Genome Organization, Phylogenies, Expression Patterns, and Three-Dimensional Protein Models of Two Acetylcholinesterase Genes from the Red Flour Beetle

Yanhui Lu[1,3], Yuan-Ping Pang[2]*, Yoonseong Park[3], Xiwu Gao[1], Jianxiu Yao[3], Xin Zhang[3], Kun Yan Zhu[3]*

1 Department of Entomology, China Agricultural University, Beijing, People's Republic of China, 2 Computer-Aided Molecular Design Laboratory, Mayo Clinic, Rochester, Minnesota, United States of America, 3 Department of Entomology, Kansas State University, Manhattan, Kansas, United States of America

Abstract

Since the report of a paralogous acetylcholinesterase (AChE, EC3.1.1.7) gene in the greenbug (*Schizaphis graminum*) in 2002, two different AChE genes (*Ace1* and *Ace2*) have been identified in each of at least 27 insect species. However, the gene models of *Ace1* and *Ace2*, and their molecular properties have not yet been comprehensively analyzed in any insect species. In this study, we sequenced the full-length cDNAs, computationally predicted the corresponding three-dimensional protein models, and profiled developmental stage and tissue-specific expression patterns of two *Ace* genes from the red flour beetle (*Tribolium castaneum*; *TcAce1* and *TcAce2*), a globally distributed major pest of stored grain products and an emerging model organism. *TcAce1* and *TcAce2* encode 648 and 604 amino acid residues, respectively, and have conserved motifs including a choline-binding site, a catalytic triad, and an acyl pocket. Phylogenetic analysis show that both *TcAce* genes are grouped into two insect *Ace* clusters and *TcAce1* is completely diverged from *TcAce2*, suggesting that these two genes evolve from their corresponding *Ace* gene lineages in insect species. In addition, *TcAce1* is located on chromosome 5, whereas *TcAce2* is located on chromosome 2. Reverse transcription polymerase chain reaction (PCR) and quantitative real-time PCR analyses indicate that both genes are virtually transcribed in all the developmental stages and predominately expressed in the insect brain. Our computational analyses suggest that the TcAce1 protein is a robust acetylcholine (ACh) hydrolase and has susceptibility to sulfhydryl agents whereas the TcAce2 protein is not a catalytically efficient ACh hydrolase.

Introduction

Acetylcholinesterase (AChE, EC3.1.1.7) is an essential enzyme at the synapses of cholinergic neurons in the central and peripheral nervous systems in all animals. It catalyzes the hydrolysis of the neurotransmitter acetylcholine (ACh), thus terminating neuro-transmission. AChE has long been of academic and industrial interest and studied extensively at the biochemical, biophysical, and molecular levels in mammals because this enzyme is a target of palliative Alzheimer drugs, nerve agents, and insecticides [1]. In insects, AChE has also been extensively studied because it serves as the target site for organophosphate and carbamate insecticides, and involves in insecticide resistance known as target-site insensitivity [2–8].

The first insect AChE gene (*Ace*) was sequenced from *Drosophila melanogaster* in 1986 [9]. After the first *Ace* paralogous gene was reported in the greenbug (*Schizaphis graminum*) in 2002 [10], the *D. melanogaster Ace* gene was designated as *Ace* orthologous gene. It is now clear that *D. melanogaster* has only one *Ace* gene as confirmed by its genome sequence [11], whereas most other insect species have two different *Ace* genes (i.e., *Ace1* and *Ace2*) [12]. *Ace1* commonly refers to the *Ace* paralogous (AP-*Ace*) gene and *Ace2* the

Ace orthologous (AO-*Ace*) gene in relation to the *D. melanogaster Ace* [8].

To date, the cDNAs encoding AChEs have been sequenced from at least 43 insect species. Among them, both *Ace1* and *Ace2* have been reported from each of 27 species, including *Bombyx mandarina* (GenBank accession numbers: EU262633 for *BmAce1* and EU262632 for *BmAce2*); *Sitobion avenae* [13]; *Rhopalosiphum padi* [13]; *Anopheles gambiae* [14]; *Liposcelis decolor* (GenBank accession numbers: FJ647186 for *LdeAce1* and FJ647187 for *LdeAce2*); *Orchesella villosa* [15]; *Liposcelis entomophila* (GenBank accession numbers: EU854149 for *LeAce1* and EU854150 for *LeAce2*); *Blattella germanica* [16]; *Bemisia tabaci* [17]; *Culex pipiens quinquefasciatus* (GenBank accession numbers: XM_001847396 for *CqAce1* and XM_001842175 *CqAce2*); *Bombyx mori* [18]; *Acyrthosiphon pisum* (GenBank accession numbers: XM_001948618 for *ApAce1* and XM_001948953 for *ApAce2*); *Nasonia vitripennis* (GenBank accession numbers: XM_001600408 for *NvAce1* and XM_001605518 for *NvAce2*); *Pediculus humanus corporis* [19]; *Cydia pomonella* [20]; *Helicoverpa assulta* [21]; *Aedes albopictus* (GenBank accession numbers: AB218421 for *AaAce1* and AB218420 for *AaAce2*); *Aphis gossypii* [22]; *Culex tritaeniorhynchus* [23]; *Myzus persicae* [24]; *Culex pipiens* [25]; *Plutella xylostella* [26–28]; *Chilo suppressalis* [29]; *Pediculus*

humanus capitis [19]; Aedes aegypti [30]; Liposcelis bostrychophila (GenBank accession numbers: FJ647185 for LbAce1 and EF362950 for LbAce2) and Alphitobius diaperinus [31]. The remaining insect species may also have two Ace genes but only one Ace gene (Ace1 or Ace2) has been documented.

The existence of two Ace genes in insects has attracted much attention to the study of their functions, particularly their roles in insecticide resistance [8,32] and as targets for developing new insecticides [12,33–37]. Beetles (coleopterans) are the most evolutionarily successful metazoans, accounting for 25% of all known animal species, far more than any other taxonomic orders [31]. Despite the diversity and economic importance of coleopterans, Ace genes have been reported from only two species: Leptinotarsa decemlineata [5,6] and Alphitobius diaperinus [31]. Although T. castaneum (the red flour beetle) is one of the most notorious stored grain pests in the world and is now regarded as an emerging model organism, its Ace genes were only predicted from genomic sequence and detailed information on these genes has been limited.

In this paper, we report two Ace genes from T. castaneum. Our study of the two genes focuses on the genome organization, three-dimensional (3D) protein models, phylogenies, and expression patterns of the two genes at different developmental stages of the insect, in an effort to better understand the functions of the two genes and obtain insights into better strategies for insect pest control.

Results

AChE cDNA and deduced amino acid sequences

Based on the predicted sequences of two T. castaneum Ace genes in NCBI (XM_968369 and XM_965681), we designed specific primers (Table 1) to determine the full-length cDNAs of the two genes from the brain of T. castaneum. Each of several polymerase chain reaction (PCR) primer pairs was able to generate overlapping fragments for each gene and it was then possible to assemble them into its full-length cDNA of the protein coding region. The two deduced amino acid sequences show significant similarities to AChE1 (AP-AChE) and AChE2 (AO-AChE) proteins of other insects in GenBank according to our BLASTP analysis. Therefore, the two T. castaneum Ace genes are named TcAce1 (AP-Ace) and TcAce2 (AO-Ace), and their protein products are named TcAce1 and TcAce2, respectively. The TcAce1 cDNA contains 2148 base pairs (bp) and has an open reading frame (ORF) of 1944 bp, encoding a protein of 648 amino acid residues, whereas the TcAce2 cDNA contains 1,890 bp and has an ORF of 1,812 bp, encoding a protein of 604 residues. However, we were not able to obtain the 5′-untranslated region (5′-UTR) of the TcAce2 cDNA (Fig. 1).

TcAce1 and TcAce2 belong to typical Ace1- and Ace2-type genes, respectively, as judged by their sequence similarities with other known insect Aces (Fig. 2). The deduced amino acid sequences (TcAce1 and TcAce2) of TcAce1 and TcAce2 exhibit six and four N-glycosylation sites (N-X-S or N-X-T) [38], respectively (Fig. 1). Predicted isoelectric points (pI) and molecular masses of TcAce1 and TcAce2 are 6.58 and 5.39, and 72.81 and 68.15 kDa, respectively (http://www.scripps.edu/~cdputnam/protcalc.html). Both TcAce1 and TcAce2 are predicted to contain a cleavable signal peptide, which suggests that these proteins can be secreted and function in an extracellular environment. Both proteins have a C-terminal Cys residue (C617 in TcAce1 and C600 in TcAce2) that is likely to form an intermolecular disulfide bond. According to the analysis using PredGPI (http://gpcr.biocomp.unibo.it/

Table 1. PCR primers used to amplify cDNA sequences of both TcAce1 and TcAce2 genes and to analyze their gene expressions.

Primer name	Sequence (5′-3′)	Tm (°C)	Product size (bp)	Location[a]
PCR for cDNA sequences				
TcAce1-F[b]	CGGCCTGCTTACTAGTGATTCTAC	60.66	1714	17-1730
TcAce1-R	ACATCGAGGGTGAGAAACTCC	60.50		
TcAce1S-F	CAACGACCGTTGTGCAAATA	60.56	320	N/A-219
TcAce1S-R	CAGGGGATCATCTTCGGAGT	61.39		
TcAce1E-F	CGTTTGGACACCCACTTTCT	60.01	402	1665-N/A
TcAce1E-R	GTCGTGTTGATTTTGAATACCTCAC	60.99		
TcAce2-F	AGACCTCATCACGCTGTTTG	55.2	1034	750-1783
TcAce2-R	CTGGGTTATCCCGAAGCTTG	61.86		
TcAce2S-F	GTCGTAGAGGCGTCGTCGT	61.44	1014	N/A-895
TcAce2S-R	TCTCCCCCGACATGTAACTC	59.93		
TcAce2E-F	AACCAGTGACAGACGACGTG	59.78	267	1627-N/A
TcAce2E-R	CGCAACCGATGCGTTTAATA	61.86		
Quantitative real-time PCR				
TcAce1(Q)-F	CCGTTCGTCCCAGTCATTG	55.3	121	1069-1189
TcAce1(Q)-R	AGTAGTAGCCTTCTTCTGTGTTAG	55.4		
TcAce2(Q)-F	AGACCTCATCACGCTGTTTG	55.2	179	750-928
TcAce2(Q)-R	CCTCCACCAGGACCTTCC	54.9		
TcRps3-F	CCGTCGTATTCGTGAATTGAC	54.8	130	279-408
TcRps3-R	TCTAAGAGACTCTGCTTGTGC	54.7		

[a]Product location refers to the PCR fragment corresponding to the Ace gene nucleotide sequence of T. castaneum from NCBI database (TcAce1: XM_968369; TcAce2: XM_965681). N/A refers to the sequence based on the genomic sequence in Beetlebase (http://beetlebase.org/).
[b]F and R refer to forward and reverse primers, respectively.
doi:10.1371/journal.pone.0032288.t001

