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

Yan Pi¹, Zhihua Liao¹, Yourong Chai¹, Hainian Zeng¹, Peng Wang¹, Yifu Gong²,
Yongzhen Pang¹, Xiaofen Sun¹ and Kexuan Tang^{1,2,*}

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

Introduction

Cytochrome P450 (P450 or CYP) occurs in all living organisms, from archaebacteria to human beings, 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://dmnelson.utm.edu/CytochromeP450.html>), most of which have not yet been functionally characterized (Matsuno *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 agroindustrial 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

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

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; Goffiner *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% HgCl₂ for 15 min. After 4 rinses with double sterile distilled water, the sterilized seeds were incubated at 30°C between several layers of sterilized wet filter paper for germination. The germinated seeds were sown in jars containing sterilized substrate of a 1 : 1 : 1 mixture of loam, vermiculite and perlite, and grown at 28°C under a 14 h light /10 h dark photoperiod condition. Fresh flowers of *C. acuminata* were collected from *C. acuminata* trees growing on the campus of Fudan University as well and stored at -70°C.

RNA and DNA isolation. Total RNAs of different tissues including root, stem, leaf and flower from *C. acuminata* plant were extracted using TRIzol Reagent (GIBCO BRL) according to the manufacturer's instruction (Jaakola *et al.*, 2001). The genomic DNA was isolated using a CTAB-based method (Rechards, 1995). The quality and concentration of RNA and DNA samples were examined by EB-stained agarose gel electrophoresis and spectrophotometer analysis.

3' cDNA end amplification of *CaSS*. The 3' RACE System for Rapid Amplification of cDNA Ends (GIBCO BRL) was used for 3' cDNA end amplification of *CaSS*. First of all, a group of hydroxylase/cytochrome P450 proteins and mRNA sequences from *Catharanthus roseus*, *Sesamum indicum*, *Pinus taida*, *Ocimum*

basilicum, *Glycine max* and others were multi-aligned on Vector NTI Suite 6 to design two degenerated homologous PCR primers FCASS3-1 (5'-ATTGG(A/C/T)CT(T/C) CT(A/C/T)TGGGA(T/C) ATGAT-3') and FCASS3-2 (5'-GT(N)GA(A/G)TGGGC(N)ATG GC(N)GA(G/A)CTG-3') corresponding to their conserved amino acid sequences IGLLWDMI and VEWAMAEL respectively. Approximately 2.5 µg of extracted total RNA was used for reverse transcription using AP (5'-GGCCACGCGTCGACTAGTACT₁₇-3'). Reverse transcription (RT)-PCR was conducted as the protocol except that the RT step was prolonged for a further reaction at 50°C for 30 min in case of a 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 *Taq* DNA polymerase, mixed with the primers FCASS3-1 and AUAP (5'-GGCCACGCGTCGACTAGTAC-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, dC tailing 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'-ATTGGAATTGGCCTTGTGTGG-3') with an extra 30 min of RT at 50°C after standard RT at 42°C. Primers RCASS5-2 (5'-CATGAGTGGGGTTGGAGGGTG-3') and AAP (5'-GGCCACGC GTCGACTAGTACGGGIIIGGGIIGGGIIG-3') were used for primary amplification with 5 µl of dC-tailed cDNA as template in a total volume of 50 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'-CGGTCTAGCTCCTCTTGAGCC-3') and AUAP under the following PCR condition: 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'-AATCTCCTCTGCTAAGTGCTATCCA-3') and RCASS (5'-GGCGATAGAACCCACTTTATTCTAA-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 μ L reaction solution containing 5 μ L 10 $\times$ LA *Taq* buffer (plus Mg^{2+}), 2 μ L 10 mM each of dNTPs, 2 μ L 10 μ M FCASS, 2 μ L 10 μ M RCASS, 2 μ L 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 bands was ligated to pGEM-T Easy vector (Promega) and transformed via hot-shock by incubation at 42°C for 90 s into *E. coli* strain DH5 α competent cells prepared with $CaCl_2$ double suspension method (Seidman *et al.*, 1995). Three independent positive clones harboring the amplified DNA were sequenced using T7/SP6 primers (Sangon Co. Ltd.). Sequence alignments, open reading frame (ORF) translation and molecular mass calculation of predicted protein were carried out with Vector NTI Suite 6. BLAST was done at the NCBI server (<http://www.ncbi.nlm.nih.gov/blast/Blast.cgi>), while structural analysis of predicted CASS-P450 protein was carried out on the website (<http://cn.expasy.org>).

