Identification and Expression of Two Novel Cytochrome P450 Genes, CYP6CV1 and CYP9A38, in Cnaphalocrocis medinalis (Lepidoptera: Pyralidae)

Authors: Chen, Jun, Li, Chuan, and Yang, Zhifan

RESEARCH

Identification and Expression of Two Novel Cytochrome P450 Genes, *CYP6CV1* and *CYP9A38*, in *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae)

Jun Chen,[1,2] Chuan Li,[3] and Zhifan Yang[2,4]

[1]College of Chemical Engineering and Technology, Wuhan University of Science and Technology, Wuhan 430081, China
[2]Hubei Collaborative Innovation Center for Green Transformation of Bio-Resources, College of Life Sciences, Hubei University, Wuhan 430062, China
[3]School of Resources and Environment, Hubei University, Wuhan 430062, China
[4]Corresponding author, e-mail: sailyangzhf@qq.com

ABSTRACT. *Cnaphalocrocis medinalis* Güenée can cause severe losses in rice. Cytochrome P450s play crucial roles in the metabolism of allelochemicals in herbivorous insects. Two novel P450 cDNAs, *CYP6CV1* and *CYP9A38*, were cloned from the midgut of *C. medinalis*. *CYP6CV1* encodes a protein of 500 amino acid residues, while *CYP9A38*-predicted protein has 531 amino acid residues. Both cDNA-predicted proteins contain the conserved functional domains for all P450s. Phylogenetic analyses showed that CYP6CV1 is grouped in the cluster containing CYP6B members, while CYP9A38 is in the cluster including CYP9 members. However, both clusters are contained in the same higher lineage. Homologous analysis revealed that CYP6CV1 is most similar to CYP6B8, CYP6B7, CYP6B6, CYP6B2, and CYP6B4 with the highest amino acid identity of 41%. CYP9A38 is closest to CYP9A17, CYP9A21, CYP9A20, and CYP9A19 with the highest amino acid identity of 66%. Studies of temporal expression profiles revealed that *CYP9A38* showed a steady increase in mRNA level during the five instar stages, but a low-expression level in pupae, and then presented at a high-expression level again in adults. Similar expression patterns were obtained with *CYP6CV1*. In the fifth instar larvae, *CYP6CV1* was mainly expressed in midgut and fat bodies, whereas *CYP9A38* was mainly expressed in midgut. Expression studies also revealed a 3.20-fold over-expression of *CYP6CV1* and 3.54-fold over-expression of *CYP9A38* after larval exposure to host rice resistance. Our results suggest that both *CYP6CV1* and *CYP9A38* may be involved in detoxification of rice phytochemicals.

Key Words: *Cnaphalocrocis medinalis*, CYP6CV1, CYP9A38, host rice resistance, expression profile

The rice leaffolder *Cnaphalocrocis medinalis* (Güenée) (Lepidoptera: Pyralidae) is a species of the Crambidae family. It is considered a major pest of rice (*Oryza sativa* L.) (Riley et al. 1995). The attacked rice plants dry up and the vigor of plants reduces. The yield loss may vary up to 10–50%. And it is more problematic at booting stage (Riley et al. 1995). Serious outbreaks of *C. medinalis* have been reported in many Asian countries including India, Korea, Japan, China, Malaysia, Sri Lanka, and Vietnam (Senthil et al. 2006). And *C. medinalis* can undertake migrations which makes the pest control difficult (Riley et al. 1995).

Some rice varieties such as TKM-6 (*O. sativa* L.) show resistance to *C. medinalis* to a certain extent, while some like Taichung Native 1 (TN1) (*O. sativa* L.) are susceptible to this pest. In Asia, TN1 and TKM-6 have been the most frequently used in crossing programs (Upadhyay et al. 2011). TN1 was developed in Taiwan. It was the first semidwarf indica to respond to nitrogen fertilizer up to ~100 kg/ha. TN1 is short-statured (83–85 cm) and high tillering. But the area planted gradually declined, because TN1s principal deficiency is susceptibility to several insects and diseases (Dalrymple 1978). TKM-6 was developed in India. It has a weak stem. It was used as a parent for several high yielding Indian and International Rice Research Institute (IRRI) varieties. The discovery of resistant gene *Bph* 1 in TKM-6 is significant for rice breeders because the variety is resistant to several other diseases and insects (Dalrymple 1986, Khan and Joshi 1990). Therefore, both of TN1 and TKM-6 are important germ plasm resources in rice breeding, as well as vital materials in fundamental research. In this work, we use the two rice varieties as materials to rear or treat *C. medinalis* larvae.

The mechanism of resistance of TKM-6 to *C. medinalis* was attributed to a comparatively broader and thicker sclerenchymatous hypodermis in the stem, to the closer disposition of vascular bundles, the presence of more silicated cells in the leaf epidermis, more wax on leaf surface and narrower, and more hairy leaves (Chandramani et al. 2009). One other factor responsible for resistance in TKM-6 seems to be of biochemical nature. The changes of various biochemical constituents such as leaf soluble protein, phenol, ortho-dihydroxy phenol, tannin, and enzymes viz., peroxidase, phenyl alanine ammonia lyase (PAL) were ever assessed spectrophotometrically in the rice genotypes before and after *C. medinalis* infestation. The protein profile was analyzed using sodium dodecyl sulphate-poly acrylamide gel electrophoresis (SDS-PAGE) method. A significant constituent of biochemical content such as tannin, phenol, and ortho-dihydroxy phenol was proved increased along with enzyme activities of peroxidase and PAL in the infested TKM-6 rice (Punithavalli et al. 2013). It is also evident that there are more biochemicals such as phenol, orthodihydroxy phenol, and tannin in TKM-6 than in TN1 plants, which were negatively correlated with *C. medinalis* damage. However, leaf protein content was less in TKM-6 than in TN1, which was positively correlated with the damage by *C. medinalis* (Punithavalli et al. 2013). Meanwhile, the increased death of early instars, the slower development of larvae, the reduced size or pupal stage, and abnormal behavior of *C. medinalis* were watched on TKM-6, when compared with that on the other rice varieties (Khan et al. 1989, Masoud et al. 1996).

It is well accepted that planting resistant rice varieties is an effective practice to control some of rice pests (Sogawa et al 2003). But no rice varieties with sufficient resistance level have been developed to control *C. medinalis* in practical application to date (Rao et al. 2010). Why *C. medinalis* has so strong adaptation to its host rice remains unknown. Therefore, to research the molecular interaction of *C. medinalis* and its host rice resistance may obtain some crucial information to develop novel integrated pest managements (IPMs).

