A THESIS
FOR THE DEGREE OF MASTER OF SCIENCE

Evolutionary origin and status of two insect acetylcholinesterases and their structural conservation and differentiation

두 곤충 아세틸콜린에스터레이즈의 진화적 기원과 양상 및 구조적 보전과 분화 분석

By
Deok Jea Cha

Department of Agricultural Biotechnology
Seoul National University
February, 2014

A THESIS
FOR THE DEGREE OF MASTER OF SCIENCE

Evolutionary origin and status of two insect acetylcholinesterases and their structural conservation and differentiation

두 곤충 아세틸콜린에스터레이즈의 진화적 기원과 양상 및 구조적 보전과 분화 분석

By
Deok Jea Cha

Department of Agricultural Biotechnology
Seoul National University
February, 2014

Evolutionary origin and status of two insect acetylcholinesterases and their structural conservation and differentiation

두 곤충 아세틸콜린에스터레이즈의 진화적 기원과 양상 및 구조적 보전과 분화 분석

UNDER THE DIRECTION OF ADVISER SIHYEOCK LEE
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL OF
SEOUL NATIONAL UNIVERSITY

By
Deok Jea Cha

DEPARTMENT OF AGRICULTURAL BIOTECHNOLOGY
SEOUL NATIONAL UNIVERSITY
February, 2014

APRROVED AS A QUALIFIED THESIS OF DEOK JEA CHA
FOR THE DEGREE OF MASTER OF SCIENCE BY THE COMMITTEE
MEMBERS

CHAIRMAN	**Seunghwan Lee**	_____
VICE CHAIRMAN	**Si Hyeock Lee**	_____
MEMBER	**Kwang Pum Lee**	_____

Evolutionary origin and status of two insect acetylcholinesterases and their structural conservation and differentiation

Major in Entomology
Department of Agricultural Biotechnology, Seoul National University

Deok Jea Cha

Abstract

Acetylcholinesterase (AChE) plays a pivotal role in the synaptic transmission in the cholinergic nervous system of most animals, including insects. Insects possess two distinct acetylcholinesterases (AChE1 vs. AChE2), which are encoded by two paralogous loci originated from the duplication that occurred long before the radiation of insects. In this study, phylogenetic analysis and structural modeling were performed to understand when the *ace* duplication occurred and what structural features have been associated with the differentiation of two AChEs during evolution. The phylogenetic analysis was conducted for the AChE-like genes from all known lower animals with their genome sequenced together

with all known arthropod *ace1* and *ace2* orthologs, including those from a number of insects that were newly cloned in this study. The result suggested that the last common ancestor of *ace1* and *ace2* shared its origin with those of platyhelminthes, which further implies that the lineage of arthropod *ace1* and *ace2* has undergone a divergent evolution along with those of platyhelminthes. In addition, it appears that the *ace* duplication event resulting in the split of the *ace1* and *ace2* clades occurred after the divergence of Ecdysozoa and Lophotrochozoa from their Protostomian common ancestor but before the split of Ecdysozoa into their descendents. The *ace1* lineage showed a significantly lower evolutionary rate (*d*, or distance) and higher purifying selection pressure (*dN/dS*) compared to *ace2* lineage, suggesting that the *ace1* lineage has maintained relatively more essential functions following duplication. Given that the putative functional transition of *ace* in some Hymenopteran could not be explained by such difference in evolutionary rate, this event appears to have occurred in relatively recent time by only a few numbers of mutations resulting in dramatic alteration of AChE function.

The amino acid sequence comparison between AChE1 and AChE2 from a wide variety of insect taxa revealed a high degree of sequence conservancy in the functionally crucial domains, suggesting that presence of strong purifying selection pressure over these essential residues. Interestingly, the EF-hand motif

was mostly found in the AChE1 lineage but not in AChE2. In contrast, LRE-motif was partially conserved in the AChE2 lineage but not in AChE1. In addition, the AChE2-specific insertion domain appeared to have been introduced relatively more recently, perhaps during the radiation of insects, after the duplication. The comparison of five essential domains [i.e., the catalytic anionic site (CAS), peripheral anionic site (PAS), acyl binding pocket (ABP), oxyanion hole and catalytic triad] in the active-site gorge showed unique differences in amino acid residues of the PAS (Asp72 vs. Tyr72, Tyr121 vs. Met121; amino acid numbering of *Torpedo californica* AChE, hereafter) and the ABP (Cys288 vs. Leu288) between AChE1 and AChE2. Three-dimensional modeling of active-site gorge from insect AChEs with a particular focus on the PAS revealed that a subtle but consistent structural alteration in the active-site gorge topology was caused by the PAS amino acid substitution, likely resulting in a remarkable functional differentiation between two AChEs. Although *ace1* appears to have evolved at significantly slower rates to retain its essential function, it is likely that a few specific amino acid substitutions in *ace1* causing a dramatic reduction of enzyme activity may have occurred locally in more recent time and resulted in the functional transition from AChE1 to AChE2 as observed in some Hymenopteran insects.

Key words: Insect acetylcholinesterase, gene duplication, evolutionary distance, selection pressure, active-site gorge, phylogenetic analysis, three-dimensional modeling

Student Number: 2011-21321

CONTENTS

LIST OF TABLES

CHAPTER 1.

The origin and evolutionary history of two insect acetylcholinesterases

CHAPTER 2.

Structural conservation and differentiation of two insect acetylcholinesterases during evolution

LIST OF FIGURES

CHAPTER 1.

The origin and evolutionary history of two insect acetylcholinesterases

CHAPTER 2.

Structural conservation and differentiation of two insect acetylcholinesterases during evolution

CHAPTER 1.

The origin and evolutionary history of two insect acetylcholinesterases

The origin and evolutionary history of two insect acetylcholinesterases

Abstract

Acetylcholinesterase (AChE) plays a pivotal role in the synaptic transmission in the cholinergic nervous system of most animals, including insects. Insects have two different *ace* (*ace1* and *ace2*) loci that encode two distinct AChEs (AChE1 and AChE2), which were originated by duplication events long before the radiation of insects. However, little is known about when the *ace* duplication occurred and how each duplicated *ace* locus has evolved to retain the original functions. In this study, the phylogenetic analysis was performed for acetylcholinesterase-like genes from all known lower animals with their genome sequenced together with all known arthropod *ace1* and *ace2*, including those from a number of insects that were newly cloned in this study. Several independent duplications found in lower animal lineages were not directly related with that of arthropod *ace1* and *ace2*, suggesting that these lower animals were not the direct origin of *ace1* and *ace2*. The common ancestor of *ace1* and *ace2* shared their origin with those of platyhelminthes, which implies that the lineage of arthropod

ace1 and *ace2* has undergone a divergent evolution along with those of platyhelminthes. In addition, it appears that the *ace* duplication event resulting in the split of the *ace1* and *ace2* clades occurred after the divergence of Ecdysozoa and Lophotrochozoa from their Protostomian common ancestor but before the split of Ecdysozoa into their decendents. Comparison of the evolutionary rate (*d,* or distance) and selection pressure (*dN/dS*) of two *aces* from different insect groups relative to those from insect common ancestors revealed that *ace1* has evolved with a significantly lower evolutionary rate compared to *ace2*, suggesting that the *ace1* lineage has maintained relatively more essential functions following duplication. Given that the putative functional transition of ace in some Hymenopteran could not be explained by such difference in evolutionary rate, this event appears to have occurred in relatively recent time by only a few numbers of mutations resulting in dramatic alteration of AChE function.

