

# Co-amplification explains linkage disequilibrium of two mosquito esterase genes in insecticide-resistant *Culex quinquefasciatus*

Ashley VAUGHAN\*, Nicola HAWKES and Janet HEMINGWAY†

School of Pure and Applied Biology, University of Wales Cardiff, P.O. Box 915, Cardiff CF1 3TL, U.K.

The mosquito *Culex quinquefasciatus* (Say) is a vector of human disease and a world-wide biting nuisance. Organophosphorus insecticides (OPs) have been widely used to control *C. quinquefasciatus* populations and this has led to the emergence of OP-resistance. Predominantly, resistance is caused by increased production of two non-specific carboxylesterases, Est $\alpha$ 2<sup>1</sup> and Est $\beta$ 2<sup>1</sup>. Increased abundance of these esterases is associated with the amplification of their respective genes. The *est $\alpha$ 2<sup>1</sup>* and *est $\beta$ 2<sup>1</sup>* genes were cloned and sequenced from OP-resistant Sri Lankan *C. quinquefasciatus*; the two adjacent genes are in a head to head configuration, within a single amplification unit (amplicon). The homology between the two genes suggests that they arose from an ancient duplication event. The two genes have different numbers of exons (*est $\alpha$ 2<sup>1</sup>* has seven and *est $\beta$ 2<sup>1</sup>* has four); however, the intron/exon boundaries in *est $\beta$ 2<sup>1</sup>* are all conserved in *est $\alpha$ 2<sup>1</sup>*. The two genes are co-amplified in three other mosquito strains with the elevated Est $\alpha$ 2<sup>1</sup>/Est $\beta$ 2<sup>1</sup> phenotype. Their complete

linkage disequilibrium is explained by the location of the two genes involved in resistance within a single amplicon. In insecticide-susceptible *C. quinquefasciatus*, the non-amplified *est $\alpha$*  and *est $\beta$*  gene loci are also found in a similar head to head configuration, but the size of the intergenic non-coding region is approx. 1 kb less than in the amplicon. The smaller intergenic spacer is also found in a strain with amplified *est $\beta$ 1<sup>1</sup>*, which suggests that extensive laboratory selection for this amplified esterase has not eliminated the non-amplified genes. The intergenic spacer regions have been subcloned and sequenced. They contain numerous possible TATA boxes, promoters and a number of possible regulatory elements with high homology to the consensus sequence of the Barbie box. These latter putative regulatory elements are more numerous in the larger intergenic spacer, which differs from the non-amplified spacer by two large (> 420 bp) and one small (5 bp) insertions.

## INTRODUCTION

The development of insecticide-resistance has caused problems in control of agricultural, medical and veterinary insect pests. In a wide range of species, the elevation of esterase activity is the major mechanism of organophosphorus insecticide (OP) resistance [1–8]. The molecular basis of this resistance mechanism has only been extensively studied in two species. In the mosquito *Culex quinquefasciatus* and the aphid *Myzus persicae*, amplification of one or more esterase genes underlies the increase in esterase activity [3,9,10]. The esterases involved in both species are all serine or B esterases according to the classification of Aldridge [11]. Members of this family of enzymes have been characterized from a large number of mammals and insects and they are able to hydrolyse carboxylester, amide and thioester bonds in a variety of compounds.

A revised classification for the esterases involved in OP-resistance in *C. quinquefasciatus* has recently been put forward [9]. In accordance with this classification, the most common elevated esterase pattern found in OP-resistant *C. quinquefasciatus* populations is the co-elevation of Est $\alpha$ 2 and Est $\beta$ 2 esterases [12–15]. The elevated Est $\alpha$ 2<sup>1</sup> and Est $\beta$ 2<sup>1</sup> esterases are generally acknowledged to be in complete linkage disequilibrium in mosquito populations, although there have been at least two calculations of linkage distances on the basis of rare recombination events [16,17].

Increased expression of Est $\alpha$  and Est $\beta$  esterases involved in OP-resistance is associated with gene amplification [9,10,18,19]. Amplification of a single esterase, coded for by the *est $\beta$ 1<sup>1</sup>* gene, was first shown in the Californian TEM-R strain of mosquito [18]. Immunological and molecular studies have since suggested that the *est $\beta$ 1<sup>1</sup>*, *est $\beta$ 1<sup>2</sup>* and *est $\beta$ 2<sup>1</sup>* genes were originally alleles of the same locus [10,20,21]. More recently, cloning of the *est $\alpha$ 2<sup>1</sup>* cDNA has enabled comparison of the coding regions of the *est $\alpha$*  and *est $\beta$*  esterase genes [9]. The high homology (47% similarity) between the assumed amino acid sequences of the two esterases suggests that the two genes arose from an ancient duplication event.

We now report the gene sequences of the *est $\alpha$ 2<sup>1</sup>* and *est $\beta$ 2<sup>1</sup>* genes and the genomic organization of the two esterase loci in both insecticide-resistant and -susceptible mosquito strains from various continents.

## EXPERIMENTAL

### Mosquito strains

Details of the PeRR and PeSS mosquito strains are as reported previously [9,10]. The Tanzanian strains of *C. quinquefasciatus* were collected from Dar es Salaam (DAR91) and Tanga (TANGA). The DAR91 strain originated from a resistant field population which had been selected with chlorpyrifos for almost

Abbreviation used: OP, organophosphorus insecticide.

\* Present address: Department of Entomology, University of Madison-Wisconsin, 1630 Linden Drive, Madison, WI 53706, U.S.A.

† To whom correspondence should be addressed.

The nucleotide sequence data for *est $\alpha$ 2<sup>1</sup>* and *est $\beta$ 2<sup>1</sup>* will appear in the EMBL, GenBank and DDBJ Nucleotide Sequences Databases under the accession numbers Z47988 and Z86069 respectively.

ACTGATAAAATAGTGTATTTAACCTAACACAATCTCAACAGGAIAATTTTAAAAGTCTCCCTCAAAACAGCTCATCGTTCACGTACATGCTGCCACATGTCCAATCGGTCCGCATCCGGATAATCAACGAACGCCAC 140  
 F L E D N V Y M A D W M D L R D A D P Y D V F A V  
 B2 exon 4

CCCGTCGTGGCAATGTTGACAGCATTGAACGTGGGGTTCGCTGCGCGTTCGGCTCGAACACACACACCCGCTCTTCGCGCATGCCACAGTTGGATCCCGTGTGATGACGAACCGGGTGAACACATCGACCAGCGTTT 280  
 G D N A I N L C K F T P K T Q A N P E F V V G S K A T H G C N P D G N I V F A T F V D V L T Q  
 B2 exon 4