*TcAce*1

```
ACGACCGTTGTGCAAATACATTAAATCCAAACAAGTGTTTAAATCATCGGCCATTCCGTTTATTTATAGGCCGCTTTGTTGCAGAATTATTGAAAAGCAC ATG ACG GGA GCG TGG GCG GCC TGC TTA CTA GTG ATT CTA   138
                                                                                                       M   T   G   A   W   A   A   A   C   L   L   V   I   L    13

CTT CCC TCA TGC ATC CCT TCG CCA CAC CGA GGC CGC CAC CAC CCC CCC GAG CCT CAC GCC GAG GCC TAC CAC ATG TCC CGC GAC CCC TTC GAC CCG CAC CGG GAC TCG GAG GAG   252
 L   P   S   C   I   P   S   P   H   R   G   R   H   H   P   P   E   P   H   A   E   A   Y   H   M   S   R   D   P   F   D   P   H   R   D   S   E   E          51

TTC CGA CGA GAC GCC CCG GAC GAC AAG CGG GAG TTT ACG CGT CGC GAC TCC GAA GAT GAT CCC CTG GTC ATC AGC ACA AAG AAG GGA AAG GTC CGG GGC ATC TCC CTC ACG GCA   366
 F   R   R   D   A   P   D   D   K   R   E   F   T   R   R   D   S   E   D   D   P   L   V   I   Q   T   K   K   G   K   V   R   G   I   S   L   T   A          89

GCC ACG GGA AAG AAG GTC GAC GCG TGG CTG GGA ATC CCC TAC GCG CAG AAA CCG CTC GGC AAC CTC CGG TTC CGG CAC CCG CGG CCT GCG GAA AAG TGG GAG GGC GTG ATG AAC   480
 A   T   G   K   K   V   D   A   W   L   G   I   P   Y   A   Q   K   P   L   G   N   L   R   F   R   H   P   R   P   A   E   K   W   E   G   V   M   N          127

ACG ACT AGC CAG CCG AAT TCC TGC GTG CAG ATT ATA GAC ACA GTG TTT GGG GAT TTC CCA GGT GCG ACC ATG TGG AAC CCC AAC ACA CCG TTG AAC GAA GAC TGT CTT TAC GTG   594
 T   T   S   Q   P   N   S   C   V   Q   I   I   D   T   V   F   G   D   F   P   G   A   T   M   W   N   P   N   T   P   L   N   E   D   C   L   Y   V          165

AAC GTG GTG CCG AAG CCG AGG CCG ACC AGT GCC GCT GTC ATG GTG TGG GTG TTC GGT GGT GGA TTC TAC TCC GGG ACG AAC ACT TTA GAA GTT TAC GAC CAC AAC ATA CTA   708
 N   V   V   P   K   P   R   P   T   S   A   A   V   M   V   W   V   F   G   G   G   F   Y   S   G   T   N   T   L   E   V   Y   D   H   N   I   L          203

GTG TCC GAA GAG AAC ATT ATT TTA GTT TCG ATG CAG TAT AGA GTT GCT TCG CTC GGA TTC CTG TAT TTC GGG ACG CCG GAC GTT CCG GGA AAT GCA GGA TTG TTT GAC CAA ATG   822
 V   S   E   E   N   I   I   L   V   S   M   Q   Y   R   V   A   S   L   G   F   L   Y   F   G   T   P   D   V   P   G   N   A   G   L   F   D   Q   M          241

ATG GCG TTG CAA TGG GTA CGT GAT AAC ATC GCC GCT TTC GGA GGA AAC CCC AAC AAT ATC ACT CTA TTC GGA GAA TCA GCA GGA GCT GTT TCA GTC TCC CTG CAT CTA CTC TCT   936
 M   A   L   Q   W   V   R   D   N   I   A   A   F   G   G   N   P   N   N   I   T   L   F   G   E   S   A   G   A   V   S   V   S   L   H   L   L   S          279

CCA TTA TCG AGA GAC CTT TTC TCA CAA GCT ATC ATG GAA TCT GGG AGT GCG ACA CCT TGG GCT ATA ATC TCC CGT GAA GAA AGC ATT TTG CGA GGA TTG TTG AGA CTA GCA GAA   1050
 P   L   S   R   N   L   F   S   Q   A   I   M   E   S   G   S   A   T   A   P   W   A   I   I   S   R   E   E   S   I   L   R   G   L   R   L   A   E          317

GCT GTG GGT TGT CCG CAT GAG CGC CAC GAG CTT TCC GCT GTT ATA GAC TGC TTA AAA AAG AAA GAC CCT ATT GAT TTA GTC AAT AAC GAA TGG GGA ACA CTC GGC ATA TGT GAG   1164
 A   V   G   C   P   H   E   R   H   E   L   S   A   V   I   D   C   L   K   K   K   D   P   I   D   L   V   N   N   E   W   G   T   L   G   I   C   E          355

TCT CCG TTC GTC CCA GTC ATT GAT GGA GCC TTT CTG GAC GAA TCC CCC ACA CGG GCT CTG GCA AAC AAA AAC TTC AAA AAA ACC AAT ATC CTC ATG GGT TCT AAC ACA GAA GAA   1278
 S   P   F   V   P   V   I   D   G   A   F   L   D   E   S   P   T   R   A   L   A   N   K   N   F   K   K   T   N   I   L   M   G   S   N   T   E   E          393

GGC TAC TAC TTC ATA ATC TAC TAC TTG ACC GAA CTT TTC CGA AAG GAG GAG AAC GTT TAC GTC AAC CGG CAG GAA TTC CTG CGG GCT GTA ACC GAA CTG AAC CCG TAC TTC AAC   1392
 G   Y   Y   F   I   I   Y   Y   L   T   E   L   F   R   K   E   E   N   V   Y   V   N   R   Q   E   F   L   R   A   V   T   E   L   N   P   Y   F   N          431

GCG ATT TCG CGC CAA GCA ATC GTC TTC GAA TAC ACC AAC TGG TTA AAC CCC GAC GAC CCG GTG AGC AAC CGC GAC TCC CTC GAT AAA ATG GTA GGC GAC TAC CAC TTC ACT TGC   1506
 A   I   S   R   Q   A   I   V   F   E   Y   T   N   W   L   N   P   D   D   P   V   S   N   R   D   S   L   D   K   M   V   G   D   Y   H   F   T   C          469

AAC GTC AAC GAA TTC GCC CAC CGG TAC GCC GAA ACC GGC AAC ACG GTC TAC ATG TAC TAC TAC AAG CAC CGG ACA GTG GCG AAT CCC TGG CCC TCC TGG ACC GGC GTG ATG CAC   1620
 N   V   N   E   F   A   H   R   Y   A   E   T   G   N   T   V   Y   M   Y   Y   Y   K   H   R   T   V   A   N   P   W   P   S   W   T   G   V   M   H          507

GCC GAC GAA ATC AAC TAC GTC TTC GGA GAA CCA CTC AAT CCC ACT AAA AGC CAC ACA GCA CAA GAA GTC GAT CTG AGT AAG ATC ATG AGA TAC TGG GCC AAT TTC GCC AAG   1734
 A   D   E   I   N   Y   V   F   G   E   P   L   N   P   T   K   S   H   T   A   Q   E   V   D   L   S   K   I   M   R   Y   W   A   N   F   A   K          545

ACG GGC AAC CCG AGT CAG TCG CCG AAC GGC GTT TGG ACC CCG ACC TTC TGG CCC CCG CAC ACA GCT TTC GGA AGG GAG TTT CTC ACC CTC GAT GTC AAC TCC ACT GCC ACA GGG   1848
 T   G   N   P   S   Q   S   P   N   G   V   W   T   P   T   F   W   P   P   H   T   A   F   G   R   E   F   L   T   L   D   V   N   S   T   A   T   G          583

AGG GGA CCT AGA CTC AAG CAG TGT GCC TTC TGG AAG AAA TAC CTA CCT CAA CTT CAG CAG CAA ACG AGC GAG CTT CTG AAC CAA CCG CCC CGG CAA AAC TGT ACC GAC GCG GCC   1962
 R   G   P   R   L   K   Q   C   A   F   W   K   K   Y   L   P   Q   L   Q   Q   Q   T   S   E   L   L   N   Q   P   P   R   Q   N   C   T   D   A   A          621

AGT TCG TTG CGA TGG TCA AGA GAC GGC GCT GCC GGC CTC CTG ATG GTG TCG ACA GTG GCG GCG CTC CTC GCG GGG CCT TTC TAA CTTAGAGAAAGTCCGTTTAGGAGTTCCCCTGAAGCGCA   2085
 S   S   L   R   W   S   R   D   G   A   A   G   L   L   M   V   S   T   V   A   A   L   L   A   G   P   F   *                                                 648

TTAGAAACTAGCGCCTTGAAACATATCTCTAGTATTGTGTGTGAGGTATTCAAAATCAACACG                                                                                            2148
```

*TcAce*2

```
ATG GGT TCG AAT CTC GTC GTC GTC GTG GTC GTC GTG GTG GTG GTG GTG GCG TCG CTG AGT GCT TCC GTC CGC GCC TAC TCC TGG CCC AGC GAG GAG ACG ACG ACG CGT CCG CCC   114
 M   G   S   N   L   V   V   V   V   V   V   V   V   V   V   V   V   A   S   L   S   A   S   V   R   A   Y   S   W   P   S   E   E   T   T   T   R   P   P    38

CAA GCC AGA GAC TTT CAC AGC GAC CCG CTG GTG GTG GAG ACC ACC AGC GGC CTG GCG CGG GGC AAG GCC AAG ACG GTG CTG GGC CGC GAG GTG CAC GTC TTC ACC GGG ATT CCC   228
 Q   A   R   D   F   H   S   D   P   L   V   V   E   T   T   S   G   L   A   R   G   K   A   K   T   V   L   G   R   E   V   H   V   F   T   G   I   P          76

TTC GCC AAG CCG CCC ATC GAA CAG CTG AGG TTC AGG AAG CCC GTC CCC ATC GAT CCC TGG CAC GGC ACC TTA GAT GCC ACC AAG CTG CCC AAC TCG TGC TAC CAG GAG CGC TAC   342
 F   A   K   P   P   I   E   Q   L   R   F   R   K   P   V   P   I   D   P   W   H   G   T   L   D   A   T   K   L   P   N   S   C   Y   Q   E   R   Y          114

GAG TAC TTC CCC GGC TTT GAG GGC GAG GAG ATG TGG AAC CCC AAC ACG AAC ATC TCC GAG GAC TGC CTC TAC CTC AAC ATC TGG GTG CCC CAG CGC TTG CGC ATC CGG CAC CAC   456
 E   Y   F   P   G   F   E   G   E   E   M   W   N   P   N   T   N   I   S   E   D   C   L   Y   L   N   I   W   V   P   Q   R   L   R   I   R   H   H          152

GGC GAG AAG CTC CCC CAG GAC CGG CCC AAA GTC CCC GTC CTC GTG TGG ATC TAC GGC GGG GGC TAC ATG AGC GGC ACC TCC ACC CTG GAC ATC TAC GAC GCC GAC ATC ATC GCC   570
 G   E   K   L   P   Q   D   R   P   K   V   P   V   L   V   W   I   Y   G   G   G   Y   M   S   G   T   S   T   L   D   I   Y   D   A   D   I   I   A          190

GCC ACC TCC GAC GTC ATC GTC AGC ATG CAA TAT CGC GTG GGC GCC TTC GGC TTC CTC TAC TTG AGC AAG TAC TTC CCT GGC GAG AGT GAG GAG GCG CCC GGT AAC ATG GGC   684
 A   T   S   D   V   I   V   A   S   M   Q   Y   R   V   G   A   F   G   F   L   Y   L   S   K   Y   F   P   R   G   S   E   E   A   P   G   N   M   G          228

ATG TGG GAC CAA GCC CTG GCC ATC CGC TGG ATC AAG GAG AAC GCG GCG GCC TTT GGG GGC GAC CCA GAC CTC ATC ACG CTG TTT GGG GAG TCG GCT GGA GGC GGC TCC GTC AGC   798
 M   W   D   Q   A   L   A   I   R   W   I   K   E   N   A   A   A   F   G   G   D   P   D   L   I   T   L   F   G   E   S   A   G   G   G   S   V   S          266

ATC CTG CTC CTG AGT CCG GTT ACT AAA GGC CTG GCC AGG AGG GGG ATT CTG CAG TCG GGG ACT ATG AAC GCC CCT TGG AGT TAC ATG TCG GGG GAG AGG GCG CCG CAA ATC GGG   912
 I   L   L   L   S   P   V   T   K   G   L   A   R   R   G   I   L   Q   S   G   T   M   N   A   P   W   S   Y   M   S   G   E   R   A   P   Q   I   G          304

AAG GTC CTG GTG GAG GAC TGC GGG TGC AAC GTC TCC TTG TTG GAG ACG AGG CCG CAT GAG GTC ATT GAT TGC ATG AGG GCG GTG GAG GCC AAG ACG ATT TCG CTG CAA CAG TGG   1026
 K   V   L   V   E   D   C   G   C   N   V   S   L   L   E   T   R   P   H   E   V   I   D   C   M   R   A   V   E   A   K   T   I   S   L   Q   Q   W          342

AAT TCG TAT TCG GGG ATT TTG GGC TTC CCC TCA GAC CCT ACG GTT GAT GGC GTC TTC ATG CCC AAG CAT CCC ATG GAT ATG CTG GCG GAA GGG GAT TAC GAG GAT ATG GAG ATC   1140
 N   S   Y   S   G   I   L   G   F   P   S   T   P   T   V   D   G   V   F   M   P   K   H   P   M   D   M   L   A   E   G   D   Y   E   D   M   E   I          380

CTG GTC GGG AGT AAC CAA GAT GAA GGC ACT TAC TTC TTA CTT TAC GAT TTT ATC GAT TTC TTC GAA AAG GAT GGC CCT AGC TTC CTC CAA CGA GAC AAA TAC CAC GAC ATT ATC   1254
 L   V   G   S   N   Q   D   E   G   T   Y   F   L   L   Y   D   F   I   D   F   F   E   K   D   G   P   S   F   L   Q   R   D   K   Y   H   D   I   I          418

GAT ACG ATA TTC AAA AAT ATG AGT CGG TTG GAA CGT GAT GCC GTC ATA GTA TTT CAG TAT ACT GAT TGG GAG CAC GTC AAC GAC GGC TAC TTG AAC CAA AAG ATG GTG GGC GAC GTC   1368
 D   T   I   F   K   N   M   S   R   L   E   R   D   A   I   V   F   Q   Y   T   D   W   E   H   V   N   D   G   Y   L   N   Q   K   M   V   G   D   V          456

GTC GGA GAT TAT TTT TTC ATT TGT CCA ACC AAC GAT TTC GCC GAG CTG GCA GCA GAG CGC GGA ATG AAA GTC TAC TAT TAT TTT TTC ACA CAC AGG ACA AGC ACG TCG TTG TGG   1482
 V   G   D   Y   F   F   I   C   P   T   N   D   F   A   E   L   A   A   E   R   G   M   K   V   Y   Y   Y   F   F   T   H   R   T   S   T   S   L   W          494

GGC GAA TGG ATG GGG GTG ATG CAC GGG GAT GAG ATA GAA TAC GTG TTT GGC CAT CCT TTG AAC ATC TCG TTG CAG TTT AAC TCA AGG GAA CGG GAA CTC AGT CTG AAG ATA ATG   1596
 G   E   W   M   G   V   M   H   G   D   E   I   E   Y   V   F   G   H   P   L   N   M   S   L   Q   F   N   S   R   E   R   E   L   S   L   K   I   M          532

CAA GCC TTT GCA AGA TTT GCA GCA ACG GGG AAA CCA GTG ACA GAC GAC GTG AAT TGG CCA CTA TAT ATC AAA GAC CAA CCG CAG TAT TTC ATC TTC AAC GAC GAC AAA AAC GGC   1710
 Q   A   F   A   R   F   A   A   T   G   K   P   V   T   D   D   V   N   W   P   L   Y   I   K   D   Q   P   Q   Y   F   I   F   N   A   D   K   N   G          570

ATC GGC AAA GGT CCT CGA GCG ACA GCG TGC GCG TTT TGG AAC GAT TTC CTG CCC AAG CTT CGG GAT AAC CCA GGT AAA TTC AAT CGT TGC AGT CTG CAT CAT TAA ACGGATCGATT   1826
 I   G   K   G   P   R   A   T   A   C   A   F   W   N   D   F   L   P   K   L   R   D   N   P   G   K   F   N   R   C   S   L   H   H   *                       604

GTGTGGCCGACGAGATCAATTTCAAAACAAGCTTAGAATGCGCTTAATTATTAAACGCATCGGT                                                                                            1890
```

predgpi/), TcAce1 appears to contain a GPI-anchor, which is linked to the C-terminal residue after a proteolytic cleavage at the ω site (D619 in TcAce1), whereas TcAce2 doesn't seem to contain a GPI-anchor. Both proteins have relatively high sequence identity (47% for TcAce1 and 59% for TcAce2) to the C-terminal residues 526–543 (*i.e.*, QTCAFWNRFLPKLLSAT) of the recombinant mouse AChE (mAChE) [39,40]. This level of sequence identity is identical or higher than the corresponding sequence identity (47%) between mAChE and *D. melanogaster* AChE (DmAChE). All these suggest that residues 589–605 in TcAce1 and residues 578–594 in TcAce2 are likely responsible for the formation of a dimeric four-helix bundle at the C-terminus as seen in mAChE [40] and DmAChE [41].

