

XENOBIOTIC DETOXIFICATION IN THE NAVEL ORANGEWORM *AMYELOIS*
TRANSITELLA (LEPIDOPTERA: PYRALIDAE)

BY

KATHERINE G. NOBLE

THESIS

Submitted in partial fulfillment of the requirements
for the degree of Master of Science in Entomology
in the Graduate College of the
University of Illinois at Urbana-Champaign, 2013

Urbana, Illinois

Adviser:

Professor May R. Berenbaum

ABSTRACT

The highly polyphagous navel orangeworm, *Amyelois transitella*, is a major pest of high-value tree crops such as pistachios and almonds. These non-native crops have been grown commercially in the United States for less than 150 years, and the acquisition by the navel orangeworm of these and other novel hosts may have been facilitated by its metabolic capacity to detoxify xenobiotics. Using Illumina RNA sequencing, I characterized the cytochrome P450 genes transcribed in the midgut of the caterpillar. A total of 319 million paired-end reads were sequenced from larval midgut tissue, and these were assembled *de novo* into 186,000 contigs. Identification of putative P450s in this assembly was carried out using BLAST homology. As detoxification of phytochemicals, mycotoxins, and pesticides is mediated at least in part by P450 enzymes, developing genomic resources for characterizing P450s and elucidating their function in this insect will aid in management efforts for this pest.

An important component of elucidating P450-mediated detoxification in this highly polyphagous species is to understand how ingestion of multiple types of toxins influences detoxification. One pyrethroid pesticide used in almond orchards, α -cypermethrin, is likely detoxified by P450 genes transcribed in the navel orangeworm midgut. Dietary exposure to two phytochemicals found in almonds and pistachios, chlorogenic acid and quercetin, increases survival of navel orangeworms exposed to α -cypermethrin. These phytochemicals may decrease toxicity of the pesticide by inducing the P450s involved in its detoxification. Decreased effectiveness of α -cypermethrin is a concern to almond and pistachio growers. Characterization of the P450s involved in pesticide detoxification in the navel orangeworm on its multiple host plants can inform pesticide management strategies for this pest.

ACKNOWLEDGMENTS

I would like to gratefully thank Kim Walden for all of her guidance and assistance; her knowledge and patience were essential in the completion of this work. Thank you to my adviser, Dr. May Berenbaum, for providing support and a great many late-night edits. Thank you to my committee members, Dr. Hugh Robertson and Dr. Mary Schuler, for allowing me to use your lab space and offering your time and expertise so generously. I would also like to thank Berenbaum lab members current and previous (Dr. Guodong Niu and Dr. Reed Johnson), for their help in all aspects of this project. Thank you to UIUC alumnus Joel Siegel for providing endless help with all things navel orangeworm. Funding was provided by the Cell and Molecular Biology Training Grant from UIUC, the Almond Board of California and the California Pistachio Research Board.

TABLE OF CONTENTS

CHAPTER I: CHARACTERIZATION OF CYTOCHROME P450 TRANSCRIPTION IN THE NAVEL ORANGEWORM MIDGUT

Introduction.....	1
Materials and Methods.....	4
Results.....	10
Discussion.....	14
Figures.....	19
Tables.....	25
References.....	27

CHAPTER II: INCREASED PESTICIDE TOLERANCE OF NAVEL ORANGEWORMS WITH DIETARY EXPOSURE TO PLANT COMPOUNDS

Introduction.....	31
Materials and Methods.....	32
Results.....	33
Discussion.....	34
Figures.....	35
References.....	37

APPENDIX.....	38
---------------	----

I. CHARACTERIZATION OF CYTOCHROME P450 TRANSCRIPTION IN THE NAVEL ORANGEWORM MIDGUT

Introduction

The navel orangeworm (NOW) *Amyelois transitella* (Walker) (Lepidoptera: Pyralidae) is a major agricultural pest in the central valley of California, where it causes great damage to high-value tree crops (Siegel et al., 2010). The native range of NOW extends from the southern United States through South America, and the larval hosts in this range include numerous native species belonging to a diversity of host families, including Asparagaceae, Fabaceae, Rubiaceae and Sapindaceae (Heinrich, 1956). Described originally from fallen citrus fruits in the 1920s, NOW was already considered an orchard pest in California by the 1940s (Wade, 1961). Broadly polyphagous, NOW causes economic damage to pomegranates (Lythraceae), figs (Moraceae), walnuts (Juglandaceae), almonds (Rosaceae) and pistachios (Anacardiaceae) in this region, demonstrating an ability to consume plants from a wide range of chemically disparate plant families (Campbell et al., 2003; Molyneux et al., 2007).

NOW feeds internally in mummy fruits, consuming the unsound, fallen fruits in a manner typical of its closest relatives in the tribe Phycitini (Heinrich 1956). The larvae render unharvested fruits and nuts unusable through physical means (consumption of nutmeat, contamination with webbing and frass) and through vectoring and predisposing nuts to inoculation by fungi, primarily *Aspergillus* spp. The presence of fungal infection in these nuts is closely associated with NOW infestation (Campbell et al., 2003) and production of mycotoxins such as aflatoxin B1 causes further economic injury.

A. transitella has accommodated an expansion onto a variety of novel hosts in a relatively short period of time. The apparently wide ancestral host range of this moth and its more recent

introduction to plants such as the fig, *Ficus carica*, suggests an ancestral physiological preadaptation to for survivorship on these hosts (Janz 2011). Though polyphagy can increase the likelihood of shifting to a new host plant that is taxonomically or phytochemically similar to the ancestral host, the NOW host expansion has occurred in a wide group of chemically distinct host plants (Nylin and Janz 2009). In its realized host range in the California Central Valley, *A. transitella* encounters many biochemically diverse xenobiotics, such as synthetic organic pesticides, mycotoxins, and host phytochemicals (Niu et al., 2012; <http://www.ars-grin.gov/duke/> and Duke, 1992). Plants in the ancestral host range of *A. transitella*, such as members of the family Fabaceae, have both cyanogenic glycosides and furanocoumarins; these chemical classes occur in almonds and figs, respectively (Veitch et al., 2010; Seigler 1998). The internal feeding behavior on unsound fruits, which is shared with other members of the tribe Phycitini, also suggests a capacity for mycotoxin detoxification (Heinrich 1956), inasmuch as feeding damage predisposes fruits to colonization by fungi. An ancestral capacity for detoxification of these xenobiotics may have facilitated the NOW transition to the California agroecosystem.

The extent to which the host range expansion of *A. transitella* is governed by its detoxificative capacity is an area of interest for agricultural research. Biochemical characterization of detoxificative processes in NOW may assist in efforts to restrict further expansion and reduce the likelihood of pesticide resistance acquisition. The midgut is an area of xenobiotic interface in insects and the presence and activity of cytochrome P450 enzymes involved in detoxification in this tissue are well-studied (Feyereisen, 2011). Cytochrome P450 enzyme-mediated furanocoumarin detoxification has been characterized in at least six species in four families of Lepidoptera (reviewed in Schuler 2011). Moreover, three P450s previously characterized in *A. transitella* demonstrate that P450-mediated furanocoumarin detoxification

occurs in this species (Niu, 2011). Although a specific P450 that metabolizes xanthotoxin, a linear furanocoumarin that occurs in figs, has not been identified in *A. transitella*, this compound is likely metabolized by P450 enzymes because its toxicity to NOW is synergized by piperonyl butoxide, a specific P450 inhibitor (Niu et al., 2012). In this study, we performed RNA sequencing of the NOW midgut after exposure to xanthotoxin in order to shed light on the P450-mediated biochemical capacity that might mediate furanocoumarin detoxification as an element of NOW host range expansion.

Materials and Methods

Navel orangeworm colony:

A laboratory colony of *A. transitella* designated as SPIRL-1966 has been maintained at the USDA San Joaquin Valley Agricultural Research Center since 1966. Larvae from this source were reared on standard wheat bran diet at the University of Illinois Urbana-Champaign at 28 °± 4°C and photoperiod of 16:8 (L:D) h. (Siegel et al., 2010).

RNA extraction and sequencing:

Fifth instar male and female larvae that had molted within a 24-hr period were placed on artificial diet for two days. This diet was supplemented with 2 mg/g xanthotoxin, a concentration that causes less than 20% mortality of early NOW instars (Niu, 2010).

Midguts from 33 larvae were dissected into cold 0.1M sodium phosphate buffer (pH 7.8) and flash-frozen in liquid nitrogen. Homogenization and RNA extraction were carried out using the RNeasy Mini Kit and manufacturer's protocol (Qiagen, Valencia, CA). The RNA concentration was measured with an Agilent 2100 Bioanalyzer at the University of Illinois Roy J. Carver Biotechnology Center. To check for genomic DNA contamination, 2 uL were run on a 1% agarose gel with a 100bp plus DNA ladder (Thermo Scientific, Waltham, MA). A total of 52.44 ug/30 ul RNA in RNase-free water was sent to the UIUC Roy J. Carver Center for construction of an RNA-Seq library with the Mate Pair Library Prep Kit (Illumina) and sequencing on a single paired-end lane on an Illumina HiSeq2000 instrument.

De novo assembly of reads:

The first two 5' bases of each sequenced read were trimmed using the fastx_trimmer script from the FASTX toolkit [http://hannonlab.cshl.edu/fastx_toolkit/]. The 3' ends were trimmed to remove low-quality bases using the fastq_quality_trimmer with a quality score cutoff

of 20 and with a modification to retain reads of length 0 to conserve mate pairs for downstream assembly. This was carried out in both forward and reverse directions for the paired-end reads. *De novo* transcriptome assembly was carried out in Trinity, an RNA-Seq data assembly program that is particularly useful for large datasets and identification of splice isoforms (Grabherr et al., 2011).

Midgut transcriptome analysis:

To annotate the midgut transcriptome, all contigs with a minimum length of 500 nt from the Trinity assembly were processed in Blast2go (Conesa et al., 2005). Using default parameters, I searched transcript contigs against the June 2012 non-redundant protein database at NCBI using BLASTX. Hits with an e value less than 1×10^{-6} were mapped to simplified gene ontology (GO-Slim) terms, the InterPro protein database (v. 42.0) for further functional annotation or the Kyoto Encyclopedia of Genes and Genomes (KEGG v.66.1) for pathway annotation (Kanehisha et al., 2012).

Putative P450s in the midgut transcriptome:

To identify potential transcripts of P450 genes, the 186,000 assembled *A. transitella* midgut transcripts were formatted into a BLAST database and queried with a compiled set of 156 annotated lepidopteran cytochrome P450s. This set was composed of amino acid sequences from lepidopteran P450s available at NCBI including *Bombyx mori*, *Manduca sexta*, *Helicoverpa zea*, *Mamestra brassicae*, *Heliothis virescens* and three previously cloned P450s from *Amyelois transitella* [<http://www.ncbi.nlm.nih.gov/>]. Representatives from each lepidopteran CYP family were included.

Due to the divergence of P450 protein sequences, even low score (S) BLAST hits between midgut transcripts and annotated P450s were explored. The presence of any of the five

conserved protein motifs in insect P450s, the WxxxR motif, the GxE/DTT/S motif, the ExLR motif, the PxxFxPE/DRF motif, and the PFxxGxRxCxG/A motif, were also used to identify potential P450 transcripts (Feyereisen, 2011.) Trinity-assembled contigs containing potential P450s were translated into the longest ORF in DNASTrider, and amino acid sequences and full-length cDNAs from the first start and stop codon were noted (Marck, 1988).

RNA-Seq is a tool for quantitative transcription analysis and the read counts of a particular transcript should represent the gene expression level. Transcription levels of full-length P450s were compared within the single midgut sample using the Burrows-Wheeler Aligner (BWA) software package (Li and Durbin, 2010). This program is designed for comparison of transcript expression levels among different samples; as the *A. transitella* midgut transcriptome is based on one sample of pooled individuals, only large-scale expression differences can be described with this program. The first two 5' bases of each sequenced read were trimmed using the fastx_trimmer script from the FASTX toolkit [http://hannonlab.cshl.edu/fastx_toolkit/]. The 3' ends were trimmed to remove low-quality bases using the fastq_quality_trimmer with a quality score cutoff of 20. A total of 319,781,126 reads were mapped in single-end mode. To control for length differences between transcripts, expression was measured as #reads/kb transcript.

Extension of partial P450 transcripts:

The amino acid sequences of 61 putative P450 proteins were aligned in ClustalW to detect any translated Trinity contigs lacking N- or C- termini (ClustalW v.2, 2007; Ai et al., 2011). In total, 43 partial cDNA sequences were identified and searched against the current assembly of the *A. transitella* genome formatted into a BLAST database. Once a truncated transcript was

located on a genomic scaffold, upstream or downstream scaffold sequence was searched against the NCBI nr database using BLASTX to detect coding regions.

Comparison with intron boundaries of other *A. transitella* putative P450 genes guided refinement of P450 transcript extension. Extended transcripts were verified by a BLASTN search against the 100 bp RNA-Seq raw reads to confirm that each putative full-length cDNA was supported with overlapping reads. Nine P450s were successfully lengthened, and the remaining 35 partial sequences can be resolved with improved genome assemblies.

P450 gene introns:

Putative midgut P450 transcripts were searched against the current 36X coverage assembly of the NOW genome in order to discern scaffold position, intron position, length and phase to the extent presently possible. Phase 0 introns are inserted between codons, phase 1 introns are inserted between the first and second nucleotides in a codon, and phase 3 introns are inserted between the second and third nucleotides in a codon. Multiple sequence alignment of intron position and phase among full-length *A. transitella* P450s was carried out using a modified exon alignment Perl script from Dr. Reed M. Johnson.

P450 nomenclature:

The final set of putative 62 P450s is composed of all full-length transcripts that could be mapped to unique positions within the current genome assembly. Naming of putative P450s was carried out by Dr. David R. Nelson. Names were based on percent amino acid identity of *A. transitella* P450s compared to previously annotated P450s, some of which are available only in his database. P450 genes are assigned to a particular family with >40% identity and subfamily with >55% amino acid identity (Feyereisen, 2012). Different alleles of a gene are designated as subscripts v1, v2, and so on. Identity-based nomenclature of CYP alleles may introduce errors;

while a small number of nucleotide differences may distinguish two alleles, it may also represent the difference between two genes, exemplified in the *Drosophila melanogaster* and *Anopheles gambiae* genomes (Feyereisen, 2005). To verify Dr. Nelson's designations, the positions of each proposed P450 transcript in the current genomic scaffolds were compared; if two proposed transcripts overlapped at the same location, they were confirmed as allelic variants at the same locus.

Phylogenetic analysis:

Alignment of full-length *A. transitella* P450s with all 78 functional *B. mori* P450s was carried out in ClustalW. Allelic variants of *B. mori* P450s were not included, nor were previously cloned NOW P450s. Phylogenetic tree reconstruction was performed in MEGA 4 using the neighbor-joining algorithm with a bootstrap of 1000 replicates (Tamura et al., 2007). The same alignment and tree reconstruction techniques were repeated for a comparison among full-length *A. transitella* P450s. Phylogenetic trees were graphically edited in FigTree v1.4 [<http://tree.bio.ed.ac.uk/software/figtree/>].

Resolution of P450 isoforms:

The 37 putative P450s identified as separate isoforms in the Trinity assembly were remarkably similar, differing by a few nucleotides or by only the 3' UTR. Alignment of these transcripts to the current genome did not clarify from which genes these are transcribed, as a tBLASTn comparison locates them on the same scaffold within the same 57kb of genomic DNA and sequence gaps remain in this region. Removal of the 3' UTR and alignment in Sequencher 5.1 (Gene Codes Corporation, Ann Arbor, MI) with the "dirty data" assembly algorithm established only three potentially distinct transcripts. The intron positions of only one of these

transcripts could be discerned, as the other two could not be reliably aligned to the genome due to scaffold gaps.

Estimation of total number of P450s in the genome:

The number of P450s present in the NOW genome was approximated by searching the set of putative P450 transcripts against a previous genome assembly with BLASTn. Any unique scaffolds and base positions with low e-value hits or matching conserved insect P450 motifs were noted as potential P450 gene loci. This method provides only a rough estimate of the number of P450s in the genome, and each potential locus will be evaluated in future genome assemblies for gene prediction.

Results

NOW midgut transcriptome:

Sequencing of *A. transitella* midgut RNA produced 319 million paired-end reads, each 100 nt in length. Paired-end sequencing is recommended for a *de novo* transcriptome assembly, and the NOW midgut transcriptome has particularly deep coverage owing. Informative *de novo* transcriptomes have been assembled from 43 million raw reads in the silverleaf whitefly *Bemisia tabaci*, and 39 million in the case of the Asiatic rice borer *Chilo suppressalis* (Wang et al., 2010; Sparks et al., 2013). Trinity assembly of the NOW midgut transcriptome resulted in 186,000 contigs with an average length of 700 nt, the shortest being 101 nt and largest being 46,644 nt. Each of these contigs represents a single transcript, with distinction between various isoforms. The depth of sequencing of this transcriptome allows for representation of isoforms that are expressed even at low levels in the midgut.

Only a subset (N=41,356) of these contigs were processed in Blast2go; contigs shorter than 500 nt were not included and the program discarded 1,837 contigs that exceeded the maximum size limitation of 8 kb (Conesa et al., 2005). From this subset, 41% of contigs had significant similarity to the NCBI non-redundant protein, KEGG, InterPro, or GO databases with an e value cutoff of 1×10^{-6} (Kanehisa et al., 2012). The top BLASTx hits of *A. transitella* transcripts with the NCBI non-redundant protein database were made chiefly with proteins from other insect species (Fig. 1.1). Of these, the monarch butterfly *Danaus plexippus* and the domestic silkworm *B. mori* had the most top hits. KEGG pathways were associated with 888 different Trinity-assembled transcripts from the NOW midgut . InterPro and GO annotation were associated with 20,135 transcripts, corresponding with GO levels 1-7 in molecular function,

cellular component, and biological process terms (Fig.1.2). The *A. transitella* midgut transcriptome appears to be comprehensive.

Identification and phylogeny of midgut P450s:

Analysis of the NOW midgut transcriptome yielded 61 putative P450s, a conservative set comprised only of transcripts which could be mapped to unique locations on the NOW genome (Table 1.1). While 35 unextended partials remain, at least 17 of these are located on the same genomic scaffold interrupted by sequence gaps and may actually be truncated contigs from the same gene. The amino acid sequences of *A. transitella* P450s have the five motifs conserved in insects, although in CYP301A1 there is a Y instead of F in the PFxxGxRxCxG/A motif; this substitution is found in CYP301A1 orthologs (Feyereisen, 2011). The N-terminal sequence targets of CYPs in Clan 2, 3 and 4 are rich in hydrophobic amino acids, such as M, V, I, W, L, and F. In the compiled set of lepidopteran P450s, this motif is also found in CYP304 and CYP4M representatives from *H. zea*, *M. brassicae* and *M. sexta*.

The set of 61 putative P450s transcribed in the NOW midgut comprises the four insect CYP clans: CYP2, CYP3, CYP4 and mitochondrial (Table 1.1). The CYP3 clan has the most representatives (N=32) transcribed in the midgut, while the CYP2 clan and mitochondrial Clan have six and seven members, respectively. The CYP6 family has the most members (N=20) at this level of classification. Three P450s were named as allelic variants: CYP321C1v2, CYP6AB11v2, and CYP6B44v2; these are allelic variants of three previously cloned *A. transitella* P450s (Niu et al., 2011). A reconstruction of the NOW P450 phylogeny using neighbor-joining methods shows clade separation between each CYP clan, with Clan 3 members undergoing the greatest levels of expansion (Fig. 1.3). At the subfamily level, CYP6AE and

CYP6AB groups appear to be undergoing the greatest levels of expansion among *A. transitella* P450s transcribed in the midgut.

