Biochemical and Biophysical Research Communications 443 (2014) 938–943

Contents lists available at ScienceDirect

Biochemical and Biophysical Research Communications

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

# Molecular cloning and characterization of a cytochrome P450 taxoid 9á-hydroxylase in *Ginkgo biloba* cells

Nan Zhang [a,1], Zhentai Han [b,1], Guiling Sun [c,1], Angela Hoffman [d], Iain W. Wilson [e], Yanfang Yang [a], Qian Gao [b], Jianqiang Wu [c], Dan Xie [f], Jungui Dai [f], Deyou Qiu [a,*]

[a] State Key Laboratory of Tree Genetics and Breeding, The Research Institute of Forestry, Chinese Academy of Forestry, Beijing 100091, China
[b] Key Laboratory of Forest Ecological Environment of Ministry of Forestry, Research Institute of Forest Ecology Environment and Protection, Chinese Academy of Forestry, Beijing 100091, China
[c] Key Laboratory of Economic Plants and Biotechnology, Kunming Institute of Botany, Chinese Academy of Sciences, Kunming 650201, China
[d] Department of Chemistry, University of Portland, Portland, OR 97203, USA
[e] CSIRO Plant Industry, PO Box 1600, Canberra, ACT 2001, Australia
[f] State Key Laboratory of Bioactive Substances and Functions of Natural Medicines, Institute of Materia Medica, Chinese Academy of Medical Sciences & Peking Union Medical College, Beijing 100050, China

ARTICLE INFO

*Article history:*
Received 4 December 2013
Available online 28 December 2013

*Keywords:*
Taxol
Taxoid 9α-hydroxylase
*Ginkgo biloba*
CYP450
Cell-free protein synthesis system
LC–MS

ABSTRACT

Taxol is a well-known effective anticancer compound. Due to the inability to synthesize sufficient quantities of taxol to satisfy commercial demand, a biotechnological approach for a large-scale cell or cell-free system for its production is highly desirable. Several important genes in taxol biosynthesis are currently still unknown and have been shown to be difficult to isolate directly from *Taxus*, including the gene encoding taxoid 9α-hydroxylase. *Ginkgo biloba* suspension cells exhibit taxoid hydroxylation activity and provides an alternate means of identifying genes encoding enzymes with taxoid 9α-hydroxylation activity. Through analysis of high throughput RNA sequencing data from *G. biloba*, we identified two candidate genes with high similarity to *Taxus* CYP450s. Using *in vitro* cell-free protein synthesis assays and LC–MS analysis, we show that one candidate that belongs to the CYP716B, a subfamily whose biochemical functions have not been previously studied, possessed 9α-hydroxylation activity. This work will aid future identification of the taxoid 9α-hydroxylase gene from *Taxus sp*.



## 1. Introduction

Taxol was first extracted from the Pacific Yew bark in 1963 by American chemists Wani and Wall [1]. Its significant antineoplastic effect on ovarian, breast and other cancers is accomplished by a unique mechanism of preventing tubulin dissociation [1]. Given its important application in medicine, natural production of taxol is unable to meet demand, and new methods for commercial production are needed. Understanding the biosynthesis of this important metabolite is a prerequisite for directly manipulating Taxol yields in *in vitro* and *in vivo* systems. Taxol biosynthesis in *Taxus* involves 19 steps beginning with geranylgeranyl diphosphate [2], but only 13 of these taxol biosynthetic genes have been identified. Importantly, the genes responsible for C9 oxidation, C1 hydroxylation, oxetane formation, and C2′ hydroxylation of taxol biosynthetic pathway in *Taxus* sp, are unknown.

Taxoid 9α-hydroxylation is thought to occur early in the biosynthesis pathway [3]. It was reported that a putative candidate for the cytochrome P450 taxoid 9a-hydroxylase had been identified [4], however, the sequence for the clone was not released and no enzymatic validation was performed. Given the difficulties in identifying genes from *Taxus sp*, alternative systems for studying taxoid synthesis have been explored. Hu et al. [5] found that filamentous fungi could specifically hydroxylate taxane skeletons at 1β and 9α position, similar to *Taxus chinensis* suspension cells that preferentially produce 9α-hydroxylation rather than 9β-hydroxylation. Dai et al. [6] found that *Ginkgo biloba* suspension cells could hydroxylate 2α,5α,10β,14β- tetraacetoxy-4(20),11-taxadiene (sinenxan A, SIA) at 9α position with a yield close to 70% (Fig. 1) [7]. Thus, *G. biloba* cells possess taxoid 9α-hydroxylase activity and could provide an alternate source for obtaining genes with Taxoid 9α-hydroxylase activity. Using RNA sequencing data from *G. biloba*, candidate genes with high similarity to a *Taxus* CYP450 were identified. Using *in vitro* cell free protein synthesis assays, we demonstrate by LC–MS analysis that one candidate possessed 9α-hydroxylation activity. The identification of a *G. biloba* gene encoding a protein with taxoid 9α-hydroxylase activity will

* Corresponding author.
*E-mail address:* qiudy@caf.ac.cn (D. Qiu).
[1] These authors contributed equally to this work.