The P450s are important metabolic systems in insects because of their involvement in the syntheses of endogenous hormones, fatty acids

and steroids, and in the catabolism of xenobiotics, such as drugs, pesticides, and plant toxins (Scott 2008). Multiple P450 family genes were found in herbivorous insects, which represent an adaptation in the "animal–plant warfare" (Gonzalez and Nebert 1990). Ingestion of plant toxins frequently induces insects P450 genes responsible for catabolism of plant toxins. Some examples of insect P450 genes induced by phytochemicals are summarized in Feyereisen (2005), Schuler (2011), and Scott (2008). As noted previously, we found that *CYP6AE28* and *CYP6AE30* of *C. medinalis* are induced in response to resistant rice variety TKM-6 (Liu et al. 2010). In this study, we cloned another two P450 genes *CYP6CV1* and *CYP9A38* from *C. medinalis* and studied their expression profiles, in order to establish a foundation for further study of their functions.

Materials and Methods

Experimental Insects and RNA Isolation. Insect cultures: *C. medinalis* larvae were collected from paddy fields in the Wuchang district, Wuhan City, China, and reared in containers on TN1 rice in booting stage, a leaffolder susceptible rice variety, at 25°C under a photoperiod of 16:8 (L:D) h, and 80% relative humidity.

Treatment with TKM-6 rice: For induction analyses of the P450 genes, the newly molted fifth instar larvae were transferred from TN1 plants to TKM-6 rice in booting stage and kept for 24 h. Whole insect bodies were collected and deeply frozen in liquid nitrogen then stored in −80°C refrigerator for further use.

RNA isolation and cDNA synthesis: For temporal expression analyses of the P450 genes, whole bodies of the newly developed larvae in the first, second, third, fourth, fifth instar stages, pupae, and adults were collected for total RNA isolation by using TRIzol reagent (Invitrogen, Carlsbad, CA). For spatial expression analyses, head, midgut, fat body, and carcass dissected from the fifth instar larvae treated with TKM-6 were used for RNA isolation. For induction analyses, the fifth instar larvae fed with TN1 and TKM-6 plants were used for RNA extraction. Half microgram of total RNA was used for cDNA synthesis using the iScriptTM cDNA synthesis kit (BioRad, Hercules, CA).

Cloning of the Full-Length cDNAs. A pair of degenerate oligonucleotide primers (dCYP6-S and dCYP6-AS, Table 1) was designed for amplification of cytochrome P450 family six genes (Kasai et al. 2000). Another pair of degenerate primers (dCYP9-S and dCYP9-AS, Table 1) was used to amplify cytochrome P450 family nine genes (Stevens et al. 2000). The cDNAs transcribed from the midgut, RNA were used as template in the reverse transcription-polymerase chain reactions (RT-PCRs). Two P450 clones representing two novel P450 genes were selected after cloning, sequence and analyzing for CYP6 and nine homologies with the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST)

Network Server. Gene-specific primers (GSPs) for 5′ and 3′ rapid-amplification of cDNA ends (RACE) were designed according to the two cDNA clones. The primers were rCYP6-S, rCYP6-AS, rCYP9-S, and rCYP9-AS (Table 1). The cDNA synthesis and RACE were performed exactly according to the instruction manual of the SMART RACE cDNA Amplification Kit (BD Bioscience Clontech, Palo Alto, CA).

Amplified fragments were routinely cloned into pGEM-T vector (Promega, Madison, WI) and sequenced. The nucleotide sequences of the 5′ and 3′ RACE products were aligned to form two full-length cDNA sequences and the cDNA-predicted proteins were called as CYP6CV1 and CYP9A38 by the P450 nomenclature committee (D.R. Nelson), respectively. Two pairs of primers for long distance-PCR (LD-PCR) were designed to amplify the internal sequences of the full cDNAs, respectively. The primers were lCYP6-S, lCYP6-AS, lCYP9-S, and lCYP9-AS (Table 1). The PCR system was heated at 95°C for 1 min and then amplified for 34 cycles (95°C for 30 s, 62°C for 30 s, and 72°C for 3 min). Amplified fragments were cloned and sequenced.

Computer-Assisted Analysis of P450 cDNAs. Molecular mass and isoelectric point were predicted by Compute pI/Mw tool (http://us.expasy.org/tools/pi_tool.html). The transmembrane anchors of the deduced P450s were predicted by the TMpred server (http://www.ch.embnet.org/software/TMPRED_form.html). ClustalX v1.8 (Thompson et al. 1997) was used to analyze the alignment. A molecular phylogenetic tree was constructed by the ClustalW Server (http://crick.genes.nig.ac.jp/homology/clustalw-e.shtml) by using the bootstrap N-J tree option (number of bootstrap trials = 1,000; Page 1996).

Real-Time Quantitative PCR Analysis. The transcript levels of *CYP6CV1* and *CYP9A38* were determined by real-time quantitative PCR (qPCR), using iQ SYBR Green Supermix (Bio-Rad, Hercules, CA) in the iCycler iQ Real Time PCR Detection System (Bio-Rad, Hercules, CA). To standardize qPCR inputs, a master mix that contained iQ SYBR Green Supermix and forward and reverse primers was prepared (final concentration = 100 nM per reaction; primer sequences [qCYP6-S, qCYP6-AS, qCYP9-S, and qCYP6-AS] are listed in Table 1). The qPCRs were conducted with the same quantity of 10-fold diluted cDNA for each instar larva, pupa, and adults for analyses of temporal gene expression profiles. The reactions were also carried out with equivalent cDNA input for the organs, including head, midgut, fat body, and carcass from the fifth instar larvae for analyses of spatial gene expression profiles. In addition, equal cDNA input of the fifth instar larvae ingested with TN1 and TKM-6 plants were used for analyses of host resistance induction. PCR profiles were: 95°C for 5 min; 40 cycles of 95°C for 15 s, 55°C for 10 s, and 72°C for 10 s with a plate read at the end of each cycle. All reactions were performed in duplicates and three replicates were used to estimate variation.