Southern blotting analysis. Southern blot was carried out to analyze the copy number of *CaSS* present in *C. acuminata* genome. Total genomic DNA was isolated from 1.00 g of fresh *C. acuminata* leaf materials following the standard protocol (Rechards, 1995). Aliquots of DNA (20 μ g) were digested overnight at 37°C with *Bgl*II, *Bam*HI, *Xba*I, *Sac*I and *Eco*RV, which did not cut within the probe region, respectively, fractionated by 0.85% 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 FCASS3-2 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 manufacture'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. acuminata*, 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. taeda*, 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 possessed typical low G + C content and a putative polyadenylation signal AATAAA 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. erythrorhizon*. 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/gorf/>)

```

01      aatctcctctgctaagtgtctatccacccctatctctcg
41  ATGCGGGTCTCTATTATTCCTCCCTCTCATCTTCATATTCCTT
86  GCTTACAACCTCTATCAACGGCTTAGATTCAAGCTCCCGCGGGA
131  CCCC GCCCTTTGCCGTTCTCGTGGAAACCTTTACAGCGTGAAACCG
176  GTGAAGTCCGGTCTTCGCCGAATGGGCACAAGTTTATGGTCCG
221  ATCATGTCGGTGTGGTTCGGTTCGACTTAAACGTAGTGATCTCA
266  AACTCAGAGTTGGCTAAAGAAGTACTCAAGGAAATGATCAACAT
311  TTAGCAGATAGGGAAGAAATAGATCATCAATTCATTACGAGA
356  GGTGGGAAGGATCTGATATGGGCTGATTATGGCCCTCACTATGC
401  AAGGTTAGAAAGCTGTGACTGTTGAACCTTCACTCAGAAAGG
446  CTTGAAGCTCTAGACCTATTGGAAGATGAAGTTACTGCCATG
491  GTTGACTCCATTTCAAAGCTCGGCAATCCTGATAATTATGGG
536  AAAAGCTGTTGGTGAAGCAGTACATAGGAGCAGTAGCATTCAAC
581  AACATGAACGAGCTGTTATTTGGGAAGCGGTTTGAAGCTCGGAG
626  GGGGTGATGGATGAGCAAGGAAGGAATTCAGAGGATTGTGGCC
671  AATGGCGTCAAGATTTGGTGGATTGCGTTTCATTGGAGAGCATT
716  CCATGGCTTGGTGGTATGTCACAGGAGGAAGATATTGGT
761  AAACATGAAGCAGTGGGATCGACTCACCAGGCGCATCGAA
806  GAGCACACACTTCCCGCAAGCACTGTGGCGTGGCAAGCAGCAT
851  TTTGCGATGCCTTGTCTACTCTCAGAAGGAATATGAAGCTCAGT
896  GATGACACTGTCTTGGCTCGCTTTGGGACATGATCACTGACGGC
941  ATGACACACGTCGATCTCAGTTGAGTGGCGATGGCCGAGCTG
986  ATCAAGAACCCAGAGTGCAACAGAGGCTCAAGAGGAGCTAGAC
1031  CGCGTGATGGATGAGCATATCATGTCTGAACTGACTTCCCG
1076  AACCTCCATCACTACAATGTGTAGCTAAGGAAGCGCTGAGGTTG
1121  CACCTCCACCCCACTCATGTTCCACACAGGCAATTCGAAT
1166  GTCAAGATTGGTGGCTGATACATCCCTAAGGCTCTATTGTGCAC
1211  GTAATGTATGGGCATTGCTGCTGATCCAGCAACCTGGAAGAC
1256  CCTCAGAGTTCCGGCTGAAGCTTTCCTTGAGGAAGATGTTGAC
1301  ATGAAGGCTCATGATTATAGGCTACTACCTTTGGTGCAGGAAGG
1346  CGTATCGCCCTGGTCTCAGATGCCATTAACTTGATCAGATCG
1391  ATGCTTGGTCACTGTTGCACCATTTTCTTGGGCCACCACTGAA
1436  GGGGTCAAGCAGAAAGATCGACTGAGAACCTGGATTG
1481  GTTACTTTTATGAAGACACAGTACAAGCTGTTGCTAAACCAAGA
1526  TTGCTTCACTATTGTACAAAGCTGTGGCTGTGGGTGCTGTG764
1571  ttttttaatcatgctgtcatgcattgatctccttctagtagg
1616  gtttggattttgaattctgtggtgattgttaattgttaatta
1661  cgtatctctgattaaagatcgattgaggaagacaattatttat
1706  gatgattagaataaagtggttctatcgcaaaaaaaaaaaaaa
1751  aaaaaaaaaaaaaaaaaaaaaaaaaaaaaa

```

Fig. 1. 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 AATAAA were boxed. The primers RCASS and FCASS 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.

gorf.html) and the predicted CASS protein precursor was 509 aa in length with a theoretical molecular weight of 57.87 kDa and a pI value of 8.84. The charged amino acids constituted 31.43% of the protein, while the hydrophobic amino acid content reached 37.52% with high leucine content (10.02%).