0006-291X/$ - see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.bbrc.2013.12.104

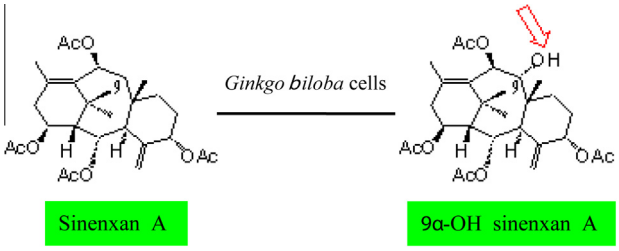


**Fig. 1.** *G. biloba* cells can transform sinenxan A (SIA) to 9OH-SIA.

aid efforts to increase semi-synthetic synthesis of taxol as well as identify the equivalent gene in *Taxus*.

## 2. Materials and methods

### 2.1. Enzymes, substrates, and reagents

*G. biloba* cells, substrate SIA and the standard of 9OH-SIA were provided by Prof. Jungui Dai from the Chinese Academy of Medical Sciences and Peking Union Medical College. Enzymes and vectors were purchased from Tiangen Biotech Co., Ltd. cDNA Synthesis Kit was purchased from Invitrogen (Carlsbad, CA). Gel Extraction Kits and Plasmid Extraction Kits were obtained from Axgen (Tewksbury, MA). DNA marker, Restriction Endonucleases and T4 DNA Ligase were obtained from Takara (Japan). Other chemicals were purchased from Sigma (St. Louis, MO).

### 2.2. Identification and cloning of 9α-hydroxylase candidates

*G. biloba* transcriptome datasets from our own RNA-sequencing work with *G. biloba* cells [8], Roche 454 sequencing dataset of *G. biloba* leaf tissue generated by Dr. Shilin Chen's group at the Institute of Medicinal Plant Development, China (Accession: SRX022356) and *G. biloba* transcriptome datasets generated by the Prof. C.R. Buell laboratory at the Department of Plant Biology, Michigan State University [9] were used in our study. *G. biloba* secondary stem RNA-Seq data (Accession: SRX087427), *G. biloba* sterile seedling RNA-Seq data (Accession: SRX087425), *G. biloba* ripe fruit (with seed) RNA-Seq data (Accession: SRX087424), *G. biloba* lateral roots RNA-Seq data (Accession: SRX087422 and *G. biloba* mature leaf (fully expanded) RNA-Seq data (Accession: SRX087421) were downloaded from GenBank. The data in sra format were converted to data in fastaq format using the fasta-dump in NCBI SRA Toolkit. The 454 transcriptome dataset (SRX022356) was also downloaded from GenBank and assembled using MIRA [10] with default settings. All of the Illumina transcriptome datasets were separately assembled with ABySS [11]. The assembly with ABySS used the optimal parameter for k-mer as determined for each dataset. The assembly data from each dataset were merged into one assembly using CAP3 [12]. The candidate CYP450s were obtained by BLAST search (tblastn) of this final assembly against the *Taxus* known p450s that are involved in taxol biosynthesis with *E* value set to $1e^{-5}$.

To amplify the candidate p450 clones identified by bioinformatics analysis, total RNA was extracted from cultured *G. biloba* cells using the CTAB procedure from Chang [13]. For contig5382, one specific pair of primers was designed for its amplification: 5′-GACATTCACATCCTCCAGTAGT-3′ (Forward)/5′-CCATAGAGGTCCTACATTCATC-3′ (Reverse). Another specific pair of primers was designed for amplification of contig5926: 5′-GAGATGAGTATTTTGTCGAGGATTA-3′ (Forward)/5′-CCTCTATACGATTGGGATCGGTACA-3′ (Reverse).

The fragments were amplified by PCR conditions of 94 °C for 5 min, then 30 cycles at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 4 min, and finally, 72 °C for 10 min. The reaction system was 25 μl containing a mixture of 12.5 μl 2 × Pfu PCR Master Mix, 1 μl primer (10 μM) respectively, 2 μl cDNA and 8.5 μl sterile dd$H_2O$. According to the instructions of DNA gel extraction kit, PCR products were purified and undertook the second round of PCR in the same program except for changing Tm to 58 °C. The resulting amplicon was excised from a gel after electrophoresis, gel purified, added dA to the blunting end and cloned into pGM-T (Tiangen, Beijing). This was used to transform *Escherichia coli* TOP 10 F' cells, which were selected on ampicillin. Positive transformants were grown for plasmid preparation and sequence verification. pGM-T plasmids containing right cDNA inserts were used as PCR template. A subsequently designed primer set was used for second round PCR amplification of contig5382 (5′-CGGGGTACCCCGTCAATCCCTTGATTCAAGAT-3′/5′-TTGGATATCCAATGGTTTGGAGTGGGGAGAC-3′). While for second PCR amplification of contig5926, the following primer set was used: 5′-TTGGATATCCAAATGGGTATTTTGTTGTGGAT-3′/5′-CCGCTCGAGCGGTCAGGATCTGGGAAACAATTTG-3′.

