

# Isolation and Characterization of P450 Gene from the Formosan Subterranean Termite, *Coptotermes formosanus* (Isoptera: Rhinotermitidae)

by

Yue Zhou<sup>1</sup>, Jianchu Mo<sup>2</sup>, Keping Chen<sup>1</sup> & Qinggang Xu<sup>1\*</sup>

## ABSTRACT

A cytochrome P450 gene belonging to family9 was isolated from the midgut transcriptome of the termite *Coptotermes formosanus* Shiraki, for screening enzymes related to biomass degeneration. Some studies show that insect P450 enzymes have ligninase activities for catalyzing lignin degradation. We employed the RACE method to clone this cytochrome P450 gene, named CYP9AX1 (GenBank accession No.JN969113). To the best of our knowledge, CYP9AX1 is the first member of the CYP9 family cloned from this termite. The full-length CYP9AX1 cDNA was 2242 bp long and included a 1599bp open-reading-frame (ORF), a 61-bp 5'-untranslated region (UTR) and a 592-bp 3'-UTR (excluding the poly-A tail). The CYP9AX1 protein deduced from the ORF contains 532 amino acids with a predicted signal peptide composed of 20 amino acid at its N-terminal and the classic heme-binding domain FXXGXXXCXG (residues 468-477). At position 473, residue Arg (R) changes to Gln (Q), this suggests that CYP9AX1 is a new type of CYP subfamily 9A. The phylogenetic tree showed that *C. formosanus* has high genetic relationship with *Blattella germanica* and *Diploptera punctata*. Quantitative RT-PCR assays demonstrated that CYP9AX1 was expressed most abundantly in malpighian tubules, and slightly lower in the head, foregut, midgut and hindgut. The results suggested that CYP9AX1 may be involved in enzymatic detoxification systems of the delignification process in *C. formosanus*.

<sup>1</sup>Institute of Life sciences, Jiangsu University, Zhenjiang 212013, China

<sup>2</sup>Ministry of Agriculture Key Laboratory of Molecular Biology of Crop Pathogens and Insects, Institute of Insect Sciences, Zhejiang University at Zijingang campus, Yuhangtang Road 388, Hangzhou, Zhejiang 310012, China

\*Correspondence: Qinggang Xu, Institute of Life Sciences, Jiangsu University, Zhenjiang, 212013, China. email: xuqg@ujs.edu.cn

Keywords: Characterization, *Coptotermes formosanus* Shiraki, Cytochrome P450, Real-time PCR

## INTRODUCTION

The cytochrome P450 (CYP) monooxygenases have numerous functional roles in oxidative transformation of endogenous substrates and xenobiotics (Nelson *et al.* 1996). In insects, cytochrome P450s metabolize hormones and pheromones but are best known for their roles in the metabolism of insecticides and host plant chemicals (Feyereisen 1999; Scott *et al.* 1998). Many studies focus on the function of P450 on insect development and detoxification contributes to insecticide resistance (Scott 1999, Scott & Wen 2001). The first member of the CYP9 family, CYP9A1 from *Heliothis virescens*, is constitutively over-expressed in resistant, thiodicarb-selected tobacco budworms (Rose *et al.* 1997), and may play a role in pesticide metabolism (Stevens *et al.* 2000). Many studies showed that CYP9 genes such as CYP9A2,4,5, CYP9M10, CYP9A12,17, CYP9Q, are mainly involved in detoxification of plant allelochemicals and pesticides (Hardstone *et al.* 2010; Itokawa *et al.* 2010; Mao *et al.* 2011; Stevens *et al.* 2000; Zhou *et al.* 2010). Interestingly, some insect P450 enzymes show ligninase activities for catalyzing lignin degradation (Geib *et al.* 2008; Scharf & Boucias 2010). The termite is an insect can survive on highly lignocellulosic materials. The termite cytochrome P450 enzyme may be involved in lignin degradation. It is a candidate pre-treatment enzyme for delignification of lignocelluloses (Scharf & Boucias 2010). Recent evidence suggested that lignin degradation is a key pre-treatment process in biomass transformation into biofuels (Geib *et al.* 2008; Ke *et al.* 2011). Therefore, revealing the nature and roles of the P450 monooxygenases of *C. formosanus* will facilitate the understanding of the mechanism of this enzyme in lignocellulose pre-treatment.

In this paper, a new P450 gene belonging to the CYP9 family named CYP9AX1 was isolated from the termite *C. formosanus* Shiraki. Our study shows that this CYP450 is mainly expressed in malpighian tubules. This newly identified CYP9AX1 appears to be important for *C. formosanus* to detoxify plant allelochemicals from digestion of lignocelluloses.

## MATERIALS AND METHODS

### Termites

Termites were collected from Hangzhou city of Zhejiang province. The termites were maintained in the laboratory in a dark container on wet wood at 27°C~30°C and 70% relative humidity.

### RNA isolation and gene cloning

Ten worker specimens were washed in sterile phosphate-buffered saline (PBS: 5mM Na<sub>2</sub>HPO<sub>4</sub>, 5mM NaH<sub>2</sub>PO<sub>4</sub>, 130mM NaCl), then the digestive organs (consisting of the salivary gland, foregut, midgut and hindgut) were dissected and homogenized from the worker termites, and total RNAs were extracted from the different homogenates using an RNApure isolation Kit (BioTeke, China).

