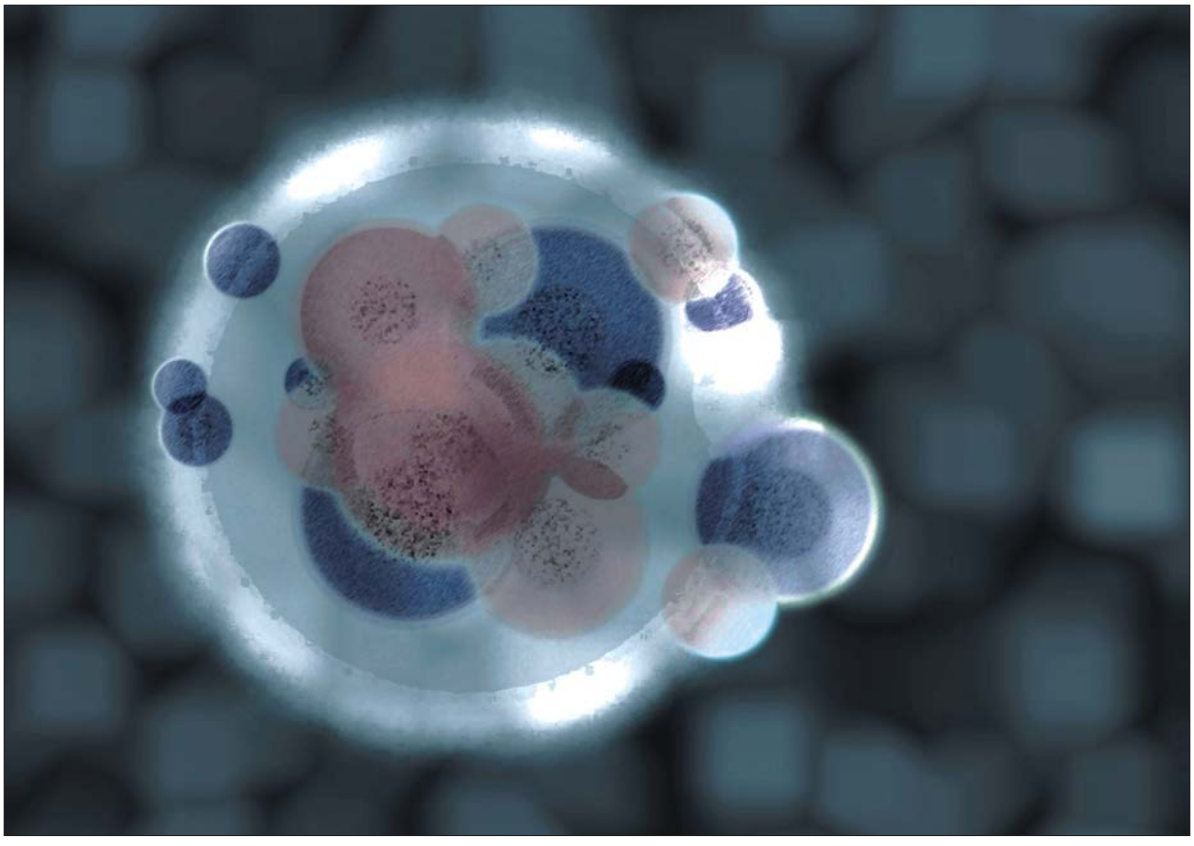

Journal of Biological Sciences

ISSN 1727-3048

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Biological Sciences 6 (4): 651-658, 2006
ISSN 1727-3048

Molecular Cloning and Characterization of cDNA Encoding Taxane 2α-O-benzoyltransferase, Catalyzing Taxol Biosynthesis from *Taxus media*

[1,2]Dongli Zhao, [1]Qian Wang, [3]Guoyin Kai, [1]Xinglong Wang, [1]Juan Lin,
[1]Yan Pi, [3]Zhiqi Miao, [1]Xiaofen Sun and [1,3]Kexuan Tang
[1]State Key Laboratory of Genetic Engineering, School of Life Sciences,
Fudan-SJTU-Nottingham Plant Biotechnology R&D Center,
Morgan-Tan International Center for Life Sciences, Fudan University,
Shanghai 200433, People's Republic of China
[2]Center for Medical Research, Medical College, Dalian University,
Dalian 116622, People's Republic of China
[3]Plant Biotechnology Research Center,
Fudan-SJTU-Nottingham Plant Biotechnology R&D Center,
School of Agriculture and Biology, Shanghai Jiao Tong University,
Shanghai 200030, People's Republic of China

Abstract: By RACE-PCR cloning, a cDNA encoding taxane 2α-O-benzoyltransferase (designated as *TmTBT*) was isolated from *Taxus media*, which catalyzes the conversion of 2-debenzoyl-7,13-diacetylbaccatin III, a semisynthetic substrate, to 7,13-diacetylbaccatin III. The full-length cDNA of *TmTBT* was 1478 bp and contained a 1320 bp Open Reading Frame (ORF) encoding a protein of 440 amino acid residues with molecular weight of 50,090 Da and an isoelectric point (pI) of 6.25, similar to taxane 2α-O-benzoyltransferase from *Taxus cuspidata*. Sequence comparison analysis revealed that *TmTBT* had high similarity with other members of plant transferase family. Phylogenetic tree analysis showed that *TmTBT* had close relationship with taxane 2α-O-benzoyltransferase from *T. cuspidata*. Tissue expression pattern analysis revealed that *TmTBT* expressed only in leaves and no expression could be detected in fruits and stems, indicating that *TmTBT* was a tissue-specific gene.

