6-7-2018

Comparative CYP-omic analysis between the DDT-susceptible and -resistant Drosophila melanogaster strains 91-C and 91-R

Keon Mook Seong
Michigan State University, seongkeo@msu.edu

Brad S. Coates
Michigan State University

May R. Berenbaum
University of Illinois at Urbana-Champaign

John M. Clark
University of Massachusetts Amherst

Barry R. Pittendrigh
Michigan State University

Research Article

Received: 31 October 2017 Revised: 28 March 2018 Accepted article published: 15 April 2018 Published online in Wiley Online Library: 7 June 2018

(wileyonlinelibrary.com) DOI 10.1002/ps.4936

Comparative CYP-omic analysis between the DDT-susceptible and -resistant *Drosophila melanogaster* strains *91-C* and *91-R*

Keon Mook Seong,[a*] Brad S Coates,[b] May R Berenbaum,[c] John M Clark[d] and Barry R Pittendrigh[a]

Abstract

BACKGROUND: Cytochrome P450 monooxygenases (P450s) are involved in the biosynthesis of endogenous intracellular compounds and the metabolism of xenobiotics, including chemical insecticides. We investigated the structural and expression level variance across all P450 genes with respect to the evolution of insecticide resistance under multigenerational dichlorodiphenyltrichloroethane (DDT) selection.

RESULTS: RNA-sequencing (RNA-seq) and reverse transcriptase–quantitative polymerase chain reaction (RT-qPCR) indicated that the transcript levels of seven P450 genes were significantly up-regulated and three P450 genes were down-regulated in the DDT-resistant strain *91-R*, as compared to the control strain *91-C*. The overexpression of *Cyp6g1* was associated with the presence of an *Accord* and an *HMS-Beagle* element insertion in the 5′ upstream region in conjunction with copy number variation in the *91-R* strain, but not in the *91-C* strain. A total of 122 (50.2%) fixed nonsynonymous (amino acid-changing) mutations were found between *91-C* and *91-R*, and 20 (8.2%) resulted in amino acid changes within functional domains. Three P450 proteins were truncated as a result of premature stop codons and fixed between strains.

CONCLUSION: Our results demonstrate that a combination of changes in P450 protein-coding regions and transcript levels are possibly associated with DDT resistance, and thereby suggest that selection for variant function may occur within this gene family in response to chronic DDT exposure.

Supporting information may be found in the online version of this article.

Keywords: cytochrome P450 monooxygeneases; *Drosophila melanogaster*; duplication; genome-wide evolution; loss-of-function mutation; transposable elements

1 INTRODUCTION

Insects have evolved various adaptive strategies, including the diversification of their detoxification mechanisms in response to different classes of chemical insecticides.[1] The subsequent loss of effective control has led to challenges in controlling these insects with regard to the production of agriculturally important crop plants as well as the protection of human health. For example, recurrent exposure of the Western corn rootworm, *Diabrotica virgifera virgifera*, to carbamate, organophosphate, and organochlorine insecticides in agricultural fields imposed a selective pressure that led to the evolution of resistant populations.[2,3] Analogously, insecticide resistance within insect populations that are vectors of disease pose public health risks to humans.[4] Multigenerational xenobiotic selection pressures within insect populations can lead to adaptation via heritable changes in detoxification pathways. Increased survivorship often occurs among individuals in response to environmental toxins,[5,6] and arises by the action of directional selection to alter gene function by changes to gene expression patterns or amino acid sequences. Understanding aspects of genomic selection for insecticide resistance will aid in molecular screening methods for detecting field-evolved resistance as well as in methods devised to delay or circumvent the fixation of causal alleles within insect populations.[7,8]

Cytochrome P450 genes constitute a large gene family and are found in virtually all living organisms, including bacteria, fungi, plants, and animals.[9,10] P450 proteins are involved in endogenous and exogenous metabolism, including the oxidative degradation of many xenobiotic compounds.[11] There is great diversity of P450 genes within and between sequenced genomes, where

* *Correspondence to: KM Seong, Department of Entomology, Michigan State University, East Lansing, MI 48824, USA. E-mail: seongkeo@msu.edu*

a *Department of Entomology, Michigan State University, East Lansing, MI, USA*

b *USDA-ARS, Corn Insects & Crop Genetics Research Unit, Iowa State University, Ames, IA, USA*

c *Department of Entomology, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

d *Department of Veterinary & Animal Science, University of Massachusetts, Amherst, MA, USA*

the number in this large multi-gene family ranges from 37 in the human body louse, *Pediculus humanus humanus*,[12] to 170 in the genome of *Culex quinquefasciatus*.[13] This high number of P450 genes within insect genomes inspired the term 'P450 family blooming'.[14] The *Drosophila melanogaster* (*D. melanogaster*) reference genome contains 90 P450 genes and five to seven pseudogenes.[15,16] The P450 gene superfamily has diversified into 25 different families within the *D. melanogaster* genome grouped into four large clades (the CYP2, CYP3, CYP4, and mitochondrial clades) across P450 families,[17] which have evolved through copy number expansions that mostly occurred within the *Cyp4* and *Cyp6* families within the CYP3 clade.[15] Gene duplication and loss have been recognized as the predominant evolutionary factors in determining levels of P450 gene family complexity and diversity.[18,19] This pattern is highlighted by the ancestral *Cyp12d1/2/3* duplication and diversification that subsequently led to the loss of *Cyp12d2* in *D. melanogaster*,[20] and also includes 114 P450 gene copy losses and 74 gains across 12 *Drosophila* species.[16] The level of redundancy in any given genome may allow leeway in the evolutionary trajectories of P450 gene copies (orthologs and paralogs) to shift between varying structural, signaling, and metabolic roles.[21] The evolutionary diversity of the different P450s implicated within insecticide resistance mechanisms across a range of phylogenetically diverse insects has been used to highlight their importance in the adaptation of insect populations to xenobiotics.[22]

Previous studies have focused on the role of P450s in the detoxification of intracellular and xenobiotic compounds.[23–26] Several P450s are implicated in the degradation of caffeine,[27] and the up-regulation of *Cyp4e3* in the Malpighian tubules is associated with permethrin resistance and surmised to impact cellular response to stress.[28] To date, the functional role of the *Cyp6g1* gene has been particularly focused upon in order to identify the mechanism by which this P450 metabolizes insecticides in *D. melanogaster*. Low levels of DDT insecticide resistance in field populations of *D. melanogaster* are influenced by the constitutive or inducible up-regulation of *Cyp6g1* and *Cyp12d1*,[29,30] and microarray analysis indicated that DDT resistance is associated with over-transcription of *Cyp6g1* in *D. melanogaster*.[31] Schmidt *et al.* also found that *Cyp6g1* expression was influenced by copy number variation and an *Accord* transposable element insertion upstream of the *Cyp6g1* gene.[32] Furthermore, six alleles of *Cyp6g1* were identified on the basis of the duplication events and insertion of different types of transposable elements. These six alleles were strongly correlated with DDT resistance level.[32] Recent work by Battlay *et al.* has also shown that *Cyp6g1* is involved in organophosphate resistance through transgenically overexpressing the *Accord* element from a resistant strain (Hikone-R) in *D. melanogaster*.[33] Most recently, two P450s, *Cyp6g1* and *Cyp6g2*, have been shown to contribute to the metabolization of imidacloprid and are involved in conferring resistance to this insecticide in *D. melanogaster*.[34]

Additionally, multiple P450s associated with DDT resistance were over-transcribed in the highly DDT-resistant *D. melanogaster* strain as compared to DDT-susceptible strain, suggesting that resistance may be in part attributable to differential expression of multiple P450s.[35,36] Retention of a functional Nrf2/Maf (NF-E2-related factor 2/Muscle aponeurosis fibromatosis) transcription factor binding site was previously linked to the over-expression of *Cyp6a2* associated with DDT resistance in *91-R*,[37] as was *trans*-regulatory control by the *D. melanogaster* CNC-bZIP (Cap'n'collar-basic leucine zipper) transcription factor ortholog to Nrf2 that regulates the transcriptional response of several genes encoding detoxification enzymes that contain an upstream antioxidant response element (ARE),[38,39] including *Cyp6a2* and *Cyp6a8*.[40] As mutations in the *D. melanogaster* estrogen-related receptor (ERR) caused changes in the expression of P450s, there is the possibility that components of signal transduction pathways might also play a role in the evolution of DDT insecticide resistance mechanisms.[41]

Lamb *et al.* designated the collective set of P450 genes within the genome of an organism a 'CYPome'.[42] As many previous studies have focused on the role of individual P450 genes in the xenobiotic metabolism of insects,[43] the structural and function responses of the CYPome to insecticide selection are not yet fully understood. With reductions in the cost of sequencing whole genomes, the opportunity to take a systems approach to investigating genes associated with insecticide resistance becomes increasingly more feasible, facilitating comparative analyses of the structural and functional variation among P450 superfamily members. We introduce the term 'comparative CYP-omic analysis' to describe the study of the similarities and differences in genomes, transcriptomes, proteomes, or a combination thereof, between individuals, strains, populations, species, or higher level taxa, as they relate to cytochrome P450s. Within the context of this study, we used a comparative CYP-omics approach to assess gene polymorphisms and transcriptional differences toward a better understanding of the evolutionary processes associated with prolonged DDT selection under laboratory conditions.

Two *D. melanogaster* strains, the highly resistant DDT-selected strain *91-R* and its corresponding control, *91-C*, have experienced differential exposures to DDT in the laboratory for more than six decades.[44–46] These two strains were derived from a known common origin, but subsequent differences in the DDT selection schemes have provided a valuable resource for comparative analysis within the context of a model for studying xenobiotic resistance mechanisms on a genomic-wide scale. Indeed, these two strains formed the basis of a comparative genomic study that identified genome regions with significant variance in nucleotide fixation putatively attributable to the effects of differential selection, which included one region encoding *Cyp4g1*[47,48] and ATP-binding cassette transporters (ABC transporters) open reading frames.[47,49] Previously, Brandt *et al.* showed that the P450 inhibitor piperonyl butoxide (PBO) drastically reduces (but does not eliminate) DDT resistance in *D. melanogaster*, providing further evidence that differential expression of P450 genes is a component of the DDT resistance phenotype.[29] The current study expands on prior work performed in the DDT-resistant strain *91-R* by investigating the structural and expression level (functional) variance across all P450 gene family members compared with that of its control strain, *91-C*. These studies associating specific P450 variants with DDT resistance are important for the understanding of evolutionary processes involved in genome-wide responses to chronic xenobiotic exposure.

2 MATERIALS AND METHODS

2.1 *Drosophila melanogaster* strains

The *D. melanogaster* strains DDT-susceptible *91-C* and DDT-resistant *91-R* were created > 60 years ago.[46] The *91-R* strain has been shown to be ~250-fold more resistant to DDT than the susceptible *91-C* strain through the use of contact bioassays.[30] The *91-R* strain has been continually selected by maintaining the flies in a colony bottle in the presence of a 150-mg DDT filter paper disk while *91-C* has been maintained without any exposure to DDT.