The second round of PCR amplification of the fragments was done using the following PCR conditions; 94 °C for 5 min; 25 cycles at 94 °C for 30 s, 55 °C for 30 s, 72 °C for 3 min. and finally, 72 °C for 10 min. The reaction system was 25 μl consisting of; 12.5 μl 2 × Pfu PCR Master Mix, 1 μl primer (10 μM) respectively, 0.5 μl plasmids and 10 μl sterile dd$H_2O$. Directionally ligated pEU-E01-MCS vector with directional ligation was used to transform *E. coli* TOP 10 F' cells by established conventional methods. The two clones were constructed into a vector for cell-free expression.

### 2.3. Cell free protein expression

Cell free protein expression of the two candidate genes was performed using the Premium PLUS Expression Kit (CellFree Sciences Co., Ltd, Japan), following the manufacturer's instruction.

### 2.4. SDS–PAGE analysis

10 ml 12% SDS–polyacrylamide separating gel and 10 ml SDS–polyacrylamide stacking gel were prepared for protein electrophoresis. 10 μl expression products or control samples was mixed with 2 × SDS–PAGE loading buffer respectively and boiled for 5 min prior to loading. 10 μl protein marker and 20 μl sample was loaded into wells by tips. Gel electrophoresis was run at a constant voltage of 20 V. The gel was put into 25 ml Coomassie Blue staining solution and shaken for 5–10 min in order stain thoroughly. After removing the staining solution, the gel was rinsed with distilled water and then destaining solution was added until bands were visualized.

### 2.5. Enzymatic reaction and analytical procedures

The cell free synthesis expression products derived from KF773141 and KF773140 and vector only control (no gene insert) were incorporated into to same enzymatic reaction mixtures. Assays were performed in 10 ml centrifugal tubes containing 0.5 mg of the taxoid substrate sinenxan A (SIA), 10 μl DMSO as solvent, 450 μl (225 μl per tube × 2) cell free synthesis expression products, 0.25 g glucose-6-phosphate, 3 unit of glucose-6-phosphate dehydrogenase, 2.3 mg FAD, 0.9 mg FMN, 22.3 mg NADPH in a total volume of 1 ml of Tris–HCl (pH7.5). Following incubation for 70 min in the dark at 32 °C with gentle shaking, the reaction mixture was saturated with 1 ml NaCl and mixed gently by pipetting. Then extracted with 2 ml of hexane:ethyl acetate (4:1, v/v) and blended fully. This mixture was centrifuged at 5000 rpm for 5 min. The supernatant was transferred to a new tube and the extraction and centrifugation were repeated. All the supernatant (~4 ml) was collected and freeze dried for 30 min. Finally 100 μl

acetonitrile were added and transferred to sample vials for LC–MS analyses.

### 2.6. LC–MS analyses

LC–MS analyzes were performed on a Thermo Scientific LTQ FT LC–MS system using a SunFire TM column (4.6 mm inner diameter × 150 mm, 5 μm particle size, C18). The temperature of the column was constant at 40 °C. The mobile phase was acetonitrile in water using a $H_2O$:$CH_3CN$ gradient from 5% to 100% CH3CN at 1 ml/min over 50 min, followed by 5 min hold at 100%. The flow rate was 1 ml per minute. The sample injection volume was 10 μl.

### 2.7. Phylogenetic analysis

Homolog sequences were collected from plants and bacteria by a similarity search against NCBI non-redundant (nr) database. Multiple sequence alignment was performed using ClustalX2 [14] and the ambiguous regions and gaps were removed. The most optimal model of protein substitution matrix with rate heterogeneity was determined by ModelGenerator (v_851) [15]. The phylogenetic tree was reconstructed with PHYML 3.0 [16]. Bootstrap analyses used 100 pseudo-replicates. The tree topology structure was viewed and edited with NJplot [17].