1. Introduction

Acetylcholinesterase (AChE, EC 3.1.1.7) plays a pivotal role in the synaptic transmission by hydrolyzing the neurotransmitter acetylcholine at cholinergic synapses and neuromuscular junctions in the cholinergic nervous system of most animals, including insects (Toutant 1989). In addition to its classical synaptic functions, AChE is known to play other non-classical roles, including neurite outgrowth, synapse formation (Olivera et al., 2003), modulation of glial activation, tau phosphorylation (Ballard et al., 2005) and learning/memory (Gauthier et al., 1992; Shapira et al., 2001).

True cholinesterases (ChEs) with highly selective substrate specificity and catalytic activity, including AChE and butyrylcholinesterase (BChE, EC 3.1.1.8), have been suggested to appear in the early bilaterians, such as platyhelminthes having the rudimentary cholinergic system (Pezzementi and Chatonnet 2010). To date, two *ace* loci (*ace1*, encoding AChE1, which is paralogous to *Drosophila melanogaster ace*; *ace2*, encoding AChE2, which is orthologous to *D. melanogaster ace*) have been identified in various insect species (Baek et al., 2005; Gao et al., 2002; Kim et al., 2006; Lee et al., 2006; Li and Han 2002; Nabeshima et al., 2003; Weill et al., 2002). In contrast, Cyclorrhaphan flies, including *D. melanogaster* (Weill et al., 2002) and *Musca domestica* (Fournier et al., 1988), are

known to possess only a single *ace* locus (*ace2*) by losing the *ace1* copy during the course of evolution (Weill et al., 2002).

Among the two AChEs, AChE1 has been proposed as a major catalytic enzyme based on its higher expression level and frequently observed point mutations associated with insecticide resistance. A relatively large scale survey to determine the main catalytic AChE across a wide array of insect species, however, revealed that AChE1 was the main catalytic enzyme in 67 species out of 100 insect species examined, whereas AChE2 was predominantly expressed as the main catalytic enzyme in the remaining 33 species, ranging from Palaeoptera to Hymenoptera (Kim and Lee 2013). In Diptera, all of the Orthorrhaphan flies were determined to have two AChEs, with AChE1 being the main enzyme, whereas only AChE2 was detected in all Cyclorrhaphan flies examined. In social Hymenoptera, however, many bees and wasps were found to use AChE2 as the main catalytic enzyme. These findings challenged the common notion that AChE1 is the only main catalytic enzyme in insects with the exception of Cyclorrhapha, and further demonstrate that the specialization of AChE2 as the main enzyme or the functional transition from AChE1 to AChE2 (or replacement of AChE1 function with AChE2) or vice versa were rather common events, having multiple independent origins during insect evolution.

Although the exact physiological functions of AChE1 and AChE2 remain to

be elucidated, functional specialization between AChE1 and AChE2 appears to be common in insects. AChE1 was proposed to have the main synaptic function, whereas AChE2 to play a limited synaptic role and/or other non-synaptic roles in Culicidae mosquitoes (Huchard et al., 2006; Weill et al., 2002). Likewise, AChE1 plays a major role in synaptic transmission whereas AChE2 has a limited neuronal function and/or other functions in *B. germanica* (Kim et al., 2006; Kim et al., 2010; Revuelta et al., 2009). Furthermore, knockdown of AChE1 resulted in 100% mortality and increased susceptibility to OP insecticides in *T. castaneum* (Lu et al., 2012). More recent study on honey bee AChE revealed that, unlike most other insects, AChE2 has neuronal functions, whereas AChE1 has non-neuronal functions, including chemical defense. Taken together, any AChE possessing the neuronal function is likely more ancient and essential, resulting in immediate disruption of physiology if its activity or expression is suppressed.

It has been hypothesized that the *ace1* and *ace2* in insects originated from a duplication event that occurred long before the radiation of insects (Weill et al., 2002) but it is unclear yet its phylogenetic origin. The first true ChEs with high catalytic activity in the early bilaterians was most likely to have neuronal functions. Thus, if one can narrow down the time point of *ace* duplication during the evolution of bilaterians, including Protostomia, it may be possible to confirm whether the last common ancestor of insect AChE1 and AChE2 had the neuronal

function and whether the functional diversification of a duplicated *ace* copy was evolved to acquire other non-neuronal functions. With rapid expansion of genome information of various organisms, particularly lower animals, it has been feasible to determine the presence or absence of *ace* paralogs within a genome of interest. Therefore, phylogenetic analysis of insect *ace* genes with the recently identified *ace* genes from many lower animals, such as ones belonging to Choanoflagellata, Porifera, Placozoa, Cnidaria, Ctenophora, Platyhelminthes, Annelida and Mollusca, would enable to determine when the two *ace* paralogs appear and which *ace* clade belongs to a more ancient lineage.

The evolutionary rate of a gene can be estimated by calculating the number or rate of substitutions per position between orthologous sequences and has been suggested to be determined by the significance of the gene for the fitness of the organism. Based on the "knockout rate" hypothesis, the greater the impacts of a gene alteration on fitness is, the slower the evolutionary rate is. Therefore, the evolutionary rate of genes having essential functions is assumed to be significantly slower than that of genes with less essential functions (Wilson et al., 1977). Under the assumption that the primary function of AChE is neuronal and the secondary function gained after duplication is non-neuronal, AChE having more essential function [i.e., neuronal (synaptic) function] could have been resistant to mutation, particularly to non-synonymous one, resulting in a relatively

lower rate of evolution. In contrast, the other surplus copy of AChE, which may be relatively less essential, could have been evolved relatively faster and been specialized (or subfunctionalized) to gain non-neuronal (non-synaptic) functions.