To estimate the frequency of large deletions in the *Cyp4p1* and *Cyp4p2* regions among *D. melanogaster* strains, an additional

54 wild-type lines were obtained from the Bloomington *Drosophila* Stock Center (Bloomington, IN, USA) to genotype a 3.3-kb nucleotide deletion for *Cyp4p1* and *Cyp4p2*. Strains included: *Amherst 3*, *BER 2*, *Berlin K*, *BOG 2*, *BOG 3*, *Canton-S*, *Canton-S-iso2B*, *CO 4*, *CO 7*, *Crimea*, *EV*, *Florida-9*, *Harwich*, *Hikone-A-S*, *Hikone-A-W*, *Hikone-R*, *KSA 4*, *Lausanne-S*, *MO 1*, *MWA 1*, *NO 1*, *Oregon-R-P2*, *Oregon-R-SNPiso2*, *Oregon-R-SNPiso3*, *Oregon-R*, *Oregon-R-modENCODE*, *Oregon-R-S*, $pi_2 < P >$, *PYR 3*, *RC 1*, *Reids 1*, *Reids 2*, *Reids 3*, *RVC 2*, *RVC 4*, *Samarkand*, *Swedish-C*, *TW 1*, *TW 2*, *TW 3*, *Urbana-S*, *VAG 2*, *VAG3*, *Wild 1A*, *Wild 2A*, *Wild 1B*, *Wild 3B*, *Wild 5A*, *Wild 5B*, *Wild 5C*, *Wild 10E*, *Wild 11C*, *Wild 11D*, and w^{1118}. The UCSC Genome Browser track for Drosophila Genetic Reference Panel (DGRP) Freeze 2 database was used to investigate all insertions/deletions (indels) including *Cyp4p1* and *Cyp4p2* deletions (http://dgrp2.gnets.ncsu.edu/).[50]

2.2 Illumina sequencing and variation of P450 gene expression between strains *91-C* and *91-R*

All RNA-sequencing (RNA-seq) read data were previously generated from *91-C* and *91-R* in triplicate and these data were submitted to the National Center for Biotechnology Information (NCBI) Short Read Archive (SRA) database (accession number: SRX2611754-SRX2611759) previously.[51] Specifically, a total of 1000 5-day-old *D. melanogaster* flies (500 males and 500 females) were pooled for each of three biological replicates per strain. Total RNA was extracted from each pool ($n = 6$), not exposed to DDT within that generation, using the Qiagen RNeasy Maxi Kit according to the manufacturer's instructions (Qiagen, Valencia, CA, USA). Bidirectional Illumina RNA-seq libraries for transcriptome analysis were sequenced using Illumina HiSeq™ 2500 at the W. M. Keck Center for Comparative and Functional Genomics (University of Illinois Urbana-Champaign, Urbana, IL). With 160-bp read length of paired-end reads. All raw read data were imported to the CLC Genomic Workbench 9.5 (Qiagen, Valencia, CA, USA) and all reads were processed to remove low-quality and adaptor sequences to obtain clean reads for analysis. The trimmed reads of each replicates were mapped against annotated *D. melanogaster* genes and transcripts extracted from *D. melanogaster* genome assembly release 6.07 (file dmel-all-chromosome-r6.07.fasta downloaded from Flybase.org).[52] The differential expression of P450 transcripts was evaluated between technical replicates of *91-C* and *91-R*. Specifically, the number of reads per kilobase of transcript per million reads (RPKM) mapped against the annotated *D. melanogaster* genome assembly v.6.07 was used as a proxy for gene expression. The false discovery rate (FDR) method was used to determine the threshold of the *P*-value corresponding to differential expression,[53] and P450 genes with an FDR ≤ 0.05 and $\log_2$ fold-change ≥ 1.0 were considered as differentially expressed.

RT-qPCR was applied to validate the expression of putatively differentially expressed P450 transcripts between strains *91-C* and *91-R*. Thirty adult flies (15 males and 15 females) were collected per replicate and three biological replicates were analyzed. In order to determine constitutive gene expression, total RNA was extracted from each replicate ($n = 6$), not exposed to DDT within that generation, using the Qiagen RNeasy Mini Kit according to the manufacturer's instructions (Qiagen, Valencia, CA, USA) with resulting extracts treated with DNAse I (Qiagen, Valencia, CA, USA) to remove contaminating genomic DNA.

The cDNA first strand was synthesized from 1.0 μg of total RNA using the SuperScript III First-Strand Synthesis System (Invitrogen, Carlsbad, CA, USA). Subsequent RT-qPCR was performed using a StepOnePlus Real-Time PCR system (Applied Biosystems Inc., Foster City, CA, USA) according to the manufacturer's instructions. The ribosomal protein 49 (rp49), ribosomal protein L32 (rpl32) and *tubulin* genes were used as an internal control (primers used for RT-qPCR analysis are listed in Supporting Information Table S1).[54] Primer efficiencies were determined from dilution curves using the formula: $E = 10^{-1/\text{slope}}$,[55,56] with the slope determined by the StepOnePlus Real-Time PCR software (Table S1). Melting curve analysis and gel electrophoresis of the PCR products were also conducted to verify that a single gene-specific product was produced. Relative expression levels and fold change were analyzed using the comparative Ct method ($2^{-\Delta\Delta Ct}$).[57] Statistical analysis was performed using Student's *t*-test to compare the results between the two strains by XLSTAT (XLSTAT 2008; Addinsoft, New York, NY, USA).

2.3 Transposable elements (TEs) and duplication analysis in *Cyp6g1*

Genomic DNA was extracted from ten adult flies (five males and five females) for the *91-C* and *91-R* strains, using the Qiagen DNeasy Mini Kit according to the manufacturer's instructions (Qiagen, Valencia, CA, USA). The diagnostic PCR for the absence or presence of the TE (*P*-element, *Accord* and *HMS-Beagle*) insertion was carried out as described in Schmidt *et al.*[32] using the primer sets, that flank the *Accord* insertion sites and the primer that anneals to the *HMS-Beagle* sequence. Thus, the *Accord* insertion produce a single band of 604 bp, while a single band of the 114-bp fragment is expected for no *Accord* insertion. Moreover, if only the *Accord* insertion is present in *Cyp6g1*, the forward primer that anneals to the *HMS-Beagle* sequence fails to produce an amplified fragment. For duplication analysis, primers were designed as described by Emerson *et al.*,[58] with a set of primers being designed to face inversely within the *Cyp6g1* gene to produce an amplicon across the tandem duplication junction. In the presence of tandem duplications, the forward and reverse tandem duplication primers are expected to produce a band > 2 kb for *Cyp6g1*; if only a single copy of the gene is present, the primers fail to produce an amplified product. All the primers used for PCR are summarized in Table S1.

2.4 Detection of variants in P450s between strains *91-C* and *91-R*

The single nucleotide polymorphisms (SNPs) and deletion insertion polymorphisms (DIPs) were investigated within P450 gene regions in order to identify potential structural mutations and any association with DDT resistance by comparing the *91-C* and *91-R* strains. Putative sequence variations were detected by separately mapping RNA-seq read data for *91-C* and *91-R* to the *D. melanogaster* P450 reference genes obtained from Flybase.org using the CLC Genomic Workbench 9.5 (Qiagen, Valencia, CA, USA) with the 'Map Reads to Reference' tool (minimum length fraction 0.9, minimum similarity fraction 0.8, insertion/deletion cost = 3, and mismatch cost = 3). Mapping data were used to construct consensus sequences from the 87 P450 genes annotated in assembly v. 6.07 (noise threshold of 0.1 and minimum nucleotide count of 20). The consensus base frequency of mapped reads compared to the reference sequence was used to call fixed ($P = 1.0$) or unfixed ($P < 1.0$) alleles within the depth of reads. The derived amino acid sequences from each P450 gene sequence of *91-C* and *91-R* were used in a multiple sequence alignment with the corresponding homologous sequence in strain *Canton-S* using the Alignments and Tree tool of the CLC Genomic Workbench. Putative substrate recognition sites and other protein signature motifs were identified based on domains defined by Gotoh.[59]

Table 1. Annotations for differentially expressed P450 transcripts between the DDT-resistant strain *91-R* and DDT-susceptible strain *91-C*

Gene symbol	Functional annotation	Annotation symbol	Log_2 fold change[a]	FDR[b]
Cyp6a8	Cytochrome P450-6a8	CG10248	5.59	7.72E-92
Cyp6g1	Cytochrome P450-6g1	CG8453	4.08	1.14E-23
Cyp6w1	Cytochrome P450-6w1	CG8345	3.56	6.73E-94
Cyp4p1	Cytochrome P450-4p1	CG10842	5.02	1.19E-27
Cyp6g2	Cytochrome P450-6g2	CG8859	2.33	7.06E-03
Cyp6a2	Cytochrome P450-6a2	CG9438	1.71	1.60E-12
Cyp12a4	Cytochrome P450-12a4	CG6042	1.02	3.54E-07
Cyp4e2	Cytochrome P450-4e2	CG2060	0.81	1.80E-03
Cyp28d1	Cytochrome P450-28d1	CG10833	0.75	8.71E-03
Cyp9f2	Cytochrome P450-9f2	CG11466	−0.49	7.85E-03
Cyp6a23	Cytochrome P450-6a23	CG10242	−0.89	4.26E-02
Cyp9b1	Cytochrome P450-9b1	CG4485	−1.29	2.12E-02
Cyp4e3	Cytochrome P450-4e3	CG4105	−3.18	6.72E-18

[a] Fold change was calculated as $\log_2$ *91-C/91-R*.
[b] FDR, false discovery rate. Differentially expressed genes were identified at the threshold [FDR < 0.05 and $\log_2$(fold change) ≥ |1.0|] of *91-C/91-R*.

Illumina genome sequencing data sets (SRA accession number SRP041176) from *91-C* and *91-R* were generated in order to determine the presence of known or unique allelic variants putatively affecting the evolution of DDT resistance.[48] To identify functional polymorphisms in the *cis*-regulatory ARE and Nrf2 transcription factor binding sites that regulate the transcription of P450s with upstream AREs, sequences were searched for within approximately 300-bp upstream gene regions obtained from FlyBase.org. Genome data sets of *91-C* and *91-R* were mapped to the upstream regulatory regions of significantly up- and down-regulated P450s using the 'Map Reads to Reference' tool from CLC genomic workbench (Qiagen, Valencia, CA, USA). The resultant positional SNP and DIP data were provided in table format. AREs were identified within the upstream sequences of differentially expressed P450 genes using the online query tool at the JASPAR database (http://jaspar.genereg.net) with the Nfe2l2 matrix model (ID MA150.2).[60]

2.5 Allele frequencies of a block deletion in *Cyp4p1* and *Cyp4p2*

The *91-C* and *91-R* genome data sets (SRA accession number SRP041176) were mapped to the reference genome of *D. melanogaster* for *Cyp4p1* and *Cyp4p2* to determine the presence of deletions using the 'Map Reads to Reference' tool from the CLC genomic workbench (Qiagen, Valencia, CA, USA). PCR amplification of *Cyp4p1* and *Cyp4p2* and Sanger sequencing were performed in order to verify deleted regions predicted from the original mapping. Specifically, the gene-specific PCR primers were designed to target exon 1 of *Cyp4p2* and exon 4 of *Cyp4p1* and were used to amplify these regions from 56 *D. melanogaster* strains (Table S1).

