# Molecular Cloning of Cytochrome P450 4 Family Gene Fragments from Midgut of the Beet Armyworm, Spodoptera exigua

Pyeongjae Lee, Iksoo Kim<sup>1</sup>, Heui Sam Lee<sup>1</sup>, Il Je Cho and Jae Yu Moon\*

School of Biological Resources and Material Engineering, College of Agriculture and Life science, Seoul National University, Suwon 441-744, Korea.

<sup>1</sup>Department of Sericulture and Entomology, The National Institute of Agricultural Science & Technology, Rural Development Administration, Suwon 441-100, Korea.

(Received 15 April 2002; Accepted 10 May 2002)

Cytochrome P450 (CYP) gene has been known to play one of the most important roles in metabolizing the exogenous materials. In insect, CYP is particularly known to detoxify toxic materials by adding oxygen molecule to the hydrophobic region of the materials. Thus, CYP-dependent metabolism is associated with the adaptation of insect to host plant chemicals. This in turn is known to be one of the driving forces for CYP diversification. In the present study, we cloned seven gene fragments of CYP 4 (CYP4) family from the midgut of the beet armyworm, Spodoptera exigua, through RT-PCT. Sequence analysis of the product showed the gene fragment to contain an open reading frame of ~ 150 amino acids, consisted of ~450 bp. The cloned gene fragments contained typical, conserved regions found in CYP4 family. Pairwise comparison of the deduced amino acid sequences among seven clones ranged in divergence from 0% to 52.86% and resulted in five distinct clones. The other two clones were identical or differ by one amino acid respectively to the corresponding clone, although each differed by ten nucleotides. Analysis of correlation between GenBank-registered, full length CYP4 and the cloned fragments resulted in statistically significant relationship ( $r^2 = 0.96085$ ; p < 0.001), suggesting utility of the partial sequences as such full-length sequences. Phylogenetic analysis of the clones with GenBank-registered insect and

mammal CYP4 family sequences by parsimony and several distance methods subdivided the clones into two groups: ones belonging to CYP4S and the others to CYP4M families.

**Key words**: *Spodoptera exigua*, Armyworm, Cytochrome P450, *CYP4* family, RT-PCR, Phylogeny

### Introduction

Living organisms are always exposed to external materials that are toxic to the host as well as essential for energy production. Living organisms, thus, have evolved to develop a system of detoxifying toxic materials. Insect is one good example that had successfully adjusted to the environmental change by developing detoxifying system against host plant (Pichersky and Gang, 2000). Once plants invent a second metabolite, which kills insects as a defense system, insects again find a key system against this metabolite. This is the mechanism by which insects and plants have co-evolved. One of such key found by insects is a system, which transforms hydrophobic second metabolite of plant into hydrophilic compound (Schuler, 1996; Scott *et al.*, 1998).

In general, hydrophobic material is absorbed into the tissue faster and remains longer in the tissue than hydrophilic material. Thus, detoxifying procedure means a process that changes the hydrophobic material into the hydrophilic material. Detoxifying procedure is divided into Phase I and Phase II. One of the enzymes representing Phase I is Mixed-Functional Oxidase (MFO). MFO includes all the enzymes that make the substrate hydrophilic by adding one molecule of oxygen to substrate (e.g., cytochrome P450 and cytochrome P450

School of Biological Resources and Material Engineering, College of Agriculture and Life science, Seoul National University, Suwon 441-744, Korea.

Tel: +82-31-290-2483; Fax: +82-31-291-8697;

E-mail: mjy@snu.ac.kr

<sup>\*</sup>To whom correspondence should be addressed.

dependent reductase). Cytochrome P450 (CYP) exits as families and insect has families CYP4, CYP 6, CYP 9, CYP12, and CYP28 (Scott et al., 1998). Although families 6, 9, 12, and 28 are specific to insect, family 4 is found commonly in vertebrates, plants and microorganisms. This wide distribution of CYP4 family allowed ones to believe that all the CYP families were evolved from common origin. At present, many of CYP genes have been identified in a diverse insect species, but their detailed function is still in question. CYP6D1 in housefly (Lee and Scott, 1992; Liu and Scott, 1996; Korytko and Scott, 1998) and CYP6B family in the Papilio species are known to have a function related to resistance to pyrethroid insecticide and to metabolism of furanocoumarin in host plant, respectively (Cohen et al., 1992; Hung et al., 1995). In case of CYP4 family, it has been known that gene expression in Drosophila is enhanced by treatment of alkaloid, suggesting that CYP4 family is also related to metabolism of exogenous material (Danielson et al., 1998).

The beet armyworm, *Spodoptera exigua*, is a pest found in a welsh onion and has more than one hundred host plants. Thus, we think that *CYP4* family may be important genes in adaptation of *S. exigua* to various host plants. In this study, *CYP4* family gene fragments were cloned from the midgut of the species and relations of the genes to other *CYP4* subfamilies were discussed.

### **Materials and Methods**

The larvae of the beet armyworm, *Spodoptera exigua*, were obtained from the Department of Crop Protection, The National Institute of Agricultural Science & Technology, Rural Development Administration, Korea.

