Molecular cloning and characterization of the complete acetylcholinesterase gene (*Ace1*) from the mosquito *Aedes aegypti* with implications for comparative genome analysis

Akio Mori, Neil F. Lobo, Becky deBruyn, David W. Severson*

Center for Global Health and Infectious Diseases, Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556-5645, USA

Abstract

Insensitive acetylcholinesterase (AChE) has been shown to be responsible for resistance to organophosphates and carbamates in a number of arthropod species. Some arthropod genomes contain a single *Ace* gene, while others including mosquitoes contain two genes, but only one confers insecticide resistance. Here we report the isolation of the full-length cDNA and characterization of the complete genomic DNA sequence for the *Ace1* gene in the yellow fever mosquito, *Aedes aegypti*. The *Ace1* homolog in other mosquito species has been associated with insecticide resistance. The full-length cDNA consists of 2721 bp and contains a 2109 bp open reading frame that encodes a 702 amino acid protein. The amino acid sequence is highly conserved with that of other mosquitoes, including greater than 90% identity with *Culex* spp. and about 80% identity with *Anopheles gambiae*. The genomic DNA sequence includes 138,970 bp and consists of eight exons with seven introns ranging from 59 to 114,350 bp. Exons 2 and 8 show reduced amino acid conservation across mosquito species, while exons 3–7 are highly conserved. The *Ace1* introns in *Ae. aegypti* reflect a high frequency of repetitive sequences that comprise about 45% of the total intron sequence. The *Ace1* locus maps to the *p*-arm of chromosome 3, which corresponds to the orthologous genome regions in *Culex* spp. and *An. gambiae*.
© 2007 Elsevier Ltd. All rights reserved.

Keywords: Acetylcholine; Gene annotation; Insecticide resistance; Linkage mapping; Repetitive elements

1. Introduction

Acetylcholinesterase (AChE) is critical for hydrolysis of acetylcholine at cholinergic nerve synapses and is a target for organophosphate (OP) and carbamate insecticides. Insensitive AChE has been shown to be responsible for resistance to OPs and carbamates in a number of arthropod species (for review, see Fournier and Mutero, 1994). The first insect AChE gene (*Ace*) was isolated from *Drosophila melanogaster* (Hall and Spierer, 1986). Subsequently, *Ace* mutations conferring resistance were documented in Cyclorrhapha dipteran species including, *D. melanogaster* (Mutero et al., 1994), *Musca domestica* (Kozaki et al., 2001; Walsh et al., 2001), and *Lucilia cuprina* (Chen et al., 2001). However, efforts to identify *Ace*

mutations conferring resistance were unsuccessful in Orthorrhapha dipterans, including the mosquitoes *Culex pipiens* (Malcolm et al., 1998) and *Culex tritaeniorhynchus* (Mori et al., 2001), despite documentation of insensitive AChE-mediated resistance in both species. Of note, the resistance phenotype mapped to a different genome location than the *Ace* locus, suggesting the existence of an independent molecular mechanism.

The existence in some insects of a paralogous AChE gene (*Ace1*) was first determined for the green bug, *Schizaphis graminum* (Gao et al., 2002), although some insect genomes such as *D. melanogaster* carry only one AChE gene (now called *Ace2*). Insensitive AChE-mediated resistance due to nonsynonymous substitutions in the *Ace1* gene has since been documented in a number of insects, including the mosquitoes *Anopheles gambiae* and *Cx. pipiens* (Alout et al., 2007; Weill et al., 2002, 2003), *Anopheles albimanus* (Weill et al., 2004a), and *Cx. tritaeniorhynchus* (Nabeshima

et al., 2004). The amino acid coding sequence was reported for the mosquito, *Aedes aegypti*, *Ace1* gene, although insensitive AChE-mediated insecticide resistance is rare in natural *Ae. aegypti* populations and no resistance-specific mutations have been identified (Bisset et al., 2006; Weill et al., 2004b).

Among mosquito species, both cDNA and genomic DNA sequence information are available only for *An. gambiae*. Because of the importance of *Ace1* in insecticide resistance and the interest in comparative genome analysis between *An. gambiae* and *Ae. aegypti*, here we report on the isolation and characterization of the full-length cDNA and complete genomic DNA sequence organization for *Ae. aegypti Ace1* and its comparative organization with *An. gambiae*.

2. Materials and methods

2.1. Mosquitoes

Ae. aegypti strains used in this study included: Liverpool (cDNA and genomic DNA isolation), MOYO-R and RED (genetic mapping). Larvae were reared on a bovine liver powder suspension. Adults were maintained on a 5% sugar solution. Mosquitoes were maintained in an environmental chamber at 26 °C, a relative humidity of 84%, and under a 16-h light/8-h dark cycle with a 1-h crepuscular period at the beginning and end of each light cycle.

2.2. Full-length Ace1 cDNA sequence determination

To isolate full-length cDNA, total RNA was extracted from pooled *Ae. aegypti* adults (Liverpool strain) using TRI-REAGENT (Molecular Research Center, Inc.). Near-full-length cDNA was synthesized using *Ace1* specific primers (AaAce1-F7: 5′-CCGAGACCAAAGAACT-CAGC and AaAce1-R8: 5′-ATGCAGTGACACCGAGA-CAG) based on a partial *Ae. aegypti Ace1* sequence available in GenBank (AJ621915) and the SMART PCR cDNA synthesis kit (Clontech). Full-length cDNA was obtained by 5′–3′ RACE using the Gene Racer Core Kit (Invitrogen). Total RNA was treated with calf intestinal phosphatase to remove 5′ phosphates, and then with tobacco acid pyrophosphatase to remove the 5′ mRNA cap structure which leaves a 5′ phosphate required for ligation to the GeneRacer RNA Oligo (5′-CGACUG-GAGCACGAGGACACUGACAUGGACUGAAGGA-GUAGAAA-3′). The GeneRacer RNA Oligo was ligated to the 5′ end of the decapped mRNA using T4 RNA ligase. The ligated mRNA was reverse-transcribed using Super-Script II reverse transcriptase and the GeneRacer Oligo dT Primer (5′-GCTGTCAACGATACGCTACGTAACGG-CATGACAGTG(T)$_{18}$) to synthesize complete cDNA carrying GeneRacer kit-specific oligonucleotide sequences at the 5′- and 3′-ends. The 5′-, 3′-cDNA ends were amplified with gene specific primers: 5′-RACE (AaAce1-R7: 5′-GGAGCTTCGAGTGTGAGTCC), 3′-RACE

(AaAce1-F5: 5′-TGGGCTCAACACAACCTATG) and GeneRacer primers: 5′-Primer (5′-GCACGAGGACAC UGACAUGGACUGA: Position 9–33 of GeneRacer RNA Oligo), 3′-Primer (5′-GCTGTCAACGATACGC-TACGTAACG: Position 1–25 of GeneRacer Oligo dT Primer). After 20 fold dilution of full-length RACE cDNA, 1 μl was added to a 25 μl PCR reaction (35 cycles of 94 °C for 1 min, 63 °C for 1 min and 72 °C for 2 min) with 5 min preheating at 94 °C and 10 min final extension at 72 °C. Amplified 5′- and 3′-RACE DNA was cloned using the TOPO TA cloning kit (Invitrogen) following the supplier's protocol. Sequencing reactions were performed using Big Dye Terminators and analyzed using an ABI 310 or ABI 3700 sequencer (Perkin-Elmer Applied BioSystems). The full-length cDNA sequence was submitted to the BLASTX program (http://www.ncbi.nlm.nih.gov) for homology searches and confirmation as the *Ace1* gene. Multiple sequence alignment of the Ace1 protein to homologs from the mosquitoes *Cx. pipiens*, *Culex pipiens (pallens)*, *Cx. tritaeniorhynchus*, and *An. gambiae* (GenBank accession nos. AJ489456, AY762905, AB122152, AJ488492, respectively) was performed using ClustalW (http://pbil. univ-lyon1.fr/pbil).