## 3. Results

### 3.1. 9α-Hydroxylase candidate identification and sequence analyses

To find potential CYP450s candidates with 9α-hydroxylation activity in *G. biloba*, RNA sequence libraries of our own [8] and those publically available were combined and assembled into one large dataset. This large transcriptome assembly was then interrogated by performing BLAST searches (tblastn) against known *Taxus* P450s that are involved in taxol biosynthesis (Supplemental Table 1). A total of 325 assembled sequences were found that possessed significant similarities to these proteins (Supplemental Table 2). Two candidates (contig5926 and contig5382) with the greatest similarities to *Taxus* P450s from our *G. biloba* cell line [8] (*E* value < $10^{-147}$ and $10^{-120}$ respectively) were selected for further characterization. The sequences of contig5382 and contig5926 are both full length, and have been deposited in GenBank (accession numbers KF773141-KF773140, respectively). The ORF of KF773141 is 1458 bp and encodes a 485-residue hydroxylase with a calculated molecular weight of 54,842 and exhibited characteristic sequence elements of cytochrome P450 including a signal peptide (amino acids 1–27), a typical N-terminal membrane anchor (amino acids 10–28), the oxygen binding domain, the reductase binding domain, the conserved PSRF (amino acids 407–410, generally is PERF) motif, the highly conserved PFG element binding to heme (amino acids 422–424), the ETLR salt bridge (amino acids 353–356, the motif is EXXR), and the essential cysteine at position 430 (Supplemental Fig. 1). The ORF of KF773140 is 1512 bp and it codes for a 503-residue hydroxylase with a calculated molecular weight of 57,202 and exhibits characteristic sequence elements of a cytochrome P450 including a typical N-terminal membrane anchor (amino acids 21–43), the oxygen binding domain, the reductase binding domain, the conserved PSRF (amino acids 425–428, generally is PERF) motif, the highly conserved PFG element binding to heme (amino acids 441–443), the ETLR salt bridge (amino acids 371–374, the motif is EXXR), and the essential cysteine at position 449 (Supplemental Fig. 2). Phylogenetic analysis comparing the two candidate genes to 39 known plant and bacterial taxoid hydroxylases revealed that KF773140 and KF773141 grouped with the known CYP725A and CYP716B subfamily of P450s and thus we designated them as GbCYP725A and GbCYP716B respectively (Supplemental Fig. 3).

### 3.2. Protein gel electrophoresis analyses

The candidate 9α-hydroxylation clones were cloned and *in vitro* transcribed and expressed using a cell free protein expression system based on wheat germ extracts. Protein products were separated on a SDS–PAGE gel followed by Coomassie blue staining to visualize protein bands. GFP and DHFR clones inserted into the expression vector pEU-E01 vector were included used as positive controls (~27 KDa and 23 KDa respectively), and an empty pEU-E01 vector was used as a negative control. Based on the BSA reference standard, the concentration of KF773141 and KF773140 expression products was ~150 ng/μl. The size of the expressed protein of the clone KF773141 was as expected to be ~55 KD (Supplemental Fig. 4).

### 3.3. LC–MS assay and product identification

To determine whether the proteins encoded by KF773140 and KF773141 possessed taxoid 9α-hydroxylase activity, the expressed proteins were added to a reaction mixture containing the substrate, taxoid sinenxan A (SIA). The molecular weight of SIA is 504 with a retention time of 39.77 min which was clearly detected in the total HPLC diagram (Supplemental Fig. 5). $[M + NH_4]^+$ and $[M + Na]^+$ ions for SIA were observed simultaneously, while $[M + H]^+$could not be seen in LC–MS spectra. The molecular ion peaks of SIA in LC/MS were *m/z* 522 $[M + NH_4 + OH\text{-}H]^+$, and 527 $[M + Na + OH\text{-}H]^+$ (Supplemental Fig. 6). It generated typical MS/MS peaks *m/z* 467 $[M + Na\text{-}AcO + OH\text{-}H]^+$, 407 $[M + Na\text{-}2AcO + OH\text{-}H]^+$, and 347 $[M + Na\text{-}3AcO + OH\text{-}H]^+$ (Supplemental Fig. 7). Proteins with 9α-hydroxylase activity should result in the addition of an OH group at the C9 position to SIA, a chemical change that can be detected using LC–MS analysis. The product of the reaction catalyzed by the protein translated from KF773141 was analyzed on a LC–MS, and was found to have a retention time of 31.86 min, which corresponded to the standard substance 9OH-SIA (retention time 31.84 min) (Fig. 2). The empty vector controls did not yield detectable product at the same retention time (Supplemental Fig. 8). Again, $[M + NH_4]^+$ and $[M + Na]^+$ ions for 9OH-SIA were observed simultaneously, while $[M + H]^+$ could not be seen in LC–MS spectra. The molecular ion peaks of 9OH-SIA in LC/MS were *m/z* 538 $[M + NH_4 + OH\text{-}H]^+$, and 543 $[M + Na + OH\text{-}H]^+$ (Fig. 3A). Since 9OH-SIA has four acetyl groups, it generated typical MS/MS peaks *m/z* 483 $[M + Na\text{-}AcO + OH\text{-}H]^+$, 423 $[M + Na\text{-}2AcO + OH\text{-}H]^+$, and 363 $[M + Na\text{-}3AcO + OH\text{-}H]^+$ (Fig. 4A). The data from MS and MS/MS analysis of 9OH-SIA (Figs. 3A and 4A) and the product of the protein translated from KF773141 (Figs. 3B and 4B) clearly indicated that KF773141 has taxoid 9α-hydroxylase activity in *G. biloba* cells. The reaction product catalyzed by the protein encoded by KF773140 did not match the standard substance 9OH-SIA (Supplemental Fig. 9). Base on this result it appears that the expression product of KF773140 did not have taxoid 9α-hydroxylase activity.