#### RT-PCR, cloning and sequencing

Midgut of *S. exigua* was extracted and washed in 1 × SSC. It was frozen in the liquid nitrogen and ground into fine powder. Total RNA was isolated from 30 mg of midgut power by using Total RNA Extraction Kit (Promega). Total RNA was identified by electrophoresis and quantified by measuring UV absorption at 260 nm. Degenerative primer was designed on the basis of the conserved region of CYP4 family protein (Table 1). The primer sequences are as follows: *CYP4* up-primer, 5-GA(C/T) ACITT(C/T)ATGTT(C/T)GA(A/G)GGIGA(C/T)AC-3 and *CYP4* down-primer, 5-GC(A/G)AA(T/C)TGICC(A/G/T) TTIC(T/G)IGGICC-3, where I indicates inosine.

The first strand of cDNA was synthesized by PCR at the condition of 45°C for 45 min and 94°C for 2 min, followed by 40 cycles of 94°C for 30 sec, 45°C for 1 min, and 68°C for 2 min. The final extension was performed at 68°C for 9 min. PCR products were identified on 0.7% agarose gel. The expected product bands were excised from the gel and dialyzed for elusion. The PCR products for sequencing were cloned into pGEM-T vector (Promega). Plasmids were extracted by using Wizard plus SV Miniprep Kit (Promega) and successful cloning was detected both by 0.7% agarose gel and enzyme digestion (EcoR I). The plasmids were sequenced using an automatic sequencer (model 377, PE Applied Biosystems).

#### Data analysis

Nucleotide sequences of *S. exigua* clones were compared to protein sequences incorporated into GenBank by using BLASTX program on the NCBI Blast Server. Sequence identity among full-length *CYP4* and among partial *CYP4* were performed in the Clustal W (ver. 1.8; Jeanmougin *et al.*, 1998). PAUP (Phylogenetic Analysis using Parsimony) ver. 3.1 (Swofford, 1993) was used to infer possible phy-

**Table 1.** Conserved regions of CYP4 gene of insect species utilized for primer design

Species	Upstream	Family type	GenBank Accession number	References
Tropical cockroach (Blaberus discoidals)	N-DTFMFEGHDT-C C-GPRNCIGQKFA-N	CYP4C1	P29981	Bradfield et al. (1991)
Pacific beetle cockroach (Diploptera punctata)	N- <b>DTFMFAGHDT</b> -C C- <b>GPRNCIGQKFA</b> -N	CYP4C7	AAC69184	Sutherland et al. (1998)
Fruit fly (D. melanogaster)	N- <b>DTFMFKGHDT</b> -C C- <b>GPRNCIGQKFA</b> -N	CYP4D2	AAB71182	Unpublished
Fruit fly (D. melanogaster)	N- <b>DTFMFEGHDT</b> -C C- <b>GPRNCIGQKFA</b> -N	CYP4D2	Q27589	Frolov and Alatortsev (1994)
Tabacco hornworm (Manduca sexta)	N- <b>DTFMFEGHDT</b> -C C- <b>GPRNCIGQKFA</b> -N	CYP4M2	S66374	Snyder et al. (1995)

Bold-faced letters indicate identical amino acid residues among insect species compared.

Sequences at the upper line at each column represent conserved sequences in the upstream region and those at lower line represent conserved sequences in the down stream region.

logenetic relationships among CYP4 family sequences cloned in this study and the GeneBank-registered CYP4 family sequences of some insects and mammals. The

homologous CYP4 family gene of house mouse, Mus musculus (GenBank accession number Q64462; Imaoka et al., 1995), was used as an outgroup. The reliability of