Multi-alignment of CASS with other plant hydroxylases/cytochrome P450 was conducted (Fig. 2). According to NCBI protein-protein BLAST, the deduced CASS amino acid sequence showed 77% (380/489) identities and 86% (425/489) positives in local alignments to *G. max* cytochrome P450 monooxygenase 98A2p. It also shared very broad and high local identities and positives to hydroxylases/cytochrome P450 from species of broad plant classes such as *S. bicolor* (with 67% identity and 81% positive), *L. erythrorhizon* (with 70% identity and 83% positive), *S. indicum* (with 66% identity and 80% positive), *O. basilicum* (with 77% identity and 88% positive), *S. scutellarioides* (with 70% identity and 84% positive), *A. thaliana* (with 74% identity and 86% positive), *P. taeda* (with 74% identity and 85% positive).

According to NCBI conserved domain search, the conserved domains p450 (F₅₁-T₄₈₂) and CypX (P₂₈-P₄₇₈) were found to penetrate through the whole molecule of predicted CASS protein. Signal P1.1 analysis showed a most likely cleavage site between K₁₈ and L₁₉, while TargetP V1.0 Prediction result indicated a secretory pathway of subcellular location, suggesting that the predicted CASS protein had a signal peptide of 18 aa in the N terminal. By cutting away the 18 aa signal peptide, the mature polypeptide is 491 aa with a molecular weight of 55.73 kDa and a pI of 8.74.

The secondary and three-dimensional structure predictions of CASS protein were conducted (Fig. 3). Based on the Hierarchical Neural Network (Combet *et al.*, 2000) result, CASS mature peptide was composed of 40.33 % alpha helix, 15.89 % extended strand and 43.79 % random coil (Fig. 3a). Alpha helices resided within the main body especially in the middle region of CASS, while random coils penetrated through most part of this sequence accompanied by a small amount of extended strands. TMpred prediction yielded 4 strong transmembrane helices with a total score of 5826: 1) 1-17 bp, o-i (outside to inside), with a score of 2604; 2) 57-74 bp, i-o (inside to outside), with a score of 1279; 3) 167-188 bp, o-i, with a score of 832; and 4) 437-459 bp, i-o, with a score of 1111. The heme-binding sequence PFGXGRRXCX, highly conserved among the P450s of higher plants, was also found between the amino acid position 430 and 439 of CASS. Swiss-Model structure prediction (Guex and Peitsch, 1997) of CASS mature protein resulted in a similar folding mode and spatial configuration to oxidoreductases such as mammalian cytochrome P450 2C5 and cytochrome P450 BM-3 heme domain. The putative CASS protein was composed of two proteins composed of 194 aa and 449 aa respectively (Fig. 3b).