GCAGACCCGCTACTCGAAGTTTCTTTCGCGGGGACCTGCTGGTGAAGTTGAAAAACAGATAGGACAGCTCGTGGCATGGGCGGTGCCGCGAGCTTCGGTGCATCATGATGCGGTAGTGGTGTAGAATCC 420  
 L G R Y E F T E K G P V Q O T F N S F L Y S L E D A H A T G R L K P D I M M I R Y H N Y F E  
 B2 exon 4

GAATCCAGACAGATCCGGTACACGAAGGTGCGGGCGCGCATGAGCGGCTCGGGCAGGATGGTGGGTCAGGCGGTCCAGAAGACCCGGTCCGACATCATCGAAATTTTGTGGAGGGAGTAATATGTGATGTG 560  
 S D L C I R Y V F T R A R S R A A R A L I T R H L G H W F V R D S M M  
 B2 exon 4

GTTATTGAGGTTGAATGGAACCTACATGAACGTATCCAGGTTGTTCCATCGAAGGACTGCTATCGGGGTAGTAACGTTGCTTCAGTTTGGCAGCAAACTCGATTCTGTTTTCCATGCTGATCTTCAAGTTGGAGG 700  
 H V Y G L N N E H S P S S D P Y Y R O K L K A A F E I R K E M S I K L N P P  
 B2 exon 3

AACATTCCCGAAGACAGATGAGGATGGGACAGTAGTTCGGTTCGAACCTGATCTTTTTCGAGCAGCAGCGTCTTCCCTCAGAAGTTCCCTCCGATCATGATATCTATCTTGTCCACCCAAAGCAGTTCGAGCCATCTCAA 840  
 V N G L F L H P H S L L E P Q L K I K O L L G E E S T G G I M I L D I K D G W A T R A M E F  
 B2 exon 3

ACGGTTCCTTCGGAATCATGCATGTTCCGTCAGGTACGGTTCAACGGTAGGTTCAACGGTAGGTTCAACGGTAGGTTCAACGGTAGGTTCAACGGTAGGTTCAACGGTAGGTTCAACGGTAGGTTCAACGGTAGGTT 980  
 P E K P I M C Q E T L Y P E V T P G F P T F I D Q M D Q O D T L L K E Q N A V I D E P K A A  
 B2 exon 3

TTCAAGAATCAACGACCCGGACTCACACCTGTCCATCCCAACCGATGGCTTCGCGCAATTCACACCCAGTTCGGTTCGCTGGTCAACGACCAACTGTTATACGACTCCACAGACATACGATAGCCCGTGA 1120  
 K L F R L A G S E G G Q G D W G I A K A L K E V W N R O R T L S W S N Y T S G S M V I A R O F  
 B2 exon 3

CAAGTCTTGAAGACATCCGAGATCAGATGATACGACGAAAGGCGACCTGCGCATGACGACCGAGGTCACGGCTTCGGTCCCTCCAAAGCGGCAATGTTCTCCAGAACCACCGAATGCCAAGTTCGTAT 1260  
 L D K S A D S I L H Y Q V S A A G A S H G V L T V R K P D G G F A A I N E L V W R I A L N O D  
 B2 exon 3

CTTTGAGTCGGGCATTACCGGGTACGCCATCTGCTCCGATTGACAACACAGGAATCCCAACGCCCAATCGGTTAGTGAACGACACCAAAACGATATCTTCTGAACCAAGAAATCCGGACCGTACAGTTCGGTTCG 1400  
 K L G A N G P V G D O E S Q C C L F G L A G I R Y N F S V L V I D K Q V L F D P G Y L E T G  
 B2 exon 3

CTGGTTCCTTCGGTGAAGCCACCGCGTAATGTACAGCATCACCGAAGCGGCTTGAAGGGTGTATCGGAGAAATACCACATAAAAACATAGATTTTTTAATTTTAAAATCCAACCTCACCTCTTCGCAACACG 1540  
 S T G E T F G G G Y I Y L M V P L P K S P N I  
 B2 exon 3

TTGATCTTCAGACTGCTCGACGCGCAGTCTTCGGAGGCGCCGATGCAAGTATAGCAGGGTTCGCAATGCTGCGAGCAGTCCACGTTTCGGTCCAGTTTTCGGTGGAACTGGTCCGGAATGTTGAAATCTT 1680  
 N I K L S D E C G V I K Q L R R D F H Y C P E C Q Q S C D L T E T W N Q P P V P A  
 B2 exon 2

AGTTTTTTTTTAAATTTCCAGTGGTTCATGTTTATTTTCGTTGAAGTCTCCGGGAAATACAGTGTGGCCAAAATATTTGACACAGTTAGGTGTAATTTCAATGTTCTTTTTTGAATAATTTATCGCAAAAT 1820  
 CTCTAAGGATCTAGGAAAGTTTAGTTTTTCGAGTCGATTATTTTAAACATTTTCAAAATGTCAAAACATATAAATTTTACTGAAAATATATGAAAATTAAGTTTATTCGTAGCCATTGGAAGTGTATTGAAAT 1960  
 GAAAATGATAGGATGCTTAAATTTTAAATTTCTTTTACACAAATGAAATTTTAAATAAGTAAATTTCTGAGGGCTTGAAGAAAGCAGCTAATGTTCAATAACTTCTACTGCTGACTTCATTATCAATT 2100  
 GTACACAAACAAATTTGGATTAAAGTTTATGATCTATCATTTTCAAAATGGAATGACAGTATGTTCTAATTTTGGCTTGTCACTTCAATTTCAATTCACGCTAGACCATGACTGGTTGTACGTTGCC 2240  
 AAGCAGAAAATTTTATACAGTTTTTTTTTAAACAAATATGTCATGTCATGAGGTGAAGTATCAACCTCGCGGAGGTCACACTTTAGATACAGCGCTAATCAGCAGCAGTCCGGTCTTTTATTTGAACAAATGACTT 2380  
 GCACATCCATAAAATGTGAACAGCACTTAAAACAAATTTTCCCTTCTCACCTTAAACCGCAGCTCCCTTCTGGTCCCGGGCGTACGGAATCCCTGGAAGCTGACGTACTCTGCCCAGCAGACAGCTCCG 2520

primer B2 003R  
 K F R L E G E P A R A Y P I G Q F S V Y E Q G L L S V S R  
 B2 exon 1

TTTGGCCCGGACCGGACCGTATTTGGTCTGAACGGTAAAGCTTCCAAACTCATCTACGTAGATTGGAGTTTTTTTGTCTTAAATTTTCGCTATCAGCTGGGTGTGACTCTCGAAGAGTTTTTTTTTGTCTGACT 2660  
 K G R V P G Y K T Q V T L S E L S M  
 B2 exon 1