clone 1 clone 2 clone 3 clone 4 clone 5 clone 6	CATACOTTCA CACCOTTA MITACOTTCA MITACOTTCA MITACOTTCA CALCOTTCA	TO TEAGGE TO TEAT C G G TO TEAG C A G T G G	CONTACC CONTACC TO TGCT CONTACT CONTACT CONTACT CONTACT	ACGG GAGTG ACGG GAGTG ACAG AATGG ATG AACA ATG ATCAG ACAA CTTAG	GGT GACGTT GGT GACTTT CTC CGTTTT TGC CTGTTT GAC TACTTT CTC TACTTT	TTGCTTCATO TTGCTTTATO TGGACTTATC CATGATCATO CTTATTATC TGGCTTGATG
clone 7	GACAL GENTA	CENGES	TACT	AGAATG	CACCCTT	TGGACTTATG
						120
clone 1	CTGTTACCAC	ATCATAAAGA	TGCA <b>CAGGA</b> C	AAAA#CGTAG	ATGA ATTGA	AGAAGTCTTG
clone 2	CTGTTALAC	ATC TAAA	TGCACACAC	AAAATCGTAG	ATGA TTGAA	AGAAGTCT <b>T</b> G
clone 3 clone 4	TTGCTAGCAG AGACTG CGA	ACCATCCA. ACGAACCTCA	AGTA <b>CAGTA</b> A AATA <b>CAGTA</b> T	CGGANTTACG TCGGTAGCTG	AAGA CTAAA	GACGARCCEG AAGTATATTT
clone 5	TTGCTACCA	ATCATCCA	AATTCACCAC	AAAGTTGTAA	GTCACCTCA	TGACATATEC
clone 6	CTGT <b>T</b> G <b>CC</b> TG	ATCATGAAGA	TGTACAL AG	CGCA#CTACG	AGGA TGCC	AACGATCCTG
clone 7	TTGCTACCAG	ACCETCCAGA	agta <b>cacea</b> a	CGGATTTACG	AAGAATGTCA	GACGATCCTG
						180
clone 1	GCAGATTTTA	AACGTCCAAT	cacaa <b>t</b> tgaa	GATCTACCGA	aga <b>tgaagta</b>	CT <b>T</b> GGAGCGT
clone 2	GCAMATTTTA	ACGTCCAAT	CACAATTGAA	GATCTACCGA	AGATEARGTA	CTTGGAGCGT
clone 3 clone 4	GGTGATTCAG GGCGACAGTC	ACACATCTCC AACGACCTCC	AACGA®GTCT TACCCTTGAA	GATCTGGCTG GACTTGAGTC	aaatgaagta Agatgaagta	TC#GGAGGCA CC#GGACTGT
clone 5	CALACTOGO	AACGTTGGGC	ATGCATGAAC	GACTTGCCAA	AGATCAACTA	CTTAGACCGG
clone 6	<b>GG</b> T <b>CA</b> CTCAG	AACACGT	GACGATGTCC	A TCTGTCCG	Tarres Agris Agris Agris	TT <b>TGGA</b> GGCA
clone 7	GGTGATTCAG	ACACATOTOO	AACGATGTCT	GATCTGGCTG	ANTONOUS	TC <b>EGGA</b> GGCA
						240
clone 1	cg <b>tgttaagg</b>	<b>A</b> ATCGTTGC <b>G</b>	GTTG <b>TA</b> C <b>CC</b> T	cca <b>gtgc</b> at <b>t</b>	TTATCAGTAG	GAGCTTGCAC
clone 2	TGTGTTAAGG	AATCGTTGCG	GTTGTACCCT	CCAGTGCATT	TTATCAGTAG	GAGCTTGCAC
clone 3 clone 4	GTTATCAAAG TGTATAAAAGG	AAATCTTAAG AGTCGCTCCG	ACTGTACCCG TCTCTACCCC	agt <b>gtgc</b> cg <b>t</b> agc <b>gtgc</b> ac <b>t</b>	TCATTGCCCG TTATGTCCAG	GGAGGTGACT ATGTTTTACT
clone 5	TGTALAMAG	GTCACTGC	TATGUACCCG	CCGCTCCACT	TCATTAGTCG	GAAACTGAAC
clone 6	gt <b>tatcaa</b> g <b>g</b>	AATACTGCG	ATTG CCCC	AGC <b>OTS</b> CCT.	CATTGGACE	AGAGATCACT
clone 7	GTTATCAAAG	AATCTTAAG	actg <b>tatec</b> g	agt <b>gtgc</b> cg <b>t</b>	TTATTGCCCG	GGAGG <b>T</b> GACT
						300
clone 1	<b>GA</b> AGATGTGA	TATTGACTGA	CTACTTAG <b>T</b> A	CCAGCTGGAA	CATTCTGCCA	TATACATATC
clone 2 clone 3	GAAGATGTGA GAGGACTTTA	Tategaetga Tgttageaga	CTACTTAGTA TGTCCTGGTA	CCAGCTGGA AAGAAAGGT	ATTCTGCCA CAGAAGTATC	TATACATATC AATCCACATA
clone 4	GAGGATGTGA	AGTTAGCCGA	TGTCACAGTA	CCATACGACA	CAATGTGCCA	TTTCAACGTG
clone 5	<b>ga</b> c <b>ga</b> aacag	TAT <b>T</b> AA <b>G</b> TA <b>A</b>	TCACACAATA	CCTGCCAGTA	CATTGTGTCA	cataccaata
clone 6 clone 7	GAAGATTTTA GAAGACTTTA	agc <b>t</b> gg <b>g</b> tg <b>a</b> tgt <b>t</b> gg <b>g</b> ag <b>a</b>	CATAACAGTG TGTCCTGGTA	AAGAAGGGT <b>A</b> AAGAAAGGT <b>A</b>	CGACAGTGGA CAGAAGTATC	CGTCCACATC AATCCACATA
Crone /	GAAGACIIIA	IGI BGGGAGA	IGICCIGGIA	AHGAAAGG1 A	CAGAAGIAIC	ANICCACAIA
		122A		23 Mars - 1886 - 1884 - 18		360
clone 1 clone 2	TATGACTTGC TATGACTTGC	ATCGCCAACC ATCGCCAACC	GCATTTATTC GCATTTATTC	CCCAACCCTA CCCAACCCTA	ACAAGTTCGA ACAAGTTCGA	CCCTGACAGA CCCTGACAGA
clone 3	ACCACC	ATCGTCTGCC	AMATTGTAT	CCAGECCCGG	AGGCATTCAA	GCCGGAACGG
clone 4	TTCGACATAC	<b>ATC</b> GCAATCC	G <b>GA</b> TATATTC	CCGGACCCGG	<b>A</b> GAAG <b>TT</b> CAT	ACCGGAACGA
clone 5	TACGACCTGC	ATCATCGAGA	GCATTTGTTC	CCGAACCCTG	AAGTATTTGA	CCCTGACCGG
clone 6 clone 7	TACGAGC TGC TACGACC TGC	ACCGTCGCGC ATCGTCTGCC	TCACATGTTC AGAATTGTAT	CCGGAGGCGG CCAGACCCGG	AGAAGTTCAT AGGCTTTCAA	GCCGGAGCGG GCCGGAACGG
	37 · · · 200 · · · · · · · · · · · · · · ·		-	3000		
clone 1	TTCCTGCCTG	AGAACAGTGT	TGGAAGACAT	CCCTATGCTT	ACATCCCCTT	420 TAGTGCGGGC
clone 2	TTCOTGCCTG	AGAACAGTGT	TGGAAGACAT	CCCTATGCTT	ACARCCCCTT	TAGTGCGGGC
clone 3	TTCCTCAA	TCAACAACC-	GACACAT	CCCTACCCCT	TTGTGCCTTT	TAGTGCTGGC
clone 4	TTTGTGCCCG	AAAACTGTGT	TTCGCGT <b>CAT</b>	CCATATGCTT	aca <b>taccatt</b>	TAGTGTTGGC
clone 5	TTCCTACCAG	AAAATTGTGA	aggaaga <b>cat</b>	CCTTATGCAT	aca <b>tacc</b> g <b>tt</b>	TAGTGCTGGC
clone 6	TTCCTCGG	CACGGAGAT-	GAAACAT	CCTTACCCTT	ACGTGCCGTT	CAGCGCTGGC
clone 7	TTCCTCAA	TCAACAGCC-	GACACAT	CECTACGCCT	TTG <b>T</b> G <b>CCTTT</b>	TAGTGCTGGC
			449			
clone 1	COCCATT	GCAZAGGCCA	GAAATTCGC			
clone 2	COCACCASTY	GTATTERCUA	AAAATICCC			
clone 3 clone 4	TT	GTATA SECA	GAAGTTTEC			
clone 5	CECCONTT CECANONICE	CT A APPECA CTACCOCCA	A A A STORE			
clone 6	CONCLUS TT	GCATAGGCCA	a <b>ra</b> attoc G <b>aa</b> attosc			
clone 7	CCCCCCANTT	GTATCEGCCA	GARATTEC			
	ACCORDING TO MAKE TO CONTINUE	o establish a september 2000	marries (Marries Marries)			