## 4. Discussion

The result of our SDS–PAGE analysis and enzyme assay indicated that KF773141 can be successfully expressed in a wheat germ cell-free protein synthetic system and the encoded enzyme could 9α-hydroxylate taxoid SIA. The *in vivo* bioactivity of the taxoid 9α-hydroxylase of KF773141-encoding enzyme (GbCYP716B) in *G. biloba* cells may require additional cofactors or proteins that are required to achieve its full enzyme activity. Further studies

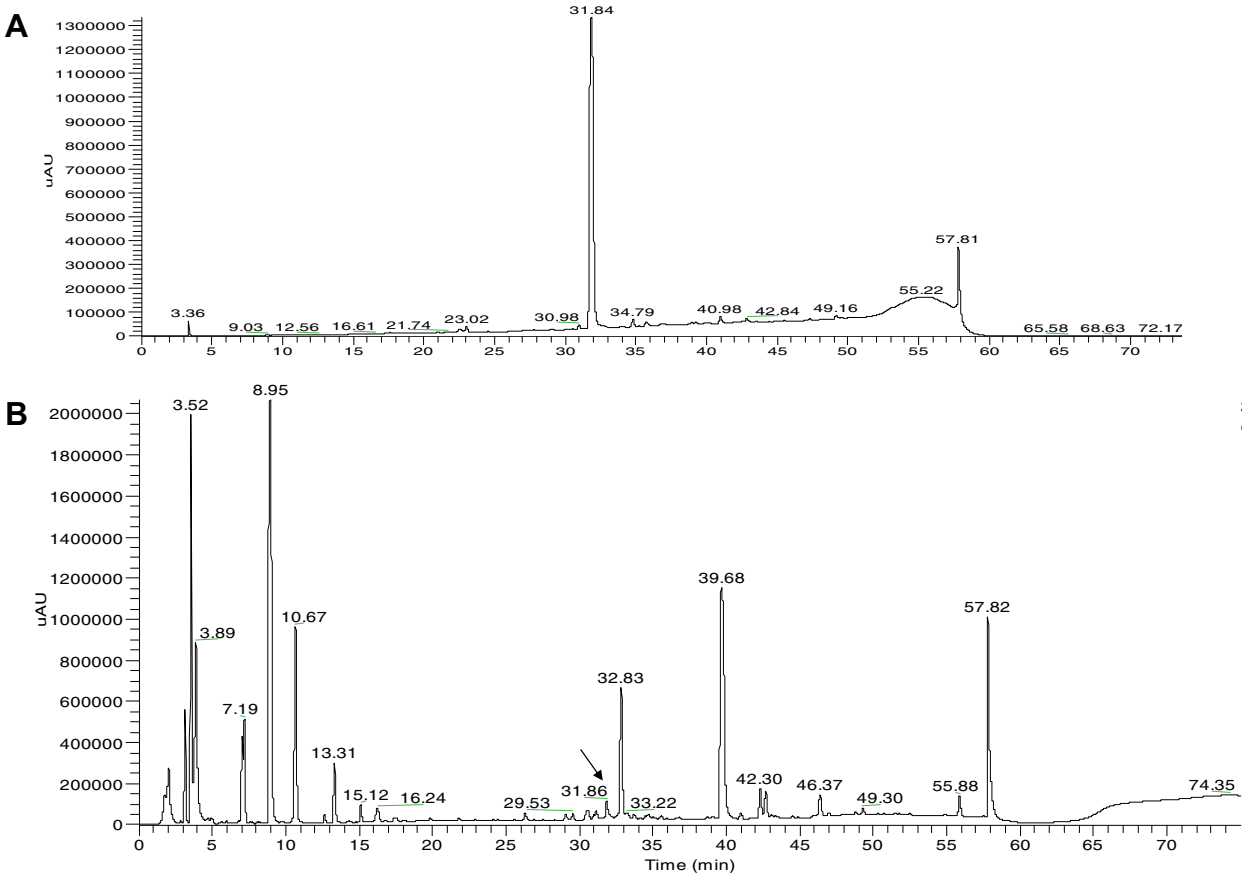


**Fig. 2.** The HPLC chromatogram of 9OH-SIA and the product catalyzed by the enzyme encoded by KF773141. (A) 9OH-SIA (B) the product catalyzed by the enzyme encoded by KF773141.