Table 1. Primers used in the present study

Primer set	Primer sequence (5′-3′)	Application type	Product length [bp]
dCYP6-S	GA(A or G)AC(A or G or C or T)(A or C or T)(C or T)(A or G or C or T)(A or C) G(A or G or C or T)CC(A or G or C or T)(G or T)C	RT-PCR	239
dCYP6-AS	GG(A or G or C or T)CC(A or G or C or T)(G or T)C(A or G or C or T)CC(A or G)AA(A or G or C or T)GG		
dCYP9-S	TACGA(AG)(CT)IGCI(AG)(AT)IAA(CT)CC(CT)GA	RT-PCR	404
dCYP9-AS	CCIA(GT)(AG)CA(AG)TTIC(GT)IGGICC		
rCYP6-S	TGGCTGATTACACGTTTCCTGGAACTGA	RACE	856
rCYP6-AS	TCGGAAAATATTCTGGGTCGGCGTTA	RACE	1,258
rCYP9-S	CTGGTATCGCGATGGACCGGATCTGC	RACE	569
rCYP9-AS	CTCGGAGAAGCGCTCGGGGTCAAACT	RACE	1,427
lCYP6-S	CTCCAACATGGCGCTGCTCGTG	LD-PCR	1,636
lCYP6-AS	CTCAATTCACGCCATCGCCTTC		
lCYP9-S	GCGCAAACCGGCCTGAGCCATG	LD-PCR	1,554
lCYP9-AS	GGACAGCTGGGCAGGGATGCTG		
qCYP6-S	ACGTTTCCTGGAACTGATGTCA	qPCR	198
qCYP6-AS	CAAACCAATGCAATTCCGAGGT		
qCYP9-S	CAGATCCTCACCTTCTTCGCTT	qPCR	188
qCYP9-AS	CGACCTTTCTTAGCCTCCATGA		

S and AS indicate forward primer and reverse primer, respectively. Y = C or T, K = T or G.

Beta-*actin* was used as an internal reference. Relative accumulation of *CYP6CV1* and *CYP9A38* normalized against beta-*actin* was calculated from the formula $2^{-\Delta\Delta Ct}$ where two is the reaction efficiency and $\Delta\Delta Ct$ is the difference in beta-*actin* Ct values between a defined control and the rest samples in an assay, or the difference in *CYP6CV1* or *CYP9A38* Ct values between a defined control and the rest samples in an assay.

Results

cDNA Cloning and Characterization. Two pairs of degenerate primers for the conserved regions of insect P450 proteins were used to amplify reverse-transcribed midgut mRNA of *C. medinalis*. The PCR products were cloned and sequenced. A clone with 239 bp in length encoding a reading frame of 79 amino acids was amplified by using dCYP6-S and dCYP6-AS; another clone with 404 bp encoding 134 amino acids was obtained by using dCYP9-S and dCYP9-AS as primers. Both clones scored highly with other P450 proteins by BLAST Network searches. Based on the two cDNA clones, GSPs were designed. Two P450 cDNAs with 2,041 and 1,853 bp in length were amplified by RACE and LD-PCR strategies. Close alignment with known insect CYP6 and CYP9 members showed that the two new P450s belonged to the two families, respectively. The former sequence (accession number FN421127) was called CYP6CV1, and the later (accession FN421128) was called CYP9A38.

CYP6CV1 nucleotide sequences analysis revealed that this cDNA contains a putative ORF of 1,503 bp, a 21 bp 5′-untranslated region (5′-UTR), and a 517 bp 3′-UTR, with a putative polyadenylation signal sequences (AATAAA) upstream of the poly(A) tract (Fig. 1). CYP9A38 cDNA is composed of a 1,593 bp ORF, with a 63 bp 5′-UTR

and a 197 bp 3′-UTR with a polyadenylation signal (AATAAA) upstream of the poly(A) tract (Fig. 2).

Characterization of the cDNA-Predicted Proteins. Based on the predicted amino acid sequence, CYP6CV1 has a theoretical *pI* value of 8.95 and molecular mass of 56,886, and CYP9A38 has a theoretical *pI* value of 7.67 and molecular mass of 61,454. Two strong inside-to-outside transmembrane helices from amino acid 2 to 20 and 293 to 315 in CYP6CV1 sequences, and three transmembrane helices from amino acid 1 to 18, 215 to 234 and 319 to 342 in CYP9A38 sequences were predicted, suggesting that both P450s are endoplasmic reticulum membrane-bound proteins. The putative proteins contain the typical motifs of an insect P450 protein (Nelson et al. 1993), including the proline and glycine rich hinge region (xxPxPxxGxx), helix-C (WxxxR), helix-I (AGxE or DT), helix-K (ExxR), PExF (PxxFxPExF), and the heme-binding domain (FxxGxxxCxG) (Figs. 1 and 2). Figure 3 shows the alignment of CYP6CV1, CYP6AE28, CYP6AE30, and the other four CYP6 family members (CYP6A2, CYP6B2, CYP6AE1, and CYP6AE12). All of them contain the six residue sequence PExFxP (PENFSP, position 419 to 424 in CYP6CV1) upstream of the heme-binding domain, specific to family six members (Nelson et al. 1993). Likewise, the SR(F or I or L)(A or G)xx(Q or E) sequence immediately following the heme-binding domain, specific to family 9, is found in CYP9A38 (SRFALCE, position 478 to 484; Maïbèche-Coisne et al. 2005).