Fig. 1. Alignment of cDNA segment of seven CYP4 gene fragments from Spodoptera exigua. Identical sequences are boxed with gray color.

the trees was tested by 100 iterations of bootstrapping (Felsenstein, 1985). As an alternative to the parsimony analysis, we used subprograms such as NEIGHBOR, KITSCH, and FITCH incorporated in PHYLIP (Phylogeny Inference Package) ver. 3.5c (Felsenstein, 1993). To obtain phylogenetic tree, the data set was first iterated 100 times using the subprogram SEQBOOT. Next, the iterated data set was run using the subprogram PRODIST to obtain distance matrix between pairs of amino acid sequences. Individual trees from each distance matrix were obtained using the subprograms and a consensus tree representing reliability at each branch was obtained using the subprograms CONSENSE, respectively. Input method was chosen for random input in the NEIGHBOR and global rearrangement for FITCH, but jumbling option was omitted to save computer hour in all cases.

#### **Results and Discussion**

# Cloning, sequencing and characterization of the *CYP4* gene fragments

Purified CYP protein from a diverse organism showed the presence of a conserved region in the protein, called heme site (Scott et al., 1994; Snyder et al., 1996; Feyereisen, 1999). The site is consisted of four of its residues (F--G---C-G) and recognized as a signature motif of cytochrome P450. Along with this site, CYP4 family has one more conserved region. This region corresponds to a portion of helix I in the tertiary structure of cytochrome P450cam, which is distal to the heme and is thought to be important in the activation of molecular oxygen and catalysis (Scott et al., 1994; Snyder et al., 1996; Feyereisen, 1999). Sequence alignment of a few cytochrome P450 CYP4 subfamilies from the cockroach (Blaberus discoidals), the pacific cockroach (Diploptera punctata), the fruit fly (Drosophila melanogaster), and tabacco hormworm (Manduca sexta), in fact, revealed these two highly conserved stretches of amino acid (Table 1). In order to clone the genes of cytochrome P450 CYP4 family from the midgut of S. exigua, degenerative primer was designed on the basis of these conserved sequences (Table 1).