Key words: Phylogenetic tree analysis, RACE, taxane 2α-O-benzoyltransferase, *Taxus media*, *TmTBT*

INTRODUCTION

Taxol (generic name paclitaxel), a diterpenoid alkaloid produced by yew (*Taxus*) species (Baloglu and Kingston, 1999), is one of the most efficient anticancer drugs approved by FDA (Food and Drug Administration), especially for the treatment of breast and ovarian cancers (Kohler and Goldspiel, 1994). However its content in yew trees is very small, approximately 0.01% in Pacific yew bark, which limits its universal usage (Croom, 1995).

Recently, alternative methods for Taxol production, such as organic synthesis, tissue and cell culture, genetic transformation of *Taxus* species, are being tried extensively (Han *et al*., 1994; Holton *et al*., 1995; Furmanowa and Syklowska-Baranek, 2000; Shin *et al*., 2000; Yukimune *et al*., 2000; Jennewein and Croteau, 2001; Walker and Croteau, 2001; Yuan *et al*., 2002). Among them, transgenic technology is considered as one of the most potentially efficient methods for improving Taxol production. Up to now, the enzymes catalyzing chain reaction for Taxol biosynthesis have been mostly elucidated and more than 10 genes encoding these proteins including the rate-limiting steps enzymes, have been cloned (Walker and Croteau, 2001; Hezari *et al*., 1995; Wildung and Croteau, 1996; Walker and Croteau, 2000; Walker *et al*., 2002), which provides a possibility to engineer taxol biosynthetic pathway for enhancing Taxol production.

Being a key enzyme of acetyl transferases in Taxol biosynthetic pathway, taxane 2α-O-benzoyltransferase

Corresponding Author: Dr. Kexuan Tang, State Key Laboratory of Genetic Engineering, School of Life Sciences,
Fudan-SJTU-Nottingham Plant Biotechnology R&D Center,
Morgan-Tan International Center for Life Sciences, Fudan University,
Shanghai 200433, People's Republic of China
Tel: 021 65642772 Fax: 021 65643552

(*TBT*)-encoding gene has been cloned from *Taxus cuspidata* and functionally expressed in *Escherichia coli* (Walker and Croteau, 2000). Until now, there is no report on the cloning of *TBT* from *Taxus media*, a hybrid presumably arisen from *T. cuspidata* and *T. baccata* containing comparatively higher Taxol content in its needles than other *Taxus* species and as one of the major sources currently for commercial production of Taxol (Kai *et al*., 2004). Molecular cloning and characterization of *TBT* from *Taxus media* may be helpful for further understanding the biosynthetic pathway and the molecular basis for higher Taxol content in *Taxus media*. In this study, we describe the cloning and characterization of the taxane 2α-O-benzoyltransferase gene from *Taxus media* (*TmTBT*), as an initial step to investigate the physiological role of *TmTBT* in *Taxus media*. The expression pattern of *TmTBT* in various tissues including leaves, fruits and stems was also studied.

MATERIALS AND METHODS

***Taxus media* plants and RNA isolation:** *Taxus media* Rehder plants, provided by Professor Feng Tan from Southwest Normal University, China, were grown in pots in the greenhouse under 25°C with 16 h light period in 2004.

Total RNA was extracted from tissue materials (1 g) including leaves, fruits and stems of *Taxus media* plants, using cetyltrimethylammonium bromide (CTAB) based RNA isolation method as described previously (Kai *et al*., 2004; Jaakola *et al*., 2001). The quality and concentration of the extracted RNA were checked by agarose gel electrophoresis and by spectrophotometer (DU-640, Beckman, USA) analysis. The RNA samples were stored at -80°C prior to RACE and RT-PCR analysis.

3' RACE of *TmTBT*: The first strand cDNA (3'-ready cDNA) was synthesized from 5 μg of total RNA according to the manual of the SMART™ RACE cDNA Amplification Kit (CLONTECH Laboratories INC., USA) using the 3' RACE CDS Primer A (5'-AAGCAGTGGTATCAACGCAGAGTAC $(T)_{30}N_{-1}N$-3') provided within the kit. Primer TBT-F2 (5'-AGACATTAT GTACCTCCAGTTTGA-3') was designed and synthesized according to the conserved regions of benzoyltransferase cDNAs of plant origin deposited in GenBank for the cloning of the 3' cDNA ends. The 3' RACE was performed using primer TBT-F2 as the forward primer and the Universal Primer A Mix (UPM, Long: 5'-CTAATACGACTCACTATAGGGCAAGCAGTGGTA TCAACGCAGAGT-3'; Short: 5'-CTAATACGACTCACT ATAGGGC-3') as the reverse primer. PCR reaction was carried out in a total volume of 50 μL containing 2.5 μL cDNA, 20 pmol primer TBT-F2, 5 μL UPM, 41.5 μL Master Mix (34.5 μL PCR-Grade Water, 5 μL 10 x Advantage 2 PCR buffer, 1 μL 10 mM dNTP Mix, 1 μL 50 x Advantage 2 Polymerase Mix) under the following condition: the template was denatured at 94°C for 3 min followed by 35 cycles of amplification (94°C for 1 min, 56°C for 1 min, 72°C for 1 min) and by 10 min at 72°C. The amplified product was purified and cloned into pGEM-T vector (Promega, USA) and transformed into *E. coli* DH5α. Based on the color reaction using Xgal-IPTG System and PCR identification, the positive clones were picked out and sequenced by ABI 377 Sequencer (Perkin-Elmer, USA).