3 RESULTS

3.1 Variation in P450 gene expression between *91-C* and *91-R*

Analysis of RNA-seq data revealed that the expression levels of 12 P450 genes were significantly different between *91-C* and *91-R* (cutoff FDR ≤ 0.05 and $\log_2$ fold-change ≥ |1|; Table S2). Among the 12 transcripts predicted by RNA-seq to be differentially expressed (Table 1), RT-qPCR showed that transcript levels of ten of these P450 genes were significantly different,

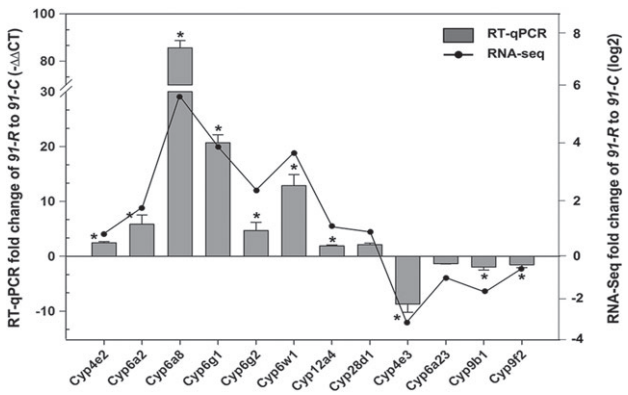

Figure 1. RT-qPCR validation of 12 P450 transcripts putatively differentially expressed between DDT-resistant *91-R* and the control strain *91-C* at a low-stringency threshold [FDR < 0.05 and $\log_2$(fold change) ≥ |1.0|]. The left y-axis indicates the fold change in *91-R* relative to *91-C* by RT-qPCR, and the right y-axis indicates the log2Ratio of *91-R/91-C* by RNA-seq.

with seven P450 transcripts being up-regulated (*Cyp4e2*, *Cyp6a2*, *Cyp6a8*, *Cyp6g1*, *Cyp6g2*, *Cyp6w1*, and *Cyp12a4*), and three transcripts down-regulated (*Cyp4e3*, *Cyp9b1*, and *Cyp9f2*) in *91-R versus 91-C* (Fig. 1). Specifically, RT-qPCR analyses showed that the relative transcript levels were increased 85.69-, 20.7-, 12.89-, 5.82-, 4.69-, 2.46-, and 1.88-fold, respectively, for *Cyp6a8*, *Cyp6g1*, *Cyp6w1*, *Cyp6a2*, *Cyp6g2*, *Cyp4e2*, and *Cyp12a4*, in the *91-R* strain as compared to *91*-C. Analogously, the relative levels of *Cyp4e3*, *Cyp9b1* and *Cyp9f2* were respectively 8.72-, 1.94-, and 1.54-fold lower in *91-R* (Fig. 1). Among the P450 genes up-regulated in *91-R*, five were in the *Cyp6* gene family, which belongs to the CYP3 clade (*Cyp6a2*, *Cyp6a8*, *Cyp6g1*, *Cyp6g2* and *Cyp6w1*). The up-regulated *Cyp4e2* and *Cyp12a4* genes belong to the CYP4 and mitochondrial clades, respectively. Analogously, the three down-regulated transcripts (*Cyp4e3*, *Cyp9b1*, and *Cyp9f2*) in *91-R* belong to the two larger P450 clades, CYP4 and CYP3. Expression of three reference genes was analyzed by RT-qPCR to verify that all three reference genes were consistently expressed among all samples. Reference and target primers exhibited comparable efficiencies as determined using a dilution series of target cDNA. We compared the expression levels of *rpl32* and *tubulin* using *rp49*

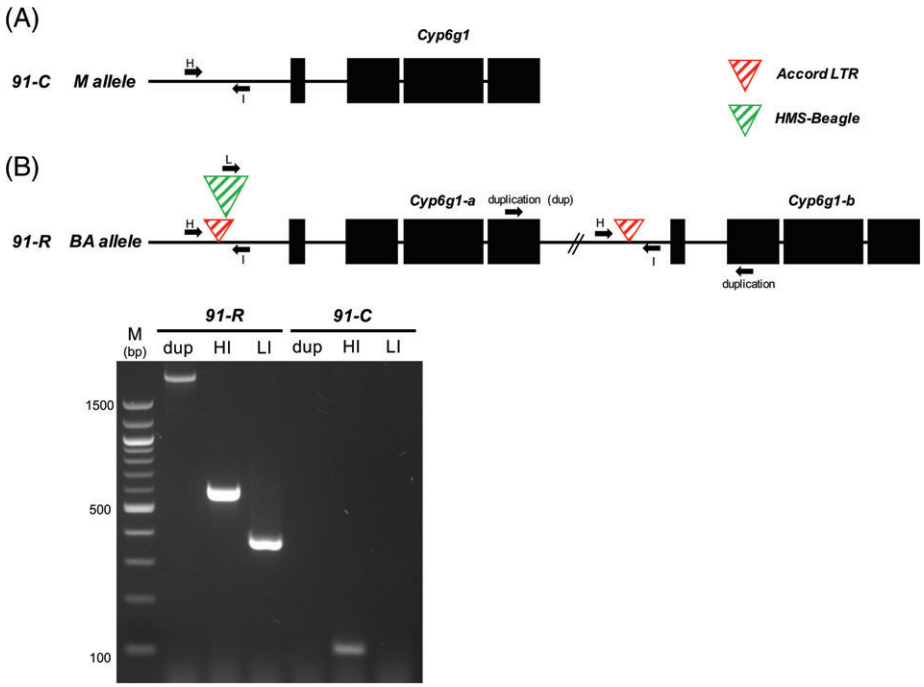

Figure 2. The molecular structure of the *Cyp6g1* locus of the *91-C* and *91-R* strains. (A) The *M* allele in the *91-C* strain has a single copy without TE insertions. The *BA* allele in the *91-R* strain has two copies with an *Accord* and a *HMS-Beagle* element insertion in the upstream region. Exons are depicted as filled black boxes, and introns and 5′/3′ flanking sequences as black lines. The *Accord* and *HMS-Beagle* elements inserted into the 5′ flanking region are depicted as inverted triangles. The annealing positions of the diagnostic PCR primers are shown as horizontal arrows with a single letter. Note that primer L anneals to the *HMS-Beagle* sequence whereas primers H and I flank the *Accord* element insertion sites. (B) A gel demonstrating the diagnostic PCRs is shown. The tandem duplication primer set only produces an amplification fragment when the duplication is present.

as a reference between *91-C* and *91-R* strains; the ΔΔCt for *rpl32* is −0.1809 with a *P*-value of 0.31; for *tubulin*, the ΔΔCT is 0.1401 and the *P*-value is 0.17. We observed that all three control genes were equivalently expressed among all samples within statistical limits. Therefore, we confirmed that the normalization with *rp49* was a suitable way to analyze differences in the expression levels of all P450 genes between the *91-C* and *91-R* strains.

3.2 Copy number variation (CNV) of differentially expressed P450s in strains *91-C* and *91-R*

As previously reported by Schmidt *et al.*[32] and Harrop *et al.*,[61] *Cyp6g1* was tandemly duplicated in the *D. melanogaster* genome and the duplication event of *Cyp6g1* is responsible for overexpression and increased resistance to DDT. We used the mapping data of whole-genome sequencing generated by Steele *et al.*[48] to determine the CNV for differentially expressed P450s in *91-C* and *91-R*. Among 12 differentially expressed P450s, high variation in sequencing coverage between *91-C* and *91-R* was found in *Cyp6g1* and *Cyp6g2* (Table S3), supporting the theory that there may be CNV for these two P450s. Therefore, the CNV for *Cyp6g1* and *Cyp6g2* was verified using a PCR amplification. The *91-R* strain produced a single 2.5-kb amplicon for *Cyp6g1* (Fig. 2), supporting the hypothesis that tandem duplication of the *Cyp6g1* gene is present in the *91-R* strain. For the *91-C* strain, however, amplification with the tandem duplication primers failed to generate an amplified product, demonstrating that the *91-C* strain does not contain tandem duplications of the *Cyp6g1* gene. For *Cyp6g2*, although a 1.5-kb fragment was amplified with primers covering the tandem duplication junction, we sequenced an amplified fragment, which revealed double peaks showing both *Cyp6g1-Cyp6g2* alleles on the sequence chromatograms (data not shown). We inferred that this may be attributable to partial chimeric repeats of *Cyp6g1-Cyp6g2* which were identified in Schmidt *et al.*[32]

3.3 *Cyp6g1* allele identification in strains *91-C* and *91-R*

Previously, six different alleles of *Cyp6g1* (*M*, *A*, *AA*, *BA*, *BP*, and *BP*Δ) were described and defined on the basis of the presence of the duplication and the insertion of different types of TE.[32] PCR analysis for the presence of insertion of a partial *P*-element and the *Accord* and *HMS-Beagle* elements in the 5′ untranslated region (UTR) of the *Cyp6g1* locus and a duplication event using genomic DNA extracted from the *91-C* and *91-R* strains showed that the *91-R* strain has not only the *Accord* element insertion in the 5′ region of the *Cyp6g1* locus but also insertion of the *HMS-Beagle* element in one of the *Accord* elements. However, the partial *P*-element was not detected within the *Accord* element in the *91-R* strain (Fig. 2). In contrast, the *91-C* strain revealed no duplication and TE insertions of the *Cyp6g1* locus. These observations suggest that the *91-R* strain has a *BA* haplotype, in which both *Cyp6g1* and *Accord* element are duplicated, whereas *91-C* has an *M* haplotype which is an ancestral allele without duplication and TE insertion events (Fig. 2).

3.4 Variant detection between P450s in strains *91-C* and *91-R*: nonsynonymous changes

Mutations that were fixed differently between *91-C* and *91-R* can be considered as candidates for involvement in DDT resistance, especially when predicted to have a functional consequence. As a consequence of the short reads used for re-sequencing, the phase (haplotype of each allele) was not determined. A total of 243 nucleotide nonsynonymous changes were predicted in the open reading frames (ORFs) of 73 of the 87 P450s across the

D. melanogaster genome, of which nonsynonymous changes were found in all of these genes when comparing between *91-C* and *91-R* (Tables 2 and S4). Specifically, no amino acid sequence differences were found in 14 out of the 87 derived P450 sequences (16.1%; CYP18A1, CYP305A1, CYP306A1, and CYP307A2 for the CYP2 clade; CYP6A20, CYP6G1, CYP6D5, and CYP28C1 for the CYP3 clade; CYP4AD1, CYP4D2, CYP4G15, and CYP318A1 for the CYP4 clade; CYP12C1 and CYP49A1 for the mitochondrial clade). Out of 87 P450 genes, 63 (72.4%) and 10 (11.5%) were found to have nonsynonymous SNPs and DIPs, respectively. A total of 243 nonsynonymous mutations were found within the 73 genes with polymorphisms; 233 SNPs (95.9%) and 10 DIPs (4.1%), giving an estimated 2.8 mutations per gene. Approximately 50.2% of mutations were fixed differently between *91-C* and *91-R* (homozygous within strain), whereas the remaining 49.8% were segregating (heterozygous) within both strains.