From the midgut of *S. exigua*, we obtained a PCR product with ~ 450 bp (base pair) by RT-PCR (data not shown) and this fragment was cloned into pGEM-T vector (Promega). Twenty colonies were selected, and sequenced, and eventually provided us seven gene fragments with the size of 440 bp, 443 bp, and 449 bp (Fig. 1). Among the seven clones, clones 1 and 2 were variable in ten nucleotide sites (data not shown), but differ by only one residue in terms of amino acid (Table 2). Clones 6 and 7 also showed ten nucleotide difference (data not shown),

**Table 2.** Pairwise comparison among partial amino acid sequences of the seven clones of *CYP4* family gene from *Spodoptera exigua* midgut

	1	2	3	4	5	6	7
Clone 1	-	0.67	43.24	36.00	28.67	47.14	44.90
Clone 2	1	-	43.24	35.33	29.33	47.14	47.14
Clone 3	65	65	-	44.59	47.97	28.57	28.57
Clone 4	54	53	67	-	36.67	47.86	47.86
Clone 5	43	44	72	55	-	52.86	52.86
Clone 6	71	71	43	72	79	-	0
Clone 7	71	71	43	72	79	0	-

Numbers above the diagonal are percent divergence; numbers below the diagonal are actual numbers of amino acid difference.

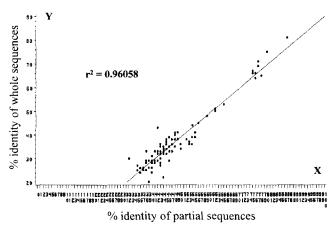
but identical in their deduced amino acid sequences (Table 2). Excluding these similar clones (clones 2 and 7), amino acid divergence among them ranged from 28.57% (43 amino acid residues) to 52.86% (79 amino acid residues). Considering substantially large amino acid difference among the five clones, the very similar clones, 2 and 7, appear to be allelic variants to clones 1 and 6, respectively. The length of the protein ranged from 147 to 150 amino acid residues including primer sequences (Fig. 2). All of the deduced amino acid sequences contained the regions found typically in all *CYP4* families: ΦΦΚΕ-LRΦ-P, where Φ is hydrophobic amino acid and a PERF motif (P-RF) (Fig. 2).

## Correlation between partial and total fragments of CYP4

From the beginning, because we aimed to clone only the fragments of CYP4 family using conserved sequences in the CYP4 family, utility of such PCR products may not necessary reveal accurate levels of sequence homology among them and among organisms as the full-length genes do. More fundamentally, designation of the partially coned sequences may not confidentially provide the basis for the assignment of the given sequences to the existing type of the gene. This may be especially true when the partial data are utilized for the inference of evolution and multigene family diversity. Thus, we obtained ~ 40 Gen-Bank-registered full-length amino acid sequences of CYP4 family and investigated correlation of the levels of partial amino acid sequence homology to the full-length sequence by statistical analysis (Fig. 3). The result showed a high correlation between total and fragment amino acid sequences of CYP4 family ( $r^2 = 0.96085$ ) and the estimate was statistically significant at the level of P < 0.001. Thus, this partial sequence of CYP4 family can be confidentially used to predict full-length sequence identity and this

(primer sequent N-DTFMFEGHD)								40
clone 1 clone 2 clone 3 (7) clone 4 clone 5 clone 6	ASGE EASGE EAMA EATVE EASGE ETTLAE	THCFH THCFH VHGLIN CHMAN THLFH THGLIN	LINH LINH LIND RINE LIND LIND	KDAOD KDAOD PEVOE PEAOD PEIOD EDVOE	KIVDE KIVDE RIYEE SVAES KVVSE RIYEE	LKEVL LKEVL CQTIL LKSIF LNDIF CQTIL	COCKR COCKR COSDT COSQR COSQR COSEH	PITTE PITTE SPTMS PPTLE WACMN VTMSN
clone 1 clone 2 clone 3 (7) clone 4 clone 5 clone 6	DLPKM DLPKM DLAEM DLSQM DLPKM LSD-M	ER EC EA DC DC DR EA	RVIES RVIES VIES CICS VIES VIES	ELY ELY LY ELY EMY EMY ELT LRO-P	P整H豐I P壁H豐I S蒙H豐I S蒙H豐I BY	SMSLH SMSLH AMEVT SMCFT SMKLN RMEIT	EDVIS EDVIS EDVKS DETVS EDFKS	80 SDYLV SDYLV GDVLL GOVTV SNHTI GDITV
clone 1 clone 2 clone 3 (7) clone 4 clone 5 clone 6	PAGTF PAGTF KKGTE PYDTM PASTL KKGTT	CHIHI CHIHI VSIHI CHFNV CHIPI VDVHI	YDL資R YDL發R YDL發R FDL發R YDL錄H YEL錄R	QPDLF QPDLF LPELY NPDIF REDLF RADMF	登い選択 登い選択 登い 登い 登し 登 を を を を を を を を を を を を を	FDPDR FDPDR FKER ELPER FDPDR FMPER P-R	PEN LPEN LQQ FIPEN FIPEN FIGTA F	120 SVGR報 SVGR報 PT報 CVSR認 CEGR稿 MK最
clone 1 clone 2 clone 3 (7) clone 4 clone 5 clone 6	6.79 AI 6.79 AI 6.79 AI 6.78 AI 6.78 AI 6.78 AI		PRNICGQK rimer seque					

**Fig. 2.** Alignment of deduced amino acid sequences for the CYP4 gene fragments from  $Spodoptera\ exigua$  isolated by RT-PCR. Identical sequences are boxed with gray color. The common "ΦΦΚΕ-LRΦ-P" and "P-RF" amino acids in the CYP4 gene are bold-faced, where Φ is hydrophobic amino acid.