The goal of this chapter is two folds. First, to identify the phylogenetic origin of insect *ace* genes and predict when the duplication occurred in the evolution of bilaterians, phylogenetic analysis was conducted for a variety of invertebrates, ranging from Choanoflagellata to Arthropoda. Secondly, to test the hypothesis whether the functionality of AChE (i.e., neuronal vs. non-neuronal) can be predicted by comparing the evolutionary rates between the *ace1* and *ace2* orthologs within a given insect taxon, we first cloned *ace1* and *ace2* genes from underrepresented groups of Hexapoda, such as Collembola, Archaeognatha, Zygentoma, Ephemeroptera, Odonata and Dermaptera and retrieved *ace1* and *ace2* gene sequences of a wide variety of insect species from online database. Then, we estimated and compared the evolutionary rates of the *ace1* and *ace2* orthologs in Coleoptera, Diptera (exception of cyclorrhaphan flies), Hemiptera, Hymenoptera and Lepidoptera displaying distinct propensity of AChE usage.

2. Materials and methods

2.1. Insects sampling and Sequence preparation

Tomocerus kinoshitai (Collembola), *Haslundichilis viridis* (Archaeognatha), *Ctenolepima longicandata* (Zygentoma), *Lyriothemis pachygasta* (Odonata) and *Anechura japonica* (Dermaptera) were collected from the Gwanak-mountain, Gwanak-gu, Seoul and *Ecdyonurus levis* (Ephemeroptera) was collected from the Ungcheon stream, Ungcheon-eup, Boryung, Choongchungnamdo, Korea.

Other metazoans (i.e., Arthropoda, Nematoda, Mollusca, Annelida Platyhelminthes, Ctenophora, Cnidaria, Placozoa, Porifera and Choanoflagellata) *aces* gene and *ace-like* gene sequences were retrieved from ESTHER (http://bioweb.ensam.inra.fr.esther), ENSEMBL (http://www.ensembl.org/), Uniprot (http://www.uniprot.org/), VectorBase (https://www.vectorbase.org/), DOE Joint Genome Institute (http://genome.jgi-psf.org/), GENBANK (http://www.ncbi.nlm.nih.gov/) and NCBI Genome (http://www.ncbi.nlm.nih.gov/genome/). All of the *aces* and *ace-like* sequence data were recorded in the appendix.

2.2. Total RNA extraction and cDNA synthesis

Total RNA was extracted from head and thorax tissues with appropriate

volume of TRI reagent (MRC, Cincinnati, OH, USA). First-strand cDNA synthesis was carried out according to the Superscript^{TM} III first-strand synthesis system or ThermoScript reverse transcripatase (Invitrogen, Carlsbad, CA) protocol using an oligo(dT)^{20} primer and subsequently used as an internal PCR template.

2.3. Homology-probing PCR

Internal fragments of *ace1* and *ace2* genes were PCR-amplified from cDNA template with a set of degenerate primers (Table 1) designed from the conserved motifs specific to insect *ace1* (5'ace1-de, -MWNPNT-; ace1-de-5(R), -GGGFYSG-; ace1-de-3(F), -MVGDYHF-; 3'ace1-DF, -MRYWANF-) and *ace2* (5'ace2-de, -IPYAKP-; 5'ace2-DF, -GEEMWNP-; 3'ace2-DF2, -WIYGGG-; 3'ace2-de2, -WGEWMGV-), respectively. From the first round of PCR using sets of degenerate primers, internal cDNA fragments of *ace1* (246 - 248 bp) and *ace2* (126 -136 bp) were obtained, from which gene-specific primers were designed. Subsequent PCR with a pair of gene-specific primer vs. degenerate primer (Table 2) produced the extended fragments of *ace1* (1042 - 1205 bp) and *ace2* (873 – 1094 bp). For all PCRs, Advantage 2 Polymerase mix (Clontech, Palo Alto, CA, USA) was used under the thermal program with multiple annealing steps: following an initial denaturation at 98°C for 5 min, a total of 9 cycles of gradient

Table 1. Degenerated primers used for the primary internal PCRs of insect *ace1* and *ace2*

Name	Sequence	Amino acid sequence	Degeneracy	Length (bp)
ace1				
5'ace1-de	ATGTGGAAYCCNAAYWC	MWNPNT	32	17
ace1-de-5(R)	GGNGGHGGNTTYTAYTCNGG	GGGFYSG	768	20
ace1-de-3(F)	ATGGTVGGNGAYTAYCABTT	MVGDYHF	144	20
3'ace1-DF	ATGMGRTAYTGGDCNAAYTT	MRYWANF	192	20
ace2				
5'ace2-de	THCCNTWYGCVAARCC	IPYAKP	288	16
5'ace2-DF	GGNGARGARATRTGGAAYCC	GEEMWNP	64	20
3'ace2-DF2	TGGATHTAYGGNGGHGGBT	WIYGGG	432	19
3'ace2-de2	GGGNGADTGGATGGGNGT	WGEWMGV	48	18

N: A or C or G or T, M: A or C, R: A or G, W: A or T, Y: C or T, V: A or C or G; not T, H: A or C or T; not G, D: A or G or T; not C B: C or G or T; not A

Table 2. Gene-specific primers used for the secondary internal PCRs of insect *ace1* and *ace2*

Name		Sequence	Amino acid sequence	Length (bp)
Tomocerus kinoshitai				
ace1	TK-A1-3	GAGCTCAACCTCCTGAGTTG	TQEVEL	20
ace2	TK-A2-5	GGATCCCTGCTTCGGTATTC	IPASVF	20
Haslundichilis viridis				
ace1	HV-A1-3	GCTCTATTTCACTGCTGGTG	TSSEIE	20
ace2	HV-A2-5	GTACCTGAACATTTGGAGGC	YLNIWR	20
Ecdyonurus levis				
ace1	EL-A1-3	GCTCAATTTCTTGCGGCTGG	QPQEIE	20
ace2	EL-A2-5	GGGCACCTAAGAAAAGGAAAG	APKKRK	21
Lyriothemis pachygasta				
ace1	LP-A1-3	CTCTTGCTGAGGTCGATTTC	EIDLSK	20
ace2	LP-A2-5	ACCAACATCTCCGAGGACTG	TNISED	20
Anechura japonica				
ace1	AJ-A1-3	CTCGAGAGTTCGATCTCATG	HEIELS	20
ace2	AJ-A2-5	ACCAACATTTCCGAGGACTG	TNISED	20
Ctenolepima longicandata				
ace1	Zy-A1-3	CATCATTCGCTTGCTAAGTTCC	ELSKRMM	22
ace2	Zy-A2-5	ACTGTCTATATCTCAATCTTTGG	CLYLNLW	23

annealing [94°C for 30 sec, 60, 55, and 50°C for 30 sec (3 cycles for each annealing temperature), and 68°C for 2 min] was followed by 30 main cycles (94°C for 30 sec, 45°C for 30 sec, 68°C for 2 min) and final 1 cycle of single extension (68°C for 2 min). The PCR products were extracted by using gel extraction kit (QIAgen, Valencia, CA, USA) and then cloned into pGEM®-T easy vector (Promega, Madison, WI, USA) and sequenced.

