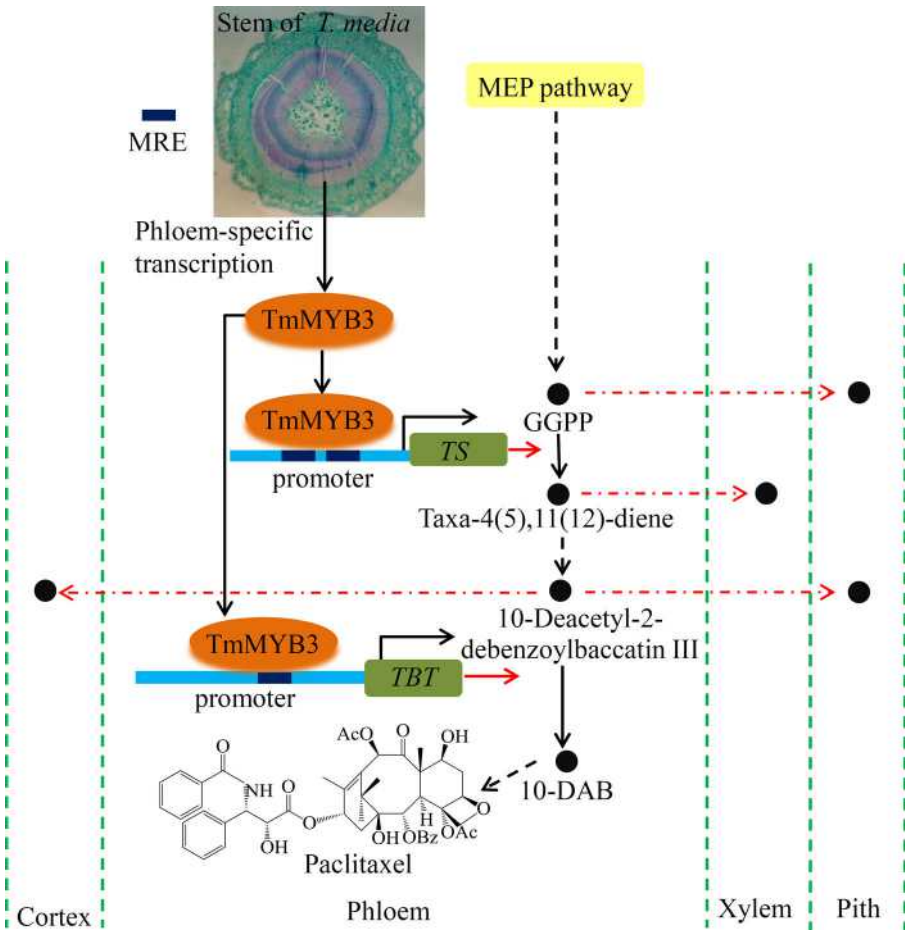

Figure 9. A model for the role of TmMYB3 in paclitaxel biosynthesis.
The phloem-specific TmMYB3 has a significant effect on paclitaxel biosynthesis by activating the expression of the *TS* and *TBT* genes. GGPP, geranylgeranyl diphosphate; MEP, 2-*C*-methyl-D-erythritol 4-phosphate; MRE, MYB recognition element; 10-DAB, 10-deacetylbaccatin III.

Isolation of full-length *TmMYB3* by 5′/3′-RACE

The partial *TmMYB3* gene sequence was identified from the *T. media* transcriptome, and full-length *TmMYB3* cDNA was obtained using a SMARTer RACE 5′/3′ Kit (Clontech) according to the manufacturer's manual. In brief, total RNAs were extracted from young twigs of *T. media* using TRIzol reagent (Qiagen, https://www.qiagen.com/). A mixture containing 1 μg of RNA, 1 μl of 5′ CDS primer, 1 μl of SMARTer II A oligonucleotide and double-distilled H_2O to the required final volume was served as the 5′-RACE-Ready cDNA solution. Master Mix solution was prepared using 5× First Strand buffer (4 μl), 100 mM dithiothreitol (0.5 μl), 20 mM dNTP mix (1 μl), RNase inhibitor (20 U) and SMARTScribe Reverse Transcriptase (200 U). The Master Mix was added to the 5′-RACE-Ready cDNA solution, and the mixture was incubated at 42°C for 90 min. The reaction was stopped by heating at 70°C for 10 min. Sequences of primers used for 5′-RACE are listed in Table S1.