**Fig. 3.** Interrelationship between the percent identities of whole CYP4 proteins and the percent identity of the fragment of the proteins.

resultantly can be further utilized for the inference of the phylogenetic relationships of the cloned *CYP4* family sequences.

# Relationships among clones and other insect CYP4 family

For the understanding of the phylogenetic relationships

among cytochrome P450 CYP4 family from the diverse organisms, GenBank-registered data were extensively searched, and those with enough length were selected and trimmed for phylogenetic analysis. This effort resulted in a total of 33 amino acid sequences. These were 26 CYP4 family gene sequences from insects, one from crayfish, and six from mammal species. Their GenBank accession number, scientific name, type of CYP4 family, and their references are listed in Table 3. Pairwise comparison among these CYP4 family genes including six clones from this study resulted in amino acid sequence divergence ranging from 0.33% (5 amino acid residues) to 62.9% (96 amino acid residues) (data no shown). The least divergence was observed in a comparison between two sequences of Drosophila melanogaster (GenBank number AAF47649 and AAB71182), belonging to the same family, CYP4D. The highest divergence was observed in a comparison between one of the D. melanogaster (Gen-Bank number AAF48426; CYP4S) and Mus musculus (GenBank number NP\_31848; CYP4A).

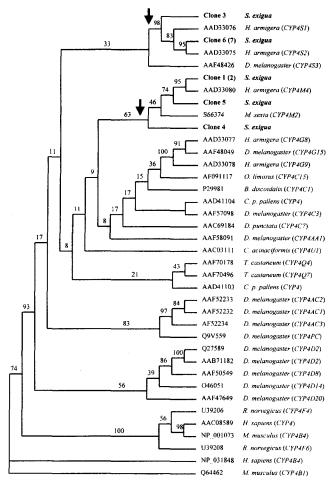
To infer phylogenetic relationships of the cloned *CYP4* family sequences to the previously reported gene family, 33-GenBank-registered *CYP4* family genes and five cloned *CYP4* family sequences (clones 1 and 2 and clones 6 and 7, respectively, were treated as each one clone because of

**Table 3.** List of CYP4 family gene sequences obtained from GenBank

	GenBank Accession Number	Common name	Scientific name	Family type	Reference
1.	AAD33076	Cotton bollworm	Helicoverpa armigera	CYP4S1	Unpublished
2.	AAD33075	Cotton bollworm	H. armigera	CYP4S2	Unpublished
3.	AAD33080	Cotton bollworm	H. armigera	CYP4M4	Unpublished
4.	AAD33077	Cotton bollworm	H. armigera	CYP4G8	Unpublished
5.	AAD33078	Cotton bollworm	H. armigera	CYP4G9	Unpublished
6.	S66374	Tabacco hornworm	Manduca sexta	CYP4M2	Snyder et al. (1995)
7.	AAC03111	Australian termite	Coptotermes acinaciformis	CYP4U1	Unpublished
8.	P29981	Tropical cockroach	Blaberus discoidalis	CYP4C1	Bradfield et al. (1991)
9.	AAF70178	Red flour beetle	Tribolium castaneum	CYP4Q4	Unpublished
10.	AAC69184	Pacific beetle cockroach	Diploptera punctata	CYP4C7	Sutherland et al. (1998)
11.	AAF70496	Red flour beetle	Tribolium castaneum	CYP4Q7	Unpublished
12.	AAD41103	Red house mosquito	Culex pipiens pallens	CYP4	Unpublished
13.	AAD41104	Red house mosquito	C. p. pallens	CYP4	Unpublished
14.	AAF48049	Fruit fly	Drosophila melanogaster	CYP4G15	Adams et al. (2000)
15.	AAF58091	Fruit fly	D. melanogaster	CYP4AA1	Adams et al. (2000)
16.	AAF57098	Fruit fly	D. melanogaster	CYP4C3	Adams et al. (2000)
17.	AAF48426	Fruit fly	D. melanogaster	CYP4S3	Adams et al. (2000)
18.	AAF52233	Fruit fly	D. melanogaster	CYP4AC2	Adams et al. (2000)
19.	AAF52232	Fruit fly	D. melanogaster	CYP4AC1	Adams et al. (2000)
20.	AF52234	Fruit fly	D. melanogaster	CYP4AC3	Adams et al. (2000)
21.	Q9V559	Fruit fly	D. melanogaster	CYP4P3	Adams et al. (2000)
22.	Q27589	Fruit fly	D. melanogaster	CYP4D2	Frolov and Alatortsev (1994)
23.	AAF50549	Fruit fly	D. melanogaster	CYP4D8	Adams et al. (2000)
24.	O46051	Fruit fly	D. melanogaster	CYP4D14	Adams et al. (2000)
25.	AAF47649	Fruit fly	D. melanogaster	CYP4D20	Adams et al. (2000)
26.	AAB71182	Fruit fly	D. simulans	CYP4D2	Unpublished
27.	AF091117	Spinycheek crayfish	Orconectes limosus	CYP4C15	Dauphin-Villemant et al. (1999)
28.	U39206	Norway rat	Rattus norvegicus	CYP4F4	Kawashima and Strobel (1995)
29.	U39208	Norway rat	R. norvegicus	CYP4F6	Kawashima and Strobel (1995)
30.	AAC08589	Human	Homo sapiens	CYP4	Unpublished
31.	NP_001073	Human	H. sapiens	CYP4B4	Simpson (1997)
32.	NP_031848	House mouse	Mus musculus	CYP4A14	Heng et al. (1997)
_33.	Q64462	House mouse	M. musculus	CYP4B1	Imaoka <i>et al.</i> (1995)