5' RACE and full-length cDNA cloning of *TmTBT*: The first strand cDNA (5'-ready cDNA) synthesis in 5' RACE was performed according to the manual of the SMART™ RACE cDNA Amplification Kit (CLONTECH Laboratories INC., USA) using the 5'-RACE CDS Primer (5'-$(T)_{25}N_{-1}N$-3') provided within the kit. Based on the sequence of the 3' RACE product, the complementary reverse gene specific primer TBT-R2 (5'-CTTCTCAAGATTAAGAGG TGGGTGTATG-3') was designed and synthesized. Using the 5'-ready cDNA as the template, TBT-R2 as the reverse primer and UPM as the forward primer, PCR reaction was performed in a total volume of 50 μL containing 2 μL cDNA, 10 pmol each of primers, 10 mM dNTPs, 1 x cDNA reaction buffer and 2.5U *Taq* polymerase by denaturing the cDNA at 94°C for 4 min followed by 35 cycles of amplification (94°C for 1 min, 62°C for 1 min and 72°C for 1 min) and by 7 min at 72°C. The PCR product was purified and cloned into pGEM-T vector followed by sequencing.

By comparing and aligning the sequences of the 3' RACE and 5' RACE products with Vector NTI 8.0 software, the full-length cDNA sequence of *TmTBT* was deduced, which was subsequently amplified via PCR using a pair of primers TBT-F1 (5'-GGCAGTATTGAAGG AGAAGAGAGTC-3') and TBT-R1 (5'-GAGCATACAAT AATCAGAAGTACCAA-3') and the 5'-ready cDNA as the template, in the same PCR reaction system as 5' RACE described above by denaturing the cDNA at 94°C for 4 min followed by 35 cycles of amplification (94°C for 1 min, 57°C for 1 min and 72°C for 2 min) and by 10 min at 72°C. The PCR product was purified and cloned into pGEM-T vector followed by sequencing. The PCR amplification using primers TBT-F1 and TBT-R1 under the same PCR condition as mentioned above and sequencing for the full-length cDNA of *TmTAT* was repeated three times. The full-length *TmTBT* sequence was subsequently analyzed for molecular characterization such as sequence homology, the presentation of conserved motifs, the secondary structure and molecular evolution.

Computer analyses: DNA sequences and associated molecular information were analyzed using software Vector NTI Suite 8.0 and programs from PSI-BLAST, SOPMA and CLUSTAL W1.82.

Molecular evolution analysis: Phylogenetic analysis of *TmTBT* and transferases from other *Taxus* species retrieved from GenBank was carried out by CLUSTAL W (1.82) using default parameters. A phylogenetic tree was constructed by the neighbor-joining method (Thompson *et al*., 1994).

Tissue expression pattern analysis: Semi-quantitative one-step RT-PCR was carried out to investigate the expression profile of *TmTBT* in different tissues including leaves, fruits and stems of *Taxus media*. Aliquots of total RNA (0.5 μg) extracted from leaves, fruits and stems of *Taxus media* were used as templates in one-step RT-PCR with the forward primer TBT-F1 and reverse primer TBT-R1 using one-step RNA PCR kit (Takara, Japan). Meanwhile, Two primers, 18SF (5'-ATGATAACTCGACGGATCGC-3') and 18SR (5'-CTTGGATGTGGTAGCCGTTT-3'), were also used to amplify 18S rRNA gene in the semi-quantitative RT-PCR as an internal control.

The amplifications were performed under the following condition: 50°C for 30 min, 94°C for 2 min followed by 25 cycles of amplification (94°C for 30 sec, 60°C for 30 sec and 72°C for 150 sec). The amplified products were separated on 1% agarose gel and analyzed with Gene analysis software package (Gene Company, USA).

RESULTS AND DISCUSSION

Cloning and sequence analysis of the full-length cDNA of *TmTBT*: Based on cDNA sequences of the conserved regions of plant benzoyltransferases, primer TBT-F2 was designed and synthesized for the amplification of 3' ends of *TmTBT*. A single fragment of about 900 bp was amplified using TBT-F2 and UPM and a 3' untranslated region (UTR) of 94 bp was found downstream from the stop codon. Complementary reverse specific primer TBT-R2 was subsequently designed and synthesized based on the sequence of 3' RACE product and used in the 5' RACE. A single and specific fragment of about 700 bp was amplified in which a 5' UTR of 37 bp was found upstream of the first ATG codon. By alignment and deletion of overlapping fragment with Vector NTI 8.0 software, the two fragments of 3' and 5' RACE products were assembled into a full-length cDNA, which was amplified using a pair of primer TBT-F1 and TBT-R1 and confirmed by sequencing for three times. The cloned full-length cDNA of 2α-O-benzoyltransferase from *Taxus media* (designated as *TmTBT*, Genbank accession number: AY675557) was 1478 bp with a polyA tail of 24 bp, a size with 66 bp longer than 2α-O-benzoyltransferase from *T. cuspidata* (1388 bp) after cutting off polyA sequence (Walker and Croteau, 2000). The cDNA contained a 1320 bp open reading frame encoding a protein of 440 amino acids with a calculated molecular weight of 50,090 Da and an isoelectric point of 6.25 (Fig. 1). The molecular size of *TmTBT* is very similar to that of *T. cuspidata* (50,085 Da), as well as to those of Taxadien-5α-ol O-acetyltransferases from *Taxus chinensis* (49,180 Da) (unpublished data, GenBank accession no. AAL78754), *T. cuspidata* (49,076 Da) (Walker *et al*., 2000) and *Taxus media* (49,219 Da) (Kai *et al*., 2004).