2.3. Determination of Ace1 genomic DNA sequence

PCR primers were designed based on a partial *Ae. aegypti Ace1* sequence available in GenBank (accession no. AJ428049) using Primer3 (http://frodo.wi.mit.edu/cgi-bin/ primer3/primer3_www.cgi). A 204 bp genomic DNA fragment was amplified in a total volume of 25 μl containing 50 mM KCl, 10 mM Tris (pH 9.0), 0.1% Triton X, 1.5 mM MgCl$_2$, 200 μM dNTPs, 2.5 pmol each of forward primer AaAce1-F2 (5′-ATGGTCGGAGATTATCACTTCA) and reverse primer AaAce1-R3 (5′-ATCCTCCATGTACCC-CAGATCA), 5 ng genomic DNA (Liverpool strain) and 1 unit of *Taq* polymerase. PCR conditions were 5 min at 94 °C, followed by 30 cycles of 1 min at 94 °C, 1 min at 62 °C, and 2 min at 72 °C, with a final extension at 72 °C for 10 min. The PCR amplicon was used as a radiolabeled probe to screen an *Ae. aegypti* BAC library (Liverpool strain) as previously described (Jiménez et al., 2004; Severson, 1997). BAC plasmid from one positive clone was isolated and purified following our standard procedures (Jiménez et al., 2004), and was subjected to direct cycle sequencing on both strands by primer walking with gene specific primers and the ABI Prism Big Dye Terminator kit v. 3.1. Sequencing reactions were analyzed using an ABI 310 or ABI 3700 automated sequencer (Perkin-Elmer Applied BioSystems). Sequences were assembled using CAP3 (Huang and Madan, 1999).

BLASTn analysis of the full-length cDNA sequence (see Section 2.2) against the *Ae. aegypti* whole genome sequence assembly (Liverpool-IB12 strain) that became available at The Broad Institute (http://www.broad.mit.edu/annotation/ disease_vector/aedes_aegypti/) identified an assembly

(supercontig1.9, GenBank accession no. CH477194) that included the first exon and the large first intron sequences for the complete *Ace1* gene. The first intron contained two gaps of unknown size that spanned three contig assemblies (GenBank accession nos. AAGE02000752, AAGE02000 753, AAGE02000754). Gap sequences were determined following PCR amplification of a genomic fragment spanning the gap with specific primers anchored in the linking contig sequences or by isolation and cloning of a restriction fragment containing the gap from the *Ace* gene-containing BAC clone described above.

2.4. Ace1 gene annotation

The assembled *Ae. aegypti* genomic sequence containing the complete *Ace1* gene was compared to gene build AaegL1.1 annotations on VectorBase (www.vectorbase.org). Repeat sequences were identified using VectorBase and CENSOR (Jurka et al., 1996). The repeat-masked as well as the unmasked sequences were annotated for gene content with GENESCAN 1.0 (Burge and Karlin, 1997) using the vertebrate training data set, and GENEID 1.1 (Parra et al., 2000) using the human and *D. melanogaster* training data set. ESTs corresponding to this genomic region were identified on VectorBase.

2.5. Genetic mapping

Genetic data are based on F_1 intercross progeny ($n = 96$) from pairwise matings between the *Ae. aegypti* RED and MOYO-R strains. The RED strain carried multiple morphological markers including *red-eye* (*re*) on chromosome 1, *spot-abdomen* (*s*) on chromosome 2, and *black-tarsus* (*blt*) on chromosome 3. DNA extractions from individual mosquitoes, digestion with Eco*RI*, Southern blotting and hybridizations were performed as previously described (Severson, 1997). Informative marker loci distributed across the genome were identified by pre-screening individual mosquitoes from both parental strains with a panel of markers selected from the existing genetic map (Severson et al., 2002). For mapping the *Ace1* locus, a PCR amplicon of the *Cx. tritaeniorhynchus Ace1* gene was obtained using specific primers (GenBank accession no. AB122152, Forward: 5′-GGAGGA-GAACGTGATCGTGGTT and Reverse: 5′-GATCTT-CCGGCTAAAGTCCTTC), and hybridized as a heterologous probe to Southern blots under low stringency conditions at 65 °C (post-hybridizations washes: 15-min at room temperature followed by 15-min at 65 °C in 2X SSC/0.1% SDS). Multipoint linkage mapping was performed using MAPMAKER/EXP (3.0b) with a minimum LOD threshold of 3.0 (Lander et al., 1987). Map distances were calculated as Kosambi centiMorgans (Kosambi, 1944).

3. Results and discussion

3.1. Ace1 cDNA

The full-length *Ae. aegypti Ace1* cDNA sequence was determined to consist of 2720 bp and contains a 2109 bp open reading frame that encodes the 702 amino acid protein (Fig. 1). The 5′-UTR is 465 bp long, while the 3′UTR is 147 bp and contains two copies of the standard polyadenylation signal (AATAA). The sequence has been deposited in GenBank (accession no. EF209048). The amino acid sequence is 100% identical to that reported by Weill et al. (2004b), (GenBank accession no. AJ621915). The *Ae. aegypti* amino acid sequence is also highly conserved with the Ace1 homolog in other mosquito species, sharing 93.3% identity with *Cx. pipiens*, 93.3% with *Cx. pipiens pallens*, 90.8% with *Cx. tritaeniorhynchus*, and 79.9% with *An. gambiae* (Fig. 2).

To date, three amino acid substitutions have been associated with insensitive AChE-mediated insecticide resistance in mosquitoes. A G119S substitution (using *Torpedo californica* nomenclature; Toutant, 1989) is associated with resistance in *Cx. pipiens*, *An. gambiae*, and *An. albimanus* (Weill et al., 2003, 2004a), but was not evident among samples of *Ae. aegypti* representing 28 populations collected worldwide (Weill et al., 2004b). The second amino acid substitution associated with insensitive AChE-mediated insecticide resistance in mosquitoes, an F455W substitution (*T. californica* nomenclature), was reported for *Cx. tritaeniorhynchus* (Nabeshima et al., 2004). A third amino acid substitution, F290V, associated with insensitive AChE-mediated insecticide resistance, was recently reported for *Cx. pipiens* (Alout et al., 2007). Weill et al. (2004b) suggested that the G119S substitution is less likely to occur in *Ae. aegypti* due to the observed codon usage (GGA), which requires two independent mutation events to generate a serine codon (applies to TCN or AGY). However, the F455W substitution in *Cx. tritaeniorhynchus* (Nabeshima et al., 2004) does reflect two independent mutation events (TTT to TGG). No broad investigation of *Ace1* polymorphism among *Cx. tritaeniorhynchus* populations has been conducted and, therefore, it is not known whether one or both of the necessary T–G mutations are actively segregating and effect resistance through a single-step mutation process. With *Ae. aegypti* Liverpool strain, codon 119 (TTC) also would require two independent mutations to generate a F455W substitution and possible resistance. Thus, the random likelihood of either the G119S or F455W substitutions seems equally probable in *Ae. aegypti* populations. Of note, the F290V substitution with *Ae. aegypti* requires only a single mutation event (TTT to GTT).