A reverse transcription reaction was carried out to synthesize the first strand cDNA using Reverse Transcript Kit2.0 (TaKaRa, China). Briefly, reverse transcription was performed for 1 h at 42°C in 10μL reaction mixture, containing 1 μg of total RNA, 25 pmol Oligo dT primer, 0.5 mmol/l dNTPs, 10 units of recombinant RNasin ribonuclease inhibitor and 50 units of PrimerScript™ reverse transcriptase (TaKaRa, China). The cDNA was diluted to 30μL. The PCR reactions were performed in 25μL reaction mixture containing 1μL of single strand cDNA, 1.5 mmol/l MgCl<sub>2</sub>, 30 pmol of each primer, and 1 unit of LA Taq DNA polymerase (TaKaRa, China) using Eppendorf 2724 thermal cycler (USA).

One pair of PCR primers was designed based on the P450 gene sequences from the transcriptome of termite midgut (unpublished). The primers for amplifying the *C. formosanus* P450 gene (CfP450) were as follows: P450F: 5' -ATGGACTGGAAGTGGGCGCTGTCAGAC-3' and P450R: 5' -AGAACTTCTGAGCTCCAGCCC-3'. The PCR reaction was performed as follows: 94°C for 1 min, 30 cycles of 94°C for 30 s, 60°C for 30 s, and 72°C for 1 min, and a final extension for 5 min at 72°C. The PCR product of *C. formosanus* P450 was about 1500 bp long. The amplified cDNA fragments were separated by 1% (w/v) agarose gel and purified with a DNA recovery kit (Axygen, China), subcloned into pMD18-T vector (TaKaRa, China), and sequenced by Sangon (Shanghai, China).

## Rapid-amplification of cDNA ends of Cfp450

A RACE (rapid-amplification of cDNA ends) reaction was conducted to amplify the 5' and 3' UTR of the Cfp450 gene using a SMARTer™ RACE cDNA Amplification Kit (Clontech, USA). The gene-specific primer (GSP) (5' RACE primer: P450-GSP1: 5' -GGATCGCATGTCCTTCCAACGCTGACCTGAC-3' and 3' RACE primers: P450-GSP2: 5' -GCTTGAACCTGGGGACAGACTCTTC-3') was designed to amplify the full-size Cfp450 gene. The PCR reaction was performed at 94°C for 5 min for DNA predenaturation, then 35 cycles of 94°C for 1 min, 50°C for 1 min and 72°C for 1 min 30 s, followed by one final extension at 68°C for 10 min. Both fragments of 5'- and 3'-RACE were analyzed by 1% agarose gel, then purified and cloned into pMD-18T vector (TaKaRa, China). Positive clones were sequenced.

## Bioinformatics analysis

Some online bioinformatics software was used to characterize the Cfp450 gene. SignalP 4.0 server was used to predict N-terminal signal peptide cleavage sites in amino acid sequences (<http://www.cbs.dtu.dk/services/SignalP/>). The transmembrane domain was predicted with the TMHMM tool (<http://genome.cbs.dtu.dk/services/TMHMM/>). The protein post-translational modification of glycosylation sites and phosphorylation sites were analyzed with DictyOGlyc 1.1 Server and NetPhos 2.0 Server (<http://www.cbs.dtu.dk/services/DictyOGlyc/>; <http://www.cbs.dtu.dk/services/NetPhos/>). The secondary structure was predicted by GOR method ([http://npsa-pbil.ibcp.fr/cgi-bin/npsa\\_automat.pl?page=npsa\\_gor4.html](http://npsa-pbil.ibcp.fr/cgi-bin/npsa_automat.pl?page=npsa_gor4.html)).

## Phylogenetic analysis

The relationship between Cfp450 and other insect P450s was determined using phylogenetic analysis. Amino acid sequences of insect P450 family 9 were aligned by the ClustalW program (Larkin *et al.* 2007). A maximum likelihood (ML) tree was constructed by MEGA5 with 1000 bootstrap replications (Tamura *et al.* 2011). Bayesian inference tree was also reconstructed using MrBayes v3.1.2 (Ronquist & Huelsenbeck 2003). Four independent Markov Chain Monte Carlo (MCMC) chains were run for 1,000,000 generations with the default temperature of 0.1 and parameter sampling occurring every 500 generations. Posterior probabilities for internal node were calculated

from the posterior density of trees.

### **mRNA expression assay by real-time PCR**

mRNA expression levels of CfP450 in different digestive organs were assessed by real-time PCR using a Bio-Rad CFX96 Real-Time PCR Detection system and a SYBR real-time premixture Kit (Biotek, China). The reverse transcription reaction was carried out as previously described to synthesize the first strand cDNA. CfP450 cDNA fragments were amplified by PCR with forward primer P450FT 5' - AAAGCCCGACATTATCCAGCAC-CTCAT -3' and reverse primer P450RT 5'- GTATCCAGTCCGGCCA-CAAGAAAC -3'. The product of cDNA fragments is 159 bp. Serial 5× dilutions of cDNA templates was performed in order to assess amplification efficiency of PCR. A pair of primer was designed to amplify termite actin as a reference gene to normalize the target gene expression levels. The primers were actinF, 5' - AAGCGCCTCTGAACCCAAAAGCAAA -3' and actinR, 5' - TGGCATGGTGCACAGCATAACCTTC- 3'. The product of termite actin is 204 bp.

Real-Time PCR of each cDNA sample and template free was performed in three independent replicates. The PCR reaction volume was 20 μL. The PCR protocol consisted of an initial denaturation at 95°C for 2 min, followed by 40 cycles: 95°C for 10 s, 55°C for 30 s, and 72°C for 30 s. To verify the specificity of the PCR amplification, a melting curve analysis was carried out after amplification by taking continuous fluorescence readings while increasing the temperature from 65 to 95°C in steps of 0.5°C with 5 s at each step. BioRad CFX Manager 2.0 software was used to determine the threshold cycle (Ct) value. The relative expression levels of target genes were calculated by the comparative CT method as described by Livak and Schmittgen (2001).