Database retrieval with PSI-Blast (Altschul *et al*., 1997) and multiple alignments of *TmTBT* with other known acyltransferases by CLUSTAL W1.82 revealed that the deduced amino acid sequence of *TmTBT* most resembled that of 2α-O-benzoyltransferase from *T. cuspidata* (*TcuTBT*, GenBank accession no. AAG38049) with 97% of identity, indicating that the key structure and functional sites of this type of transferases were strongly conserved. At amino acid level, *TmTBT* had 65-57% identity and 79-74% similarity with seven other acyl/aroyltransferases involved directly in Taxol biosynthesis, including taxadienol acetyl transferases (TATs) from *T. chinensis* (*TchTAT*, GenBank accession no. AAL78754), *T. cuspidata* (*TcuTAT*, GenBank accession no. AAF34254) and *Texus media* (*TmTAT*, GenBank accession no. AAS49031.1), phenylpropanoyltransferase (BAPT) from *T. cuspidata* (*TcuBAPT*, GenBank accession no. AAL92459.1), 10-deacetylbaccatin III 10-O-acetyltransferases (DBATs) from *T. cuspidata* (*TcuDBAT*, GenBank accession no. AAF27621) and *T. baccata* (*TbDBAT*, GenBank accession no. AAL57617.1), 3'-N-debenzoyl-2'-deoxytaxol N-benzoyltransferase (DBTNBT) from *T. canadensis* (TcaDBTNBT, GenBank accession no. AAM75818) (Fig. 2), indicating that *TmTBT* belonged to acyltransferase superfamily.

NCBI conserved domain search results indicated that *TmTBT* belonged to transferase family. Like most of the transferase enzymes of plant origin, *TmTBT* also possessed the typical acyltransferase motif HXXXDG (H158, D162, G163, Fig. 2) characteristic of other acyltransferases and the histidine residue of this element was essential for catalytic activity of these enzymes, suggesting that this histidine may function in acyl group transfer from acyl-CoA to the substrate alcohol (Walker and Croteau, 2000; Walker *et al*., 2002; Kai *et al*., 2004; Walker *et al*., 2000).

The secondary structure of *TmTBT* was analyzed by SOPMA (Geourjon and Deléage, 1995) and the result showed that the putative *TmTBT* peptide contained