2.4. Rapid amplification of cDNA ends (RACE)

Gene-specific RACE primers (Table 3) were designed for rapid amplification of cDNA ends (RACE). The 5' and 3' RACE reactions were performed using the 5'-full Core set (TaKaRa, Shiga, Japan) and the SMART RACE kit (Clontech), respectively. Both experimental procedures were followed by the Manufacturer's instructions with some modification. For 5' RACE reactions, 5' RACE ready-cDNA as template was synthesized by using 5' phosphorylated primers that were designed from the internal cDNA fragment (Table 3) with AMV Reverse Transcriptase XL and was converted to concatemer or circular form by T4 RNA ligase (NEB, UK). The 5' RACE PCR was conducted with following thermal condition: following initial denaturation at 98°C for 8 min, consecutive 10 cycles of gradient annealing [95°C for 30 sec, 65~63°C and 61~63°C for 30 sec (5 cycles for each annealing temperature), and 68°C for 2~3 min] was followed by 30 main

Table 3. Primers used for 5'RACE and 3'RACE of insect *ace1* and *ace2*

Name			Sequence	Amino acid sequence	Length (bp)
Tomocerus kinoshitai					
5'RACE	*ace1*	Tk-A1-RT	GTACTGCATGGAGAC	VSMQY	15
		Tk-A1-S1	GGGGACTTCTACGCTGGACG	GTSTLD	20
		Tk-A1-S2	TTGGCCGTGACCGAGCAGG	LAVTEQ	19
		Tk-A1-A1	CGCCAACTTTGGTGGTCTATCG	DRPPKLA	22
		Tk-A1-A2	AACAGGCAGTCCTCTGAAACTTG	QVSEDCLF	23
	ace2	Tk-A2-RT	TGGTCGAGTAAACCG	VGLLDQ	15
		Tk-A2-S1	TGGTATCAACTCAGTATCGAGTGG	VVSTQYRV	24
		Tk-A2-S2	TTCGGAACAGACGATGCTCCC	FGTDDAP	21
		Tk-A2-A1	TCAGCTGCTAGGACGGCTCC	GAVLAAE	20
		Tk-A2-A2	GCACTGCTTCCACTGAAGAATCC	GFFSGSSA	23
3'RACE	*ace1*	Tk-A1-GSP	TGTATTACTTCACCCACCGGTCCTCCC	MYYFTHRSS	27
		Tk-A1-NGSP	GGTGAGCCACTTGACCCTTCG	GEPLDPS	21
	ace2	Tk-A2-GSP	AATTGGGAGAACCCAAATGACGGGTACC	NWENPNDGY	28
		Tk-A2-NGSP	GACGCTACGCTGACACTTGGTC	RRYADTWS	22

Table 3. Continued

Name			Sequence	Amino acid sequence	Length (bp)
Haslundichilis viridis					
5'RACE	*ace1*	Hv-A1-RT	TGTTGTCGTGAACCC	WVHDNI	15
		Hv-A1-S1	TATTTCGACACCGCCGACGTTC	YFDTADV	22
		Hv-A1-S2	TTCGATCAGCTGATGGCTCTGC	FDQLMAL	22
		Hv-A1-A1	GCGATACTGCATCGACACCAG	LVSMQYR	21
		Hv-A1-A2	TGTGATCGTAGACGTCGAGCG	TLDVYDH	21
	ace2	Hv-A2-RT	ATTACCAGGAGCCTC	EAPGN	15
		Hv-A2-S1	CGACGGCTTGGCTGCAGAAG	DGLAAE	20
		Hv-A2-S2	GTCGTGGCATCAATGCAGTATAG	VVASMQYR	23
		Hv-A2-A1	GTCGACGTACCACTCATATAGCC	GYMSGTST	23
		Hv-A2-A2	CCTTGCCTTCTCTTGGAGCC	APREGK	20
3'RACE	*ace1*	Hv-A1-GSP	TTTGCGTATCGCTATGCCGAAACGGGC	FAYRYAETG	27
		Hv-A1-NGSP	CTACATCTTCGGCGAACCTCTC	YIFGEPL	22
	ace2	Hv-A2-GSP	ACCTTCCTCCAGCGCGACAAGTACCTG	TFLQRDKYL	27
		Hv-A2-NGSP	TCAGAGAACAAGCCTCAACCTGTG	SENKPQPV	24

Table 3. Continued

Name			Sequence	Amino acid sequence	Length (bp)
Ecdyonurus levis					
5'RACE	*ace1*	El-A1-RT	TGCTACGCGATACTG	QYRVA	15
		El-A1-S1	ATTCCGGTACTTCCACTTTAGACG	YSGTSTLD	24
		El-A1-S2	TGGCAGAAGAAAAGGTCATACTTG	VAEEKVILV	24
		El-A1-A1	TCCTCCGCCGTATATCCACAC	VWIYGGG	21
		El-A1-A2	TTGGTCTTGGTTTTGGAACTACAAC	VVVPKPRP	25
	ace2	El-A2-RT	TGCCTCGCTGTCTC	ETARH	14
		El-A2-S1	GCTATGCAATATCGAGTTGGTGC	AMQYRVGA	23
		El-A2-S2	ATACATGCCTGGTGAAGATGGTAG	YMPGEDGS	24
		El-A2-A1	ACATCCAGGGTAGTAGTGCCG	GTTTLDV	21
		El-A2-A2	ACCACCACCATATACCCATACAAG	LVWVYGGG	24
3'RACE	*ace1*	El-A1-GSP	GTCAATGAATTGGCTCACCGTTACGCCG	VNELAHRYA	28
		El-A1-NGSP	ACCGGAGTCATGCATGGAGATG	TGVMHGD	22
	ace2	El-A2-GSP	ATTCCAGTACACCGCTTGGGAACACATGG	FQYTAWEHM	29
		El-A2-NGSP	CTGAGCATGGAGTACCAGTCTAC	AEHGVPVY	23

Table 3. Continued

Name			Sequence	Amino acid sequence	Length (bp)
Lyriothemis pachygasta					
5'RACE	*ace1*	Lp-A1-RT	TCACCAAACAGCGTG	TLFGE	15
		Lp-A1-S1	GGTTGTACGATCAGGTTATGGC	GLYDQVMA	22
		Lp-A1-S2	ACAACATCCATGCCTTTGGAGG	DNIHAFGG	22
		Lp-A1-A1	GAGAAGCGACGCGGTATTGC	QYRVAS	20
		Lp-A1-A2	GACGTGCCGGAATAGAACCC	GFYSGTS	20
	ace2	Lp-A2-RT	TAGGGTCAGCGATTC	ESLTL	15
		Lp-A2-S1	TCGGTTTCCTCTACCTCGGC	FGFLYLG	20
		Lp-A2-S2	CAGGCTTTGGCCATCCGTTG	QALAIRW	20
		Lp-A2-A1	CTTCGGCAGCAGCCATCATG	MMAAAE	20
		Lp-A2-A2	CTCCTCCGTAGATCCACACC	VWIYGG	20
3'RACE	*ace1*	Lp-A1-GSP	GATGCTCTGGACAAGATGGTGGGAGAC	DALDKMVGD	27
		Lp-A1-NGSP	CCGATGAGATACCTTATGTGTTCGG	ADEIPYVFG	25
	ace2	Lp-A2-GSP	GATCGGCAGTAACCAGGACGAAGGGAC	IGSNQDEGT	27
		Lp-A2-NGSP	CTTCCGAAGAGAACAGGCGAGC	FRREQAS	22