3.2. Ace1 genomic sequence

Genomic sequence analysis of the *Ae. aegypti Ace1* gene was determined from a BAC clone by primer walking and

```
              10        20        30        40        50        60        70        80        90        100
               |         |         |         |         |         |         |         |         |         |
      aaatatattctcgacctctgacaccgtgcccgtgtggtgtttcgctggttggcacataagggagtagttcacgctaacgaattgatttcaaagccgaggc   100
      aagagacagaagcaccatctgtcacacaccgactggaaaacgtcgtcactcagagaagaaaagttaacccagaaaatttaagtagataataaacccccc     200
      aactcaactcagtgagttacgtactacagtgagtgaacccgcatatatacaagtggaacctgttattgtctccggtggtgctcctagcaagtggatgtg    300
      tagtggcaattagcgaatttcacacaaaaggagtcggaaaacacagcagccagttgagtgaaccacgcaaccagaagagcggccagcacttgagaaattg   400

      gtgtcatcatcgttaaatgctcctctctacagtcttatagctttatttgtgatttctccaagccATGGAGATCCGAGGCCTAATAACCCGATTGCTAGGT   500
                                                                       M  E  I  R  G  L  I  T  R  L  L  G

      CCATGTCACATCCGACATCTGATACTGTGCAGTTTGGGAATATACTCCATCCTGGTGCAGTCGGTCCATTGCCGGCATCATGACATCGGTAGTTCCACGG   600
      P  C  H  L  R  H  L  I  C  S  L  G  L  Y  S  I  L  V  K  S  V  H  C  R  H  H  D  I  G  S  S  V  A

      CGCATCAGTTAGGCTCGAAGTACTCACAGTCATCCTCGTTATCGTCATCCTCGGCAATCGTCATCATCGGTTAGTCGAAGATCCAGTGCTGCTAACAAAGATTC   700
       H  Q  L  G  S  K  Y  S  Q  S  S  S  L  S  S  S  S  Q  S  S  S  S  L  A  E  E  A  T  L  N  K  D  S

      AGATGCATTTTTTACACCATATATAGGTCACGGACATTCTGTTCGAATTGTAGATGCCGAATTAGGTACCCTAGAGCCGCGAACATGTTCACAGCACTACC   800
      D  A  F  F  T  P  Y  I  G  H  G  D  S  V  R  I  V  D  A  E  L  G  T  L  E  R  E  H  V  H  S  T  T

                                                                                              <------
      ACGCGACGCGCGTGGCTTAACGCGAAGGGAGTCTAGCTCAGATGCTACCGACAATGATCCGTTGCTGATCACCACGGATAAAGGTAAGGTACGCGACTCA   900
       T  R  R  R  G  L  T  R  R  E  S  S  S  D  G  T  D  N  D  P  L  L  I  T  T  D  K  G  K  V  R  G  L  T

      --SRACE------
      CACTCGAAGCTCCAAGTGGAAAGAAAGTAGACGCGATGGCTTGGCATCCCATACGCACAGCCTCCTCTGGGGCCATTGAGATTCCGGCATCCCCGACCAGT   1000
       L  E  A  P  S  G  K  K  V  D  A  W  L  G  I  P  Y  A  Q  P  P  L  G  P  L  R  F  R  H  P  R  P  V

      CGAAAAATGGACGGGAGTGCTGAACGCCACAACGCCACCGAACTCTTGCGTCCAGATTGTGGACACAGTATTTGGTGACTTTCCGGGTGCCACTATGTGG   1100
       E  K  W  T  G  V  L  N  A  T  T  P  P  N  S  C  V  Q  I  V  D  T  V  F  G  D  F  P  G  A  T  M  W

      AACCCGAATACACCACTATCCGAAGACTGTCTATACATCAACGTGGTTGTGCCGCACCCGAGACCAAAGAACTCAGCCGTTATGCTGTGGATATTCGGGG   1200
       N  P  N  T  P  L  S  E  D  C  L  Y  I  N  V  V  V  P  R  P  K  N  A  A  V  M  L  W  I  F  G  G

      GCCGGATTCTACTCTGGTACTGCTACTCTAGACGTGTACGACCATCGCACGCTCGCATCGGAAGAGAACGTTATCGTAGTTTCACTGCAGTATCGAGTGGC   1300
      G  F  Y  S  G  T  A  T  L  D  V  Y  D  H  R  T  L  A  S  E  E  N  V  I  V  V  S  L  Q  Y  R  V  A

      TAGTTTGGGATTTTTTGTTCCTAGGTACTCCGGAAGCTCCGGGAAATGCCGGATTATTCGATCAAAACTTAGCATTGAGATGGGTACGCGACAACATTCAC   1400
       S  L  G  F  L  F  L  G  T  P  E  A  P  G  N  A  G  L  F  D  Q  N  L  A  L  R  W  V  R  D  N  I  H

      AAGTTCGGCGGAGATCCGTCCCGGGTGACCCTGTTTGGTGAAAGTGCAGGTGCTGTCTCGGTGTCACTGCATTTGCTTTCGGCCTTATCCCGGGATCTGT   1500
       K  F  G  G  D  P  S  R  V  T  L  F  G  E  S  A  G  A  V  S  V  S  L  H  L  L  S  A  L  S  R  D  L  F

      TCCAGAGAGCCATTCTGCAAAGCGGATCTCCGACTGCCCCATGGCGCTTTGGTTTCCCGCGAAGAAGCTACGCTCAGAGCTCTTCGCTTAGCCGAAGCCGT   1600
       Q  R  A  I  L  Q  S  G  S  P  T  A  P  W  A  L  V  S  R  E  E  A  T  L  R  A  L  R  L  A  E  A  V

      CAATTGTCCGCACGATGCCGACAAAGCTCACTGACACCGTAGAGTGTCTCCGCACTAAAGATCCAAACGTGCTGGTCGATAACGAATGGGGAACGCTAGGA   1700
       N  C  P  H  D  A  T  K  L  T  D  T  V  E  C  L  R  T  K  D  P  N  V  L  V  D  N  E  W  G  T  L  G

      ATCTGCGAGTTTCCATTTGTACCCGTCGTTGACGGTGCATTCCTCGACGAAACACCCCAACGTTCGCTAGCCAGTGGTAGGTTTAAGAAGACGGACATCC   1800
       I  C  E  F  P  F  V  P  V  V  D  G  A  F  L  D  E  T  P  Q  R  S  L  A  S  G  R  F  K  K  T  D  I  L

      TAACCGGCAGTAATACGGAGGAAGGTTATTACTTCGATAATATACTACTTGACTGAACTATTGCGGAAAGAGGAGGGTGTCACAGTTTCACGGGAGGAGTT   1900
       T  G  S  N  T  E  E  G  Y  Y  F  D  N  I  Y  Y  L  T  E  L  L  R  K  E  E  G  V  T  V  S  R  E  E  F

      CTTGCAGGCCGTTAGAGAACTGAATCCTTACGTGAACGGAGCCGCGAGGCAGGCTATCGTGTTCGAGTACACCGACTGGACAGAACCGGAAAATCCCAAC   2000
       L  Q  A  V  R  E  L  N  P  Y  V  N  G  A  A  R  Q  A  I  V  F  E  Y  T  D  W  T  E  P  E  N  P  N

      AGCAATCGGGATGCATTGGACAAAATGGTCGGAGATTATCACTTCACGTGTAATGTGAATGAGTTTGCCCAGCGATATGCAGAAGAAGGCAACAATGTGT   2100
       S  N  R  D  A  L  D  K  M  V  G  D  Y  H  F  T  C  N  V  N  E  F  A  Q  R  Y  A  E  E  G  N  N  V  Y

                                                                                          ------
      ACATGTATCTGTACACTCATAGAAGCAAAGGTAACCCCTGGCCACGGTGGACCGGTGTGATGCATGGTGACGAGATCAATTATGTGTTCGGTGAGCCTCT   2200
       M  Y  L  Y  T  H  R  S  K  G  N  P  W  P  R  W  T  G  V  M  H  G  D  E  I  N  Y  V  F  G  E  P  L

      --3RACE-->
      GAACTCTGATCTGGGGTACATGGAGGATGAAAAAGACTTCAGTAGGAAGATTATGAGATACTGGTCTAACTTTGCTAAAACTGGCAATCCAAATCCAAGT   2300
       N  S  D  L  G  Y  M  E  D  E  K  D  F  S  R  K  I  M  R  Y  W  S  N  F  A  K  T  G  N  P  N  P  S

      CCTCCAAATAGTGACTTCCCAGAATGGCCGAAACATACAGCGCATGGTAGACATTATCTAGAACTTGGGCTCAACACAACCTATGTGGGCCGTGGCCCAC   2400
       P  P  N  S  D  F  P  E  W  P  K  H  T  A  H  G  R  H  Y  L  E  L  G  L  N  T  T  Y  V  G  R  G  P  R

      GATTGCGGCAATGTGCGTTCTGGAAAAAATACCTGCCGCAACTAGTAGCAGCTACCTCTAATCTACAAGCAACTCCTGCTCCAAGCGAGCCATGCGGAAG   2500
       L  R  Q  C  A  F  W  K  K  Y  L  P  Q  L  V  A  A  T  S  N  L  Q  A  T  P  A  P  S  E  P  C  G  S

      CAGTGCAACGCCCTATCGACCCCTTCTATTTTTGATAGTCTCGCTAGTTCTAGTAACATGTTTCAAAATATAAaattttttatgtttgttaggctttatac   2600
       S  A  T  P  Y  R  P  L  L  F  L  I  V  S  L  V  L  V  T  C  F  K  I  *

      aatggtcagtttgaaaatagctgaatgatggttgcaaaataacatcaatactactataaataccaatactattaattastaaaaaatgaaaatggaacaa   2700
      taaagctagattttattaaaaaaaaaaaaaaaaaaaaa                                                                2740
```
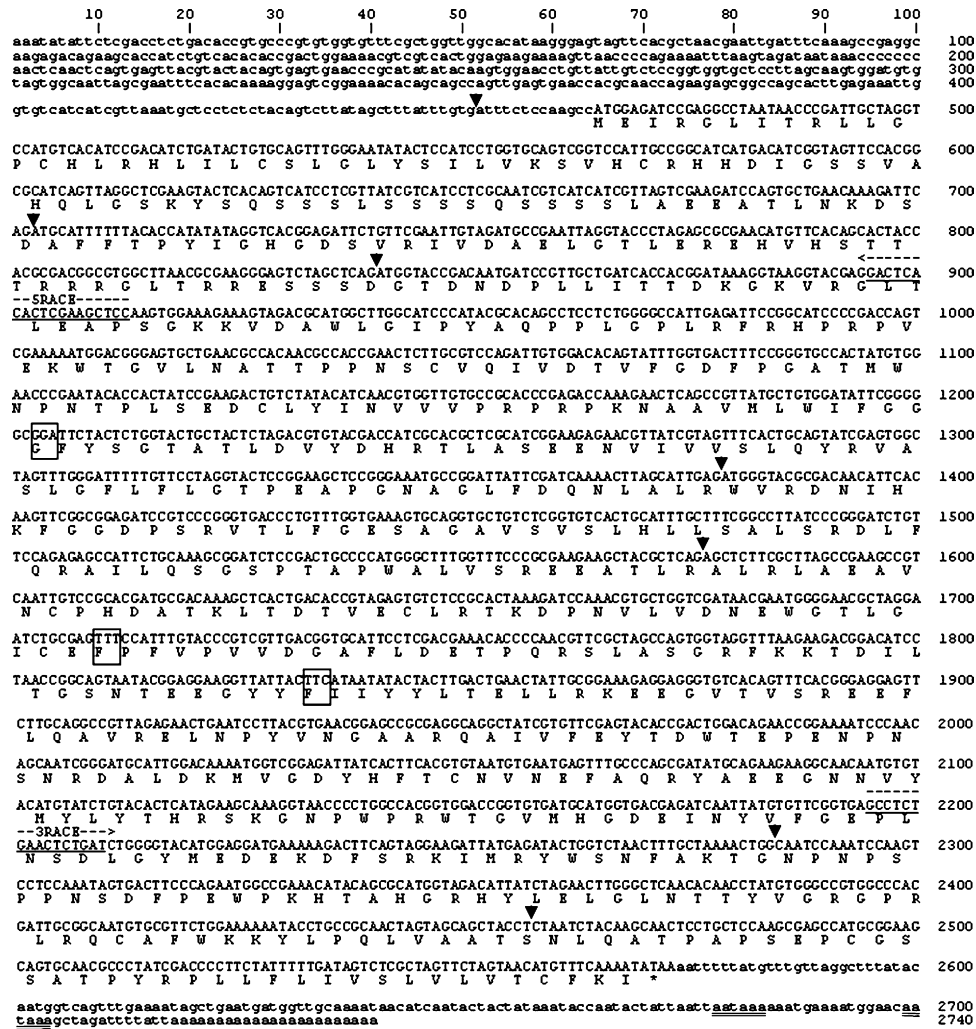
Fig. 1. Sequence and deduced amino acid translation products of the *Ae. aegypti* full-length *Ace1* cDNA. The deduced amino acid sequence is shown under the nucleotide sequence in a single letter code; the poly(A) signal site is underlined; primer sites for 5′ and 3′ RACE analysis are noted above the sequence; triangles indicate intron positions. Positions for amino acid substitutions associated with insensitive AChE-mediated insecticide resistance are indicated in boxes. The previously described partial *Ace1* cDNAs encompass nucleotides 125,763–126,051 (Genbank accession no. AJ428049) and 114,816–138,413 (Genbank accession no. AJ621915) in the compete genome sequence (Genbank accession no. BK006052). The full-length cDNA sequence has been submitted to Genbank under accession no. EF209048.