The relatedness of CYP6CV1, CYP6AE28, CYP6AE30, CYP9A38, and some CYP6 family P450s is revealed by the fact that CYP6CV1 and CYP6CV2 form an independent cluster contained in the clade including CYP6A2 and CYP6B members. CYP9A38 is in the cluster including other six CYP9 members. A higher lineage containing
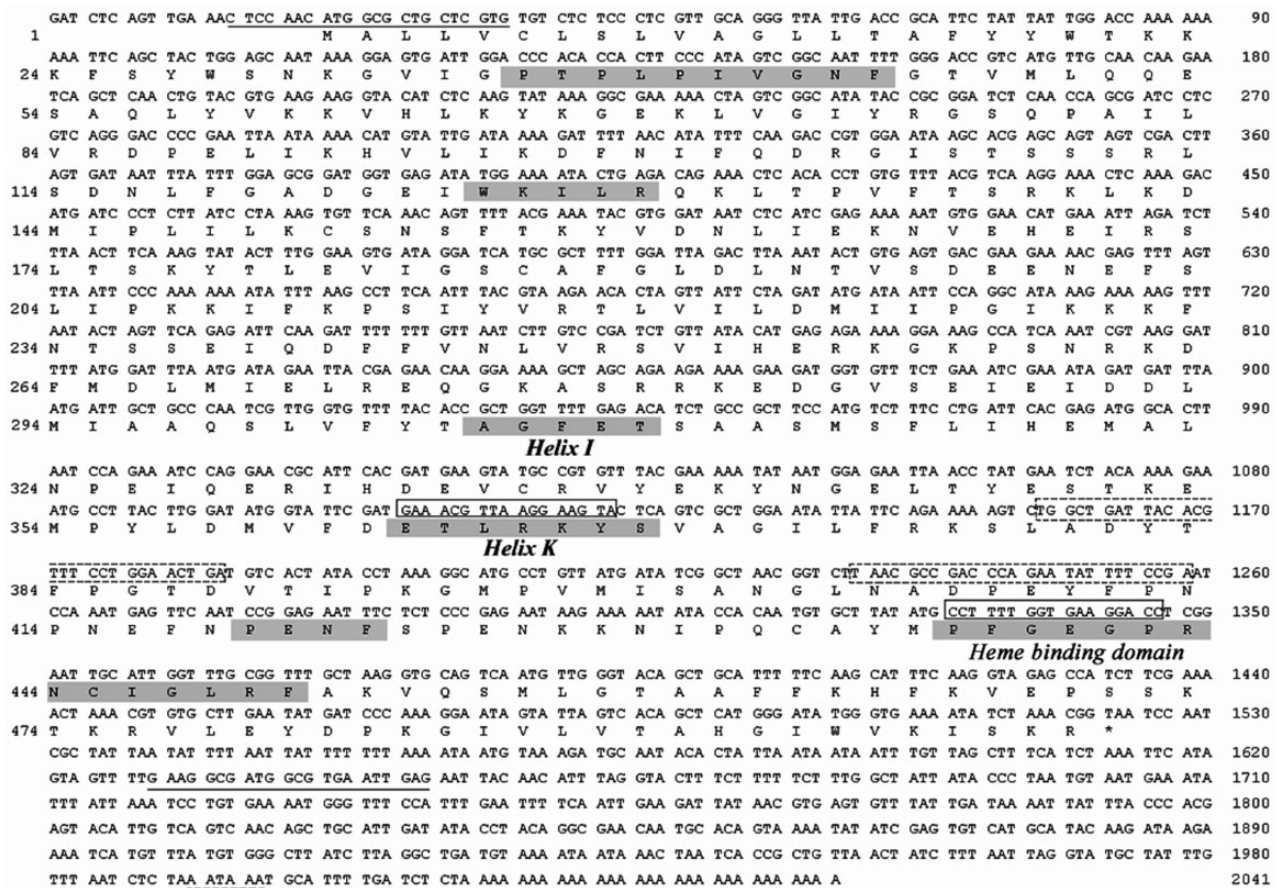
Fig. 1. Nucleotide sequence and deduced amino acid sequence of *CYP6CV1* in *C. medinalis*. The proline and glycine rich hinge region, C-helix, I-helix, K-helix, PExF, and heme-binding domain are shaded. The consensus polyadenylation signal is indicated by a discrete underline. The degenerate primer corresponding regions are framed. The GSPs are framed with discrete line. The LD-PCR primers are underlined.

```
      GTG TGT TAA GGT TGA AAA TTG TGA CAA GTT GTA TTT GTG TCA GTG CGC AAA CCG GCC TGA GCC ATG TTG ATG TAC GTA TGG CTG GCG GCT   92
   1                                                                                M   L   M   Y   V   W   L   A   A
      ATG GCC GCC GCA GTG TGG CTG TAC TGC AGG CAG ATC TAC TCG AAG TTC TCC AAG GCG GGG GTG AAG CAC TTC AAG CCG CTG CCG ATC GTG   182
  10  M   A   A   A   V   W   L   Y   C   R   Q   I   Y   S   R   F   S   K   A   G   V   K   H   F   K   P   L   P   I   V
      GGC AAC ATG CTG AAG ATC CTA CTG AGG ATG GAT CAC TTC ACG GAT AAC ATC GAG AGC TTG TAC TTC GCT TAT CCT GAT GAG AAG TTC GTA   272
  40  G   N   M   L   K   I   L   L   R   M   D   H   F   T   D   N   I   E   S   L   Y   F   A   Y   P   D   E   K   F   V
      GGG AGG TAC GAG TTC ATC AAC CCC ATG GTC ATG ATT AAG GAT GTG GAC CTG CTG AAG AAG ATC ACC GTC AAG GAC TTT GAG CAC TTC ATC   362
  70  G   R   Y   E   F   I   N   P   M   V   M   I   K   D   V   D   L   L   K   K   I   T   V   K   D   F   E   H   F   I
      GAC CAC AGG ACC CTC GTC AAT GAG AAG ACT GAC CCG TTC TTC GGG AGG AAC TTG TTC TCT TTA AAA GGT GAC GAA TGG AAA GAC ATG CGC   452
 100  D   H   R   T   L   V   N   E   K   T   D   P   F   F   G   R   N   L   F   S   L   K   G   D   E   W   K   D   M   R
      TCT ACC CTG AGC CCG GCC TTC ACC AGC TCC AAG ATG AAG CTG ATG CTG CCC TTC ATG GTG GAA GTG GGC AAT CAG ATG GTG GAC TCA TTG   542
 130  S   T   L   S   P   A   F   T   S   S   K   M   K   L   M   L   P   F   M   V   E   V   G   N   Q   M   V   D   S   L
      AAG AAA AAG ATT AAG GAA TCA AAC GCC ACC CAC ATC GAA ATC GAC GCT AAG GAC CTT ACC ACC CGC TAC GCC AAT GAC GTC ATC GCT TCC   632
 160  K   K   K   I   K   E   S   N   A   T   H   I   E   I   D   A   K   D   L   T   T   R   Y   A   N   D   V   I   A   S
      TGC GCC TTC GGT CTG AAG GTG GAC TCG CAG ACG GAC ACT GAG AAC CAT TTC TAT GTG ATG GGG AAG AAC ACC ACC GCT TTC AAC TTC CGT   722
 190  C   A   F   G   L   K   V   D   S   Q   T   D   T   E   N   H   F   Y   V   M   G   K   N   T   T   A   F   N   F   R
      CAG ATC CTC ACC TTC TTC GCT TTA GCC AAC TTC CCT AAG ATT ATG GAT AAA TTT AAA GTG ACG CTG TTC ACG GAA AAC ACC AGG CAG TTC   812
 220  Q   I   L   T   F   F   A   L   A   N   F   P   K   I   M   D   K   F   K   V   T   L   F   T   E   N   T   R   Q   F
      TTC CAG CAT CTG GTG ATG GAC ACG ATG AAC GAG CGA GAG CTG AAG AAG ATC ATC AGA CCT GAC ATG ATC CAT CTT CTC ATG GAG GCT AAG   902
 250  F   Q   H   L   V   M   D   T   M   N   E   R   E   L   K   K   I   I   R   P   D   M   I   H   L   L   M   E   A   K
      AAA GGT CGC CTG ACG CAC GAT GAC AAG GCT TCT CAC GAC CCT GAT GCA GGC TTT GCT ACT GTT GAA GAA TCT AAC ATC GGA AAG AAA AAC   992
 280  K   G   R   L   T   H   D   D   K   A   S   H   D   P   D   A   G   F   A   T   V   E   E   S   N   I   G   K   K   N
      ATT AAC CGA GAA TGG TCC GAC GAC CTA ACA GCC CAA GCC TTC CTC TTC TTC GCC GGC TTC GAA ACC GTC TCC TCT GCC ATG TCC   1082
 310  I   N   R   E   W   S   D   D   L   T   A   Q   A   F   L   F   F   A   G   F   E   T   V   S   S   A   M   S
                                                                          Helix I
      TTC GCC CTG TAC GAG CTG GCC GTT AAC CCT GAC GTG CAG GAG AAG TTT CTA GAA GAG ATC AAG GAG GCC GAC GCC AAG AAT GGC GGA AAG   1172
 340  F   A   L   Y   E   L   A   V   N   P   D   V   Q   E   K   F   L   E   E   I   K   E   A   D   A   K   N   G   G   K
      TTT GAC TAC AAC TCC ATA CAG AAC ATG ACT TAT ATG GAC ATG GTT GTC TCA GAG GTG CTC AGG CTG TGG CCG CCT GGT ATC GCG ATG GAC   1262
 370  F   D   Y   N   S   I   Q   N   M   T   Y   M   D   M   V   V   S   E   V   L   R   L   W   P   P   G   I   A   M   D
                                                                          Helix K
      CGG ATC TGC GTT AAA GAC TAC AAC CTT GGA AAA CCC AAT GAC AAG GCT ACT TCT GAC TAC ATT ATC CGC AAA GGC GAG TCC ATT ATG ATC   1352
 400  R   I   C   V   K   D   Y   N   L   G   K   P   N   D   K   A   T   S   D   Y   I   I   R   K   G   E   S   I   M   I
      CCA GCT GGC ATA GAT ATA CAC GAC CCG AAG TTC TTC CCA AAC CCA ATG AAG TTT GAC CCC GAG CGC TTC TCC GAG GAG AAC AAA CAC AAC   1442
 430  P   A   W   A   I   H   H   D   P   K   F   F   P   N   P   M   K   F   D   P   E   R   F   S   E   E   N   K   H   N
      ATG GAT GTA ACT GCG TAC ATG CCC TTT GGA GTT GGC CCT AGG AAT TGC ATT GGT TCC AGG TTC GCT CTC TGC GAG GTG AAG GTG ATG TTA   1532
 460  M   D   V   T   A   Y   M   P   F   G   V   G   P   R   N   C   I   G   S   R   F   A   L   C   E   V   K   V   M   L
                                                      Heme binding domain
      TAC CAG CTG CTC CTC CAC ATG GAG ATC TCT CCA TCG CCC AGG ACC AGC ATC CCT GCC CAG CTG TCC AAG GAA ACC TTC AAC GTC CGA ATT   1622
 490  Y   Q   L   L   L   H   M   E   I   S   P   S   P   R   T   S   I   P   A   Q   L   S   K   E   T   F   N   V   R   I
      AAG GGA GGA CAC TGG CTC AAC TTC AGG AGT AGG ACT TAG AGC GCA TTT ACA TTG GTC GGT GCG CAA TTA CAT TGG TCT TTA CCG TAC ACT   1712
 520  K   G   G   H   W   L   N   F   R   S   R   T   *
      AAA TGA TTA GCA CGA ATA GAG GGA GAA GCA GGA AAG CCC TCA TTT GAG CTA ATA ACC GAC TTG TCA ACC TGA ATA AAT GCC CCC CTT TTA   1802
      ACG AAA ATG CTG ATA TTA GCA TCC AAC ACA AAA AAA AAA AAA AAA AAA AAA   1853
```