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

		1		50
<i>Sorghum bicolor</i>	(1)	MDASLLLSVALAVLIPLSLALNRLRLGRLP:GP:PWVVLGNL		RQ
<i>Lithospermum erythrorhizon</i>	(1)	MALPAIPLAIIIFLIISYKLYQKRL-KLP:GP:PLPIIGNIYQ		
<i>Sesamum indicum</i>	(1)	MALPLLIISFLT-IFIAYKLFQRLRY-KLP:GP:PLPVVGNLYD		
<i>Glycine max</i>	(1)	MALLLIIPISLVT-LWLGYTLYQRLRF-KLP:GP:PWVVLGNLYD		
<i>Ocimum basilicum</i>	(1)	MAALLLLLLPLLLPA-IFLLHHLYRLRF-RLP:GP:PLPIVGNLYD		
<i>Solenostemon scutellarioides</i>	(1)	MTSLLAHPLPPP-RRRPLPPLPPP-SHPAGP:PGRWWRNLYD		
<i>Arabidopsis thaliana</i>	(1)	MSWFLIAVATIA-AVVSYKLIQRLRY-KFP:GP:SPKPIVGNLYD		
<i>Pinus taeda</i>	(1)	MSVPEMGPFVVTGIVA-IAIVYKLVQRWF-KLP:GP:PWVVLGNLQ		
<i>Camptotheca acuminata</i>	(1)	MAVLLLFPLIF-IFLAYKLYQRLRF-KLP:GP:PLPFVGNLYS		
Consensus	(1)	MALLLI L LI I LSYKLYQRLRF KLPPGPRP PVVGNLYD		
		51		100
<i>Sorghum bicolor</i>	(47)	IKPIRCRCQEWAEYGPVIVSWFGSGLTIVVSTSELAKEVLKENDQOLA		
<i>Lithospermum erythrorhizon</i>	(44)	VKPVKFCFYNYSKYGPISFYYSQMNVIIVSTTELAKVLEKNDQOLA		
<i>Sesamum indicum</i>	(43)	IKPLLVRCTEWSQVYGPISVYLGSHLSVVNSAEAKEVLKNDQOLA		
<i>Glycine max</i>	(44)	IKPVRFRCFAEWAQSYGPIISVWFGSTLNVIVSNSAEAKEVLKENDQOLA		
<i>Ocimum basilicum</i>	(47)	VKPVRFRCFADWAQSYGPIISVWFGSTLNVIVSNTELAKVLEKENDQOLA		
<i>Solenostemon scutellarioides</i>	(44)	VKPVRFRCFAEWAQFSGTFSVWFGSTLNVIVSSSELAKEVLKENDQOLA		
<i>Arabidopsis thaliana</i>	(43)	IKPVRFRCYEWASQYGPISVWFGSTLNVIVSSAEAKEVLKENDQOLA		
<i>Pinus taeda</i>	(49)	IKPVRFRCFWDWSKYGPISVWFGSTLNVIVSNTELAKVLEKENDQOLA		
<i>Camptotheca acuminata</i>	(43)	VKPVKFCFAEWAQVYGPISVWFGSTLNVIVSNSAEAKEVLKENDQOLA		
Consensus	(51)	IKPVRFRCF EWAQ YGPIISVWFGSTLNVIVSSSELAKEVLKE DQOLA		
		101		150
<i>Sorghum bicolor</i>	(97)	DRPRRSTORFSRNGODLIWADYGPHYIKVRKLCNLELFTPKRLEALRPI		
<i>Lithospermum erythrorhizon</i>	(94)	DRFRTRSSASMSRGGKDLIWADYGPHYVVKRKL CNVELFSPKRLEALRPM		
<i>Sesamum indicum</i>	(93)	NRNRTRQIAKFSKNGADLIWSDYGPHYVVKRKLCTLELFSMKRLEGLRPI		
<i>Glycine max</i>	(94)	DRHRSRSAKF SRDGKDLIWADYGPHYVVKRKLCTLELFSMPKRLEALRPI		
<i>Ocimum basilicum</i>	(97)	DRHRSRSAKF SRDGODLIWADYGPHYVVKRKLCTLELFSMPKRLEALRPI		
<i>Solenostemon scutellarioides</i>	(94)	DPPRSRAVKLSRDGKDLIWADYGPHYVVKRKLCTVLFSPKRLEALRPI		
<i>Arabidopsis thaliana</i>	(93)	DRHRNRSTEARNGODLIWADYGPHYVVKRKLCTLELFTPKRLESRLPI		
<i>Pinus taeda</i>	(99)	DRPRSRAKF SRNGODLIWADYGPHYVVKRKLCTLELFSMPKRLEALRPI		
<i>Camptotheca acuminata</i>	(93)	DRERNRSSNSF SRGGKDLIWADYGPHYVVKRKLCTVELFTPKRLEALRPI		
Consensus	(101)	DR RRSRA KFSR G DLIWADYGPHYVVKRKLCTLELFSMPKRLEALRPI		
		151		200
<i>Sorghum bicolor</i>	(147)	REDEVTALVESVYRAATAPGNECKPMVVRNHL SMVAFNNITRLAFGKRFM		
<i>Lithospermum erythrorhizon</i>	(144)	REDEVTALVESIYKDOTKPELKCKSLLVREYLSVAFNNITRLAFGKRFM		
<i>Sesamum indicum</i>	(143)	REDEVTALVESIFKDCNKPENCKALVREYLGMMAFLLHARLTFGKRFM		
<i>Glycine max</i>	(144)	REDEVTALVESVDSYVNHCTSTENLCKGILLRKH LGVAFNNITRLAFGKRFV		
<i>Ocimum basilicum</i>	(147)	REDEVTALVESIYHDC TAPDNACKSLLVKKYLGAVAFNNITRLAFGKRFV		
<i>Solenostemon scutellarioides</i>	(144)	REDEITVIVESIYQD SAASCKSVVYKYLASMAFNNITRLVFGKRFV		
<i>Arabidopsis thaliana</i>	(143)	REDEVTALVESVFRDNL PENRAGLQLRKYLGAFAFNNITRLAFGKRFM		
<i>Pinus taeda</i>	(149)	REDEVAAMVESIFNDGCKOEGICKPLVVKRYL SGVAFNNITRPAFGKRFV		
<i>Camptotheca acuminata</i>	(143)	REDEVTALVDSIFKDSANPDNYCKSLLVKQYIGAVAFNNITRLVFGKRFM		
Consensus	(151)	REDEVTAMVESIYKDC PEN GKSLLVKRYLGAFAFNNITRLAFGKRFM		
		201		250
<i>Sorghum bicolor</i>	(197)	NANGIDDEGGREFKTIIVNGIKIGASLSVAEFTWYLRWLCPNLEELYKTH		
<i>Lithospermum erythrorhizon</i>	(194)	DSNGVNEGGQEFKTI THDGIKITAKLSIAEYIPWIRMMFKVEQDALDKF		
<i>Sesamum indicum</i>	(193)	DSNGVVDGQGEELKALHNGIKLGTKKSWAEFLPWRFLFKTENQLLAEH		
<i>Glycine max</i>	(194)	NSEGVMDEGGVEFKATVNGIKLGASLMAEHI PWLRMMFPLEEGAFKHX		
<i>Ocimum basilicum</i>	(197)	NSEGIIDKQGLEFKATVSGIKLGASLMAEHI PSLRMMFPLEDAFAKH		
<i>Solenostemon scutellarioides</i>	(191)	NSEGEVDKQGEFKATVINGIKLGASLAVSEHI PWLRMMFPLEDAFTQH		
<i>Arabidopsis thaliana</i>	(193)	NAEGVDEQGLEFKATVSGIKLGASLSIAEHI PWLRMMFPADEKAFKHX		
<i>Pinus taeda</i>	(199)	NEEGMDPQGEVEFKIVATGKLGASLTMAEHI PYLRMMFPLEEGAFKHX		
<i>Camptotheca acuminata</i>	(193)	NSEGVMDEQGEFRGIVANGVKTGGLRFIGEHI PWLRMMFPLEEGAFKHX		
Consensus	(201)	NSEGVVDEQGE FKAIV NGLKLGASLSIAEHI PWLRMMFPLEDAFAKH		

