Molecular Cloning and Characterization of a Novel Stem-specific Gene from *Camptotheca acuminata*

Yan Pi, Zhihua Liao, Yourong Chai, Hainian Zeng, Peng Wang, Yifu Gong, Yongzhen Pang, Xiaofen Sun and Kexuan Tang

1State Key Laboratory of Genetic Engineering, Institute of Genetics, School of Life Sciences, Morgan-Tan International Center for Life Sciences, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, Fudan University, Shanghai 200433, People’s Republic of China
2Plant Biotechnology Research Center, Fudan-SJTU-Nottingham Plant Biotechnology R&D Center, School of Agriculture and Biology, Shanghai Jiaotong University, Shanghai 200030, People’s Republic of China

Received 12 September 2005, Accepted 30 October 2005

In higher plants, P450s participate in the biosynthesis of many important secondary metabolites. Here we reported for the first time the isolation of a new cytochrome P450 cDNA that expressed in a stem-specific manner from *Camptotheca acuminata* (designated as CaSS), a native medicinal plant species in China, using RACE-PCR. The full-length cDNA of CaSS was 1735 bp long containing a 1530 bp open reading frame (ORF) encoding a polypeptide of 509 amino acids. Bioinformatic analysis revealed that CaSS contained a heme-binding domain PFGXGRRXCX and showed homology to other plant cytochrome P450 monooxygenases and hydroxylases. Southern blotting analysis revealed that there was only one copy of the CaSS present in the genome of *Camptotheca acuminata*. Northern blotting analysis revealed that CaSS expressed, in a tissue-specific manner, highly in stem and lowly in root, leaf and flower. Our study suggests that CaSS is likely to be involved in the phenylpropanoid pathway.

Keywords: *Camptotheca acuminata*, CaSS, Cytochrome P450, RACE

Database Accession No: AY466856

Abbreviations: CTAB: Cetyl trimethyl ammonium bromide, ORF: open reading frame, PCR: Polymerase chain reaction, RACE: rapid amplification of cDNA ends, RT-PCR: reverse transcriptase-polymerase chain reaction.

*To whom correspondence should be addressed.
Tel: +86-21-65642772, Fax: +86-21-65643552
Email: kxtang1@yahoo.com or kxtang1@163.com

Introduction

Cytochrome P450 (P450 or CYP) occurs in all living organisms, from archaea to humans, and oxygenates organic compounds. In higher plants, P450 plays crucial roles in the biosynthesis of lignins, fatty acids, gibberellic acids, brassinosteroids, jasmonic acid, flavonoids, cyanogenic glucosides, and alkaloids (Yamada et al., 2000). The scanning and analysis of the whole Arabidopsis thaliana genome reveal that more than 270 cytochrome P450 genes clustered in 45 CYP families exist in this species (http://drnelson.utmem.edu/CytochromeP450.html), most of which have not yet been functionally characterized (Matsumoto et al., 2002).

In recent years, the phenylpropanoid pathway, through which some important natural products such as lignins and flavonoids in vascular plants are ultimately synthesized, has attracted considerable attention (Croteau et al., 2000). A phylogenetic analysis of P450 in *A. thaliana* indicates that the *CYP98* family is most closely related to *CYP73A5*, which is supported by the high frequency of *CYP98* ESTs detected in a variety of plants species and tissues. *CYP73A5* encodes cinnamic-acid 4-hydroxylase, the second enzyme and the first P450, in the phenylpropanoid pathway. The *CYP98* family of cytochrome P450 genes encodes the 3'-hydroxylase of coumaroyl esters, which catalyzes an essential step in the synthesis of lignin monomers and chlorogenic acid (Schoch et al., 2001; Franke et al., 2002a; Franke et al., 2002b). Engineering the P450 family in plants has important agronomic applications, including enhancement of plants’ defense against microbial and herbivore attack, and modification of lignin composition to improve forage digestibility and wood pulping (Baucher et al., 1998; Boudet, 2000). Access to *CYP98* genes from major crops, forage plants and most common woody species is the initial step for modifying *CYP98* expression. At present, there is a very limited
understanding of how both monolignol formation and lignin biosynthesis in developing vascular plants are regulated, particularly of how metabolic flux (carbon allocation) in the pathway and lignin compositions are controlled (Lewis et al., 1999). Furthermore, the lack of detailed knowledge of pathway regulation has led to the speculations that every enzymatic step in the monolignol biosynthesis pathway to the lignins has either a “key” or regulatory role (Croteau et al., 2000; Gowri et al., 1991; Goffner et al., 1994). The cloning and characterization of the genes involved in monolignol formation and lignin biosynthetic pathway would greatly help to reinforce the knowledge in higher plants. So far, only a few P450 genes involved in the phenylpropanoid pathway in woody species have been isolated.