genome sequence data that became available at The Broad Institute via the *Ae. aegypti* whole genome sequencing project. We determined that the *Ace1* gene is contained within a 138,970 bp genome region, and consists of eight exons, with seven introns that vary markedly in length from 59 to 114,350 bp (Fig. 3). The exon–intron splice junctions all have the conserved GT-AG motifs (Table 1). The sequence has been deposited in GenBank (accession no. BK006052). The first intron sequence as represented in the *Ae. aegypti* genome assembly (GenBank accession no. CH477194) contained two gaps. We determined that Gap1, which spanned contigs AAGE02000752 and AAGE020 00753, was an artifact of the whole genome shotgun sequencing assembly process that reflected misassembly of repetitive sequence that linked the two contigs. Correct sequence orientation was determined only following sequence analysis of a BAC subclone containing the

putative Gap1 region. For Gap2, which spanned contigs AAGE02000753 and AAGE02000754, we were able to design unique primers spanning the gap, and direct PCR product sequencing determined that the gap sequence consisted of only 19 bp. Here too, the gap in sequence assembly was likely a consequence of the complex repetitive nature of the genome in that region. Of note regarding gap closure, our efforts were greatly hampered by the difficulty in identifying single copy genome sequence in the regions, which made even identification of suitable PCR primer sequences problematic. This is a likely characteristic of most gaps in the *Ae. aegypti* genome assembly, as ~68% of the sequence consists of known transposable elements and other interspersed repeat sequences (Nene et al., 2007).