Phylogenetic comparison between *A. transitella* P450s transcribed in the midgut and *B. mori* genomic P450s presents several 1:1 orthologs and many lineage-specific expansions at the subfamily level (Fig. 1.4). In particular, the CYP6B and CYP6AB subfamilies in *A. transitella* appear to be undergoing expansion relative to the *B. mori*. In both species, the mitochondrial clan appears to be undergoing less expansion than other CYP clans, consistent with its putative conserved function in hormone metabolism.

P450 transcript relative expression levels:

BWA analysis and subsequent corrections showed three orders of magnitude difference in expression levels among P450s, with #reads/kb ranging from 306, 179 for CYP6AE58 and 82 for CYP304F12 (Table 1.2). Four members of the CYP3 clan and one member of the mitochondrial clan are the six most highly expressed P450s. Relative expression levels of P450 enzymes in the midgut transcriptome may be a function of their importance in NOW midgut as a site of detoxification.

P450 genes in the NOW genome:

The coding regions of NOW P450 genes from the midgut transcriptome have 0-9 introns; four of these genes are entirely intronless in the coding region (Fig. 1.6). With the exception of one GC intron donor site in CYP304F12, all NOW P450 introns follow the canonical GT-AG rule. There exist several shared intron sites among different clans: 31 CYP genes from Clans 3 and 4 share a phase 1 intron at the second glycine of the heme-binding motif, and 10 CYP genes from Clan 4 and the mitochondrial clan share a phase 1 intron at variable amino acids in the heme-binding motif. While intron analysis of the *B. mori* P450s shows no paralogous gene pair

in the mitochondrial clan that shares intron organization, this is not true for *A. transitella* P450s (Ai et al., 2011). In the mitochondrial clan, CYP333A8 and CYP333A10 share all eight introns, and CYP333B19 and CYP333B20 share five. In general, the CYP6 family shares intron organization, as do the CYP341 and CYP4 families.

Based on BLASTn comparison of NOW midgut P450s and a draft genome assembly, an estimated 14 distinct scaffold locations have P450s that lack annotation. With addition to the 18 partial P450s that can be located on distinct scaffolds and the set of 61 P450s from the transcriptome, a rough approximation of the NOW CYPome is 93 P450s.

Discussion

With 312 million sequenced reads and 186,000 Trinity-assembled contigs, the *A. transitella* midgut transcriptome has high sequence coverage, imparting coverage confidence even for transcripts expressed at low levels. Characterization of the *A. transitella* midgut transcriptome is a descriptive exercise; quantitative analyses within our sample are biased because transcript counts of isoforms identified by Trinity may cause overestimation of the number of gene loci. To mitigate this potential source of error, a functional annotation of the *Lymantria dispar* transcriptome clusters putatively unique transcripts (PUTs) into a single likely unigene (Sparks et al., 2013). Even so, the large quantity of descriptive information from the *A. transitella* midgut transcriptome may be useful for describing general patterns across the increasing number of published transcriptomes. Although only 41% of *A. transitella* transcripts in our Blast2go subset were functionally annotated, this is comparable to the 45% transcripts in a reduced subset of a *Plutella xylostella* midgut transcriptome that mapped to GO terms and the nr database at NCBI (Xie et al., 2012).

In general, the *A. transitella* midgut transcriptome shares GO terms with the whole-body *de novo* transcriptome of the olive fruit fly *Bactrocera oleae* and the greenhouse whitefly *Trialeurodes vaporariorum*, where binding, cell, and metabolic process are the most abundant GO level 2 categories (Pavlidi et al., 2013; Karatolos et al., 2011). *De novo* characterization of the midgut transcriptome has been carried out in three other species of Lepidoptera, *P. xylostella*, *M. sexta*, and *C. suppressalis*, but global comparison between these datasets is tentative at best because of the varying levels of analysis that can be reported through annotation (Sparks et al., 2013; Pauchet et al., 2010; Ma et al., 2012). There now exist several large datasets of GO terms for midgut transcriptomes that await large-scale analysis.

Quantification of P450 transcription via transcriptome annotation is not yet accurate; based on Blast2go sequence identity annotation the *A. transitella* midgut transcriptome contained 114 transcripts annotated as different P450 genes. This overestimation may be a product of the isoforms in the Trinity assembly or due to the proliferation of misleading GO and InterPro terms associated with cytochrome P450s that do not involve the five conserved insect motifs (Feyereisen, 2012). The more detailed approach to P450 annotation employed in the *A. transitella* project resulted in annotation of 61 full-length P450 transcripts that could be mapped to unique locations on the genome assembly. The *C. suppressalis* midgut transcriptome reports a total of 70 P450s, and the *L. dispar* midgut transcriptome 100 using the Pfam PF00067.16 accession number that Feyereisen (2011) recommends as “precise and unambiguous.” However, while both approaches involved measures to reduce isoform bias, neither was based on full-length transcript identity and did not account for partial reads that may cause overestimation. While the set of 61 P450s described from the *A. transitella* midgut is conservative, this can be addressed with future assemblies of the genome.

Enumeration of P450s in a transcriptome or genome has been the focus of studies of P450 gene evolution (Karatalos 2011; Claudianos et al., 2006). The estimate of 93 members of the *A. transitella* CYPome is comparable to the 87 members in the *B.mori* genome, and 86 in the *D. plexippus* genome [<http://drnelson.uthsc.edu/CytochromeP450.html>]. However, though P450 number is informative for some aspects of evolution, there is not necessarily a straightforward correlation between capacity to withstand xenobiotic exposure and size of the genomic complement of P450s. Notably, the human body louse *Pediculus humanus* exhibits P450-mediated resistance to permethrin but has a relatively small 36-member CYPome (Audino et

al.,2005; Lee et al., 2010). CYPome diversity at the CYP clan, family and subfamily level may be the most informative site for comparison among different species.

Transcriptome level analysis of P450s in the NOW midgut revealed that a wide diversity of P450s are transcribed in this tissue. P450s from the mitochondrial clan are known to be expressed in the midgut; one example is the Halloween gene *CYP314A1* (*Shade*), which in *Drosophila melanogaster* is expressed in midgut, fatbody, and epidermis (Petryk et al., 2003). Seven mitochondrial P450s are expressed in the NOW midgut, including the *Shade* ortholog CYP314A1, and CYP301A1, an enzyme implicated in cuticle development with presumed 1:1 orthologs in all sequenced insect genomes (Sztal et al., 2011). Although many insect mitochondrial P450s carry out endogenous processes such as ecdysteroid biosynthesis, CYP314 in the cladoceran water flea *Daphnia magna* is upregulated with exposure to the herbicide glyphosate and the organophosphate insecticide methidathion (Le et al., 2011). The biochemical characterization of mitochondrial P450s in the NOW midgut may shed light on whether they modify endogenous or exogenous substrates.

As a xenobiotic interface, the NOW midgut suitably expresses a range of P450s associated with detoxification. The blooms evident in the CYP6AE and CYP6AB lineages may be related to the NOW capacity for withstanding a wide range of plant secondary compounds, including the demonstrably insecticidal furanocoumarins, in their diet. Members from these subfamilies have been characterized in other Lepidoptera; CYP6AE14 in *H. armigera* shows activity against the phenolic aldehyde gossypol, and CYP6AB5 in *B. mori* is induced by the flavonoid quercetin (Mao et al., 2007; Zheng et al., 2012). Even allelic variants of an enzyme may offer considerably different detoxification functions; in *D. pastinacella*, another microlepidopteran capable of metabolizing furanocoumarins (Mao et al. 2006), CYP6AB3v1 and

CYP6AB3v2 enzymes metabolize imperatorin at significantly different rates (Mao et al. 2007). CYP6AB11, the CYP6AB subfamily member previously characterized from *A. transitella*, displayed highly specific activity against the linear furanocoumarin imperatorin (Niu et al., 2011).

The CYP6B lineage in *A. transitella* may also be related to furanocoumarin detoxification. This group has undergone extensive expansion in members of the family Papilionidae, which has generalist and specialist consumers of furanocoumarin-containing host plants (reviewed in Schuler, 2011). Notably, CYP6B1 from *Papilio polyxenes* exhibits a very high rate of furanocoumarin turnover, particularly of xanthotoxin (Wen et al., 2003). CYP6B8 in *H. zea* exhibits a broader substrate range, detoxifying six biosynthetically diverse phytochemicals including xanthotoxin as well as the three structural classes of insecticides including diazinon, cypermethrin, and aldrin (Li et al., 2004). CYP6B44 was previously cloned from *A. transitella*, but the substrate remains unidentified and it has no activity against xanthotoxin or aflatoxin B1 when heterologously expressed (Niu, 2010).

Members of the CYP321 family in *A. transitella* are also particularly interesting with regard to xenobiotic detoxification. In *H. zea*, CYP321A1 has a broad range of substrates, including aflatoxin B1, xanthotoxin, and angelicin, as well as the insecticides α -cypermethrin, diazinon, and aldrin (Niu et al., 2008; Rupasinghe et al., 2007). The broad range of substrates detoxified by CYP321A1 may reflect adaptation to polyphagy in *H. zea*. In an attempt to characterize the P450 that allows *A. transitella* to withstand even greater amounts of aflatoxin than *H. zea*, CYP321C1 was cloned and characterized from the NOW midgut (Niu et al., 2011). Although CYP321C1 did not metabolize aflatoxin B1, other NOW P450s in the CYP321 family may be worth characterizing in this pursuit.

While P450 protein diversity is a factor in an insect's capacity to withstand xenobiotic exposure, the level of P450 enzyme expression can also contribute to this end. In populations of *Musca domestica* that are resistant to deltamethrin, overexpression of CYP6A1 confers this trait (Cariño et al., 1995). Similarly, overexpression of CYP9A12, CYP9A14, and CYP6B7 contribute to pyrethroid resistance in populations of *H. zea* (Yang et al., 2006). Relative expression levels of P450s in *A. transitella* show ten-fold changes in expression between groups measured in base 10, base 100, base 1000, and base 10,000 reads/kb. Three of the most highly expressed P450 transcripts in the *A. transitella* midgut, CYP6AB43, CYP6B54, and CYP9A66, are related to sequences encoding enzymes that metabolize furanocoumarins and pesticides in other lepidopterans and may well be involved in detoxification of these substrates in *A. transitella*.

Through analysis of protein diversity and expression levels, several P450s of interest have been characterized. Future work characterizing their metabolic activity may clarify which P450 enzymes play a role in the extraordinary capacity of *A. transitella* for xenobiotic detoxification, an attribute that may underlie the ability of this species to colonize a wide range of crop species recently introduced into California and may also facilitate the evolution of pesticide resistance in this economically important pest species.

Figures

Figure 1.1. Species associated with lowest e-value BLAST hits between NCBI nr protein database and Trinity-assembled *Amyelois transitella* contigs in Blast2go. The monarch butterfly, *Danaus plexippus* had the most top-hits (lowest e-value) with *A. transitella*, followed by the silk moth *Bombyx mori*.

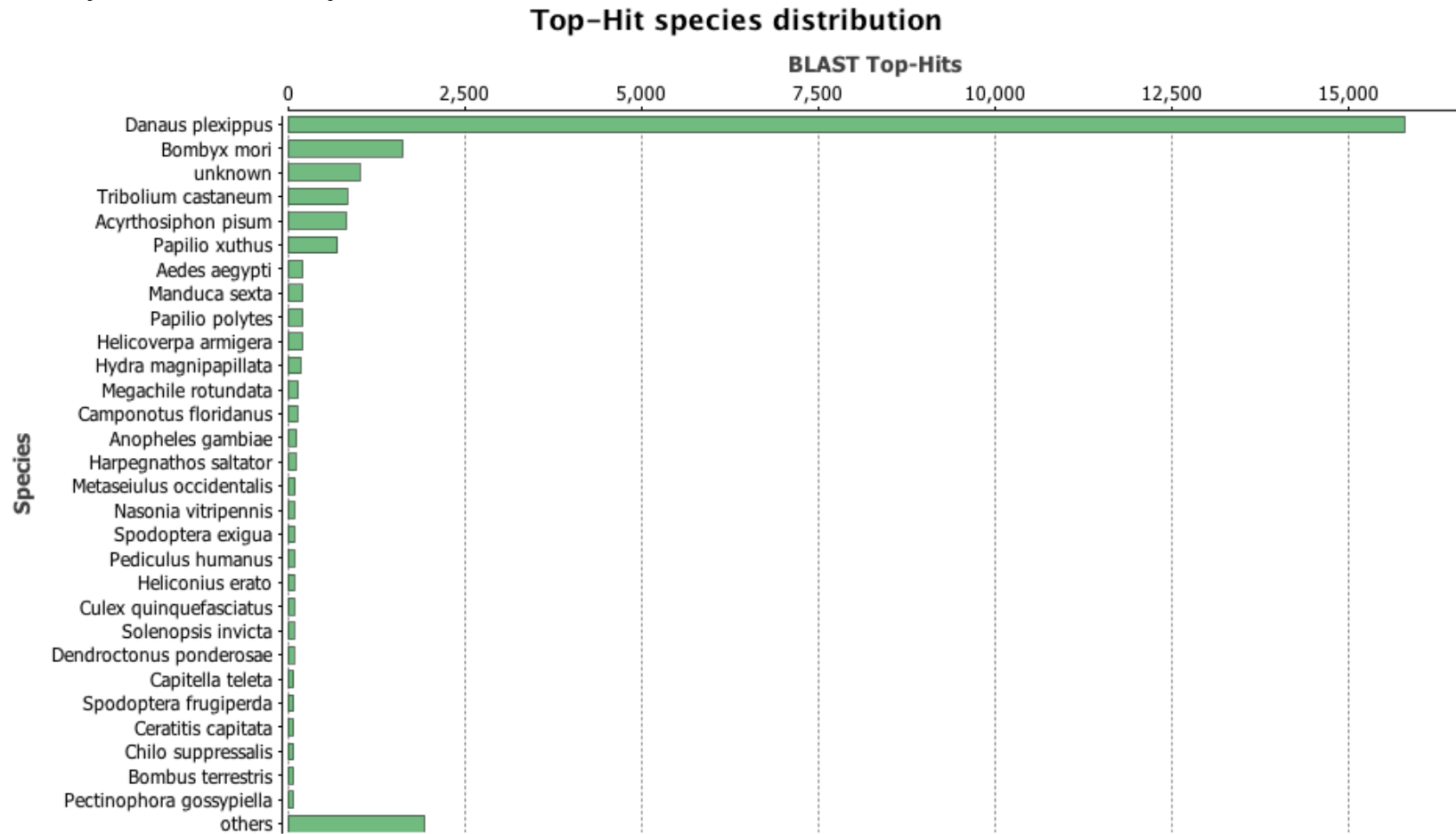


Figure 1.2. GO-Level distribution frequency of *A. transitella* midgut transcripts analyzed in Blast2go. Levels 1-8 are represented in biological process, molecular function, and cellular component terms.

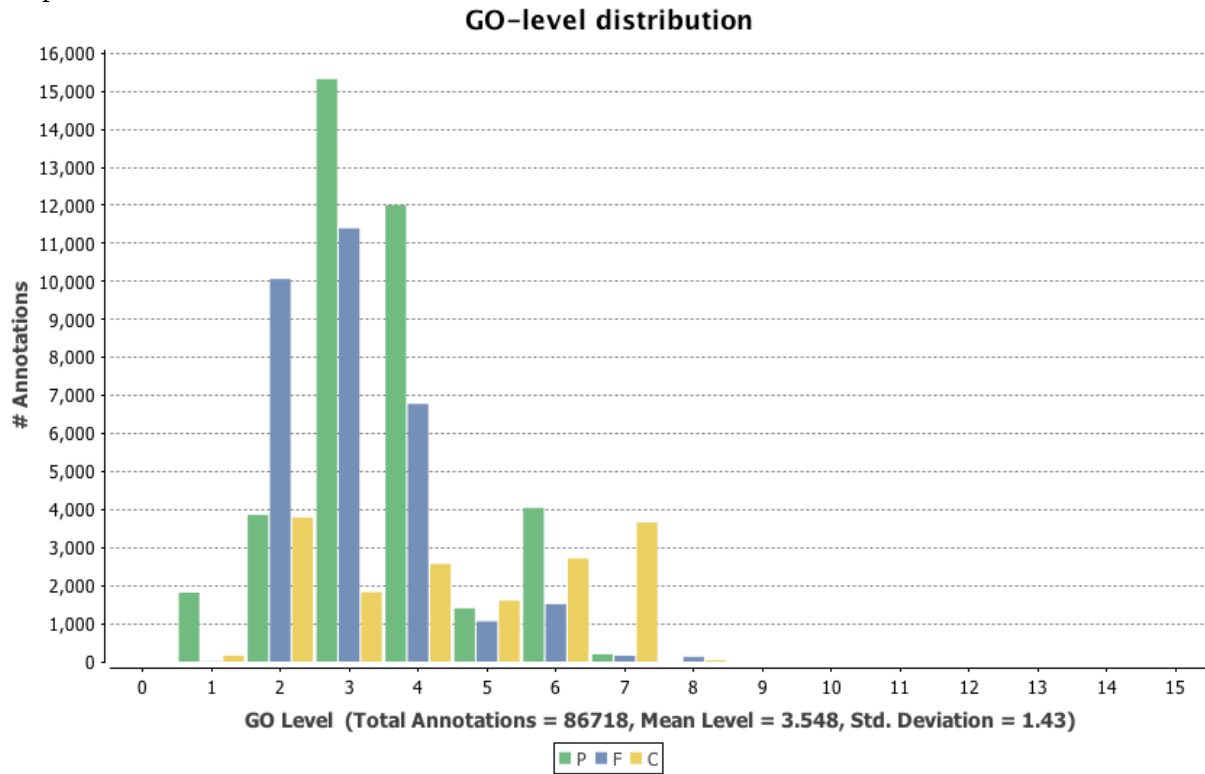


Figure 1.3. Level 2 GO term distribution of midgut transcripts for biological process (BP), cellular component (CC), and molecular function (MF) categories. Binding, cell, and metabolic process are the most abundant categories.