In this study, *Camptotheca acuminata*, a kind of Chinese traditional medicinal plant, was chosen as the model tree species in our plant research program. We hope to use it to extend our knowledge on phenylpropanoid pathway. In this study we report the cloning and characterization of a novel cytochrome P450 gene (*CaSS*) from *Camptotheca acuminata*, the Chinese Happy tree with pharmaceutical use. Bioinformatic analysis and molecular modeling indicated that the *CaSS* was probably a hydroxylase-like cytochrome P450 gene involved in the phenylpropanoid pathway. The expression pattern of *CaSS* in various tissues of *C. acuminata* was also investigated.

**Materials and Methods**

**Plant materials.** Mature seeds of *Camptotheca acuminata* were collected from the plants growing in the campus of Fudan University, Shanghai, China. The seeds were pretreated with 75% alcohol for 1 min, washed with double sterile distilled water for 3 times, followed by treating the seeds with 0.1% HgCl$_2$ for 1 min. After 4 rinses with double sterile distilled water, the sterilized seeds were incubated at 30°C for 3 min and in a box with high GC content and complex secondary structure. The 3’ cDNA end primary amplification was performed using 2 µl of RT products and 2.5 units of Tag DNA polymerase, mixed with the primers FCASS3-1 and AUAP (5'-GGCCACCACGCGTCACTGAC-3'), as the universal amplification primer, other component were applied in a standard reaction as recommended. The PCR reaction was performed under the following conditions: cDNA was denatured at 94°C for 3 min followed by 35 cycles of amplification (94°C for 45 sec, 55°C for 1 min, 72°C for 1 min) and by 72°C for 10 min. The PCR product was checked by agarose gel electrophoresis, and then diluted 50-fold and used as the template for nested PCR amplification using primers FCASS3-2 and AUAP under the following PCR condition: the template was denatured at 94°C for 3 min followed by 28 cycles of amplification (94°C for 1 min, 55°C for 1 min, 72°C for 1 min) and by 72°C for 10 min. The PCR products were purified and cloned into pGEM-T Easy vector (Promega, USA), followed by sequencing.

**5’ cDNA end amplification of CaSS.** The 5’ RACE System for Rapid Amplification of cDNA Ends (GIBCO BRL) was used for 5’ cDNA end amplification of *CaSS*. According to the sequencing result of 3’ cDNA end, three primers were designed and synthesized for 5’ RACE of *CaSS*. RT, dCTailing and PCR amplifications were conducted as the standard protocol with minor modifications. The 5 μg of total RNA was reversely transcribed using primer RCASS5-1 (5'-ATTGGATATGGCTTGTTGCG-3') with an extra 30 min of RT at 30°C after standard RT at 42°C. Primers RCASS5-2 (5'-CATAGTGGTGGTGGAGGTG-3') and AAAP (5'-GGCCACGGC TTCAGACTGACGTACG-3') were used for primary amplification with 5 µl of cT-tailed cdNA as template in a total volume of 30 µl reaction mixture under the following PCR condition: the cdNA was denatured at 94°C for 3 min followed by 30 cycles of amplification (94°C for 1 min, 55°C for 1 min, 72°C for 2 min) and by 72°C for 10 min. After confirmation on agarose gel electrophoresis, the PCR product was diluted 50-fold and used as the template for nested PCR amplification using primers RCASS5-3 (5'-GGGGTCTTGGTGGGTGGAGGTG-3') and AUAP (5'-GGCCACGGC TTCAGACTGACGTACG-3') were used for primary amplification with 5 µl of cT-tailed cdNA as template in a total volume of 30 µl reaction mixture under the following PCR condition: the cdNA was denatured at 94°C for 3 min followed by 28 cycles of amplification (94°C for 1 min, 62°C for 1 min, 72°C for 2 min) and by 72°C for 10 min. The PCR products were purified and cloned into pGEM-T Easy vector (Promega, USA), followed by sequencing.