Comparison of the *Ae. aegypti* genome sequence to that for the *An. gambiae Ace1* homolog revealed considerable

```
Aaeg   ------MEIRGLITRLLGPCHIRHLILCSLGIYSILVQS-----VHCRHHDIG-----------SSTAHQLGSKYSQSSSLSS-------   61
Aalb   ------MEIRGLITRLLGPCHLRHLILCSLGIYSILVKS-----VHCRHHDIG-----------SSTAHQLGSKYAQSSSLSS-------   61
Cppi   ------MEIRGLITRLLGPCHLRHLILCSLGLYSILVKS-----VHCRHHDIG-----------SSVAHQLGSKYSQSSSLSS-------   61
Cppa   ------MEIRGLITRLLGPCHLRHLILCSLGLYSILVQS-----VHCRHHDIG-----------SSVAHQLGSKYSQSSSLSS-------   61
Ctri   MRPLQAMEIRGLITRLLGPCHLRHLILCSLGLYSILVLVQS---VHCRHHDIG-----------SSAAHQLGSKYSQSSSLSS-------   67
Agam   RSPPVPMEIRGLLMGRLR-LGRRMVPLGLLGVTALLLILPPFALVQGRHHELNNGAAIGSHQLSAAAGVGLASQSAQSGSLASGVMSSVP   89
              ******:    *   * : *  **: ::*:         *: ***::.        ::.. *.*: :**.**:*

Aaeg   ----SSQSSSSLVEDP--------VLNKDSDAFFTPYIGHGDSVRIVDAELGTLEREHVHS-TTTRRRGLTRRESSSSDGTDNDPLLITTD  138
Aalb   ----SSQSSSSLVEDP--------ALNKDSDAFFTPYIGHGDSVRIVDAELGTLEREHVHS-TTTRRRGLTRRESSSSDANDNDPLLITTD  138
Cppi   ----SSQSSSSLAEEA--------TLNKDSDAFFTPYIGHGDSVRIVDAELGTLEREHIHS-TTTRRRGLTRRESSSSDATDSDPLVITTD  138
Cppa   ----SSQSSSSLAEEA--------TLNKDS-------GHGDSVRIVDAELGTLEREHIHS-TTTRRRGLTRRESSSSDATDSDPLVITTD  130
Ctri   ----SSQSSSSLAEEA--------TLNKDS--------GHGDSVRIVDAELGTLEREHIHS-TTTRRRGLTRRESNSDSTDSDPLVITTD  136
Agam   AAGASSSSSSSLLSSSAEDDVARITLSKDADAFFTPYIGHGESVRIDAELGTLE--HVHSGATPRRRGLTRRESNSDANDNDPLVVNTD  177
              **.*****  ...       .*.**:       ***:****.********  *:** :*.*********.**..*.***::.**

Aaeg   KGKVRGLTLEAPSGKKVDAWLGIPYAQPPLGPLRFRHPRPVEKWTGVLNATTPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVV  228
Aalb   KGKVRGLTLEAPSGKKVDAWLGIPYAQPPLGPLRFRHPRPVEKWTGVLNATTPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVV  228
Cppi   KGKIRGTTLEAPSGKKVDAWMGIPYAQPPLGPLRFRHPRPAEKWTGVLNATKPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVV  228
Cppa   KGKIRGTTLEAPSGKKVDAWMGIPYAQPPLGPLRFRHPRPAERWSGVLNATKPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVV  220
Ctri   KGKIRGTTLEAPSGKKVDAWMGIPYAQPPLGPLRFRHPRPAEKWTGVLNATKPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVV  226
Agam   KGRIRGITVDAPSGKKVDVWLGIPYAQPPVGPLRFRHPRPAEKWTGVLNTTTPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVV  267
       **::**  *::*******.*:.********:*********.*:*:****:*.**********.*******************:*********

Aaeg   VPHPRPKNSAVMLWIFGGGFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQNLALRWVRDNIHKFGGDP  318
Aalb   VPHPRPKNSAVMLWIFGGGFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQNLALRWVRDNIHKFGGDP  318
Cppi   VPRPRPKNAAVMLWIFGGGFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQNLALRWVRDNIHRFGGDP  318
Cppa   VPRPRPKNAAVMLWIFGGGFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQNLALRWVRDNIHRFGGDP  310
Ctri   APRPRPKNAAVMLWIFGGGFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQNLALRWVRDNIHRFGGDP  316
Agam   APRPRPKNAAVMLWIFGGGFYSGTATLDVYDHRALASEENVIVVSLQYRVASLGFLFLGTPEAPGNAGLFDQNLALRWVRDNIHRFGGDP  357
        .*:*****:*************************:*****************************************************

Aaeg   SRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTAPWALVSREEATLRALRLAEAVNCPHDATKLTDTVECLRTKDPNVLVDNEW  408
Aalb   SRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTAPWALVSREEATLRALRLAEAVNCPHDATKLSDAVECLRTKDPNELVDNEW  408
Cppi   SRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTAPWALVSREEATLRALRLAEAVNCPHDATKLSDAVECLRTKDPNELVDNEW  408
Cppa   SRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTAPWALVSREEATLRALRLAEAVNCPHDATKLSDAVECLRTKDPNELVDNEW  400
Ctri   SRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTAPWALVSREEATLRALRLAEAVNCPHDATKLSRAVECLRTKDPNELVDNEW  406
Agam   SRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTAPWALVSREEATLRALRLAEAVGCPHEPSKLSDAVECLRGKDPHVLVNNEW  447
       *******************************************************.***..**::.***** ***. **:***

Aaeg   GTLGICEFPPVPVVDGAFLDETPQRSLASGRFKKTDILTGSNTEEGYYFIIYYLTELLRKEEGVTVSREEFLQAVRELNPYVNGAARQAI  498
Aalb   GTLGICEFPPVPVVDGAFLDETPQRSLASGRFKKTDILTGSNTEEGYYFIIYYLTELLRKEEGVTVSREEFLQAVRELNPYVNGAARQAI  498
Cppi   GTLGICEFPPVPVVDGAFLDETPQRSLASGRFKKTDILTGSNTEEGYYFIIYYLTELLRKEEGVTVTREEFLQAVRELNPYVNGAARQAI  498
Cppa   GTLGICEFPPVPVVDGAFLDETPQRSLASGRFKKTDILTGSNTEEGYYFIIYYLTELLRKEEGVTVSREEFLQAVRELNPYVNGAARQAI  490
Ctri   GTLGICEFPPVPVVDGAFLDETPQRSLASGRFKKTEILTGSNTEEGYYFIIYYLTELLRKEEGVTVSREEFLQAVRELNPYVNGAARQAI  496
Agam   GTLGICEFPPVPVVDGAFLDETPQRSLASGRFKKTEILTGSNTEEGYYFIIYYLTELLRKEEGVTVTREEFLQAVRELNPYVNGAARQAI  537
       ***********************************:*:*************************:***:***********************

Aaeg   VFEYTDWTEPENPNSNRDALDKMVGDYHFTCNVNEFAQRYAEEGNNVYMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNSDLGYMED  588
Aalb   VFEYTDWTEPENPNSNRDALDKMVGDYHFTCNVNEFAQRYAEEGNNVYMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNSDLGYMED  588
Cppi   VFEYTDWIEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAEEGNNVFMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNSALGYQDD  588
Cppa   VFEYTDWIEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAEEGNNVFMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNSALGYQDD  580
Ctri   VFEYTDWIEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAEEGNNVFMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNAALGYQED  586
Agam   VFEYTDWTEPDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAEEGNNVYMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNPTLGYTED  627
       ******* **:*********************************************:****************************. *** :*

Aaeg   EKDFSRKIMRYWSNFAKTGNPNPSPPNSDFPEWPKHTAHGRHYLELGLNTTYVGRGPRLRQCAFWKKYLPQLVAATSNLQATPAPSEPCG  678
Aalb   EKDFSRKIMRYWSNFAKTGNPNPSPPNSDFTEWPKHTAHGRHYLELGLNTTYVGRGPRLRQCAFWKKYLPQLVAATSNLQATPAPSEPCG  678
Cppi   EKDFSRKIMRYWSNFAKTGNPNPSTPSVDLPEWPKHTAHGRHYLELGLNTTFVGRGPRLRQCAFWKKYLPQLVAATSNLQVTPAPSVPCE  678
Cppa   EKDFSRKIMRYWSNFAKTGNPNPSTPSVDLPEWPKHTAHGRHYLELGLNTTFVGRGPRLRQCAFWKKYLPQLVAATSNLQVTPAPSVPCE  670
Ctri   EKDFSRKIMRYWSNFAKTGNPNPSPPNADLPEWPKHTAHGRHYLELGLNTSFVGRGPRLRQCAFWKKYLPQLVAATSNLQVTPAPSVPCE  676
Agam   EKDFSRKIMRYWSNFAKTGNPNPNTASSEFPEWPKHTAHGRHYLELGLNTSFVGRGPRLRQCAFWKKYLPQLVAATSNLPGPAPPSEPCE  717
       ***********************....  ::.*******************:.********************************  ...** **

Aaeg   SSATPYR-PLLFLIVSLVLVTCFKI-  702
Aalb   SSAALYR-PLLFLIVSLVLVTRFKI-  702
Cppi   SSSTSYR-STLLLIVTLLLVTRFKI-  702
Cppa   SSSTSYR-STLLLIVTLLLVTRFKI-  694
Ctri   SSSTSYRSSTLLLIVTLLLVIRFKI-  701
Agam   SSAFFYRPDLIVLLVSLLTATVRFIQ  743
       **:  **  :.*:*:*:.    *
```