## **RESULTS**

### **Characterization of the CfP450**

The full-length *C. formosanus* CYP gene revealed by RACE was 2242 bp long. It includes a 1599bp open-reading-frame (ORF), a 36-bp 5'-untranslated region (UTR) and a 592-bp 3'-UTR (excluding the poly-A tail) which includes the canonical AATAAA polyadenylation signal at nucleotides 2208-2213. The CfP450 protein deduced from the ORF contains 532 amino

ATGGGAGCAAAAGTACCCGTCAGCAGCGTGATC

37 ATGGACTGGAACTGGGCGCTGTCAGACTGGATCGTCCTTTCGGCTGCCCTCGTGACAGCAGCATAC  
M D W N W A L S D W I V L L A A L V T A A Y

103 CTGTGGCGTACGTGGAGCCCAACCACTTCAAGAAGAGGAACTGGCTTACATCAGACTGCTCCA  
L C G T W S H N H F K K R N V P Y I R P A P

169 TTCTTTGGTAACAAGGGCCAGCAATACGCAGCAGAAACAAGGAACACTTCCCGACTACATACTG  
F F G N K R P A I R S R N K E H F P D Y I L

235 AGGACGTACAGAGAGCTACAGGGCCACGCGTACGGCGAAGCTTCAACTTATGCAGCCAGAGATA  
R T Y R E L Q G H A Y G G T F N F M Q P E I

301 ATCCTTCGGGACCTGAACTGATCAAAACATCACAGTGAAGGACTTCGAACATTCACGAACCCAC  
I L R D P E L I K T I T V K D F E H F T N H

367 GGGTCCTCTTAAACAACGCTACTGAGCCGATATGGGAACAATCTCTTCAAGTTTGCAGGCTCAG  
G S F L N N A T E P I W D K S L F S L S G Q

433 CGTTGGGAAGGACATGCCATCCACACTCAGTCTGCCTTCACATCCAGCAAGATGAAGGCAATGTT  
R W K D M R S T L L S P A F T S S K M K A M F

499 GGGCTGGTGTACAGTGTGCCAACAGTGGTAGACCATCTGGAGCAGTGTACGAACAGCCAAATC  
G L V S Q C C Q Q L V D H L E Q C Y E Q P I

565 GAACAGGGGTGTATATGCAGAAAGATGGCGAAATGTTAATACTGGAAGTGAACGACTTATATACC  
E Q G C D M Q K D G E M L I L E L N D L Y T

631 CGATATACCAATGACGTATCGCTACAGCCGCAATTTGGAAATGGAGTGGACTCTCTGAAACACCCG  
R Y T N D V I A T A A F G M E L D S L K H P

697 ACAAAAGAAATTTATATGATGGCAGAGAGCTGTTAAATTTGGAGTCTTATGGTGGCCAGGTC  
T N E F Y M M A Q R A V K I G S L S A A K S

763 AGCGGTTATTTGATCTCTCAAAGCTTACGAGTTATCGGGTTCACCAACATGTGGAAAGTGTG  
S G Y L I S P K L T Q L S G V T T M S K T V

829 ACCGAGTTTTTCAGATCAGTCAATGATGACACTCTCAAGACGAGAGAAGGGACACTGTAAAG  
T E F F R S V I D D T I S R R E K G R T V K

895 CCCGACATTTCAGCACCTCATAAGCGAAGAAAGGAGACACTAGAGATGCAACTCCACGGAA  
P D I I Q H L I Q A K K G D T R D A N S T E

961 AATGCCAGGACATCGATAATACACAAACTGAATGATGACGACATCGCTGCTCAAATTCCTGGG  
N A K D I D N T H K L N D D I A A Q I L A

1027 TTCTTGTGGCCGGACTGGATACAACATCCACACTTCTGTCCTTCGGCTCTCACCGATGGCAGTC  
F L V A G L D T S T L L S F A S H Q L A V

1093 TATCTGAAATACAAAGCCGACTTCAAGAAGAGATTGATGAAACCTCGAGGACATCGAGGGGAA  
Y P E I Q S R L Q E E I D E T L Q E H A G K

1159 TTCATACGAGGCTGTCAACGACATGAAGTACTGGGCGATGGTGTATCAGAGACACTGAGGATG  
F T Y E A V N S M K Y L G M V V S E T L R

1225 TTCCCACTCAGTGACAGCAGAAAGACTTGCATCAAACCTCACACCTGGACATCAATCTCTCT  
F P P T V T A E R L C I K P Y T L D I N P P

1291 CTGGAGCTTGAACCTGGGACAGACTCTTCAATCCAGTTTACGGACTGCACCATGACCCCATGTAC  
L E L E P G D R L F I P V Y G L H H D P M Y

1357 TATCTGATCCTGAGCGGTTTGACCCAGAACGCTTCTGTGATGAAACAAGCTGCATATCAACACT  
Y P D P E R F D P E R F C D E N K L H I N T

1423 TCAGCATACCTGCCCTTTGGATCAGGACCACAAAGCTGATTTGGCAACAGTTGCCACTGGTGGCA  
S A Y L P F G S G P Q S C I G N Q F A L V A

1489 TCCAAGCTGGTGGTGGTGCACCTGTTGCCGGTTCAACATCAGAGTACTGCAAGCAGCCAGCTG  
S K L V L V H L L S R F N I R V T A K T P L

1555 CCTATGAAAATAACAGACAGGATTCAACATGTCTGTGGAAAGGGGTTCTGGTTTGGGCTGGAG  
P M K I I Q T G F N M S V E G G F W F G L E

1621 CTCAGAAGTTCTGAAATGCAGCATCAAAATCACACCTGTCTCCCAAGCAGTAAATGCATTTAC  