TAAATTTGACTTGCACAACAGCTCGGTGCGCTGCTCACTGCACTGATGATTTGTAGAGCTGAAAATTTGCGATGTTGGTAGTGTGCTGACTTGCACAAAATATAGATTTTTGAGCACAAAACCGTACTCAAATGAAAT 2800  
 GATTTCTTTAGTTTCCGAAAGCTTTAAAATAATCCCTCAAACCTTGTCCAAAACCTAAAATAAAGATTTTCGGTTGCTTAAATTTTAAAGAAAATAATGATATCCAAATTTTCCGCAAAAACATTTCCCTATGTTACTA 2940  
 CTTTGAAGAGTAGGTTTTTAAAGCCGCTTTTATAACGTTTTTACAGTACAGCAACCGGCGATGGATGTAACCGGCTGAGTGCAAAATGCGCGTAGCAAAAATCTGCTAAAATATTTTAAAGTTTACTAA 3080  
 GTTGTATCGGACATTTAGTTTGAATAACGTTTTTAAATGGTTCTCAAAGGTTTATTTGATACTGAGAAAATTTAGTCAGTACCGCTATCTGAGATGAGTTTAAAATTTAAAATAAAAAAGGCGTGTAAATTTT 3220  
 TGGTACATACCAATTTTTTCACTGAAAATTCAGTAAAGTTTACTAAAATTTTCACTACTGAAATTTTCACTAAACGATTTTCACTAAAGTTTACTGAAATTTCCAGTTAACTGACATTTGCTACTTTGACAGATATTT 3360  
 ACTGAAAATTTAGTAAATTTGGCTGTTAAAAAAGCTGAAAATTTAGTAAAAGAAAGCGCTATTTTTTTTGTGCAAAAACCTAGTAATTTTACTTGTGTGCTAGTTGACAGAACTTGTAGAAATTTGCTAG 3500  
 TATTTCAAGTTTCAAAAATTTGATTTTGTCTGACCAAAAATTTTATGCTTGTAGTACATAAATGTTTCAACCAAAACCAAAACGTTTGTGTTTACATTTGCTGCTTTTCGTTTTTGTTCATTTCAAGGTCAAAACAT 3640  
 AAAACCGCTTTTTTCCGGAACGTCGAAAATGGCGGGTGGCAGAGATAGCAGCAGGATTCGCGATCAAAGCTCAAATAATCCGACAGATGGCGAAAAGCATCGAAGATGACGATTTCAATTTTTCGTTTTCGTTTACAT 3780  
 AGGAATGCAAACTTAAAATTTGAAATTTACAGCTCATAGAACCACTTTTGAAGAAAATTTCAAGTAGTACGATTTGTAACCTTTTCCCTTATAAATTAACCGCAATAAAAAAATTTGAAAAAATTTTGTAA 3920  
 TAATTTGTTTAAAATGATTTGAGCATCAAAGTTTAAACAAAGCAGCAGCTGCAATTTTGTCAAAGTTTGTCTAACGTTGGAATTTAAACAGCAGAAATTCGATACATAACCTCAAAGGGCGGAAAATTTCAATCGACA 4060  
 TCTTTCGCTGTTCTGCTACTCAACACTAGGTTGTCGCTGTCGTTTGGCTCTTGAACGGCAATGTCGATTTTATGTTTATCAAAGTAAACATCAAACCGAGATTTCTCGAAAACGTCAAAAGGACGCTAGGGGGAG 4200  
 AGGGGTAATATGCAACCCCGGGGCAAAAATGCAACCCCTGCTTTTTCGGTTTTTGAAGAAATTTCCGGGGAAAAATCATAGAAAATTTGAAGCTTAAACATTTGCAACCATGCTAGAAAAATTTGAGCTTCGCAATC 4340  
 AAAACAGCTCAAGTTATTTGCAAAAATTTAAAATGTTGTTTTCATCTAATTTTATGAACTTTCATTTTGGGTTTTGGACAATAAAAAAGCCATTTTATGATTTGTAAGTGTGAGTTATAAAGTTTTTACCATAA 4480  
 TCTTCTTATCTGATAGATGAGTCCAAAACATCAAATAATGCAATTTTAAATAAATTTCTGCAAAAACAGCGTGGTGGGGCAGAAATGCAACCGCTTTGAAAGCTCTTTGTAAAAACAGCGGTGCTCATCTAGTGGT 4620  
 GACTAGTAAAACCTGCTTTTACCCTGCTATTTTCCCAATGTGCGGTGCATTTTGGCCGTTGTTGAGTGCATTTTGGCCCAATTTTGGCCCAATTTTGGCCCAATTTTGGCCCAATTTTGGCCCAATTTTGGCCCAATTT 4760  
 TAGAAAATTTGATTTTTGCAACAATTTTGAAGTTTTTAAAGTTTTTTCGATGGATGACGAAGAATAATAGTTGTTTGGGCTGAAAATAATGATAATTTGTAAGGCAATTTTGTTTTCAAATATAGGAAAACCTAA 4900  
 AACGCGTTTTGCTCGAACTTGTATTTGGCTTAAAGTTTCCAAATGAGCAGCTCAAGCAATCAATAATCCAGCTTTTTCAGAGATTTTACAGATAATCACACCAACTCGCGCGCTCACCAGAGCCAACTCAGTTG 5040  
 ATACCGCGTACTGGTATACGTCAGAGTACAGCAGGGTATCCATCATTCGCTCGCAAGATAACAAATAAAAAAAGCAGCTGCTATGATAGGAGGGGGCTTCTGGTGGGGAACAGATGTTTATCTATGACACA 5180  
 ATTGCCATACCACCGTCCATGAAAAGCCCTTATCGAAGTCACTGAGCAACACACTGATTTTAAACATCTCGGAACATGAGACCTCGGTTGTCGCCACCCGTTGTCGCCACCCAGTACCGTCCAGTTAGGGGGTTCGT 5320  
 M D V E H P V V P T Q Y G P V R G V R  
 A2 exon 1