their sequence similarity or identity) were subjected to PAUP and PHYLIP analyses. Parsimony-based PAUP analysis (Fig. 4) showed that the five clones obtained in this study were subdivided into two groups: one including clones 1 (2), 4, 5 and the other including 3, 6 (7). The subdivision and clustering were somewhat robust in that each group was supported by relatively high bootstrap values (63 and 98, respectively). The first group included two types of *CYP4S* from cotton bollworm (GenBank numbers AAD33076 and AAD33075) and one type of fruit fly

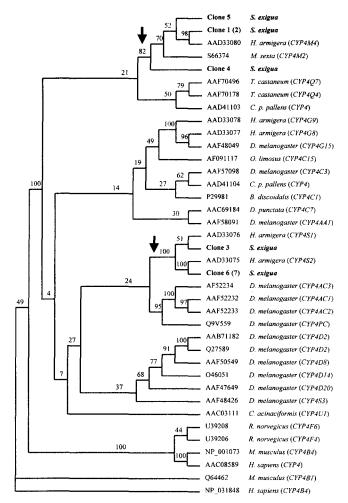
(AAF48426), and the second group included each one type of *CYP4M* from cotton bollworm (AAD33080) and tabacco hornworm (S66374) (Fig. 4). This result suggested us that the midgut of *S. exigua* clearly possesses at least two subfamilies of *S. exigua* CYP4 sequences by this analysis. PHYLIP analysis by neighbor-joining method also supported presence of two subfamilies among the clones obtained in this study (Fig. 5). However, *CYP4S* from fruit fly (AAF48426) was placed at a different group containing several other *CYP4* subfamilies (Fig. 5). This sepa-



**Fig. 4.** PAUP analysis of 38 *CYP4* family genes. The tree shown is an unordered tree obtained with the option of "retain groups with frequency <50%" by majority-rule consensus of one most parsimonious tree from the heuristic search. The numbers shown on the branches represent bootstrap values for 100 replicates. Tree length is 1,424 steps, Consistency Index is 0.515, and Retention Index is 0.535. The arrow indicates the place of clustering of the cloned *CYP4* gene with the gene from other species.

ration instead clustered the remaining members more robustly (100% of bootstrapping value). These phylogenetic analyses consistently suggest that the clones 1 (2), 4, 5 are belonged to *CYP4M* family and clones 3 and 6 (7) are belonged to *CYP4S* family. Although we also performed other distance-based analyses (*e.g.*, FITCH and KITCH), we did not present here because the topology obtained from these analyses were almost identical to the analyses provided here.

In conclusion, we report here molecular cloning of two subfamilies of *CYP4* from the midgut of the beet armyworm, *Spodoptera exigua* by RT-PCR. Seven fragments were successfully amplified using the degenerative prim-



**Fig. 5.** PHYLIP analysis of 38 *CYP4* family genes. The tree was obtained using the subprogram NEIGHBOR incorporated in PHYLIP. The tree was rooted using Q64462 (*CYP4B1*; *Mus musculus*). The number shown on branches, which represent bootstrap values for 100 replications, were obtained using the subprogram CONSENSUS. The arrow indicates the place of clustering of the cloned *CYP4* gene with the gene from other species.

ers designed from conserved regions of other homologous, insect gene sequences, and the presence of typical conserved region was also found. Phylogenetic analysis confirmed that clones 3 and 6 (7) are belonged to *CYP4S* family, whereas clones 1 (2), 4, and 5 are belonged to *CYP4M* family.

### References

Adams, M. D. *et al.* (2000) The genome sequence of *Droso-phila melanogaster*. *Science* 287, 2185-2195.
Bradfield, J. Y., Y. H. Lee and L. L. Keeley (1991) Cytochrome