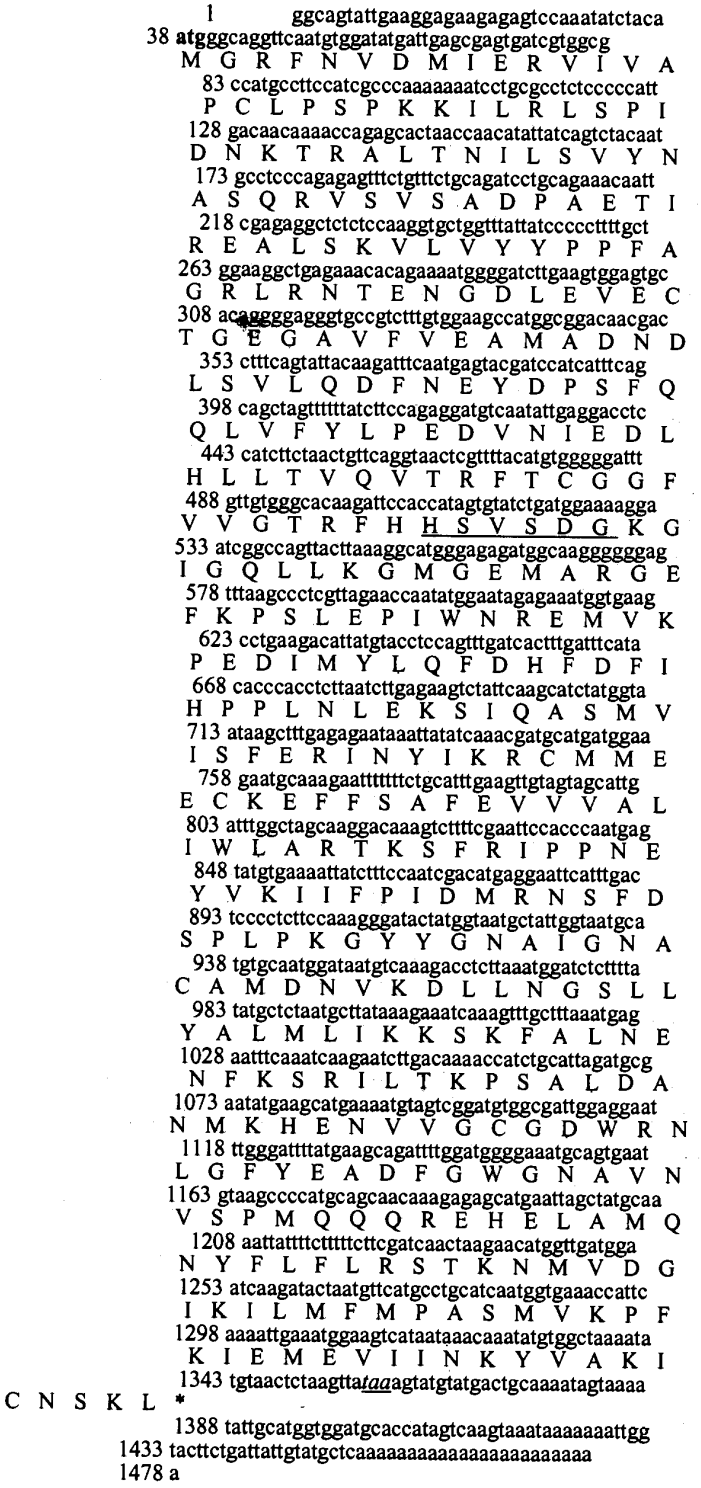

Fig. 1: The full-length cDNA sequence and deduced amino acid sequence of *Taxus media* taxane 2α-O-benzoyltransferase (*TmTBT*). The start codon (atg) was bolded and the stop codon (taa) was underlined italically. The conserved motifs were underlined

```
TchTAT          MEKT----DLHVNLNEKVMVGPSLPLPKTTLQLSSIDNLPGVRGSIFNALLIYNASPSPT 56
TmTAT           MEKT----DLHVNLIEKMMVGPSLPPPKTTLQLSSIDNLPGVRGSIFNALLIYNASPSPT 56
TcuTAT          MEKT----DLHVNLIEKVMVGPSPPLPKTTLQLSSIDNLPGVRGSIFNALLIYNASPSPT 56
TcaDBTNBT         MEKAG-STDFHVKKFDPVMVAPSLPSPKATVQLSVVDSLTICRG-IFNTLLVFNAPDN--
56
TcuDBAT          MAGS---TEFVVRSLERVMVAPSQPSPKAFLQLSTLDNLPGVRENIFNTLLVYNASDR-- 55
TbDBAT           MAGS---TESVVRSLERVMVAPSQPSPKAFLQLSTLDNLPGVRENIFNTLLVYNASDR-- 55
TmTBT            ------MGRFNVDMIERVIVAPCLPSPKKILRLSPIDNKT---RALTNILSVYNASQRVS 51
TcuTBT           ------MGRFNVDMIERVIVAPCLQSPKNILHLSPIDNKT---RGLTNILSVYNASQRVS 51
TcuBAPT          MKKTGSFAEFHVNMIERVMVRPCLPSPKTILPLSAIDNMA---RAFSNVLLVYAANMDR-
56
                        *  : ::* *.   **  : ** :*. .     : * * :: *

TchTAT
MVSADPAKLIREALAKILVYYPPFAGRLRETENGDLEVECTGEGAMFLEAMADNELSVLG 116
TmTAT
MISADPAKLIREALAKILVYYPPFAGRLRETENGDLEVECTGEGAMFLEAMADNELSVLG 116
TcuTAT          MISADPAKLIREALAKILVYYPPFAGRLRETENGDLEVECTGEGAMFLEAMADNELSVLG
116
TcaDBTNBT         -ISADPVKIIREALSKVLVYYFPLAGRLRSKEIGELEVECTGDGALFVEAMVEDTISVLR
115
TcuDBAT
-VSVDPAKVIRQALSKVLVYYSPFAGRLRKKENGDLEVECTGEGALFVEAMADTDLSVLG 114
TbDBAT
-VSVDPAKVIRQALSKVLVYYSPFAGRLRKKENGDLEVECTGEGALFVEAMADTDLSVLG 114
TmTBT
-VSADPAETIREALSKVLVYYPPFAGRLRNTENGDLEVECTGEGAVFVEAMADNDLSVLQ 110
TcuTBT
-VSADPAKTIREALSKVLVYYPPFAGRLRNTENGDLEVECTGEGAVFVEAMADNDLSVLQ 110
TcuBAPT
-VSADPAKVIREALSKVLVYYYPFAGRLRNKENGELEVECTGQGVLFLEAMADSDLSVLT 115
              :*.**.: **:**:*:**** *:***** .* *:******** *.:*:***.:  :***

TchTAT
DFDDSNPSFQQLLFSLPLDTNFKDLPLLVVQVTRFTCGGFVVGVSFHHGVCDGRGAAQFL 176
TmTAT
DFDDSNPSFQQLLFSLSLDTNFKDLPFLVVQVTRFTCGGFVVGVSFHHGVCDGRGAAQFL 176
TcuTAT
DFDDSNPSFQQLLFSLPLDTNFKDLSLLVVQVTRFTCGGFVVGVSFHHGVCDGRGAAQFL 176
TcaDBTNBT
DLDDLNPSFQQLVFWHPLDTAIEDLHLVIVQVTRFTCGGIAVGVTLPHSVCDGRGAAQFV 175
TcuDBAT          DLDDYSPSLEQLLFCLPPDTDIEDIHPLVVQVTRFTCGGFVVGVSFCHGICDGLGAGQFL
174
TbDBAT           DLDDYSPSLEQLLFCLPPDTDIEDIHPLVVQVTRFTCGGFVVGVSFCHGICDGLGAGQFL
174
TmTBT            DFNEYDPSFQQLVFYLPEDVNIEDLHLLTVQVTRFTCGGFVVGTRFHHSVSDGKGIGQLL
170
TcuTBT           DFNEYDPSFQQLVFNLREDVNIEDLHLLTVQVTRFTCGGFVVGTRFHHSVSDGKGIGQLL
170
TcuBAPT          DLDNYNPSFQQLIFSLPQDTDIEDLHLLIVQVTRFTCGGFVVGANVYGSACDAKGFGQFL
175
               *::: .**::**:*     *. ::*:  : **********:.**. .  . .*. * .*::

TchTAT           KGLAEMARGEVKLSLEPIWNRELVKLDDPKY-LQFFHFEFLRAPSIVEKIVQTYFIIDFE
235
TmTAT            KGLAEMARGEVKLSLEPIWNRELVKLDDPKY-LQFFHFEFLRAPSIVEKIVQTYFIIDFE
235
TcuTAT           KGLAEMARGEVKLSLEPIWNRELVKLDDPKY-LQFFHFEFLRAPSIVEKIVQTYFIIDFE
235
TcaDBTNBT          TALAEMARGEVKPSLEPIWNRELLNPEDPLH-LQLNQFDSICPPPMLEELGQASFVINVD
234
TcuDBAT          IAMGEMARGEIKPSSEPIWKRELLKPEDPLYRFQYYHFQLICPPSTFGKIVQGSLVITSE
234
TbDBAT           IAMGEMARGEIKPSSEPIWKRELLKPEDPLYRFQYYHFRLIRPPSTFGKIVQGSLGITSE
234
TmTBT            KGMGEMARGEFKPSLEPIWNREMVKPEDIMY-LQFDHFDFIHPPLNLEKSIQASMVISFE
229
TcuTBT           KGMGEMARGEFKPSLEPIWNREMVKPEDIMY-LQFDHFDFIHPPLNLEKSIQASMVISFE
229
TcuBAPT          QSMAEMARGEVKPSIEPIWNRELVKLEHCMP-FRMSHLQIIHAPVIEEKFVQTSLVINFE
234
               .:.******.* * ****.**::: :.     :: :: : .*    :  *  : *  :

TchTAT           TINYIKQSVMEECKEFCSSFEVASAMTWIARTRAFQIPESEYVKILFGMDMRNSFNPPLP
295
TmTAT            TINYIKQSVMEECKEFCSSFEVASAMTWIARTRAFQIPESEYVKILFGMDMRNSFNPPLP
295
TcuTAT           TINYIKQSVMEECKEFCSSFEVASAMTWIARTRAFQIPESEYVKILFGMDMRNSFNPPLP
295
TcaDBTNBT
TIEYMKQCVMEECNEFCSSFEVVAALVWIARTKALQIPHTENVKLLFAMDLRKLFNPPLP 294
TcuDBAT          TINCIKQCLREESKEFCSAFEVVSALAWIARTRALQIPHSENVKLIFAMDMRKLFNPPLS
294
TbDBAT           TIKWIKQCLREESKEFCSAFEVVSALAWIARTRALQIPHSENVKLIFAMDMRKLFNPPLL
294
TmTBT            RINYIKRCMMEECKEFFSAFEVVVALIWLARTKSFRIPPNEYVKIIFPIDMRNSFDSPLP 289
TcuTBT           RINYIKRCMMEECKEFFSAFEVVVALIWLARTKSFRIPPNEYVKIIFPIDMRNSFDSPLP 289
```