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* (BAC44836), *Sesamum indicum* (AAL47545), *Glycine max* (AAB94587), *Ocimum basilicum* (AAL99200), *Solenostemon scutellarioides* (CAD20576), *Arabidopsis thaliana* (NP_850337), *Pinus taeda* (AAL47685). The completely identical amino acids were dark shaded and the heme-binding sequence PFGXGRRXCX was underlined and also boxed. The conserved amino acid sequences IGLLWDMI and VEWAMAEL equivalent to the two degenerated homologous PCR primers for 3'cDNA amplification of *CaSS* were shown in grey boxes.

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

		251	300
Sorghum bicolor	(247)	NERRDRLTMKII EEHAKSLKESG-AKOHFVDALFTLKQOYDLS	EDTVIGL
Lithospermum erythrorhizon	(244)	AARDHLTRVIM EEHIIKS-----GNTKOHFVDALLTLQOYD	SEDITIGL
Sesamum indicum	(243)	DALADSFYKKIMQEHTLARQKTGNTKNHFVDALLTLQOYD	SDDTVIGL
Glycine max	(244)	GARRDRLTRAIMAEHTEARKKSGGAKOHFVDALLTLQOYD	SEDITIGL
Ocimum basilicum	(247)	GARRDQLTREIM EEHTRAREESSGGAKOHFVDALLTLQOYD	SEDITIGL
Solenostemon scutellarioides	(241)	GVRMERLTREIMQEHTLARQKTGGAKOHFVDALLTLQOYD	SEDITIAL
Arabidopsis thaliana	(243)	GARRDRLTRAIM EHTLARQKSSGAKOHFVDALLTLQOYD	SEDITIGL
Pinus taeda	(249)	GARRDNVTKAIM EHTLARQTSG-AKOHFVDALLTLQOYD	SEDITIGL
Camptotheca acuminata	(243)	EARRDRLTRAIM EHTLARKHSGGAKOHFVDALLTLQOYD	SDDTVIGL
Consensus	(251)	GARRDRLTR IMEHTLAR SGGAKOHFVDALLTLQD YD	LSEDITIGL
		301	350
Sorghum bicolor	(296)	LWDMITAGMDTTIVSVEWAMAE LVRNPRVQKKLQEE	DRVVGRDRVMLET
Lithospermum erythrorhizon	(290)	LWDMIAAGMDTATISTEWAMAE LVRNPRVQKQOEE	DRVVGPRIMTEA
Sesamum indicum	(293)	LWDMISAGMVTITITVEWAMAE LVRNPRVQCKVQEE	DRVVGS DRVMT EA
Glycine max	(294)	LWDMITAGMDTTAISVEWAMAE LVRNPRVQCKVQEE	DRVVGLERVMT EA
Ocimum basilicum	(297)	LWDMITAGMDTTAISVEWAMAE LVRNPRVQCKVQEE	DRVVGYERVMT EL
Solenostemon scutellarioides	(291)	LWDMIAAGMDTPAISVEWAMAE LVRNPRVQCKVQEE	DRVVGHIERIMTEL
Arabidopsis thaliana	(293)	LWDMITAGMDTTAITAEWAMAE LVRNPRVQCKVQEE	DRVVGLDRILTEA
Pinus taeda	(298)	LWDMITAGMDTTAITVEWAMAE LVRNPRVQCKVQEE	DRVVGRDRVMNET
Camptotheca acuminata	(293)	LWDMITAGMDTTISISVEWAMAE LVRNPRVQCKVQEE	DRVVGYERIMSET
Consensus	(301)	LWDMITAGMDTTAISVEWAMAE LVRNPRVQCKVQEE	DRVVG DRVMT E
		351	400
Sorghum bicolor	(346)	DFPNLPYLQAVVKESLR HPPTPLMLPHKASTIVK	GGYDIPKGANVWVN
Lithospermum erythrorhizon	(340)	DVPKLPYLQCVVKESLR HPPTPLMLPHRASANVK	GGYDIPKGSIVHVN
Sesamum indicum	(343)	DIPNLPYLQCVTKCFRR HPPTPLMLPHKASTIVK	GGYDIPKGATVSVN
Glycine max	(344)	DFSNLPYLQCVTKAMR HPPTPLMLPHRANANVK	GGYDIPKGSNVHVN