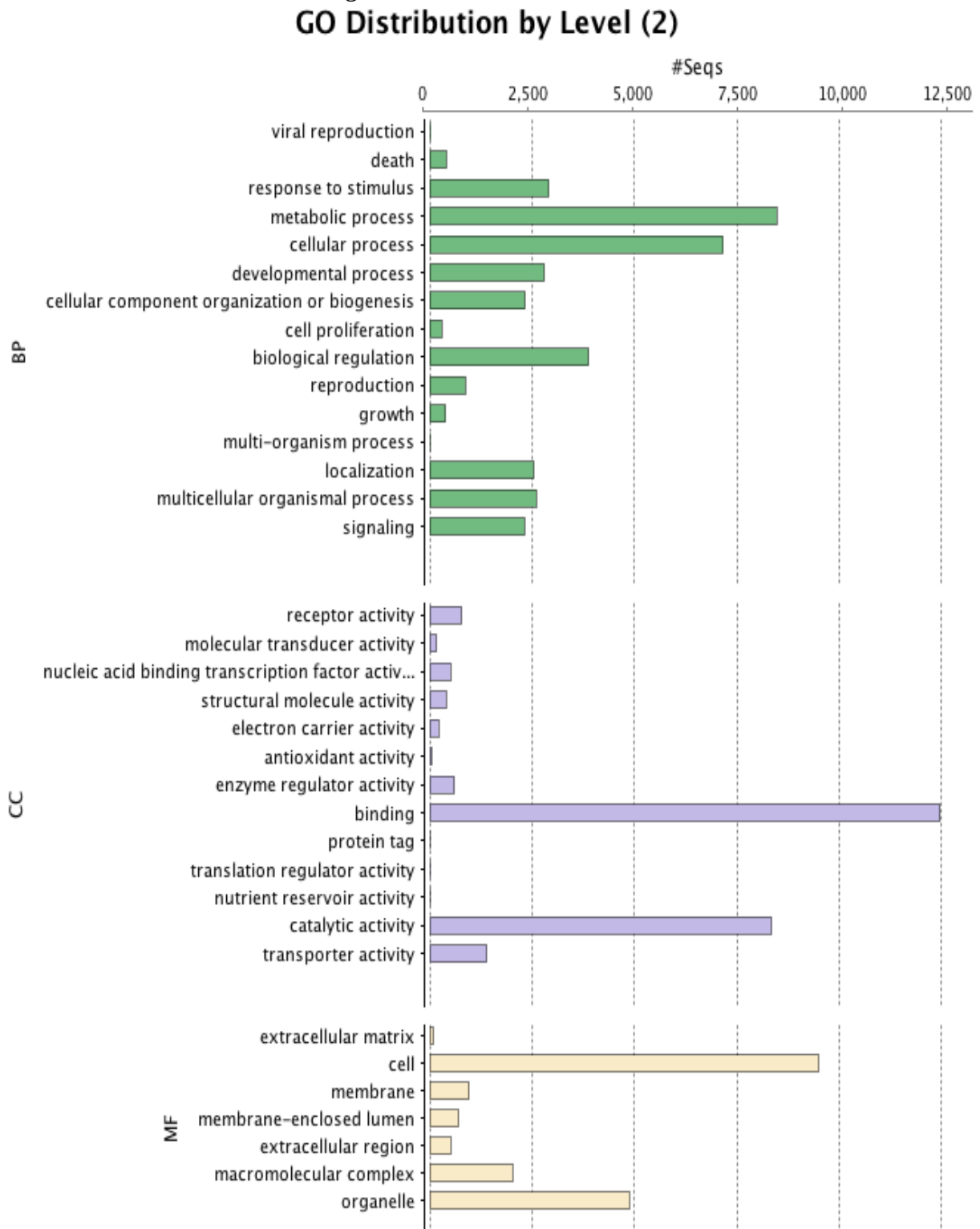


Figure 1.4. Neighbor-joining tree of *A. transitella* P450s with phylogeny inferred from 1000 bootstrap replicates. Branches with bootstrap values <50% are collapsed.

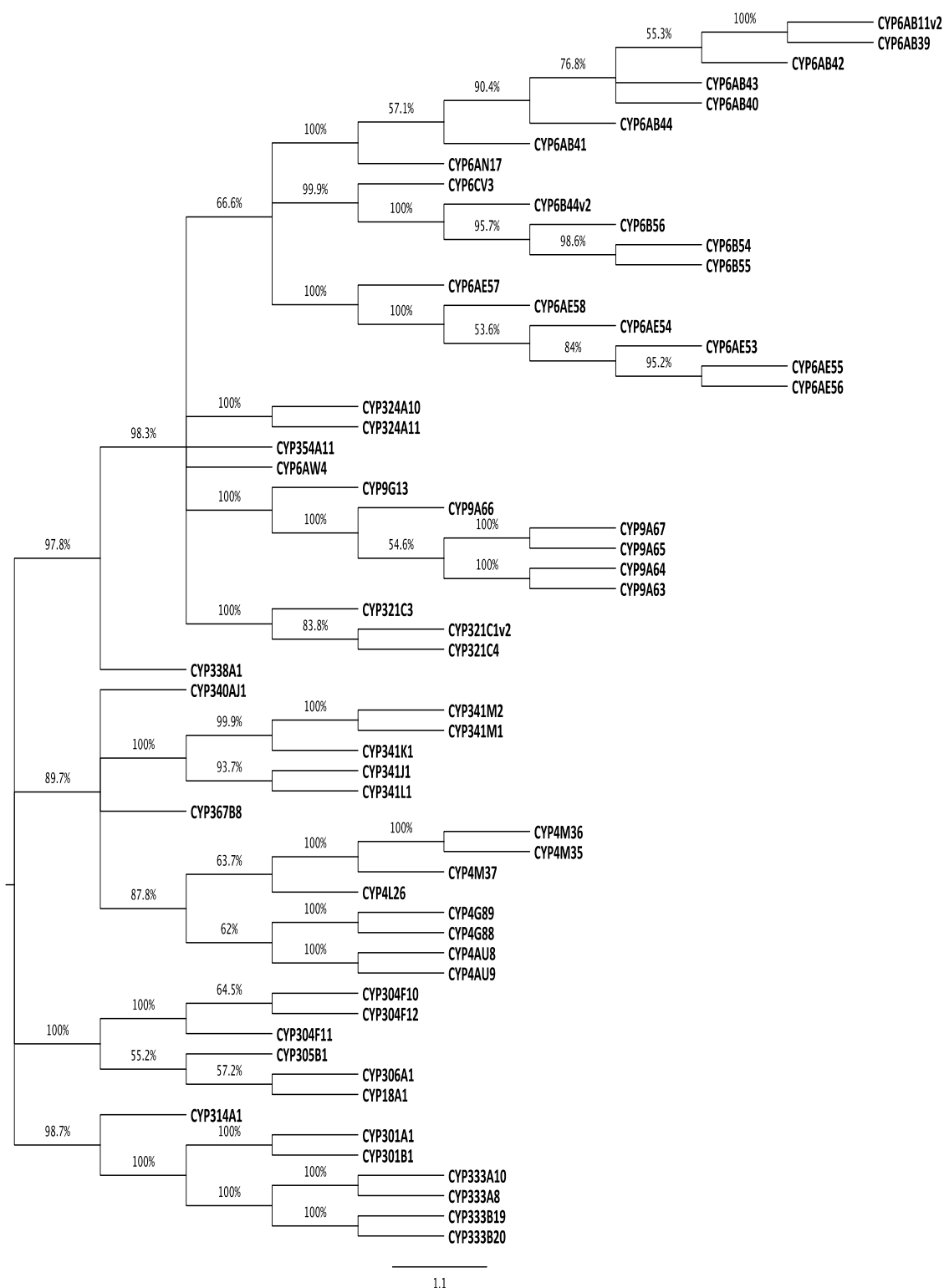


Figure 1.5. Neighbor-joining tree of *Amyelois transitella* midgut transcriptome and *Bombyx mori* genomic P450s. Phylogeny inferred from 1000 bootstrap replicates.

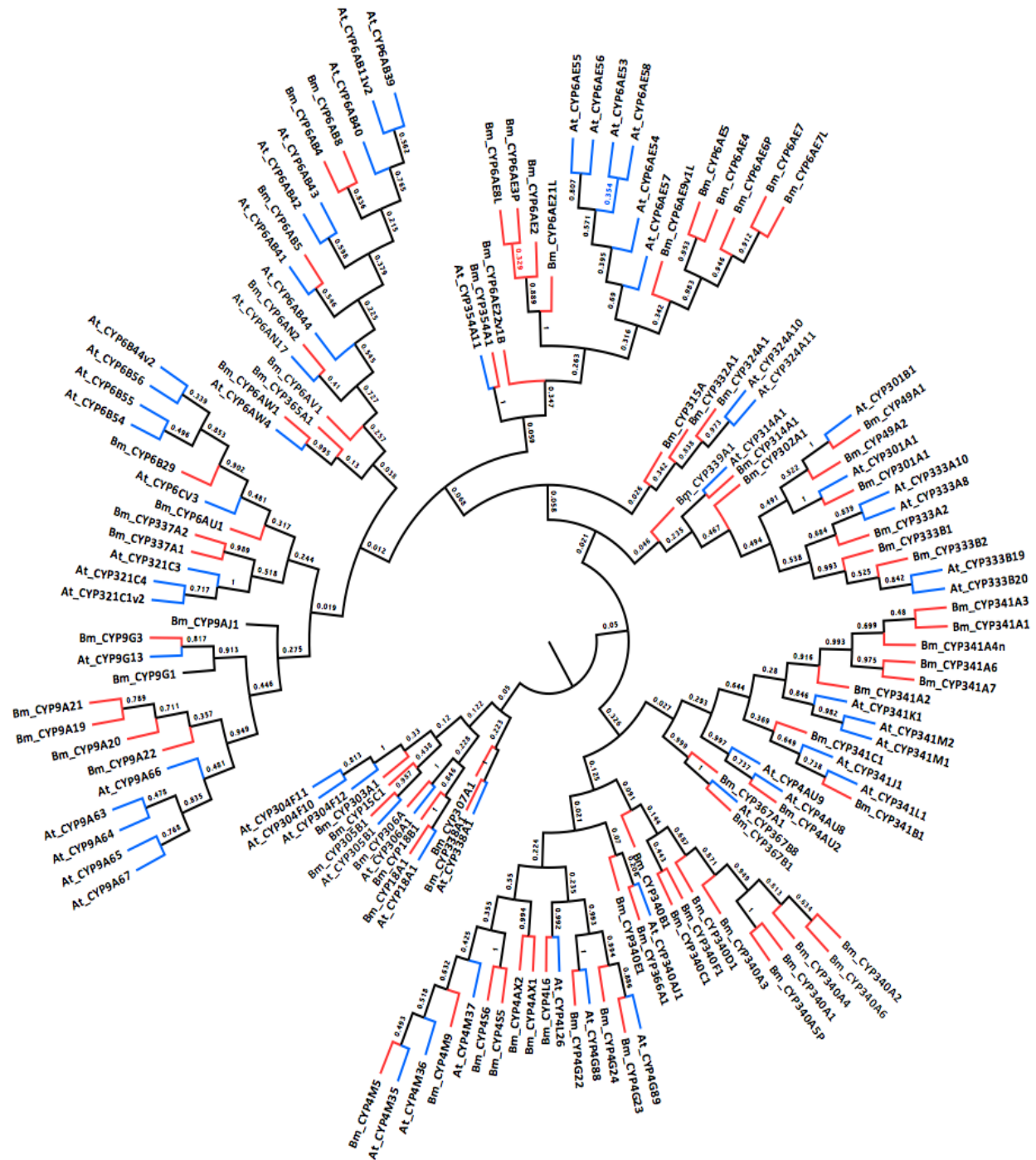
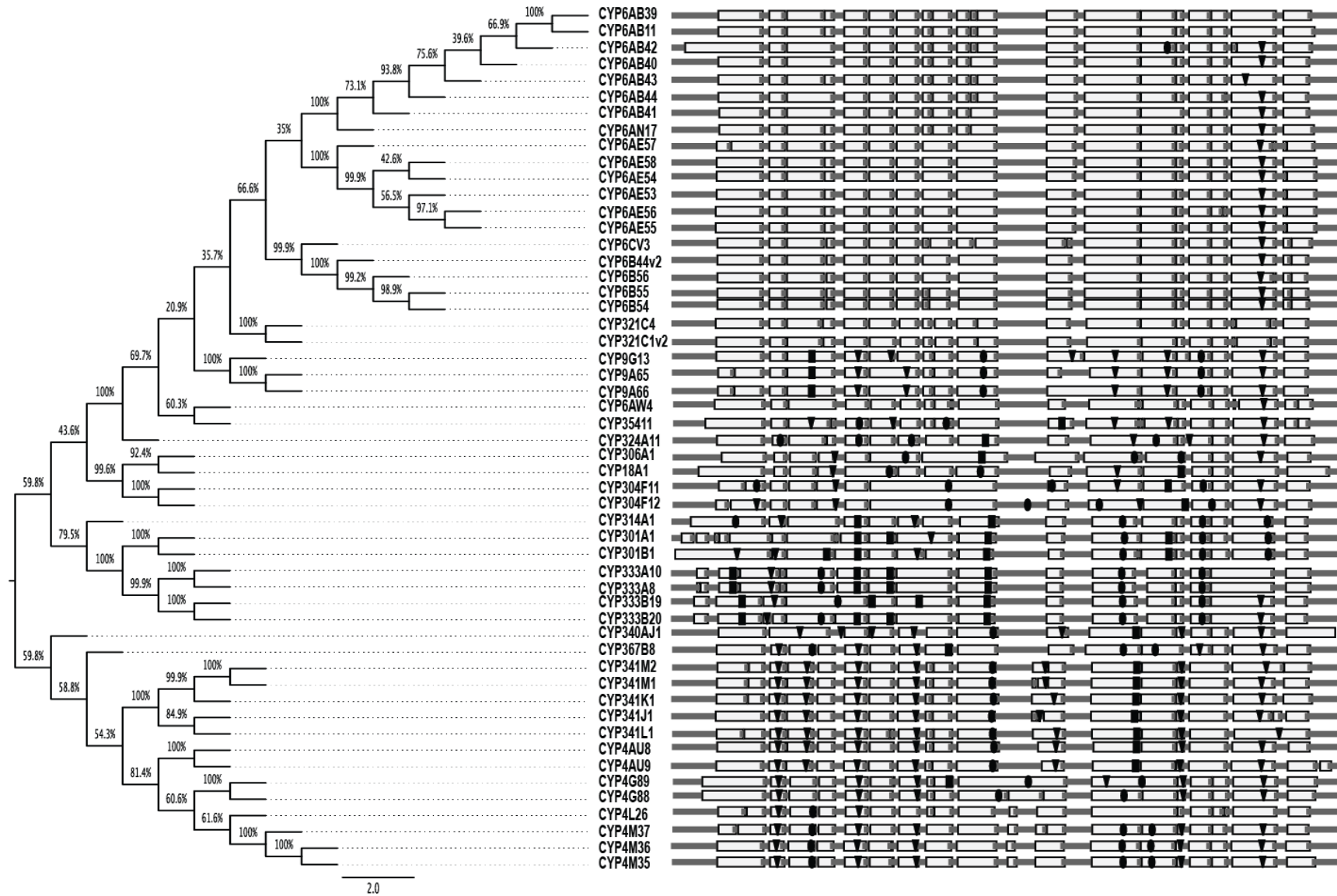


Figure 1.6. Alignment of *A. transitella* P450 amino acids transcribed in the midgut. Ovals represent phase 0 introns, triangles are phase 1, and squares are phase 2. Accompanying tree built with neighbor-joining method with 1000 bootstrap replicates.



Tables

Table 1.1. Nomenclature assigned by Dr. David Nelson to putative P450s transcribed in the *A. transitella* midgut. All four cytochrome P450 clans (CYP2, CYP3, CYP4, and mitochondrial) have representatives in this set. At a total of 32 transcribed P450s, the CYP3 Clan has the greatest number of P450 proteins transcribed; 20 of these belong to the CYP6 family, which has the greatest number of P450 proteins among families in this set.

CYP Clan	P450 Family	Number	P450 Subfamily Members
CYP2	CYP304	5	CYP304F10, CYP304F12, CYP304F11
CYP2	CYP305	1	CYP305B1
CYP2	CYP306	1	CYP306A1
CYP2	CYP18	1	CYP18A1
CYP3	CYP6	20	CYP6AB11v2, CYP6AB39, CYP6AB42, CYP6AB43, CYP6AB40, CYP6AB44, CYP6AB41, CYP6AN17, CYP6CV3, CYP6B44v2, CYP6B56, CYP6B54, CYP6B55, CYP6AE57, CYP6AE58, CYP6AE54, CYP6AE53, CYP6AE55, CYP6AE56, CYP6AW1
CYP3	CYP9	6	CYP9G13, CYP9A66, CYP9A67, CYP9A65, CYP9A64, CYP9A63
CYP3	CYP354	1	CYP354A11
CYP3	CYP321	3	CYP321C3, CYP321C1v2, CYP321C4,
CYP3	CYP338	1	CYP338A1
CYP3	CYP324	2	CYP324A10, CYP324A11
CYP4	CYP4	8	CYP4M36, CYP4M35, CYP4M37, CYP4L26, CYP4G89, CYP4G88, CYP4AU8, CYP4AU9
CYP4	CYP340	1	CYP340AJ1
CYP4	CYP367	1	CYP367B8
CYP4	CYP341	5	CYP341M2, CYP341M1, CYP341K1, CYP341J1, CYP341L1,
Mitochondrial	CYP314	1	CYP314A1
Mitochondrial	CYP301	2	CYP301A1, CYP301B1
Mitochondrial	CYP333	4	CYP333A10, CYP333A8, CYP333B19, CYP333B20

Table 1.2. Relative expression levels of P450s corrected for sequence length, expressed as #reads/kilobase.

P450	Clan	#Read/kb
CYP6AB43	3	228,391
CYP333B19	Mito	163,115
CYP6B54	3	108,907
CYP9A66	3	108,146
CYP6B55	3	76,050
CYP6B44v2	3	65,235
CYP6AB11v2	3	56,948
CYP321C1v2	3	49,418
CYP321C3	3	47,267
CYP6AE55	3	47,057
CYP6AN17	3	41,318
CYP9A67	3	35,480
CYP6AE54	3	31,633
CYP6AB42	3	30,781
CYP6AB41	3	19,400
CYP333A10	Mito	14,237
CYP6AB40	3	10,697
CYP306A1	2	10,635
CYP333B20	Mito	8,312
CYP6CV3	3	8,063
CYP354A11	3	7,988
CYP341M1	4	7,320
CYP324A10	3	6,615
CYP6AB44	6	6,245
CYP314A1	Mito	5,547
CYP9A64	3	5,230
CYP333A8	Mito	2,897

P450	Clan	#Read/kb
CYP4G89	4	2,389
CYP4M37	4	2,382
CYP4AU8	4	2,101
CYP341J1	4	1,757
CYP6AE56	3	1,537
CYP6AE53	3	1,170
CYP324A11	3	1,169
CYP4G88	4	1,131
CYP321C4	3	1,094
CYP4AU9	4	1,030
CYP9G13	3	941
CYP4L26	4	915
CYP341L1	4	828
CYP304F11	2	754
CYP305B1	2	666
CYP4M36	4	605
CYP367B8	4	561
CYP301A1	2	541
CYP18A1	2	478
CYP4M35	4	428
CYP301B1	2	328
CYP340AJ1	4	300
CYP6AE57	3	284
CYP338A1	Mito	175
CYP341K1	4	171
CYP304F12	2	82

References

- Ai J, Zhu Y, Duan J, Yu Q, Zhang G, Wan F, Xiang Z (2011) Genome-wide analysis of cytochrome P450 monooxygenase genes in the silkworm, *Bombyx mori*. *Gene* 480, 42–50.
- Campbell BC, Molyneux RJ and Schatzi TF (2003) Current research on reducing pre- and post-harvest aflatoxin contamination of US almond, pistachio, and walnut. *Journal of Toxicology: Toxin Reviews* 22, 225–266.
- Cariño F, Koener JF, Plapp FW, Feyereisen R (1994) Constitutive overexpression of the cytochrome P450 gene CYP6A1 in a house fly strain with metabolic resistance to insecticides. *Insect Biochemistry and Molecular Biology* 24, 411–8.
- Claudianos C, Ranson H, Johnson RM, Biswas S, Schuler MA, Berenbaum MR, Feyereisen R, Oakeshott JG (2006) A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee, *Insect Molecular Biology* 15, 615–636.
- Conesa A, Götz S, Garcia-Gomez JM, Terol J, Talon M, and Robles M (2005) Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21, 3674–3676.
- Connell JH (2001) Leading edge of plant protection for almond. *Hort Technology* 12, 619–622.
- Crawford JE, Guelbeogo WM, Sanou A, Traoré A, Vernick KD, Sagnon N, Lazzaro BP (2010) De novo transcriptome sequencing in *Anopheles funestus* using Illumina RNA-Seq technology. *PloS One* 5, e14202.
- Duke, JA. Dr. Duke's Phytochemical and Ethnobotanical Databases, <http://www.ars-grin.gov/duke/>.
- Duke, JA (1992) Handbook of Phytochemical Constituents of GRAS Herbs and Other Economic Plants: Herbal Reference Library. Boca Raton, FL: CRC Press, LCC.
- Feyereisen, R (2005) Insect cytochrome P450. In L. I. Gilbert, K. Iatrou, & S. S. Gill (Eds.), *Comprehensive Molecular Insect Science* (4, 1–77). Oxford, UK: Elsevier.
- Feyereisen, R (2006) Evolution of insect P450. *Biochemical Society Transactions* 34, 1252–1255.
- Feyereisen, R (2011) Insect CYP genes and P450 enzymes. *Insect Molecular Biology and Biochemistry* 450, 236–316. Oxford, UK: Elsevier.
- Feyereisen, R (2011) Arthropod CYPomes illustrate the tempo and mode in P450 evolution. *Biochimica et Biophysica Acta* 1814, 19–28.
- González AP, Barrios S, Vassena C, Cueto GM, Zerba E, Picollo M (2005) Increased Monooxygenase Activity associated with resistance to permethrin in *Pediculus humanus capitis* (Anoplura: Pediculidae) from Argentina. *Journal of Medical Entomology* 42, 342–345.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology* 29, 644–52.
- Janz, N (2011) Ehrlich and Raven revisited: mechanisms underlying codiversification of plants and enemies. *Annual Review of Ecology, Evolution, and Systematic* 42, 71–89.