Table 3. Continued

Name			Sequence	Amino acid sequence	Length (bp)
Anechura japonica					
5'RACE	*ace1*	Aj-A1-RT	GAGTGATGCAACACG	RVASL	15
		Aj-A1-S1	GTGGTGGCTTCTATTCAGGAAG	GGGFYSGS	22
		Aj-A1-S2	CAACATTGGACGTGTATGATCCG	STLDVYDP	23
		Aj-A1-A1	ATCCACACCATGACTGCGGC	AAVMVWI	20
		Aj-A1-A2	ACGCTAACGTAAAGGCAATCCTC	EDCLYVSV	23
	ace2	Aj-A2-RT	TGAATCGAAGCGACG	VASIQ	15
		Aj-A2-S1	GGCACGGCCACATTGGATATC	GTATLDI	21
		Aj-A2-S2	ATCTGATGGTCGCCGCCAATG	DLMVAAN	21
		Aj-A2-A1	TCCGTAGACCCACACAAGAATTG	PILVWVYG	23
		Aj-A2-A2	GCGGATTCGGTTTCGTTGAGC	AQRNRIR	21
3'RACE	*ace1*	Aj-A1-GSP	TCAACGAGTTTGCCCACAGATACGCGG	VNEFAHRYA	27
		Aj-A1-NGSP	ATATTCGGGGAGCCACTCGACC	IFGEPLD	22
	ace2	Aj-A2-GSP	ACAGCAGATGATCGGACAAATGGTTGGCG	QQMIGQMVG	29
		Aj-A2-NGSP	CTCAACTAATGACAGACCATGGGTC	AQLMTDHGS	25

Table 3. Continued

Name			Sequence	Amino acid sequence	Length (bp)
Ctenolepima longicandata					
5'RACE	*ace1*	Zy-A1-RT	GGAGAGCCATTAATTGG	QLMAL	17
		Zy-A1-S1	TGTGTCTATGCAATACAGGGTCG	VSMQYRV	23
		Zy-A1-S2	TTCTTTGATACCTCGGACGTTCC	FFDTSDVP	23
		Zy-A1-A1	CCAGAGTTTTATGATCATACACATCC	DVYDHKTL	26
		Zy-A1-A2	ATATATCCATACCATTACTGCCGAG	SAVMVWIY	25
	ace2	Zy-A2-RT	TTGTTCTTCAACCGCTC	ERLKN	17
		Zy-A2-S1	CGCTTCTATGCAATATCGTGTTGG	ASMQYRV	24
		Zy-A2-S2	GAAGAAGCCCCAGGTAACGTG	EEAPGNV	21
		Zy-A2-A1	CCACCGTAGATCCACACCAAG	LVWIYG	21
		Zy-A2-A2	ACCCAAAGATTGAGATATAGACAGTC	DCLYLNLWV	26
3'RACE	*ace1*	Zy-A1-GSP	GATCAGCAGGAAATCCATGGCCTAGTTGG	SAGNPWPSW	29
		Zy-A1-NGSP	AACCCGGAAAAGAACTACTTATCAAGC	NPEKNYLSS	27
	ace2	Zy-A2-GSP	AGTATACTAACTGGGAGAACCTCGAAGATGG	YTNWENLED	31
		Zy-A2-NGSP	CAAACTACTTTGCACAAACCGTGGC	NYFAQTVA	25

cycles (95°C for 30 sec, 59~61°C for 30 sec, 68°C for 2~3 min) and final 1 cycle of single extension (68°C for 2~3 min). The 3' RACE reactions for insect *ace1* and *ace2* were performed using 3' RACE ready-cDNA synthesized by 3'-RACE CDS primer A and universal or nested primer A mix (offered from kit). The thermal program for the primary 3' RACE PCR was 95°C for 3 min, followed by 5 cycles of 95°C for 30 sec, 70°C for 3 min, 5 cycles of 95°C for 30 sec, 70°C for 30 sec, 72°C for 3 min, 25 cycles of 95°C for 30 sec, 68°C for 30 sec, 72°C for 2~3 min and final 3 min at 72°C. The secondary PCR was conducted by 1 cycle of 95°C for 3 min, followed by 35 cycles of 95°C for 30 sec, 68°C for 30 sec, 72°C for 3 min and 1 cycle of 72°C for 3 min. The 5' RACE and 3' RACE PCR products were extracted using gel extraction kit (QIAgen, Valencia, CA, USA) and then cloned into pGEM®-T easy vector (Promega, Madison, WI, USA) and sequenced. The assembly of sequence contigs was performed by Lasergene Seqman Pro software (DNASTAR, Madison, WI, USA)

2.5. Phylogenetic analysis

A total of 181 *ace* genes and 11 *ace*-like genes sequences were used for phylogenetic analysis. Alignments of *ace* and *ace-like* genes were performed by the alignments of nucleotides using method of MUSCLE in Molecular Evolutionary Genetic Analysis (MEGA) v5.0 program (parameter default

condition) (Tamura et al., 2011). Signal peptide and C-terminal highly variable amino acid sequence of the alignments were trimmed. As a result, these alignments have been included in the core region (amino acid residues 65–520) of *D. melanogaster* AChE (NM_057605.5). The evolutionary history was inferred on the basis of the Neighbor-Joining method, which was conducted using MEGA v5.0 program. The evolutionary distances were computed using the Poisson correction method with 2000 bootstrap replicates. All ambiguous positions were removed for each sequence pair.

2.6. Estimation of evolutionary rate

Evolutionary rates of *ace1* and *ace2* were estimated in order to determine which *ace* has retained relatively more essential function in each insect group. Since there is no available information on the ancient *ace* gene sequences of the putative common ancestor that existed before the *ace* duplication event, the average evolutionary rate of respective *ace* gene from individual insect groups (Coleoptera, Hemiptera, Diptera, Hymenoptera and Lepidoptera) was estimated by calculating the nucleotide substitution number per site (*d,* evolutionary distance) between the *ace1* or *ace2* sequences of each insect species (a total of 33 species: 3 Coleoptera, 5 Hemiptera, 8 Diptera, 8 Hymenoptera, and 9 Lepidoptera) in a target insect group and the corresponding *ace* sequences of several out-group

species (i.e., *Orchesella villosa*, *Haslundichilis viridis*, *Ctenolepima longicandata*, *Ecdyonurus levis*, *Lyriothemis pachygasta*) that are apparently more ancient compared to insects. The *d* value was calculated using the kimura's 2-parameter method with SEs being calculated by the MEGA v5.0 program with 500 bootstrap replicates. Within each insect group, all *d* values calculated per each out-group were combined to obtain mathematical mean.