Ocimum basilicum	(347)	DFSNLPYLQCVAKEALR HPPTPLMLPHRSNSNVK	GGYDIPKGSNVHVN
Solenostemon scutellarioides	(341)	DIPNLPYLQCVVKESLR HPPTPLMLPHRSNADVK	GGYDIPKGSNVHVN
Arabidopsis thaliana	(343)	DFSNLPYLQCVVKESFR HPPTPLMLPHRSNADVK	GGYDIPKGSNVHVN
Pinus taeda	(348)	DFPHLPYLQCVTKALR HPPTPLMLPHKATONVK	GGYDIPKGSNVHVN
Camptotheca acuminata	(343)	DFPNLPYLQCVAKEALR HPPTPLMLPHKANSIVK	GGYDIPKGSIVHVN
Consensus	(351)	DFPNLPYLQCV KEALRHPPTPLMLPHRANANVK	GGYDIPKGSNVHVN
		401	450
Sorghum bicolor	(396)	VWAVARDPAVWKNPLEFRPERFLEEDVDIKG	DFRLLPFGAGRRVCPGAQ
Lithospermum erythrorhizon	(390)	VWAVARDPAVWKNPEFRPERFLEEDVDIKG	DFRLLPFGAGRRVCPGAQ
Sesamum indicum	(393)	VWAVARDPAVWKNPLEFRPERFLEEDVDIKG	DFRLLPFGAGRRVCPGAQ
Glycine max	(394)	VWAVARDPAVWKNPLEFRPERFLEEDVDIKG	DFRLLPFGAGRRVCPGAQ
Ocimum basilicum	(397)	VWAVARDPAVWKNPEFRPERFLEEDVDIKG	DFRLLPFGAGRRVCPGAQ
Solenostemon scutellarioides	(391)	VWAVARDPAVWKNPLEFRPERFLEEDVDIKG	DFRLLPFGAGRRVCPGAQ
Arabidopsis thaliana	(393)	VWAVARDPAVWKNPEFRPERFLEEDVDIKG	DFRLLPFGAGRRVCPGAQ
Pinus taeda	(398)	VWAVARDPAVWKNPEFRPERFLEEDVDIKG	DFRLLPFGAGRRVCPGAQ
Camptotheca acuminata	(393)	VWAVARDPAVWKNPEFRPERFLEEDVDIKG	DFRLLPFGAGRRVCPGAQ
Consensus	(401)	VWAVARDPAVWKNPLEFRPERFLEEDVDIKG	DFRLLPFGAGRRVCPGAQ
		451	500
Sorghum bicolor	(446)	LGINLVASMLGHLHFFWTPPEGTRPEDVNMES	PGLVTFMGTPLCAVA
Lithospermum erythrorhizon	(440)	LAINLITSSMLGHLHFFWTPPEGTRPEDVNMES	PGLVTVMRNPVKAIV
Sesamum indicum	(443)	LAIYLVTSMLGHLHFFWTPPEGTRPEDVNMES	PGLVTVMRTPLCAVP
Glycine max	(444)	LGINLVASMLGHLHFFWTPPEGTRPEDVNMES	PGLVTVMRTPLCAVP
Ocimum basilicum	(447)	LGINLVTSMLGHLHFFWTPPEGTRPEDVNMES	PGLVTVMRTPLCAVP
Solenostemon scutellarioides	(441)	LGINLVTSMLGHLHFFWTPPEGTRPEDVNMES	PGLVTVMRTPLCAVP
Arabidopsis thaliana	(443)	LGINLVTSMLGHLHFFWTPPEGTRPEDVNMES	PGLVTVMRTPLCAVP
Pinus taeda	(448)	LGINLVQSMGLGHLHFFWTPPEGTRPEDVNMES	PGLVTVMRTPLCAVP
Camptotheca acuminata	(443)	LAINLITSSMLGHLHFFWTPPEGTRPEDVNMES	PGLVTVMRTPLCAVP
Consensus	(451)	LGINLVTSMLGHLHFF WTPP GVKPEEIDM ENPGLVTVMRTPLCAVP	
		501	518
Sorghum bicolor	(496)	KPRLEKEELYNRPVEM-	
Lithospermum erythrorhizon	(490)	SPRLS-AVHLYKRVEDSM	
Sesamum indicum	(493)	TPRLP-ADLFKRVASGNV	
Glycine max	(494)	SPRLP-SHLYKRVPAEI-	
Ocimum basilicum	(497)	TPRLP-SDLYKRIAVDL-	
Solenostemon scutellarioides	(491)	TPRLP-ANLYKRVAVVDI	
Arabidopsis thaliana	(493)	TPRLP-SDLYKRVPYDM-	
Pinus taeda	(498)	TPRLP-DHLYKROPLN-	
Camptotheca acuminata	(493)	KPRLP-SHLYKRVAVGAV	
Consensus	(501)	TPRLP S LYKRV V	