Chromosomal locations of *TcAce1* and *TcAce2*

The exon-intron organizations of *TcAce1* and *TcAce2* were revealed by comparisons of the full-length cDNAs with their corresponding genomic sequences (http://beetlebase.org/). The lengths of *TcAce1* and *TcAce2* genomic DNA sequences are 2,986 bp and 32,243 bp, respectively. Genome structure analysis shows that the two genes are located on different chromosomes of *T. castaneum*; *TcAce1* is located on chromosome 5, whereas *TcAce2* on chromosome 2. *TcAce1* has two exons and one intron, whereas *TcAce2* has six exons and five introns (Fig. 3).

Phylogenetic relationship of *T. castaneum* AChE to other AChEs

The phylogenetic tree of the deduced amino acid sequences of AChEs from the Pacific electric ray (*Torpedo californica*), twospotted spider mite (*Tetranychus urticae*), and all the insect species available in GenBank was generated using the neighbor-joining method. Phylogenetic analysis suggests that there are two major groups (Ace1 and Ace2); TcAce1 and TcAce2 belong to the *Ace1*- and *Ace2*-type genes, respectively (Fig. 4). As expected, the cDNA-deduced TcAce1 has high protein sequence identities to BgAce1 (70%), HaAce1 (64%), AgAce1 (61%), SaAce1 (57%), and SgAce1 (57%). The cDNA-deduced TcAce2 has high protein sequence identities to LdAce2 (84%), HaAce2 (69%), BgAce2 (68%), AgAce2 (60%), SaAce2 (57%), and DmAce (55%) (Table 2). TcAce1 and TcAce2 exhibited 39% and 38% protein sequence identities to *T. californica* AChE, respectively. However, the protein sequence identities of Ace1 and Ace2 in the same insect species are 36% (between TcAce1 and TcAce2), 31% (between SaAce1 and SaAce2), 35% (between AgAce1 and AgAce2), 35% (between BgAce1 and BgAce2), and 32% (between HaAce1 and HaAce2) (Table 2).

It is worth noting that, per the present nomenclature, the reported PhAce1, PhcAce1, AgoAce1, MpAce1 and CtAce1 in the Ace2 group should change to PhAce2, PhcAce2, AgoAce2, MpAce2 and CtAce2, respectively, whereas the reported PhAce2, PhcAce2, AgoAce2, MpAce2 and CtAce2 in the Ace1 group should be named PhAce1, PhcAce1, AgoAce1, MpAce1 and CtAce1, respectively (Fig. 4).

Three-dimensional models

TcAce1 (TcAP-AChE). Eighteen 10-ns molecular dynamics simulations of the substrate-bound TcAce1 protein homology

model derived from a human butyrylcholinesterase (hBChE) crystal structure (Protein Data Bank ID: 2J4C [42]) resulted in a time-averaged model with a distorted catalytic triad and partial unfold of the omega loop. This result indicates that the homology model is structurally unstable, presumably due to the low sequence identity of the omega loop between TcAce1 and hBChE that results in gaps in the omega loop. Another homology model was therefore built from a simulation-refined model of the African malaria mosquito (*A. gambiae*) AP-AChE (Protein Data Bank ID: 2AZG) [33] that has a higher sequence identity to TcAce1 (73%) than the identity between hBChE and TcAce1 (46%). Twenty-two 10-ns molecular dynamics simulations of this model liganded with its substrate yielded a time-averaged model without distortions in the catalytic triad and the omega loop. In this final model of TcAce1 (Fig. 5), C354 is exposed to solvent and accessible to covalent bonding at the opening of the active-site gorge [33]; R407 is enclosed by F143, F146, and F406 via cation-pi interactions; Y189, Y396, and Y400 adopt conformations that make the gorge relatively open; ACh adopts the fully extended conformation with its ammonium group forming a cation-pi interaction with the indole ring of W152 and its carbonyl oxygen atom anchored at the oxyanion hole comprising of G186, G187, and A268; E266 forms a hydrogen bond with Y198 at the bottom of the active-site gorge; E393, H507, and S267 form the catalytic triad. The TcAce1 active site is very similar to those of *A. gambiae* and *S. graminum* AP-AChEs [33,34], and different from the human AChE active site in that Y449 in the human enzyme is replaced by Asp creating void space at the bottom of the TcAce1 active site (Fig. 6). In *T. californica* AChE, rotation of Y442, which corresponds to Y449 of human AChE, reportedly controls the opening of a 3.4-Å-wide channel that enables rapid clearance of substrate hydrolysis products [43]. In DmAChE, the counterpart Y449 is also mutated to Asp; hence the crystal structure of DmAChE (Protein Data Bank ID: 1DX4 [41]) has a channel with a diameter of ~5 Å that is formed by G79, W83, W472, L479, and D482 and connects the active-site gorge to solvent [44]. Because of the mutation Tyr of human to Asp of TcAce1, the final model of TcAce1 has a similar channel that is comprised of G148, W152, W499, M506, and D509. The TcAce1 channel is, however, partially blocked by M151 and P498.

TcAce2 (TcAO-AChE). A homology study identified the crystal structure of DmAChE (PDB ID: 1DX4 [41]) as a template with a sequence identity of 60% and generated a homology model of TcAce2. This model has a protruded large loop conformation for residues 145–162, which is due to the omission of the corresponding loop (residues 103–136) in the 1DX4 crystal structure. An initial set of 21 10-ns molecular dynamics simulations of the TcAce2 homology model without residues 145–162 resulted in a time-average model with the catalytic triad distorted.