Subcellular localization and phylogenetic analysis of TmMYB3

The full-length sequence of *TmMYB3* was cloned into the pH7FWG2.0 vector with an artificial green fluorescent protein (GFP) gene fused at the N-terminus. Sequences of primers used for the GFP-TmMYB3 construct are listed in Table S1. All vectors were transiently expressed in *N. benthamiana* epidermal cells by *Agrobacterium tumefaciens* (GV3101)-mediated transformation. The GFP fluorescence of the fusion protein was detected using a LSM710 confocal microscope (Carl Zeiss).

The TmMYB3 protein sequence and other tryptic MYB proteins from different plants were used for multiple sequence alignments using ClustalW with default parameters. An unrooted phylogenetic tree was constructed using MEGA 6.1 (http://www.megasoftware.net/) employing the neighbor-joining method.

Preparation of TmMYB3-specific polyclonal antibody and Western blotting

The full-length *TmMYB3* gene was inserted into the pET30a vector and transformed into *Escherichia coli* Rosetta (DE3) to produce recombinant TmMYB3 protein with a His tag. After inducing with 1 mM isopropyl β-D-1-thiogalactopyranoside at 30°C for 5 h, the recombinant protein was isolated and purified using His60 Ni Superflow Resin (Clontech) according to the manufacturer's instructions. His-tagged fusion proteins were separated by SDS-PAGE and bands of interest were used for preparation of TmMYB3-specific polyclonal antibody. The expression of the

TmMYB3 protein in distinct stem tissues was tested by Western blotting analysis. Total proteins were extracted from different stem tissues (100 mg each). About 10 μg of total soluble proteins for each tissue were separated on 12% SDS-PAGE gels and transferred to a polyvinyl difluoride membrane. The Western blotting analysis was performed according to a previous work (Liu *et al.*, 2013).

Electrophoretic mobility shift assay and ChIP-qPCR analysis

Electrophoretic mobility shift assay was performed as previously described (Hellman and Fried, 2007). Briefly, probes containing MYB-specific *cis*-elements (CAGTTA and TGGTTA) derived from the promoters of *TBT*, *TS* and *DBTNBT* genes were labeled with 5′6-FAM (FITC) fluorescent dye. Unlabeled probes were used in competition assays. MYB-specific *cis*-elements within probes were changed into CCCGGG and used in mutation assays. The probe sequences used for EMSA are listed in Table S1.

Taxus media samples were treated with 1% formic acid (FA) to cross-link genomic DNA and TmMYB3 protein. Chromatin was cut into pieces with an average length of 500 bp by sonication on ice, and immunoprecipitation of TmMYB3 cross-linked with DNA was carried out using affinity-purified TmMYB3 antibody at 4°C overnight. Preimmune IgG was used as a negative control. The cross-linked complex was isolated on Protein A/agarose beads following a 1 h incubation at 4°C. Beads were washed for 10 min at 4°C with low-salt, high-salt, LiCl and TRIS-EDTA washing buffers, and the immunoprecipitated sample was eluted by gentle rotation for 15 min at 65°C. Cross-linking of immunoprecipitated DNA was reversed by incubation in 0.2 M NaCl at 65°C overnight. The amount of each precipitated DNA fragment was determined by quantitative real-time PCR (qRT-PCR) (Yu *et al.*, 2017).