**The full-length cdNA amplification of CaSS.** The sequences of 3’ and 5’ cdNA end products were assembled to generate a putative full-length cdNA of *CaSS*. A pair of PCR primers, FCASS (5'-AATCCTCTTCTGCTAAGTGGCTAC-3') and RASS (5'-GGCCATAGAACCCACCATTTACTCAA-3'), were designed for the amplification of full-length cdNA according to the assembled...
sequence of CaSS. PCR was carried out in a total volume of 50 μl reaction solution containing 5 μl 10 x LA Taq buffer (plus Mg2+), 2 μl 10 mM each of dNTPs, 2 μl 10 μM FCASS, 2 μl 10 μM RCASS, 2 μl cDNA (3' RACE RT product) and 2.5 units LA Taq DNA polymerase (TaKaRa, Japan) using the following protocol: 94°C for 3 min followed by 30 cycles of amplification (94°C for 1 min, 58°C for 1 min, 72°C for 2 min) and by 72°C for 10 min.

Sequencing and bioinformatic analysis. After agarose gel electrophoresis detection and gel extraction, the DNA of target Sequencing and bioinformatic analysis. After agarose gel electrophoresis detection and gel extraction, the DNA of target

To investigate the

Northern blotting analysis. Aliquots of DNA (20 µg) were digested overnight at 37°C with BglII, BamHI, XhoI, SacI and EcoRI, which did not cut within the probe region, respectively, fractionated by 0.8% agarose gel electrophoresis, and transferred to a positively charged Hybond-N+ nylon membrane (Amersham Pharmacia). The 780 bp long nest-amplified 3'RACE cDNA end was used as template for probe labeling using primers FCASS2-1 and RCASS. The PCR condition and amplification procedure were the same as those for 3'end nested PCR. Probe labeling (biotin), hybridization and signal detection were performed using Gene images random prime labeling module and CDP-Star detection module following the manufacturer's instructions (Amersham Pharmacia). The filter was washed under high stringency conditions (65°C) and the hybridized signals were visualized by exposure to Fuji X-ray film at room temperature for 1.5 h.

Southern blotting analysis. Southern blot was carried out to analyze the copy number of CaSS present in C. acuminate genome. Total genomic DNA was isolated from 1.00 g of fresh C. acuminate leaf materials following the standard protocol (Rechults, 1999). Aliquots of DNA (20 µg) were digested overnight at 37°C with BglII, BamHI, XhoI, SacI and EcoRI, which did not cut within the probe region, respectively, fractionated by 0.8% agarose gel electrophoresis, and transferred to a positively charged Hybond-N+ nylon membrane (Amersham Pharmacia). The 780 bp long nest-amplified 3'RACE cDNA end was used as template for probe labeling using primers FCASS2-1 and RCASS. The PCR condition and amplification procedure were the same as those for 3'end nested PCR. Probe labeling (biotin), hybridization and signal detection were performed using Gene images random prime labeling module and CDP-Star detection module following the manufacturer's instructions (Amersham Pharmacia). The filter was washed under high stringency conditions (65°C) and the hybridized signals were visualized by exposure to Fuji X-ray film at room temperature for 1.5 h.

Northern blotting analysis. To investigate the CaSS expression pattern in different parts of C. acuminate, total RNAs were extracted from root, stem, leaf and flower, and used in Northern blotting analysis. Aliquots (20 mg/sample) of total RNA were denatured and separated on a 1% agarose gel containing 0.66 M formaldehyde, blotted onto a positively charged Hybond-N+ nylon membrane (Amersham Pharmacia) and hybridised with the same biotin-labeled probe used in Southern analysis according to the instructions of Gene Images module (Amersham Pharmacia). The hybridized signals were visualized by exposure to Fuji X-ray film at room temperature for 1.5 h.