Fig. 2: Continued

```
TcuBAPT          IINHIRRRIMEERKESLSSFEIVAALVWLAKIKAFQIPHSENVKLLFAMDLRRSFNPPLP 294
                 *: ::: : ** :*  *:**:. *: *:*: ::::** .* **.:* :*:*. *:.**

TchTAT           SGYYGNSIGTACAVDNVQDLLSGSLLRAIMIIKKSKVSLNDNF-KSRAVVKPSELDVNMN
354
TmTAT            SGYYGNSIGTACAVDNVQDLLSGSLLRAIMIIKKSKVSLNDNF-KSRAVVKPSELDVNMN
354
TcuTAT           SGYYGNSIGTACAVDNVQDLLSGSLLRAIMIIKKSKVSLNDNF-KSRAVVKPSELDVNMN
354
TcaDBTNBT
NGYYGNAIGTAYAMDNVQDLLNGSLLRAIMIIKKAKADLKDNYSRSRVVTNPYSLDVNKK 354
TcuDBAT           KGYYGNFVGTVCAMDNVKDLLSGSLLRVVRIIKKAKVSLNEHF-TSTIVTPRSGSDESIN
353
TbDBAT            KGYYGNFVGTVCAMDNVKDLLSGSLLRVVRIIKKAKVSLNEHF-TSTIVTPCSGSDESIN
353
TmTBT
KGYYGNAIGNACAMDNVKDLLNGSLLYALMLIKKSKFALNENF-KSRILTKPSALDANMK 348
TcuTBT
KGYYGNAIGNACAMDNVKDLLNGSLLYALMLIKKSKFALNENF-KSRILTKPSTLDANMK 348
TcuBAPT          HGYYGNAFGIACAMDNVHDLLSGSLLRTIMIIKKSKFSLHKEL-NSKTVMSSSVVDVNTK
353
                 ***** .* . *:***:***.**** .: :***:*  *:..   *  :      * . :

TchTAT
HENVVAFADWSRLGFDEVDFGWGNAVSVSPVQ--QQCELAMQNYFLFLKPSKNKPDGIKI 412
TmTAT
HENVVAFADWSRLGFDEVDFGWGNAVSVSPVQ--QQCELAMQNYFLFLKPSKNKPDGIKI 412
TcuTAT
HENVVAFADWSRLGFDEVDFGWGNAVSVSPVQ--QQSALAMQNYFLFLKPSKNKPDGIKI 412
TcaDBTNBT
SDNILALSDWRRLGFYEADFGWGGPLNVSSLQR-LENGLPMFSTFLYLLPAKNKSDGIKL 413
TcuDBAT
YENIVGFGDRRRLGFDEVDFGWGHADNVSLVQHGLKDVSVVQSYFLFIRPPKNNPDGIKI 413
TbDBAT
YENIVGFGDRRRLGFDEVDFGWGHADNVSLVQHGLKDVSVVQSYFLFIRPPKNNPDGIKI 413
TmTBT
HENVVGCGDWRNLGFYEADFGWGNAVNVSPMQQQREHELAMQNYFLFLRSTKNMVDGIKI 408
TcuTBT
HENVVGCGDWRNLGFYEADFGWGNAVNVSPMQQQREHELAMQNYFLFLRSAKNMIDGIKI 408
TcuBAPT
FEDVVSISDWRHSIYYEVDFGWGDAMNVSTMLQQQEHEKSLPTYFSFLQSTKNMPDGIKM 413
                 ::::: .*  .  : *.***** . .** :    :    : . * :: .**  ****:

TchTAT           LM-FLPLSKMKSFKIEMEAMMKKYVAKV----- 439
TmTAT            LM-FLPLSKMKSFKIEMEAMMKNYVAKV----- 439
TcuTAT           LM-FLPLSKMKSFKIEMEAMMKKYVAKV----- 439
TcaDBTNBT         LLSCMPPTTLKSFKIVMEAMIEKYVSKV----- 441
TcuDBAT          LS-FMPPSIVKSFKFEMETMTNKYVTKP----- 440
TbDBAT           LS-FMPPLIMKSFKFEMETMTNKYVTKP----- 440
TmTBT            LM-FMPASMVKPFKIEMEVIINKYVAKICNSKL 440
TcuTBT           LM-FMPASMVKPFKIEMEVTINKYVAKICNSKL 440
TcuBAPT          LM-FMPPSKLKKFKIEIEAMIKKYVTKVCPSKL 445
                 *   :*   :* **: :*.  ::**:*
```