The degree of fixation for predicted nonsynonymous mutations was evaluated with respect to any association with DDT resistance. For example, among the seven significantly up-regulated P450 genes in *91-R*, five nonsynonymous SNPs were found in *Cyp6w1* at bp locations 106, 497, 616, 901, and 1109, and respectively led to amino acid changes V36F, T166S, Y206H, I301V, and V370A (Table 2). Such mutations were defined as fixed and may be putatively associated with DDT resistance in *91-R*. In contrast,

Table 2. Nonsynonymous nucleotide and associated amino acid changes for differentially expressed P450 genes in the DDT-resistant strain *91-R* compared with the control strain *91-C*

Gene	Gene location	CDS length	SNP position	Allelic variant (frequency)		Amino acid change	Fixation
				91-C	*91-R*		
Overexpressed genes							
Cyp6w1		1545	106	G (1.0)	T (1.0)	V36F	Fixed
			497	C (1.0)	G (1.0)	T166S	Fixed
	2R: 6173684-6176195		616	T (1.0)	C (1.0)	Y206H	Fixed
			901	A (1.0)	G (1.0)	I301V[a]	Fixed
			1109	T (1.0)	C (1.0)	V370A[a]	Fixed
Cyp6a8		1522	136	T (1.0)	A (1.0)	L46M	Fixed
	2R: 14887171-14889010		481	G (1.0)	A (1.0)	A161T	Fixed
			1150	A (1.0)	G (1.0)	T384A[a]	Fixed
Cyp6g2		1560	1169	C (1.0)	C(0.77)/T(0.23)	T390I[a]	Unfixed
	2R: 12188338-12190377		1267	C (1.0)	C(0.66)/T(0.34)	P423S	Unfixed
			1325	G (1.0)	A(0.18)/G(0.82)	R442H	Unfixed
Cyp28d1		1509	720	T (1.0)	A(0.42)/T(0.58)	N240K[a]	Unfixed
	2L: 5210460-5212445		779	G (1.0)	A(0.41)/G(0.59)	R260Q	Unfixed
			812	T (1.0)	C(0.4)/T(0.59)	V271A	Unfixed
Underexpressed genes							
Cyp4e3	2L: 9747839-9750071	1582	358 359	T (1.0) A (1.0)	A(0.18)/T(0.82) A(0.82)/C(0.18)	Y120T[a]	Unfixed
			454 456	G (1.0) A (1.0)	A(0.17)/G(0.83) A(0.83)/C(0.17)	A152T	Unfixed
			596	G (1.0)	C (1.0)	C199S	Fixed
			766	G (1.0)	G(0.28)/T(0.72)	D256Y	Unfixed
			791	T (1.0)	A(0.69)/T(0.31)	M264K	Unfixed
			797	G (1.0)	A(0.68)/G(0.32)	G266D	Unfixed
			799 801	T (1.0) T (1.0)	C(0.69)/T(0.31) C(0.69)/T(0.31)	Y267H	Unfixed
			878	T (1.0)	C (1.0)	I293T	Fixed
			1018	A (1.0)	A(0.34)/G(0.66)	I340V	Unfixed
			1027 1028	A (1.0) G (1.0)	A(0.34)/C(0.66) A(0.65)/G(0.35)	S343H	Unfixed
			1036	G (1.0)	A(0.66)/G(0.34)	D346N	Unfixed
			1169	A (1.0)	A(0.36)/G(0.64)	N390S	Unfixed
			1474	A (1.0)	A(0.31)/T(0.69)	M492L	Unfixed
Cyp9f2	3R: 12403340-12406028	1551	18	C(0.72)/G(0.28)	C (1.0)	L6F	Unfixed
			37	A(0.31)/G(0.69)	G (1.0)	T13A	Unfixed
			123	G(0.31)/T(0.69)	T (1.0)	L41F	Unfixed
			163	C(0.3)/T(0.7)	T (1.0)	L55F	Unfixed
			165	G(0.3)/T(0.7)	T (1.0)		
			202	A(0.72)/G(0.28)	A (1.0)	G68K	Unfixed
			203	A(0.72)/G(0.28)	A (1.0)		
			204	C(0.28)/G(0.72)	G (1.0)		
Gene on selective sweep region							
Cyp4g1[b]	X: 467057-469335	1671	43	T (1.0)	G (1.0)	S15A	Fixed

[a] The amino acid changes that are located in the SRS region.
[b] Three regions in the genome of *91-R* were previously defined as having undergone a minor selective sweep that reduced nucleotide diversities under the influence of DDT selection.[47]

the ORF for the up-regulated *Cyp6g2* gene in *91-R* had three predicted nonsynonymous sites at bp locations 1169, 1267, and 1325 and, respectively, resulted in amino acid changes T390I, P423S, and R442H. These three mutations were unfixed in *91-R* and fixed in *91-C*, and putatively defined as unfixed allele differences. Among the down-regulated P450 transcripts in *91-R*, *Cyp4e3* was predicted to have two mutations fixed (C199S and I293T) and 11 mutations unfixed (Y120T, A152T, D256Y, M264K, G266D, Y267H, I340V, S343H, D346N, N390S, and M492L; Table 2) between strains. *Cyp4g1* is not differentially expressed based on RNA-seq estimates and has a single nonsynonymous mutation fixed between strains (S15A). However, it was previously identified within genome regions under the influence of a selective sweep in *91-R*.[47] In contrast, the down-regulated *Cyp9f2* showed no predicted fixed nonsynonymous changes between strains.

Among the nonsynonymous substitutions fixed differently between *91-C* and *91-R*, 20 amino acid variations occurred within or adjacent to functional domains of 16 different P450 genes and 13 of these mutations were fixed between strains in ten P450 proteins (CYP4D20, CYP6A8, CYP6V1, CYP6W1, CYP302A1, CYP303A1, CYP304A1, CYP309A1, CYP313B1, and CYP315A1; Tables 2 and S4). Specifically, 12 of these fixed amino acid alterations between strains were within the functional substrate recognition site 1–6 (SRS 1–6) domains, and one was within the 'meander'-binding sequence (Tables 2 and S4). For example, the I301V alteration in CYP6W1 was detected within SRS 4 and near the oxygen-binding motif. However, the mutation giving rise to the valine residue at position 301 in the *91-R* CYP6W1 was identical to that in the susceptible strain *Canton-S*, suggesting that this mutation may not be associated with DDT resistance. Also, the V370A alteration was detected within SRS 5 and was very close to the 'meander'-binding sequence and was unique to *91-R*, suggesting that this amino acid variant may play a role in regulating substrate entry into the catalytic site and thus potentially impact function (Fig. 3). For CYP6G2, the three mutations giving rise to the T390I, P423S, and R442H alterations were detected as unfixed alleles in *91-R* when compared to the fixed alleles in both susceptible strains, *91-C* and *Canton-S*. The T390I alteration is located in the SRS 5 region and all three mutations giving rise to the T390I, P423S and R442H alterations were detected very close to the highly conserved 'meander region' motif or heme-binding motif (Fig. S1). These unfixed mutations in proximity to signature motifs may be associated with catalytic activity. For CYP6A8, furthermore, the three amino acid mutations L46M, A161T, and T384A were detected as fixed alleles between the *91-C* and *91-R* strains, and the T384A alteration was located in the SRS 5 region (Fig. S2). However, the three amino acid residues in *91-R* were identical to those in the susceptible strain *Canton-S*, suggesting that the three mutations may not be involved in DDT resistance.

In order to identify putative SNPs in the upstream regulatory regions, specifically within AREs that are known to direct P450 transcript levels via the Nrf2 transcription factor,[38,39] a combined mapping and computational transcription factor binding site prediction strategy was used. From this, we identified nine functional ARE sequences and potential Nrf2 transcription factor binding sites within the upstream *cis*-regulatory regions of nine differentially expressed P450 genes (*Cyp4e2*, *Cyp4e3*, *Cyp6a2*, *Cyp6a8*, *Cyp6a23*, *Cyp6g1*, *Cyp6g2*, *Cyp6w1*, and *Cyp12a4*; Table S5). The corresponding positions for predicted SNPs fixed between *91-C* and *91-R* did not coincide with any of the identified AREs.

3.5 Detection of variants between P450s in strains *91-C* and *91-R*: loss-of-function mutations

In total, nine P450 genes (six from *91-C*, two from *91-R*, and one from both *91-C* and *91-R*) revealed indels causing putative premature stop codons that were fixed or unfixed (Table 3). Of particular significance when compared to *91-R*, the control *91-C* strain had one fixed deleted region in *Cyp6a2*, and one fixed insertion within *Cyp316a1*. Analogously, one fixed deletion was predicted for *Cyp6a14* in *91-R*. Specifically, a single-base fixed insertion in exon 4 of the *Cyp316a1* allele was predicted to cause a protein truncation in the *91-C* strain. In the *91-R* strain, an allele homozygous for *Cyp6a14* contained multiple nucleotide deletion mutations, which caused a frameshift and subsequent generation of a premature stop codon. In all cases, these fixed indel mutations were predicted to have caused premature stop codons, ostensibly resulting in the absence of functional motifs. Specifically, these fixed indel mutations created a premature truncated protein, lacking SRS 6 for CYP6A2, SRS1-6, the oxygen-binding motif, the heme-binding motif, and the conserved putative 'meander'-binding sequence for CYP6A14, and the heme-binding motif, the conserved putative 'meander'-binding sequence, and SRS 6 for CYP316A1. Additionally, alleles at *Cyp4e1*, *Cyp4e2*, *Cyp9b1*, *Cyp12a4*, and *Cyp28d2* loci were predicted to contain insertions that led to frame shifts and protein truncations but were not fixed in both *91-C* and *91-R*. To expand our analysis of indel variations within *D. melanogaster* species, we utilized the recent DGRP Freeze 2 population, a set of inbred lines which provides polymorphism data from 205 *D. melanogaster* lines.[50] Analysis of DGRP genome data revealed that the fixed deletion (AGGG allele) within *Cyp6a14* that was found in the *91-R* strain was also detected in four DGRP lines (RAL 93, RAL 732, RAL 790, and RAL 892) out of 205 lines, suggesting that multiple nucleotides deletion events within *Cyp6a14* exist in the *D. melanogaster* populations. Consequently, the fixed deletion (AGGG allele) in *Cyp6a14* was the only fixed mutation found in the *91-R* and DGRP lines, whereas the fixed deletion (G allele) and insertion (G allele) within *Cyp6a2* and *Cyp316a1* were found in only the *91-C* and *91-R* strains and were not predicted in the DGRP lines.

3.6 Allele frequencies of a block deletion in *Cyp4p1* and *Cyp4p2*

Mapping of genome sequencing reads against the *D. melanogaster* genome assembly (v. 6.07) predicted a deleted region of approximately 3.3 kb that contained most of the exons of *Cyp4p1* and *Cyp4p2* but was only found in *91-C* (Fig. S3). Specifically, the *Cyp4* family forms a single clade with *Cyp4p1*, *Cyp4p2*, and *Cyp4p3* (Fig. 4a). Three novel deletions in the *Cyp4p1* and *Cyp4p2* alleles from *91-C* result in the omission of functional domains or regulatory upstream promoter elements in this strain (Fig. 4(b)). A 167-bp deletion started within intron 1 and spanned the region to exon 1 of *Cyp4p2* (Deletion 1), and a 12-bp deletion occurred in intron 3 of *Cyp4p1* (Deletion 3). Deletion 2 was estimated to be 3121 bp, and started within intron 2 of *Cyp4p2* and ended in exon 3 of *Cyp4p1* in *91-C*. This large deletion event eliminated most signature sequences of P450 proteins, such as the oxygen-binding motif, the heme-binding motif, and the putative 'meander'-binding sequence from the *Cyp4p1* and *Cyp4p2* transcripts.