Results and Discussion

Amplification of 3' cDNA end of CaSS. Agarose gel analysis revealed that amplification with primers FCASS3-1 and AUAP resulted in a DNA band of about 900 bp accompanied by some faint bands and smearing. Nested amplification of the 3' end resulted in a bright band of about 850 bp which was also accompanied by 2 faint bands and some smearing. When the 20-bp AUAP primer sequence was cut away, a fragment of 817 bp 3' cDNA end with 43 bp of poly (dA) tail was generated. BLAST-n analysis indicated that it had wide similarities to known hydroxylase/cytochrome P450 sequences from O. basilicum, S. scutellarioides, A. thaliana, S. indicum, P. taeza, implying that it was probably a part of a hydroxylase-like cytochrome P450 gene.

Amplification of 5' cDNA end of CaSS. Electrophoresis analysis of primary PCR product of 5' RACE showed a specific band of about 1.2 kb. Nested amplification was carried out directly using this PCR product as the template and a specific band of about 1.1 kb was amplified. By cutting away the 20 bp AUAP and 17 bp poly (dG) sequences, the resulting 5' cDNA end product of CaSS turned out to be 1032 bp. BLAST of this sequence resulted in similar outcome as that from the 3' cDNA end BLAST. Based on the 71 bp overlap between the 5' and the 3' fragments, the full-length CaSS cDNA of 1735 bp (poly (dA) not included) was deduced and a pair of PCR primers, FCASS and RCASS, was designed for the full-length cDNA amplification.

The full-length cDNA amplification and nucleotide analysis of CaSS. As anticipated, the specific 1735 bp DNA band was amplified by PCR using primers FCASS and RCASS. Sequencing result of this sequence completely coincided with the deduced full-length cDNA sequence (Fig. 1). The cDNA of CaSS possessed a 1530-bp open reading frame (ORF) from 41 bp to 1570 bp of the sequence, except for a 40-bp 5'UTR and a 165-bp 3'UTR. The 3'UTR possesses typical low G+C content and a putative polyadenylation signal AAATAA was detected 15 bp upstream from the polyadenylation site.

Nucleotide-nucleotide BLAST of the CaSS full-length cDNA sequence on NCBI website indicated that its conserved segments had similarities to G max cytochrome P450 monooxygenase CYP98A2p and hydroxylase genes from O. basilicum, S. indicum, and L. erythrobacteur. Vector NTI Suite 6.0 pairwise alignments of CaSS cDNA with the above homologous genes on the whole molecule level resulted in similarities of about 55% to 65%.

Characterization of the CASS protein. The ORF was obtained by ORF Finder on NCBI (http://www.ncbi.nlm.nih.gov/orf/)
A Novel Cytochrome P450 Gene from *Camptotheca acuminata*

71

The full-length cDNA sequence and deduced amino acid sequence of Cytochrome P450 gene (*CaSS*) from *Camptotheca acuminata*. The start codon (ATG) was underlined and the stop codon (TAG) was underlined italically. The heme-binding sequence PFGXGRRXCX and putative polyadenylation signal AAATAA were boxed. The primers RCAAS and FCAAS were underlined, with the direction of priming shown by the arrows. The upright arrow indicated the cleavage site of the deduced signal peptide between K18 and L19.

---

**Southern blotting analysis.** It was reported earlier that all P450 genes in the phenylpropanoid pathway that had been characterized so far in *A. thaliana* were present in single copy.

**Note:** The pagination and section markers have been removed for clarity. The text continues with analysis and results related to the gene sequence and its expression. Further details include the length, molecular weight, and amino acid composition, as well as BLAST results for similarity to other plant proteins. The gene is predicted to be a signal peptide, and its three-dimensional structure is predicted using various computational tools.
in the genome (Schoch et al., 2001). It was an unusual situation. In order to investigate the copy number of CaSS present in *C. acuminata* genome, Southern blot analysis was performed under high stringency condition. Only one hybridization band was observed in five different restriction enzyme-digested DNA lanes, indicating that there was only