Fig. 2. Nucleotide sequence and deduced amino acid sequence of *CYP9A38* in *C. medinalis*. The conserved domains, polyadenylation signal, and primers are marked as in Fig. 1.

the above two clusters is formed in the phylogenetic tree. CYP6AE28 and CYP6AE1 form another independent cluster, while CYP6AE30 is in the cluster containing CYP6AE9 and CYP6AE47 as shown in Fig. 4.

Homologous analysis revealed that CYP6CV1 shows the highest percentage amino acid identities (41%) to *Helicoverpa zea* CYP6B8 (AF285828), *Helicoverpa armigera* CYP6B7 (ABE60887), CYP6B6 (ABE60886), CYP6B2 (ABE60885), and *Papilio glaucus* CYP6B4 (AAB05892), respectively. CYP9A38 is closest to members of the CYP9A subfamily. It shares the highest identity (66%) with CYP9A17 (ACJ37388) from *H. armigera*, CYP9A21 (ABN71369), CYP9A20 (ABO07439), CYP9A19 (ABQ18318) from *Bombyx mori*, respectively.

Expression Analyses of CYP6CV1 and CYP9A38. We analyzed the mRNA levels of *CYP6CV1* and *CYP9A38* across insect life stages by quantitative RT-PCR (Fig. 5A). Both gene mRNA level exhibited a cyclic pattern. *CYP9A38* showed a steady increase in mRNA level during the five instar stages, but a low-expression level in pupae, and then presented at high-expression level again in adults. Similar expression patterns were obtained with *CYP6CV1*. *CYP6CV1* was significantly activated in the fourth, fifth instar larvae, and the adults. Namely, *CYP9A38* expression was 4.43-fold, 4.89-fold, 8.53-fold, 9.48-fold, 6.77-fold, and 4.50-fold comparatively to pupae for first, second, third, fourth, fifth larval stages and adults, respectively. *CYP6CV1* was expressed by 0.61-fold, 1.46-fold, 1.61-fold, 3.50-fold, 3.35-fold, and 3.44-fold comparatively to pupae for first, second, third, fourth, fifth larval stages and adults, respectively.
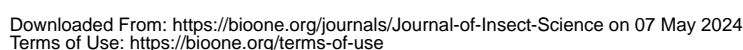
In the fifth instar larvae, *CYP6CV1* was mainly expressed in midgut and fat bodies, with 0.14-fold, 1.41-fold, and 0.27-fold comparatively to fat bodies for head, midgut, and carcass, respectively. Whereas *CYP9A38* was mainly expressed in the midgut, with 0.38-fold, 4.72-fold, and 0.32-fold comparatively to fat bodies for head, midgut, and carcass, respectively (Fig. 5B).
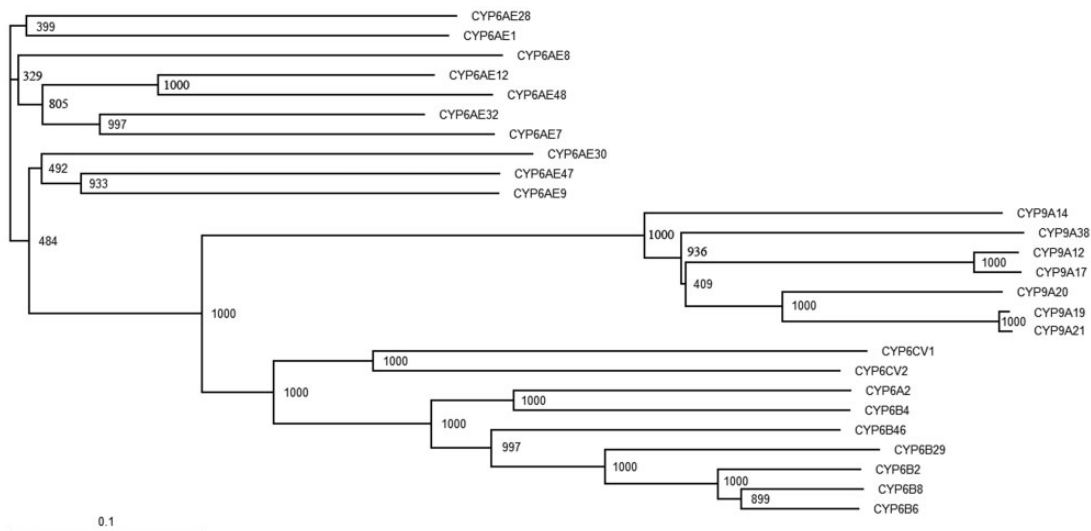
In the fifth instar larvae reared on the susceptible rice TN1 plants, both *CYP6CV1* and *CYP9A38* were constitutively expressed at low levels, but significantly induced by exposure to the resistant rice variety TKM-6 (Fig. 6). Expression studies revealed a 3.20-fold over-expression of *CYP6CV1* and 3.54-fold over-expression of *CYP9A38* after larval exposure to host rice resistance.