In the loop of residues 145–162, there are two histidine residues, four arginine residues, two lysine residues, one aspartate residue, and one glutamate residue. At the physiological pH of 7.4, this loop has a net charge of +4. To avoid a possible effect of the highly-charged loop on the catalytic triad distortion, a second set of 21 10-ns simulations of the TcAce2 homology model possessing residues 145–162 was performed. The triad was still distorted in an average conformation of all trajectories saved at 1.0-ps intervals

```
DmAce   ----------------------------------------------MAISCRQSRVLPMSLPLPLT--------IPLPLVLVL------------------------------------------  29
AgAce2  -----------------------------------------------MASAYYHQSAVGVGN--------------VLVLLLG-------------------------------------------  22
TcAce2  ----------------------------------------MGSNLVVVVVVVVVVVVASLSASVRAYSWP--------SEETT-------------------------------------------  34
LdAce2  ----------------------------------------MG--QLSILCLFVTVCASVCGYSWP--------SDETT-----------------------------------------------  28
HaAce2  ----------------------------------MPVRTTIFEMISNTKIVFTKLLLCCFVSGAVARSWANH--HDTTTSTTQTTP----------------------------------------  50
BgAce2  ----------------------------------MATQCLLILAGCICTAMAIGAGSKPHFNSSSASSTP---HSTIGGGSSP-------------------------------------------  46
SaAce2  ----------------------------------MSVDCVYTSAVTLLLCCSAVLGRPSSNGGADGGGGGGGGGAGAGGGGGGGG-----------------------------------------  53
TcAce1  ----------------------MTG-AWAAC------LLVILLPSCIPSPHRGRHHPPEPHAEAYHMSRDPFDPHRDS---EEFRRDAPDDK----------------------------------  60
BgAce1  ----------------------MDVGDLVG-GLSVGSGDTDPLRDIHIREHVIKDHVHSHHGSGGGDPSYMSRGDPSIVGRGEPMDMGRRDPPVDFGRRHSIDMDR-------------------  83
HaAce1  ---------MRVVLAALTALAARALAGPHEHRARHHAPEHPPHFPAPAPPQPYRGHGEAVRYNPELDTILPRLEDHETSSKRAKFEDAETSSKRAKYD--------------------------  89
SaAce1  ----------------------MDQWLLWFGYLVASTYGLSLRHARHQSVGTPTVEEILEPQILIEDTDHVFRQRASDMFAQEPEYTEKRNLHRRRSEFSGNQ-------------------  82
SgAce1  ----------------------MDQWLLWFGYLVASTYGLSLRHARHQSVGTPTAEEILEPQILIEDTDHVFRQRASDMFAQEPEYTEKRNLHRRRSEFSGNQ-------------------  82
AgAce1  MEIRGLLMGRLRLGRRMVPLGLLLGVTALLLILPPFALVQGRHHELNNGAAIGSHQLSAAAGVGLASQSAQSGSLASGVMSSVPAAGASSSSSSSLLSSSAEDDVARITLSKDADAFFTPYIGHGESVRIIDAELGT  136
TcaAce  ------------------------------------------MNLLVTSSLGVLLHLVVLCQADDHS-------------------------------------------------------  25
```

```
                                                                                    1    ⇓                 ⇓          1
DmAce   -----------------SLHLSGVCGVIDRLVVQTSSGPVRGR-SVTVQGREVHVYTGIPYAKPPVEDLRFRKPVPAEPWHG-----VLDATGLSATCQERYEYFPGFSGEEIWNPNTNVSEDLLYINVWAPAKAR  143
AgAce2  -----------------ATVICPAYAIIDRLVVQTSSGPIRGR-STMVQGREVHVFNGVPFAKPPVDSLRFKKPVPAEPWHG-----VLDATRLPPSCIQERYEYFPGFAGEEMWNPNTNVSEDLLYLNIWVPTKTR  136
TcAce2  -----------------TRPPQARDFHSDPLVVETTSGLARGK-AKTVLGREVHVFTGIPFAKPPIEQLRFRKPVPIDPWHG-----TLDATKLPNSCYQERYEYFPGFEGEEMWNPNTNISEDCLYLNIWVPQRLR  148
LdAce2  -----------------TKPSQFKDFHTDPLVVETTSGLVRGY-SKTVLGREVHVFTGIPFAKPPIEQLRFRKPVPIDPWHG-----ILDATKQPNSCFQERYEYFPGFEGEEMWNPNTNISEDCLYLNIWVPQRLR  142
HaAce2  -----------------TTSPVPKNFHNDPLVVETKSGLVKGY-AKTVMGREVHIFTGIPFAKPPLGPLRFRKPVPIDPWHG-----VLEATAMPNSCYQERYEYFPGFEGEEMWNPNTNISEDCLYLNIWVPQHLR  164
BgAce2  -----------------NTNPGKLDYHDDPLVVETQSGLVRGS-ARVVLGKEVHIFTGIPFAKPPVGPLRFRRPVPVEPWHG-----VLDATALPNSCYQERYEYFKGFEGEEMWNPNTNVSEDCLYLNIWVPQRVR  160
SaAce2  -----------------GGGGSAVDDTDESPVVVTSSGMVQGY-TKIIANRGVRVYTGIPFAKPPVGPLRFRRPVPVDPWTG-----VLNATRLPNTCYQERYEYFPGFVGEEMWNPNTNKLSEDCLYLNIWIPKKQR  167
TcAce1  -----------------REFTRRDSEDDPLVIQTKKGKVRGISLTAATGKKVDAWLGIPYAQKPLGNLRFRHPRPAEKWEG---VMNTTSQPNSCYQIIDTVFGDFPGATMWNPNTPLSEDCLYVNVVVP-KPR  173
BgAce1  VGHGGDPLDLAHESRMSEEGSGHGMQDDPLVIDTKKGKVRGITLTAATGKLVDAWLGIPYAQKPLGPLRFRHPRPVDRWDKHG-EIYNATKMPNSCVQIIDTVFGDFPGALIWNPNTPLSEDCLYINVVVP-KPR  216
HaAce1  ERFYSNHERADEEPMLADEPQLGPEDEDPLVVRTRKGRVRGITLTAATGKKVDAWFGIPYAQKPVGDLRFRHPRPAESWGD---EILNTTTLPHSCVQIIDTVFGDFPGAMMWNPNTDMQEDCLFINIVTP-RPR  220
SaAce1  -----------------DNDFESSGETYSAYKSDDPLVIHTNKGKIRGITQAASSGKLVDAWLGIPYAKKPIGDLRFRHPRPIDRWDKTSPETILNCTTPPNTCVQIFDTLFGDFPGATMWNPNSPVSEDCLYINVVVP-KLR  207
SgAce1  -----------------DTDFASSGETYSAYTSDDPLIHTNKGKIRGITQTATTGKLVDAWLGIPYAKKPIGDLRFRHPRPIDRWDTTTPETILNCTTPPNTCVQIFDTLFGDFPGATMWNPNSPVSEDCLYINVVVP-KLR  207
AgAce1  LEHVISGATPRRRGLTRRESNSDANDNDPLVVNTDKGRTRGITVDAPSGKKVDVWLGIPYAQPPVGPLRFRHPRPAEKWTG---VLNTTTPPNSCVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVVAP-RPR  266
TcaAce  -----------------ELLVNTKSGKVMGTRVPVLSS-HISAFLGIPFAEPPVGNMRFRRPEPKKPWSG---VWNASTYPNNCQQYVDEQFPGFSGSEMWNPNREMSEDLLYLNIWVP-SPR  126
             :  .  *.            .  :**:: .  :  :.   *:**::*.   .  *  .         :  .  **    .***   **  :****::  :.*****:::.**
```

```
                   ⇓          ⇓         ⇓
DmAce   LRHGRGANGGEHPNGKQADTDHLIHNGNPQNTTNGLPILIWIYGGGFMTGSATLDIYGGGFMSGSTSTLDIYNAEIMAAVGNVIVASFQYRVGAFGFLHLAPEMPSEFAEEAPGNVGLWDQALAIRWLKDNAHAFGGPEWMTLFGESAGS  279
AgAce2  LRHGRGLNFG--SNDYFQDDDDDFQR---QHQSKGGLAMLVWIYGGGFMSGTSTLDIYNAEILAAVGNVIVASMQYRVGAFGFLYLAPYIN-GYEEDAPGNMGMWDQALAIRWLKENAKAFGGDPDLITLFGSAGG  266
TcAce2  TRHHIGEKLPQ-----------------DRPKVPVLVWIYGGGYMSGTSTLDIYDADIIAATSDVIVASMQYRVGAFGFLYLSKYFP-RGSEAPGNMGMWDQALAIRWIKENAAAFGGDPDLITLFGESAGG  262
LdAce2  IRHHADKPTI------------------DRPKVPVLIWIYGGGYMSGTATLDVYDADIIAATSDVIVASMQYRLGSFGFLYLNRYFP-RGSDETPGNMGLWDQLALAIRWIKDNAAAFGGDPDLITLFGESAGG  256
HaAce2  VRHHQDKPLA------------------ERPKVPILVWIYGGGYMSGTATLDYLYKADIMASSSDVIVASMQYRVGAFGFLYLNKYFS-PGSEEAPGNMGLWDQQLALAIRWIKDNARAFGGDPELITLFGESAGG  278
BgAce2  LRHSG---------------------ERPKVPVLVWIYGGGYMSGTSTLDVYDADIVVATSDIIVASMQYRVGSFGFLYLKPFFG-PDSEEAPGNMGMWDQALAIRWIKDNAAAFGGDPDLITLFGESAGG  269
SaAce2  TRHHSNNAHH---------AKIPVLVWIYGGGYMSGTSTLDIYDGDLLAVTFDVNYRAIGKALAGFGSLVLTPELP-EDSDDAPGNMGLWDQALAFGSLVLTPELP-EDSDDAPGNMGLWDQALAIRWLKENATVFGADPETITLFGESAG  279
TcAce1  PT-----------------------SAAVMVWVFGGGFYSGTNTLEVYDHNILVSEENIILVSMQYRVASLGFLYFG-----TPDVPGNAGLFDQMMALQWVRNIAAFGGNPNNITLFGESAGA  270
BgAce1  PK-----------------------NAAVMLWVFGGGFYSGATLDVYDHKTLVSEENIIVSMQYRVASLGFLFFD-----TIDVPGNAGLFDQLMALQWVHDNIQAFGGNPNNVTLFGESAGA  313
HaAce1  PK-----------------------NAAVMLWVFGGGFYSGTATLDVYDPKILVSEENVIVSMQYRVASLGFLFFD-----TPDVPGNAGLFDQLMALQWVRDNIAYFGGNPNHITLFGESAGA  317
SaAce1  PQ-----------------------NAAVVWIFGGGFYSGSATLDIYDPKVLVSEENIILSMQYRVASLGFLYFD-----TEDVPGNAGLFDQLMALQWVHENIKLFGGNPNNVTLFGESAGA  304
SgAce1  PQ-----------------------NAAVMWIFGGGFYSGSATLDIYDPKILVSEENVIVSMQYRVASLGFLYFD-----TEDVPGNAGLFDQLMALQWVHENIKLFGGNPNNVTLFGESAGA  304
AgAce1  PK-----------------------NAAVMLWVFGGGFYSGTATLDVYDHRALASEENVIVSLQYRVASLGFLFLG-----TPEAPGNAGLFDQNLALRWVRDNIHRFGGDPSRVTLFGESAGG  363
TcaAce  PK-----------------------STTVMVWIYGGGFYSGSSTLDVYNGKYLAYTEEVVLSLSYRVGAFGFLALHG-----SQEAPGNVGLLDQRMALQWVHDNIQFFGGDPKTVTIFGESAGG  224
                 .   *             :::**:.:.****: :*.  **:   :  .   .: .*:.*:**::: .**   :  ::.*  . *  .  :::. *: :****
```

```
                                ⇓                2                2      ⇓   ⇓      ⇓   ⇓   ⇓
DmAce   SSVNAQLMSPVTRGLVKRGMMQSGTMNAPWSHMTSEKAVEIGKALINDCNASMLKTNPAHVMSCMRSVDAKTISVQQWNSYS--GILSFPSAPTIDGAFLPADPMTLMKTADLKDYDILMGNVRDEGTYFLLYD  413
AgAce2  SSVSLHLLSPVTRGLSKRGILQSGTLNAPWSHMTAEKALQIAEGLIDDCNLTMLKESPSTVMQCMRNVDAKTISVQQWNSYS--GILGFPSAPTIDGVFMTADPMTMLREANLEGIDILVGSNRDEGTYFLLYD  400
TcAce2  GSVSILLLSPVTKGLARGILQSGTMNAPWSYMSGERAPQIGKVLVEDCGNVSLLETRPHEVIDCMRAVDAKTISLQQWNSYS--GILGFPSTPTVDGVFPKHPMDMLAEGDYEDMEILVGSNQDEGTYFLLYD  396
LdAce2  GSISIHLLSPVTKGLVRRGIMQSGTMNAPWSYMSGEAEQIGKILIQDCGNVSLLENSPRKVMDCMRAVDAKTISLQQWNSYS--GILGFPSTPTIEGVLLPKHPMDMLAEGDYEDMEILLGSNHDEGTYFLLYD  390
HaAce2  GSVSLHMLSPEMKGLFKRGILQSGTLNAPWSWMTGERAQDIGKVLVDDCNSSLLAADPSLVMDCMRGVDAKTISVQQWNSYT--GILGFPSAPTVDGVFLPKDPDTMMKEGNFHNTEVLLGSNQDEGTYFLLYD  412
BgAce2  GSVSLHLMSPITKGLVRMGIMQSGTMNAPWSYMSGEKAADIGRVLVEDCGNSTQLSEAPSKVMACLRSVDSKAISVQWNSYW--GILGFPSAPTIDGVFLPKHPMDLIKELDWEDTEILIGSNQDEGTYFLLYD  403
SaAce2  GSVSVHLLISPETRGMVRRGIIQSGTVNAPWSYMTGERAVDIAKKLLDDCNSTSLDSNPIGTMSCLRSVDASTISKKQWNSYS--GILGFPSAPTVDGVLLPEHPDMLAKANFSDIDILIGSNLNEGTYFLLYD  413
TcAce1  VSVSLHLLSPLSRNLFSQAIMESGSATAPWAIISREESILRGLRLAEAVGCPH--ERHELSAVIDCLKKKDPIDLVNNEWGTL--GICESPFVPVIDGAFLDESPTRALANKNFKKTNILMGSNTEEGYFIIYY  401
BgAce1  VSVSLHLLSPLSRNLFSQAIMESGSPTAPWAIISREESILRGIRLAEAVGCPR--SRSDIRAVIDCLRKKNATDLVNNEWGTL--GICEPPFVPIIDGILDGPPGRSFFEGTELPAKIKNFKKTNILMGSNTEEGYYFIIYY  444
HaAce1  VSVSLHLLSPLSRNLFSQAIMQSGAATAPWAIISREESILRGIRLAEAVHCPH--SRTDMGPMIECLRKKSPDELVNNEWGTL--GICEFPFVPIIDGSFLDELPVRSLVHQNFKKTNILMGSNTEEGYYFILYY  448
SaAce1  VSVSLHLLSPLSRNLFNQAIMESGSSTAPWAILSREESYNRGLKLAKAMGCPD--DRNAIHKTVECLRKVNSSAMVEKEWDHV--AICFFPFVPVVDGAFLDDYPQKSLSTNNFKKTNILMGSNSEEGYYSIFYY  435
SgAce1  VSVSLHLLSPLSRNLFNQAIMESGSSTAPWAILSREESFNRGLKLAKAMGCPD--DRNTIHKTVECLRKANSSVMVEKEWDHV--AICFFPFVPVVDGAFLDDYPQKSLSTNNFKKTNILMGSNSEEGYYSIFYY  435
AgAce1  VSVSLHLLSALSRDLFQRAILQSGSPTAPWALVSREEATLRALRLAEAVGCPH--EPSKLSDAVECLRGKDPHVLVNNEWGTL--GICEPPFVPVVDGAFLDETPQRSLASGRFKKTEILTGSNTEEGYYFIIYY  494
TcaAce  ASVGMHILSPGSRDLFRRAILQSGSPNCPWASVSVAEGRRRAVELGRNLNL--NSD--EELIHLRKEKKPQELIDVEWNVLPFDSIRFSFVPVIDGEFFPTSLESMLNSGNFKKTQILLGVNKDEGSFLLYG  356
             *:.   .::*.  .:.  .  .:.:::**:..**:  ::   .,  :  *       : :..    :  :  .*.  ::  . .* * :: :  :::*  . :**:  :**  .:*:
```

```
                                                     3                                        ⇓        ↓        ⇓
DmAce   FIDYFDKDDATALPRDKYLEIMNNIFGKATQAEREAIIFQYTSWE---GNPGYQNQQQIGRAVGDHFFTPTNEYAQALAERGASVHYYYFTHRTSTSLWGEWMGVLHGDEIEYFFGQPLNNSLQYRPVERELGKR  546
AgAce2  FIDYFEKDAATSLPRDKFLEIMNTIPNKASEPEREAIIFQYTGWES---GNDGYQNQHQVGRAVGDHFFIQTNEFALGLTERGASVHYYYFTHRTSTSLWGEWMGVLHGDEIEYVFGPMNASLQYRQRERDLSRR  534
TcAce2  FIDFFEKDGPSFLQRDKYHDIIDTIFKNMSRLERDAIVFQYTDWEH---VNDGYLNQKMIGDVVGDYFFVCPTNDFAELAAERGMKVYYYFTHRTSTSLWGEWMGVMHGDEIEYVFGHPLNMSLQFNSRERELSLK  530
LdAce2  FIDFFEKDGPSFLQREKYHDIIDTIFKNMSRLERDAIVFQYTNWEH---VHDGYLNQKMIGDVVGDYFFVCPTNNFAEVAADRGMKVFYYFTHRTSTSLWGEWMGVIHGDEVEYVFGHPLNMSLQFNSRERELSLK  524
HaAce2  FLDYFEKDGPSFLQREKFLEIVDTIFKDFSKIKREAIVFQYTDWEE---ITDGYLNQKMIADVVGDYFFVCPTNVFAEVLADSGVDVYYYFTHRTSTSLWGEWMGVMHGDEMEYVFGHPLNMSLQYHTRERDLAAH  546
BgAce2  FMDYFEKDSPTFLQRDKYLDIVNLIFKNMTRLERDAIVFQYTDWEH---LADGYKNQRMIADVVGDYFFIQPTNLFAEQFAEHGTKVYYYYFTQRTSLNLWGQWMGVMHGDEIEYVFGHPLNMSLIDYNDNERDLSLR  537
SaAce2  FVDFFDRTSATALPKEKFVQIVNNIFKDRTQLERDAIIYQYSGWEKKEVDDKYSNQKQLSDVVADYFFVCPTNLFANIVSSRGARVYYYFFTHRTDSHLWGDWMGVLHGDEMQYVFGHPLNMSMPYNARERDLSIR  549
TcAce1  LTELFRKEENVYVNKQEFLRAVTELNPYFNAISRQAIVFEYTNWLN--PDDPVSNRDALDKMVGDYHFTNVNEFAHRYAETGNTVYMYYYKHRTVANPWPSWTGVMHADEINYVFGEPLNPKTKSHTAQEVDLSKR  535
BgAce1  LTELFRKEENVYVNREEFLRSVQELNPYVNNVARQAIVFEYTDWLN--PDDPITRNRDALDKMVGDYHFTNVNEFAHRYAETGNNVYMYYYKHRTVADPWPSWTGVMHGDEINYVFGEPLNPAKNYQPQEIELSRR  578
HaAce1  LTELFPKEENVGISREQYLQAVRELNPYVVNDAGRQAIVFEYTDWLN--PEDPIKNRNALDKMVGDYHFTCGVNEFAHRYAETGNNVFTYHYKHRSKNNPWPSWTGVMHADEINYVFGEPLNPGKNYSPEEVEFSKR  582
SaAce1  LTELFKKEENVVVSRENFIKAIGQLNPNADAAVKSAIEFEYTDWFS--PNDPEKNRNALDKMVGDYQFTCNVNEFAHKYALTGNNVYMYYFKHRSLNNPWPKWTGVMHGDEISYVFGDPLNPNKRYEIEEIELSKS  569
SgAce1  LTELFKKEENVMVSRENFIKAIGQLNPNADAAVKSAIEFEYTDWFS--PNDPEKNRNALDKMVGDYQFTCNVNEFAHKYALTGNNVYMYYFKHRSLNNPWPKWTGVMHGDEISYVFGDPLNPNKRYEIEEIELSKS  569
AgAce1  LTELLRKEEGVTVTREEFLQAVRELNPYVNGAARQAIVFEYTDWTE--PDNPNSNRDALDKMVGDYHFTCNVNEFAQRYAEEGNNVYMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNPTLGYTEDEKDFSRK  628
TcaAce  APG-FSKDSESKISREDFMSGVKLSVPHANDLGLDATVLQYTDWMD--DNNGIKNRDGLDDIVGDHNVLPLMHFVNKYTKFGNGTYLYFFNHRASNLVWPEWMGVIHGYEIEFVFGLPLVKELNYTAEEEALSRR  489
             . :   .    ..  .:      .  :::   .**:. ::      *   *  ..  *.  *::   .   .  .* *. :  .    .*:** *   ::* *      :*** :  .  :  .  :  :
```

```
                                              3
DmAce   MLSAVIEFAKTGNPAQD------GEEWPNFSKEDPVYYIFSTDD----KIEKLARGPLAARTSFWNDYLPKVR---------SWAG----TCDGDSGSASISPRLQLLGIAALIYICAALRTKRVF  649
AgAce2  MVLSVSEFARTGNPALE------GEIHWPLYTRENPIYFIFNAEGEDDLRGEKYGRGPMATSCAFWNDFLPLRLR--------AWSVPLKDPCKLDDHTSIASTARAAPTVALLIALSLAVARLVAA  645
TcAce2  IMQAFARFAATGKPVTD------DVNWPLYTKDQPQYFIFNAD----KNGIG--KGPRATACAFWNDFLPKLR------DNPGK----FN----RCSLHH  604
LdAce2  IMQAFARFATTGKPVTD------DVNNPLYTKDQPQYFIFNAD----KNGIG--KGPRATACAFWNDFLPKLR------DNSGSEEAPCVNTYLSKIRSSSNELLPPSTSLVLIWIMTLLNAL  629
HaAce2  IMQSFTRFALTGKPHKP------DEKWPLYSRTSPHYYTYTADG---TSGPAGPRGPRASACAFWNDFLNKLN------ELEHIMPCDGAVT--------GPYSSVAGTTLPIVLLTTLATTVAL  647
BgAce2  IMEIYSRFALTGKPIAN------EADWPTYTKDHPKYYILNSD----TYGTG--KGPRTTNCAFWNDFLPKLANHP-----DPDYACASGVAETANLGVNSATSCLLVGPKHLLYFVTIVLARLLLSL  648
SaAce2  IMEAFTRFSLTGTPVSD------DIDWPLYNESNPIYHVWNAA-----ELIHVGYGPRAAECAFWNQ IAQALRETTKITCEDYPDSMPTTNENCTFTSSFATINPRISFTIIFIFVLPAHGLF  664
TcAce1  IMRYWANFAKTGNPSQSPNGVWTPTFWPPHTAFGREFLTLDVN-----STATGRGPRLKQCAFWKKYLPQLQ--------QQTSELLNQPPRQNCTDAASSLRWSRDGAAGLLMVSTVAALLAGPF---  648
BgAce1  MMRYWANFAKTGNPSMSEDGTWTATYWPVHTAYGREYLTLDVN-----STETGRGPRLKQCAFWKKYLPQLI--------AVTSNLNQQP---ESCPTDTSGAEMNKESSLLLICLVMKIILLWGPRGWV  692
HaAce1  LMRYWANFARSGNPSINPSGDSTKIHWPVHTASGREYLSLAVN-----SSSIGHGLRVKECAFWQKYLPLQN--------ATTSKILLELFIKNL  664
SaAce1  MMRYWTNFAKTGNPSKTFEGSWVTPKWPVHTAYGKEFLTLDTN-----NTSIGVGPRLEQCAFWKNYVPDLT--------AISKSMKSDK---NCTTISGGTKTNMIKLSLWTIVMTTAVLVL-  676
SgAce1  MMRYWTNFAKTGNPSKTLEGSWVTPKWPVHTAYGKEFLTLDTN-----NTSIGVGPRLEQCAFWKNYVPDLT--------AISKSMKSDK---NCTTISGGTKTNVIELSVWTIVMTTAVLML-  676
AgAce1  IMRYWSNFAKTGNPN--TASSEFPWPKHTARDHRHYLELGLN------TSFVGRGPRLRQCAFWKEYLPQLT--------AATSNLPGPAP-PSEPCESSAFFYRPDLIVLLVSLLTATVRFIQ-  737
TcaAce  IMHYWATFAKTGNPNEP------HSQESKWPLFTTKEQKFIDLNTE----PMKVHQRLRVQMCVFWNQFLPKLL--------NATETIDEAERQWKTEFIHRWSSYMMHWKNQFDHYSRHESCAEL  596
             :    *:.**:* *         . *      * . :      :  :        ***  .:. :              * **:  :.  :
```