- Karatolos N, Pauchet Y, Wilkinson P, Chauhan R, Denholm I, et al (2011) Pyrosequencing the transcriptome of the greenhouse whitefly, *Trialeurodes vaporariorum* reveals multiple transcripts encoding insecticide targets and detoxifying enzymes. *BMC Genomics* 12, 56.
- Kanehisa M, Goto S, Sato Y, Furumichi M, and Tanabe M (2012) KEGG for integration and interpretation of large-scale molecular datasets. *Nucleic Acids Research* 40, 109-114.
- Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ and Higgins DG (2007) ClustalW and ClustalX version 2. *Bioinformatics* 23, 2947-2948.
- Lee SH, Kang JS, Min JS, Yoon KS, Strycharz JP, Johnson R, Mittapalli O, Margam VM, Sun W, Li HM, Xie J, Wu J, Kirkness EF, Berenbaum MR, Pittendrigh BR, Clark JM (2010) Decreased detoxification genes and genome size make the human body louse an efficient model to study xenobiotic metabolism. *Insect Molecular Biology* 19, 599-615.
- Li H and Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler Transform. *Bioinformatics* 25, 1754-60.
- Ma W, Zhang Z, Peng C, Wang X, Li F, Lin Y (2012) Exploring the midgut transcriptome and brush border membrane vesicle proteome of the rice stem borer, *Chilo suppressalis* (Walker). *PloS one* 7, 1-7.
- Mao YB, Cai WJ, Wang JW, Hong GJ, Tao XY, et al (2007) Silencing a cotton bollworm P450 monooxygenase gene by plant-mediated RNAi impairs larval tolerance of gossypol. *Nature Biotechnology* 25, 1307-1313.
- Mao W, Rupasinghe SG, Zangerl AR, Berenbaum MR, Schuler MA (2007) Allelic variation in the *Depressaria pastinacella* CYP6AB3 protein enhances metabolism of plant allelochemicals by altering a proximal surface residue and potential interactions with cytochrome P450 reductase. *The Journal of Biological Chemistry* 282, 10544-52.
- Marck C (1988) 'DNA Strider': a 'C' program for the fast analysis of DNA and protein sequences on the Apple Macintosh family of computers, *Nucleic Acids Research* 5, 1829-36.
- Molyneux RJ, Mahoney N, Kim JH, Campbell BC (2007) Mycotoxins in edible tree nuts. *International Journal Food Microbiology* 119, 72-8.
- Niu G, Rupasinghe SG, Zangerl AR, Siegel JP, Schuler MA, Berenbaum MR (2011) A substrate-specific cytochrome P450 monooxygenase, CYP6AB11, from the polyphagous navel orangeworm (*Amyelois transitella*). *Insect Biochemistry and Molecular Biology* 41, 244-53.
- Niu G, Pollock HS, Lawrance A, Siegel JP, Berenbaum MR (2012) Effects of a naturally occurring and a synthetic synergist on toxicity of three insecticides and a phytochemical to navel orangeworm (Lepidoptera: Pyralidae) *Journal of Economic Entomology* 105, 410-417.
- Niu G, Wen Z, Rupasinghe SG, Zeng RS, Berenbaum MR, Schuler MA (2008) Aflatoxin B1 detoxification by CYP321A1 in *Helicoverpa zea*. *Archives of Insect Biochemistry and Physiology* 69, 32-45.
- Niu G (2010) Toxicity of mycotoxins to insects and underlying molecular and biochemical mechanisms. PhD dissertation. University of Illinois, Urbana Champaign.
- Nylin S, Janz N (2009) Butterfly host plant range: an example of plasticity as a promoter of speciation? *Evolutionary Ecology* 23, 137-46.

- Pauchet Y, Wilkinson P, Vogel H, Nelson DR, Reynolds SE, Heckel DG, ffrench-Constant RH (2010) Pyrosequencing the *Manduca sexta* larval midgut transcriptome: messages for digestion, detoxification and defence. *Insect Molecular Biology* 19, 61–75.
- Pavlidis N, Dermauw W, Rombauts S, Chrisargiris A, Van Leeuwen T, Vontas J (2013) Analysis of the olive fruit fly *Bactrocera oleae* transcriptome and phylogenetic classification of the major detoxification gene families. *PLoS One* 8, e66533.
- Petryk A, Warren JT, Marqués G, Jarcho MP, Gilbert LI, Kahler J, Parvy JP, Li Y, Dauphin-Villemant C, O'Connor MB (2003) Shade is the *Drosophila* P450 enzyme that mediates the hydroxylation of ecdysone to the steroid insect molting hormone 20-hydroxyecdysone. *Proceedings of the National Academy of Sciences* 100, 13773–8.
- Petersen R, Zangerl R, Berenbaum MR, Schuler M (2001) Expression of CYP6B1 and CYP6B3 cytochrome P450 monooxygenases and furanocoumarin metabolism in different tissues of *Papilio polyxenes* (Lepidoptera: Papilionidae). *Insect Biochemistry and Molecular Biology* 31, 679–90.
- Rewitz KF, O'Connor MB, Gilbert LI (2007) Molecular evolution of the insect Halloween family of cytochrome P450s: Phylogeny, gene organization and functional conservation. *Insect Biochemistry and Molecular Biology* 37, 741–753.
- Rupasinghe SG, Wen Z, Chiu TL, Schuler MA (2007) *Helicoverpa zea* CYP6B8 and CYP321A1: different molecular solutions to the problem of metabolizing plant toxins and insecticides. *Protein Engineering, Design & Selection* 20, 615–24.
- Seigler, David S (1998) *Plant Secondary Metabolism*, (134). Norwell, MA: Kluwer Academic Publishers.
- Siegel JP, Kuenen LPS, Ledbetter C (2010) Variable development rate and survival of navel orangeworm (*Amyelois transitella*, Lepidoptera: Pyralidae) on wheat bran diet and almonds. *Journal of Economic Entomology* 103, 1250–1257.
- Sparks ME, Blackburn MB, Kuhar D, Gundersen-Rindal DE (2013) Transcriptome of the *Lymantria dispar* (gypsy moth) larval midgut in response to infection by *Bacillus thuringiensis*. *PloS One* 8, e61190.
- Sztal T, Chung H, Berger S, Currie PD, Batterham P, et al (2012) A cytochrome P450 conserved in insects is involved in cuticle formation. *PLoS One* 7, e36544.
- Le T, Min J, Lee S, Kim Y (2011) Gene Expressions of the Dhb, Vtg, Arnt, CYP4, CYP314 in *Daphnia magna* induced by toxicity of glyphosate and methidathion pesticides. In Dr. Margarita Stoytcheva (Ed.), *Pesticides in the modern world - pests control and pesticides exposure and toxicity assessment*. InTech. ISBN: 978-953-307-457-3.
- Veitch NC, Elliott PC, Kite GC, Lewis GP (2010) Flavonoid glycosides of the black locust tree, *Robinia pseudoacacia* (Leguminosae). *Phytochemistry* 71, 479–86.
- Wade WH (1961) Biology of the navel orangeworm, *Paramyelois transitella* (Walker), on almonds and walnuts in northern California. *Hilgardia* 31, 129–171.
- Wang, XW, Luan JB, Li JM, Bao YY, Zhang CX, Liu SS (2010) De novo characterization of a whitefly transcriptome and analysis of its gene expression during development. *BMC genomics* 11, 400.
- Wang Z, Gerstein M, Snyder M (2009) RNA-Seq: a revolutionary tool for transcriptomics. *Nature Reviews Genetics* 10, 57–63.

- Xie W, Lei Y, Fu W, Yang Z, Zhu X, Guo Z, Wu Q, Wang S, Xu B, Zhou X, Zhang Y (2012) Tissue-specific transcriptome profiling of *Plutella xylostella* third instar larval midgut. *International Journal of Biological Sciences* 8, 1142-1155.
- Yang Y, Chen S, Wu S, Yue L, Wu, Y (2006) Constitutive overexpression of multiple cytochrome P450 genes associated with pyrethroid resistance in *Helicoverpa armigera*. *Journal of Economic Entomology* 99, 1784-1789.
- Zhang YE, Ma HJ, Feng DD, Lai XF, Chen ZM, Xu MY, Yu QY, Zhang Z (2012) Induction of detoxification enzymes by quercetin in the silkworm. *Journal of Economic Entomology* 105, 1034-42.

II. INCREASED PESTICIDE TOLERANCE OF NAVEL ORANGEWORMS WITH DIETARY EXPOSURE TO PLANT COMPOUNDS

Introduction

The navel orangeworm *Amyelois transitella* (Walker) (Lepidoptera: Pyralidae) is currently the primary pest affecting North American almond production (Siegel 2010). Larvae cause damage by entering split hulls, where they consume nuts and increase the chance of infection by mycotoxin-producing *Aspergillus* spp. Contamination by aflatoxin and ochratoxins is a major concern of the almond industry as most markets strictly regulate mycotoxin content (Cleveland et al., 2003). Pyrethroid pesticides are the chief chemical control method employed during the growing season (Leal et al., 2009).

Navel orangeworms are broadly polyphagous and are exposed to a wide variety of phytochemicals when they feed on figs (*Ficus carica* L) and pomegranates (*Punica granatum*) grown near almond orchards. In another polyphagous lepidopteran (*Helicoverpa zea*), dietary exposure to phytochemicals induces cytochrome P450 activity and increases survivorship after exposure to the pyrethroid α -cypermethrin (Wen et al. 2009). To assess the effect of dietary phytochemicals on pesticide tolerance, navel orangeworms were placed on a diet containing α -cypermethrin and two compounds encountered in a typical almond orchard: chlorogenic acid, present in almond hulls, and quercetin, present in nutmeats (Chen et al. 2007; Takeoka 2003).

Materials and Methods

Insect Colony:

A laboratory colony of *A. transitella* was established and repopulated with specimens from Dr. Joel Siegel at the USDA-ARS in Parlier, CA. Larvae were reared on standard wheat bran diet at the University of Illinois Urbana-Champaign at $28 \pm 4^\circ\text{C}$ and photoperiod of 16:8 (L:D) h. (Siegel et al., 2010).

Pesticide and phytochemical treatment:

Newly emerged first instar *A. transitella* larvae were transferred to 1-oz (29.6 ml) plastic deli cups that contained 5 g of an agar-based semi-defined artificial diet with pesticide and phytochemical treatments (Niu et al., 2012). For the toxicity assays, artificial diet was supplemented with 5 mg/g of quercetin, 2 mg/g of chlorogenic acid, 50 ng/g α -cypermethrin dissolved in methanol, or a methanol control to establish mortality rates on these substances alone. In previous navel orangeworm bioassays, these concentrations caused less than 50% mortality after 24 hours (Niu, 2010). Treatments of α -cypermethrin combined with either quercetin or chlorogenic acid were monitored for synergistic effects. Mortality was measured every 6 hours over the next 48 hours. Bioassays were replicated three times using the same α -cypermethrin stock solution. Mortality data from all three trials were pooled and analyzed with one-way analysis of variance (ANOVA) and Tukey's LSD test, using SPSS 21 (IBM Corp, Armonk, NY).

Results

Mortality rates on the control, quercetin and chlorogenic acid diets did not differ significantly. The pesticide α -cypermethrin alone is an effective mortality agent at a concentration of 50 ng/g. Dietary exposure to the phytochemicals quercetin and chlorogenic acid significantly decreased mortality of *A. transitella* neonates exposed to α -cypermethrin (Fig. 2.1). Analysis of the mean number of survivors per treatment showed a significant difference ($p < 0.05$) between three groups of treatments after 48 hours (Fig. 2.2). There was no significant difference between the mean numbers of survivors on the control, quercetin, or chlorogenic acid diets (Group A). The mean number of survivors on the cypermethrin treatment (Group B) is significantly different from the number of survivors in other treatments. Mortality rates in treatments with α -cypermethrin and chlorogenic acid or quercetin (Group C) were also significantly different from other treatment groups.

Discussion

In *H. zea*, quercetin, chlorogenic acid and α -cypermethrin are detoxified at least in part by the same cytochrome P450 enzyme, CYP6B8 (Li et al., 2004). In the navel orangeworm, four P450 enzymes in the same CYP6B subfamily have been sequenced from the midgut; metabolic characterization of these enzymes will determine if one or more is responsible for metabolism of α -cypermethrin, quercetin and chlorogenic acid. Induction of P450 enzymes after consumption of quercetin and chlorogenic acid might increase metabolism of α -cypermethrin, reducing the efficacy of this pesticide. A similar process occurs in the honey bee *Apis mellifera*, where dietary exposure to quercetin reduces the toxicity of the acaricide tau-fluvalinate (Johnson et al., 2012). Decreased efficacy of pesticide treatment after exposure to plant chemicals encountered in the *A. transitella* host range is a concern to the almond-producing industry. Identification of the mechanism underlying the synergistic interaction between quercetin, chlorogenic acid and α -cypermethrin can inform management strategies of the navel orangeworm that involve use of pesticides.

Figures

Figure 2.1. First instar navel orangeworm mortality expressed as a proportion at multiple time points following dietary exposure to 50 ng/g α -cypermethrin, 2 mg/g chlorogenic acid, 5 mg/g quercetin, and controls. Mortality was measured at 6 hour intervals for 48 hours. Error bars represent mean standard error of raw data.

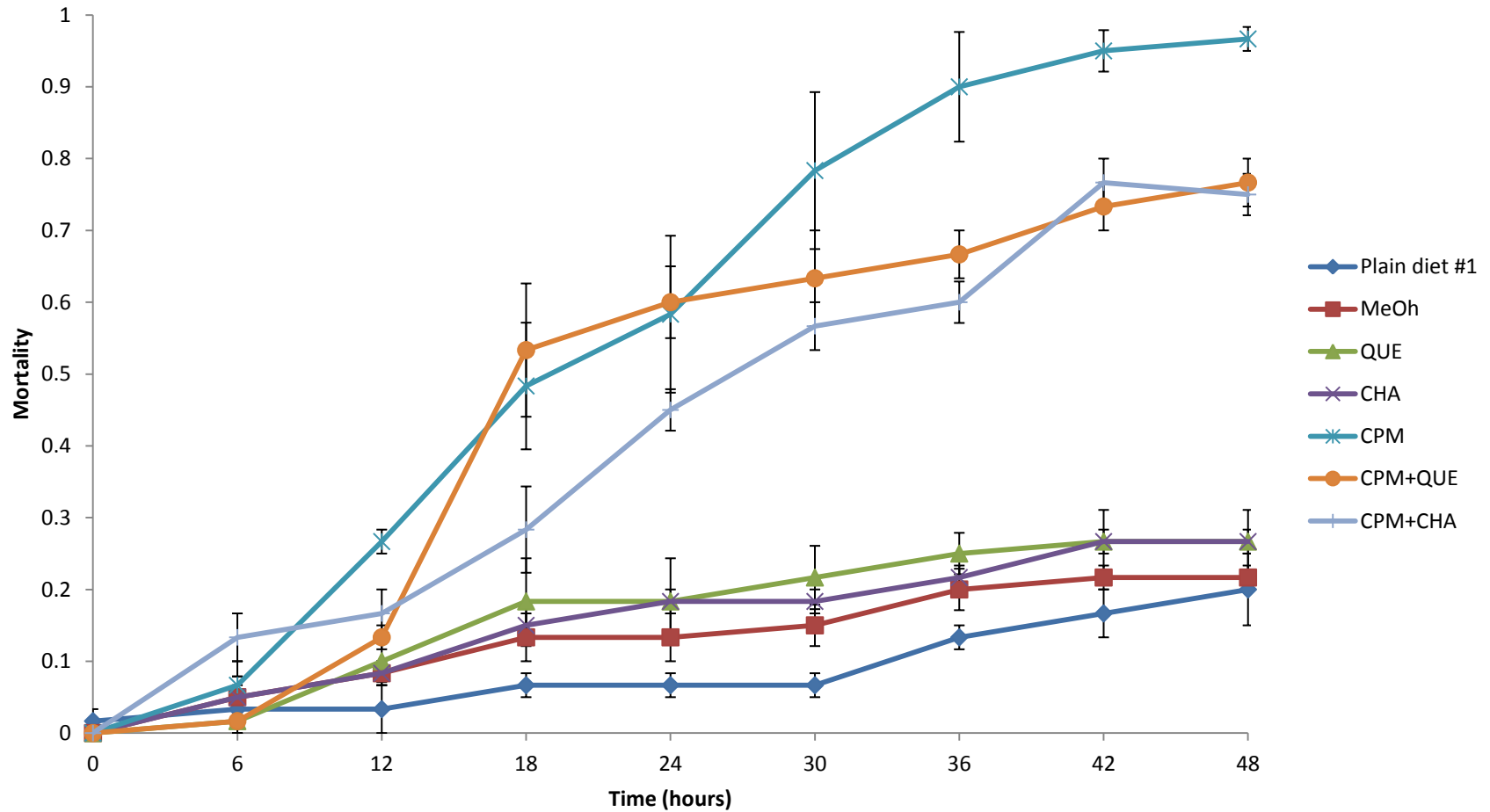
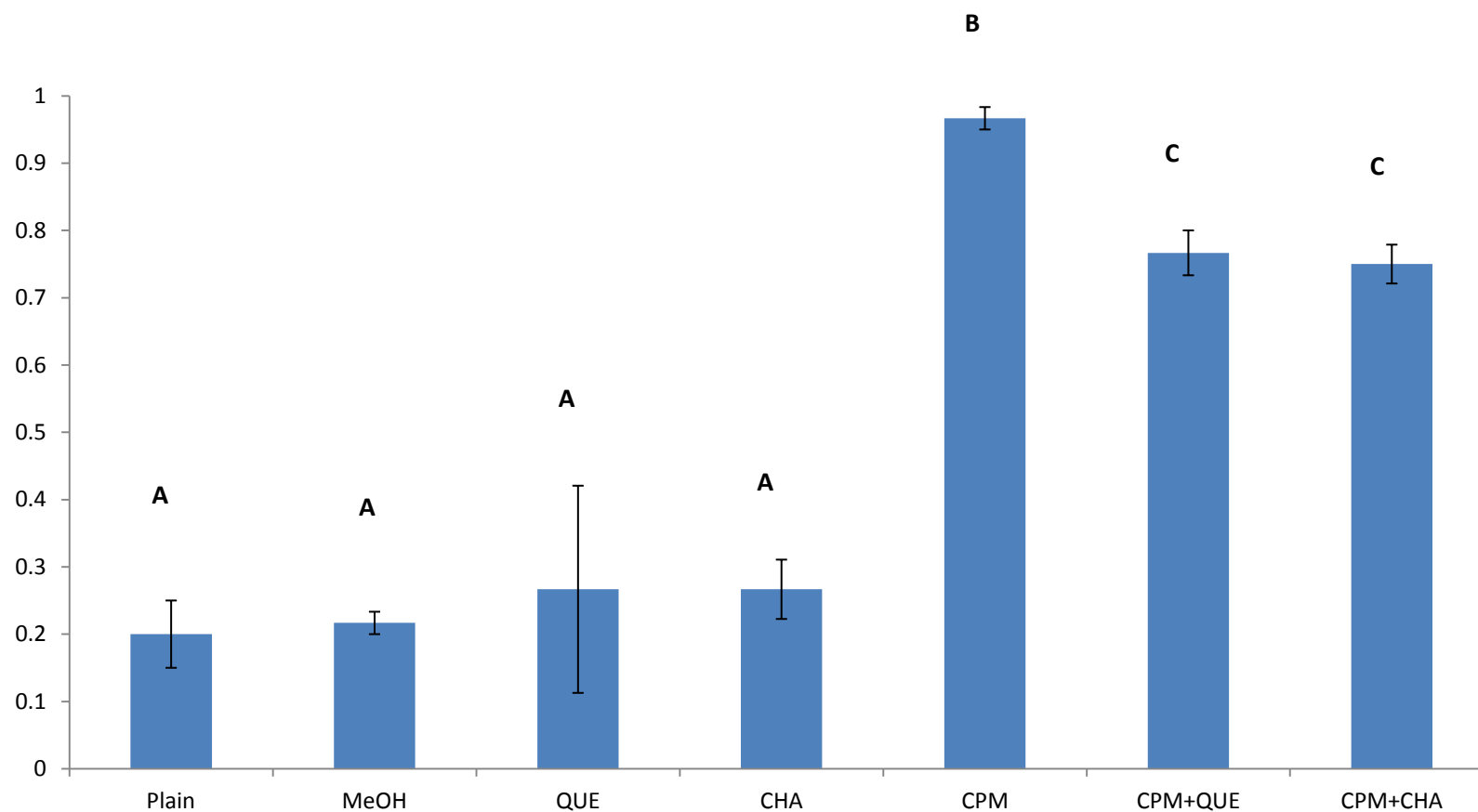


Figure 2.2. Synergistic effects of 2 mg/g chlorogenic acid (CHA), 5 mg/g quercetin (QUE) on 50 ng/g α -cypermethrin pesticide treatment. Mortality of first instar larvae represented as a proportion after 48 hours of dietary exposure. Bioassays were replicated three times; the mean mortality was compared between treatments with post-hoc LSD tests of one-way analysis of variance (ANOVA). Letters A, B, and C represent significantly different groups ($p < 0.05$).