Discussion

During the last three decades, people have focused interest in insect P450s on their role in fundamental physiological functions, such as growth, development or reproduction through the biosynthesis and the catabolism of key hormones, such as juvenile hormone (JH) or 20-hydroxyecdysone, in the oxidative metabolism of various xenobiotics including insecticides and plant phytochemicals (Feyereisen 2005). Most of the CYP6 family members studied to date, especially in crop pests, are generally expressed in the digestive tract and fat body, and were found to be mainly responsible for insecticide metabolism and inactivation of phytochemicals (Li et al. 2002, Feyereisen 2005, Scott 2008). The insect CYP9 family is most closely related to the CYP6 family, and together with the CYP28 family, they are grouped in the CYP3 clan including some mammal P450s (Nelson 1998). *CYP9A1* from *Heliothis virescens*, the first member of this family, was found to be constitutively over-expressed in thiodicarb-selected tobacco budworms and may play a role in pesticide metabolism (Rose et al. 1997). *CYP9A2* was activated by the wild tomato compound 2-undecanone, indole-3-carbinol, phenobarbital, 2-tridecanone, and xanthotoxin. *CYP9A4* and *CYP9A5* were induced differentially by clofibrate and xanthotoxin (Stevens et al. 2000). *CYP9A13* may probably involve in the metabolism of odorant compounds and play a role in taste in the moth *Mamestra brassicae* (Maïbèche-Coisne et al. 2005). *CYP9A12* and *CYP9A17* mRNA proved to be affected by deltamethrin, gossypol

Fig. 3. Amino acid alignment of CYP6CV1 (CAZ65618.1), CYP6AE28 (CAX94849.1), and CYP6AE30 (CBB07053.1) with four CYP6 members: *Papilio xuthus* CYP6A2 (BAM18141.1), *H. armigera* CYP6B2 (ABE60885.1), *Depressaria pastinacella* CYP6AE1 (AAP83689.1), *H. armigera* CYP6AE12 (ABB69054.1). The proposed substrate recognition sites, denoted as SRS1–SRS6, are covered with lines (Gotoh, 1992). The helix I, helix K, the heme-binding motif, and the six residue sequences specific to CYP6 members are indicated and denoted with the corresponding names in italics. Identical amino acids are indicated by asterisks and conservative substitutions by dots.

and phenobarbital (Zhou et al. 2010). However, for absence of heterologeously expression and biochemical characteristics of the CYP9 members, little is known of the substrates of these P450s.

In this work, two novel P450 genes were cloned from the rice pest *C. medinalis.* Both cDNA-predicted proteins contain the conserved structural and functional regions characteristic of all insect P450s (Figs. 1 and 2). Both genes were inducible when larvae were exposed to resistant rice TKM-6 (Fig. 6). *CYP6CV1* was mainly expressed in the midgut and fat body, whereas *CYP9A38* was mainly expressed in the midgut, but with a slightly lower level in the fat body (Fig. 5B). These results coincided with the previous studies showing activation of P450 enzymatic activity in different insects feeding either on host plants or artificial diets containing plant allelochemicals (Li et al. 2007).

In another developmental study of the two P450 gene expressions, *CYP6CV1* showed a significant increase in mRNA levels in the fourth, fifth instar larvae, and adults. However, *CYP9A38* showed a steady increase in mRNA levels during the five instar stages and the adult stage (Fig. 5A). The earlier studies revealed that the patterns of expression of individual P450s can vary between life stages (Rewitz et al.

2007). In general, total P450 levels are undetectable in eggs, up and down in each instar larvae, are very low in pupae and are expressed at high levels in adults. It is nothing surprising that *CYP6CV1* and *CYP9A38* were expressed in such a different pattern, except both genes presented a low-expression pattern in the pupal stage.

CYP6CV1 and CYP6B8 are similar to each other with the highest amino acid identity of 41%. Because CYP6B8 is expressed in *H. zea* in response to plant phytochemicals and plant signaling molecules metabolize these compounds with varying efficiencies (Rupasinghe et al. 2007), we can suppose reasonably that these proteins may catalyze similar substrates and can be induced by them. Simultaneously, CYP9A38 and CYP9A17 are highly homologous to each other with the highest amino acid identity of 66%. The previous study revealed that *CYP9A17* mRNA is affected in dose-dependent and tissue-specific manners by deltamethrin, gossypol, and phenobarbital (Zhou et al. 2010), it is not surprising that *CYP9A38* should respond to similar phytochemicals from rice and be activated by them. However, the host resistance induction of *CYP6CV1* and *CYP9A38* does not exclude a possible physiological function on endogenous substrates, such as hormone metabolism or mating behavior.