Fig. 2: Multiple alignment of *TmTBT* with eight most homologous acyltransferases. The symbols "*", ":" and "." represented completely identical, conservative and semi-conservative amino acid residues respectively among the aligned sequences. *TchTAT*: *T. chinensis* taxadienol acetyltransferase (AAL78754). *TmTAT*: *Taxus media* taxadienol acetyltransferase (AAS49031). *TcuTAT*: *T. cuspidata* taxadienol acetyltransferase (AAF34254). TcaDBTNBT: *T. canadensis* 3'-N-debenzoyltaxol N-benzoyltransferase (AAM75818). *TcuDBAT*: *T. cuspidata* 10-deacetylbaccatin III-10-O-acetyl transferase (AAF27621). *TbDBAT*: *T. baccata* 10-deacetylbaccatin III-10-O-acetyl transferase (AAL57617). *TcuTBT*: *T. cuspidata* taxane 2α-O-benzoyltransferase (AAG38049). *TcuBAPT*: *T. cuspidata* phenylpropanoyltransferase (AAL92459). The typical acyltransferase motif HXXXDG in eight of the nine sequences was shaded

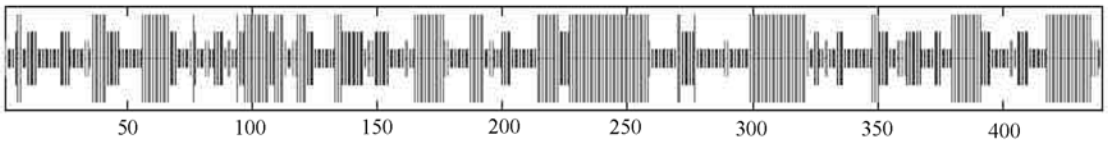

Fig. 3: Prediction of the secondary structure of *TmTBT*. The helix, sheet, turn and coil were indicated respectively with the longest, second longest, medium and the shortest vertical lines

36.59 helix, 19.09 extended strand (sheet), 7.95 β turn and 36.36% random coil (Fig. 3). The random coil and alpha helix constituted interlaced domination of the main part of the secondary structure. From the above sequence analyses, *TmTBT* was found to have many characteristics common to the acyltansferase superfamily. This implies

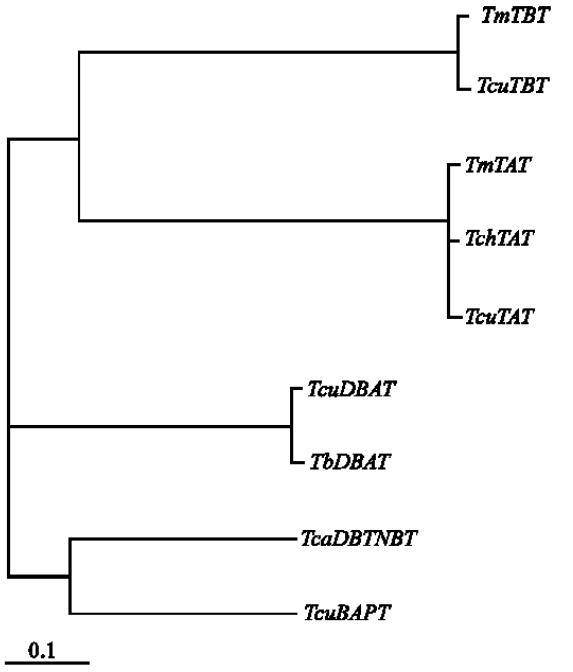

Fig. 4: Phylogenetic tree analysis of *TmTBT* and other acyltransferases from *Taxus*

that the conserved motifs may play an important role in the biological functions and thus are preserved in evolution, while some variations on non-conserved domains can provide the molecular foundation for the diversity of protein structures and functions. The X-ray crystal structure analysis will further help to elucidate the detailed structure of *TmTBT* in the future.

Molecular evolution analysis: To investigate the evolutionary relationships among transferases in *Taxus*, a phylogenetic tree was constructed based on the deduced amino acid sequences of *TmTBT* and transferases from other *Taxus* species. The result showed that *TmTBT* and *TcuTBT* were grouped into a cluster with the shortest distance (group 1) (Fig. 4). TATs (*TchTAT*, *TcuTAT* and *TmTAT*) formed another cluster (group 2). DBATs (*TbDBAT* and *TcuDBAT*) naturally formed the third cluster (group 3). *T. cuspidata* phenylpropanoyl transferase (*TcuBAPT*) and *T. cuspidata* 2-debenzoyl-7,13-diacetylbaccatin III-2-O-benzoyl transferase (TcuDBTNBT) formed the forth cluster (group 4). The four groups of transferases were derived from a common ancestor in the evolution, suggesting that transferases in *Taxus* including TBTs share a common evolutionary origin based on their conserved characteristics such as sequence homology and conserved domain motifs (Fig. 2).