- P450 family 4 in a cockroach: molecular cloning and regulation by regulation by hypertrehalosemic hormone. *Proc. Natl. Acad. Sci. USA* **88**, 4558-4562.
- Cohen, M. B., M. A. Schuler and M. R. Berenbaum (1992) A host-inducible cytochrome P450 from a host-specific caterpillar: molecular cloning and evolution. *Proc. Natl. Acad. Sci. USA* 89, 10920-10924.
- Danielson, P. B., J. L. M. Foster, M. M. Mamahill and M. K. Smith (1998) Induction by alkaloids and Phenobarbital of family 4 cytochrome P450s in *Drosophilla*: evidence for involvement in host plant utilization. *Mol. Gen. Genet.* 259, 54-59.
- Dauphin-Villemant, C., D. Bocking, M. Tom, M. Maibeche and R. Lafont (1999)
- Cloning of a novel cytochrome P450 (CYP4C15) differentially expressed in the steroidogenic glands of an arthropod. *Biochem. Biophys. Res. Commun.* **264**, 413-418.
- Felsenstein, J. (1993) PHYLIP (Phylogenetic Inference Package). Department of Genetics, Univ. of Washington, Seattle.
- Feyereisen, R. (1999) Insect P450 enzymes. *Ann. Rev. Entmol.* **44**, 507-533.
- Frolov, M. V. and V. E. Alatortsev (1994) Cluster of cytochrome P450 genes on the X chromosome of *Drosophila melanogaster*. *DNA Cell Biol.* **13**, 663-668.
- Heng, Y. M., C. S. Kuo, P. S. Jones, R. Savory, R. M. Schulz, S. R. Tomlinson, T. J. Gray and D. R. Bell (1997) A novel murine P-450 gene, Cyp4a14, is part of a cluster of Cyp4a and Cyp4b, but not of CYP4F, genes in mouse and humans. *Biochem. J.* 325, 741-749.
- Hung, C. F., M. R. Berenbaum and M. A. Schuler (1995) CYP6B3: a second furanocoumarin-inducible cytochrome P450 expressed in *Papilio polyxenes*. *Insect Mol. Biol.* 4, 149-160.
- Imaoka, S., T. Hiroi, Y. Tamura, H. Yamazaki, T. Shimada, M. Komori, M. Degawa and Y. Funae (1995) Mutagenic activation of 3-methoxy-4-aminoazobenzene by mouse renal cytochrome P450 CYP4B1: cloning and characterization of mous CYP4B1. *Arch. Biochem. Biophys.* 321, 255-262.
- Jeanmougin, F., J. D.Thompson, M. Gouy, D. G. Higgins and T. J. Gibson (1998) Multiple sequence alignment with Clustal X. *Trends Biochem. Sci.* **23**, 403-405.
- Kawashima, H. and H. W. Strobel (1995) cDNA cloning of three new forms of rat brain cytochrome P450 belonging to

- the CYP4F subfamily. *Biochem. Biophys. Res. Commun.* **217**, 1137-1144.
- Korytko, P. J. and J. G. Scott (1998) CYP6D1 protects thoracic ganglia of house flies from the neurotoxic insecticide cypermethrin. Arch. Insect Biochem. Physiol. 37, 53-57.
- Lee, S. S. T. and J. G. Scott (1992) Tissue distribution of microsomal cytochrome P-450 monooxygenases and their inducibility by phenobarbital in house fly *Musca domestica* L. *Insect Biochem. Mol. Biol.* 22, 699-711.
- Liu, N. and J. G. Scott (1996) Genetic analysis of factors controlling elevated cytochrome P450, CYP6D1, cytochrome b, P450 reductase and monooxygenase activities in LPR house flies, *Musca domestica*. *Biochem. Genet.* **34**, 133-148.
- Pichersky, E. and R. Gang (2000) Genetics and biochemistry of secondary metabolites in plants: an evolutionary perspective. *Trends Plants Sci.* **5**, 439-445.
- Schuler, M. A. (1996) The role of cytochrome P450 monooxygenases in plant-insect interactions. *Plant Physiol.* **112**, 1411-1419.
- Scott, J. A., F. H. Collins and R. Feyereisen (1994) Diversity of cytochrome P450 genes in the mosquito, *Anopheles albi*manus. Biochem. Biophys. Res. Commun. 205, 1452-1459.
- Scott, J. G., N. Liu and Z. Wen (1998) Insect cytochromes P450: diversity, insecticide resistance and tolerance to plant toxins. *Comp. Biochem. Physiol.* 121, 147-155.
- Simpson, A. E. (1997) The cytochrome P450 4 (CYP4) family. *Gen. Pharmacol.* **28**, 351-359.
- Snyder, M. J., J. L. Stevens, J. F. Andersen, R. Feyereisen (1995) Expression of cytochrome P450 genes of the CYP4 family in midgut and fat body of the tobacco hornworm, *Manduca sexta*. *Arch. Biochem. Biophys.* **321**, 13-20.
- Snyder, M. J., J. A. Scott, J. F. Anderson and R. Feyereisen (1996) Sampling of P450 diversity by cloning polymerase chain reaction products obtained with degenerate primers. *Method Enzymol.* **272**, 304-312.
- Sutherland, T. D., G. C. Unnithan, J. F. Andersen, P. H. Evans, M. B. Murataliev, L. Z., E. A. Mash, W. S. Bowers and R. Feyereisen (1998) A cytochrome P450 terpenoid hydroxylase linked to the suppression of insect juvenile hormone synthesis. *Proc. Natl. Acad. Sci. USA* **95**, 12884-12889.
- Swofford, D. L. (1993) PAUP (phylogenetic analysis using parsimony), ver 3.1. Illinois Natural History Survey, Champaign, IL.