L R S S \*

1687 TTTTACATTTTTAATTTCTTTATAGATGTAATAACAACACTACCTGCTCACATTTTTAATAACTGG  
1753 CTCTACGCAGAAATTTTTACTGTGATTGTTTCAGGATTCACCAAACTCTACAGCAGGGTTTA  
1819 CTGATATGCGTCTGCAGACAGAAAGCAGGCGCTAACCATGGTCAACTAAGCTAGAATAA  
1885 GGCATAAGCCTTATTTGGGATTCTGTGAAGATTGAGTCTCTGCTAGAAACAAGCTGCTGCC  
1951 TTCACAGCTCACCATGTATATGTGTACTTGTACTAACAGGAATGCTAATCACTTAGTTTCT  
2017 TGTATTTAAACAAAGGCTTCAACAGCAATTAACACAGAGTCCACATAAAACGATTTAATCTAC  
2083 ATCTATGCCACAACACACACCCCTAACCTTTCAGGCAGGAGGTAATGTATGCACCACCTAATT  
2149 AACATTAAGACTTTGCAATTTGCCAGAGTGTACATATATATGTGTACAATTTCAAAATAAAG  
2215 TGTTTATTTCTGCAAAAAAAAAAAAAA

Fig.1. The full-length cDNA sequence of *CYP9AX1* and its deduced amino acid sequence. The start codon ATG is underlined and the stop codon, TGA is indicated in bold and by an asterisk. The five conserved motifs of insect P450 enzymes are indicated by the boxed amino acids. The amino acids thickly underlined are the positions of transmembrane region, which also belongs to the region of Signal P prediction.

acids with a 20 amino acid predicted signal peptide at its N-terminal. The comparative analysis of the CfP450 protein sequence with other insect CYP genes showed that CfP450 has the highest homology to the CYP9 family. The CfP450 protein sequence was named CYP9AX1 (GenBank accession No. JN969113) by the P450 nomenclature committee (D. Nelson, personal communication).

The P450 amino acid sequences have some conserved signature motifs, though they have tremendous sequence diversity. In CYP9AX1, an FXXGXXXCXG (Nelson 1998) motif is presented at amino acid residues 468-477, and the cysteine residue in this motif is known as the important ligand for heme binding (Nebert & Gonzalez 1987). This motif is also characterized as FXXGXRXCXG (Feyereisen 1999), but in CYP9AX1, there is one amino acid residue change from Arg (R) to Gln (Q) at position 473 (Fig. 1). The helix C, WXXR (Rewitz *et al.* 2006), present at amino acid residues 134-138, appears as WKDMR. The helix I motif believed to create an oxygen-binding pocket, A/GGXE/DTT/S (Rewitz *et al.* 2006), present at amino acid residues 335-339, appears as AGLDTT. The Helix K motif, EXLR (Rewitz *et al.* 2006), present at amino acid residues 385-395, appears as YLGMVVSETLR. The PERF motif (Rewitz *et al.* 2006) present at amino acid residues 441-452, appears as YDPERFDPERF.

### **Bioinformatic analysis of CfCYP9AX1**

As in other microsomal proteins, in CYP9AX1, a signal sequence of approximately 20 amino acids is predicted by the Signal P analysis (Fig. 2). This signal sequence is from amino acid residues 1 to 20, and the cleavage site is between position 20 and 21. In this signal sequence, 13 hydrophobic residues out of 20 are extremely hydrophobic.

TMHMM software analysis showed that the CYP9AX1 protein has a highly hydrophobic N terminus transmembrane domain that functions as a membrane-anchor signal (Sakaguchi *et al.* 1987), with an area number from 7 to 29 amino acids (Fig. 3). The transmembrane domain of CYP9AX1 may also play a role as a membrane receptor or ion channel protein (Zhang *et al.* 2011).

A glycosylation site was predicted by DictyOGlyc projections at POS.282

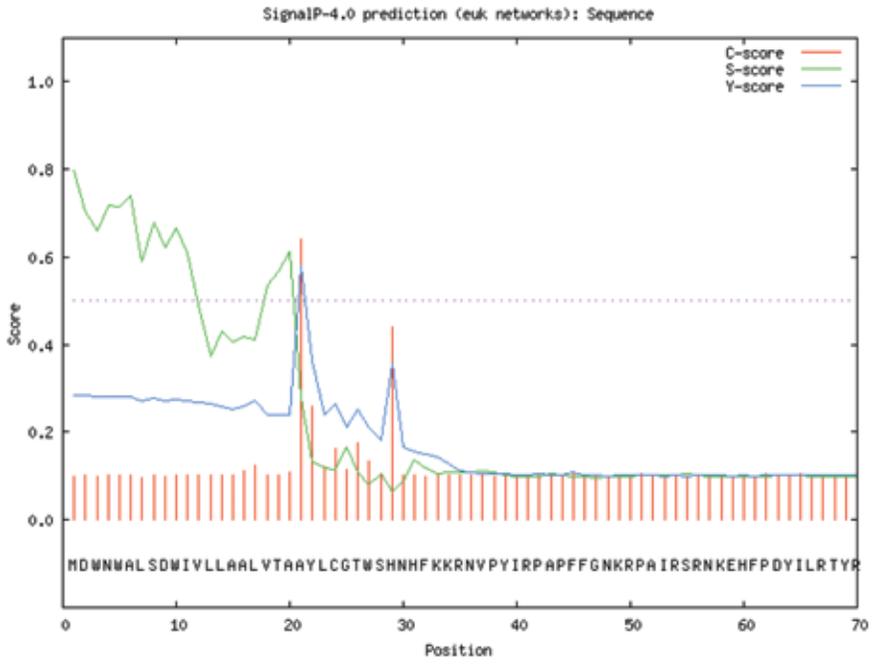


Fig. 2. Signal sequence prediction of CYP9AX1 by SignalP server.