during the last 1-ns period of all 21 simulations or in an average of each cluster of the trajectories generated by a cluster analysis. As to the loop conformation of residues 145–162, the cluster analysis showed that 24% of the trajectories had residues 145–162 folded in contact with surface residues such as R21, D148, D501, and D203.

A third set of 22 10-ns simulations of TcAce2 with residues 145–162 adopting the folded conformation was then carried out. The initial conformation of the third set of simulations was obtained from averaging all the trajectories of TcAce2 with the folded loop conformation of residues 145–162 followed by manual adjustment of the side-chain torsions to restore the hydrogen bond network of the catalytic triad. A cluster analysis of all the trajectories saved at 1.0-ps intervals during the last 1.0-ns period of all 22 simulations showed that residues 145–162 remain the folded conformation. However, the average of all the trajectories had a distorted catalytic triad. Visual inspection of all 22 simulations found that conformations of the first of the 22 simulations have a catalytic triad engaging in a hydrogen-bond network. The final model of TcAce2 was then obtained from averaging all trajectories saved at 1.0-ps intervals during the last 1.0 ns period of the first simulation.

Although the homology model of TcAce2 was based on the DmAChE crystal structure, the active site of the simulation-refined TcAce2 model (Fig. 7) is very different from that of the DmAChE crystal structure. It is also different from those of human AChE and insect AP-AChEs (*i.e.*, AChE1s). In the refined model of TcAce2 (Fig. 7), Y114, Y345, Y395, and W342 form an aromatic cluster that completely block the entrance of the active site; E388, H502, and S259 form a catalytic triad; ACh has a cation-pi interaction with W126, but it does not adopt the fully extended conformation, nor is its carbonyl oxygen atom located in the oxyanion hole. In contrast to the TcAce1 model with residues 146–154 and 493–509 that partially shield ACh from interacting with solvent, the TcAce2 model has its corresponding residues (120–128 and 488–504) adopt conformations that leave ACh to be exposed to solvent (Fig. 8).
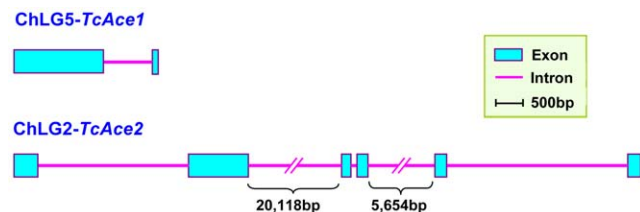
Figure 3. Schematic diagram of the organization of two *Ace* **genes from** *Tribolium castaneum.* The full lengths of the two genomic sequences were 2,986 bp for *TcAce1* and 32,243 bp for *TcAce2*. Genome structure showed that two different *Ace* genes in *T. castaneum* located on different chromosomes. *TcAce1* located on chromosome 5 and *TcAce2* on chromosome 2. *TcAce1* has two exons and one intron, whereas *TcAce2* has six exons and five introns.
doi:10.1371/journal.pone.0032288.g003

TcAce gene expression profiles

The transcript levels of *TcAce1* and *TcAce2* were evaluated by reverse transcription PCR (RT-PCR) and quantitative real-time PCR (qPCR) in tissues of *T. castaneum* at different developmental stages (Fig. 9). Both *TcAce* genes were transcribed in all the stages examined, including 1-day (d) and 3-d eggs; 5-d and 20-d larvae; 1-d, 3-d and 6-d pupae; and 2-d and 14-d adults. The lowest expression levels of these genes were found in eggs, particularly for *TcAce2* whose expression level was undetectable by RT-PCR in 1-d eggs (Fig. 9A and 9B). The expression patterns of *TcAce1* and *TcAce2* were very similar. In addition, the *TcAce1* and *TcAce2* genes also exhibited similar tissue-specific expression patterns (Fig. 9C and 9D). As expected, these genes were predominately expressed in the brain, although their expressions were also detected in the gut and carcass after the brain and ventral nerve cord were removed.

Discussion

Since the first insect AChE orthologous gene (*i.e.*, the one later named *Ace2* or AO-Ace) and the first insect AChE paralogous gene (*i.e.*, the one later named *Ace1* or AP-Ace) were reported in *D. melanogaster* in 1986 [9] and in *S. graminum* in 2002 [10], respectively, cDNAs encoding both *Ace1* and *Ace2* have been sequenced from each of at least 27 insect species. However, the gene models and the genomic organizations of *Ace1* and *Ace2* have not been well established in insects. In this study, we confirmed the two *Ace* gene (*TcAce1* and *TcAce2*) models in *T. castaneum* by sequencing the coding regions of their cDNAs followed by comparative analyses of their cDNA and genomic sequences. The *TcAce1* and *TcAce2* genes are significantly different not only in the length of their genomic DNA (*TcAce1* with 2,986 bp and *TcAce2* with 32,243 bp) but also in the intron/exon organizations. Specifically, *TcAce1* possesses only one intron whereas *TcAce2* has five introns. Furthermore, *TcAce1* is located on chromosome 5 (ChLG5), whereas *TcAce2* is on chromosome 2 (ChLG2; Fig. 3). Apparently, the intron/exon organizations and the chromosomal locations of these genes in *T. castaneum* are different from their counterparts in other insect species [8,14,18,23,45].

Despite the significant differences in genomic structures and the chromosomal locations of *TcAce1* and *TcAce2*, the deduced protein sequences of the two genes exhibit all the common features of an AChE sequence wise. These features include (1) a conserved active-site triad, including S267 in TcAce1 and S259 in TcAce2 (S200 in *Torpedo*), E393 in TcAce1 and E388 in TcAce2 (E327 in *Torpedo*), and H507 in TcAce1 and H502 in TcAce2 (H440 in *Torpedo*); (2) a choline binding site, W152 in TcAce1 and W126 in TcAce2 (W84 in *Torpedo*); (3) three pairs of cysteines putatively forming intramolecular disulfide bonding (C135~C162, C321~C334, and C469~C591 in TcAce1; and C109~C136, C313~C328, and C464~C580 in TcAce2); (4) a cysteine forming intermolecular disulfide bonding (C617 in TcAce1 and C600 in TcAce2); (5) 10 conserved aromatic amino acid residues out of 14
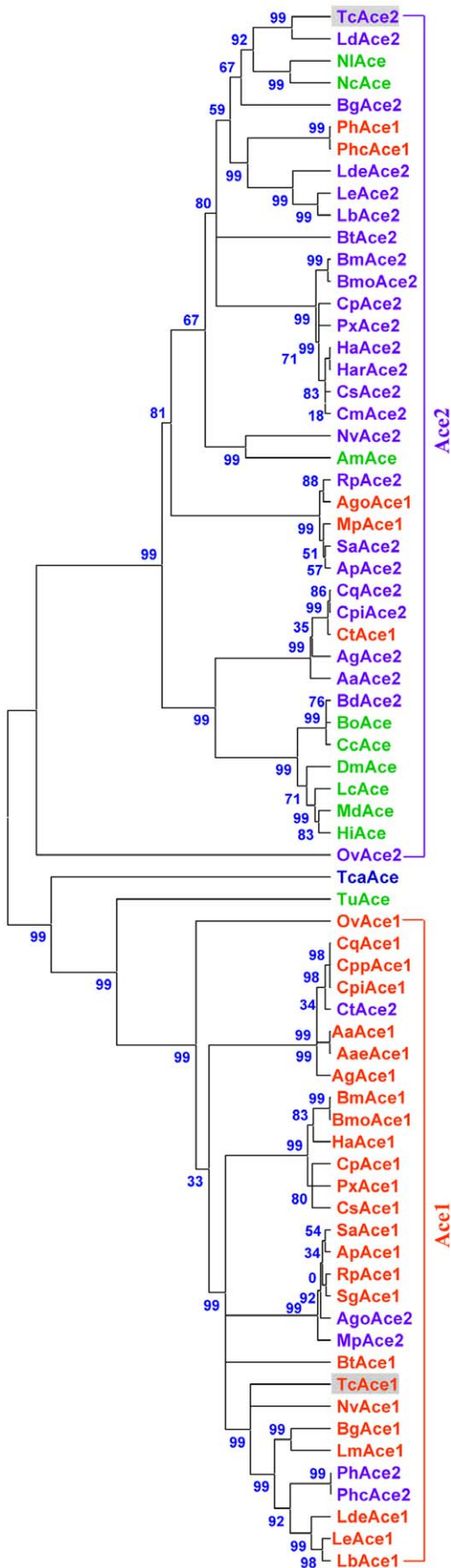
Figure 4. Rooted phylogenetic tree of deduced *Ace* amino acid sequences from the Pacific electric ray (*Torpedo californica*), two-spotted spider mite (*Tetranychus urticae*) and 43 insect species constructed by the neighbor-jointing method. The name is made up of a species abbreviation (first letter of the genus followed by the first one or two letters of the specific name). Sequences used: TcaAce (CAA27169, *T. californica* Ace); TcAce1 (HQ260968, *Tribolium castaneum Ace1*, this paper); TcAce2 (HQ260969, *T. castaneum Ace2*, this paper); BmAce1(EU262633, *Bombyx mandarina* Ace1); BmAce2 (EU262632, *B. mandarina* Ace2); SaAce1 (AY819704, *Sitobion avenae* Ace1); SaAce2 (AY707319, *S. avenae* Ace2); RpAce1 (AY667435, *Rhopalosiphum padi* Ace1); RpAce2 (AY707318, *R. padi* Ace2); DmAce (X05893, *Drosophila melanogaster* Ace); AgAce1 (XM_321792, *Anopheles gambiae* Ace1); AgAce2 (BN000067, *A. gambiae* Ace2); LdeAce1 (FJ647186, *Liposcelis decolor* Ace1); LdeAce2 (FJ647187, *L. decolor* Ace2); OvAce1 (FJ228227, *Orchesella villosa* Ace1); OvAce2 (FJ228228, *O. villosa* Ace2); LeAce1 (EU854149, *Liposcelis entomophila* Ace1); LeAce2 (EU854150, *L. entomophila* Ace2); BgAce1 (DQ288249, *Blattella germanica* Ace1); BgAce2 (DQ288847, *B. germanica* Ace2); BtAce1 (EF675188, *Bemisia tabaci* Ace1); BtAce2 (EF675190, *B. tabaci* Ace2); CqAce1 (XM_001847396, *Culex quinquefasciatus* Ace1); CqAce2 (XM_001842175, *C. quinquefasciatus* Ace2); BmoAce1 (NP_001037380 *Bombyx mori* Ace1); BmoAce2 (NP_001108113 *B. mori* Ace2); ApAce1 (XM_001948618, *Acyrthosiphon pisum* Ace1); ApAce2 (XM_001948953, *A. pisum* Ace2); NvAce1 (XM_001600408, *Nasonia vitripennis* Ace1); NvAce2 (XM_001605518, *N. vitripennis* Ace2); PhAce1 (AB266605, *Pediculus humanus corporis* Ace1); PhAce2 (AB266606, *P. humanus corporis* Ace2); CpAce1 (DQ267977, *Cydia pomonella* Ace1); CpAce2 (DQ267976, *C. pomonella* Ace2); HaAce1 (DQ001323, *Helicoverpa assulta* Ace1); HaAce2 (AY817736, *H. assulta* Ace2); AaAce1 (AB218421, *Aedes albopictus* Ace1); AaAce2 (AB218420, *A. albopictus* Ace2); AgoAce1 (AF502081, *Aphis gossypii* Ace1); AgoAce2 (AF502082, *A. gossypii* Ace2); CtAce1 (AB122151, *Culex tritaeniorhynchus* Ace1); CtAce2 (AB122152, *C. tritaeniorhynchus* Ace2); SgAce1 (AF321574, *Schizaphis graminum* Ace1); MpAce1 (AF287291, *Myzus persicae* Ace1); MpAce2 (AY147797, *M. persicae* Ace2); CpiAce1 (AJ489456, *Culex pipiens* Ace1); CpiAce2 (AM159193, *C. pipiens* Ace2); MdAce (AY134873, *Musca domestica* Ace); PxAce1 (AY970293, *Plutella xylostella* Ace1); PxAce2 (AY061975, *P. xylostella* Ace2); CsAce1 (EF453724, *Chilo suppressalis* Ace1); CsAce2 (EF470245, *C. suppressalis* Ace2); CmAce2 (FN538987, *Cnaphalocrocis medinalis* Ace2); PhcAce1 (AB266614, *Pediculus humanus capitis* Ace1); PhcAce2 (AB266615, *P. humanus capitis* Ace2); TuAce (AY188448, *T. urticae* Ace); AaeAce1 (EF209048, *Aedes aegypti* Ace1); LmAce1 (EU231603, *Locusta migratoria manilensis* Ace1); LbAce1 (FJ647185, *Liposcelis bostrychophila* Ace1); LbAce2 (EF362950, *L. bostrychophila* Ace2); NlAce (FM866396, *Nilaparvata lugens* Ace); CcAce (EU130781, *Ceratitis captitata* Ace); AmAce (AB181702, *Apis mellifera* Ace); BdAce2 (AY155500, *Bactrocera dorsalis* Ace2); HiAce (AY466160, *Haematobia irritans* Ace); CppAce1 (AY762905, *Culex pipiens pallens* Ace1); HarAce2 (AF369793, *Helicoverpa armigera* Ace2); BoAce (AF452052, *Bactrocera oleae* Ace); LcAce (U88631, *Lucilia cuprina* Ace); NcAce (AF145235, *Nephotettix cincticeps* Ace); LdAce2 (L41180, *Leptinotarsa decemlineata* Ace2).
doi:10.1371/journal.pone.0032288.g004

aromatic residues lining the catalytic gorge of AChE; and (6) the conserved sequence FGESAG, flanking S267 in TcAce1 and S259 in TcAce2 (Fig. 2).