Fig. 2. Continued.

Fig. 3. The secondary and three-dimensional structures of the predicted CASS polypeptide. (A) The secondary structure of CASS. α -helix and extended strand were denoted as vertical long bars and vertical short bars respectively, with the horizontal line presenting the random coil running through the whole molecule. (B) The three-dimensional structure of CASS. The α -helix was shown in grey (helix-shaped), the sheet in grey (wide ribbon-shaped) and the random coil also in grey (line-shaped). The highly conserved heme-binding sequence was shown in dark. More detailed structure of the heme-binding region was shown in the right.

one copy of the *CaSS* present in the genome of *C. acuminata* (Fig. 4), which is similar to the counterpart in *A. thaliana*.

Fig. 4. Southern blotting analysis. Total genomic DNA was isolated from fresh leaves of *Camptotheca acuminata* and was digested overnight at 37°C with *Bgl*II, *Bam*HI, *Xba*I, *Sac*I and *Eco*RV, which did not cut within the probe region, respectively, followed by hybridization with the Biotin-labeled *CaSS* 3' end cDNA fragment.

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

Fig. 5. Northern blotting analysis. Total RNA was isolated from root, stem, leaf and flower tissues of *Camptotheca acuminata* followed by hybridization with the Biotin-labeled *CaSS* 3' end cDNA fragment (upper panel). The 18s and 28s rRNAs were used as internal control in each lane (lower panel).

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 phenylpropanoid 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