Fig. 2. Alignment of mosquito AChE sequences. The deduced amino acid sequence for *Ae. aegypti* AChE1 is aligned with that from *Cx. pipiens* (Cpi), *Cx. pipiens pallens* (Cpp), *An. gambiae* (Ag), and AChE2 from *Cx. tritaeniorhynchus* (Cxt) (GenBank accession nos. AJ489456, AY762905, AJ488492 and AB122152, respectively). Identical amino acids are indicated by asterisks and conservative substitutions by dots. Arrows indicate the G119S, F290V and F455W positions previously associated with insensitive AChE-mediated resistance in mosquitoes.
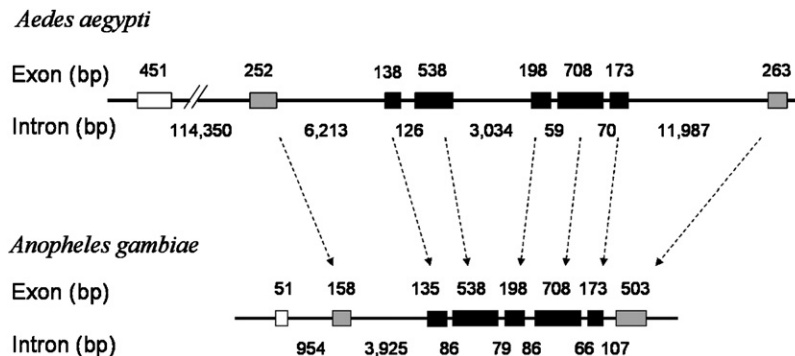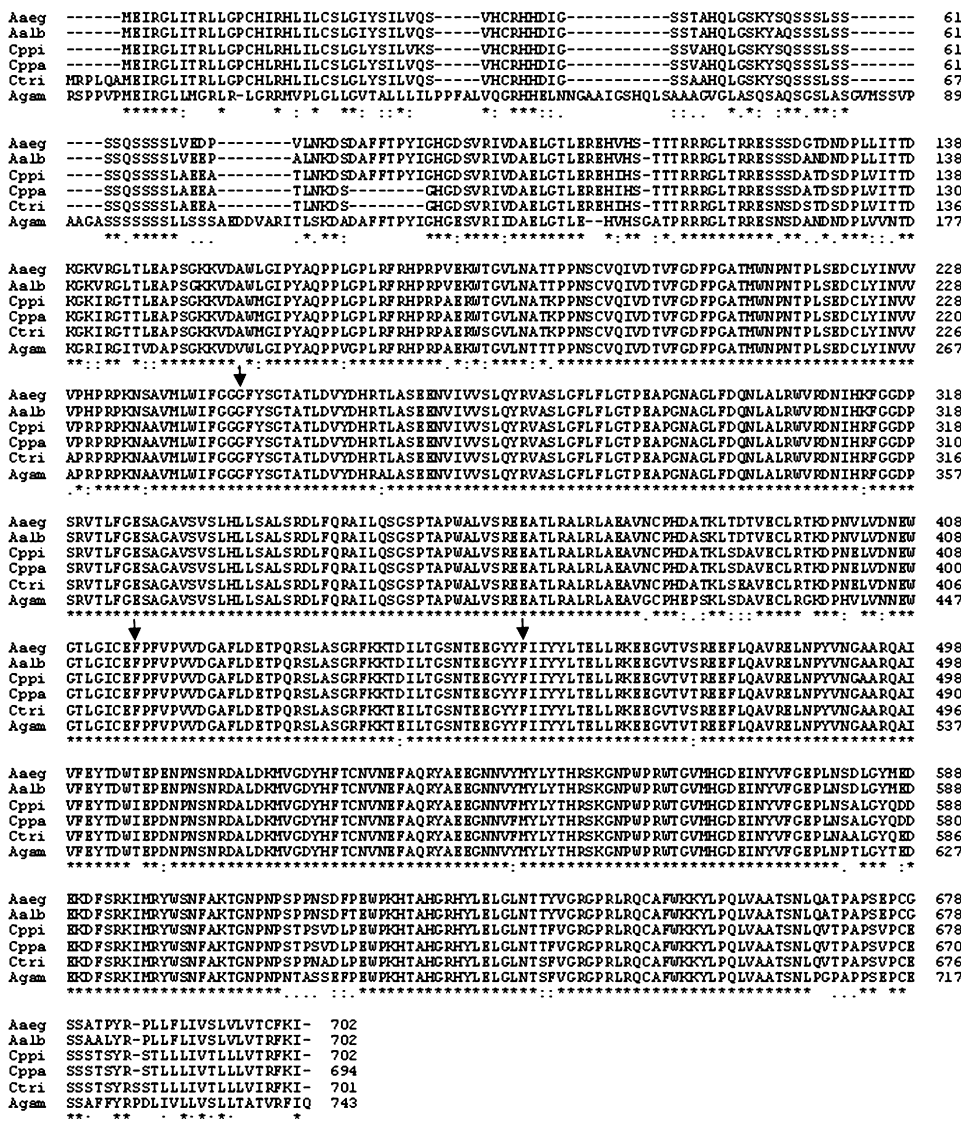

Fig. 3. Comparative genomic structure of the *Ae. aegypti* and *An. gambiae* *Ace1* genes. Exon 1 is noncoding in both species. Exon 2 contains the start codon (ATG) and exhibits partial amino acid conservation between species. Exons 3–7 are highly conserved between species. Exon 8 contains the stop codon and exhibits partial amino acid conservation between species.

Table 1
Exon–intron junctions of the *Ace1* gene

No.	Exon size (bp)	Splice-donor	Intron size (bp)	Splice-acceptor
1		−14		−13
	451	TTGTG**gt**gagtaatg	114898	tcagttgc**ag**A̲TTTCT
2		238		239
	251	GATTCAG̲**gt**aattacca	6213	aatttttc**ag**A̲TGCATTT
		asp ser asp		ala phe
3		376		377
	138	AGCTCAG̲**gt**gagatcgg	126	aacatttc**ag**A̲TGGTACC
		ser ser asp		gly thr
4		914		915
	538	GCATTGAG̲**gt**gagtgaaa	3034	ataattct**ag**A̲TGGGTA
		ala leu arg		trp val
5		1112		1113
	198	ACGCTCAG̲**gt**aagaacta	59	tattccac**ag**A̲GCTCTT
		thr leu arg		ala leu
6		1820		1821
	708	AAAACTGG̲**gt**aagtttat	70	gtctttat**ag**C̲AATCCA
		lys thr gly		asn pro
7		1993		1994
	173	GCTACCT**gt**aagttttc	11987	ttcttttc**ag**C̲TAATCTA
		ala thr ser		asn leu
8		2109		
	116	AAAATA*TAA*aatttttat		
		lys ile		