Ethanolmine and glucosamine residues are characteristic of a C-terminal glycolipid anchor in most G2 AChE [46,47]. Our analysis by using PredGPI predictor suggests that DmAChE and TcAce1 contain a GPI-anchor at the C-terminal. Although TcAce2 has higher sequence identity level with DmAChE than TcAce1, TcAce2 does not appear to contain a GPI-anchor at the C-terminal. Because the C-terminal Cys residue in AChE is reportedly for the intermolecular disulfide linkage [48], C617 in TcAce1 and C600 in TcAce2 are likely involved in the intermolecular disulfide linkage, although the corresponding Cys residue is missing in DmAChE [46]. We also analyzed the hydrophilic and hydrophobic of non-homologous amino acid residues of TcAce1 and TcAce2 and compared homologous

Table 2. Percent identities of amino acid residues among the AChEs of *Tribolium castaneum*, *Torpedo californica* and other seven insect species.

Name	Tca	TcAce1	SaAce1	AgAce1	BgAce1	SgAce1	HaAce1	DmAce	TcAce2	SaAce2	AgAce2	BgAce2	LdAce2	HaAce2
Tca	—	39	38	42	40	38	39	35	38	37	37	38	37	37
TcAce1		—	57	61	70	57	64	31	36	31	34	36	34	34
SaAce1			—	51	59	96	55	32	36	31	33	34	33	34
AgAce1				—	57	51	59	34	37	33	35	39	35	36
BgAce1					—	59	62	33	36	33	33	35	36	34
SgAce1						—	55	32	36	30	33	34	34	33
HaAce1							—	31	35	30	32	35	34	32
DmAce								—	55	48	62	51	53	51
TcAce2									—	57	60	68	84	69
SaAce2										—	50	53	55	52
AgAce2											—	55	57	56
BgAce2												—	63	60
LdAce2													—	66
HaAce2														—

NOTE: TcaAce, *Torpedo californica*; TcAce, *Tribolium castaneum*; SaAce, *Sitobion avenae*; DmAce, *Drosophila melanogaster*; AgAce, *Anopheles gambiae*; BgAce, *Blattella germanica*; HaAce, *Helicoverpa assulta*; SgAce, *Schizaphis graminum*; LdAce, *Leptinotarsa decemlineata*.
doi:10.1371/journal.pone.0032288.t002

sequences of TcAce1 and TcAce2 with mAChE and DmAChE at the C-terminal for possible formation of the four-helix bundle. Our analysis suggests that C-terminal sequences of both TcAce1 and TcAce2 may form the dimeric four-helix bundle.

According to comparisons of TcAce1 and TcAce2 with Ace proteins from other insect species, TcAce1 showed high sequence identities to the AChE1 (AP-AChE) proteins from *Sitobion avenae* (57%), *A. gambiae* (61%), *Blattella germanica* (70%), *S. graminum* (57%), and *Helicoverpa assulta* (64%; Table 2). Similarly, TcAce2 also showed high amino acid identities to the AChE2 (AO-AChE) proteins of *S. avenae* (57%), *D. melanogaster* (55%), *A. gambiae* (60%), *B. germanica* (68%), *Leptinotarsa decemlineata* (84%), and *H. assulta* (69%). Based on our analysis of deduced amino acid sequences from the two genes among respective insect species, the sequence identity levels within the paralogous (*Ace1*) or orthologous (*Ace2*) genes range from 48 to 96%, whereas the sequence identity levels between the *Ace1* and *Ace2* genes of the same insect species range only from 31 to 36% in all insect species examined (Table 2). These results support the hypothesis that the two *Ace* genes were originated from an old duplication before the diversification of insect species [32].

Furthermore, a phylogenetic tree, which was generated from the highly conserved regions of all insect and *T. urticae* AChE

Figure 5. Close-up view of the active site of TcAce1 with a perspective from the free cysteine at the opening of the active-site gorge down to ACh and the catalytic triad at the bottom of the gorge.
doi:10.1371/journal.pone.0032288.g005
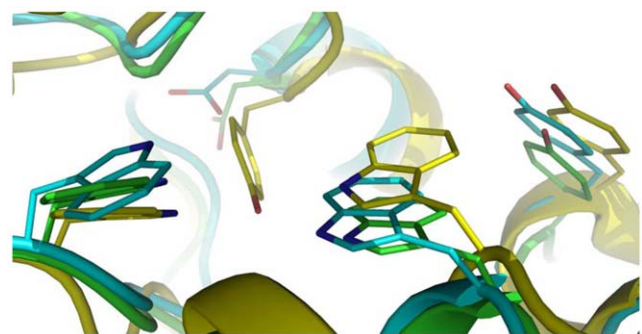
Figure 6. Comparison of the bottom of the active-site gorge in TcAce1 to those in human AChE and *Anopheles gambiae* AP-AChE. Tyr449 in human AChE (yellow) is mutated to Asp509 in TcAce1 (green) and Asp441 in *A. gambiae* AChE (cyan).
doi:10.1371/journal.pone.0032288.g006
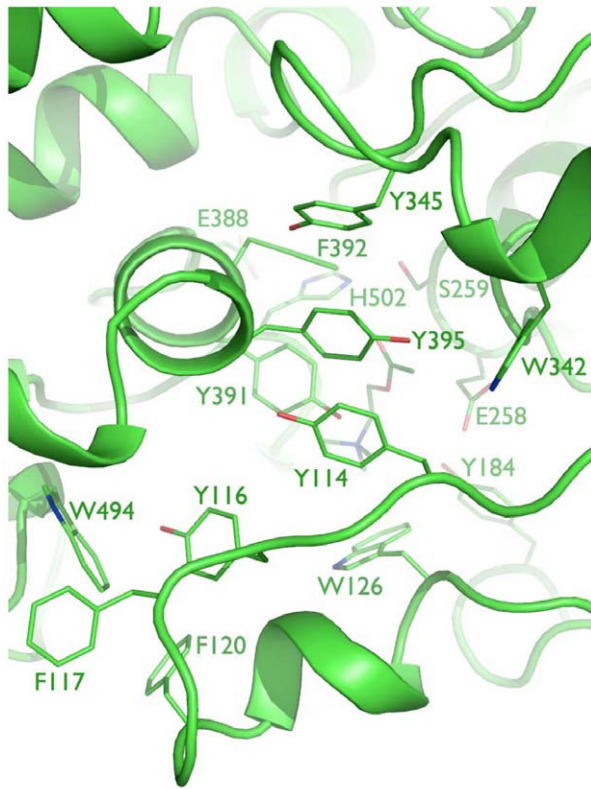
Figure 7. Close-up view of the active site of TcAce2 with a perspective looking down to acetylcholine and the catalytic triad at the bottom of the gorge.
doi:10.1371/journal.pone.0032288.g007

Figure 8. Comparison of loop conformations of residues 146–154 and 493–509 in TcAce1 with the corresponding ones of residues 120–128 and 488–504 in TcAce2.
doi:10.1371/journal.pone.0032288.g008

amino acid sequences available in GenBank and the corresponding one in *T. californica* using the neighbor-joining method, revealed two insect AChE clusters. *TcAce1* was grouped into the insect *Ace1* cluster and *TcAce2* into the insect *Ace2* cluster. The significant divergence between *TcAce1* and *TcAce2* suggests that these genes were evolved from their corresponding *Ace* gene lineages in insect species [49]. These results also suggest, for the first time, that the divergence of *Ace1* and *Ace2* might occur prior to insect speciation and that the *Ace1* gene might be lost in *D. melanogaster* and other species in Cyclorrapha suborder of Diptera during the evolutionary process [14,25,32]. Thus, it is likely that both the *Ace* genes in insects may have different functions because these genes have evolved during the evolutionary histories of these insect species.

Using a reported simulation-refined model of *A. gambiae* AP-AChE [33] and a crystal structure of DmAChE [41] with high sequence identities to TcAce1 (73%) and TcAce2 (60%), respectively, and the same multiple molecular dynamics simulation method to model and refine TcAce1 and TcAce2, we obtained models of TcAce1 and TcAce2 both which are in complex with acetylcholine. The TcAce1 model has an active site that is almost identical to those of *A. gambiae* and *S. graminum* AP-AChEs, and it has C354 at the opening of the active-site gorge just like the insect-specific C286 of *A. gambiae* and C289 of *S. graminum* that are susceptible to sulfhydryl agents [33,34]. The TcAce2 model has an active site with an entrance comprised of G122, W126, W494, M501, and D504. This entrance corresponds to the small opening at the bottom of the active-site gorge

of DmAChE [44] or *T. californica* AChE when Y442 moves away from W84 [43]. In the TcAce2 model, the region that corresponds to the entrance of TcAce1 is completely blocked by Y114, Y345, Y395, and W342. Of the four aromatic residues, Y395 and W342 correspond to Y334 and W279 of *T. californica* AChE, respectively, and belong to the 14 conserved aromatic residues that line the active-site gorge of *T. californica* AChE [50]. In other words, the entrance of the active-site gorge of TcAce2 appears to be reversed relative to that of TcAce1. Unlike ACh in the TcAce1 model, ACh does not adopt the fully extended conformation and its carbonyl oxygen atom is not placed in the oxyanion hole in the TcAce2 model.