References

- Chen CY, Milbury PE, Chung SK, Blumberg J (2007) Effect of almond skin polyphenolics and quercetin on human LDL and apolipoprotein B-100 oxidation and conformation. *The Journal of Nutritional Biochemistry* 12, 785-94.
- Johnson RM, Mao W, Pollock HS, Niu G, Schuler MA, Berenbaum MR (2012) Ecologically appropriate xenobiotics induce cytochrome P450s in *Apis mellifera*. *PloS One*, 7, e31051.
- Leal WS, Ishida Y, Pelletier J, Xu W, Rayo J, Xu X, Ames JB (2009) Olfactory proteins mediating chemical communication in the navel orangeworm moth, *Amyelois transitella*. *PLoS ONE* 4, e7235.
- Li X, Baudry J, Berenbaum MR, Schuler MA (2004) Structural and functional divergence of insect CYP6B proteins: From specialist to generalist cytochrome P450. *Proceedings of the National Academy of Sciences of the United States of America* 101, 2939-44.
- Niu G, Siegel J, Schuler MA, Berenbaum MR (2009) Comparative toxicity of mycotoxins to navel orangeworm (*Amyelois transitella*) and corn earworm (*Helicoverpa zea*). *Journal of Chemical Ecology* 35, 951-957.
- Niu G (2010) Toxicity of mycotoxins to insects and underlying molecular and biochemical mechanisms. PhD dissertation. University of Illinois, Urbana Champaign.
- Niu G, Pollock HS, Lawrance A, Siegel JP, Berenbaum MR (2012) Effects of a naturally occurring and a synthetic synergist on toxicity of three insecticides and a phytochemical to navel orangeworm (Lepidoptera: Pyralidae) *Journal of Economic Entomology* 105, 410-417.
- Siegel JP, Kuenen LPS, Ledbetter C (2010) Variable development rate and survival of navel orangeworm (*Amyelois transitella*, Lepidoptera: Pyralidae) on wheat bran diet and almonds. *Journal of Economic Entomology* 103, 1250-1257.
- Takeoka GR, Dao LT (2003) Antioxidant constituents of almond [*Prunus dulcis* (Mill.) D.A. Webb]. hulls. *Journal of Agricultural and Food Chemistry* 51, 496-50.
- Wen Z, Zeng RS, Niu G, Berenbaum MR, Schuler MA (2009) Ecological significance of induction of broad-substrate cytochrome P450s by natural and synthetic inducers in *Helicoverpa zea*. *Journal of Chemical Ecology* 35, 183-189.

Appendix

Appendix 1. Full-length amino acid sequences of 61 P450s transcribed in the *A. transitella* midgut.

>CYP4AU8

MIILGVLVLVVGAYWLYFSIKTHRMEKLTASLPSLPALPLVGNAMMFIGNTEKILHNLEY
MAEIAFKNKGCAKVWLGPPLYIAIGNPEDAQVVLNCLDKDGVYRFLRAWLGHGLFV
APVGLWKTHRKVLLPVFHNKVVEEYLGVIGEQADVLVGRLADHVDKEEFDILKYITAC
TLDIVFETAMGERMDVQHTPDTPYLRARHTVIAILNMRIKFWVLQPDCLFNLT SYAKQQ
KDNIDLTHRFTDEVVKKKRLEYESKKRKDNNESEGKLQAVLDLLFGREIEFTNEQLREHI
DSITVAGNDTTALVIAATLVLLGVHQGAQERAFQEQESIFGKSSRGANKEDLQQMHYLE
RVIKESMRLYTVVPIIARNVHKD TYLPKCGVTPAGAGAVIGSF A I H R A K D V W G D D A N T
FDPDRFLPERSVGRHPAAFLPFSFGSRNCIGRNF G M L I M K S I I S R V L R D Y I I G A K P I D R L K I E
VLLFPTNGHQIKLTRRSA

>CYP4AU9

MIVFLLVFILAVFFYVKYRGGGRKQLYELSDILPGPGCVPIVGN SHWFIGGPERILQNIKILA
NRAEKCPSGICKIWIGPSLYVVVTNPQDVQIVLDSCLEKDYTYKFLRAWLGNGLFVAPL
DLWKSHRRVILPIFQNR LIEDFIENFGKQGSILVQRLNEVVGKSEFDVFKYITSCMLDIVFE
TAMGERMDVQHNP DTRYLRARNYVMSVINMRLFKPWMQPDALFHLTKYSKLQKESID
ITHKFTEE VVTRKRKAYENGEIHKGCKDLELLLSRGTKFTDEELREHIDSITIAGNDTTA
LVISYVLQILGHHQKEQQLVYEELQQIFGSSSRLPNKEDLMKMEYMERVIKETMRLFTV
VPIIARKTLKEVKLSSITLPAGISCAVVPFAMHRSKKLWGPDVDDFNPD RFLPENCANRH
PCAFIPFSYGARNCIGRHFGMLAMKSILANILRKYKINSSACGNL KIEILLFPVSGHLISLE
KRETPVDTSVNSSMKSMKIGPLVS

>CYP4G88

MSYATENVVQSTTWAATSLFYMLLV PALVLWYAYWRMSRRHLYELAEKLDGPPGLPL
LGNALEFTGGSADIFRKL IQRSEPYDKEAVVKVWIGPRLLVFLYDPRDVELILSSHVHIDK
ADEYRFFK P W L G N G L L I S T G Q K W R S H R K L I A P T F H L N V L K S F I D L F N A N S R A V V D K M K K
EAGVFDCHDYMSECTVEILLETAMGVSKSTQDQSGFEYAMAVMKMCDILHLRHTKIWL
RPELIFNFTKYAKNQTQLLDIIHGLTKKVIKRKKEEFKSGKKPSVLSETKDVPQNSETTTK
TTSVEGLSFGQSSGLKDDLDVEDNDVGEKKRLAFLDLLLESSQSGVVISDEEIKEQVDTI
MFEGHDTTAAGSSFFLSMMGIHQDIQDKVIEELDQIFGDSDRPATFQDTLEMKYLERCL
METLRMYPPVPIIARHLKQDVVLPSSGKKVPAGTTVIIATIKLHRRPDIYPNPDKFDPDNF
LPERSANRHYAFVPFSAGPRSCVGRKYAMLK L K I I L S T I L R N F R V H S D L K E S D F Q L Q A D
IILKRAEGFKVRLTPRKRTAKAQ

>CYP4G89

MSVIEEETGYHYGYSRLIFYPLLLLATGLWLFHRWQQQSKIFILGNKIQGPSTIPFLGNTL
LALGKTPSDILELSVKLSKKYTKIARIWFGNKL VVFLTEPDDVEVILNSQVHIDKSAEYAF
FKPWLGEGLLISSGHKWRSHRKMIAPTFHINILKS FVG VFNQNSLVVEK L K P E V G K V F
DVHDYMSGVTVDILLETAMGINRKTQDESGFDYAMAVMKMCDIIHQ RHYKLWLRFDFI
FKLTSIFKQKIKLLKIIHGLTNKVIKQKKQVY MENKAKGIIPNTIEELIRSNNENDLVDDTK
TLADMVFKGYRDDLD FNDENDVGAKKRLAFLDLMIESAQNGTNNITDHEIKEEVDTIM

FEGHDTTAAGSSFVLCLLGIHQDIQTKVYDELYEIFGSDSRLATFNDTLKMKYLERVILE
SLRLYPPVPIIARKLNQDVKIVTNGHILPARSTVVIGTYKIHRSKYYYKDPEKFNPDNFLP
ENTANRHYYSYIPFSAGPRSCVGRKYAMLKLKVLLSTILRNFKTTSVVNEQGFKLQADII
LKREDGFRLMIEPRKRVPSKTIN

>CYP4L26

MITVIFISTAICLSLISWLMLLWDYYKSNIPGPLPLPILGNAHMMMAVDAQKLLIVMKELAE
KYGSVNSVFLLSNRYMIISNPNFAEALLSSTEFIDKGRSYNFLKPWLGYGLLTATGQRW
KTHRKFLTPTLHFNILQNFLPVFCKNEKILVGKLRKLADGKSIDVFPIIALAALDNVTESI
MGVSMKIQTDTESKYVKAIECTSNIMSSRMNRNSFLAIDAVFNLTLLKKRQDEALKILHSQ
TKQVINARRKEIKKSNTETITDPSKIGIKNKHAFLDLLLLAEINGNKIDDEHVREEVDTFM
FEGHDTTTSGLAFSLYCLSQYKDIQEKVYDEQKSIFGDDLDERDPSYSELQQMKYLEMVL
KEALRLYPSVPLIERKIIKRTDITGFHIPVDTSVIINIFQMQRSPFEDPMVFKPERFDSSV
HRNAFSWIPFSAGPRNCIGQKFAMMELKIIVAGIIRRFHLMPSSEPELVPILRSRNGV
YVKFLPRD

>CYP4M35

MIAYVLMTLVVLCCVIHLLINYNKKARMMRRIPGPKETFLVGNAPETIVSPVELFALGRK
YANTWKNLYRFWAYPLSAVNIYNPEDIEVIISTTKFNEKSRVYYFLKPWLQDGLLLSNG
SKWQQRRKILTPAFHFNILREFYVVLEENSDRLVAKLAQHTGKELDLVPVMSEFTLNSIC
ETAMGTQLSEETSVSAKRYKEAILSLINVVFERFIKMPYYFDNIFNLSPLGKEQQKFITTIR
EFTKKVIHDRKLYWELNKFDLNDFTESGSDDVYMNKRRKKTAMLDLLIAAQKEGMIDE
NGIQEEVDTFMFEGHDTTASGLTFCLMLLANNRDKQDKIVQELHEIFGKSLRSATIDDL
KMKYLECCIKESLRLYPPVHFISRNLNQTVELSNYTVPPGTFCHIHVYDLHRREDLYKDP
LVFDPDRFLPENCVGRHPYAYIPFSAGPRNCIGQKFAMLEMKLVSASVLRREFQLFPVTT
DDIRISADLVLRNSGPVSVRFEKRNKNNNYLL

>CYP4M36

MLFVLFAFALFLCFIHYLINENEKARILKKIPGPKDVFLLGNALEVMLTPVELFKLGRRN
AQKYNIGYRFWAFPRSAVNIYNPEDIEVVIATSKHNSKSPLYNFLKPWLWDGLLLSSGS
KWHQRRKILTPAFHFNILRQFHIILEENSNRLIKSLERTPGKAIDIVSLLSEYTLNSICETAM
GTQLSEETTGSASYKAAIYGLGNLLVQRISRIYYHTDFVFNLSLGRQQKEYLHTVREF
TEKVIKDRKNYIETNRISSSDDIHTDDTFMTKKKKKAAMLDDLILAKKEGLINNVGIQEE
VETFMFEGHDTTAAGLTFSLVLLAEKDIQNKIYLEQREIFGDSSRAATSEDLKMRYLE
CCIKECLRLYPPVPVISRALSETVKLSNYTIPVGSYCNIIHYDLHRRETFLKNPLKFDPDFR
LPENSVGRHPYSYIPFSAGPRNCIGQKFAMMEMKSALSAILRHFELEPVTRQSDLEIITDL
VLRNSQPVYVTFNKRVS

>CYP4M37

MELFLWLIVILFCLIFLARLVKHEQFDKLPGLPGLYIFGNALDFFLMDTVRLFYHFRTLGS
RYNKLFLKLMGPMKILILYNPEDVETLISGTKYNQKGFVYSFIRPWLKEGLLLSDGSKW
QQRKILTPAFHFNILRNFNILQENSWKFVERLRHEENESKTEVSSMITDFTLNSICETS
MGTQLDAETSEFGKTYKDAIHKLGCVVYRGQRCWYFEIVFNRTVGRKFNELLNLL
TSFRRKVVDKRRNDKEFKNLVFDKDDNVQYEDSLESKKRLAMLDDLHAEENGTTIDA
DGIGEEVDTFMFEGHDTTSSALQFAFMLIANNEECQEKIFEECEKIFGSSDRAADINDLSE
MKYLGCCIKETMRLYPPVPFIVRRILQPLQLKDYLPTETECVLIYDLHRRSDQFPEPLE

FRPERFLAEPTWHPYAYIPFSAGPRNCIGQKFAMMEMKLVLSVIRKYRMLPVTRPQDL
VFVTDFILRTTAPIYVKFQKR

>CYP6AB11v2

MIPIAVIVICLLLYYYGTRNFKYWQKRGIIHDKPIPFNGNIDGYLFRKSVTQIATDMYR
KYPNEKVVGIFYRANIPELVVRDPETIKRVLVTD FEHL YPRGINHKEVVEPLMRNLFFAD
GDLWRLLRQRMTPAFTSGKLKAMFPLIIDCADRLQKRALT VSAAGKKLDARDLMARYT
TDFIGACGFGVNADSLGDEDSAFRKLGAQIFQPRVQDLIVAILKEVFPDTFKHLKYL SRL
EDDFKKFVSDVLHLRNYKPSGRNDFIDLMLECKQKGTIVGESIERRDKEGRPEKATVEL
DDTLIAAQVFVFFAAGFETSSSATS YTLHELAFNPEIQERVQKEIDTVLAKHDNKLSYDA
VKEMTYLEWTFKEAMRMLPSLGFLIRQSVRPYTFPELGLSIDADV GIMIP LQALHTDPEY
FDNPMEFRPERFDPENFTRKQRDIYLPFGTGPRACIGERLGLMQSLAGLA AVL SKFSVSP
AADTKRWPEIDPRSDIVQTIVGGLPLQFNERKRN VVNFN

>CYP6AB39

MILPIVVIIAILLYYYSTRNFKYWQKRGVKHDKPIPFNGNIEVYVLRKCMSDA AVEAY
CKYPGEKIVGYRANTPELVIRDPDIMKRILVTD FEYFY ARGINHKEVVEPLMRNLFFAD
GDLWRLLRQRMTPAFTSGKLKAMFPLIIDCAERLQKRALT VVSAAGKKLDARDLMARYT
TDFIGACGFGVDANSLGDEESA FRKLGAQIFQPTVKDMLVATLKEVFPDACKNLKYLAR
FEDDLKKFVSDVLRRLRNYKPSGRNDFIDLMLECKQRGII VGESIERRAKDGT PETASIELD
DTLIAAQVFIFFAAGFETSSSVTS YTLHELAFNPEIQEKVQKEIDTVLAKYDNKLT YDAIK
EMTYLGWTMKEAMRMLPSVGFLVRQSAKP YTFPEL DISIDANVGIIPLQAMHNDPEYF
ENPKEFRPERFDPENLTRKQRDVFLPFGAGPRACIGERLGLMQSLTGLAAVLSKFSVSPA
PDTIRWPKIDPKSDVVQIIQGGLPLMFNERKITS

>CYP6AB40

MIVILALVLLLATLYVYGTRNH DYWARKGVKHDKPLPLFGTEIKRFLQQKSLTQMAVD
NYWKYPDERFVGAYRGTVPELVIRDPDLIKQIIVTD FASFYSRGLSPEHG VIEPLLKNLFF
ADGDLWRLLRQRMTPAFSSGKLKSMFPLIVDCADRLQKRALKAAASGDALDARDLMA
RYTTDFIGACGFGLNVD SLGDDESEFRKLGNKIFTITPSRALTSIFKILFPDIFKNFKFLGQE
LED SILAVVNMIQKERGNKPCGRNDFIDMMELKQK GKIVVESMERFNADGSP EMAEL
ELDDILLAAQAFIFFAAGFETSSSATSFTLHQLAYNQSVQKKVQEDIDRVLAKYDNKLSY
DAIKEMTYLQWTFQEGMRIFPSSG LLMRKCTQKYTIPGTDVTIDKGVKVTIPIQAIHNDP
KYFDNPTEFRPERFSPEEIAKRHKFVYLPFGDGPRSCIGGRLGQM QSVAGLAALLARATV
APAPSSKRELQQDPTTSIVQNVIGGIPLQFTQRHI

>CYP6AB41

MLIVFFIIIVLVLYWYGTRTFHYWKEKGVKHDEPLPFLGNNFKQFIQKASIGMMANEMY
KKYPEEKVVGFYRATAPELIIRDPEIAKRILVSDFSSFHARGLNPHRTVVEPLLQNLFFGE
GDLWRLLRQRFTFAFSSGKLKAMFPLITKRAEKLKIITEEITGRDFYDARELMARYTTDCI
GECGFGINMDSLSEENNQFRRLGKRIFQRNLRDIVCSALKFIFPELAKNL TLLTPELEKSM
NKIVHAVLKERDNKPSGRNDFIDLLLEV KQKGNMIGESIEKKNADGVPEIVEMELDDML
LIAQVFLFFGAGFETSSSASSYTLHQLAFNPDCQVKVQEEIDRVLEKYDNKLT YEAVKD
MKYLEMAFQEAMRMYPAPGYLIRMCTVPEYTFPELNITVHDGVKMIPIAAIQTDETYFS
NPEQFDPERFRDGA KDNSKKFVYLPFGEGPRACVGARLGMMKSMAGLA AVLQNYSVE
PAKCSKLKPD PDPAGIVTEGFVEGLPLKFKKRLR