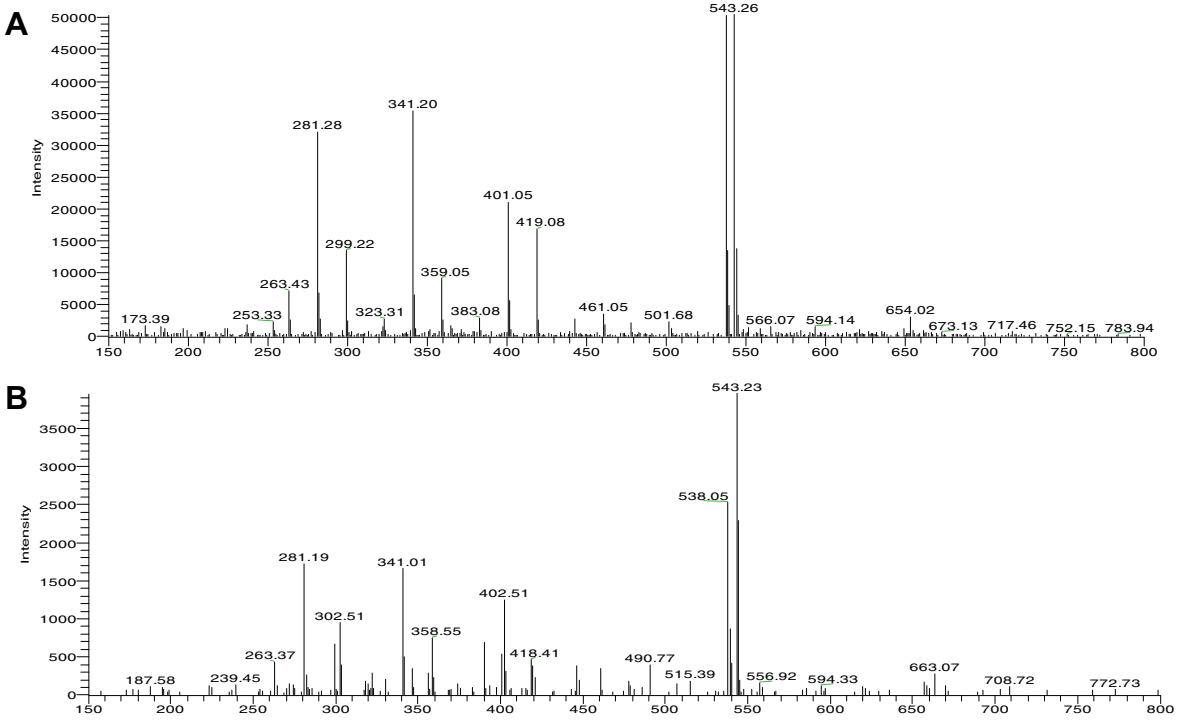


**Fig. 3.** Spectra of ion fragments in MS analysis of 9OH-SIA and the product catalyzed by the enzyme encoded by KF773141. (A) 9OH-SIA (B) the product catalyzed by the enzyme encoded by KF773141.