In case of social Hymenoptera, since many bees and wasps were found to use *ace2* as the main catalytic enzyme (i.e., more essential role) than *ace1*, suggesting a possibility of functional transition from *ace1* to *ace2* within these insect groups (Kim et al., 2012), it was desirable to compare the difference in evolutionary rate between *ace1* and *ace2* at a narrower time scale. To do this, the average evolutionary distance of Hymenoptera *ace* was compared with that of sister group *ace*, both of which were calculated by using *H. viridis* (Archaeognatha), the closest putative common ancestor, as an out-group. The evolutionary distances (*d*) of two compared groups [i.e., Hymenoptera (A) and sister-group (B: Coleoptera, Diptera, Hemiptera and Lepidoptera)] after divergence from their putative common ancestor, with reference to the out-group (O, *H. viridis*), was estimated by the equation: $d_A = (d_{AB} + d_{AO} - d_{BO})/2$ and $d_B = (d_{AB} + d_{BO} - d_{AO})/2$, respectively (Eyre-Walker and Gaut 1997; Yue et al., 2010).

Finally, to determine the selection pressure given to insect *aces,*

nonsynonymous (*dN*) and synonymous (*dS*) substitution rates (*dN/dS* ratio or Ka/Ks ratio) between the *ace1* or *ace2* sequences of each insect species (a total of 33 species: 3 Coleoptera, 5 Hemiptera, 8 Diptera, 8 Hymenoptera, and 9 Lepidoptera) in a target insect group and corresponding *ace* sequences of several out-group species (i.e., *O. villosa*, *H. viridis*, *C. longicandata*, *E. levis*, and *L. pachygasta*) were calculated according to the method of Yang, Z. and Nielsen, R (2000) using the Ka/Ks calculator (Zhang et al., 2006) and then average values were obtained.

3. Results and discussion

3.1. The origin of insect *ace1* and *ace2*

Phylogenetic analysis was performed for *ace-like* genes from all the lower animals with their genome sequenced together with all known arthropod *ace1* and *ace2*, including those from a number of insects that were newly cloned (Fig. 1). The resulting tree consisted of five clades [i.e., clade of *ace1* from Arthropoda, Nematoda, Annelida and Mollusca (*ace1* clade); clade of *ace2* from Arthropoda and Nematoda (*ace2* clade); mixed clade of *ace2*, *ace3* and *ace4* from Annelida and Mollusca; clade of Platyhelminthes *ace*; and clade of *ace-like* from lower animals, including Choanoflagellates, Porifera, Placozoa, Cnidaria and Ctenophora] (Fig. 1). The *ace1* clade is consisted of the arthropod *ace1* genes along with those from Annelida, Mollusca, Nematoda, and Ixodida. Within the *ace1* clade, however, the paraphyletic relationship between the Arthropod *ace1* group and the Annelida/Mollusca *ace1* and Nematoda *ace1* groups was not supported by bootstrap value (= 30~44) whereas the ancestral relationship of the Ixodida *ace1* group to the Arthropoda *ace1* group was determined to be reliable (bootstrap value = 84). The *ace2* clade is solely comprised of the Arthropod a*ce2* group and nematode *ace2/3/4* group. The *ace1* and *ace2* clades formed a larger monophyletic clade along with the Annelida/Mollusca *ace* clade. However, the