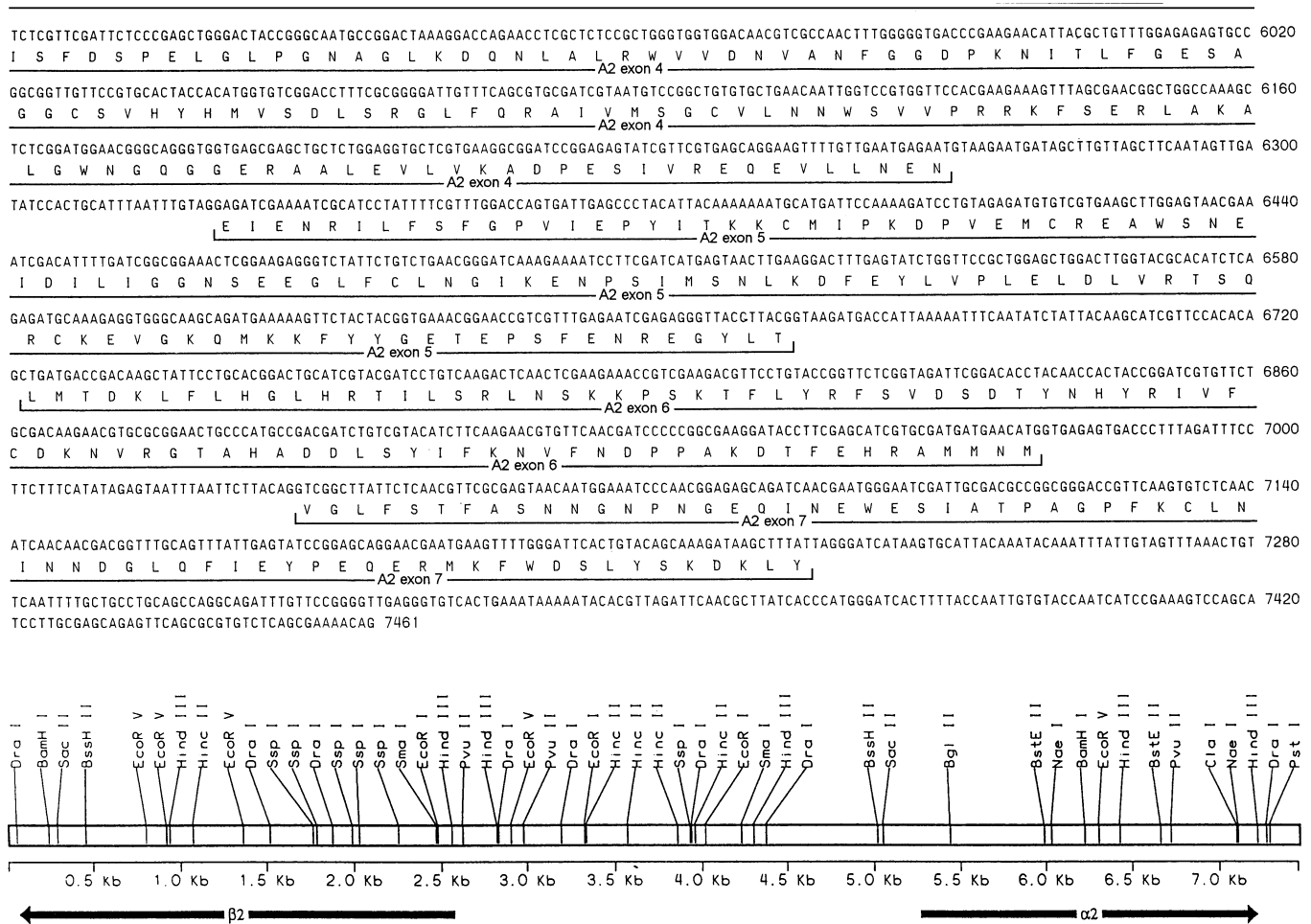
AAGTTGGCCGCAACCGGGGTCGATTACTACAGCTTTCAGCGGATTCGTCAGTGCAGCTCCGTTGGAGAGTTGAGATTTAAGTGGGATTTGGGTTATGATTTGAGGATCTTTGTAAGTTGATTTTCTAGGA 5460  
 K L A A T G V D Y Y S F Q R I P Y V Q P P V G E L R F K  
 A2 exon 1

primer A2 005R  
 TGCTAACCCGCAAAACCGTGGACGGAACCGTTGAGCTGACCGTCCAAAGGCGCAGTGGGTTCAATACAGTAAAGCTGCTGAACAAAATCATTGGAAGAGAGGACAGCTTGCACATGAATGTTTACGAAGACGTA 5600  
 A Q P P K P W T E P L D C T V Q G P G G Y Q Y S K L L N K I I G R E D S L H M N V F T K N  
 A2 exon 2

GTTGAATGAGCTTTGAACTTGTGGTGTAGCAGTCTGATTTGATGTCGATTAATAATTTAGCTCGATAGCAAGCAGTTATACCTGTGATGTTGATATCCAGGTTGGGCTTTATGAGAGGATCTAGTGGAG 5740  
 L D S K Q L L P V M L Y I H G G A F M R G S S G  
 A2 exon 3

TGGAGATGACGCTCCGACTATCTCATACAGAAGGATGTGGTGTTCGTTGCTTCAACTACGAATAGGCGACATGGGATGATGATGTTCCCTGGTAGACCACTTGAACAATATACAAATATCTTTCAGATTC 5880  
 V E M Y G P D Y L I Q K D V V F V S F N Y R I G A L G  
 A2 exon 3

Figure 1 For legend see facing page.



**Figure 1** DNA sequence of PeIRR *estα2*, *estβ2* and intergenic spacer (top) and restriction map of PeIRR *estα2/estβ2* amplicon (bottom)

Top, the complete DNA sequence of the co-amplified esterase genes, *estα2*' and *estβ2*', from a Sri Lankan OP-resistant strain, PeIRR, of the mosquito *C. quinquefasciatus* is shown. The DNA originated from a recombinant bacteriophage, φRRA2, which was isolated by screening a PeIRR genomic library with an *estα2*' cDNA probe. The two genes lie in a head to head configuration and the Figure covers 7461 bp of genomic DNA. Each exon and its predicted amino acid sequence are underlined (note *estβ2*' sequence extends from 2574 to 0066 and is given as the non-coding strand, but the amino acid translation is from the correct coding orientation). The locations of the primers used to amplify the intergenic spacer are given. Bottom, a linear illustration of the two co-amplified genes from PeIRR with the positions of a number of common restriction enzyme sites marked.