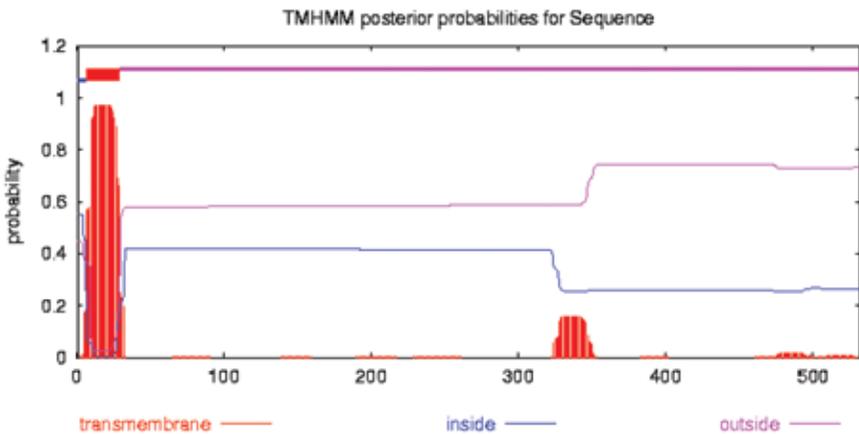


Fig. 3. Transmembrane domain analysis of CYP9AX1 with TMHMM software

Gin CYP9AX1 (Fig. 4). The phosphorylation site analysis of CYP9AX1 with NetPhos2.0 Server showed that there were 9 serine (Ser) phosphorylation sites, 9 threonine (Thr) phosphorylation sites and 5 tyrosine (Tyr) phosphorylation sites uniformly distributed throughout the polypeptide chain (Fig. 5).

GOR secondary structure analysis of CYP9AX1 protein showed that 34.77% of  $\alpha$  helix and 55.66% of random coil form the largest parts of the

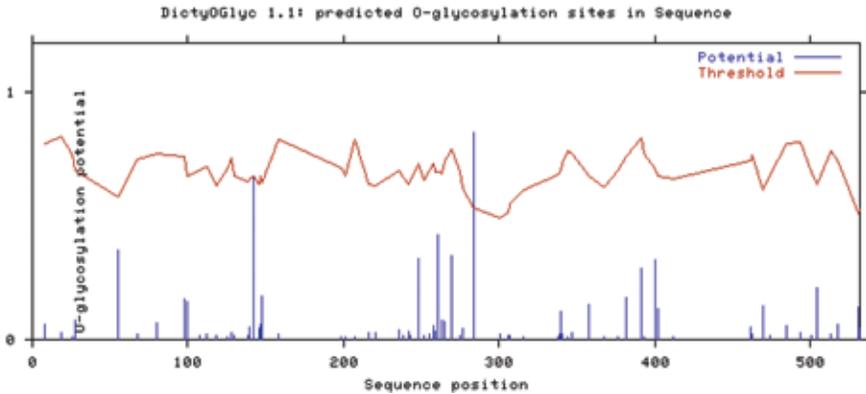


Fig. 4. O-glycosylation site prediction of CYP9AX1 by DictyOGlyc projections

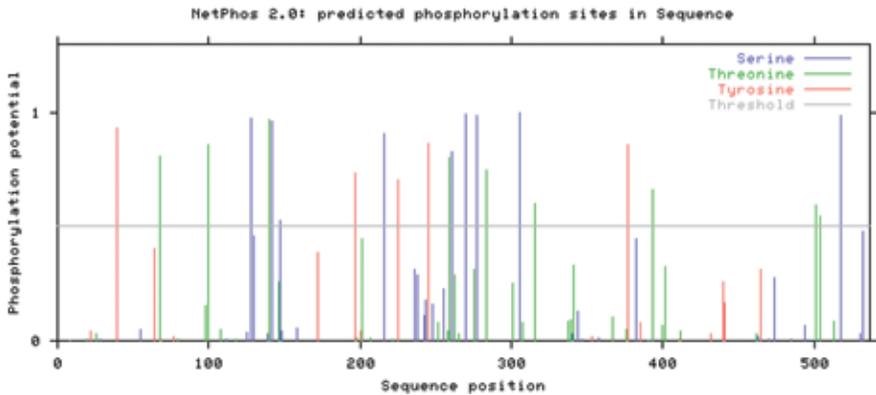


Fig. 5. Phosphorylation site prediction of CYP9AX1 by NetPhos2.0 Server

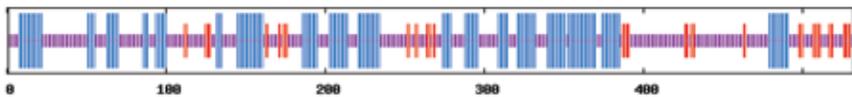


Fig. 6. GOR secondary structure analysis of CYP9AX1

structural elements of CYP9AX1 (Fig. 6). The extended strand accounting for 9.59% of CYP9AX1 was distributed throughout the protein (Fig. 6).