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Sorghum bicolor</td>
<td>(1) MDASLLL SILAVI LPSLALNLUR GRP GP GP PVLRIQ RO</td>
</tr>
<tr>
<td>Lithospermum erythrorhizon</td>
<td>(1) MALP AIA IFLI SS KIYDKRLK LKP GP PLIENIYO</td>
</tr>
<tr>
<td>Sesamum indicum</td>
<td>(1) MALP LLISFTLFIAYKLFGLRY KLP GP PVLRIQ RO</td>
</tr>
<tr>
<td>Glycine max</td>
<td>(1) MAAL IIIPL SVLTLW GLOTLYR LF KLP GP PVLRIQ RO</td>
</tr>
<tr>
<td>Ocimum basilicum</td>
<td>(1) MALL IIIPL SVLTLW GLOTLYR LF KLP GP PVLRIQ RO</td>
</tr>
<tr>
<td>Solenostemon scutellarioides</td>
<td>(1) MGIVPFFP VYI TAVKYL VEM RGP GP PVLRIQ RO</td>
</tr>
<tr>
<td>Arabidopsis thaliana</td>
<td>(1) MGVPFPP F VYI TAVKYL VEM RGP GP PVLRIQ RO</td>
</tr>
<tr>
<td>Camptotheca acuminata</td>
<td>(1) MGVPFPP F VYI TAVKYL VEM RGP GP PVLRIQ RO</td>
</tr>
<tr>
<td>Consensus</td>
<td>(1) MALL I LI T LSYKLYR LF KLPGP GRP PVLRIQ RO</td>
</tr>
</tbody>
</table>

Fig. 2. Multi-alignment made by VNTI Suite 6 of the deduced CaSS amino acid sequence with the 8 most homologous plant cytochrome P450 proteins from *Sorghum bicolor* (AAC39316), *Lithospermum erythrorhizon* (BAC4436), *Sesamum indicum* (AA47545), Glycine max (AAC94587), Ocimum basilicum (AA92000), Solenostemon scutellarioides (CAD20576), Arabidopsis thaliana (NP_830337), Pinus taeda (AA47685). The completely identical amino acids were dark shaded and the hemebinding sequence PFGXGRXXC was underlined and also boxed. The conserved amino acid sequences IGLL W DMI and VEW AMAEL equivalent to the two degenerated homologous PCR primers for 3'cDNA amplification of CaSS were shown in grey boxes.
Fig. 2. Continued.
one copy of the CaSS present in the genome of C. acuminata (Fig. 4), which is similar to the counterpart in A. thaliana.

**Northern blotting analysis.** Among the many roles lignin plays in plant growth and development are those providing structural support for land plants. Among tissues expressing high levels of CYP98 mRNA were poplar and pine xylem (Sterky et al., 1998; Allona et al., 1998), soybean hypocotyl and stem, cotton fibers, as well as A. thaliana inflorescence stems and wounded tissues (Schoch et al., 2001). Such special tissues were involved in stem and branches to ensure that the desired orientations relative to the gravitational vector were attained. In order to investigate if CaSS expression was correlated to the lignin biosynthesis pathway in C. acuminata plant, Northern blotting was performed. The result showed...
that the expression level of CaSS was much higher in stem than in root, leaf and flower (Fig. 5). It was close to other land plant such as poplar, pine, soybean, cotton and A. thaliana (Sterky et al., 1998; Allona et al., 1998; Schoch et al., 2001). Expression of CaSS in C. acuminata is high in lignin-synthesizing tissues. All these results imply that the expression of CaSS might influence the growth and development of the stems such as lignin biosynthesis.

In summary, we have successfully cloned a new gene encoding cytochrome P450, a committed-step enzyme involved in lignin biosynthesis, from C. acuminata. In earlier attempts to identify the other P450 cDNAs involved in phenypropanoid pathways, the coding sequences were expressed in yeast, such as C4H in P. taeda, CYP98A3 in A. thaliana and CYP98A13 in O. basilicum to identify their hydroxylase activities (Anterola et al., 2002; Schoch et al., 2001; David et al., 2002). Considering the similarity between hydroxylases, CaSS will be expressed in yeast in a similar way to further characterize the enzymatic activity of CASS in the future.

Acknowledgments This work was funded by China National “863” High-Tech Program, China Ministry of Education and Shanghai Science and Technology Committee.

References