In addition, R576 in the TcAce2 model is close to E388 (the separation between the side-chain N atom of R576 and the side-chain O atom of E388 is 3.9 Å), a composite residue of the catalytic triad, in contrast to the corresponding arginine residue that is away from the catalytic glutamate residue in TcAce1, *A. gambiae* AP-AChE and human AChE (Fig. 10). Analysis of all the trajectories of the third set of simulations of TcAce2 showed that 95% of the trajectories has a hydrogen bond between R576 and E388, accompanied by hydrogen bonds between H502 and E505 and between S259 and E258, leading to disruption of the catalytic triad. These computational observations suggest that TcAce1 is a robust ACh hydrolase and susceptible to sulfhydryl agents and that TcAce2 is not a catalytically efficient ACh hydrolase, although further study is needed to comprehensively elucidate physiological functions of *Ace1* and *Ace2* genes. In view of these computational results, it is logical to investigate whether TcAce2 functions more as a cholinesterase-like adhesion molecule (CLAM) [51,52] than TcAce1. However, our sequence analysis using ClustalW 2.0.12 shows that TcAce1 has a slightly higher sequence homology to *D. melanogaster* gliotactin, which is one of the three *D. melanogaster* CLAMs, than TcAce2 (Table 3); both TcAce1 and TcAce2 have dipole moments that are comparable to those of other AChEs (Table 4). The orientations of the dipole moments of TcAce1 and TcAce2 are almost the same. The two dipole moments are approximately along the beta strand that corresponds to Strand 5 of the *T. californica* AChE crystal structure [53] and nearly identical to those of other AChEs but orthogonal to that of *Galactomyces geotrichum* lipase [52,54]. These sequence and dipole moment analyses do not support the hypothesis that TcAce2 functions as a CLAM.
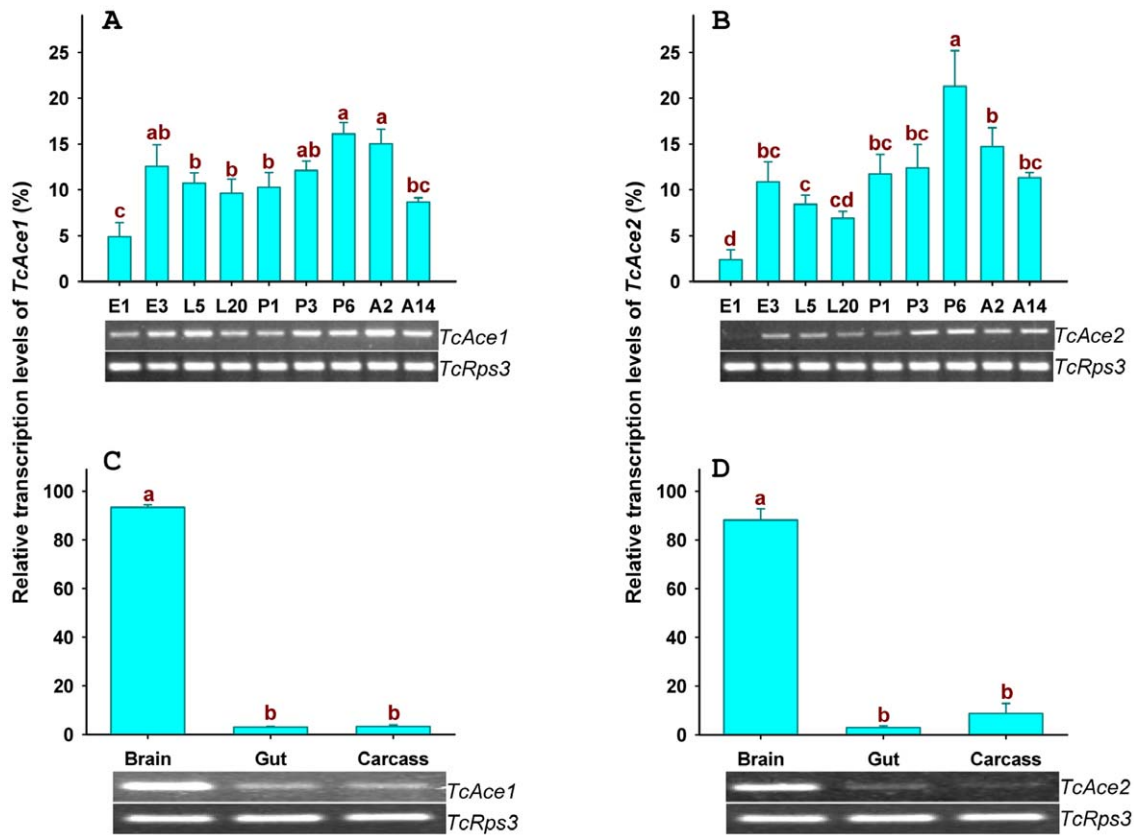
Figure 9. The two *TcAce* genes mRNA expression levels of developmental stages and different tissues were determined by RT-PCR as shown by gel pictures at the bottom of each panel and quantitative real-time PCR as shown by histograms. Real-time PCR data were normalized to *TcRps3* gene expression. E1, 1-d eggs; E3, 3-d eggs; L5, 5-d larvae; L20, 20-day larvae; P1, 1-d pupae; P3, 3-d pupae; P6, 6-d pupae; A2, 2-d adults; A14, 14-d adults. Gut (including midgut and hindgut), Carcass (not including head, gut and nerve system). Standard error bars were base on three replicates. One-Way ANOVA-Fisher's LSD was used in statistical analysis of quantitative real-time PCR data.
doi:10.1371/journal.pone.0032288.g009
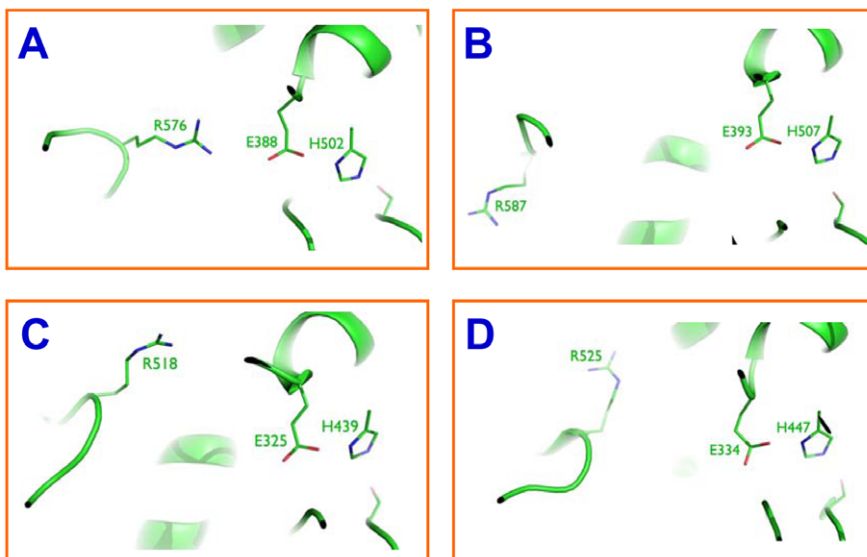
Figure 10. Separation of Arg from the catalytic triad in TcAce2 (A), TcAce1 (B), *Anopheles gambiae* AP-AChE (C), and human AChE (D).
doi:10.1371/journal.pone.0032288.g010

Table 3. Percentages of amino acid sequence identity of TcAce1 and TcAce2 to cholinesterase-like lipase and adhesion proteins.

AChE	*Gc*NeutralLipase	*Dm*Neuroligin	*Dm*Neurotactin	*Dm*Gliotactin
Torpedo californica AChE	23	19	17	25
TcAChE1	19	21	17	27
TcAChE2	20	20	18	25

The NCBI session numbers of *Gc*NeutralLipase, *Dm*Neuroligin, *Dm*Neurotactin, and *Dm*Gliotactin are P79066, AAF52450, CAA37831, and AAC41579, respectively.
doi:10.1371/journal.pone.0032288.t003

Materials and Methods

Insect culture

The Georgia-1 (GA-1) strain of *T. castaneum* was reared on whole-wheat flour containing 5% (by weight) of brewers' yeast at 30°C and 65% relative humidity under standard conditions in the laboratory of Kansas State University (Manhattan, Kansas, United States of America) based on the method of Haliscak and Beeman [55].

Total RNA isolation and reverse transcription

Total RNA was isolated from *T. castaneum* samples using TRIzol reagent following the recommended procedure by Invitrogen (Carlsbad, California, United States of America). The RNA was treated with DNase I (Fermentas, Glen Burnie, Maryland, United States of America) according to the manufacturer's instruction and the first-strand cDNA template was synthesized from 3.0 μg of total RNA by using First Strand cDNA Synthesis Kit (Fermentas) with oligo $(dT)_{18}$ as the primer.

Subcloning and sequencing of cDNA

To obtain the cDNAs corresponding to the entire protein coding regions of *TcAce1* and *TcAce2*, we designed specific primers based on *TcAce* gene predictions and their genomic organization (Table 1). The PCR products of each reaction were subjected to electrophoresis on 1% agarose gel containing ethidium bromide. The PCR bands were excised and purified using QIAEX II Agarose Gel Extraction Kit (Qiagen, Valencia, California, United States of America). The purified fragment was subcloned into a pGEM-T Easy Vector (Invitrogen) according to the manufacturer's instruction. The ligation DNA mixtures were used to transform bacterial cells by using Z-Competent *E. coli* Transfor-

mation Kit and Buffer Set™ (Zymo Research Corporation, Irvine, California, United States of America). Plasmids were isolated from the bacterial cells and used for DNA sequencing (KSU DNA Sequencing and Genotyping Facility, Manhattan, Kansas, United States of America). Signal P software was used to predict signal peptide [56].

Phylogenetic analysis of AChEs

ClustalW software (www.ebi.ac.uk/clustalw/) [57], was used to perform multiple sequence alignments prior to phylogenetic analysis. Phylogenetic analysis was done using MEGA 4.0 [58] for construction a neighbor-joining tree to examine the evolutionary relationships among *T. californica*, *T. urticae* and 43 insect species. To evaluate the branch strength of the phylogenetic tree, a bootstrap analysis of 1000 replications was performed.

AChE model prediction

Model preparation. The starting conformations of TcAce1 and TcAce2 used in the multiple molecular dynamics simulations were generated by the SWISSMODEL homology program using a computer model (Protein Data Bank ID: 2AZG [33]) and a crystal structure (Protein Data Bank ID: 1DX4 [41]) as templates, respectively. ACh was manually docked into the active site of TcAce1 or TcAce2 according to the bound ACh conformation in the crystal structure of *T. californica* AChE (Protein Data Bank ID: 2ACE [59]). All His, Glu, Asp, Arg, and Lys residues of the ACh-bound TcAce1 or TcAce2 were treated as HIP, GLU, ASP, ARG, and LYS, respectively. The topology and coordinate files were generated by the PREP, LINK, EDIT, and PARM modules of the AMBER 5.0 program [60]. The complex was refined by energy minimization using the SANDER module of the AMBER 5.0 program with a dielectric constant of 1.0 and 500 cycles of

Table 4. The electrostatic characteristics of TcAce1, TcAce2, and other proteins.

PDB ID (Species)	Total atoms	Total residues	Net charge of N terminal deletion (*e*)	Net charge of C terminal deletion (*e*)	Net charge of deletion between N and C termini (*e*)	Net charge of the structure or model (*e*)	Dipole moment (Debye)	Dipole moment per atom (Debye)
TcAce1 (*Tribolium castaneum*, AP)	4201	529	−5	0	0	−3	1490	0.35
TcAce2 (*Tribolium castaneum*, AO)	4353	545	−1	+1	0	−19	1163	0.27
1QO9 (*Drosophila melanogaster*, AO)	4273	540	0	−2	+1	−18	669	0.16
2AZG (*Anopheles gambiae*, AP)	4243	536	0	−1	0	−8	1718	0.40
2HCP (*Schizaphis graminum*, AP)	4302	540	−7	+2	0	−8	1312	0.30
2ACE (*Torpedo californica*)	4143	527	−2	−1	−1	−8	1819	0.44
1J06 (*Mus musculus*)	4177	535	+2	−2	0	−9	867	0.21
2X8B (*Homo sapiens*)	4179	536	0	−2	0	−10	1384	0.33
1THG (*Galactomyces geotrichum*, lipase)	4287	543	+1	0	0	−15	814	0.19

doi:10.1371/journal.pone.0032288.t004

steepest-descent minimization followed by 10,000 cycles of conjugate-gradient minimization. The energy-minimized ACh complex with TcAce1 or TcAce2 was solvated with 5,897 or 6,744 TIP3P water molecules [61], leading to a system of 20,703 or 23,244 atoms, respectively. The water molecules were obtained from solvating the complex using a pre-equilibrated box of 216,000 TIP3P molecules, whose hydrogen atom charge was set to 0.4170, where any water molecule was removed if it had an oxygen atom closer than 2.2 Å to any solute atom or a hydrogen atom closer than 2.0 Å to any solute atom, or if it was located further than 10.0 Å along the x-, y-, or z-axis from any solute atom.

Multiple molecular dynamics simulations. The solvated protein complex was energy-minimized for 100 cycles of steepest-descent minimization followed by 100 cycles of conjugate-gradient minimization to remove close van der Waals contacts in the system, then heated from 0 to 300 K at a rate of 10 K/ps under constant temperature and volume, and finally simulated independently with a unique seed number for initial velocities at 300 K under constant temperature and pressure using the PMEMD module of the AMBER 8.0 program [62] with the AMBER force field (ff99SB) [63,64]. All simulations used (1) a dielectric constant of 1.0, (2) the Berendsen coupling algorithm [65], (3) a periodic boundary condition at a constant temperature of 300 K and a constant pressure of 1 atm with isotropic molecule-based scaling, (4) the Particle Mesh Ewald method to calculate long-range electrostatic interactions [66], (5) a time step of 1.0 fs, (6) the SHAKE-bond-length constraints applied to all the bonds involving the H atom, (7) saving the image closest to the middle of the ''primary box'' to the restart and trajectory files, (8) formatted restart file, and (9) default values of all other inputs of the PMEMD module. All simulations were performed on a cluster of Apple Mac Pros with 80 Intel Xeon cores (3.0 GHz) and a cluster of Apple Xserves with 590 G5 processors (2.2/2.3 GHz).

Simulation analysis. Average structures were obtained by using the CARNAL module of AMBER 5.0. Cluster analyses were performed by using the PTRAJ module [67] of AMBER 10.

Dipole moment calculations. All dipole moment calculations were performed using the Protein Dipole Moments Server (http://bioinfo.weizmann.ac.il/dipol/indexj.html) [68]. Protein structures with minimal deletions were used and ligands and structural water molecules were removed before the dipole moment calculations.

Analysis of expression of TcAce1 and TcAce2 by RT-PCR and qPCR

The expression patterns of both TcAce1 and TcAce2 genes were analyzed at various developmental stages including embryos (1 day and 3 days eggs), early larvae (5 days larvae), late larvae (20 days larvae), early pupae (1 day pupae), middle pupae (3 days larvae), late pupae (6 days pupae), early adults (2 days adults) and two weeks old adults. To analyze tissue specific expression, we collected samples from the following dissected late pupa tissues (pooled from thirty late pupae): brain, gut (midgut and hind gut) and carcass (the whole body excluding brain, ganglia and gut). For all the samples, 3.0 μg of total RNA were treated with DNase I (Fermentas) to remove any genomic DNA contaminations, and then used as templates for the first strand cDNA synthesis. The cDNAs prepared from total RNA were used as templates for amplification and detection of specific TcAce sequences. The gene-specific primers were designed by using the Beacon Designer 2.0 software (Premier Biosoft International, Palo Alto, California, United States of America) and are shown in Table 1. For reverse transcription PCR (RT-PCR), cDNA fragments of each TcAce were amplified using the PCR conditions as follows: 94°C for 1.5 min followed by 30 cycles (26 cycles for TcRps3 gene) of 94°C 30 s, 55°C 30 s and 72°C 45 s. A final extension at 72°C for 5 min was added at the end of the PCR. The relative mRNA expression of each TcAce was assessed by qRT-PCR using SYBR-Green in the Bio-Rad iCycler iQ^{TM} multi-coclor real-time PCR detection system (Bio-Rad Laboratories, Hercules, CA, USA) based on the method of Giulietti et al. [69]. All the experiments were performed in triplicate and normalized to the mRNA level of ribosomal protein S3 (Rps3) as a reference gene for each sample [70]. The relative mRNA expression levels were calculated according to the $2^{-\Delta\Delta Ct}$ method [71].

Statistical analysis

The data from the qPCR analysis were subjected to ANOVA followed by Fisher's least significant difference (LSD) multiple comparisons to separate the means among the treatments by using ProStat software (Poly Software International, Pearl River, New York, United States of America).

Author Contributions

Conceived and designed the experiments: YL YPP XG YP KYZ. Performed the experiments: YL JY XZ YPP KYZ. Analyzed the data: YL YPP YP JY XZ KYZ. Contributed reagents/materials/analysis tools: YP KYZ. Wrote the paper: YL YPP KYZ. Contributed with revisions: YL YPP XG YP JY XZ KYZ.