Fig. 4. Phylogenetic tree of full-length amino acid sequences of deduced CYP6CV1 and CYP9A38 and some members from CYP6 and nine families constructed by the neighbour-joining method. The tree was constructed with the full-length sequences of the P450s. The bootstrap values were indicated on each branch. The GenBank accession numbers of the P450s in a reduced version of the neighbour-joining tree are: *Plutella xylostella* CYP6CV2 ADW27429.1, *P. xuthus* CYP6A2 BAM18141.1, *Manduca sexta* CYP6B46 ADE05579.1, *Spodoptera litura* CYP6B29 ACY41036.1, *H. zea* CYP6B8 AAM90316.1, *P. glaucus* CYP6B4 AAB05892.1, *H. armigera* CYP6B6 AAY21920.1, *H. armigera* CYP6B2ABE60885.1, *C. medinalis* CYP6AE28 CAX94849.1, *C. medinalis* CYP6AE30 CBB07053.1, *M. sexta* CYP6AE32 ADE05581.1, *H. armigera* CYP6AE12 ABB69054.1, *D. pastinacella* CYP6AE1 AAP83689.1, *Spodoptera littoralis* CYP6AE48 AFP20589.1 *B. mori* CYP6AE9 NP_001104004.1, *S. littoralis* CYP6AE47 AFP20588.1, *B. mori* CYP6AE7 NP_001104006.1, *Bombyx mandarina* CYP6AE8 ABY40426.1, *B. mori* CYP9A20 NP_001077079.1, *B. mori* CYP9A19 ABQ18318.1, *H. armigera* CYP9A12 ACB30273.2, *H. armigera* CYP9A17 AAY21809.1, *B. mori* CYP9A21 NP_001103394.1, *H. armigera* CYP9A14 ABY47596.1.
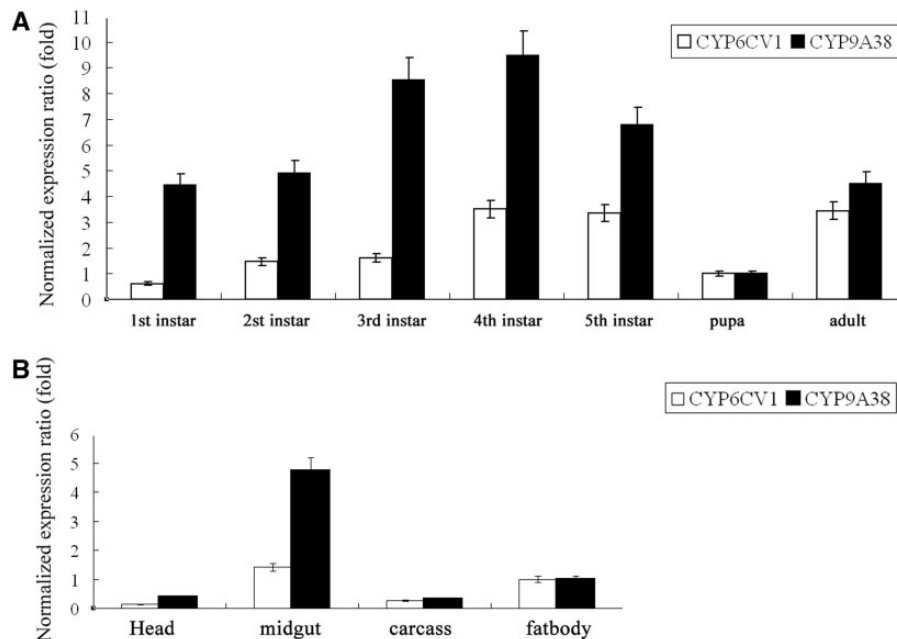
Fig. 5. Expression levels of *CYP6CV1* and *CYP9A38* mRNA in different developmental stages of *C. medinalis* (A) and in different tissues of fifth instar larvae of *C. medinalis* (B). A. Expression levels of *CYP6CV1* and *CYP9A38* in *C. medinalis* larvae during larval–larval, larval–pupal, and pupal–adult transitions were detected by qRT-PCR and normalized against *actin* transcript, and then normalized to obtain an expression ratio of 1 in pupas. B. Detection of *CYP6CV1* and *CYP9A38* expression in head, midgut, carcass, and fat body of *C. medinalis*. Equal tissue equivalents were analyzed by qRT-PCR and normalized against *actin* transcript, and then normalized to obtain an expression ratio of 1 in fat bodies. mRNA profiles are representative of three independent replicates. Vertical bars indicate the SEM ($n = 3$).

In our previous work, two P450 genes *CYP6AE28* and *CYP6AE30*, cloned from *C. medinalis* for the first time, were found to be induced by TKM-6 (Liu et al. 2010). The results together with the data in this study contribute to characterize the response of *C. medinalis* to rice dietary phytochemicals and emphasize the role of P450 genes in the adaptation of *C. medinalis* larvae to resistant rice. More works including heterologous expression of the new P450 genes, reconstruction of heterologously expressed enzyme systems, studying the biochemical
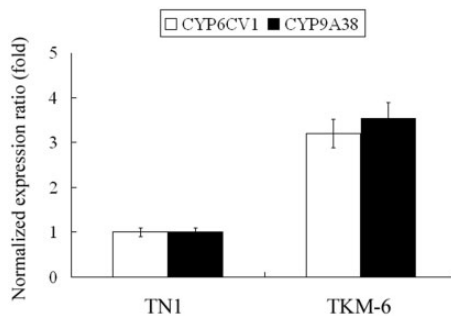
Fig. 6. Expression levels of *CYP6CV1* and *CYP9A38* mRNA in *C. medinalis* larvae feeding on different rice varieties. mRNA profiles are representative of three independent replicates. Vertical bars indicate the SEM ($n = 3$). TN1 and TKM-6 indicate *C. medinalis* larvae feeding on the rice plants of TN1 and TKM-6, respectively.

characteristics of the P450s and identification of their natural substances from rice, need to be carried out in the future to reveal the function of these P450 genes.

Acknowledgments

References Cited

Chandramani, P., R. Rajendran1, P. Sivasubramanian, and C. Muthiah. 2009. Impact of biophysical factors as influenced by organic sources of nutrients on major pests of rice. J. Biopestic. 2: 1–5.

Dalrymple, D. G. 1978. Development and Spread of High-Yielding Rice Varieties in Developing Countries, pp. 25–26. U.S. Department of Agriculture, Office of International Cooperation and Development & U.S. Agency for International Development, Washington, DC.

Dalrymple, D. G. 1986. Development and Spread of High-Yielding Rice Varieties in Developing Countries, p. 17. U.S. Department of Agriculture, Office of International Cooperation and Development & U.S. Agency for International Development, Washington, DC.

Feyereisen, R. 2005. Insect cytochrome P450. *In* L.I. Gilbert, K. Iatrou and S.S. Gill (eds.), Comprehensive Molecular Insect Science, vol. 4, pp. 1–77. Elsevier Press, Oxford, United Kingdom.