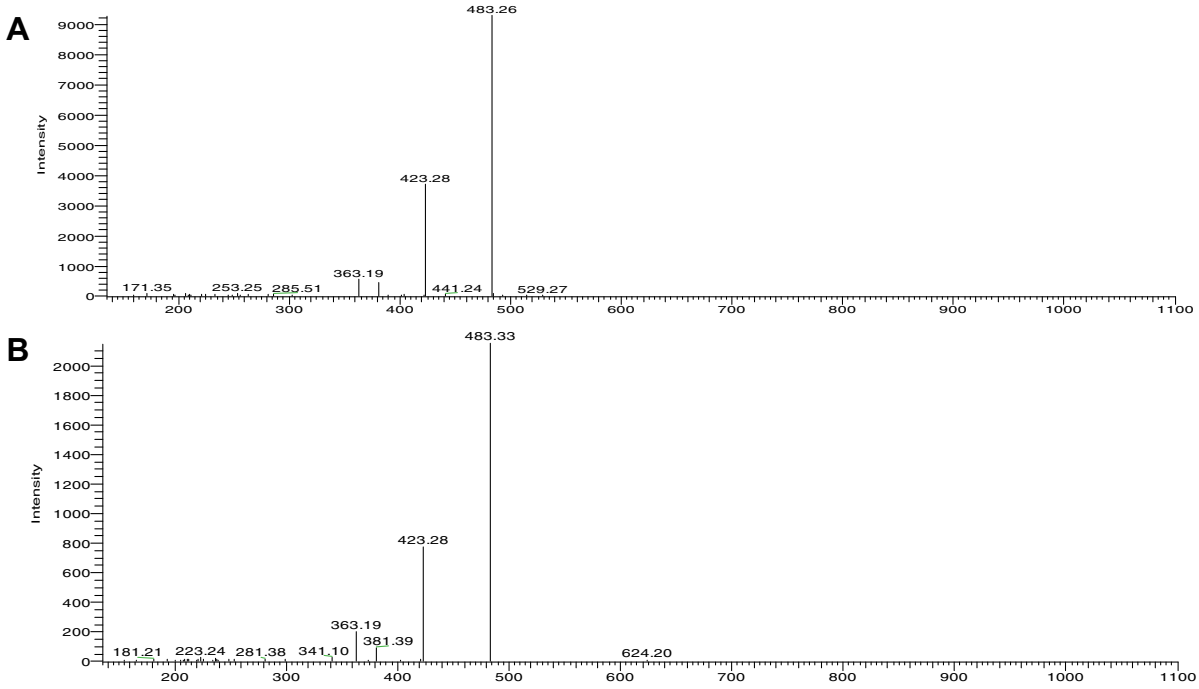


**Fig. 4.** Spectra of ion fragments in MS/MS analysis of 9OH-SIA and the product catalyzed by the enzyme encoded by KF77314. (A) 9OH-SIA (B) the product catalyzed by the enzyme encoded by KF77314.

are required to understand the conditions necessary for optimal taxoid 9α-hydroxylase activity *in vitro* with GbCYP716B.

In biosynthesis of taxol, cytochrome P450-mediated oxygenations plays a major role such that approximately one-half of the proposed 19 distinct enzymatic steps of the pathway are considered to be catalyzed by cytochrome P450 oxygenases [18]. These taxoid hydroxylases share more than 70% sequence identity but less than 35% similarity to other plant P450s. These P450s belong to CYP725 subfamily, which specifically occurs only in *Taxus* species [19]. KF773141 belongs to a CYP716-like subfamily which has high homology with CYP716B2 and CYP716A12. It has been previously reported that there is some overlap in the CYP716 and CYP725 families [20]. Specifically, CYP716B1 of *Picea sitchensis* (Sitka spruce) is 44% identical to CYP725A4 (taxadiene 5-α hydroxylase) in taxol biosynthesis in *Taxus canadensis* (Canadian yew) and some other CYP725A sequences [21,22]. The 44% sequence identity between CYP716B sequences of unknown function and CYP725A sequences that act on taxane diterpenoids suggests there are some similarities in their substrates. Based on the known evolutionary relationship, CYP716 and CYP725 have a close phylogenetic distance that both belong to the CYP85 group. In the taxol biosynthesis pathway, CYP725A is exclusively found in the *Taxus* genus. The CYP725 family has some overlap with the more common CYP716 family and seems to be an offshoot of the much older CYP716 family [23]. Recent research had shown that some CYP716 subfamily genes can catalyze oxidation steps in the biosynthesis pathway of some plant secondary metabolites. For example, *Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins [24], and CYP716A47 catalyzes the formation of protopanaxadiol during ginsenoside biosynthesis in *Panax ginseng* [25]. The CYP716B sequence, which was obtained after the CYP725 family was named, now appears to bridge CYP716A and CYP725. To date, no gene of CYP716B subfamily has been functionally characterized.

In this research, we demonstrated that GbCYP716B has the hydroxylation function at 9α-C position of taxoid. Our result also reveals the molecular mechanism underlying the specific hydroxylation activity of converting SIA to 9OH-SIA by *G. biloba* suspension cells. Our successful identification of the taxoid 9α-hydroxylase in *G. biloba* will facilitate the cloning of taxoid 9α-hydroxylase from *Taxus* in the near future. Furthermore, this new gene will also benefit the communities working towards mass production of taxol from plants and those improving the production of taxol with genetically engineered microbes.

## Acknowledgment

This work was supported by a grant for the National non-profit Research Institutions of Chinese Academy of Forestry (CAFYBB2012042).

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.bbrc.2013.12.104.

## References

[1] M.E. Wall, M.C. Wani, Camptothecin and taxol: discovery to clinic-thirteenth bruce F. cain memorial award lecture, Cancer Res. 55 (1995) 753–760.

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

[3] H.G. Floss, U. Mocek, Biosynthesis of taxol, in: M. Suffness (Ed.), Taxol-Science and Applications, CRC Press, Boca Raton, FL, USA, 1995, pp. 191–208.

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

[5] S.S. Hu, D.A. Sun, X.F. Tian, Q.C. Fang, Regio-and stereo-selective hydroxylation of taxoids by filamentous fungi, Chirality 14 (2002) 495–497.

[6] J.G. Dai, H.Z. Guo, D.D. Lu, J.H. Zheng, D.A. Guo, Biotransformation of 2a,5a,10β,14β, tetra-acetoxy-4(20),11-taxadiene by Ginkgo cell suspension cultures, Tetrahedron Lett. 42 (2001) 4677–4679.