Dual-luciferase reporter assay

For transcription activity analysis, the full-length cDNA of TmMYB3 was inserted into the pBD vector under the control of the 35S promoter as effector (pBD-TmMYB3). A GAL4-LUC reporter and an internal control REN (*Renilla* luciferase) reporter under the control of the 35S promoter were constructed as double-reporter vectors. To determine the binding activity of TmMYB3 to the promoters of paclitaxel biosynthesis-related genes, the promoters of *TS*, *TBT* and *DBTNBT* were cloned into the pGreenII 0800-LUC double-reporter vector. The resulting effector and reporter plasmids were co-transformed into tobacco leaves by *A. tumefaciens* strain GV3101 as described previously (Zhu *et al.*, 2017). The LUC and REN luciferase activities were determined using a dual luciferase assay kit (Promega, https://www.promega.com/) according to the manufacturer's instructions. The binding activities of TmMYB3 to the selected promoters were calculated from the LUC/REN ratio (assured in triplicate). Sequences of primers used for the dual-luciferase reporter assay are listed in Table S1.

Statistical analysis

For untargeted metabolomic analysis, the Wilcoxon test was performed to identify DAMs in each sample comparison, and supervised partial least squares-discriminant analysis was performed to analyze the different variables between two sample groups. For transcriptomic analysis, *P*-values, which were used as a threshold to determine differentially expressed genes in each sample comparison, were produced by false discovery rate analysis and adjusted using the Benjamini–Hochberg method. The statistical analysis was carried out using SPSS software v.19.0 (https://spss.software.informer.com/). A one-way analysis of variance was carried out to compare the content differences of taxoids between two sample groups. A P-value < 0.05 was considered statistically significant.

ACCESSION NUMBERS

The metabolomic datasets generated and analyzed during the current study are available in the 'Baidu Netdisk' (https://pan.baidu.com/s/1AIkJRQYcao-oxbx3Q_iO2Q) with extraction code ytsc. The MS proteomic data have been set to the ProteomeXchange Consortium by the PRIDE partner repository program under identifier PXD008871. The transcriptomes of *T. media* used for proteome annotation have been set to the NCBI database under accession number GSM3438662.

ACKNOWLEDGEMENTS

This work was funded by Zhejiang Provincial Natural Science Foundation of China under grant no. LY19C160001, the Open Foundation of State Key Laboratory of Subtropical Silviculture, Zhejiang A & F University (KF201708), Major Increase or Decrease Program in the Central Finance Level grant no. 2060302 and Zhejiang Provincial Key Research & Development Project grants 2017C02011 and 2018C02030. We are also grateful to LC Sciences company (Hangzhou, China) for metabolomic analysis and PTM Biolabs Inc. (Hangzhou, China) for proteomic analysis.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

AUTHOR CONTRIBUTIONS

CY, BZ, HW and CS designed the research. CY, XX and HY performed the metabolite and protein extractions. XL and SF did the sampling for quantification of four different taxoids. XL and CZ performed microsection observation. XL and XX took care of the plants. CY, XL, CZ, YC and XZ performed targeted UPLC-MS/MS analysis. CY, XL, CZ, XZ, JX, XX and SF performed proteomic analysis. CY and XX carried out promoter isolation by chromosome walking and full-length *TmMYB3* isolation by 5′/3′-RACE. CY, XL and HY performed subcellular localization and phylogenetic analysis. CY, XL, CZ, XX and XZ did TmMYB3-specific polyclonal antibody production and Western blotting. LZ performed the EMSA and ChIP-qPCR analysis. HW and CS wrote the manuscript.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article.

Figure S1. Quality control parameters of the metabolomes.

Figure S2. Protein electrophoresis photograph of isolated protein samples.

Figure S3. Quality control parameters of the proteomes.

Figure S4. Checking the statistical consistency.

Figure S5. Analysis of differentially produced proteins among the four stem tissues.

Figure S6. The movements of intermediates among different stem tissues.

Figure S7. The promoter sequences of five important paclitaxel biosynthesis-related genes.

Figure S8. Phylogenetic analysis of TmMYB3 and several known MYBs from other plants.

Figure S9. The positive controls of dual-luciferase reporter assays.

Table S1. The primer sequences used in the present study.

Table S2. Detailed information about the data from the untargeted metabolomes.