The frequency of this large 3.3-kb deletion event was estimated using PCR amplification across 56 different *D. melanogaster* strains (including *91-C* and *91-R*). The *Cyp4p2* and *Cyp4p1* deletion region primers were designed to flank the large deletion region as shown in Fig. 4. In the presence of the large 3.3-kb deletion event, the

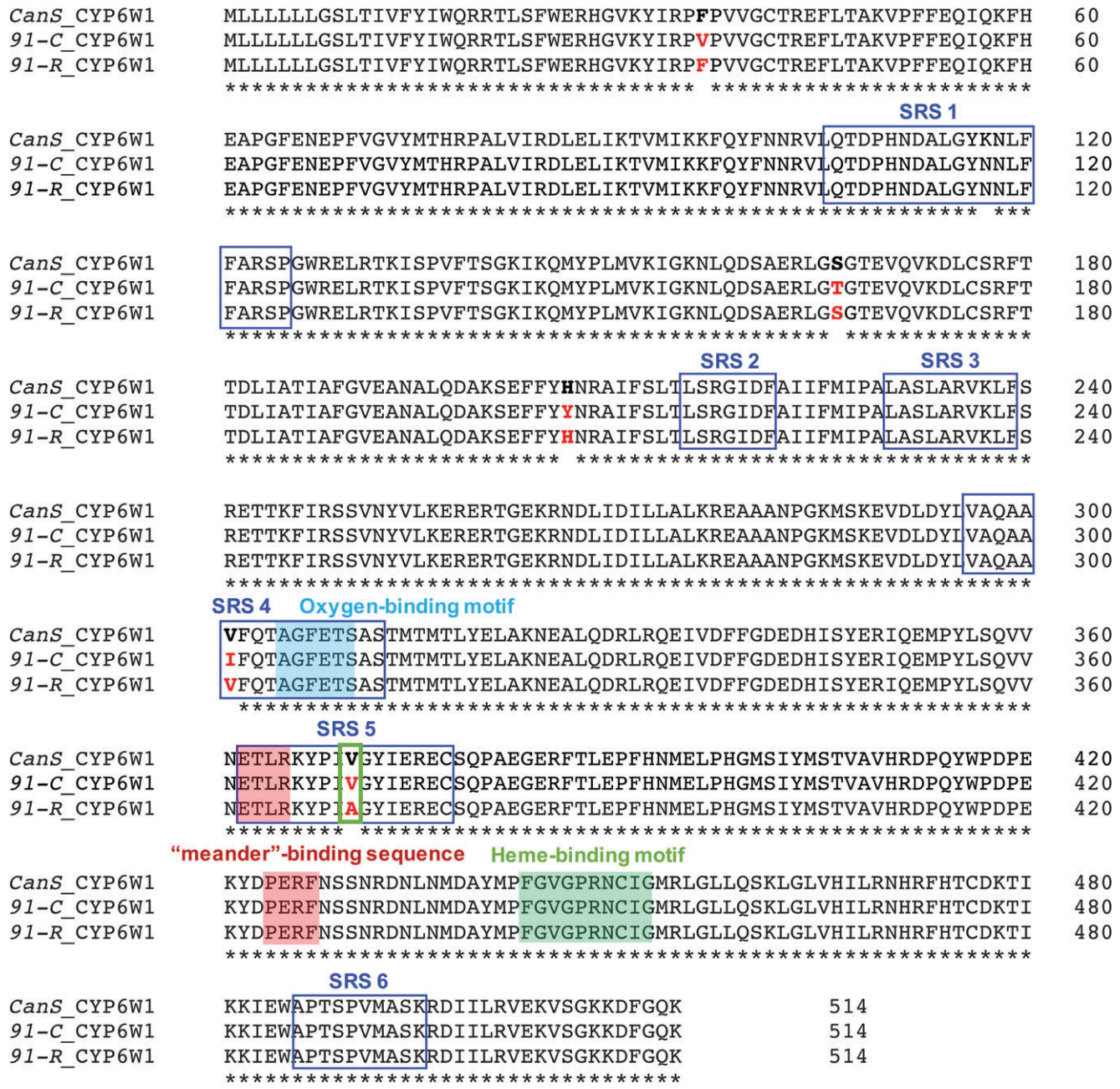

Figure 3. Alignment of deduced amino acid sequences for CYP6W1 from *D. melanogaster* strains *91-C*, *91-R*, and *Canton-S*. The five homozygous amino acid substitutions are indicated in red and changes fixed in *91-R* are indicated by the green box. The blue boxes indicate substrate recognition sites (SRSs) 1–6. Conserved domains common to P450 genes are shaded with different colors such as the oxygen-binding motif (blue), the heme-binding motif (green), and the conserved putative 'meander'-binding sequences EXXR and PXRF (red).

forward and reverse primers produced a 450-bp band; if no deletion region was present, the primer produced 3.3-kb amplified products. Analysis of all 56 strains with this primer sets revealed that 55 strains produced 3.3-kb bands, demonstrating that they do not contain the large deletion region of *Cyp4p1* and *Cyp4p2*. Only the *91-C* strain, however, generated 450-bp amplified products, supporting the hypothesis that this strain contains the large deletion region of *Cyp4p1* and *Cyp4p2* (Fig. S4). For 205 DGRP lines, furthermore, the large deletion region was entirely absent from the DGRP. Alternatively, however, a 104-bp deletion region within exon 2 of *Cyp4p1* was predicted for one DGRP line (RAL 405).

4 DISCUSSION

The current study found that multigenerational selection for survival in response to chronic high levels of DDT exposure, random genetic drift in relatively small closed laboratory populations, or a combination thereof has resulted in variation in expression and allelic diversity between P450 genes in *91-C* and *91-R*. Specifically, statistically significant differences in expression were documented in seven up-regulated and three down-regulated P450 transcripts in *91-R*, with *Cyp6* family members highly represented among up-regulated genes and *Cyp9* family members highly represented among down-regulated transcripts (Fig. 1). This finding agrees with previous evidence of significant changes

Table 3. Deletion/insertion polymorphisms predicted to cause premature stop codons (protein truncations) in *91-C* and *91-R*

Gene	Gene location	CDS length (bp)	Strain	Indel position (bp location)	Exon	Insertion/ Deletion	Allelic variant (frequency)	Fixation	Type of mutation	Introduce a premature stop codon
Cyp6a2	2R: 6779749-6781491	1520	*91-C*	1341	Exon 2	Deletion	G (1.0)	Fixed	Frame shift	Yes
Cyp6a14	2R: 8564187-8567202	1528	*91-R*	191–194	Exon 1	Deletion	AGGG (1.0)	Fixed	Frame shift	Yes
Cyp316a1	3L: 7491864-7493484	1455	*91-C*	1128	Exon 4	Insertion	G (1.0)	Fixed	Frame shift	Yes
			91-R				G(0.2)/gap(0.8)	Unfixed	Frame shift	Yes
Cyp28d2	2L: 5207267-5209345	1509	*91-R*	696 1070–1077	Exon 4	Insertion	G(0.2)/gap(0.8)	Unfixed	Frame shift	Yes
							TTTTAGAA(0.8)/ gap(0.2)			
Cyp4e1	2R: 8447451-8449745	1597	*91-C*	468	Exon 3	Insertion	C (0.2)/gap(0.8)	Unfixed	Frame shift	Yes
Cyp4e2	2R: 8444560-8447164	1581	*91-C*	343	Exon 3	Insertion	T(0.56)/gap(0.44)	Unfixed	Frame shift	Yes
Cyp12a4	3R: 19135158-19137136	1612	*91-C*	958	Exon 3	Insertion	A(0.13)/gap(0.87)	Unfixed	Frame shift	Yes
Cyp9b1	2R: 7127838-7129808	1518	*91-C*	716	Exon 3	Insertion	C(0.14)/gap(0.86)	Unfixed	Frame shift	Yes
Cyp4d8	3L: 7493557-7496179	1398	*91-C*	654–655	Exon 2	Insertion	GC(0.84)/ gap(0.16)	Unfixed		No
				657–660			GAAG(0.84)/ gap(0.16)	Unfixed		No

in expression for *Cyp6a2*, *Cyp6g1*, and *Cyp9c1* between *91-R* and the insecticide-susceptible strain *Canton-S*,[36] as well as *Cyp6a2*, *Cyp6a8*, *Cyp6g1*, and *Cyp6w1* in the DDT-resistant *91-R* and *Wisconsin* strains,[35] with this reproducibility across comparisons arguably evidence for selection. Additional lines of evidence show that high transgenic expression of *Cyp6g2* in Malpighian tubules and gut tissues mediates insecticide resistance in transformed *D. melanogaster*,[62] and up-regulation of *Cyp12a4* in wild populations of *D. melanogaster* is associated with lufenuron resistance.[63] The over-expressed *Cyp6a8* we found in *91-R* is not known to metabolize DDT[64] or confer resistance to DDT when expressed at high levels.[62] The common induction of *Cyp6a2*, *Cyp6a8*, and *Cyp6w1* in response to chemical challenge,[65] and *Cyp6a8* and *Cyp6w1* up-regulation in response to caffeine exposure[27] may suggest their involvement in general detoxification or stress response pathways. Notably, a number of previous studies reported that multiple P450 genes such as *Cyp6g1* and *Cyp12d1* were over-expressed in DDT-resistant *D. melanogaster* strains, suggesting that *Cyp6g1* and *Cyp12d1* are globally associated with DDT resistance.[29,30] Consistent with previous studies, we found that *Cyp6g1* and other P450 genes were constitutively over-expressed in DDT-resistant strain *91-R* as compared to DDT-susceptible strain *91-C*. Therefore, it seems that not just one P450 (*Cyp6g1*) but multiple P450 genes may play an important role in DDT resistance in the *91-R* strain. Alternatively, the findings may indicate that transcripts derived from different closely related gene family members may be co-regulated and suggest that endogenous expression data alone may be insufficient to link a gene with specific resistance traits. Contrasting evidence also shows that *Cyp6g1* imparts a potential and uncharacterized selective advantage to embryos from DDT-resistant *D. melanogaster*.[66] Furthermore, CYP4G16/17, orthologs of the CYP4G1/2 of *D. melanogaster*, are highly overexpressed in oenocytes of pyrethroid-resistant *Anopheles* mosquitoes to catalyze cuticular hydrocarbon production, supporting the conclusion that the cuticular-based resistance mechanism has occurred in pyrethroid-resistant *Anopheles* mosquitoes.[67] Taken together, these observations indicate that co-selection acting on regulatory modules or selection acting on P450 genes is unconnected with xenobiotic detoxification. These findings also suggest that population-specific changes in the expression levels of P450 may occur as a result of chronic exposure and subsequent directional selection in *91-R* compared to the lack of analogous adaptive response by *91-C*.

Cyp6g1 was first identified as a DDT resistance gene in *D. melanogaster* by Daborn *et al.*[68] The insertion of the *Accord* retrotransposon in the upstream region of the *Cyp6g1* gene was also associated with an increased expression level for DDT resistance.[69] Chung *et al.*[70] also showed that this *Accord* element insertion results in tissue-specific gene expression and contributes to potential functions in detoxification. Schmidt and collaborators identified the different *Cyp6g1* allelic types segregating in the global *D. melanogaster* population, and found that evolution of these different haplotypes at the *Cyp6g1* locus showed a highly significant association with DDT resistance level, from a susceptible *M* ancestor to the *AA* and *BA* haplotypes with intermediate levels of resistance.[32] Le Goff and Hilliou studied the evolution of the *Cyp6g1* allele, showing that *Cyp6g1-M* is an ancestral allele without copy number variation or TE insertion.[71] *Cyp6g1-AA* and *Cyp6g1-BA* containing a duplication of *Cyp6g1* with either the *Accord* or *HMS-Beagle* TE insertions in promoter regions are the most represented alleles in global *D. melanogaster* populations. Both *Cyp6g1-AA* and *Cyp6g1-BA* confer resistance to DDT relative to the ancestral *Cyp6g1-M* allele.[71] We confirmed that the *91-R* strain showed not only a high expression level of *Cyp6g1* but also a duplication event with the *Accord* and *HMS-Beagle* element insertions conferring significantly high levels of DDT resistance. Our results are consistent with the hypothesis that the *Cyp6g1-BA* locus in *91-R* may play a particular role in evolutionary adaptation to DDT selection.