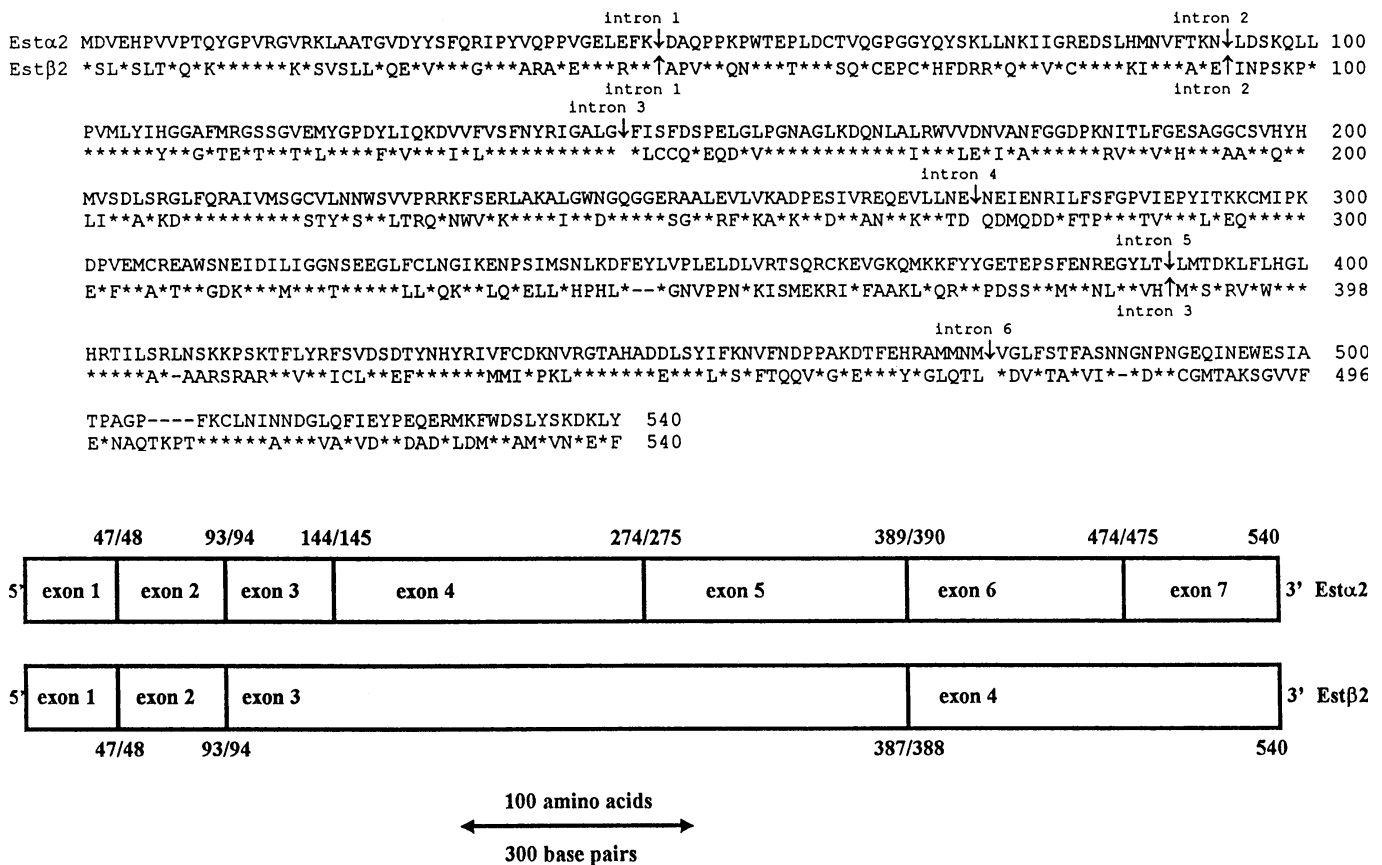
a decade and then with fenitrothion from 1988. The strain, colonized in 1991, was maintained without selection. The TANGA strain originated from a chlorpyrifos-resistant field population and was colonized in 1985. The strain has been maintained with intermittent chlorpyrifos selection pressure. The SPERM strain was collected in Jeddah, Saudi Arabia in 1989. Field selection of this population was with temephos and then a wide range of pyrethroids. Immediately after colonization it was selected for 20 generations with permethrin and subsequently selected intermittently with malathion and temephos. All three strains have the co-elevation of the *Estα2*'/*Estβ2*' esterases as their only mechanism of OP-resistance.

The TEM-R strain of *C. quinquefasciatus* from California containing the amplified *estβ1*' was kindly supplied by Professor G. Georgioui (U. C. Riverside).

#### Isolation of recombinant genomic clones possessing both the *estα2*' and *estβ2*' genes from a PeIRR genomic library

Genomic DNA was isolated from PeIRR fourth-instar larvae as previously described [10]. The λGEM-11 vector (Promega) was

used for the library construction and 10 μg of genomic DNA was partially digested with *Sau3AI* so the majority of DNA pieces were between 15 and 23 kb. In a 10 μl ligation, 1 μg of λGEM-11 arms was ligated to approx. 0.15 μg of partially digested PeIRR genomic DNA. The ligated DNA was packaged using Packagene Extracts (Promega) and approx. 0.5 × 10<sup>6</sup> clones were obtained, of which 50000 were screened. The library was hybridized with a <sup>32</sup>P-labelled *estα2*' cDNA probe [9] (specific radioactivity > 2 × 10<sup>6</sup> c.p.m./mg) at 65 °C overnight in hybridization buffer [5 × Denhardt's solution (where 1 × Denhardt's is 0.02% Ficoll/0.02% polyvinylpyrrolidone/0.02% BSA), 6 × SSC (where 1 × SSC is 0.15 M NaCl/0.015 M sodium citrate), 0.1% (w/v) SDS, 0.1% (w/v) sodium pyrophosphate, 5% (w/v) poly(ethylene glycol) 8000 and 100 mg/ml boiled sheared herring sperm DNA]. The final washes were at 65 °C in 0.1 × SSC and 0.1% (w/v) SDS for 20 min. A large number of positive plaques was seen. Four positive plaques were purified and the recombinant DNA from each was isolated. Screening of the library with an *estβ2*' probe showed that the majority of *estα2*'-positive plaques were also positive for *estβ2*', which suggests that the two genes are located close together in the genome.



**Figure 2** Alignment of predicted *est $\alpha$ 2* and *est $\beta$ 2* amino acid sequences of PeIRR

Top, an alignment of the assumed amino acid sequences of the amplified Est $\alpha$ 2<sup>1</sup> and Est $\beta$ 2<sup>1</sup> esterases from a Sri Lankan OP-resistant strain of the mosquito *C. quinquefasciatus* is shown. There is 47% similarity between the two sequences. The positions of the introns from the corresponding genes are shown in order to demonstrate further similarities between the genes. Introns 1 and 2 are at the same positions in both the *est $\alpha$ 2*<sup>1</sup> and *est $\beta$ 2*<sup>1</sup> genes. Exon 3 of the *est $\beta$ 2*<sup>1</sup> gene is split into exons 3–5 in the *est $\alpha$ 2*<sup>1</sup> gene and exon 4 corresponds to exons 6–7. Bottom, a schematic representation of the alignment shown above demonstrating the similarity between the exons from the two esterase genes in relation to their size and corresponding amino acid positions in the expressed esterases.

The DNA from each positive bacteriophage clone was doubly digested with *Sac*I (present in the multiple cloning site of the bacteriophage arms) and *Bam*HI (present in the *est $\alpha$ 2*<sup>1</sup> cDNA [9]). The fragments were separated on a 0.8% agarose gel, transferred to nylon membranes (Amersham) and hybridized with two <sup>32</sup>P-labelled *est $\alpha$ 2*<sup>1</sup> cDNA probes as above. One probe was specific to the 5' of the *est $\alpha$ 2*<sup>1</sup> cDNA *Bam*HI site and the other was 3' to this site. Recombinant bacteriophage,  $\phi$ RRA2, contained a 6 kb *Sac*I–*Bam*HI fragment, to which the 5' *est $\alpha$ 2*<sup>1</sup> cDNA probe hybridized, and a 4 kb *Sac*I–*Bam*HI fragment, to which the 3' *est $\alpha$ 2*<sup>1</sup> cDNA probe hybridized. These two fragments were subcloned into the *Sac*I and *Bam*HI sites of pBluescript II SK(+) (Stratagene) and named p5A2 (containing the 5' end of the gene) and p3A2 (containing the 3' end). With the use of both internal oligonucleotide primers and further subcloning, both p3A2 and p5A2 were sequenced using Sequenase version 2.0 (Amersham). Both strands of DNA were sequenced.