Table S3. A total of 5077 annotated metabolites were mapped into different known metabolic pathways.

Table S4. The annotation information for all the identified proteins.

Table S5. The detailed information on the phloem-specific expressed proteins.

Table S6. Proteomic analysis identified nine classical transcription factors.

REFERENCES

Agati, G., Azzarello, E., Pollastri, S. and Tattini, M. (2012) Flavonoids as antioxidants in plants: Location and functional significance. *Plant Sci.* **196**, 67–76.

Banskota, A.H., Tezuka, Y., Nguyen, N.T., Awale, S., Nobukawa, T. and Kadota, S. (2003) DPPH radical scavenging and nitric oxide inhibitory activities of the constituents from the wood of *Taxus yunnanensis*. *Planta Med.* **69**, 500–505.

Bernabeu, E., Cagel, M., Lagomarsino, E., Moretton, M. and Chiappetta, D.A. (2017) Paclitaxel: What has been done and the challenges remain ahead. *Int. J. Pharm.* **526**, 474–495.

Binenbaum, J., Weinstain, R. and Shani, E. (2018) Gibberellin localization and transport in plants. *Trends Plant Sci.* **23**, 410–421.

Brunetti, C., Fini, A., Sebastiani, F., Gori, A. and Tattini, M. (2018) Modulation of phytohormone Signaling: a primary function of flavonoids in plant-environment interactions. *Front. Plant Sci.* **9**, 1042.

Chan, K.C., Alvarado, A.B., McGuire, M.T., Muschik, G.M., Issaq, H.J. and Snader, K.M. (1994) High-performance liquid chromatography and micellar electrokinetic chromatography of taxol and related taxanes from bark and needle extracts of *Taxus* species. *J. Chromatogr. B, Biomed. Appl.* **657**, 301–306.

Chen, Y.J., Liang, Z.T., Zhu, Y., Xie, G.Y., Tian, M., Zhao, Z.Z. and Qin, J.Q. (2014) Tissue-specific metabolites profiling and quantitative analyses of flavonoids in the rhizome of *Belamcanda chinensis* by combining laser-microdissection with UHPLC-Q/TOF-MS and UHPLC–QqQ-MS. *Talanta*, **130**, 585–597.

Chezem, W.R. and Clay, N.K. (2016) Regulation of plant secondary metabolism and associated specialized cell development by MYBs and bHLHs. *Phytochemistry*, **131**, 26–43.

Choudhary, S.B., Kumar, M., Chowdhury, I., Singh, R.K., Pandey, S.P., Sharma, H.K. and Karmakar, P.G. (2016) An efficient and cost effective method of RNA extraction from mucilage, phenol and secondary metabolite rich bark tissue of tossa jute (*C. olitorius* L.) actively developing phloem fiber. *3 Biotech*, **6**, 100.

Croteau, R., Ketchum, R.E., Long, R.M., Kaspera, R. and Wildung, M.R. (2006) Taxol biosynthesis and molecular genetics. *Phytochem. Rev.* **5**, 75–97.

Cui, Y., Mao, R., Chen, J. and Guo, Z. (2019) Regulation mechanism of MYC family transcription factors in jasmonic acid signalling pathway on taxol biosynthesis. *Int. J. Mol. Sci.* **20**, pii:E1843.

Dubos, C., Stracke, R., Grotewold, E., Weisshaar, B., Martin, C. and Lepiniec, L. (2010) MYB transcription factors in *Arabidopsis*. *Trends Plant Sci.* **15**, 573–581.

Edgar, S., Li, F.S., Qiao, K., Weng, J.K. and Stephanopoulos, G. (2017) Engineering of taxadiene synthase for improved selectivity and yield of a key taxol biosynthetic intermediate. *ACS Synth. Biol.* **6**, 201–205.

Eisenreich, W., Menhard, B., Hylands, P.J., Zenk, M.H. and 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.