In addition to changes in P450 expression levels between *91-C* and *91-R*, a number of SNPs are present within and between the two strains. Although the majority of mutations were synonymous (silent), not fixed differently between strains, or predicted to have no effect on P450 enzyme function, a subset of mutations was predicted to have possible involvement in the *91-R* DDT resistance mechanism. Previous studies have found that CYP302A1, CYP306A1, CYP307A1, CYP307A2, CYP314A1, CYP315A1, and CYP18A1 are involved in the ecdysteroid biosynthetic pathway in *D. melanogaster*, and all genes highly conserved in insect

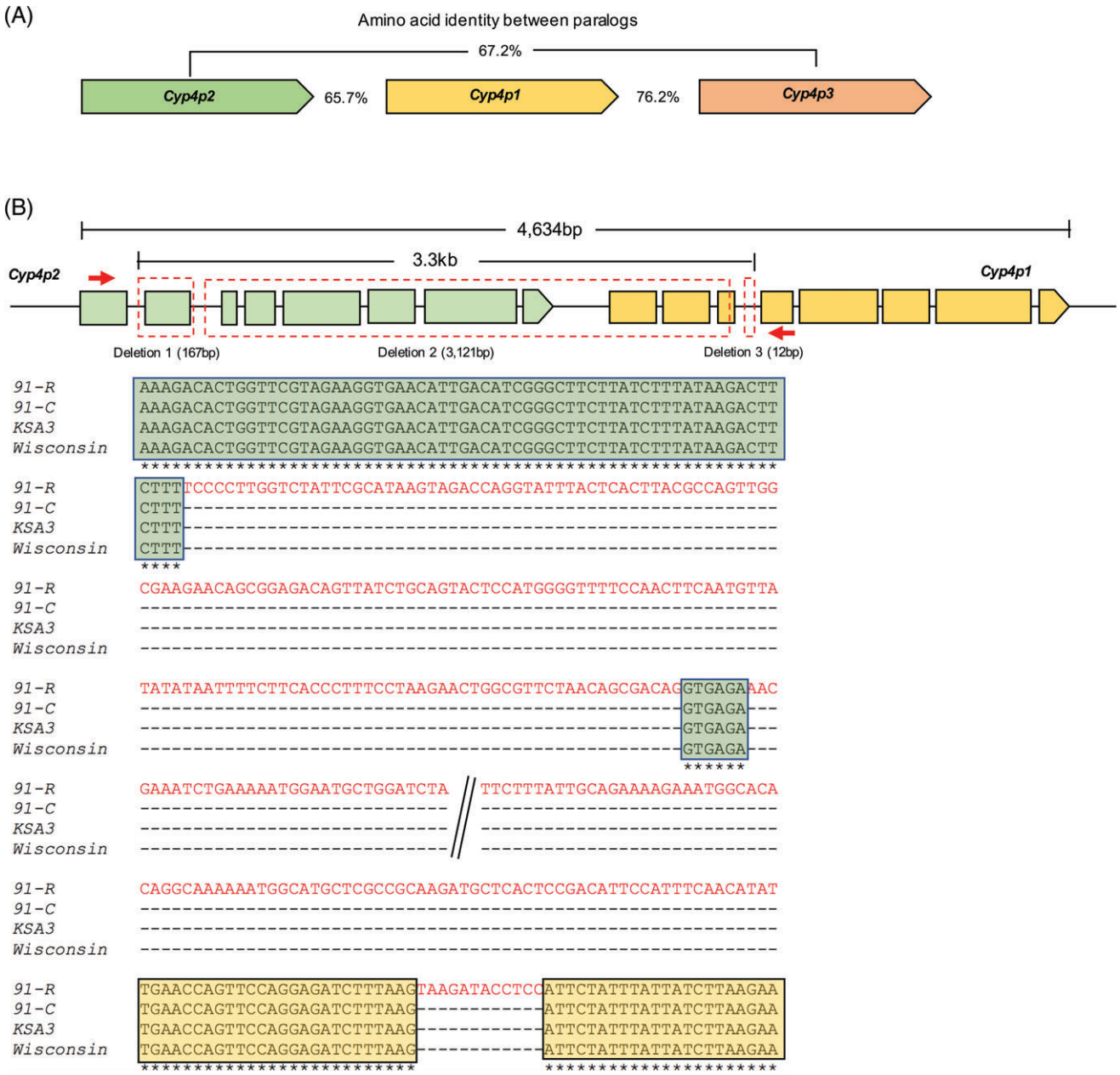

Figure 4. Schematic representation of variance between paralogs in the *Cyp4p* clade (*Cyp4p1/2/3*). (A) Amino acid identities between *Cyp4p1*, *Cyp4p2*, and *Cyp4p3*. (B) Positions for three deleted regions in the tandem duplicated *Cyp4p2* (green) and *Cyp4p1* (yellow) genes on chromosome 2 of *91-C*. The three deletions spanning approximately ~3.3 kb of DNA (Deletions 1–3) are indicated by red dashed boxes. Solid black lines indicate exons and introns. Red letters represent deleted sequences; red arrows show sites of primers used in PCR validation.

genomes.[72,73] Interestingly, we observed that the CYP2 clade (CYP18A1, CYP305A1, CYP306A1, and CYP307A2) has no amino acid difference between the *91-C* and *91-R* strains, whereas other CYP clades associated with metabolism of both endogenous and xenobiotic substrates may be more likely to have amino acid variations in order to adapt to environmental selection pressures. *Cyp305a1* is involved in ecdysteroid synthesis and lipid storage regulation,[74] and *Cyp306a1* (*Phantom*) is a *Halloween* gene encoding an ecdysteroid 25-hydroxylase.[75] Furthermore, *Drosophila Cyp18a1* is an ecdysone-responsive gene and possibly encodes an ecdysteroid 26-hydroxylase/oxidase, which coordinates major developmental transitions in *Drosophila*.[72] Therefore, our results imply that the CYP2 clade involved in hormone metabolism pathways is evolutionarily conserved in *D. melanogaster*. Our comparisons of derived P450 amino acid sequences identified that 16 P450 genes have nonsynonymous (amino acid-changing) mutations within SRS 1 – 6 and/or 'meander'-binding regions that are highly conserved among P450 sequences. Previous studies indicate that amino acids within SRS 1 – 6 contribute to the stability of the catalytic site and define the substrate specificity.[76] Additionally, an I115L alteration in the SRS 1 region of CYP6B1 is involved in substrate turnover in *Papilio polyxenes*.[77] Recent work by Schmidt *et al.*[78] demonstrated that transgenic overexpression of the CYP6W1_Ala370 allele increased DDT resistance relative to CYP6W1_Val370 and CYP6W1_Gly370. Therefore, we can infer that the overexpression of CYP6W1_Ala370 mutation that is located in

SRS 5 of CYP6W1 may play a role in the DDT resistance phenotype in the *91-R* strain.

Although SRS regions are predicted to influence the stability of P450 protein structures and substrate specificities, site-directed mutagenesis outside of the canonical SRS regions of P450 proteins has demonstrated that amino acid substitutions also affect catalytic activity and substrate reactivity.[79,80] Thus, the 13 nonsynonymous mutations from ten P450s that reside in or near functional domains, and that are fixed differently between *91-C* and *91-R*, may affect substrate selectivity and catalytic activity against xenobiotics. Thus, these candidate mutations may impact the degree of DDT resistance or susceptibility in *91-C* and *91-R*, but additional experiments are required to further investigate these hypotheses. As these mutations are fixed differently between independent laboratory lines, the assumption is that all fixed changes are the effect of positive selection as a result of the response to DDT exposure in *91-R*. Notwithstanding this, the potential remains that genetic drift within these small populations could have resulted in these changes by random chance. Thus, additional tests are required to interrogate changes in function as a consequence of protein changes caused by the predicted nonsynonymous mutations, as well as establish linkage (co-segregation) of these mutations and DDT resistance, so our results must be interpreted with caution. SNPs in transcription factor binding sites may alter the binding of transcription factors, resulting in altered transcriptional activation and variation in phenotype that may subsequently impact the susceptibility to insecticide exposure. We identified no variation within functional ARE sequences of the promoter regions from nine P450 genes that are differentially expressed between the *91-C* and *91-R* strains. Regardless, identifying potential functional SNPs in transcription factor binding sites could provide evidence for putative modulation of expression via *cis*-regulatory elements in response to environmental exposure. Alternatively, the lack of predicted mutation of AREs might lend additional support to the theory that structural changes may be more pronounced in the DDT resistance trait considered in the current study.

The accumulated differences in SNP positions between *91-C* and *91-R* also resulted in the production of an aberrant inserted stop codon (nonsense mutations) that became fixed between strains. For example, *Cyp6a2* is known to be involved in DDT resistance (see Introduction) and shows a homozygous deletion of a 'G' nucleotide at bp position 1341 in *91-C* as compared to a fixed alternate allele in resistant *91-R*. The deletion in *91-C* is predicted to cause a frameshift that introduces a premature stop codon in exon 2 (Table 3). This *Cyp6a2* mutant in *91-C* is putatively nonfunctional, as the truncated 483 protein would lack the Cys residue that is highly conserved and functions within the heme-binding domain.[59] Our analysis confirms the previous predictions for this frameshift.[37] Gross structural changes were predicted in 10 of the 87 P450s (11.5%) that would result in truncated proteins via missense mutations. Specifically, homozygous insertions or deletions that result in truncations fixed differently between strains were found in *Cyp6a2* and *Cyp316a1* from the *91-C* strain, and in *Cyp6a14* from the *91-R* strain. These premature stop codons cause aberrant translation and resultant truncated proteins that lack the highly conserved oxygen-binding motif, the heme-binding motif, and the putative 'meander'-binding sequence that are important for catalytic function.[81] Consequently, this finding further suggests that mutations fixed differently between strains *91-C* and *91-R* (e.g. three fixed alleles for DIPs in *Cyp6a2*, *Cyp316a1*, and *Cyp6a14*) might thereby represent inter-strain differences. In view of the involvement of *Cyp6a2* in xenobiotic detoxification,[65] loss of function in *91-C* suggests a loss of capacity for detoxification or general stress response. Analogous insertion/deletion events generating pseudogenes are present in P450s, such as *D. melanogaster Cyp9f3p*, which lacks a 93-nucleotide fragment that encodes 31 amino acids of the heme-binding region.[82,83]

The *Cyp4p* clade comprising *Cyp4p1/2/3* on chromosome 2R has experienced a number of duplications and gene losses across *Drosophila* species,[16] suggesting a degree of functional redundancy among the gene duplicates. The *Cyp4p1* paralog is up-regulated in response to caffeine exposure,[27] and *Cyp4p1* and *Cyp4p2* are also known to be regulated by ectopic expression of the Nrf2/Cnc transcription factor.[39] The observation that the deletion of upstream regulatory regions of *Cyp4p1* in *91-C* fails to result in any significant difference in expression compared with *91-R* confirms the RNA-seq-based estimates of low basal expression in *91-C*. Regardless, the *Cyp4p* clade remains functional in *91-R*, and *Cyp4p2* is up-regulated in DDT- and imidacloprid-resistant flies when compared with susceptible counterparts.[84] The exact role of the *Cyp4* clade in resistance remains unknown. As the gene family resides outside of genome regions of a selective sweep[47] and shows no significant up-regulation in *91-R* compared with the deficient *91-C* strain, one might conclude that *Cyp4p1/2/3* may not be involved in the DDT resistance trait of *91-R*. Such a conclusion would contradict evidence provided by fixation of the deletion differently between strains, and thus could suggest that any influence of *Cyp4p* gene family members could be additive or non-additive if paralogs retain ancestral function in response to xenobiotic exposure. In addition to the *Cyp4p1/2* deletion in the *91-C* strain, we have discovered nine P450s (six in the *91-C* strain, two in the *91-R* strain and one in both strains) that were 3′ truncated and probably nonfunctional as a consequence of indels that caused a premature stop codon/frame shift.