>CYP6AB42

MNGAGAAAVTPLRRPPRTHLTDTNYWLLLLHVLRAIMITVIVAVIALIAAYLFGTRYHD
YWLKKGIKHKDKPIPFPGTNFSVTFGLKSVTQVFAEAYRKYPTEKAVGLYFESRPMLVVR
DLEIAKRVLTTFDFANYYSRGLNHNKDVIEPMAKNLFFVDGDTWRLLRQRMTPAFTSGK
LKAMFPLIVERAEKLQTRALQAAAAGKPLDARDLMARYTTDFIGACGFGLDADSLNEE
DSAFRKLGVKIFKPDFKDIIIVIVVKQLFPELFKHAKVLGKIELEINALVHEILTRRNYEPSG
RNDFIDLLLECRKKGLMVGESIEKRKPDGTPLTVSMELDEALIAAQVFVFFAAGFETSSS
ATSYTLHELAYNPQVQKKVQDEVDRVLAKHNNKLSYDAIREMTYLECTFKEGMRIFPS
LGFLRECMVKTEIPELGITLDPRI RVAIPLQAIQTD PQYFDNPWEFRPERFMEENSSANK
FDYLPFGIGPRACIGERLGLMQSLAGLA AVLAKFTVKPAPASQRRPPVDPRGGIVQIVKD
GIPLIFEERTARVH

>CYP6AB43

MIFLTILLVGIVLLYLYNIRNFDYWKKRGVTYVKPMLFVGSTVDVYLQKKSVTQYASEV
YFKFPNEKMVG VFRSTNPGLIVRDPEI IKRILITDFYFYKRGIMPGKTEPILRNLFFAEG
DLWRLLRQRLTPAFTSGKLKAMFPLITERAEKL RTRTLNAALEGRAIDARDLMARYTTD
FIGACGFGLDADSLNDENSAFRKLGVDFHFGLRENVKLLLKDIFPELTDNITFFGRLEDK
IIDLVIDIQKQRNYQPSKRNDFIDLMLECKTKGKMVGDSIENFDKSGTPKEAKVDLDDVI
MAAQVFVFFAAGFETSSSATSFTLHQLAYYPEVQKKVQQDIDKVLAKYDDKLCYDAIK
EMTYLEWTFKEALRIFPSLGYLMRTCVPYTFPELNV TIDEDVRILIPVQALHNDPQYWT
KPEEFRPERFRPDEFNDVT KSIYLPFGEGPRSCIGERLGIMQSLAGLA AVL SRFTVEPAPN
TKRHPSVETRADV VQAIDGGLPLMFRERKK

>CYP6AB44

MIALILFLIFIAVVYLYNTRNFKYWEKKGVKHDKPVVLFGNNIKGYLMQQTKSQAAAE
MYRKYPNEKVIGYYRSARPELIVKDPELIKKIMVTDFQHFFYRGFNPNRKAVEPVFHSFLF
AAEGDLWKMLRTRMTAFTSAKLKGMFPLIIERAEKLQARTLEAAKAGKSLDSRELMA
RYTTDFIGACAFGLDSDSLNEEDSNFRKLKGKIFDAGFKDFIINLLKEMWPEAFKNLKYF
AKIEDGARQLAYEVQKARNFQPSGRNDFLDMMEIKGKGAVEFESIEKVL PNGDSEKV
VMNFTDDIFMSQLLVFFAAGFETSSNSTSM TLHLLAYNPDEQKKVHEDIDKALAKHNN
KLSYDAIRDMPYLEMAFKEAMRILPALGYVSRLCSRDTFEELGLTIDKEVRVVIPIQAL
HMDPKYWDEPEEFKPERFLPENAEKTKDVYLPFGIGPRSCIGARLGLMQSLAGLAAILSK
FEVKPAPESKRHPEIDPLSEIVQGVIGGLPLIFEERK

>CYP6AE53

MIGAALTLLILIFIVLYNVAKWKFGYWKKRGIPYLEPVPVLGNYGSYIIQREYLGQLTQR
LCDQFSKSPYFGVFYGTPTLIAHEPEI IKIIMTKDFYFSSREISDYTYKEVMTQNLFFTY
GDRWKVLRQNLTPLFSSSKMKNMFYLVDKCAHVFEDMLS NELKADNVQEV RPLMARF
TMDVIGSCAFGIDTNTMARNSETNPFTKIGCLIFSETFYRGFKIAARAIWPALFYGLKMKS
FPTEADEFFNKLMMGVFAGRKYKPTNRNDFVDLILNWKEKNYITGDSLQNMKGDDKK
VNLEVDDEILVGQCLLFFAAGFETSSTTLQYTYELAKNPEAQEKVINEIDEYLQRRDNK
LVYECFTELPYTEACFDEALRLYPVLGVL TREVVEDYTLPCGAVIGKGVRVHIPVYHLH
HNPEYFPNPEVFRPERFYGEEKRNVKPFTYMPFGEGPRLCIGMRFAKMQMFSGIITLLKK
YQVELAADTPRKLNFKSTTLVTQPKELIKLFVEREGWEQRMFNK

>CYP6AE54

MIAEVLTLILVLAITYHVSRRKFDYWKKRKIPYLKPVPLGNYGPYIIQKKYIGQTAAEL
CEKFPDTPYFGAFYGTPTLIVKDPEMVKVMKMDYFFFNSREISNYTHKEVITQNLFFT
YGDRWKVLRQNLTPIFSSAKMKNMFYLIDKSAHVFEGLDRELKMNDVQAVRTLISRY
TMDSIGSCVFGVDTETMTEDPSKNPFFQVGSIFAESYLRSLMTVSRAIWPVAVFYGLNM
KSFPKEMNDDFFDKLMKGVFEQRQYKPTNRNDFVDLILNWQKGHYITGDSMQNMKGNE
KKVKLEVDDELLVGQSILLFGAGYETSATTLQYTLYELAKNQEAQAKVHTEIDEYLLRH
NNALVYECVTELPTYEACVDEALRMYPVLGVLTVREVVVEEYTLPCGAVIEKGVRVHVPV
YYLHHNPDDFFPEPEKFRPERFYGEEKRNIVPYTYMPFGGPRCLCLGIRFAKMQUIPAITLL
KKYRVELTPDTPRKLEFTPTTIVTQPKELVKLKFVERDGWQQRLYNK

>CYP6AE55

MIVEITLSLLIIFLIIVYNISKWRFGYWKKRNVPFLEPVPVLGNYGPFILFKKFIGLVAQDL
CQKMPNKPYPFGVFGTEPTLMVQDPEMVKTVLTKDFYFSSREISKYTHREVFSQLNLF
TYGDKWKVYRQNLTPLFSSSKMKKMFYLIEKCAHVFEENMLDHEFKTNNVHGVRLHMA
RYTMDICIGSCAFGINNTMDQNSTKNPFTNIGELLFEESYIRGILVNLRAIWPVAVFYGIGLR
SFPKQVDDFFNQLMMGVFKERNYKPINRNDFVDLILNWNVGNIIIGDSIENMKGDDKK
VKLEVNDDEMLVSQCMLFFAAGFETSSTTLQYTLYELAKHPDAQKRVIQEVDEYLRHD
NKLDYDCNAELPFTEACFDEALRLYPVVSILTRETADDEYTLPCGAVIEKGIRVHIPVYHIH
HNPDLFPDVEMFRPERFYGEEKRNIPYSYMPFGGPRCLCIGMRFAKMQMFSGIITLLKK
YQVELAADTPRKLNFKSTTLVTQPKELIKLKFVEREGWEQRMFNK

>CYP6AE56

MIVEVLLIVVVTFLILYHISKRRKFGYWKERNIPYLEPVPILGNYGPFIMQRKYNGKMAQ
ELCQKFPDEPYFGVFGTEPTLMVQDPEMIKTVMKDFYFSCREISNYTHREVFTQNL
FFICGHKWRVLRQNLTPLFSSSKMKNMFYLIECAHVFEGLDRDIKVNNVQGVRLHMA
ARFTMDCIGLCAFGIETKTMEQDSRKNPFTSIAAIMFMESYIRGTMIWVRAIWPALFYGL
GIKLFPKIALNFFKKLMRGVFEDRQYKPTNRNDFVDLILNWKKNNFISGDSLQNMKGGE
NKVRLEIDDEMLVSQCMLFFAAGFETSSTTLQYLLYELAKDRKAQQKVIGEVDEYLLQH
KNKLEYECTTELPFTEACIDEALRLYPVAVGILTRELVEDYTLPCGAVIGNGIRVHIPVYHL
HHNPKFFTDPEMFRPERFYKQKRDVIPYTHIPFGGPRACIGMRFAKMQMFSGIITLLKK
YQVELAADTPRKLNFKSTTLVTQPKELIKLKFVEREGWEQRMFNK

>CYP6AE57

MIGAVFIVLCVILLLYHLVTRKFDYWNQRKIPYIKPVPLFGNYASYILLKSFIGHVTQSIC
RQFPNAPYIGTFYGTPEALLVQDPEVIKLMVKDFYFNGREITNYTHREPITTNMFFTYG
DRWKVMRQNLNPLFSSAKIKNMFYLIEKCARQFESMLDDEVQRDNILKVRALMSRYTM
DCIGSCVFGIETKTMEKGADENPFTIIGHDLFLETKVGGLKMIARSIWPVAVFYALRLRNP
ENISKFFHKLMTVEVFAGRNYKPTTRNDFVDLLLNLNEKYLTGDSIENMKGANKKISLK
VDDEMLVAQCSLFFAAGFETSSTTVQFVLYELAKKPDAAQKALEEVDEYLRHNNRVE
FDCINEMPYIEACIDEALRLYPVLSVLTRELVHKDYTLPCGAVLEKGVRVHLPVYHIQHRP
DDFFPEPEEFRPERFLGEEKRNIPQYTYFPFGGPRICIGMRFSKMQMIPGLITVLKSYRVEP
VGPAEELDDFSTTMVITAKKEIKLKFVRREDCLQRKF

>CYP6AE58

MIGAVLALLALLFVVYHASTWKFNYWKKRKVPYLKPAPLVGNYGPYIMQQQYVGHT
AQDICEKFPNESYFGVFGTEPTLMVQHPEMIKTVMTKDFYFFNSREIAKYTDREIISHSL
FFTYGDKWKLLRQNLTPIFSSAKMKNMFHLIDKCAHVFNMLDRELKTDNEQYVRPLM
ARFTMDCIGSCAFGIETNTMAPNYENNPFTVMGTLIFSDSYRNIKNIGRAIWPAIFYGLG
MQNFPIDIGKSFTKLMKEVFADRQNKPSNRNDFIDLILNWKKECITGDKIQNMKGDINE
KVSLEVNDLLVSQISSFFAAGYETSATTMHYTLYELAKNQEAQQRVLDEIDEYLLRHK
NKIGFDCITDLPFTDACFNEVLRLYPVLGILTREVDDYTLPCGAVIEKGVRVHIPVHHL
HHNPDPFPDPEEFRPERFYGDEKRNIPYTYIPFGGPRLCMGMRFKMQMSAGIMTVL
KKYRVELAADTPRKVTQFEATTMVTLPEKIKLKFVEREGWQQRQYSA

>CYP6AN17

MIVILLVFAVVAFYLYATRNYDYWNKRNKHDPPVPPFGTHFRNAFAFKSITDISQELY
NKYPNEKVVGYRWSTPELIHKDPDVRRIMSTDFNHFYFRGMARNPKVDRVLLNPFH
ADGDNWRLLRQRLTPAFTTAKLKNMFPLIVKCAEKLHSLGDDIVSRGGDCDIREMMAR
FTTEFIGACGFGVEMDTINNEKSLFREFGKLIFSRTAFDITRFAVWEAVPLVRGFLPVINP
KIEEIAIEIVKNVREQRNGKPSGRNDFIDLLLELESKGKFTAESIEKTNPDGSPVIVDMEM
TMACMAAQVFIFFAAGFETSSSATSTLHELAFNPEIQERIQEEIDEVLSEYNNKLCYEAV
AKLSLLDMAFKEAMRMFPSLGIVTRECTRPCTIPELGISIDPGVKIVIQAIQNDEKYFDN
PTEFRPERFTPEEVKNRQPFITYMPFGDGPRMCIGARLGQMQLAGLA AVLQRFTVEPNK
DTRRKPVNPRSNVVQGLVGGLPLKLT LRKNK TASK

>CYP6AW1

MWLIAILLGVILALLFLLTFFHYTTNGRKYWLHKNVPYREPCPIFGNFGPTFIMRRSYVN
MLQYFYDNYRDKKFVGIFQARRPTLMVIDLDLSKNIMSSEFHNFSDRVSVSTDTQREPL
LRNLANMSGAEWKAMRHIVTPTFSSAKMKAMFPLIADCALMLKNILLNDVTDEVDVPHK
LMTRYTTDVIGSCAFGFNPESLKDPSFLIMAGKMFRADRATILKRYCRTFFPRLFKFL
NLRTYSSDVETFTTGIIQVL TERRKTGAHRHDFLQLMLNVQKSENGFPMTDALITSNSFI
FMLAGLETSSTTSLFCLYELARLKLQDKIRNEIIECMDRYGGLTYEAVGAMKLSMQAIL
ETLRLHPPTPLTTRLCTSPCVLPETELKIKARDPVLPIHICIQRDAEHFPDPDKFEPDRFKD
DLNPPGFIAFGYGRSCPGARFAQLTTLAGLAAILTAFTVEPCARTTSRIHYDPRSVMLK
NKGIFLKF TPLDPEN

>CYP6B44v2

MIPFILVLITLITLTYYYTTKTLHFWKNKKVLGPKPVALFGNLFEMTTTRKKYLGEILKDA
YNQFPDEKVVGIYTMVTPVLLVRDLDIVKQVTIKDFDVPDRGTEFSKEGLGANLFNAD
GDTWRVLRNYFTPLFTSGKLKNMTYLMTERADNFLNYVDRIKENPEQEVHKLIIQKYT
LSSISACAFGIDMDIDKENEIVKNISSFDKLIFDISYFIEIIMMFPDLMKKFNTSLFPKSIVNF
FYDLVRVVIDERKGAFIGRKDFMDLILELRQQGEIKGTRRGDDKQLQLDFNEDVIAGQA
FAFYVAGYETSATTMTFMLYELARNLDIQKRVLDEIDEVLQRHEGKITYDSVNELNYM
GKVFDETLRMYPVLDLQRRACNYEFPDLGLKIEKGQAVFISPMMSIHDAKHYPNPQT
FDPERFSENVKNRHPCAYIPFGVGPRNCIGMRFAKVQSRVCMMKLLSKYSVEMTHSTP
KQVELSPLRLILSPMGGMHLKFNPRY

>CYP6B54

MIPVLLAFIALFYGAYYYFTKNLDYWKKKG VAGPKPIFFGNCYESFIRRK FVGDVIEEV
YNDYPEEKVVGCYTMITPVLLIRDLDLIK HVMIKDFDVFSDRGVEFSKEGLGANMFHAD
GDTWRVLRNRFTPLFTSGKLRNMTYLM TNVADKFLTYVDTLVSKQPEQNFHNLIQKYT
MSTVASCAFGIETNMSDDALIDRVTSFDNL IFQINYFQEFDMMYPGVLLKLNMSIFPKEI
KNFFNDLSKQIINERNGAPIGRNDFMDLILE LRQQGDIKTVRRSDEKELQLNFNQDVIAA
QAFVIFYVAGYETSATTLTFMLYELAKNKH IQEKVLEEIDEVLKRFDGKVTYDCISELKY
LDKVFEDETLRLHPVGSSLQRRRAETDYKF PGTDVKIKKGQICYMSIRGIHRDPKYYPDPYK
FDPERFSPENVGKRHPCAYMPFGLGSRNCIG MRFAKVQSRICAATLLSKYRVETTERTW
NNMKLDPKRVVSTPVGGMYLKLVP RR C

>CYP6B55

MFLVLLALLALVYG VFYFTSTFD FWKKKG VKGPKPIFFGNCYESTIRRK FVGKVVKE
IHCDYPDEKVVGIYTM LTPVLIIRDLDYIKHIL IKDFDFFSERGIEFSKKGLGANLFHADGE
TWRVLRNRFTPLFTSGKLRNMSYLM TNGAEHFLNYVDTLVTIQPEQNIHKL VQKYTMS
TISSCAFGIDTKMSNDG LLERVASFDNLIFGV NYFAELDMMYPGLLLKLNASLFPREISNF
FYELSKNVITQRKGAPIGRKDFMDLILELR QQREIKTVRRSDEKELHLEFNEDVIAAQT FV
FYVAGYETSATMTM TFLLYELARNKDVQEK LLEEVD EVLKKFGGKVS YDCVNELKYMD
RVFDETLRIYPVVS MLQRRRAETNYKFPGSD LEIKKGQICYISPWGIHHD PKIYPNPEKFDP
ERFSPENIGKRHPCAYLPFGLGPRNCIGMRFAKVQSRICLV SLLSKYRVETTERTWKSTM
YDPYRVVTTPTGGMYLKLVP RN

>CYP6B56

MIYLLALIALFYGLYFYFSRTL DYWKNMKVRGPEPMAFFGNFYESAMRRISAGLIVKN
IYEQFPDEKVVGVYAMVTPILLIKDL DLVKQIMIKDFDLFSERGV EFSKEGLGANLFHAD
GDTWRVLRNRFTPLFTSGKLRNMSYLM TKGADLFLEHV NKL VSKQPELDVHKL IQKYT
MSTISSCAFGIDIDMNKDDTL LDQIKAFDKLIFEVNYAQELDMMYPGILKKINMSFFPRA
LKIFFYDLTKNVIAERN GSPIGRKDFMDLILE LRQQGEINGVRRSDEKVLHLD FNEDVIAA
QAFVFYAAGYETSATTMAFMLYELARNKDVQDKLLAEID DILKKFDGKVTYDCINELK
YMEMAFEETLRLYPVVD TLQRRVETNYKFPGLDLEIKKGQMCVVSPMAIHRDAKYYPD
PEKFDPERFSPENVNRHPCA HIPFGIGPRNCIGLRFAKVQSRICLAKLLSNYRVEVT KNT
VMEFDPRRVVTS PRGGMFLKLVP R

>CYP6CV3

MVLFYLSIFVLLYAFYLF SKRKHNYWKQKGVKSPDPMPVVG NFGKVM LLKESVQKLF
QNTYKQYPNEKIVGLYRGFTPMLLLRDPELIKQVLIKDFPLFQDRGMSTSR SKLGKNLFE
ADGETWRILRQKLTPVFTSKKLKDMIPLVEQC VENYIKYVEHLTENNINHEIRSLQSKYT
LQVIGSCAFGLQLDTYNDNEDEFSIMAKRIFNPTTFSRALILLDMIIPGV RKL FNTASDLQ
NFFLNLVSKVIKQREHEPMKRKDFMDLMIELREQGKVSRKSEDGILEIDMNDGLIAAQA
LVFYAAGFETSSATMSFMIHELALHQDIQNKVYDEICSVTAKHNGKINYDCIKEMTYLE
MVFDETLRKHSVAGLLIRKSLTDYTLPGVNV SISKGTPVLISVSGLHWDEQYYKNPEVF
DPENFSAERKIARPPCTYLPFGDGPRNCIGMRFAIVQSLMGMAAFLRKFKVEPSTRTKSA
LDYKPTGLTLESKDGIWVKIVKRTPATQ