Fett Neto, A.G. and DiCosmo, F. (1992) Distribution and amounts of taxol in different shoot parts of *Taxus cuspidata*. *Planta Med.* **58**, 464–466.

Ge, G.B., Liang, S.C., Hu, Y., Liu, X.B., Mao, Y.X., Zhang, Y.Y., Luan, H.W., Qiu, M.H. and Yang, L. (2010) Rapid qualitative and quantitative determination of seven valuable taxanes from various *Taxus* species by UFLC-ESI-MS and UFLC-DAD. *Planta Med.* **76**, 1773–1777.

Hao, D.C., Xiao, P.G., Ge, G.B. and Liu, M. (2012) Biological, chemical, and omics research of *Taxus* medicinal resources. *Drug Dev. Res.* **73**, 477–486.

Hao, J., Guo, H., Shi, X., Wang, Y., Wan, Q., Song, Y.B., Zhang, L., Dong, M. and Shen, C. (2017) Comparative proteomic analyses of two *Taxus* species (*Taxus* × *media* and *Taxus mairei*) reveals variations in the metabolisms associated with paclitaxel and other metabolites. *Plant Cell Physiol.* **58**, 1878–1890.

Hellman, L.M. and Fried, M.G. (2007) Electrophoretic mobility shift assay (EMSA) for detecting protein-nucleic acid interactions. *Nat. Protoc.* **2**, 1849–1861.

Hoad, G.V. and Bowen, M.R. (1968) Evidence for gibberellin-like substances in phloem exudate of higher plants. *Planta*, **82**, 22–32.

Jyske, T., Kuroda, K., Suuronen, J.-P., Pranovich, A., Roig-Juan, S., Aoki, D. and Fukushima, K. (2016) In planta localization of stilbenes within *Picea abies* phloem. *Plant Physiol.* **172**, 913–928.

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

Koepp, A.E., Hezari, M., Zajicek, J., Vogel, B.S., LaFever, R.E., Lewis, N.G. and Croteau, R. (1995) Cyclization of geranylgeranyl diphosphate to taxa-4(5),11(12)-diene is the committed step of taxol biosynthesis in Pacific yew. *J. Biol. Chem.* **270**, 8686–8690.

Kramer, E.M. (2006) How far can a molecule of weak acid travel in the apoplast or xylem? *Plant Physiol.* **141**, 1233–1236.

Kuang, X., Sun, S., Wei, J., Li, Y. and Sun, C. (2019) Iso-Seq analysis of the *Taxus cuspidata* transcriptome reveals the complexity of Taxol biosynthesis. *BMC Plant Biol.* **19**, 210–210.

Lenka, S.K., Nims, N.E., Vongpaseuth, K., Boshar, R.A., Roberts, S.C. and Walker, E.L. (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, H., Horiguchi, T., Croteau, R. and Williams, R.M. (2008) Studies on Taxol biosynthesis: preparation of taxadiene-diol- and triol-derivatives by deoxygenation of taxusin. *Tetrahedron*, **64**, 6561–6567.

Li, S.H., Nagy, N.E., Hammerbacher, A., Krokene, P., Niu, X.M., Gershenzon, J. and Schneider, B. (2012a) Localization of phenolics in phloem parenchyma cells of Norway spruce (*Picea abies*). *ChemBioChem*, **13**, 2707–2713.

Li, S.T., Zhang, P., Zhang, M., Fu, C.H., Zhao, C.F., Dong, Y.S., Guo, A.Y. and Yu, L.J. (2012b) Transcriptional profile of *Taxus chinensis* cells in response to methyl jasmonate. *BMC Genom.* **13**, 295.

Li, S., Zhang, P., Zhang, M., Fu, C. and Yu, L. (2013) Functional analysis of a WRKY transcription factor involved in transcriptional activation of the *DBAT* gene in Taxus chinensis. *Plant Biol. (Stuttg)*, **15**, 19–26.

Lin, S.L., Wei, T., Lin, J.F., Guo, L.Q., Wu, G.P., Wei, J.B., Huang, J.J. and Ouyang, P.L. (2018) Bio-production of baccatin III, an important precursor of paclitaxel by a cost-effective approach. *Mol. Biotechnol.* **60**, 492–505.