A relaxation of selective/functional constraints is a possible explanation for our observed loss of function for several P450 genes. Both strains have been maintained under much less chemically challenging environments relative to their wild counterparts, which would encounter various plant toxins and artificial chemicals/pollutants. This may have enabled them to lose some xenobiotic-metabolizing P450s to avoid the fitness cost of their expression in the absence of other such chemical or environmental stresses. The *91-R* strain has been exposed to DDT selection, and thus it is understandable that specific P450s are probably needed more in *91-R* than in *91-C*. Additionally, in spite of evidence that selective sweeps have probably occurred in response to DDT selection in *91-R*,[47] it is possible that genetic drift could have resulted in significant genetic variation in the relatively small number of individuals in the *91-C* and *91-R* laboratory populations.[85] While P450 enzymes play a critical role in the metabolism of endogenous and exogenous chemicals, including insecticides,[86] loss-of-function mutations do not result in catastrophic loss of fitness. In *Blattella germanica*, multiple P450 pseudogenes result from deletions or frameshift introductions of stop codons in their coding DNA sequence (CDS), such as for *Cyp9e2* and *Cyp4c21*.[87] Additionally, the large deletions in cattle tick *Cyp4w1* generate a nonfunctional pseudogene.[88] This finding suggests that macro- and micro-scale changes might impact the structure and subsequent function of P450 genes, and demonstrate a remarkable degree of allelic diversity through processes such as gene duplication, substitution mutation, and generation of pseudogenes among *Drosophila* species.[16,20]

To the best of our knowledge, this study represents the first combined structural and functional analysis of variation in P450s

between highly DDT-resistant and -susceptible strains. Prior evidence suggests that high levels of DDT resistance in *91-R* are polygenic, with multiple SNPs in 13 genomic regions contributing to the DDT resistance phenotype.[47] Although comparative, the detection of SNPs and overrepresentation of nonsynonymous changes suggest putative functional changes in response to selection. Regardless, secondary validation of these structural changes and potential impacts on the DDT resistance phenotypes remain to be investigated. Specifically, the present study provides potential candidate SNPs within P450 genes for future functional studies using transgenic lines, CRISPR/Cas9, or genetic re-selection experiments to identify putative causal mutations that contribute to the observed DDT resistance phenotype in an additive or non-additive fashion.

5 CONCLUSION

The expression profiles and structural variations of a total of 87 P450 genes of *D. melanogaster* were compared between DDT-susceptible and -resistant populations. Our results represent an additional step toward understanding the evolutionary adaptive response of functional and structural variation within the P450 gene family as a result of directional selection imposed by chronic high levels of DDT exposure. In view of the widespread continuous use of pesticides in many agroecosystems, our study may provide useful insights into the rapid acquisition and stability of evolved resistance to insecticide selection, and the ability of insects to adapt to changing environments. Thus, future comparative CYP-omic studies, like the one performed here, in other pest populations or pest species, or both, may continue to expand our understanding of the evolution of detoxification systems under diverse evolutionary selection pressures.

ACKNOWLEDGEMENTS

The authors would like to thank Dr Henry Chung (Michigan State University) for excellent and insightful comments that substantially improved the manuscript. A portion of this research was supported by the United States Department of Agriculture (USDA), Agricultural Research Service (ARS), CRIS project 3625-22000-18-00D, and the Iowa Agriculture and Home Economics Experiment Station, Iowa State University, Ames, IA (Project 3543). Mention of a proprietary product does not constitute an endorsement or a recommendation by USDA for its use. USDA is an equal opportunity employer and provider.

SUPPORTING INFORMATION

Supporting information may be found in the online version of this article.

REFERENCES

1 Mallet J, The evolution of insecticide resistance: have the insects won? *Trends Ecol Evol* **4**:336–340 (1989).
2 Meinke LJ, Siegfried BD, Wright RJ and Chandler LD, Adult susceptibility of Nebraska western corn rootworm (Coleoptera: Chrysomelidae) populations to selected insecticides. *J Econ Entomol* **91**:594–600 (1998).
3 Pereira AE, Wang H, Zukoff SN, Meinke LJ, French BW and Siegfried BD, Evidence of field-evolved resistance to bifenthrin in western corn rootworm (*Diabrotica virgifera virgifera* LeConte) populations in western Nebraska and Kansas. *PLoS One* **10**:e0142299 (2015).
4 Hemingway J and Ranson H, Insecticide resistance in insect vectors of human disease. *Annu Rev Entomol* **45**:371–391 (2000).
5 Augustyniak M, Plachetka-Bozek A, Kafel A, Babczynska A, Tarnawska M, Janiak A *et al.*, Phenotypic plasticity, epigenetic or genetic modifications in relation to the duration of Cd-exposure within a microevolution time range in the beet armyworm. *PLoS One* **11**: e0167371 (2016).
6 Waddington CH, Canalization of development and the inheritance of acquired characters. *Nature* **150**:563–565 (1942).
7 Hemingway J, Hawkes NJ, McCarroll L and Ranson H, The molecular basis of insecticide resistance in mosquitoes. *Insect Biochem Mol Biol* **34**:653–665 (2004).
8 Coates BS and Siegfried BD, Linkage of an ABCC transporter to a single QTL that controls *Ostrinia nubilalis* larval resistance to the *Bacillus thuringiensis* Cry1Fa toxin. *Insect Biochem Mol Biol* **63**:86–96 (2015).
9 Nebert DW and Gonzalez FJ, P450 genes: structure, evolution, and regulation. *Annu Rev Biochem* **56**:945–993 (1987).
10 Werck-Reichhart D and Feyereisen R, Cytochromes P450: a success story. *Genome Biol* **1**:REVIEWS3003 (2000).
11 Oakeshott JG, Home I, Sutherland TD and Russell RJ, The genomics of insecticide resistance. *Genome Biol* **4**:202 (2003).
12 Lee SH, Kang JS, Min JS, Yoon KS, Strycharz JP, Johnson R *et al.*, Decreased detoxification genes and genome size make the human body louse an efficient model to study xenobiotic metabolism. *Insect Mol Biol* **19**:599–615 (2010).
13 Arensburger P, Megy K, Waterhouse RM, Abrudan J, Amedeo P, Antelo B *et al.*, Sequencing of *Culex quinquefasciatus* establishes a platform for mosquito comparative genomics. *Science* **330**:86–88 (2010).
14 Feyereisen R, Arthropod CYPomes illustrate the tempo and mode in P450 evolution. *Biochim Biophys Acta* **1814**:19–28 (2011).
15 Tijet N, Helvig C and Feyereisen R, The cytochrome P450 gene superfamily in *Drosophila melanogaster*: annotation, intron-exon organization and phylogeny. *Gene* **262**:189–198 (2001).
16 Good RT, Gramzow L, Battlay P, Sztal T, Batterham P and Robin C, The molecular evolution of cytochrome P450 genes within and between *Drosophila* species. *Genome Biol Evol* **6**:1118–1134 (2014).
17 Gotoh O, Evolution and differentiation of cytochrome P450 genes, in *Cytochrome P450*, 2nd edn, ed. by Omura T, Ishimura Y and Fujii-Kuriyama Y. Kodansha, Tokyo, pp. 255–272 (1993).
18 Nelson DR, Metazoan cytochrome P450 evolution. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* **121**:15–22 (1998).
19 Nelson DR, The cytochrome p450 homepage. *Hum Genomics* **4**:59–65 (2009).
20 McDonnell CM, King D, Comeron JM, Li H, Sun W, Berenbaum MR *et al.*, Evolutionary toxicogenomics: diversification of the *Cyp12d1* and *Cyp12d3* genes in *Drosophila* species. *J Mol Evol* **74**:281–296 (2012).
21 Sezutsu H, Le Goff G and Feyereisen R, Origins of P450 diversity. *Philos Trans R Soc London B Biol Sci* **368**:20120428 (2013).
22 Ffrench-Constant RH, The molecular genetics of insecticide resistance. *Genetics* **194**:807–815 (2013).
23 Wen Z, Zhang X and Zhang Y, P450-mediated insecticide detoxification and its implication in insecticide efficacy, in *Recent Advances in Entomological Research: From Molecular Biology to Pest Management*, ed. by Liu T and Kang L. Springer, Berlin-Heidelberg, pp. 229–245 (2011).
24 Nardini L, Christian RN, Coetzer N, Ranson H, Coetzee M and Koekemoer LL, Detoxification enzymes associated with insecticide resistance in laboratory strains of *Anopheles arabiensis* of different geographic origin. *Parasit Vectors* **5**:113 (2012).
25 Chandor-Proust A, Bibby J, Regent-Kloeckner M, Roux J, Guittard-Crilat E, Poupardin R *et al.*, The central role of mosquito cytochrome P450 CYP6Zs in insecticide detoxification revealed by functional expression and structural modelling. *Biochem J* **455**:75–85 (2013).
26 Schuler MA, P450s in plants, insects, and their fungal pathogens, in *Cytochrome P450*. Springer, Cham, pp. 409–449 (2015).
27 Coelho A, Fraichard S, Le Goff G, Faure P, Artur Y, Ferveur JF *et al.*, Cytochrome P450-dependent metabolism of caffeine in *Drosophila melanogaster*. *PLoS One* **10**:e0117328 (2015).
28 Terhzaz S, Cabrero P, Brinzer RA, Halberg KA, Dow JA and Davies SA, A novel role of *Drosophila* cytochrome P450-4e3 in permethrin insecticide tolerance. *Insect Biochem Mol Biol* **67**:38–46 (2015).
29 Brandt A, Scharf M, Pedra JH, Holmes G, Dean A, Kreitman M *et al.*, Differential expression and induction of two *Drosophila* cytochrome P450 genes near the Rst(2)DDT locus. *Insect Mol Biol* **11**:337–341 (2002).
30 Festucci-Buselli RA, Carvalho-Dias AS, de Oliveira-Andrade M, Caixeta-Nunes C, Li HM, Stuart JJ *et al.*, Expression of *Cyp6g1* and *Cyp12d1*