Coding sequences are shown in uppercase letters and noncoding regions in lower case letters. The conserved GT/AG exon/intron junctions are shown in bold. Amino acid codons split by exon/intron junctions are underlined. Amino acid residues are indicated with respect to each boundary and the stop codon (TAA) is indicated by italics. Numbers refer to the corresponding positions in the *Ace1* cDNA starting with +1 at the adenosine defining the initiation codon ATG.

conservation among most exons, but extensive divergence in intron structure that reflects overall genome contrasts between the two species (Fig. 3). At the amino acid level, exons 3–7 are largely conserved in size and content, while exons 2 and 8 share only 34.2% and 51.3% sequence identity, respectively. The *Ace1* gene in *An. gambiae* is contained within a 7767 bp genome region; therefore, the *Ae. aegypti* genome sequence is ca. 17.8-fold larger. This difference is due to the presence of three large introns in the *Ae. aegypti* gene, ranging from ca. 3–115 kb, in contrast to the fairly consistent intron size of ca. 100 bp in the *An. gambiae* gene. This is undoubtedly reflective of the overall differences in genome organization of the two species, as the *An. gambiae* genome size is ca. 243 Mbp, while *Ae. aegypti* is about 5-fold larger at ca. 1.31 Gbp. Most of this difference is due to the high frequency of repetitive sequences in the *Ae. aegypti* genome (Nene et al., 2007).

3.3. Ace1 annotation

Annotation of the *Ae. aegypti Ace1* 138,970 bp genome sequence (Fig. 4) provided insight into the basis for the increased gene size compared to *An. gambiae* and reflects problems affecting annotation accuracy in the *Ae. aegypti* genome sequencing project (Nene et al., 2007). The sequence was compared to the *Ae. aegypti* genome assembly *in silico* gene build (AaegL1.1) at VectorBase to see if the entire *Ace1* gene was identified and annotated

correctly, as well as if there were any other genes or known repetitive elements present within the introns. The VectorBase gene Q6A2E2_AEDAE (AAEL000511) represented most of the *Ace1* gene with the exception of the 5′ exon, which was not identified. In addition, two *ab initio* gene prediction programs that we tested (GENESCAN 1.0 and GENEID 1.1) did not correctly identify the entire *Ace1* gene (data not shown). Although both correctly identified most of the *Ace1* exons, they were incorrectly assembled as representing multiple genes, and the 3′ exon was missed by both programs. These observations are likely a consequence of the large introns and the unusually large number of interspersed repetitive element ORFs present in the introns. Repetitive elements identified with CENSOR (Jurka et al., 1996) coincided with those seen on VectorBase. All these were present within introns and the most abundant of these was *Feilai* (Tu, 1999). Finally, no additional genes were identified within the *Ace1* genomic sequence.

3.4. Ace1 genome location

The *Ae. aegypti Ace1* locus was mapped to a region on the *p*-arm of chromosome 3 between the LF106 and LF386 loci (Fig. 5). In comparison, the *Ace2* locus is located on chromosome 1 and is tightly linked with the sex determination locus (Severson et al., 1997). The *Ace1* homolog in *Cx. tritaeniorhynchus* (listed as *Ace2*;
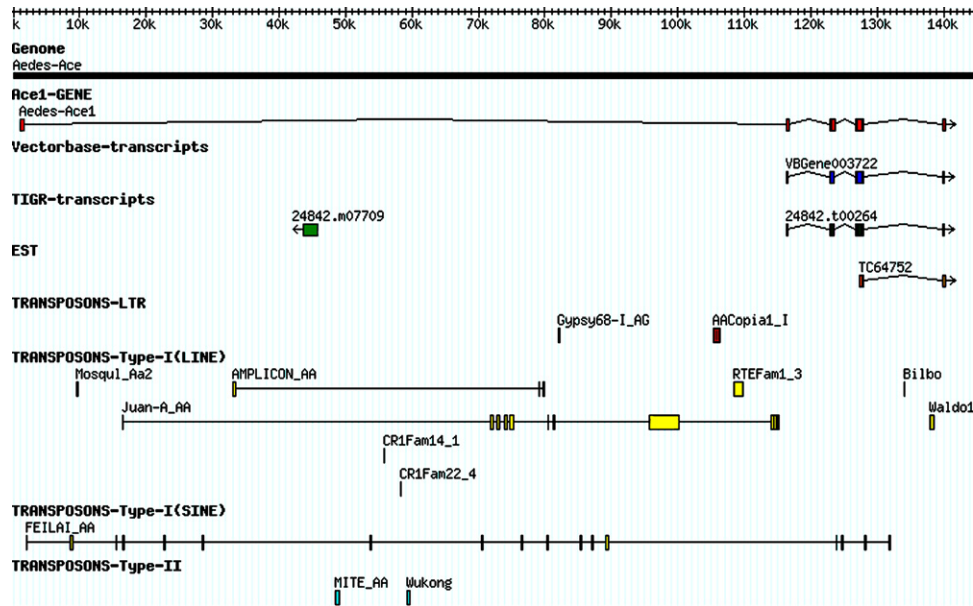
Fig. 4. *In silico* annotation of the *Ae. aegypti Ace1* genome sequence. Comparison of the Aedes *Ace1* genomic region to the genome assembly and gene build (AAegL1.1) indicates that the 5′ exon was missed. Two ESTs correspond to the 3′ end of the gene. Introns contain various transposons.
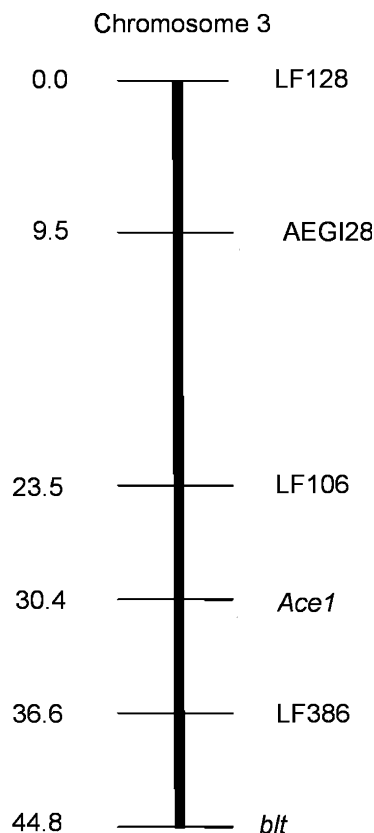

Fig. 5. Genetic map position of the *Ae. aegypti Ace1* locus. Map distances are in Kosambi centiMorgans (Kosambi, 1944).

the LF106 and LF386 loci when compared to *Ae. aegypti* (Mori et al., 1999, 2001). The *Ace1* homolog in *Cx. tritaeniorhynchus* is linked with the chromosome arm containing the LF106 locus. Similarly, the *Ace1* homolog in *An. gambiae* is located on chromosome 2R; this chromosome arm carries the genome region orthologous to the *Ae. aegypti* p-arm of chromosome 3 (Severson et al., 2004). As previously suggested, the multiple *Ace* genes identified in some insects likely arose due to gene duplication (Gao et al., 2002). Our results indicate that this event occurred prior to the evolution of the mosquito subfamily (Culicinae and Anophelinae) lineages that split about 150 million years ago (Krzywinski et al., 2006).

Acknowledgments

References

Alout, H., Berthomieu, A., Hadjivassilis, A., Weill, M., 2007. A new amino-acid substitution in acetylcholinesterase 1 confers insecticide resistance to *Culex pipiens* mosquitoes from Cyprus. Insect Biochem. Mol. Biol. 37, 41–47.

Bisset, J., Rodríguez, M.M., Fernández, D., 2006. Selection of insensitive acetylcholinesterase as a resistance mechanism in *Aedes aegypti* (Diptera: Culicidae) from Santiago de Cuba. J. Med. Entomol. 43, 1185–1189.