#### Isolation of *est $\alpha$* and *est $\beta$* intergenic sequence from insecticide-resistant and -susceptible mosquitoes

Two oligonucleotide primers were designed and used to amplify by PCR a DNA fragment containing the 5' ends of the *est $\alpha$ 2*<sup>1</sup> and *est $\beta$ 2*<sup>1</sup> genes and the intergenic region. The two 18-mers used were B2-003R (5'-CTTAAACCGCAGCTCCCC-3', complementary to the 5' end of the *est $\beta$ 2*<sup>1</sup> gene) and A2-005R (5'-

CAACAGGAGGCTGCACGT-3', complementary to the 5' end of the *est $\alpha$ 2*<sup>1</sup> gene). The product was amplified from  $\phi$ RRA2 (positive control) and from genomic DNA from numerous mosquito strains. The 50  $\mu$ l PCR mixture contained 50 ng of genomic DNA (500  $\mu$ g in the case of  $\phi$ RRA2), 50 ng of each primer, 0.5 mM dNTPs, 2 mM MgCl<sub>2</sub>, 2.5 units each of *Taq* DNA polymerase (Promega) and *Taq* antibody (Clontech), buffered in *Taq* DNA polymerase buffer [50 mM KCl, 10 mM Tris/HCl (pH 9.0 at 25 °C), 0.1% Triton X-100]. The reaction mixture was heated in an Omn-E DNA Thermal Cycler (Hybaid) to 94 °C for 3 min. The mixture was removed and 2.5 units of *Taq* extender (Stratagene) was added to the mixture which was then heated to 94 °C for a further 2 min. Thirty cycles of amplification of 52 °C for 90 s, 72 °C for 3 min and 94 °C for 30 s per cycle were carried out. The PCR products were separated on a 1% agarose gel, and the major DNA product of the amplification and PeIRR genomic DNA was subcloned into pBluescript II SK(+) (Stratagene) which had been restricted with *Eco*RV and subsequently T-tailed [22]. The insert of the recombinant plasmid was sequenced as above and compared with the sequence obtained from p5A2.

#### RESULTS

The screening of the *C. quinquefasciatus* PeIRR genomic library with an *est $\alpha$ 2*<sup>1</sup> cDNA probe resulted in a large number of

Barbie box consensus	ATCAAAAGCTGGAGG		
<b>Upstream of <math>\alpha 2</math></b>			
3703-3717	11/15	ATCAAAAGCTcaAaa	Pel RR only
3945-3959	9/15	ATCAAAAGtTaaCa	Pel RR only
Low homology region	9/15	tTgAAAAGtTGaAat	Pel SS only
2813-2827	8/15	tTCgAgAGCTtAca	Pel SS
	9/15	tTCgAAAAGCTtAaa	Pel RR
3127-3141	9/15	cTCAAAAGtTtAtt	Pel SS
	8/15	cTCAAAAGtGttAtt	Pel RR
3388-3401	9/15	AaaAAAA-CTGaAat	Pel SS
-3402	10/15	AaaAAAAGCTGaAat	Pel RR
<b>Upstream of <math>\beta 2</math></b>			
3830-3816	9/15	cTCAAAAGtTGttct	Pel RR only
4248-4233	10/15	gagAAAAGCaGGgGG	Pel RR only

**Figure 3** Barbie box homologous sequences

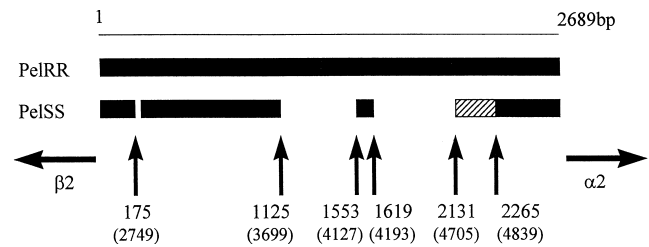
Barbie box consensus sequence and sequences sharing > 60% homology with this 15 bp sequence in the intergenic spacer between the esterases in PelRR and PelSS are shown. Note that for *est $\beta 2$* , the Barbie-box-like sequences given are the complements of those seen in Figure 1 (top) at the indicated locations.

positive plaques reflecting the amplified nature of this gene [9]. Restriction enzyme fragments from a recombinant bacteriophage  $\phi$ RRA2, containing the *est $\alpha 2$*  gene were subcloned and subsequently sequenced. The *est $\alpha 2$*  gene contained within two plasmids, p5A2 and p3A2, was compared with the sequence of the *est $\alpha 2$*  cDNA [9]. This demonstrated the presence of seven exons within the *est $\alpha 2$*  gene. The six introns were short [ranging from 52 bp (intron 5) to 71 bp (intron 2)] and all started with -GT- and ended with -AG-.

Screening of the genomic clones positive for *est $\alpha 2$*  with a probe specific for *est $\beta 2$*  showed that the vast majority of them were positive for both esterase genes, suggesting that these genes were less than 15 kb apart. Further sequencing of the p5A2 plasmid revealed the presence of *est $\beta 2$*  genomic sequence, based on the *est $\beta 2$*  cDNA sequence [10], within 3 kb of the *est $\alpha 2$*  sequence. The two genes were located in a head to head configuration. The full *est $\beta 2$*  gene sequence was compared with that of *est $\beta 1$*  [23]; both have four exons and three introns which are similar in size and location.

Figure 1 shows the 7461 bp of DNA sequence from p5A2 and p3A2 which encompasses the *est $\alpha 2$*  and *est $\beta 2$*  genes and the intergenic region between them. Identical intron/exon boundaries were found in the two genes. In addition, in *est $\alpha 2$*  the equivalent of *est $\beta 2$*  exon 3 has been divided into exons 3–5 and the equivalent of *est $\beta 2$*  exon 4 into exons 6–7 (Figure 2). Despite the common intron/exon boundaries, the homologies between equivalent *est $\alpha 2$*  and *est $\beta 2$*  introns are no greater than the homologies between the introns within the genes (~30% similarity), an ancient duplication and subsequent divergence of the two genes.