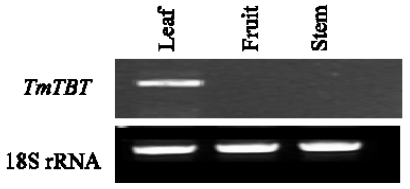

Fig. 5: Expression profile of *TmTBT* in different *Taxus media* tissues. Total RNA (0.5 μg/sample) was isolated from leaf, fruit and stem tissues respectively and subjected to one-step RT-PCR amplification (upper panel). 18S rRNA was used as the control to show the normalization of the amount of templates in PCR reaction (lower panel)

Tissue expression pattern analysis: To investigate *TmTBT* expression pattern in various tissues of *Taxus media,* total RNA was isolated from leaves, fruits and stems and subjected to one-step RT-PCR analysis using the primers TBT-F1 and TBT-R1. The result showed *TmTBT* expression could be detected only in leaves with strong signal, but no expression was detected in fruits or stems (Fig. 5). Therefore, the *TmTBT* was considered to be a tissue specific gene, similar to *TmTAT* reported previously (Kai *et al*., 2004).

In conclusion, molecular cloning and characterization of *TmTBT* showed that *TmTBT* was very similar to other *TBTs* of plant origin, indicating that *TmTBT* belonged to TBT superfamily. The purification of *TmTBT* protein from *Taxus media* and associated analysis will further elucidate the structure and function of *TmTBT*. Based on the cloning and characterization of *TmTBT*, plant expression vector containing *TmTBT* has been constructed and genetic transformation of *Taxus media* is undergoing in order to test its potential role in improving Taxol production by genetic engineering.

ACKNOWLEDGMENTS

This study was funded by China National 863 High-tech Program (No. 2005AA212191), China Ministry of Education and Shanghai Science and Technology Committee.

REFERENCES

Altschul, S.F., T.L. Madden, A.A. Schaffer, J. Zhang, W. Miller and D.J. Lipman, 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. Nucl. Acids Res., 25: 3389-3402.

Baloglu, E. and D.G.I. Kingston, 1999. The taxane diterpenoids. J. Natl. Prod., 62: 1448-1472.

Croom, E.M.J., 1995. *Taxus* for Toxol and Toxids. In: Taxol Science and Applications. Suffness, M. (Ed.), CRC Press, Boca Raton, pp: 37-70.

Furmanowa, M. and K. Syklowska-Baranek, 2000. Hairy root cultures of *Taxus media* var. Hicksii Rehd. as a new source of paclitaxel and 10-deacetylbaccatin III. Biotechnol. Lett., 22: 683-686.

Geourjon, C. and G. Deléage, 1995. SOPMA: Significant improvement in protein secondary structure prediction by consensus prediction from multiple alignments. Cabios, 11: 681-684.

Han, K.H., P. Fleming, K. Walker, M. Loper, W.S. Chilton, U. Mocek, M.P. Gordon and H.G. Floss, 1994. Genetic transformation of mature *Taxus*: An approach to genetically control the *in vitro* production of the anticancer drug. Taxol. Plant Sci., 95: 187-196.

Hezari, M., N.G. Lewis and R. Croteau, 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.

Holton, R.A., C. Somoza, H.B. Kim, F. Liang, R.J. Biediger, P.D. Boatman, M. Shindo, C.C. Smith and S. Kim, 1995. The total synthesis of paclitaxel starting with camphor. ACS. Symp. Ser., 583: 288-301.

Jaakola, L., A.M. Pirtila, M. Halonen and A. Hohtola, 2001. Isolation of high quality RNA from bilberry (*Vaccinium myrtillus* L.) fruit. Mol. Biotechnol., 19: 210-203.

Jennewein, S. and R. Croteau, 2001. Taxol: biosynthesis, molecular genetics and biotechnological applications. Applied Microbiol. Biotechnol., 57: 13-19.

Kai, G., Miao, C. Qiu, L. Zhang, L. Zhao, Z. Li, T. Xu, L. Zhang, Y. Gong, D. Zhao, D. Liu, X. Sun and K. Tang, 2004. Molecular cloning and characterization of a taxadienol acetyl transferase cDNA from *Taxus media*. Plant Sci., 167: 759-764.

Kohler, J. and B.R. Goldspiel, 1994. Evaluation of new drug Paclitaxel (Taxol). Pharmacotherapy, 14: 3-34.

Shin, S.W., Y.S. Kim and S. Lim, 2000. Elicitors for the regulation of baccatin III biosynthesis in plant cell culture system. Yakhak. Hoechi., 44: 60-65.

Thompson, J.D., D.G. Higgins, T.J. Gibson and R.D. Wolfinger, 1994. Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucl. Acids. Res., 22: 4673-4680.

Walker, K. and R. Croteau, 2001. Molecules of interest: Taxol biosynthetic genes. Phytochemistry, 58: 1-7.

Yuan, Y.J., Z.J. Wei, Z.Q. Miao and J.C. Wu, 2002. Acting paths of elicitors on taxol biosynthesis pathway and their synergistic effect. Biochem. Eng. J., 10: 77-83.

Yukimune, Y., Y. Hara, E. Nomura, H. Seto and S. Yoshida, 2000. The configuration of methyl jasmonate affects paclitaxel and baccatin III production in *Taxus* cells. Phytochemistry, 54: 13-17.

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

Walker, K., A. Schoendorf and R. Croteau, 2000. Molecular cloning of a Taxa-4(20),11(12)-dien-5a-ol-O-Acetyl transferase cDNA from *Taxus* and functional expression in *Escherichia coli*. Arch. Biochem. Biophys., 374: 371-380.

Walker, K., R. Long and R. Croteau, 2002. The final acylation step in Taxol biosynthesis: Cloning of the taxoid C13-side-chain N-benzoyltransferase from *Taxus*. Proc. Natl. Acad. Sci. USA., 99: 9166-9171.

Wildung, M.R. and R. Croteau, 1996. A cDNA cloing for taxadiene synthase, the diterpene cyclase, that catalyzes the committed step of taxol biosynthesis. J. Biol. Chem., 271: 9201-9204.