### Phylogenetic relationship of CfP450 with other insect CYP9 genes

The CfP450 and sixteen full length CYP9 genes were analyzed using MEGA 5 and MrBayes v3.1.2 software. The CYP9AX1 encoded protein has 51% identity to the German cockroaches *Blattella germanica* (AAK69410 and Q964T2) and *Diploptera punctata* (AAR97606). The phylogenetic trees showed that *C. formosanus* has high genetic relationship with *B. germanica* and *D. punctata* (Fig. 7). The maximum likelihood tree based on the CYP9

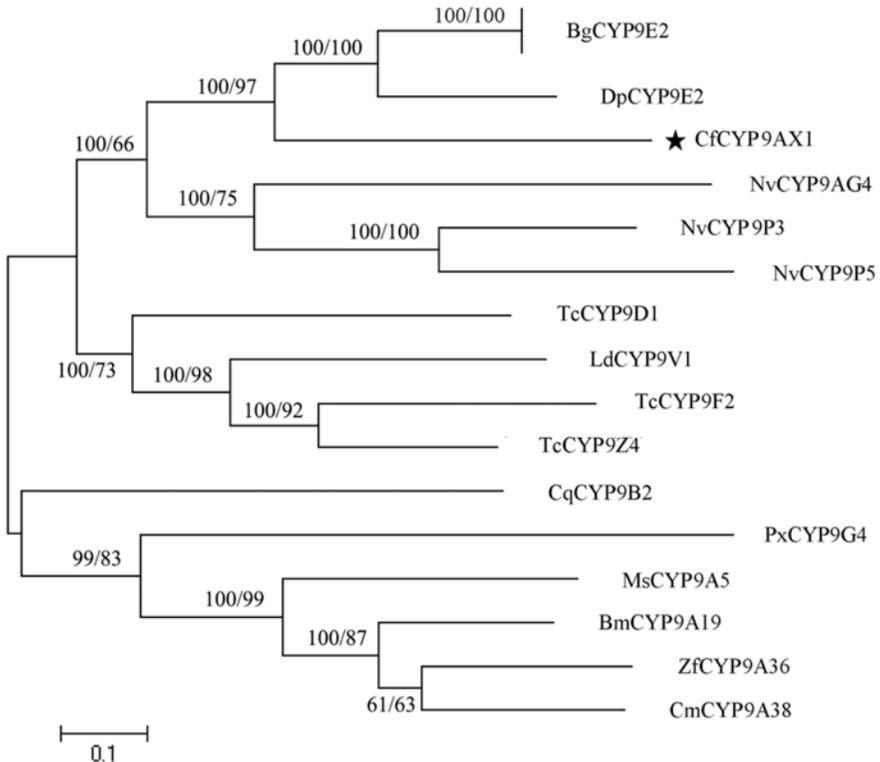


Fig. 7. Maximum likelihood tree showing the relationship between *C. formosanus* and other insect CYP9 proteins. The ML tree and Bayesian tree have identical topologies. Branch confidence (Bayesian posterior probability/ML bootstrap value) values are shown at the nodes. CfP450. The species abbreviations and the accession numbers are shown in Table 1.

proteins has the same topological structure as the Bayesian tree. The CfP450 has a long genetic distance between the subfamily that is also present in the phylogenetic tree.

### mRNA expression levels of CYP9AX1

We analysed *C. formosanus* CYP9AX1 transcription in the head, foregut, midgut, hindgut and malpighian tubules by Real Time-PCR. The relative expression levels of CYP9AX1 in different organs are illustrated in Figure 8. The mRNA levels in the foregut, midgut, hindgut and malpighian tubules were 1.6, 1.8, 2.5 fold and 10.6 fold higher than those in the head, respectively. This result indicated that CYP9AX1 was expressed most abundantly in malpighian tubules and expressed at roughly equal levels in the foregut, midgut and hindgut (Fig. 8).

## DISCUSSION

P450s are heme-thiolate proteins conducting oxidative attack of diverse

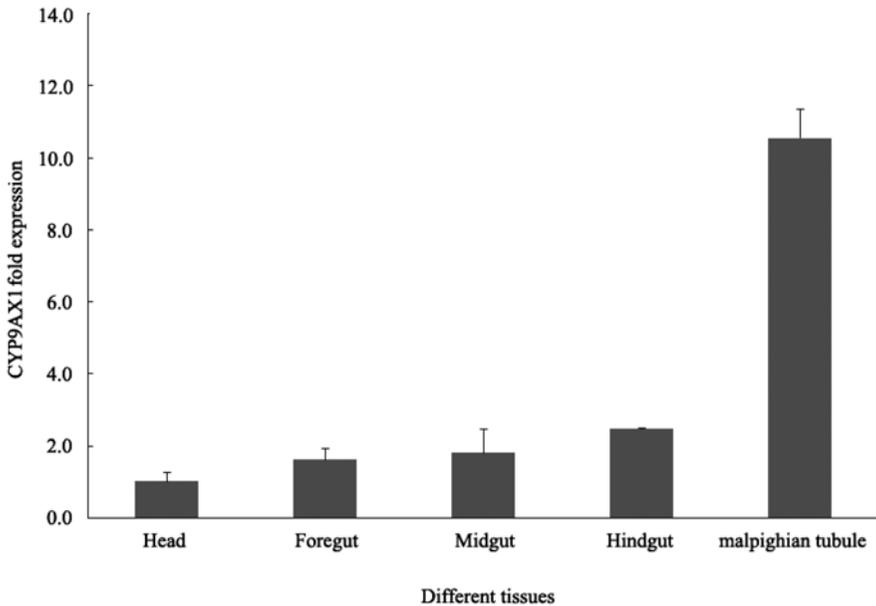


Fig. 8 Relative expression levels of CYP9AX1 in different organs. Real-time quantitative PCR data were evaluated using the formula  $2^{-\Delta\Delta Ct}$  ( $-\Delta\Delta Ct = Ct \text{ of actin} - Ct \text{ of CYP9AX1}$ ) to calculate the relative expression levels. Each column and bar represents the average of three amplification reactions.