To confirm that the co-amplification of *est $\alpha 2$*  and *est $\beta 2$*  reflected the situation in the genome and was not the product of a recombination event between bacteriophage during construction of the genomic library, PCR was carried out on DNA from a number of *Culex* strains. The primers used amplified a DNA fragment from  $\phi$ RRA2 of the expected size (approx. 3 kb) and also amplified the same fragment size from the original genomic DNA of the PelRR strain. An identical size PCR product was also produced from the TANGA, DAR91 and SPERM strains. Initial experiments suggest that this 3 kb band may be diagnostic of co-amplified *est $\alpha 2$* /*est $\beta 2$* , as in the susceptible PelSS strain a smaller 2 kb band occurred. Subsequent analysis of four strains



**Figure 4** Intergenic spacers of PelSS and PelRR

A schematic representation is shown of the alignment of the intergenic spacers between the esterases in the PelRR and PelSS strains of the mosquito *C. quinquefasciatus* showing the positions of the insertions and deletions that have occurred between the amplified and non-amplified genes. Numbering (bp) runs from 1 at the start of the *est $\beta 2$* ' sequence to 2689 at the start of the *est $\alpha 2$* ' sequence in PelRR [for ease of reference, relative positions within the 7461 bp genomic sequence of Figure 1 (top) are given in parentheses below]. The intergenic spacer in the PelRR amplicon contains two major insertions and a 5 bp insertion at positions 175–180 (2749–2753) compared with the non-amplified PelSS intergenic spacer. Insertions in PelRR from 175 to 179, 1125 to 1553 and 1619 to 2131 show 91–96% homology for the remaining sequence, with the exception of 2131–2265, where homology decreases to 28%.

of *Culex* from different continents without the amplified *est $\alpha 2$* /*est $\beta 2$* , including the TEM-R strain which has amplified *est $\beta 1$* , also gave a 2 kb product.

The DNA sequence of the products from the PelRR and PelSS genomic amplifications were subcloned and sequenced. The PelRR sequence was identical with the sequence obtained from the corresponding section of DNA from  $\phi$ RRA2. As the two esterases are in a head to head configuration, the intergenic region should contain the promoters and possible *cis*-acting regulatory elements for these genes. The intergenic region contained numerous possible TATA boxes. A number of putative promoter elements with high homology (8–11 nucleotides of 15) to the Barbie box are found within the intergenic spacer of both strains. Upstream of *est $\alpha 2$* , three Barbie-box-like elements are located in homologous positions within both PelRR and PelSS, although in two of these sequences there are point alterations within the core AAAG motif likely to abrogate recognition in PelSS, whilst the corresponding PelRR sequences retain the complete core motif. In addition, there is a single Barbie-box-like sequence unique to PelSS, and two such elements unique to PelRR. Upstream of *est $\beta 2$*  there are two Barbie-box-like sequences, both seen only in PelRR (Figure 3).

The PelRR and PelSS intergenic spacers differed by three insertions in the larger (amplified) spacer (Figure 4). There were two large insertions of 428 and 512 bp and one small insertion of 5 bp in the PelRR intergenic spacer compared with PelSS. Homology between the two intergenic spacers outside these insertions was more than 90%, with the exception of a 134 bp section adjacent to the site of the third insertion which had only 28% homology between PelRR and PelSS suggesting that a section of DNA has been deleted from the PelRR spacer during the third insertion.

## DISCUSSION

The amplified *est $\alpha 2$*  and *est $\beta 2$*  genes associated with OP-resistance have now been isolated and sequenced from the PelRR strain of *C. quinquefasciatus*. The two genes lie in a head to head configuration with only 2689 bp separating the ATG start codons of the two genes. Thus the two amplified esterase genes are present on the same amplicon. This is the first time that two functional genes have been shown to be present on the same amplicon in OP-resistant insects. The presence of both genes on

the same amplicon explains why these esterases are found in complete linkage disequilibrium. The earlier reports of linkage distances between these esterases [16,17] are therefore incorrect. The small number of insects apparently with only one of these esterases is likely to have been caused by minor localized changes in EDTA concentration in a small number of electrophoresis gels resulting in scoring of spurious nulls for one esterase. Indeed given the multiple copies of these esterases in the genome and their co-amplification it would be impossible for recombination between them to be detected through electrophoretic analysis, as this would require multiple recombination events between all copies simultaneously.

The *estα2<sup>1</sup>/estβ2<sup>1</sup>* intergenic PCR product from the amplification of genomic DNA from the Sri Lankan *C. quinquefasciatus* strain, PeLRR, was identical with its counterpart sequenced from the genomic library. This identity demonstrates that the genomic clone sequence was real and did not arise from a recombination event during the construction of the library. This could have occurred due to the large number of *estα2<sup>1</sup>/estβ2<sup>1</sup>* amplicons within the PeLRR strain and the similarities between the two genes. An intergenic product of the same size (~ 3 kb) to that from PeLRR was produced from genomic DNA of three other *C. quinquefasciatus* strains with elevated *Estα2<sup>1</sup>* and *Estβ2<sup>1</sup>*. In contrast, a smaller intergenic product (~ 2 kb) was amplified from genomic DNA of the susceptible PeLSS strain. Thus the selection of the PeLSS strain from the same parental PeL population as PeLRR, through multiple single family selection for low esterase activity [24], appears to have removed all copies of the amplified esterases. It is not possible from the current PCR results to show whether the non-amplified esterase loci are still present in the PeLRR population, as the high copy number of the amplified esterases would have swamped the PCR reaction, resulting only in a 3 kb product. We have, however, evidence that an intergenic spacer of < 2 kb is a common feature of the non-amplified esterase genes, as this occurred in four other strains of *Culex* without the amplified *Estα2<sup>1</sup>/Estβ2<sup>1</sup>*. These strains included the Californian TEM-R strain which has the amplified *estβ1<sup>1</sup>* gene. We already know that there is no *estα* locus in close association with the amplified *estβ1<sup>1</sup>* in TEM-R, hence the presence of the smaller intergenic spacer suggests that non-amplified loci for both *estα* and *estβ* still occur in TEM-R despite high levels of laboratory insecticide selection pressure over several decades.

The *estα* and *estβ* genes both code for proteins of 540 amino acids, which share 47% similarity. They are found in a linear array in both their amplified and non-amplified forms and they have conserved intron/exon boundaries, although three further introns have either been gained by *estα2<sup>1</sup>* or lost by *estβ2<sup>1</sup>* and the intronic sequences are highly diverged. This strongly suggests that the two genes arose through an ancestral gene-duplication event. This event is likely to have occurred before *Culex* speciation since elevated *Estα* and/or *Estβ* esterases have been identified in at least two further *Culex* species, *C. tarsalis* and *C. tritaeniorhynchus* [19,25,26]. The outcome of this initial duplication event could have resulted in the genes being located next to each other. This is seen in *Drosophila melanogaster*, where an ancestral duplication of esterase genes which now lie next to each other is predicted for *est6* and *estP* as well as their homologues in *Drosophila pseudoobscura* (*est5* and *est4*) and *Drosophila buzzatii* (*estJ* and *estJ*) [27,28].