Burge, C., Karlin, S., 1997. Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. 268, 78–94.

Chen, Z., Newcomb, R., Forbes, E., McKenzie, J., Batterham, P., 2001. The acetylcholinesterase gene and organophosphorus resistance in the

Nabeshima et al., 2004) and likely also in *Cx. pipiens* (Malcolm et al., 1998) is located in the orthologous genome region, although the two *Culex* spp. reflect whole chromosome arm translocations that involve a breakpoint between

Australian sheep blowfly, *Lucilia cuprina*. Insect Biochem. Mol. Biol. 31, 805–816.

Fournier, D., Mutero, A., 1994. Modification of acetylcholinesterase as a mechanism of resistance to insecticides. Comp. Biochem. Physiol. 108C, 19–31.

Gao, J.-R., Kambhampati, S., Zhu, K.Y., 2002. Molecular cloning and characterization of a greenbug (*Schizaphis graminum*) cDNA encoding acetylcholinesterase possibly evolved from a duplicate gene lineage. Insect Biochem. Mol. Biol. 32, 765–775.

Hall, L.M., Spierer, P., 1986. The *Ace* locus of *Drosophila melanogaster*: structural gene for acetylcholinesterase with an unusual 5′ leader. EMBO J 5, 2949–2954.

Huang, X., Madan, A., 1999. CAP3: a DNA sequence assembly program. Genome Res 9, 868–877.

Jiménez, L.V., Kang, B.K., deBruyn, B., Lovin, D.D., Severson, D.W., 2004. Characterization of an *Aedes aegypti* BAC library and chromosomal assignment of BAC clones for physical mapping QTL that influence *Plasmodium* susceptibility. Insect Mol. Biol. 13, 37–44.

Jurka, J., Klonowski, P., Dagman, V., Pelton, P., 1996. CENSOR—a program for identification and elimination of repetitive elements from DNA sequences. Comput. Chem. 20, 119–121.

Kosambi, D.D., 1944. The estimation of map distance from recombination values. Ann. Eugen. 12, 172–175.

Kozaki, T., Shono, T., Tomita, T., Kono, Y., 2001. Fenitroxon insensitive acetylcholinesterases of the housefly, *Musca domestica* associated with point mutations. Insect Biochem. Mol. Biol. 31, 991–997.

Krzywinski, J., Grushko, O.G., Besansky, N.J., 2006. Analysis of the complete mitochondrial DNA from *Anopheles funestus*: an improved dipteran mitochondrial genome annotation and a temporal dimension of mosquito evolution. Mol. Phylogenet. Evol. 39, 417–423.

Lander, E.S., Green, P., Abrahamson, J., Barlow, A., Daly, M., Lincoln, S., Newberg, L., 1987. MAPMAKER: an interactive computer package for constructing primary genetic linkage maps of experimental and natural populations. Genomics 1, 174–181.

Malcolm, C.A., Bourguet, D., Ascolillo, A., Rooker, S.J., Garvey, C.F., Hall, L.M., Pasteur, N., Raymond, M., 1998. A sex-linked *Ace* gene, not linked to insensitive acetylcholinesterase-mediated insecticide resistance in *Culex pipiens*. Insect Mol. Biol. 7, 107–120.

Mori, A., Severson, D.W., Christensen, B.M., 1999. Comparative linkage maps for the mosquitoes (*Culex pipiens* and *Aedes aegypti*) based on common RFLP loci. J. Hered. 90, 160–164.

Mori, A., Tomita, T., Hidoh, O., Kono, Y., Severson, D.W., 2001. Comparative linkage map development and identification of an autosomal locus for insensitive acetylcholinesterase-mediated insecticide resistance in *Culex tritaeniorhynchus*. Insect Mol. Biol. 10, 197–203.

Mutero, A., Pralavorio, M., Bride, J.M., Fournier, D., 1994. Resistance-associated point mutations in insecticide-insensitive acetylcholinesterase. Proc. Natl. Acad. Sci. USA 9, 5922–5926.

Nabeshima, T., Mori, A., Kozaki, T., Iwata, Y., Hidoh, O., Harada, S., Kasai, S., Severson, D.W., Kono, Y., Tomita, T., 2004. An amino acid substitution attributable to insecticide-insensitivity of acetylcholinesterase in a Japanese encephalitis vector mosquito, *Culex tritaeniorhynchus*. Biochem. Biophys. Res. Commun. 313, 794–801.

Nene, V., Wortman, J.R., Lawson, D., Haas, B., Kodira, C., Tu, Z., Loftus, B., Xi, Z., Megy, K., et al., 2007. Genome sequence of *Aedes aegypti*, a major arbovirus vector. Science, submitted for publication.

Parra, G., Blanco, E., Guigo, R., 2000. GeneID in Drosophila. Genome Res 10, 511–515.

Severson, D.W., 1997. RFLP analysis of insect genomes. In: Crampton, J.M., Bread, C.B., Louis, C. (Eds.), The Molecular Biology of Insect Disease Vectors: a Methods Manual. Chapman & Hall, London, pp. 309–320.

Severson, D.W., Anthony, N.M., Andreev, O., ffrench-Constant, R.H., 1997. Molecular mapping of insecticide resistance genes in the yellow fever mosquito *Aedes aegypti*. J. Hered. 88, 520–524.

Severson, D.W., Meece, J.K., Lovin, D.D., Saha, G., Morlais, I., 2002. Linkage map organization of expressed sequence tags and sequence tagged sites in the mosquito, *Aedes aegypti*. Insect Mol. Biol. 11, 371–378.

Severson, D.W., deBruyn, B., Lovin, D.D., Brown, S.E., Knudson, D.L., Morlais, I., 2004. Comparative genome analysis of the yellow fever mosquito *Aedes aegypti* with *Drosophila melanogaster* and the malaria vector mosquito *Anopheles gambiae*. J. Hered. 95, 103–113.

Toutant, J.P., 1989. Insect acetylcholinesterase: catalytic properties, tissue distribution and molecular forms. Prog. Neurobiol. 32, 423–446.

Tu, Z., 1999. Genomic and evolutionary analysis of *Feilai*, a diverse family of highly reiterated SINEs in the yellow fever mosquito, *Aedes aegypti*. Mol. Biol. Evol. 16, 760–772.

Walsh, S.B., Dolden, T.A., Moores, G.D., Kristensen, M., Lewis, T., Devonshire, A.L., Williamson, M.S., 2001. Identification and characterization of mutations in housefly (*Musca domestica*) acetylcholinesterase involved in insecticide resistance. Biochem. J. 359, 175–181.

Weill, M., Fort, P., Berthomieu, A., Dubois, M.P., Pasteur, N., Raymond, M., 2002. A novel acetylcholinesterase gene in mosquitoes codes for the insecticide target and is non-homologous to the *ace* gene in *Drosophila*. Proc. R. Soc. London B 269, 2007–2016.

Weill, M., Lutfalla, G., Mogensen, K., Chandre, F., Berthomieu, A., Berticat, C., Pasteur, N., Philips, A., Fort, P., Raymond, M., 2003. Insecticide resistance in mosquito vectors. Nature 423, 136–137.

Weill, M., Malcolm, C., Chandre, F., Mogensen, K., Berthomieu, A., Marquine, M., Raymond, M., 2004a. The unique mutation in *ace-1* giving high insecticide resistance is easily detectable in mosquito vectors. Insect Mol. Biol. 13, 1–7.

Weill, M., Berthomieu, A., Berticat, C., Lutfalla, G., Negre, V., Pasteur, N., Philips, A., Leonetti, J.-P., Fort, P., Raymond, M., 2004b. Insecticide resistance: a silent base prediction. Curr. Biol. 14, R552–R553.