Table 1. The species name, abbreviations and accession numbers of P450 genes

Species name	Gene name	Abbreviation	Accession number
<i>Blattella germanica</i>	CYP9E2	BgCYP9E2	AAK69410
	CYP9E2	BgCYP9E2	Q964T2
<i>Bombyx mori</i>	CYP9A19	BmCYP9A19	ABQ08709
<i>Coptotermes formosanus</i>	CYP9AX1	CfCYP9AX1	JN969113
<i>Cnaphalocrocis medinalis</i>	CYP9A38	CmCYP9A38	AZ65619
<i>Culex quinquefasciatus</i>	CYP9B2	CqCYP9B2	XP_001855237
<i>Diptera punctata</i>	CYP9E2	DpCYP9E2	AAR97606
<i>Leptinotarsa decemlineata</i>	CYP9V1	LdCYP9V1	AAZ94269
<i>Manduca sexta</i>	CYP9A5	MsCYP9A5	AAD51038
<i>Nasonia vitripennis</i>	CYP9P3	NvCYP9P3	NP_001165945
	CYP9P5	NvCYP9P5	NP_001166017
	CYP9AG4	NvCYP9AG4	NP_001166010
<i>Plutella xylostella</i>	CYP9G4	PxCYP9G4	ACH88357
<i>Tribolium castaneum</i>	CYP9D1	TcCYP9D1	NP_001034541
	CYP9Z4	TcCYP9Z4	NP_001164248
	CYP9F2	TcCYP9F2	NP_001127706
<i>Zygaena filipendulae</i>	CYP9A36	ZfCYP9A36	ACZ97417

substrates through catalyzing activation of molecular oxygen. They are functionally involved in carbon source assimilation, biosynthesis of hormones and degradation of xenobiotics (Werck-Reichhart & Feyereisen 2000). According to the identity at the amino acid level, P450s are grouped into different families (share >40% identity) and subfamilies (share >55% identity) (Nelson *et al.* 1996). The members of cytochrome P450 family 9 (CYP9) were well known to be related in the oxidative metabolism of insecticides (Feyereisen 1999). CYP9A1 is the first member of family 9 cloned from tobacco budworms, *Heliothis virescens* (Rose *et al.* 1997). To date, P450 variants belonging to the CYP9 family have been cloned from many insects such as cockroaches, ants, wasps and other insect species, and the majority of them are closely related to the metabolism of exogenous compounds such as plant toxins and insecticides (Hardstone *et al.* 2010; Itokawa *et al.* 2010; Mao *et al.* 2011; Stevens *et al.* 2000; Zhou *et al.* 2010). CYP9A13 from *Mamestra brassicae* is putatively involved in the metabolism of odorant compounds and xenobiotic clearance (Maibeche-Coisne *et al.* 2005). CYP9A20, a gene from *B. mori*, was presumably also related to insecticide resistance (Yamamoto *et al.* 2010).

Insect CYP genes can be detected in a wide range of tissues. Highest expres-

sion levels usually appear in the midgut, fat bodies and Malpighian tubules, but the expression level of individual CYP genes can vary (Scott *et al.* 1998; Scott & Wen 2001). For example, in *Bombyx mori*, CYP9A19 was detectable in the brain and midgut, while CYP9A20 and CYP9A21 were not detectable in the midgut (Ai *et al.* 2010). CYP9A12 mRNA was detected in all the tested tissue of *Helicoverpa armigera* larvae, whereas CYP9A17 was only detectable in the midgut and fat bodies (Zhou *et al.* 2010). The expression level can be inducible by exogenous substances such as insecticides (Stevens *et al.* 2000). In this study, CYP9AX1 was expressed most abundantly in the malpighian tubules, and less abundantly in the head, foregut, midgut and hindgut of *C. formosanus*.

The digestion of lignocellulose in termites requires enzymes to cooperate with each other. The hindgut of termites is the most important digestive location for lignocellulose. Before degeneration of the cellulose, delignification of biomass is very important. It is believed that many enzymes such as P450, laccase, catalase, and peroxidases are involved in the delignification process (Geib *et al.* 2008; Scharf & Boucias 2010). Ke *et al.* (2011) reported that the modifications of lignin began in termite foregut but happened mainly in the midgut, which indicated that the enzymes involved in lignin degradation or modification may be excreted in the foregut and midgut of the termite itself (Ke *et al.* 2011).

The role of P450s in the insect midgut has been well established in detoxification as well as possible pheromone synthesis (Aw *et al.* 2010). High transcription levels of CYP9AX1 in the Malpighian tubules could also correlate with detoxification given their role in metabolism and excretion of endogenous solutes and xenobiotics from biomass delignification processes (Chung *et al.* 2009). However, these findings need further functional validation. The cloning and characterization of CYP9AX1 will contribute to our understanding of the biological roles of P450s in lignocelluloses digestion in termites.

## ACKNOWLEDGMENTS

We thank Dr David R. Nelson for classifying the *C. formosanus* P450s before their publication. This work was supported by grants from the China

Postdoctoral Science Foundation funded project (No.20110491363), the Research Foundation for Advanced Talented Scholars of Jiangsu University (No.10JDG098), the National Basic Research Program of China (No.2012CB114600), the Natural Science Foundation of China (Project No.: 31170611), and the Zhejiang Provincial Natural Science Foundation of China (Project No.: Z3100211).