Liu, X., Yang, L., Zhou, X., Zhou, M., Lu, Y., Ma, L., Ma, H. and Zhang, Z. (2013) Transgenic wheat expressing *Thinopyrum intermedium* MYB transcription factor TiMYB2R-1 shows enhanced resistance to the take-all disease. *J. Exp. Bot.* **64**, 2243–2253.

Ma, D., Reichelt, M., Yoshida, K., Gershenzon, J. and Constabel, C.P. (2018) Two R2R3-MYB proteins are broad repressors of flavonoid and phenylpropanoid metabolism in poplar. *Plant J.* **96**, 949–965.

Mubeen, S., Li, Z.L., Huang, Q.M., He, C.T. and Yang, Z.Y. (2018) Comparative transcriptome analysis revealed the tissue-specific accumulations of taxanes among three experimental lines of *Taxus yunnanensis*. *J. Agric. Food Chem.* **66**, 10410–10420.

Nadeem, M., Rikhari, H.C., Kumar, A., Palni, L.M. and Nandi, S.K. (2002) Taxol content in the bark of Himalayan Yew in relation to tree age and sex. *Phytochemistry*, **60**, 627–631.

Nasiri, J., Naghavi, M.R., Alizadeh, H. and Moghadam, M.R. (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.

Nguyen, N.T., Banskota, A.H., Tezuka, Y., Nobukawa, T. and Kadota, S. (2003) Diterpenes and sesquiterpenes from the bark of *Taxus yunnanensis*. *Phytochemistry*, **64**, 1141–1147.

Onrubia, M., Moyano, E., Bonfill, M., Palazon, J., Goossens, A. and Cusido, R.M. (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.

Regnault, T., Daviere, J.M., Wild, M., Sakvarelidze-Achard, L., Heintz, D., Carrera Bergua, E., Lopez Diaz, I., Gong, F., Hedden, P. and Achard, P. (2015) The gibberellin precursor GA12 acts as a long-distance growth signal in *Arabidopsis*. *Nat. Plants*, **1**, 15073.

Romero, I., Fuertes, A., Benito, M.J., Malpica, J.M., Leyva, A. and Paz-Ares, J. (1998) More than 80R2R3-MYB regulatory genes in the genome of *Arabidopsis thaliana*. *Plant J.* **14**, 273–284.

Sanchez-Munoz, R., Bonfill, M., Cusido, R.M., Palazon, J. and Moyano, E. (2018) Advances in the regulation of in vitro paclitaxel production: methylation of a Y-patch promoter region alters BAPT gene expression in Taxus cell cultures. *Plant Cell Physiol.* **59**, 2255–2267.

Savage, J.A., Clearwater, M.J., Haines, D.F., Klein, T., Mencuccini, M., Sevanto, S., Turgeon, R. and Zhang, C. (2016) Allocation, stress tolerance and carbon transport in plants: how does phloem physiology affect plant ecology? *Plant Cell Environ.* **39**, 709–725.

Shen, Y.C., Prakash, C.V., Chen, Y.J., Hwang, J.F., Kuo, Y.H. and Chen, C.Y. (2001) Taxane diterpenoids from the stem bark of *Taxus mairei*. *J. Nat. Prod.* **64**, 950–952.

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

Singh, G., Singh, G., Singh, P. *et al.* (2017) Molecular dissection of transcriptional reprogramming of steviol glycosides synthesis in leaf tissue during developmental phase transitions in *Stevia rebaudiana* Bert. *Sci. Rep.* **7**, 11835.

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

Tanaka, K., Li, F., Morikawa, K., Nobukawa, T. and Kadota, S. (2011) Analysis of biosynthetic fluctuations of cultured *Taxus* seedlings using a metabolomic approach. *Phytochemistry*, **72**, 1760–1766.