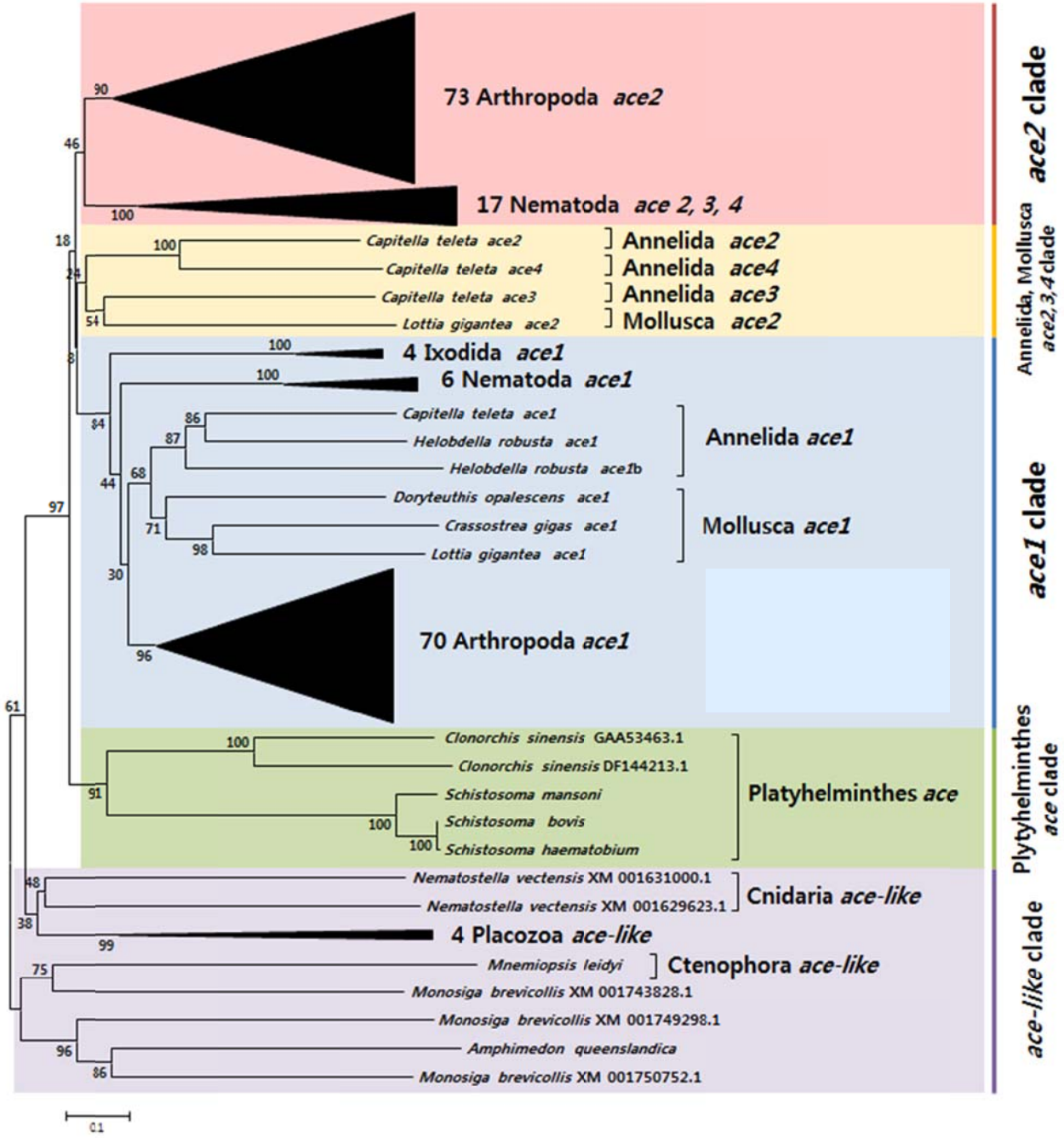

73 Arthropoda ace2
17 Nematoda ace 2, 3, 4
Capitella teleta ace2
Capitella teleta ace4
Capitella teleta ace3
Lottia gigantea ace2
Annelida ace2
Annelida ace4
Annelida ace3
Mollusca ace2
4 Ixodida ace1
6 Nematoda ace1
Capitella teleta ace1
Helobdella robusta ace1
Helobdella robusta ace1b
Annelida ace1
Doryteuthis opalescens ace1
Crassostrea gigas ace1
Lottia gigantea ace1
Mollusca ace1
70 Arthropoda ace1
Clonorchis sinensis GAA53463.1
Clonorchis sinensis DF144213.1
Schistosoma mansoni
Schistosoma bovis
Schistosoma haematobium
Platyhelminthes ace
Nematostella vectensis XM 001631000.1
Nematostella vectensis XM 001629623.1
Cnidaria ace-like
4 Placozoa ace-like
Mnemiopsis leidyi
Ctenophora ace-like
Monosiga brevicollis XM 001743828.1
Monosiga brevicollis XM 001749298.1
Amphimedon queenslandica
Monosiga brevicollis XM 001750752.1
ace2 clade
Annelida, Mollusca ace2,3,4 clade
ace1 clade
Plytyhelminthes ace clade
ace-like clade
0.1

Figure 1. Phylogenetic relationships of the insect *aces* and *ace-like* genes from other metazoans. The optimal tree with the sum of branch length (32.22093198) is shown. The percentages of replicate trees, in which the associated taxa clustered together in the bootstrap test (2000 replicates), are shown next to the branches. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The total dataset for phylogenetic analysis was comprised of 181 *aces* and 11 *ace-like* sequences (Nematoda: *Caenorhabditis brenneri, Caenorhabditis briggsae, Caenorhabditis elegans, Caenorhabditis remanei, Dictyocaulus viviparus, Bursaphelenchus xylophilus and Heligmosomoides polygyrus bakeri*; Annelida: *Capitella teleta, Helobdella robusta*; Mollusca: *Crassostrea gigas, Lottia gigantea* and *Doryteuthis opalescens*; Platyhelminthes: *Schistosoma mansoni, Schistosoma haematobium, Schistosoma bovis*, and *Clonorchis sinensis*; Ctenophora: *Mnemiopsis leidyi*; Cnidaria: *Nematostella vectensis*; Placozoa: *Trichoplax adhaerens*; Porifera: *Amphimedon queenslandica*; Choanoflagellates: *Monosiga brevicollis*)

paraphyly between the *ace1* and *ace2* clades was not supported by bootstrap value (= 18), not allowing the accurate prediction of phylogenetic origin of the Arthropod *ace1* and *ace2*. Nevertheless, the Platyhelminthes *ace* clade clearly formed a monophyly with the clades containing the Arthropod *ace1* and *ace2* clades (bootstrap value = 97), suggesting that the last common ancestor of *ace1* and *ace2* shared its origin with those of Platyhelminthes. This finding further implied that the lineage of Arthropod *ace1* and *ace2* has undergone a divergent evolution along with those of Platyhelminthes. Several duplicate copies of *ace-like* gene were found in the genomes of some lower animal groups investigated (Cnidaria, Placozoa, and Ctenophora), indicating the occurrence of more ancient independent duplications of *ace-like* gene in these lineages. However, these duplication events were not directly related with that of arthropod *ace1* and *ace2*, suggesting that these *ace-like* of lower animals are not the direct origin of *ace1* and *ace2*.

Previous studies regarding *ace* duplication suggested that *aces* from Arthropods and Nematodes originated from a duplication occurred before the split of Ecdysozoa (Pezzementi and Chatonnet 2010; Weill et al., 2002). Taken together, since the Annelida/Mollusca *ace* clade exhibited a monophyletic relationship with the Arthropoda *ace* clade, though it was not supported with high bootstrap values, it can be suggested that the *ace* duplication event resulting in the split of the *ace1*

and *ace2* clades occurred after the divergence of Ecdysozoa and Lophotrochozoa from their Protostomian common ancestor but before the split of Ecdysozoa into their descendents.

3.2. Estimation of evolutionary rate of insect *ace1* and *ace2*

Comparison of the average evolutionary distance (*d*) of two *aces* from five insect groups (Coleoptera, Diptera, Hemiptera, Hymenoptera and Lepidoptera) relative to those from insect common ancestors (Collembola, Archaeognatha, Zygentoma, Ephemeroptera, and Odonata) revealed that *ace1* has evolved with a significantly slower rate compared to *ace2*. The *d* values of *ace2* were 1.16~1.51-fold higher than those of *ace1*. It is widely known that the evolutionary rate of genes having essential functions is slower than that of genes with less essential functions (Wilson et al., 1977). Therefore, given that the primary function of AChE is associated with the neuronal synaptic transmission, *ace1* with a slower rate of evolution appears to have retained more essential neuronal functions (i.e., synaptic transmission) than *ace2* (Fig. 2). This result is in good agreement with previous reports showing that *ace1* encodes the more essential AChE as judged by its catalytic efficiency and abundance in several insect species except Hymenoptera (Baek et al., 2005; Gao et al., 2002; Kim et al., 2006; Kim and Lee 2013; Lee et al., 2006; Li and Han 2002; Nabeshima et al., 2003). The slower