Gonzalez, F. J., and D. W. Nebert. 1990. Evolution of the P450 gene superfamily: animal-plant "warfare", molecular drive, and human genetic differences in drug oxidation. Trends Genet. 6: 182–186.

Gotoh, O. 1992. Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. J. Biol. Chem. 267: 83–90.

Kasai, S., I. S. Weerashinghe, T. Shono, and M. Yamakawa. 2000. Molecular cloning, nucleotide sequence and gene expression of a cytochrome P450 (CYP6F1) from the pyrethroid-resistant mosquito, *Culex quinquefasciatus* Say. Insect Biochem. Mol. Biol. 30: 163–171.

Khan, Z. R., and R. C. Joshi. 1990. Varietal resistance to *Cnaphalocrocis medinalis* (Guenee) in rice. Crop Prot. 9: 243–251.

Khan, Z. R., B. P. Rueda, and P. Caballero. 1989. Behavioral and physiological responses of rice leafolder *Cnaphalocrocis medinalis* to selected wild rices. Entomol. Exp. Appl. 52: 7–13.

Li, X. C., M. R. Berenbaum, and M. A. Schuler. 2002. Plant allelochemicals differentially regulate *Helicoverpa zea* cytochrome P450 genes. Insect Mol. Biol. 11: 343–351.

Li, X. C., M. A. Schuler, and M. R. Berenbaum. 2007. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. Ann. Rev. Entomol. 52: 231–253.

Liu, X. L., J. Chen, and Z. F. Yang. 2010. Characterization and induction of two cytochrome P450 genes, *CYP6AE28* and *CYP6AE30*, in *Cnaphalocrocis medinalis*: possible involvement in metabolism of rice allelochemicals. Zeitschrift für Naturforschung Section 65C: 719–725.

Maïbèche-Coisne, M., C. Merlin, M. François, P. Porcheron, and E. Jacquin-Joly. 2005. P450 and P450 reductase cDNAs from the moth *Mamestra brassicae*: cloning and expression patterns in male antennae. Gene 346: 195–203.

Masoud, S. A., X. Ding, L. B. Johnson, F. F. White, and G. R. Reeck. 1996. Expression of corn bifunctional inhibitor of serine proteinases and insect α-amylase in transgenic tobacco plants. Plant Sci. 115: 59–69.

Nelson, D. R. 1998. Metazoan cytochrome P450 evolution. Comp. Biochem. Physiol. Part C 121: 15–22.

Nelson, D. R., T. Kamataki, D. J. Waxman, F. P. Guengerich, R. W. Estabrook, R. Feyereisen, F. J. Gonzalez, M. J. Coon, I. C. Gunsalus, O. Gotoh, et al. 1993. The P450 superfamily: update on new sequences, gene mapping, accession numbers, early trivial names of enzymes, and nomenclature. DNA Cell Biol. 12: 1–51.

Page, R. M. W. 1996. TREEVIEW: an application to display phylogenetic tree on personal computers. Comput. Appl. Biosci. 12: 357–358.

Punithavalli, M., N. M. Muthukrishnan, and M. Balaji Rajkuma. 2013. Defensive responses of rice genotypes for resistance against rice leafolder *Cnaphalocrocis medinalis*. Rice Sci. 20: 363–370.

Rao, Y. C., G. J. Dong, D. L. Zeng, J. Hu, L. J. Zeng, Z. Y. Gao, G. H. Zhang, L. B. Guo, and Q. Qian. 2010. Genetic analysis of leafolder resistance in rice. J. Genet. Genomics 37: 325–331.

Rewitz, K. F., M. B. O'Connor, and L. I. Gilbert. 2007. Molecular evolution of the insect Halloween family of cytochrome P450s: phylogeny, gene organization and functional conservation. Insect Biochem. Mol. Biol. 37: 741–753.

Riley, J. R., D. R. Reynolds, A. D. Smith, A. S. Edwards, X. X. Zhang, X. N. Cheng, H. K. Wang, J. Y. Cheng, and B. P. Zhai. 1995. Observations of the autumn migration of the rice leaf roller *Cnaphalocrocis medinalis* (Lepidoptera: Pyralidae) and other moths in eastern China. Bull. Entomol. Res. 85: 397–414.

Rose, R. L., D. Goh, D. M. Thompson, K. D. Verma, D. G. Heckel, L. J. Gahan, R. M. Roe, and E. Hodgson. 1997. Cytochrome P450 *(CYP)9A1*: the first member of a new CYP family. Insect Biochem. Mol. Biol. 27: 605–615.

Rupasinghe, S. G., Z. Wen, T. L. Chiu, and M. A. Schuler. 2007. *Helicoverpa zea* CYP6B8 and CYP321A1: different molecular solutions to the problem of metabolizing plant toxins and insecticides. Protein Eng. Des. Sel. 20: 615–624.

Schuler, M. A. 2011. P450s in plant-insect interactions. Biochimica et Biophysica Acta 1814: 36–45.

Scott, J. G. 2008. Insect cytochrome P450s: thinking beyond detoxification. *In* N. Liu (ed.), Recent Advances in Insect Physiology, Toxicology and Molecular Biology, pp. 117–124. Research Signpost, Kerala, India.

Senthil, N. S., K. Kalaivani, K. Sehoon, and K. Murugan. 2006. The toxicity and behavioural effects of neem limonoids on *Cnaphalocrocis medinalis* (Güenée), the rice leafolder. Chemosphere 62: 1381–1387.

Sogawa, K., G. J. Liu, and J. H. Shen. 2003. A review on the hyper-susceptibility of Chinese hybrid rice to insect pests. Chin. J. Rice Sci. 17: 23–30.

Stevens, J. L., M. J. Snyder, J. F. Koener, and R. Feyereisen. 2000. Inducible P450s of the CYP9 family from larval *Manduca sexta* midgut. Insect Biochem. Mol. Biol. 30: 559–568.

Thompson, J. D., T. J. Gibson, F. Plewniak, F. Jeanmougin, and D. G. Higgins. 1997. The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24: 4876–4882.

Upadhyay, P., V. K. Singh, and C. N. Neeraja. 2011. Identification of genotype specific alleles and molecular diversity assessment of popular rice (*Oryza sativa* L.) varieties of India. Int. J. Plant Breeding Genet. 5: 130–140.

Zhou, X. J., C. X. Ma, M. Li, C. F. Sheng, H. X. Liu, and X. H. Qiu. 2010. *CYP9A12* and *CYP9A17* in the cotton bollworm, *H. armigera*: sequence similarity, expression profile and xenobiotic response. Pest Manag. Sci. 66: 65–73.