- Allona, I., Quinn, M., Shoop, E., Swope, K., Cyr, S. S., Carlis, J., Riedle, J., Retzel, E., Campbell, M. M., Sederoff, R. and Whetten, R. W. (1998) Analysis of xylem formation in pine by cDNA sequencing. *Proc. Natl. Acad. Sci. USA* **95**, 9693-9698.
- Anterola, A. M., Jeon, J., Davin, L. B. and Lewis, N. G. (2002) Transcriptional Control of Monolignol Biosynthesis in *Pinus taeda*. *J. Biol. Chem.* **277**, 18272-18280.
- Baucher, M., Monties, B., Montagu, M. V. and Boerjean, W. (1998) Biosynthesis and genetic engineering of lignin. *Crit. Rev. Plant Sci.* **17**, 125-197.
- Boudet, A. M. (2000) Lignins and lignification: selected issues. *Plant Physiol. Biochem.* **38**, 1-16.
- Combet, C., Blanchet, C., Geourjon, C. and Deleage, G. (2000) Network protein sequence analysis. *Trends Biochem. Sci.* **25**, 147-150.
- Croteau, R., Kutchan, T. M. and Lewis, N. G. (2000) Natural products (secondary metabolites); in *Biochemistry and Molecular Biology of Plants*, Buchanan, B., Jones, R., and Gruissem, W. (eds), pp. 1250-1318, American Society of Plant Physiologists, Rockville, USA.
- David, R. G., Till, B., Pascaline, U., Danièle, W. and Eran, P. (2002) Differential production of *meta* hydroxylated phenylpropanoids in sweet basil peltate glandular trichomes and leaves is controlled by the activities of specific acyltransferases and hydroxylases. *Plant Physiol.* **130**, 1536-1544.
- Franke, R., Humphreys, J. M., Hemm, M. R., Denault, J. W., Ruegger, M. O., Cusumano, J. C. and Chapple, C. (2002a) The *Arabidopsis* *REF8* gene encodes the 3-hydroxylase of phenylpropanoid metabolism. *Plant J.* **30**, 33-45.
- Franke, R., Hemm, M. R., Denault, J. W., Ruegger, M. O., Humphreys, J. M. and Chapple, C. (2002b) Changes in secondary metabolism and deposition of an unusual lignin in the *ref8* mutant of *Arabidopsis*. *Plant J.* **30**, 47-59.
- Goffner, D., Campbell, M. M., Campargue, C., Clastre, M., Borderies, G., Boudet, A. and Boudet, A. M. (1994) Purification and characterization of cinnamoyl-coenzyme A: NADP oxidoreductases in *Eucalyptus gunnii*. *Plant Physiol.* **106**, 625-632.
- Gowri, G., Bugos, R. C., Campbell, W. H., Maxwell, C. A. and Dixon, R. A. (1991) Stress responses in alfalfa (*Medicago sativa* L.): X. Molecular cloning and expression of S-Adenosyl-L-Methionine: caffeic acid 3-o-methyltransferase, a key enzyme of lignin biosynthesis. *Plant Physiol.* **97**, 7-14.
- Guex, N. and Peitsch, M. C. (1997) SWISS-MODEL and the Swiss-PdbViewer: An environment for comparative protein modeling. *Electrophoresis* **18**, 2714-2723.
- Jaakola, L., Pirttilä, A. M., Halonen, M. and Hohtola, A. (2001) Isolation of high quality RNA from bilberry (*Vaccinium myrtillus* L.) fruit. *Mol. Biotechnol.* **19**, 201-203.
- Lewis, N. G., Davin, L. B. and Sarkanen, S. (1999) The nature and function of lignins; in *Comprehensive Natural Products Chemistry*, Barton, D.H.R., Nakanishi, K. and Meth-Cohn, O. (eds), pp. 617-745, Elsevier Science Publishers B.V. Amsterdam, The Netherlands.
- Matsuno, M., Nagatsua, A., Ogihara, Y., Ellis, B. E. and Mizukami, H. (2002) CYP98A6 from *Lithospermum erythrorhizon* encodes 4-coumaroyl-4P-hydroxyphenyllactic acid 3-hydroxylase involved in rosmarinic acid biosynthesis. *FEBS Lett.* **514**, 219-224.
- Reichards, E. J. (1995) Preparation and analysis of DNA; in *Short Protocol in Molecular Biology*, Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (eds), pp. 36-38, John Wiley and Sons, New York, USA.
- Schoch, G., Goepfert, S., Morant, M., Hehn, A., Meyer, D., Ullmann, P. and Werck-Reichhart, D. (2001) CYP98A3 from *Arabidopsis thaliana* is a 3'-hydroxylase of phenolic esters, a missing link in the phenylpropanoid pathway. *J. Biol. Chem.* **276**, 36566-36574.
- Seidman, C. E., Struhl, K. and Sheen, J. (1995) Escherichia coli, plasmids, and bacteriophages; in *Short Protocols in Molecular Biology*, Ausubel, F. M., Brent, R., Kingston, R. E., Moore, D. D., Seidman, J. G., Smith, J. A. and Struhl, K. (eds), pp. 22-24, John Wiley and Sons, New York, USA.
- Sterky, F., Regan, S., Karlsson, J., Hertzberg, M., Rohde, A., Holmberg, A., Amini, B., Bhalerao, R., Larsson, M., Villarroel, R., Van Montagu, M., Sandberg, G., Olsson, O., Teeri, T. T., Boerjan, W., Gustafsson, P., Uhlen, M., Sundberg, B. and Lundeberg, J. (1998) Gene discovery in the wood-forming tissues of poplar: analysis of 5,692 expressed sequence tags. *Proc. Natl. Acad. Sci. USA* **95**, 13330-13335.
- Yamada, T., Kambara, Y., Imaishi, H. and Ohkawa, H. (2000) Molecular cloning of novel cytochrome P450 species induced by chemical treatments in cultured tobacco cells. *Pesticide Biochem. Physiol.* **68**, 11-25.