>CYP9A63

MIFIWTLVVI AALTLYLRQAYS RFSKYGVKHLKPM PFFGDMFRVMTRLEHFTRNFERV
YDFPNERFVGM YQLVKPIVMIRDVELVKKITVKDFEYFLDHQEIFDGADQFFGRNLLSL
KGQEWKDMRSTLSPAFTSSKIRHMLPFMVEVGDQMVMCLKEQIKQSRVGYVDIEGKDL
TTRYANDVIASCAFGLKVN SHADQNNEFYEK GKEASTFKFKQLLKFFGFSTCPKLMSLL
KQSLFSEEIKRFFQTLVLSTMRNREEQSIIRPDMIHLLMEAKKGQLVHDEKTQDADAGFA
TVEESSVGKKETNRVWSDTDLIAQAVLFFVAGFETVSSAMS FALHELTANPEVQEKL AH
EIKENEMKNGGKFDYNSIQNMTYMDMVVSEVLRLWPPAILLDRKCIKDYNLGKPNENA
KHDYIIRSGENLSIPAWG FHRDPKFFPNPLKFDPERFSEENRHKIKPFTYMPFGFGPRNCIG
SRFALCEVKVM LYQLLQHFELSPCEKTSIPAKLSTETFNIRIKGGQWMRLKIRN

>CYP9A64

MMLLIWGLVLVAALALYLHQVYSKFSILGVKHFKPVPLLGNMARVTLRMNHFSDDDVE
RLYNAPSERFVGRYEFLNPMLIIQDIELVKKITVKDFEHFLDHRA FVDEQADPFFGRNLF
SLKGQEWKDMRSTLSPAFTSSKMRHMLPFMVEVGDQMVM T LKNKILESGVGYTDVEG
KDLTTRYANDVIASCAFGLKVN SHVEQDNEFYKMGKIASTFN YIQLLKFFGFVSFPKLM
TLLKVSLFTEESKNFFKNVVMSTMHEREVNNIIRPDMIHLLMEAKKVWSDTDLIAQAVL
FFVAGFETVSSAMS FALHELTANPEVQEKL AHEIKENEMKNGGKFDYNSIQNMTYMDM
VVSEVLRLWPPAILLDRKCIKDYNLGKPNENAKHDYIIRSGENLSIPAWG FHRDPKFFPN
PLKFDPERFSEENRHKIKPFTYMPFGFGPRNCIGSRFALCEVKVM LYQLLQHFELSPCEK
TSIPAKLSTETFNIRIKGGQWMRLKIRN

>CYP9A65

MLFIIWTV ALIAALTLYFRQKYSKFSKYGVKHFRPV PFLGNMARIIFRMDHFTKDFERLY
NEFPNERFIGKYEFVKPSLLIKDLELVKKITVKDFEYFLDHRSFFDENTDPFFFRNLVALK
GQEWKDMRSTLSPAFTSSKIRLM LPFMTEVGDQMVM T LKNKIQENGVS YVDIEGKDLT
TRYANDVIASCAFGLKVN SHVEQDNEFYRMGKEATTFTFRQILKLIGFVFLPKVMALLK
VSLFSERIKKFFYSIVMNTMRDREERNIVRPDMIHLLMEAKKVWSDSDLIAQAILFFMAG
FETIASAMS FGLNELTLNPEVQEKL VQEIRENEIKNGGKFDYNSIQKMTYMDMVVSEILR
LWPPAIALDRVCLNDYNLGKANDKAEYDYIIRKGEILAIPVWSFHRDPQYFSDPLKFDPE
RFSEENKHKINPLTYMPFGVGPRNCIGSRFALCEVKVM LYQLLQHFELSPCEKTSIPVVL
SMESVNMAIKGGQWMRLKIRK

>CYP9A66

MIFIWALLIVAGLTL YFRQIYSRFTKYEVKHFKPV PFFGNMVKFIFRYEHMTQDMDNLY
KAFPHERFVGKYEFTVPKIVVNDLELIKKITVKDFEHFIDHRIFADGSLDPIIGRNLLALKG
QEWKDMRSTLSPAFTSSKMKHMLPFMVEVGDQMVM T LKNQIQQSGVGHTDVDGKDL
STRCTNDVIASCAFGLKVN SHAENNEFYRK GKDAITFNIIHILKIFAYSVC PQLMAKLLK
VSVSPNTKQFFESIVMNTIRDRKEKRIVRPDMIHLLMEAKKGTLTHDDQTADTNAGFA
TVEESFVGKKQIHKEWSDDDLLAQAMIFFLAGFETVSTALSMGLHELAVNPEVQEKL LR
EIRENEVKNGGKFDYNSIQHMTYMDMVVSEIIRLWPPVPSLDRMCVKDYNMGKPN DK
AKQDYIIRKGE GEFIPVWC IHRDSQYYPNPQKFDPERFSEENKH NINPLAYMPFGAGPRN
CIGSRFALCELKVILYQLLQHFELSPCEKTSIPPVFSTKTVNLTIKGGQWVRLKMRK

>CYP9A67

MIFIWALLIVAGLTLYFRQIYSRFTKYEVKHF KPVPFFGNMVKFIFRYEHMTQDMDNLY
KAFPHERFVGKYEFTVPKIVVNDLELIKKITVKDFEHFIDHRIFADGSLDPIIGRNLLALKG
QEWKDMRSTLSPAFTSSKMKHMLPFMVEVGDQMVM T LKNKIQENGVS YVDIEGKDLT
TRYANDVIASCAFG LKVNSHVEQDNEFYRMGKEATTFTFRQILKLIGFVFLPKVMALLK
VSLFSERIKKFFYSIVMNTMRDREERNIVRPDMIHLLMEAKKGQLVHDDQPMGDSHAGF
ATVEESSVGKKQISRVWSDSDLIAQAILFFMAGFETIASAMSFG LNELTLNPEVQEKL VQ
EIRENEIKNGGKFDYNSIQKMTYMDMVSEILRLWPPAIALDRVCLNDYNLGKANDKA
EYDYIIRKGEILAI PVWSFHRDPQYFSDPLKFDPERFSEENKHKINPLTYMPFGVGPRNCI
GSRFALCEVKVMLYQLLQHFE LSPCEKTSIPVVLSMESVNMAIKGGQWMRLKIRK

>CYP9G13

MIIEILIFIVTFV VAYYLYVYKRNHWFF EKRG I KYVPGVPIFGNVYDTMLLKKHMIDSIDD
VYKAFPDEKYVGFIDAIDPIIMIRDPEIKNITVKDFDHFVN HKVFFQEEAEPLFGGSLIQM
KGEKWRDMRTT LSPAFTASKMRLMIPFMFEACKNIIDDLKERVKTDD ETDVEDLMRRY
TNDVIASAGFGLQVNSVKERDNEFFIMGRDIFVFNTFQKIIFFFSTQIPWLFKALNIKVFPQ
KTMNFFRN LVRTTIESRNTNKVYRPDMIQLLMEASKGKLTSTDADDGESVGFATTQEAL
KPRGVAREWTEDELIGQVFIFFVAGFESSAGAL T LSIHELALNPEVEEKLYQE CRQFKETE
NTLTFENVGQLKYLD CVINETMRKWAVALIMDRCTCKPYELPPPREGGKPVLLRPGDV
VYNMVNCIHMDKQYFPNPEVFD PDRFSDENKHNIKPF TFM PFGMGPRNCIASRFALLEL
KVLLYNLIVNFKIKKCKKTLDPIVLNPGDFNIKVHGGTFVKLELRA

>CYP18A1

MITMLTNSKILWGLWQVVS YCVSRTSVPLLLAAGAAYLVARLVVLMRDMRKLPPGPW
GPPVVG YLPFLGVRHKT FLELARSYGT LFSARLGNQLTVVLSDYKLIREAFRREEFTGRP
STPLMHTLDGLGIINSEGRLWKNQRRFLHEKLREFGMTYMGNGKKIMEGRIKNEVYELI
ENLHRSEGSPIDANPLLALVVS NVICGITMSVRF SHGDARFERLNFLIEEGMRLFGEVHY
GEYIPLYNYLP GKALAKEKVAKNREEMFEFYQTLIDEHRETLDLNNARDLIDVYLIEIEN
AKTEGRGGELFEGRDHELQLKQILGDLFSAGMETIKSSLLWMIVFMLRNP DVKKRVQDE
LDAVVGNERLPTIEDMPHLPYTETTILETLRMSSIVPLATTHSPTKDVHLNGYRIPAGSQV
VPLINCVHMDPNLWDEPKKFNPSRFIDENGKIKRPEFFMPFGVGRRMCLGDVLARMEM
FMFFACMMHQFDIQMEAGDRLPSLEGIVGATIAPKA FRVKFIPRGDPMPPPAPAPDHPHL
RHVGAH

>CYP301A1

MGRSLHSFTAYARPYVAKNARYATTGCPFSKRQRSQIAPT AELNEEIFANAKPYSEVPGP
KPIPLLGN TWRMVPVIGQFDISEFAKVTKMFMDKYGRIVRLGGLIGRPDLLFVYDADEIE
RMYRREGPTPFRPSMPCLVKYKSEVRKEFFGDLPGVVG VHGEQWRKFRSKVQRPILQP
QTVKKYVTPIEMVTEDFIKYMEDARDENGDLPHEFDNDIHRWSLECIGRVALDVRLGCL
TSKVSNDSEPQRIIDAAKFALRNVAVLELKAPYWRYIPTPLWTKYVNNMNF FVELCTRY
INEALERLKT KKVTSNDNLSLLERVLKSEGDPKIATVMALDLILVGIDTISMAVCSILYQA
SIRLKQQEKMAEEIRRVLPDPSKPLTYADLDKLQYTKAFVREVF RMYSTVIGNGR TLQE
DDVICGYHIPKG VQVVFPTIVTGNMEEFVS NPEEFKPERWLES GGKLHPFASLPYGFGAR
ICLGRRFADLEIQVLLAKLLRRYRLEYHHEPLDYAVTFMYAPDGPLRLRMIER

>CYP301B1

MRNRIAVIKWPLKSLSARQVGTTAHCYQAAEGAPTQTVPPLARCPIGHRSRSTHAAIDS
TVFDTLSVPSVKSWDEVPGPKPLPLLGNTWRFIPYVGGYSVEHLDDKVCLSLRQQYGKCV
KMTGLLGRPDMLFVFDAGEVERVFRGEDAAPHRPSMPSLNYYKHKLRKDFFGVEENC
AGVIAVHGDAWAAFRSKVSRVALSASAAAQYTEPVAQVADAFVDRLRQVRTDAMETP
EDFLNEVHKWSLESGLIALDTRLGCFEATDGSESQLRIDAVQTFFLCVGELELRAPWW
RIYPTAMFRRYVAALDTILSVTLSHVDRALEETRSGGTSLLQDLVAAAGPRVAAVAAL
DMFLVGIDTTSNAVASTLYQLSLRPEVQDKLYEEISKVLQGRPMRPGDINQMPYLRACI
KEVMRMFPVVIGNGRQLTKDTVICGYNIPKGTQVVFQHYVMGNSDEFFSNSSEFRPER
WLRSPNQAHHPFASLPFGFGKRMCLGRRFAELEIHTVIAKVQAFKLEYHHEPMEYHI
HPMYTPNGPIRLKLIER

>CYP304F10

MFGAVILVLLSIIVNYLYKNATKRPTKNFPPGPPSLPVYGAYWMVLANGYNNLSKAFR
KLGQTYKTKIVGLYLGSFPVIVVNDPVMIKQMLYHEDFDGRTDILGRLRAFWKRLGIFF
TDSYFWYVQRRFSLRYMRDFGFGRRDETLEAVVESEVKTMLDMRLNGPQYPAAEEKIVN
GELVYLPYFLTIPFINGMLQVLCRSTLPRSQYEVLEWEMSRAGLLFQRSSDDVGRALSLTP
WLKDLAPGWSGYSGLRKGNQYFLDFFTNLINQEMSTFDESHQRHFLDVYIKKMREEQR
QHDKSTFSVEQLILICTDYMFPAAATAVQETLAMLLERLVIQPELQDRLHEEVDRVVGSG
RLPNLDDRKNMPYTEACIREIMRFETLVPLGVPHRAISDTIFNGYDIPQGTMVSPNLVML
HMDKQIWGDPEFRPERFIENGKLNALDKSLPFGAGRRLCAGETYARQTMFQVFAGIV
QNFRVERAPGSGPMPSPRIQGIITTIPDFWIKLTARI

>CYP304F11

MIGALIFFSVSVYFLVIFYKNAFKKLPNFPPGPPSLPVYGSYWIILAAKNNLALACKRIG
SWYNTKIIGTFVGKIPTVIVNDPELIRDMLHKPELDGRMDITLGRMRAYWKRLGLFFTD
GFFWHVQRRFSLRYMRDYGFGRRDETLEQVVESEIKEIFDIRLNGPKNDAEKDMVKGD
LVYMPHFALATPFINGILHVVSRLTLPREKYGYVWELAKGTLLFQRSSNDLGGALTFTPW
LKDVMPNYSGYNDLKKGHQYLLDFFTNLVKEIISTFDSSHPRHFLDEYIRQMKKEQNEK
RSTYSVEQLILTCVDYMFPAATAPQAVLAMLLERLVMQPELQDRLHEEVDRVVGSGRL
PNLDDKIKMPFTEACIREQMRMDTLVPLGVPHRATEPTTFHGYNIPEGTMFTANYPMLH
KDEVIWGDPEVFRPDRFIVDGKLDVSKDKSLPFGAGRRLCAGETYARQTMFQVFAALV
QNFRVERAPGSAPMPAPRMQGIITTIPDYWIKLTPRH

>CYP304F12

MLGIIGSVVFLAVLLRSLYKNAYRRPSEKFPPGPPSLPVYGAYWIVLWYGYNDLANAFR
KLGEIYKTKVVGLYLGSVPTVVINDPVMKEMLSDDAFDGRMDVILNRVRSYWKRLGI
FFVDSYFWYVQRRFSLRCFRDYGFGRRHQTTLESVVENEIKTLLDIRLNGPQNPAEKGVV
DGDLVQFPHFVVTQNGMMHVLCQTTVARDEYEILWELSRKGMMFQRSSDDVGGAL
SLTPWLKDLAPNWSGYNNLRKGIQPFITYFEKIVEREMETYDKTNERHFLDAYITKMNE
EMREKGRSTFSVEQLVMICVDYLFATTVAETTLTMMFERLLIQPELQDKMQEEVDRVV
GSGRLPNIDDRKNMPYTEACIREAMRWDTPITLGVPHRATRDTVFNGYDIPQDTMVAPN
YVMLHMDKDVGDPENFRPERFIVDGKLDVSKDRSLPFGAGRRLCAGETFARNTMFCI
VSSIMQSFRVERAPGTGPLPAPRIQGIITTIPDFWVKLTLRKKQ

>CYP305B1

MLFHFTKRNTQVQQSMILVTGPRWYPFFGSSYVVQQMSRKYGTQWKGLSKLARDYST
QVLGLKLGRELLVVVFGEKNLHQVFTQKEFEGRPDSFFLRLRCLGKRLGGITFTDGPLW
KEHRQFTLKLHLKNVGFGRMEKEIQKEMLTVLEYIDRKKRKSNNVKEVLPLAVMNVLW
RYVAGEKIKEDRLRSLLEILNMRSRAFSMAGGWLNQMPWCRYLMPRISGYALIAKMNK
EISDIIIEAIIKKHKCNAVEGTDFIYSFLREIDENKETYTEEQLKTICLDLLIAGSQTTSNVVE
FAILAALRNQGVQQKMYEEIKNTLGDDVPSWGDYRLTYTSAFLLEVQRFFFTIVPLAGP
RRVLEDTQIDGYLIPKNTTVLISVADLHFDAGIWDKPHEFKPERFIDETGALKNSEHMYH
FGLGRRRCPGDSLAKSFIFITVVGILQKYRIDCSNGILPSDEPNVGLIAAAKPFTVDFVPRT
SCDSSN

>CYP306A1

MDFFLLWLFTVVFGYWIFKKIRLWQSLPPGPWGLPILGYLPFLDRHQPHLTLTKLSKEY
GPIYGIQMGSIYSVVVSDHKLVRDAFAKDSFAGRAPLYLTHGIMHGNGIICAEGPLWKD
QRKLVTTWLKSIGMTKQAATREKLEKRIASGVFELLDNIDKAGKDRMNLAHMLSNSIG
NVINEIIFGFKFPPEDKTWNWFRQIQEEGCEMGVAGVVNFLPFIRFFSASTRKTIEVLVR
GQAQTQRLYASIIRRRRKMLDLPAPEEAKYPPHADLFTENPNGYLKCITYTKYAAVEEH
FFDPNIIETEGECILDNFLREQRRFEEGEEAARFATDEQLLYLLADMFGAGLDTTSVTL
AWFLLYMALYPEEQENVRKEILSEYPKEEEVDMARVPYLMAAVCETQRIRSIVPGIPH
GCLEDTQLGDYVVPKGAMVIPLQWAMHMDPNVWESPNEFKPSRFIAADGSLLKPQEFI
PFQTGKRMCPGDEL SRMMACSLVSRLLRKRIRLASDPPTTEAMRGNVGVTLTPPKVEF
YCDDV

>CYP314A1

MALPGVLLFSQYMENFWGTTPPLVDWSVGIPTLVLLVVALVVAATALMARAADGRQA
TRLPGPPALPFLGTRWLFWSRYRMNKLHEAYEDMFRRYGPVVFVETTPGGAADVVSIAQR
EALETVLRTPAKRPYRPTEIVQVYRRSRPDYASTGLVNEQGEKWHHLRRLTAELTS
PQTMQNFLPELNNICDDFLNLLQSSRKP DGSVLGFDQLTNRMGLESVCGLMLGTRLGFL
DRMSGRAATLATAVKAHFRAQRDSYYGAPLWKFAPTRLYRTFVWSEETIHTIVSDLM
EEAKSRTCSSQDDGMQEIFLKILENPALDMRDKKAVIDFITAGIETLANSLVFLLYLLS
GRPDWQQKIRSELPVCGVLKAEEMAAAPAVKAATNEAFRLPTAPFLARLLEAPMTISG
HKLPAGTFVLAHTGAACRREENFWRAQEYLPERWLECREPHAASLVAPFGRGRRMCPG
KRFVELELHLLAKILQRWRVEFDGELDIQDFLLSPKSPVSLRLVEW

>CYP321C1v2

MFVQIILLIFIVFLFIYTTGKHNNENYWKKRGVKLHKKNKVTGMSWDFFTTDKPVFQIIG
DLYKEYREEPAVGFGNFLTSSLVNDATNIQHIFQTDQSFNHRGFEINEGDHLADNILF
MNGNRWKLVRQSMTPIFTSTKLKQMFHIMDKSGRDFITYLKDNPERLNGNAIDSISTFC
CAAIGA AVFGVESESVFESPFLDVAKRAMKFSVTMNIKALIMNVFPWLSKKNLNFKLSE
HEDFFIGAVKRIIRQREQDKVQKHDFADMCVSLQRNGTLKDRDTGLELEPTDELLSGQA
FFFFTAGVDPSAAAIFGCLLELGRNPEHLKKVHNEIDTTFEKLNGNLSYESVMEMEYLEK
VLCESLRMWPPIGVLNRQCVADTVLSVGNIKIDKDTKVYAAVYELHHDPPKYYPNPEVF
DPERFSSTSAQDSNVYMPFGKGNRLCVGMRFARLQVKTGLIHFLRNFTVKTIDGDRIR
FKKDPVQVRPVNVDIEFIWRNI