## REFERENCES

- Ai, J., Q. Yu, T. Cheng, F. Dai, X. Zhang, Y. Zhu & Z. Xiang 2010. Characterization of multiple CYP9A genes in the silkworm, *Bombyx mori*. *Mol Biol Rep* 37: 1657-1664.
- Aw, T., K. Schlauch, C.I. Keeling, S. Young, J.C. Bearfield, G.J. Blomquist & C. Tittiger 2010. Functional genomics of mountain pine beetle (*Dendroctonus ponderosae*) midguts and fat bodies. *BMC Genomics* 11: 215.
- Chung, H., T. Sztal, S. Pasricha, M. Sridhar, P. Batterham & P.J. Daborn 2009. Characterization of *Drosophila melanogaster* cytochrome P450 genes. *Proc Natl Acad Sci USA* 106:5731-5736.
- Feyereisen, R. 1999. Insect P450 enzymes. *Annu Rev Entomol* 44: 507-533.
- Geib, S.M., T.R. Filley, P.G. Hatcher, K. Hoover, J.E. Carlson, M. Jimenez-Gasco Mdel, A. Nakagawa-Izumi, R.L. Sleighter & M. Tien 2008. Lignin degradation in wood-feeding insects. *Proc Natl Acad Sci USA* 105: 12932-12937.
- Hardstone, M.C., O. Komagata, S. Kasai, T. Tomita & J.G. Scott 2010. Use of isogenic strains indicates CYP9M10 is linked to permethrin resistance in *Culex pipiens quinquefasciatus*. *Insect Mol Biol* 19: 717-726.
- Itokawa, K., O. Komagata, S. Kasai, Y. Okamura, M. Masada & T. Tomita 2010. Genomic structures of Cyp9m10 in pyrethroid resistant and susceptible strains of *Culex quinquefasciatus*. *Insect Biochem Mol Biol* 40: 631-640.
- Ke, J., D. Singh & S.L. Chen 2011. Aromatic compound degradation by the wood-feeding termite *Coptotermes formosanus* (Shiraki). *International Biodeterioration & Biodegradation* 65: 744-756.
- Larkin, M.A., G. Blackshields, N.P. Brown, R. Chenna, P.A. McGettigan, H. McWilliam, F. Valentin, I.M. Wallace, A. Wilm, R. Lopez, J.D. Thompson, T.J. Gibson & D.G. Higgins 2007. Clustal W and Clustal X version 2.0. *Bioinformatics* 23: 2947-2948.
- Livak, K.J. & T.D. Schmittgen 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2<sup>-</sup>( $\Delta\Delta C(T)$ ) Method. *Methods* 25: 402-408.
- Mao, W., M.A. Schuler & M.R. Berenbaum 2011. CYP9Q-mediated detoxification of acaricides in the honey bee (*Apis mellifera*). *Proc Natl Acad Sci U S A* 108: 12657-12662.
- Nebert, D.W. & F.J. Gonzalez 1987. P450 genes: structure, evolution, and regulation. *Annu Rev Biochem* 56: 945-993.
- Nelson, D.R. 1998. Cytochrome P450 nomenclature. *Methods Mol Biol* 107: 15-24.

- Nelson, D.R., L. Koymans, T. Kamataki, J.J. Stegeman, R. Feyereisen, D.J. Waxman, M.R. Waterman, O. Gotoh, M.J. Coon, R.W. Estabrook, I.C. Gunsalus & D.W. Nebert 1996. P450 superfamily: update on new sequences, gene mapping, accession numbers and nomenclature. *Pharmacogenetics* 6: 1-42.
- Rewitz, K.F., R. Rybczynski, J.T. Warren & L.I. Gilbert 2006. Identification, characterization and developmental expression of Halloween genes encoding P450 enzymes mediating ecdysone biosynthesis in the tobacco hornworm, *Manduca sexta*. *Insect Biochem Mol Biol* 36: 188-199.
- Ronquist, F. & J.P. Huelsenbeck 2003. MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* 19: 1572-1574.
- Rose, R.L., D. Goh, D.M. Thompson, K.D. Verma, D.G. Heckel, L.J. Gahan, R.M. Roe & E. Hodgson 1997. Cytochrome P450 (CYP)9A1 in *Heliothis virescens*: the first member of a new CYP family. *Insect Biochem Mol Biol* 27: 605-615.
- Scharf, M.E. & D.G. Boucias 2010. Potential of termite-based biomass pre-treatment strategies for use in bioethanol production. *Insect Science* 17: 166-174.
- Scott, J.G. 1999. Cytochromes P450 and insecticide resistance. *Insect Biochem Mol Biol* 29: 757-777.
- Scott, J.G., N. Liu & Z. Wen 1998. Insect cytochromes P450: diversity, insecticide resistance and tolerance to plant toxins. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* 121: 147-155.
- Scott, J.G. & Z. Wen 2001. Cytochromes P450 of insects: the tip of the iceberg. *Pest Manag Sci* 57: 958-967.
- Stevens, J.L., M.J. Snyder, J.F. Koener & R. Feyereisen 2000. Inducible P450s of the CYP9 family from larval *Manduca sexta* midgut. *Insect Biochem Mol Biol* 30: 559-568.
- Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei & S. Kumar 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28: 2731-2739.
- Werck-Reichhart, D. & R. Feyereisen 2000. Cytochromes P450: a success story. *Genome Biol* 1: REVIEWS3003.
- Zhang, Y.L., M. Kulye, F.S. Yang, L. Xiao, Y.T. Zhang, H. Zeng, J.H. Wang & Z.X. Liu 2011. Identification, characterization, and expression of a novel P450 gene encoding CYP6AE25 from the Asian corn borer, *Ostrinia furnacalis*. *J Insect Sci* 11: 37.
- Zhou, X., C. Ma, M. Li, C. Sheng, H. Liu & X. Qiu 2010. CYP9A12 and CYP9A17 in the cotton bollworm, *Helicoverpa armigera*: sequence similarity, expression profile and xenobiotic response. *Pest Manag Sci* 66: 65-73.