The co-amplified nature of the *estα2<sup>1</sup>* and *estβ2<sup>1</sup>* means that individual mosquitoes contain equal numbers of copies of both genes. Under standard laboratory rearing conditions, the amounts of active esterase produced from the two loci are, however, different [29]. The possibility exists that the two genes

are differentially regulated. The intergenic region between *estα2<sup>1</sup>* and *estβ2<sup>1</sup>* should contain the *cis*-acting promoters and regulatory elements related to these genes. Many regulatory elements, for example ARE mediating induction by monofunctional inducers and Barbie box forming the barbiturate-responsive element, contain a core AAAG motif [30]. A 15 bp consensus element has been proposed as the element responsible for barbiturate inducibility, with most inducible genes from a broad range of species containing between 8 and 12 of the 15 consensus nucleotides. A computer-aided search of the intergenic spacers of PeLRR and PeLSS revealed a total of six sequences upstream of *estα2<sup>1</sup>* and two upstream of *estβ2<sup>1</sup>*. Of these, four are seen only in PeLRR, in the regions of insertion, and one is unique to PeLSS, in the region of low homology. As far as we are aware there are no reports that barbiturates or any other xenobiotics can regulate the expression of these or other esterase genes. Treatment of F344/NCr rats with phenobarbital and a range of structurally diverse compounds including 1,1,1-trichloro-2,2-bis(*p*-chlorophenyl)ethane (DDT) results in induction of several enzymes. For example, the activities of CYP2B1 and epoxide hydrolase are increased, as is the level of glutathione S-transferase mRNA [31]. Each of these genes is known to possess a Barbie box within its promoter. It is feasible therefore that other xenobiotics such as OPs may also interact with Barbie boxes, and these interactions would be affected by the number and arrangement of these elements within a promoter. Over and above the high level of genomic amplification of *estα2<sup>1</sup>* and *estβ2<sup>1</sup>* in PeLRR, it is possible that the insertions within the intergenic spacer contribute to altered regulation of one or both of these genes via increased Barbie box copy number. There are no other obvious regulatory features within the two regions of insertion. Conclusive identification of *cis*-acting regulators and promoter sequences awaits further molecular analysis of these esterase genes.

This work was supported by the Wellcome Trust.

## REFERENCES

- Chang, C. K. and Jordan, T. W. (1983) *Pest. Biochem. Physiol.* **19**, 190–195
- Chen, W.-L. and Sun, C.-N. (1994) *Ins. Biochem. Mol. Biol.* **24**, 347–355
- Field, L. M., Williamson, M. S., Moores, G. D. and Devonshire, A. L. (1993) *Biochem. J.* **294**, 569–574
- Kao, L. R., Motoyama, N. and Dauterman, W. C. (1985) *Pest. Biochem. Physiol.* **23**, 228–239
- Malkenson, N. C. D., Wood, E. J. and Zerba, E. N. (1984) *Ins. Biochem.* **4**, 481–486
- Nagata, T. and Ohira, Y. (1986) *Appl. Entomol. Zool.* **21**, 216–219
- Parker, A. G., Russell, R. J., Delves, A. C. and Oakeshott, J. G. (1991) *Pest. Biochem. Physiol.* **41**, 305–318
- Peiris, H. T. R. and Hemingway, J. (1993) *Bull. Entomol. Res.* **83**, 127–132
- Vaughan, A. and Hemingway, J. (1995) *J. Biol. Chem.* **270**, 17044–17049
- Vaughan, A., Rodriguez, M. and Hemingway, J. (1995) *Biochem. J.* **305**, 651–658
- Aldridge, W. N. (1953) *Biochem. J.* **53**, 110–117
- Hemingway, J., Callaghan, A. and Amin, A. M. (1990) *Med. Vet. Entomol.* **4**, 275–282
- Raymond, M., Callaghan, A., Fort, P. and Pasteur, N. (1991) *Nature (London)* **350**, 151–153
- Villani, F., White, G. B., Curtis, C. F. and Miles, S. J. (1983) *Bull. Entomol. Res.* **73**, 153–170
- Villani, F. and Hemingway, J. (1987) *Pest Biochem. Physiol.* **27**, 218–228
- Callaghan, A., Hemingway, J. and Malcolm, C. A. (1993) *Biochem. Genet.* **31**, 459–472
- Wirth, M. C., Marquine, M., Georghiou, G. P. and Pasteur, N. (1990) *J. Med. Entomol.* **27**, 202–206
- Mouches, C., Pasteur, N., Berge, J. B., Hyrien, O., Raymond, M., De Saint Vincent, B. R., De Silvestri, M. and Georghiou, G. P. (1986) *Science* **233**, 778–780
- Raymond, M., Beyssat-Arnaouty, V., Sivasubramanian, N., Mouches, C., Georghiou, G. P. and Pasteur, N. (1989) *Biochem. Genet.* **27**, 417–423
- Karunaratne, S. H. P. P., Jayawardena, K. G. I., Hemingway, J. and Ketterman, A. J. (1994) *Biochem. Soc. Trans.* **22**, 127S

- 
- 21 Karunaratne, S. H. P. P., Jayawardena, K. G. I., Hemingway, J. and Ketterman, A. J. (1993) *Biochem. J.* **294**, 575–579
- 22 Marchuk, D., Drumm, M., Saulino, A. and Collins, F. S. (1991) *Nucleic Acid Res.* **19**, 1154
- 23 Mouches, C., Magnin, M., Berge, J. B., De Silvestri, M., Beysat, V., Pasteur, N. and Georghiou, G. P. (1987) *Proc. Natl. Acad. Sci. U.S.A.* **84**, 2113–2116
- 24 Amin, A. M. and Peiris, H. T. R. (1990) *Med. Vet. Entomol.* **4**, 269–273
- 25 Prabhaker, N., Georghiou, G. P. and Pasteur, N. (1987) *J. Am. Mosq. Control Assoc.* **3**, 473–475
- 26 Takahashi, M. and Yasutomi, K. (1987) *J. Med. Entomol.* **24**, 595–603
- 27 Brady, J. P., Richmond, R. C. and Oakeshott, J. G. (1990) *Mol. Biol. Evol.* **7**, 525–546
- 28 Collet, C., Nielson, K. M., Russell, R. J., Karl, M., Oakeshott, J. G. and Richmond, R. C. (1990) *Mol. Biol. Evol.* **7**, 9–28
- 29 Karunaratne, S. H. P. P. (1994) Ph.D. Thesis, University of London
- 30 Hayes, J. D. and Pulford, D. J. (1995) *Crit. Rev. Biochem. Mol. Biol.* **30**, 445–600
- 31 Lubet, R. A., Dragnev, K. H., Chauhan, D. P., Nims, A., Diwan, B. A., Ward, J. M., Jones, C. R., Rice, J. M. and Miller, M. S. (1992) *Biochem. Pharmacol.* **43**, 1067–1078

---

Received 31 October 1996/3 March 1997; accepted 7 March 1997