>CYP321C3

MNLFTPNIITSSSEDSRPIKMLIPLILLFLVTLLLFIHATGKNNENYWKRGVKFHNKKNKV
MGIFWDFLMKNRPMFQVVGDLYKQYREEPAVGFGGFFTPSLLINDATNIQHVFQTDQFQ
AFNHRGIEINEGDHLADNILLMNGNRWKLVRQSMPTIFTSAKLKKMFYIMDKSGKDFVT
YLKNNPERLKDNAFDTISTYCCAAIGAAVFGIVSESVFDSFPFLDIARKAFKSTFKSDIKTTI
SSFSPWLSKKNLKFSEHQDFIAAIKRIIRQREQENTKKYDFADMFCVSLQKSGTLRDRE
TGLELEPTDELLSGQAFFFFTAGVEPSAASIFACLMELGKHPEHLKKVHDEIDATFEKLN
GELTYESVMEMEYLEKVLCEIRLLPPIGFLGRQCVENTVLPVGNKVEKGTKVFVAVYE
LHRDPKYYPNPEVFDPERFASESAQDSSVYMPFGKGNRLCVGIRFARLQVKTGLIYFLR
NYTVKTKMNRYPPIRFRKELAQVRPNNVDIEFIPRNIKN

>CYP321C4

MWFTIILLSFIIILLMTFYMMGKYNNENYWKRGVKFHNKKNKVFGVFWDFLTCKDKAMFQI
IGDLYKQYHDEPAVGFGGLLTPSLLVNDALNIQHIFQTDQAFSHRGLEVNEGDLADNI
LFMNGNRWKLVRQSMPTIFTSTKLKQMFYIMDKSGKDFITYLKNNPKRLKGNAFDSIST
FCCAAIGAAVFGVESESVFDSFPFLDVATRAFKSTFKTIKTTIMNLSPWLSRKLKLLKFSE
YEQFFIGAIIIRQREQESGNKHDFGDMCVVLQRNGPLKDRDTGLEIEPTDELLSGQAFF
FFLAGVDPSAAAIYACLLELGRNPEHLKKVHNEIDRTFEELNDNLTYDSVMKMEHLEKV
LCESLRLPPIGFLVRQCVEDTVLPVGNIKIQKGTKVFVSVYELHHDPKYYANPEVFDPE
RFASETPQDSSIYMPFGKGNRLCVGMRFARLQVKTGLIHFLRHFTVKTCKINGDNLRFRK
DQIQVRPNNVDIEFIQRNLQ

>CYP324A10

MLVLLVSTLACLLVWLYLRWTKVKTYWAARNVPHEPPHPIWGSFTFMLQKNAGIWM
KEMYSRFSSPYVGIWLFWRPALIVNCPEIGKRILVKDSELFDRDRFLRSGRNDPIGGLNIFTI
NDPLWSSIRRLTPVFTSSKLRLSLHGLLSTKTAQLVERIRNDVKAGETVNLRMITYSDFTT
DVIGIASFGVDSDATSKGESHMRTVTKDFMTYNFWRGLAWSCEIFFPELVDFVFGFTMF
KKATDYFKKVFSHIEEQRRIEGSDKRDLMALLKIKLEAEDTNEGITEDAIIAQAAIFLQG
GFDTSGSALTFTTYELAFHPEVQDKLYHEIVELKQRIGNKDFHFDDLNLTYMTVCVINET
LRKYPPMGFLDRIAADYKIDENLTIPAGTPVYVNAIGMHVDPKYFPEPEKFNPDRLPE
NEEDIQPNTYLPFGEGPRVCIGMRFANYTMRYALASIILNFELRPLPDAPKPKDVEIEKM
GIFYMPSGVMSVEFVPLARMCRFSENVFPHNKSTSQLG

>CYP324A11

MFANLLLVGLACVLAWLFFRWRKVLYWAIRDVAHQPPHFIWGSFTFLQKNGGLW
MKEMYSRLEYPPYVGIWLFWRPALIVNCPEIGKRILSKDSEVFRDRFLRSGKTDPIGHNLF
TVKDPFWSTLRNQLNPAFTQAHLRLSLHLLFRKSKQLVQRINEDNTAKKIINLRMVYAD
FVTDVIGTAAFGVSSDATLTGDSMPRTVTKDLMNLGPWRGMAWSSIFFPELTDFFGFT
MFPKRSTDHFTKIFQSVAKKRSVEGSDKKDLMDALLKIQQDAKKTGGEINDEVLIAQAA
IFLQAGFNTTATVLTATYELAYHLEAQEKLYKEIVELKQSLGEKDIHCDLNLTYMT
CVINETLRKYPPMGFLDRVAATDYKIDKNLTIPAGIPIYINAIGIHYDPKYFPEPEKFKPER
FLPGNVRNIQPNTFLPFGDGPRRCIGQNFQDINIRYALALILLNYLLRPLPGTPKPNDCIE
KNAIFFMPSGMMPVEFVQRNKELVF

>CYP338A1

MFESFLLNLTVIIVLVVALIFDYVTKFFSYWYIRHIPYKTCVPFFGSDYHRVLGIRSNTDE
VNKFYNDYPKEKFAGAVKSRIPLIVRDPEAVKRMLSTDFANFHSRGLGLDKSQDVCLR
NNLFYADGEKWTLRLRQGVESLVKDMEWGKRDLNECLSGTNGDVIVQHLLSKVLDIFIR
DLLVDDTDAAVITSLTSAQAQRSLAGKFKTYLKNVFPSLYELFGLKTLVGEPSKETVRA
LQHSKLLLEKIKSGSVFQMGIKEKDRRNTNETDFAFSTLAMFVNEGYPCLYLLTSMFYDL
AKHPEVQKKVRQSVSNDDKYLD AVIKESLRLHPSYSVISRQCVKMYQYPENKLVIDK
KITLSIPVEAIHRDETYYEKPTVFNPERFLMNYDANNYFPFGVGPRKCVGENLAYQIIAD
VARNVLTHYEIQPCDRTPKSLPVSDFNFGRIVDQDIWLRFKPVT

>CYP333A8

MNYKAISKFIDIHRRLINIRHGVLSYSSGQPKPFESIPGLSSVPVLGTLHHFLPVIGSIGPN
LNLVQTFRVLNERYGSSVKMDGSFSKPSMLLLFEPELYEQVFRAEESNPIRPGLEALAY
YREQLRKSTLKGVTGLTASQGDWSRFRSTKVNPAKLLKLYAPALEEISEEMVDRLI
RLHREGDYLKENFDREVINWALESVAFIGMDTRLGCLNDKLGPDNPSKMLMKCATDIL
DLSFRLETLPPTMWKKYPTKTLKKLVEAFDLQWQISSQYVEEAEKRINERGYDVPPEEK
TIVERLLAVDKTVAIMMATEMLLAGIDTVSFTLISLLYDLAMNPEKQEKREGLFAEKPS
KRYLRACMKESLRLSPVVAINVRRTGKEHVVASYPHVPKGIAPNEILTNSDKFFPRPK
EFLPERWLVDKSDPLYYGNAHPMTSSPFGFGIRSCIGRRIAELEIETFVEKLLRSVKVKE
GPPELVTTTRIMKVHIKPFPGFKFEVITEYN

>CYP333A10

MKYNKSISKSLELIRRSKVRHGFRLAYSTDQPKPFKSMPGLSSLPIIGPLHHFLPLIGSIG
LRNFTETFGALGEKYGPVVKLDGIFARASMVLLFEPELYEQVYRSEETNPLRPGFETLT
YYREVMKKSTFNGVFGLTTAQGDDWSRFRSTKVNPAKLLKLYAPVLDEIADDEMVD
RLIRLQKEDGYFKENWDKEITKYSLESVAYVGLGSRGLGCLKDGLGDDHPSRQLIQCAKD
IISLSFMLELFPWLKIFPSPTYKKLVKTFDLQWEISRKYIAEAEKRIKERDHDPEEDKSII
EKLLAVDQTVAIMMANEMMLLAGIDTVSFTITTLTYHLATNPEKQEKVRQDILSKETSKPY
LRACLKESLRLPVVSANLRRTSKDHVVAGYHIPKIDVIAPNEYLSKQDKFYPRAKEFI
PERWIVDKSDPLYYGNTHPMVTSPFGMGVRSCIGRRIAELEIEIFFKLLKRVKVSWE
PLKIASKVMNSHVKPYGFKLEVLEKNLE

>CYP333B19

MQSCKLRSVLTSTVSWSKLPLQRYVAVNSAAVNEHQDVKVEGKPWKEIPGPSALPLIG
QIHHFLPGGIFYNTTTRNDYARKAYELYGPMVRLDGLLGREPMVLLFDAESAAQILRGE
NMLPVRPGFQSLAYYRSTYKYKGTGEKTGEPTGLITDHGDTWKRFRSAVNPVLLQPKTI
KLYRSALDEVAQDMIKRIKSVRGEDNMLKTKFDVEMNLWALESIGVVALGGRLNCFDP
KLPEDAPARQLIRVVHDIFKVAEELDFKPSLWRYFATPTFKKAMHIYSEQERLSIYFVNK
AKEDLKNKQKSDDEKGVLEKLLDIDEKIAHIMASDMLFAGVDTTANTIIFTLYLLAVNPE
KQEKLRREEILSTNHEDKRPYLRACVKESSLRLTPVVNGNIRQTTKEYNIMGYKIPKGMFV
AFMHQDMSCEEYYLNAQEFIPERWIAPKDDPLYYGNAHPFAYNPFGFGTRMCIGRRI
ELEIQTFLAQLIQNFKVEWFGPPLKVTLTSLNYATAPYNFILKDV

>CYP333B20

MQTCKLRPVLTASRSLLLQRNAAVNAAVNEHQDMKVEAKPWEEIPGPRALPIIGQIHH
FLPGGIFYETNDFAKKALDLYGPIVRLDRLLGRKPMVLLFDAESVAQVLRGENWLPERP

GFQSLIYYRSTYKYNEKKITEPMGLISDHGDTWKKFRSTVNPVLMQPKTIKLYTSALDEV
ANEMITRMKSIRGDDNMLKSKFDVEMNLWALESIGVVALGGRLNCLDPNLPEDSPARQ
LIKVVHEVFHIAQELDFKPSLWRYIATPTYKKAMHIYSEQERLAKYFVNKAKEDLKNNQ
RSEDEKGVLEKLLIDEKVAQTMASDMLFAGVDTTANTIMFTLYLLACNPEKQEKREE
ILSTSNEDKRPYLRACIKESLRLMPVVSGNARAANKDYNILGYKIPKGVHVILMHQDMS
KLEEHYIHAKEFLPERWIVPKGDTMYHGNAHPFANMPFGFGVRMCIGRRIAELEIQTF
ARLMQNFRVEWLGPPLKVTQTTLNYASAPYNFVLKDV

>CYP340AJ1

MVLEVLVIFVVTWAVLTYLQSGMKKENEPFVRFELPIIRHTHRILKARDGRDWWNLW
LETADDCWEKGGVLSINTIFPQFTHIVIGNPDDLIVANSTMKHYFYKAFHQHFGRHGLV
VATGNVWKNHRKQINLAFTPKFLNTVLIDLFGCCGRSLCSTWASKVVGKPFHEMQTISEA
ALEAVLGDIGKECNVSIKDYVEAVACLDRLALRLGNLFLHLDLIYFLTKDGREHIAKA
DYVSKEGRKLIEATRKKMRTNPDGTKEIPEIKTIVDFLVFMTDFNEDLLDEVNTFLAGG
HNTTASSLTSFFVILGSYPEVQEKLYKELVSVYGTSDRDVTKEDFASLPYLEAVIKEAMR
LYPAVPIVMRYSKKKLALTNYTIPGNRFYLMNIYASNRHPVWGPDRMEFKPERWLDRD
KLPPQHAFASFSLGKRSCIGKRYAMYAMKTIVTHILRRYKVIADINKFEVELAFLMKPIS
GGVITLEERNPDLATNEQKE

>CYP341J1

MLWQFCLFSLAVILLVTKWKNRKLRLKLANQLPSASFRKLPILGHAASLWGDCENLMRC
LEQLGRDSLANDGLISAWMINKLVIVVTDPDVAEVVMKVCLKKDDLNNVTRELIGYGSI
FAPVSIWRPRRKILAPIFSPKTLNSFVQIFSKQSHILLEKLEADENNGDVSIFDKVTRYSM
SVCETTLGTEVKSQSQPQHPVPQSFALHCKLAALRMVRPWYHSDAIYKWSWDYPQFKS
SAEVLKQFVRKIIISKRENFEKTNANKKITKVQSLDLLISSEDGKGFTDLELQEESSLVV
LAGTDTSSVGASFAALLMAKHPEVQKQVYQEVITVLED RPVQPEDLPKLKYLEAVVKE
TLRLYPPVPVVMKNPPKDVTLTPGVVLVPGSTALVHIWALHRNTKYWSGDADTFRPER
FLEDPPTHPYVYIPFGRGPRMCLGAEYAMMSMKTVLASLVRRYRLLPPTAGFGSQEELP
TRCEIMMKHVGNFQLQLQLRK

>CYP341K1

MWVQIVLGFILMVSFWCWRNIKFIKQYRRLGCSYTALPFIGHAHMFSGDREDIMNTFIK
FAEDAIARGGVAPLWLLGDYHLCLADPLDCEAALKNYLSKDKLLRMIRLLTGDSFIFAP
VEIWRPRRKLMAPLFSVRSINAFMSSFSKQSELMVDQLRTRADKEPFALWPYVNALTLD
AACETVLGVLDLHSQLDPEPILVAMDKLVELAGDGLTKFWFQVEFIMKNHHSYKRFLK
QQT VIRAFIDEVIRKQKERSHTITMNPITKEIYSLEDETKSVSLIEKLISLDGGLKVGDAGIR
DEALGLLLAATDTSATSISFVALMMARHPDVQEKVYNELQEVFGDSDRPVTLEDLPRLK
YLEAVVKETLRLYPPASNIVREADKDFLLPSGVKVFKGSNLIISIWGVQRNTKYWGADA
DQFRPERFLEGPLQHQN AFLAFSCGQRGCTGVQFAMMFTKTAISCILRKYRVLPAGEGD
ERKHLRLLFSVTLKEIDNFTVRLQSRENKLV

>CYP341L1

MVWLLILLVFVLWFVVFVRYRNRHKKLANLCRNNCMSLPIIGVAHHFIGTEEDRMKAL
MKIGREAIKQGGVAGGWLLNMLMVVIADPFDAEEVMRVSLRKMDGLDITEDLLGNH
MFAKVPIWRPRRKVIAPMLSQKNLEMFLEVFTRNSEN MIDKLTCLCGQPISLSKYITAYS
MDAICETSLGIKMDIQNNQDHPFLAFQTFEKMOVDRMLQSWLYPDFVYKLMPCYRIY

KKQTDILQEFVKQVILAKRIKFRDIENENRAECDQLQNQQNSRKSFLEQLIDIGYNDIDLQ
EEMLVILAGTDTSAAGASFTAVMLSQHQHVQDKLYEELMEVLGNSDRPLEPSDISRLK
YLEAVIKETLRLYPPVPAVVRMVEKEYILPSGFKLVPDVGAVIHIWAMHRNPTFWGNDA
EDFRPERFLDNNLKHPAAYMPFSFGVRNCVGARYAMTSMMTVIANLMRKFKILPPKNV
YSHGSQKRKKTPLRTKFDIIMRHVDGYEVQIEERR

>CYP341M1

MFLFSLFLIFYCYFLWLWSDRRRISLYRKIGSKYPAPVLLGHGYLYMGDREQVLQNLFK
LGDHAIANGGISTTWILHRFYLVLADAVDADYVLHHYNEKDDLMAVVRLLTRTCSLFA
TAPVWRPRRKIVTALLSHKNVSKFEPVFSKGSEILVEQLRPLAGTEAFSIWDHFNAYTLD
ATCSTILGIDFKSQLNLDEPLRNSINVMLDLAGEVVLKVWYQVEFFMKRTSFYKIFLKN
KVIKDHIDKAMKKHSDLVVKNNNEKFVLDEEKQFSLLHLFSQLKEQRGFDETMHLHDESL
GLLMASSDSTAIGISLVALVLARQPEVQEKVYQELKDVFGDSDRPVTIDDLPHLKYLEA
VIKETERLYPPAPHVSRKPAEDMVLPSGVKVPKGTSFIISIYALNRNPKYWGSDAHEFRP
ERFLDGSVPELSSYFVFGRGPRRCPGIQFSMMSTKTVLATLLRQYKLLPATPPDAPLRLFL
GITMKEVDNFRVKIQARN

>CYP341M2

MLLFILLLLLSCHVLWLWRNRRFFNLRYKVGSRFPFVPMGHGYLYLTNGEKVLNNLF
KIGDHAISNGGMSTTWLLHRFYLMADPVDAGYVLKHHLEKDELMLLIRYFTRTSSVFA
KVAIWRPRRKIVSALLSHKNVSKFEPVFAKNSEVLVEQLRPLIGKEAFSIWEHLNAYTVD
ITGSTIFGIDVNSQTNLMEPLRLAINGIMELAAEVLLQIWYHVEFLLKRSSYKQLLEYDH
VVKCYVDKALKKHTDAASRIKNDIEVPDDDHQISLLHLFAQLKEQRGIDHTTLYDESL
GLLMAASDTTAVGASTAAAMLAHWPDVQEKAYQELYEVFGDTRHMTIEDLPRLKY
MEAVIKEALRLYPPAPHVTRKALEDIVLPSGLSVPKGNTFLISAYALNRNPKYWGPAD
EFRPERFLDGSPLDQSLTIYFSFSYGPRGCPGIQFAMMSMKTSLGTLLRRYRLLPATVPNT
SRGSAGPTAPLRLFLGVVVKEADNFPVRIEARHGK

>CYP354A11

MEYFSGSRKILRFIMEDWKLLLFLTFFVFAVYFYTTSTFDFFEKKGVKYMKPTIFVGNMG
KRIASKVSFHTFQLEIYNYFKGNQFGGFFEGRRPVLYLMDPELIKAVTIRDFDYFTDRNIL
RIKEARYISQSLNLKGSEWKAVRSTLTAFSSARLKHMLPLIEHCSQQLVEFLKQYDYGK
DVEMKDTLGHFTLEVIGACAFGIKCDALTDENAHFVKVAEKFNYSMTKTRIFLLMVLMI
MPKMMQYLNMSFLNTESIAELMGILQKAKAERRATGIKKNDLQLLIDAADTENHQETS
DSEKSKVKATLDDDTIDAQTLLFLIAGYETSSTLSSFAIHALATRPDLQDKLRAHIEECTS
GKEMSYELLSQLSYLEGFLLETLMYPPVSRVDRVCVKPYTLPGTSVQVNVGDAVAIPI
YGVHMDPEYYPEPEEFRPERFIGDAKKERPSHLFLAFGAGPRNCIGLRFAMFSAKLAMV
KLLMNFKFSACSKTNNPVKIDKRAFLKADGGLWVRDLKILNLYLQLNSIYSLPVYS
KLCQSK

>CYP367B8

MIWLVLAFGLSVASFWLRWRYKNRKLLAMAKKIPGPPTLPWLGNVFMFMCDGETQV
VIKDLIRLYGDFLRFWLGPDLCVILTDPDLLKALLTNSKISIKGPQYKMYEVLGGGILS
GSGPTWRKHKRIATPNYKRAIEGYEITFNKEVDILMKKFRSCPGRTFDVYPYIVQCTSY
SVCQTLMLGTKEQTMCLPYLQSIIDKSPRMYDIVYGRMTKWYFQIDPVFWLSGNNKEL
QDFIHMLTEFSKVIVKHRREKLKTAGHTEQLMNSEEDESINTTSLVIDRFILSQELNDYEL

FKETFTIFTSSQEASAKIASYVILMMAYHPDCQEKLYKEIKQVIGDADRPVTGEDLKNMP
YLEMVFKETIRLFPIGALLQRTVAQDTVIKSGVLPGGCSLVVPIYHLHRDPRYWENPEDF
DPERFNPENTKLRHPNCYLPFSLGPMDCLGRYFGTKLIKTRVLREFEVSSPQTYQDL
KVAICISVSLNGYPVIFKKRQM