in DDT resistant and susceptible strains of *Drosophila melanogaster*. *Insect Mol Biol* **14**:69–77 (2005).
31 Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S *et al.*, A single p450 allele associated with insecticide resistance in *Drosophila*. *Science* **297**:2253–2256 (2002).
32 Schmidt JM, Good RT, Appleton B, Sherrard J, Raymant GC, Bogwitz MR *et al.*, Copy number variation and transposable elements feature in recent, ongoing adaptation at the *Cyp6g1* locus. *PLoS Genet* **6**:e1000998 (2010).
33 Battlay P, Schmidt JM, Fournier-Level A and Robin C, Genomic and transcriptomic associations identify a new insecticide resistance phenotype for the selective sweep at the Cyp6g1 locus of *Drosophila melanogaster*. *G3 (Bethesda)* **6**:2573–2581 (2016).
34 Denecke S, Fusetto R, Martelli F, Giang A, Battlay P, Fournier-Level A *et al.*, Multiple P450s and variation in neuronal genes underpins the response to the insecticide imidacloprid in a population of *Drosophila melanogaster*. *Sci Rep* **7**:11338 (2017).
35 Pedra JH, McIntyre LM, Scharf ME and Pittendrigh BR, Genome-wide transcription profile of field- and laboratory-selected dichlorodiphenyltrichloroethane (DDT)-resistant *Drosophila*. *Proc Natl Acad Sci U S A* **101**:7034–7039 (2004).
36 Qiu X, Sun W, McDonnell CM, Li-Byarlay H, Steele LD, Wu J *et al.*, Genome-wide analysis of genes associated with moderate and high DDT resistance in *Drosophila melanogaster*. *Pest Manage Sci* **69**:930–937 (2013).
37 Wan H, Liu Y, Li M, Zhu S, Li X, Pittendrigh BR *et al.*, Nrf2/Maf-binding-site-containing functional *Cyp6a2* allele is associated with DDT resistance in *Drosophila melanogaster*. *Pest Manage Sci* **70**:1048–1058 (2014).
38 Kensler TW, Wakabayashi N and Biswal S, Cell survival responses to environmental stresses via the Keap1-Nrf2-ARE pathway. *Annu Rev Pharmacol Toxicol* **47**:89–116 (2007).
39 Misra JR, Horner MA, Lam G and Thummel CS, Transcriptional regulation of xenobiotic detoxification in *Drosophila*. *Genes Dev* **25**:1796–1806 (2011).
40 Maitra S, Dombrowski SM, Basu M, Raustol O, Waters LC and Ganguly R, Factors on the third chromosome affect the level of *Cyp6a2* and *Cyp6a8* expression in *Drosophila melanogaster*. *Gene* **248**:147–156 (2000).
41 Sun W, Valero MC, Seong KM, Steele LD, Huang IT, Lee CH *et al.*, A glycine insertion in the estrogen-related receptor (ERR) is associated with enhanced expression of three cytochrome P450 genes in transgenic *Drosophila melanogaster*. *PLoS One* **10**:e0118779 (2015).
42 Lamb DC, Skaug T, Song HL, Jackson CJ, Podust LM, Waterman MR *et al.*, The cytochrome P450 complement (CYPome) of *Streptomyces coelicolor* A3(2). *J Biol Chem* **277**:24000–24005 (2002).
43 Scott JG, Liu N and Wen Z, Insect cytochromes P450: diversity, insecticide resistance and tolerance to plant toxins. *Comp Biochem Physiol C Pharmacol Toxicol Endocrinol* **121**:147–155 (1998).
44 Merrell DJ, Heterosis in DDT resistant and susceptible populations of *Drosophila melanogaster*. *Genetics* **45**:573–581 (1960).
45 Merrell DJ, Lethal frequency and allelism in DDT resistant populations and their controls. *Am Nat* **99**:411–417 (1965).
46 Merrell DJ and Underhill JC, Selection for DDT resistance in inbred, laboratory and wild stocks of *Drosophila melanogaster*. *J Econ Entomol* **49**:300–306 (1956).
47 Steele LD, Coates B, Valero MC, Sun W, Seong KM, Muir WM *et al.*, Selective sweep analysis in the genomes of the *91-R* and *91-C Drosophila melanogaster* strains reveals few of the 'usual suspects' in dichlorodiphenyltrichloroethane (DDT) resistance. *PLoS One* **10**:e0123066 (2015).
48 Steele LD, Muir WM, Seong KM, Valero MC, Rangesa M, Sun W *et al.*, Genome-wide sequencing and an open reading frame analysis of dichlorodiphenyltrichloroethane (DDT) susceptible (*91-C*) and resistant (*91-R*) *Drosophila melanogaster* laboratory populations. *PLoS One* **9**:e98584 (2014).
49 Seong KM, Sun W, Clark JM and Pittendrigh BR, Splice form variant and amino acid changes in MDR49 confers DDT resistance in transgenic *Drosophila*. *Sci Rep* **6**:23355 (2016).
50 Mackay TF, Richards S, Stone EA, Barbadilla A, Ayroles JF, Zhu D *et al.*, The *Drosophila melanogaster* genetic reference panel. *Nature* **482**:173–178 (2012).
51 Seong KM, Coates BS, Sun W, Clark JM and Pittendrigh BR, Changes in neuronal signaling and cell stress response pathways are associated with a multigenic response of *Drosophila melanogaster* to DDT selection. *Genome Biol Evol* **9**:3356–3372 (2017).
52 Dos Santos G, Schroeder AJ, Goodman JL, Strelets VB, Crosby MA, Thurmond J *et al.*, FlyBase: introduction of the *Drosophila melanogaster* release 6 reference genome assembly and large-scale migration of genome annotations. *Nucleic Acids Res* **43**:D690–D697 (2015).
53 Audic S and Claverie JM, The significance of digital gene expression profiles. *Genome Res* **7**:986–995 (1997).
54 Ponton F, Chapuis MP, Pernice M, Sword GA and Simpson SJ, Evaluation of potential reference genes for reverse transcription-qPCR studies of physiological responses in *Drosophila melanogaster*. *J Insect Physiol* **57**:840–850 (2011).
55 Pfaffl MW, A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* **29**:e45 (2001).
56 Ramakers C, Timmer D, Bakker O, Deprez RHL, Ruijter JM and Moorman AFM, Quantification of mRNA using linear regrssion of log-linear PCR data points as an alternative for the standard curve approach, in *Rapid Cycle Real-timePCR, Methods and Applications*, ed. by Meurer S, Wittwer C and Nakagawara K. Springer Press, Heidelberg, pp. 21–29 (2004).
57 Schmittgen TD and Livak KJ, Analyzing real-time PCR data by the comparative C(T) method. *Nat Protoc* **3**:1101–1108 (2008).
58 Emerson JJ, Cardoso-Moreira M, Borevitz JO and Long M, Natural selection shapes genome-wide patterns of copy-number polymorphism in *Drosophila melanogaster*. *Science* **320**:1629–1631 (2008).
59 Gotoh O, Substrate recognition sites in cytochrome P450 family 2 (CYP2) proteins inferred from comparative analyses of amino acid and coding nucleotide sequences. *J Biol Chem* **267**:83–90 (1992).
60 Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ *et al.*, JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. *Nucleic Acids Res* **42**:D142–D147 (2014).
61 Harrop TW, Sztal T, Lumb C, Good RT, Daborn PJ, Batterham P *et al.*, Evolutionary changes in gene expression, coding sequence and copy-number at the *Cyp6g1* locus contribute to resistance to multiple insecticides in *Drosophila*. *PLoS One* **9**:e84879 (2014).
62 Daborn PJ, Lumb C, Boey A, Wong W, Ffrench-Constant RH and Batterham P, Evaluating the insecticide resistance potential of eight *Drosophila melanogaster* cytochrome P450 genes by transgenic over-expression. *Insect Biochem Mol Biol* **37**:512–519 (2007).
63 Bogwitz MR, Chung H, Magoc L, Rigby S, Wong W, O'Keefe M *et al.*, *Cyp12a4* confers lufenuron resistance in a natural population of *Drosophila melanogaster*. *Proc Natl Acad Sci U S A* **102**:12807–12812 (2005).
64 Helvig C, Tijet N, Feyereisen R, Walker FA and Restifo LL, Drosophila melanogaster CYP6A8, an insect P450 that catalyzes lauric acid (omega-1)-hydroxylation. *Biochem Biophys Res Commun* **325**:1495–1502 (2004).
65 Giraudo M, Unnithan GC, Le Goff G and Feyereisen R, Regulation of cytochrome P450 expression in *Drosophila*: genomic insights. *Pestic Biochem Physiol* **97**:115–122 (2010).
66 McCart C and Ffrench-Constant RH, Dissecting the insecticide-resistance- associated cytochrome P450 gene *Cyp6g1*. *Pest Manage Sci* **64**:639–645 (2008).
67 Balabanidou V, Kampouraki A, MacLean M, Blomquist GJ, Tittiger C, Juárez MP *et al.*, Cytochrome P450 associated with insecticide resistance catalyzes cuticular hydrocarbon production in *Anopheles gambiae*. *Proc Natl Acad Sci U S A* **113**:9268–9273 (2016).
68 Daborn P, Boundy S, Yen J, Pittendrigh B and ffrench-Constant R, DDT resistance in *Drosophila* correlates with *Cyp6g1* over-expression and confers cross-resistance to the neonicotinoid imidacloprid. *Mol Genet Genomics* **266**:556–563 (2001).
69 Catania F, Kauer MO, Daborn PJ, Yen JL, Ffrench-Constant RH and Schlotterer C, World-wide survey of an *Accord* insertion and its association with DDT resistance in *Drosophila melanogaster*. *Mol Ecol* **13**:2491–2504 (2004).
70 Chung H, Bogwitz MR, McCart C, Andrianopoulos A, Ffrench-Constant RH, Batterham P *et al.*, Cis-regulatory elements in the *Accord* retrotransposon result in tissue-specific expression of the *Drosophila melanogaster* insecticide resistance gene *Cyp6g1*. *Genetics* **175**:1071–1077 (2007).
71 Le Goff G and Hilliou F, Resistance evolution in *Drosophila*: the case of CYP6G1. *Pest Manage Sci* **73**:493–499 (2017).
72 Guittard E, Blais C, Maria A, Parvy JP, Pasricha S, Lumb C *et al.*, CYP18A1, a key enzyme of *Drosophila* steroid hormone inactivation, is essential for metamorphosis. *Dev Biol* **349**:35–45 (2011).

73 Rewitz KF, O'Connor MB and Gilbert LI, Molecular evolution of the insect Halloween family of cytochrome P450s: phylogeny, gene organization and functional conservation. *Insect Biochem Mol Biol* **37**:741–753 (2007).

74 Hansen BH, Altin D, Hessen KM, Dahl U, Breitholtz M, Nordtug T *et al.*, Expression of ecdysteroids and cytochrome P450 enzymes during lipid turnover and reproduction in *Calanus finmarchicus* (Crustacea: Copepoda). *Gen Comp Endocrinol* **158**:115–121 (2008).

75 Namiki T, Niwa R, Sakudoh T, Shirai K, Takeuchi H and Kataoka H, Cytochrome P450 CYP307A1/spook: a regulator for ecdysone synthesis in insects. *Biochem Biophys Res Commun* **337**:367–374 (2005).

76 Chen JS, Berenbaum MR and Schuler MA, Amino acids in SRS1 and SRS6 are critical for furanocoumarin metabolism by CYP6B1v1, a cytochrome P450 monooxygenase. *Insect Mol Biol* **11**:175–186 (2002).

77 Wen Z, Baudry J, Berenbaum MR and Schuler MA, Ile115Leu mutation in the SRS1 region of an insect cytochrome P450 (CYP6B1) compromises substrate turnover via changes in a predicted product release channel. *Protein Eng Des Sel* **18**:191–199 (2005).

78 Schmidt JM, Battlay P, Gledhill-Smith RS, Good RT, Lumb C, Fournier-Level A *et al.*, Insights into DDT resistance from the *Drosophila melanogaster* genetic reference panel. *Genetics* **207**:1181–1193 (2017).

79 Kedzie KM, Balfour CA, Escobar GY, Grimm SW, He YA, Pepperl DJ *et al.*, Molecular basis for a functionally unique cytochrome P450IIB1 variant. *J Biol Chem* **266**:22515–22521 (1991).

80 He YA, Balfour CA, Kedzie KM and Halpert JR, Role of residue 478 as a determinant of the substrate specificity of cytochrome P450 2B1. *Biochemistry* **31**:9220–9226 (1992).

81 Oezguen N and Kumar S, Analysis of cytochrome P450 conserved sequence motifs between helices E and H: prediction of critical motifs and residues in enzyme functions. *J Drug Metab Toxicol* **2**:1000110 (2011).

82 Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG *et al.*, The genome sequence of *Drosophila melanogaster*. *Science* **287**:2185–2195 (2000).

83 Harrison PM, Milburn D, Zhang Z, Bertone P and Gerstein M, Identification of pseudogenes in the *Drosophila melanogaster* genome. *Nucleic Acids Res* **31**:1033–1037 (2003).

84 Kalajdzic P, Oehler S, Reczko M, Pavlidi N, Vontas J, Hatzigeorgiou AG *et al.*, Use of mutagenesis, genetic mapping and next generation transcriptomics to investigate insecticide resistance mechanisms. *PLoS One* **7**:e40296 (2012).

85 Santos J, Pascual M, Simoes P, Fragata I, Lima M, Kellen B *et al.*, From nature to the laboratory: the impact of founder effects on adaptation. *J Evol Biol* **25**:2607–2622 (2012).

86 Scott JG and Wen Z, Cytochromes P450 of insects: the tip of the iceberg. *Pest Manage Sci* **57**:958–967 (2001).

87 Wen Z, Horak CE and Scott JG, CYP9E2, CYP4C21 and related pseudogenes from German cockroaches, *Blattella germanica*: implications for molecular evolution, expression studies and nomenclature of P450s. *Gene* **272**:257–266 (2001).

88 Crampton AL, Baxter GD and Barker SC, Identification and characterisation of a cytochrome P450 gene and processed pseudogene from an arachnid: the cattle tick, *Boophilus microplus*. *Insect Biochem Mol Biol* **29**:377–384 (1999).

89 Wasserman WW and Sandelin A, Applied bioinformatics for the identification of regulatory elements. *Nat Rev Genet* **5**:276–287 (2004).