[7] J.G. Dai, M. Ye, H.Z. Guo, W.H. Zhu, D.Y. Zhang, H. Qiu, J.H. Zheng, D.A. Guo, Substrate specificity for the hydroxylation of polyoxygenated 4(20),11-taxadienes by Ginkgo cell suspension cultures, Bioorg. Chem. 31 (2003) 345–356.

[8] N. Zhang, G. Sun, J.G. Dai, Y.F. Yang, H.W. Liu, D.Y. Qiu, Sequencing and analysis of the transcriptome of *Ginkgo biloba* L. cells, China, Biotechnology 33 (2013) 112–119.
[9] E. Gongora, K. Childs, J. Hamilton, B. Vaillancourt, C.R. Buell, Genomics approaches for biochemical pathway discovery in medicinal plant species, Pharm. Biol. 50 (2012). 665-665.
[10] B. Chevreux, T. Pfisterer, B. Drescher, A.J. Driesel, W.E. Müller, T. Wetter, S. Suhai, Using the miraEST assembler for reliable and automated mRNA transcript assembly and SNP detection in sequenced ESTs, Genome Res. 14 (2004) 1147–1159.
[11] J.T. Simpson, K. Wong, S.D. Jackman, J.E. Schein, S.J. Jones, I. Birol, ABySS: a parallel assembler for short read sequence data, Genome Res. 19 (2009) 1117–1123.
[12] X. Huang, A. Madan, CAP3: a DNA sequence assembly program, Genome Res. 9 (1999) 868–877.
[13] S. Chang, J. Puryear, J. Cairney, A simple and efficient method for isolating RNA from pine trees, Plant Mol. Biol. Rep. 2 (1993) 113–116.
[14] J.D. Thompson, T.J. Gibson, F. Plewniak, F. Plewniak, F. Jeanmougin, D.G. Higgins, The CLUSTAL_X windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools, Nucleic Acids Res. 25 (1997) 4876–4882.
[15] T.M. Keane, C.J. Creevey, M.M. Pentony, T.J. Naughton, J.O. Mclnerney, Assessment of methods for amino acid matrix selection and their use on empirical data shows that ad hoc assumptions for choice of matrix are not justified, BMC Evol. Biol. 6 (2006) 29.
[16] S. Guindon, O. Gascuel, A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood, Syst. Biol. 52 (5) (2003) 696–704.
[17] G. Perriere, M. Gouy, WWW-query: an on-line retrieval system for biological sequence banks, Biochimie 78 (5) (1996) 364–369.
[18] R. Kaspera, R. Croteau, Cytochrome P450 oxygenases of taxol biosynthesis, Phytochem. Rev. 5 (2006) 433–444.
[19] M. Mizutani, D. Ohta, Diversification of P450 genes during land plant evolution, Annu. Rev. Plant Biol. 61 (2010) 291–315.
[20] D. Nelson, D.W. Reichhart, A P450-centric view of plant evolution, Plant J. 66 (2011) 194–211.
[21] S. Jennewein, H. Park, J.M. DeJong, R.M. Long, A.P. Bollon, R.B. Croteau, Coexpression in yeast of Taxus cytochrome P450 reductase with cytochrome P450 oxygenases involved in Taxol biosynthesis, Biotechnol. Bioeng. 89 (2005) 588–598.
[22] D. Rontein, S. Onillon, G. Herbette, A. Lesot, D. Werck-Reichhart, C. Sallaud, A. Tissier, CYP725A4 from yew catalyzes complex structural rearrangement of taxa-4(5),11(12)-diene intothe cyclic ether 5(12)-oxa-3(11)-cyclotaxane, J. Biol. Chem. 283 (2008) 6067–6075.
[23] B. Hamberger, J. Bohlmann, Cytochrome P450 mono-oxygenases in conifer genomes: discovery of members of the terpenoid oxygenase superfamily in spruce and pine, Biochem. Soc. Trans. 34 (2006) 1209–1214.
[24] M. Carelli, E. Biazzi, F. Panara, A. Tava, L. Scaramelli, A. Porceddu, N. Graham, M. Odoardi, E. Piano, S. Arcioni, S. May, C. Scotti, O. Calderini, *Medicago truncatula* CYP716A12 is a multifunctional oxidase involved in the biosynthesis of hemolytic saponins, Plant Cell 23 (2011) 3070–3081.
[25] J.Y. Han, H.J. Kim, Y.S. Kwon, Y.E. Choi, The cyt P450 enzyme CYP716A47 catalyzes the formation of protopanaxadiol from dammarenediol-II during ginsenoside biosynthesis in *Panax ginseng*, Plant Cell Physiol. 12 (2011) 2062–2073.