Uniyal, S.K. (2013) Bark removal and population structure of *Taxus wallichiana* Zucc. in a temperate mixed conifer forest of western Himalaya. *Environ. Monit. Assess.* **185**, 2921–2928.

Walker, K. and 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, Y.F., Shi, Q.W., Dong, M., Kiyota, H., Gu, Y.C. and Cong, B. (2011) Natural taxanes: developments since 1828. *Chem. Rev.* **111**, 7652–7709.

Wang, C.Y., Tang, L., He, J.W., Li, J. and Wang, Y.Z. (2019) Ethnobotany, phytochemistry and pharmacological properties of *Eucommia ulmoides*: a review. *Am. J. Chin. Med.* **47**, 259–300.

Wilson, S.A. and Roberts, S.C. (2014) Metabolic engineering approaches for production of biochemicals in food and medicinal plants. *Curr. Opin. Biotechnol.* **26**, 174–182.

Yu, C., Guo, H., Zhang, Y. *et al.* (2017) Identification of potential genes that contributed to the variation in the taxoid contents between two *Taxus* species (*Taxus media* and *Taxus mairei*). *Tree Physiol.* **37**, 1659–1671.

Yu, C., Luo, X., Zhan, X., Hao, J., Zhang, L., Song, Y.B., Shen, C. and Dong, M. (2018) Comparative metabolomics reveals the metabolic variations between two endangered Taxus species (*T. fuana* and *T. yunnanensis*) in the Himalayas. *BMC Plant Biol.* **18**, 197.

Zhang, M., Li, S., Nie, L., Chen, Q., Xu, X., Yu, L. and Fu, C. (2015) Two jasmonate-responsive factors, TcERF12 and TcERF15, respectively act as repressor and activator of tasy gene of taxol biosynthesis in *Taxus chinensis*. *Plant Mol. Biol.* **89**, 463–473.

Zhang, J., Zhou, L., Zheng, X., Zhang, J., Yang, L., Tan, R. and Zhao, S. (2017) Overexpression of *SmMYB9b* enhances tanshinone concentration in *Salvia miltiorrhiza* hairy roots. *Plant Cell Rep.* **36**, 1297–1309.

Zhang, M., Chen, Y., Nie, L., Jin, X., Liao, W., Zhao, S., Fu, C. and Yu, L. (2018a) Transcriptome-wide identification and screening of WRKY factors involved in the regulation of taxol biosynthesis in *Taxus chinensis*. *Sci. Rep.* **8**, 5197.

Zhang, M., Jin, X., Chen, Y., Wei, M., Liao, W., Zhao, S., Fu, C. and Yu, L. (2018b) TcMYC2a, a basic helix-loop-helix transcription factor, transduces JA-signals and regulates Taxol biosynthesis in *Taxus chinensis*. *Front. Plant Sci.* **9**, 863.

Zheng, W., Komatsu, S., Zhu, W., Zhang, L., Li, X., Cui, L. and Tian, J. (2016) Response and defense mechanisms of *Taxus chinensis* leaves under UV-A radiation are revealed using comparative proteomics and metabolomics analyses. *Plant Cell Physiol.* **57**, 1839–1853.

Zhou, H., Lin-Wang, K., Wang, F. *et al.* (2019a) Activator-type R2R3-MYB genes induce a repressor-type R2R3-MYB gene to balance anthocyanin and proanthocyanidin accumulation. *New Phytol.* **221**, 1919–1934.

Zhou, T., Luo, X., Yu, C., Zhang, C., Zhang, L., Song, Y.-B., Dong, M. and Shen, C. (2019b) Transcriptome analyses provide insights into the expression pattern and sequence similarity of several taxol biosynthesis-related genes in three *Taxus* species. *BMC Plant Biol.* **19**, 33–33.

Zhu, F., Luo, T., Liu, C. *et al.* (2017) An R2R3-MYB transcription factor represses the transformation of alpha- and beta-branch carotenoids by negatively regulating expression of *CrBCH2* and *CrNCED5* in flavedo of *Citrus reticulate*. *New Phytol.* **216**, 178–192.