References

1. Taylor P, Radic Z (1994) The cholinesterases: from genes to proteins. Annu Rev Pharmacol Toxicol 34: 281–320.
2. Soderlund DM, Bloomquist JR (1990) Molecular mechanism of insecticide resistance, In: Pesticide resistance in Arthropods Roush RT, Tabashnik BE, eds. Chapman and Hall, New York. pp 58–96.
3. Zhu KY, Brindley WA (1992) Enzymological and inhibitory properties of acetylcholinesterase purified from Lygus hesperus Knight (Hemiptera: Miridae). Insect. Biochem Mol Biol 22: 245–251.
4. Fournier D, Mutero A (1994) Modification of acetylcholinesterase as a mechanism of resistance to insecticides. Comp Biochem Physiol 108C: 19–31.
5. Zhu KY, Clark JM (1995) Cloning and sequencing of a cDNA encoding acetylcholinesterase in Colorado potato beetle, Leptinotarsa decemlineata (Say). Insect Biochem Mol Biol 25: 1129–1138.
6. Zhu KY, Lee SH, Clark JM (1996) A point mutation of acetylcholinesterase associated with azinphosmethyl resistance and reduced fitness in Colorado potato beetle. Pestic Biochem Physiol 55: 100–108.
7. Guedes RNC, Zhu KY, Kambhampati S, Dover BA (1997) An altered acetylcholinesterase conferring negative cross-insensitivity to different insecticidal inhibitors in organophosphate resistant lesser grain borer, Rhyzopertha dominica. Pestic Biochem Physiol 58: 55–62.

8. Kono Y, Tomita T (2006) Amino acid substitutions conferring insecticide insensitivity in *Ace*-paralogous acetylcholinesterase. Pestic Biochem Physiol 85: 123–132.

9. Hall LMC, Spierer P (1986) The *Ace* locus of *Drosophila melanogaster*: structural gene for acetylcholinesterase with an unusual 5′ leader. Embo J 5: 2949–2954.

10. Gao J-R, Kambhampati S, Zhu KY (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.

11. Myers EW, Sutton GG, Delcher AL, Dew IM, Fasulo DP, et al. (2000) A whole-genome assembly of *Drosophila*. Science 5461: 2196–2204.

12. Pang Y-P, Brimijoin S, Ragsdale DW, Zhu KY, Suranyi R (2011) Novel and viable acetylcholinesterase target site for developing effective and environmentally safe insecticides. Curr Drug Targets;(in press).

13. Chen MH, Han ZJ (2006) Cloning and sequence analysis of 2 different acetylcholinesterase genes in *Rhopalosiphum padi* and *Sitobion avenae*. Genome 49: 239–243.

14. Weill M, Fort P, Berthomieu A, Dubois MP, Pasteur N, et al. (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 Lond Ser B 269: 2007–2016.

15. Nardi F, Barazzuoli B, Ciolfi S, Carapelli A, Dallai R, et al. (2009) Acetylcholinesterase genes in the basal Hexapod *Orchesella villosa*. Insect Mol Biol 18: 45–54.

16. Kim JI, Jung CS, Koh YH, Lee SH (2006) Molecular, biochemical and histochemical characterization of two acetylcholinesterase cDNAs from the German cockroach *Blattella germanica*. Insect Mol Biol 15: 513–522.

17. Alon M, Alon F, Nauen R, Morin S (2008) Organophosphates resistance in the B-biotype of *Bemisia tabaci* (Hemiptera: Aleyrodidae) is associated with a point mutation in an ace1-type acetylcholinesterase and overexpression of carboxylesterase. Insect Biochem Mol Biol 38: 940–949.

18. Seino A, Kazuma T, Tan AJ, Tanaka H, Kono Y, et al. (2007) Analysis of two acetylcholinesterase genes in *Bombyx mori*. Pestic Biochem Physiol 88: 92–101.

19. Lee SW, Kasai S, Komagata O, Kobayashi M, Agui N, et al. (2007) Molecular characterization of two acetylcholinesterase cDNAs in *Pediculus* human lice. J Med Entomol 44: 72–79.

20. Cassanelli S, Reyes M, Rault M, Manicardi CG, Sauphanor B (2006) Acetylcholinesterase mutation in an insecticide-resistant population of the codling moth *Cydia pomonella* (L.). Insect Biochem Mol Biol 36: 642–653.

21. Lee DW, Kim SS, Shin SW, Kim WT, Boo KS (2006) Molecular characterization of two acetylcholinesterase genes from the oriental tobacco budworm, *Helicoverpa assulta* (Guenée). Biochim Biophys Acta 1760: 125–133.

22. Li F, Han ZJ (2002) Two different genes encoding acetylcholinesterase existing in cotton aphid (*Aphis gossypii*). Genome 45: 1134–1141.

23. Nabeshima T, Mori A, Kozaki T, Iwata Y, Hidoh O, et al. (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.

24. Nabeshima T, Kozaki T, Tomita T, Kono Y (2003) An amino acid substitution on the second acetylcholinesterase in the pirimicarb-resistant strains of the peach potato aphid, *Myzus persicae*. Biochem Biophys Res Commun 307: 15–22.

25. Huchard E, Martinez M, Alout H, Douzery EJ, Lutfalla G, et al. (2006) Acetylcholinesterase genes within the Diptera: Takeover and loss in true flies. Proc R Soc Lond Ser B 273: 2595–2604.

26. Ni X-Y, Tomita T, Kasai S, Kono Y (2003) cDNA and deduced protein sequence of acetylcholinesterase from the diamondback moth, *Plutella xylostella* (L.) (Lepidoptera: Plutellidae). Appl Entomol Zool (Jpn) 38: 49–56.

27. Baek JH, Kim JI, Lee DW, Chung BK, Miyata T, et al. (2005) Identification and characterization of ace1-type acetylcholinesterase likely associated with organophosphate resistance in *Plutella xylostella*. Pestic Biochem Physiol 81: 164–175.

28. Lee DW, Choi JY, Kim WT, Je YH, Song JT, et al. (2007) Mutations of acetylcholinesterase 1 contribute to prothiofos-resistance in *Plutella xylostella* (L.). Biochem Biophys Res Commun 353: 591–597.

29. Jiang X, Qu M, Denholm I, Fang J, Jiang W, et al. (2009) Mutation in acetylcholinesterase1 associated with triazophos resistance in rice stem borer, *Chilo suppressalis* (Lepidoptera: Pyralidae). Biochem Biophys Res Commun 378: 269–272.

30. Mori A, Lobo NF, de Bruyn B, Severson DW (2007) Molecular cloning and characterization of the complete acetylcholinesterase gene (*Ace1*) from the mosquito *Aedes aegypti* with implications for comparative genome analysis. Insect Biochem Mol Biol 37: 667–674.

31. Kozaki T, Kimmelblatt BA, Hamm RL, Scott JG (2008) Comparison of two acetylcholinesterase gene cDNAs of the lesser mealworm, *Alphitobius diaperinus*, in insecticide susceptible and resistant strains. Arch Insect Biochem Physiol 67: 130–138.

32. Chen HJ, Liao Z, Liu XM, Li GQ, Li F, et al. (2009) *Ace2*, rather than *Ace1*, is the major acetylcholinesterase in the silkworm, *Bombyx mori*. Insect Sci 16: 297–303.

33. Pang Y-P (2006) Novel acetylcholinesterase target site for malaria mosquito control. PLoS ONE 1: e58.

34. Pang Y-P (2007) Species marker for developing novel and safe pesticides. Bioorg Med Chem Lett 17: 197–199.

35. Pang Y-P, Singh SK, Gao Y, Lassiter TL, Mishra RK, et al. (2009) Selective and irreversible inhibitors of aphid acetylcholinesterases: steps toward human-safe insecticides. PLoS ONE 4: e4349.

36. Pang Y-P, Ekström F, Polsinelli GA, Gao Y, Rana S, et al. (2009) Selective and irreversible inhibitors of mosquito acetylcholinesterases for controlling malaria and other mosquito-borne diseases. PLoS ONE 4: e6851.

37. Lang GJ, Zhang XH, Zhang MY, Zhang CX (2010) Comparison of catalytic properties and inhibition kinetics of two acetylcholinesterases from a lepidopteran insect. Pestic Biochem Physiol 98: 175–182.

38. von Heijne G (1987) Sequence Analysis in Molecular Biology Academic Press San Diego.

39. Bourne Y, Grassi J, Bougis PE, Marchot P (1999) Conformational flexibility of the acetylcholinesterase tetramer suggested by X-ray crystallography. J Biol Chem 274: 30370–30376.

40. Bourne Y, Taylor P, Bougis PE, Marchot P (1999) Crystal structure of mouse acetylcholinesterase: A peripheral site-occluding loop in a tetrameric assembly. J Biol Chem 274: 2963–2970.

41. Harel M, Kryger G, Rosenberry TL, Mallender WD, Lewis T, et al. (2000) Three-dimensional structures of *Drosophila melanogaster* acetylcholinesterase and of its complexes with two potent inhibitors. Protein Sci 9: 1063–1072.

42. Nicolet Y, Lockridge O, Masson P, Fontecilla-Camps JC, Nachon F (2003) Crystal structure of human butyrylcholinesterase and of its complexes with substrate and products. J Biol Chem 278: 41141–41147.

43. Sanson B, Colletier JP, Xu Y, Therese Lang P, Jiang H, et al. (2011) Backdoor opening mechanism in acetylcholinesterase based on X-ray crystallography and MD simulations. Protein Sci 20: 1114–1118.

44. Nachon F, Stojan J, Fournier D (2008) Insights into substrate and product traffic in the *Drosophila melanogaster* acetylcholinesterase active site gorge by enlarging a back channel. FEBS J 275: 2659–2664.

45. Mori A, Tomita T, Hidoh O, Kono Y, Severson DW (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.

46. Gnagey AL, Forte M, Rosenberry TL (1987) Isolation and characterization of acetylcholinesterase from *Drosophila*. J Biol Chem 262: 13290–13298.

47. Haas R, Marshall TL, Rosenberry TL (1988) *Drosophila* acetylcholinesterase: demonstration of a glycoinositol phospholipid anchor and an endogenous proteolytic cleavage. Biochemistry 27: 6453–6457.

48. MacPhee-Quigley K, Vedvick TS, Taylor P, Taylor SS (1986) Profile of the disulfide bonds in acetylcholinesterase. J Biol Chem 261: 13565–13570.

49. Page RDM, Holmes EC (1998) Molecular Evolution: A Phylogenetic Approach Blackwell Science, Oxford, UK.

50. Xu Y, Colletier JP, Weik M, Jiang H, Moult J, et al. (2008) Flexibility of aromatic residues in the active-site gorge of acetylcholinesterase: X-ray versus molecular dynamics. Biophys J 95: 2500–2511.

51. Auld VJ, Fetter RD, Broadie K, Goodman CS (1995) Gliotactin, a novel transmembrane protein on peripheral glia, is required to form the blood-nerve barrier in *Drosophila*. Cell 81: 757–767.

52. Botti SA, Felder CE, Sussman JL, Silman I (1998) Electrotactins: a class of adhesion proteins with conserved electrostatic and structural motifs. Protein Eng 11: 415–420.

53. Raves ML, Harel M, Pang Y-P, Silman I, Kozikowski AP, et al. (1997) Structure of acetylcholinesterase complexed with the nootropic alkaloid (-)-huperzine A. Nature Struct Biol 4: 57–63.

54. Schrag JD, Cygler M (1993) 1·8 Å refined structure of the lipase from *Geotrichum candidum*. J Mol Biol 230: 575–591.

55. Haliscak JP, Beeman RW (1983) Status of malathion resistance in genera of beetles infesting farm-stored corn, wheat, and oats in the United States. J Econ Entomol 76: 717–722.

56. Bendtsen JD, Nielsen H, von Heijne G, Brunak S (2004) Improved prediction of signal 459 peptides: SignalP 3.0. J Mol Biol 340: 783–795.

57. Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, et al. (2007) ClustalW and ClustalX version 2.0. Bioinformatics 23: 2947–2948.

58. Tamura K, Dudley J, Nei M, Kumar S (2007) Molecular evolutionary genetics analysis (MEGA) software version 4.0. Mol Biol Evol 24: 1596–1599.

59. Sanson B, Nachon F, Colletier JP, Froment MT, Toker L, et al. (2009) Crystallographic snapshots of nonaged and aged conjugates of soman with acetylcholinesterase, and of a ternary complex of the aged conjugate with pralidoxime (dagger). J Med Chem 52: 7593–7603.

60. Pearlman DA, Case DA, Caldwell JW, Ross WS, Cheatham TE, III, et al. (1995) AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules. Comput Phys Commun 91: 1–41.

61. Jorgensen WL, Chandreskhar J, Madura JD, Impey RW, Klein ML (1982) Comparison of simple potential functions for simulating liquid water. J Chem Phys 79: 926–935.

62. Case DA, Cheatham TE, III, Darden T, Gohlke H, Luo R, et al. (2005) The Amber biomolecular simulation programs. J Comput Chem 26: 1668–1688.

63. Hornak V, Abel R, Okur A, Strockbine B, Roitberg A, et al. (2006) Comparison of multiple Amber force fields and development of improved protein backbone parameters. Proteins 65: 712–725.

64. Wickstrom L, Okur A, Simmerling C (2009) Evaluating the performance of the ff99SB force field based on NMR scalar coupling data. Biophys J 97: 853–856.

65. Berendsen HJC, Postma JPM, van Gunsteren WF, Di Nola A, Haak JR (1984) Molecular dynamics with coupling to an external bath. J Chem Phys 81: 3684–3690.

66. Darden TA, York DM, Pedersen LG (1993) Particle mesh Ewald: an N log(N) method for Ewald sums in large systems. J Chem Phys 98: 10089–10092.

67. Shao J, Tanner SW, Thompson N, Cheatham TE, III (2007) Clustering molecular dynamics trajectories: 1. Characterizing the performance of different clustering algorithms. J Chem Theory Comput 3: 2312–2334.

68. Felder CE, Prilusky J, Silman I, Sussman JL (2007) A server and database for dipole moments of proteins. Nucleic Acids Res 35: W512–521.

69. Giulietti A, Overbergh L, Valckx D, Decallonne B, Bouillon R, et al. (2001) An overview of real-time quantitative PCR: Applications to quantify cytokine gene expression. Methods 25: 386–401.

70. Li HR, Oppert B, Higgins RA, Huang FN, Buschman LL, et al. (2005) Characterization of cDNAs encoding three trypsin-like proteinases and quantitative analysis of mRNA in Bt-resistant and -susceptible strains of *Ostrinia nubilalis*. Insect Biochem Mol Biol 35: 847–860.

71. Togawa T, Dunn WA, Emmons AC, Nagao J, Willis JH (2008) Developmental expression patterns of cuticular protein genes with the R&R Consensus from *Anopheles gambiae*. Insect Biochem Mol Biol 38: 508–519.