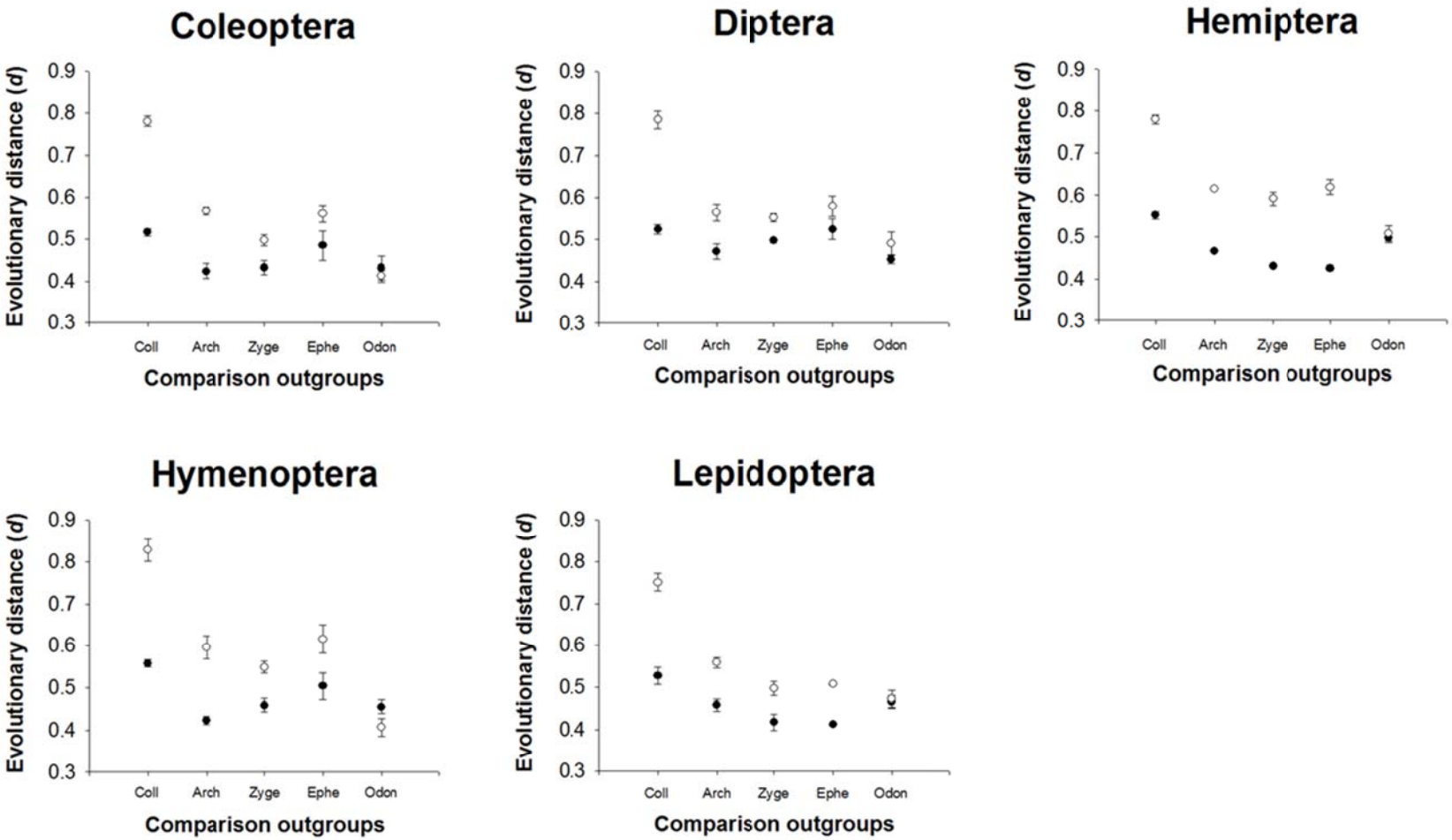

Coleoptera
Diptera
Hemiptera
Hymenoptera
Lepidoptera
Evolutionary distance (d)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
Coll
Arch
Zyge
Ephe
Odon
Comparison outgroups

Figure 2. Estimated evolutionary distance (*d*) of insect two *ace* genes when compared with several insect common ancestors. Scatter plots of *d* values between two *ace* genes of several insect groups and those of several common ancestral out-groups. Black circles and white circles indicate *ace1* and *ace2*, respectively. Y-axis indicates *d* values. X-axis indicates various insect common ancestors: Coll, Collembola (*Orchesella villosa*); Arch, Archaeognatha (*Haslundichilis viridis*); Zyge, Zygentoma (*Ctenolepisma longicaudata*); Ephe, Ephemeroptera (*Ecdyonurus levis*); Odon, Odonata (*Lyriothemis pachygastra*)

evolutionary rate of *ace1* was more apparent when *d* values were estimated from relatively more ancient common ancestors (Collembola, Archaeognatha, Zygentoma, Ephemeroptera). In contrast, the tendency of slower *ace1* evolutionary rate became less apparent when estimated from a relatively more recent common ancestor (Odonata) of insect, in which the *d* values of *ace1* were slightly lower than those of *ace2* but the differences were statistically insignificant.

In summary, overall *d* values of both *ace1* and *ace2* showed decreasing tendencies as the evolutionary distance between the test insect group and the putative common ancestor out-group decreased, and this tendency was true for all groups of insect examined. However, the decreasing rate of *d* values was more prominent in *ace2* than in *ace1*, suggesting that *ace2* has gained some essentiality or adaptive significance, perhaps by being specialized to acquire certain functions during evolution. No apparent difference was noticed in cross-comparison of *d* values between different insect orders, suggesting that individual groups of insect have retained similar levels of evolution speed in respective *ace1* and *ace2.* Taken together, these findings indicate that overall evolutionary rate of insect *ace1* gene has been constantly lower compared to that of *ace2* in all insect lineages examined but both of their values have become smaller over the evolutionary time, indicating the presence of purifying selection pressure, which is relatively more rapidly increasing during the evolution of *ace2* compared to *ace1*.

To confirm whether the trend found in evolutionary rate of *aces* is in accordance with the concept of selection pressure, the *dN/dS* ratio was estimated and compared for both *ace1* and *ace2* (Fig. 3). The resulting *dN/dS* ratios were in the range of 0.03~0.11, indicating that both *ace1* and *ace2* have been under purifying selection. Without any exception, all *dN/dS* ratios for *ace1* were significantly smaller than those for *ace2*. This finding matches well with the evolutionary distance estimation and further supports that *ace1* has evolved to retain more essential function than *ace2*. The previous studies based on genome analysis also suggested that essential genes in an organism have undergone a higher purifying selection pressure (i.e., low *dN/dS* value) than non-essential genes (Jordan et al., 2002).

Based on a finding that many bees and wasps were found to use *ace2* as the main catalytic enzyme (i.e., more essential role) than *ace1*, a possibility of functional transition from *ace1* to *ace2* within these insects was suggested (Kim and Lee 2013). To determine whether any changes in the evolutionary rate of *ace2* can be detected in these Hymenopteran insects at a narrower time scale, the evolutionary rate of each ace in some Hymenopteran insects was pairwisely compared with that in several sister group insects. In case of *ace1*, slopes of the regression line were in the ranges of 1.00-1.26, suggesting that *ace1* in Hymenoptera has evolved at relatively slower rates compared to other sister

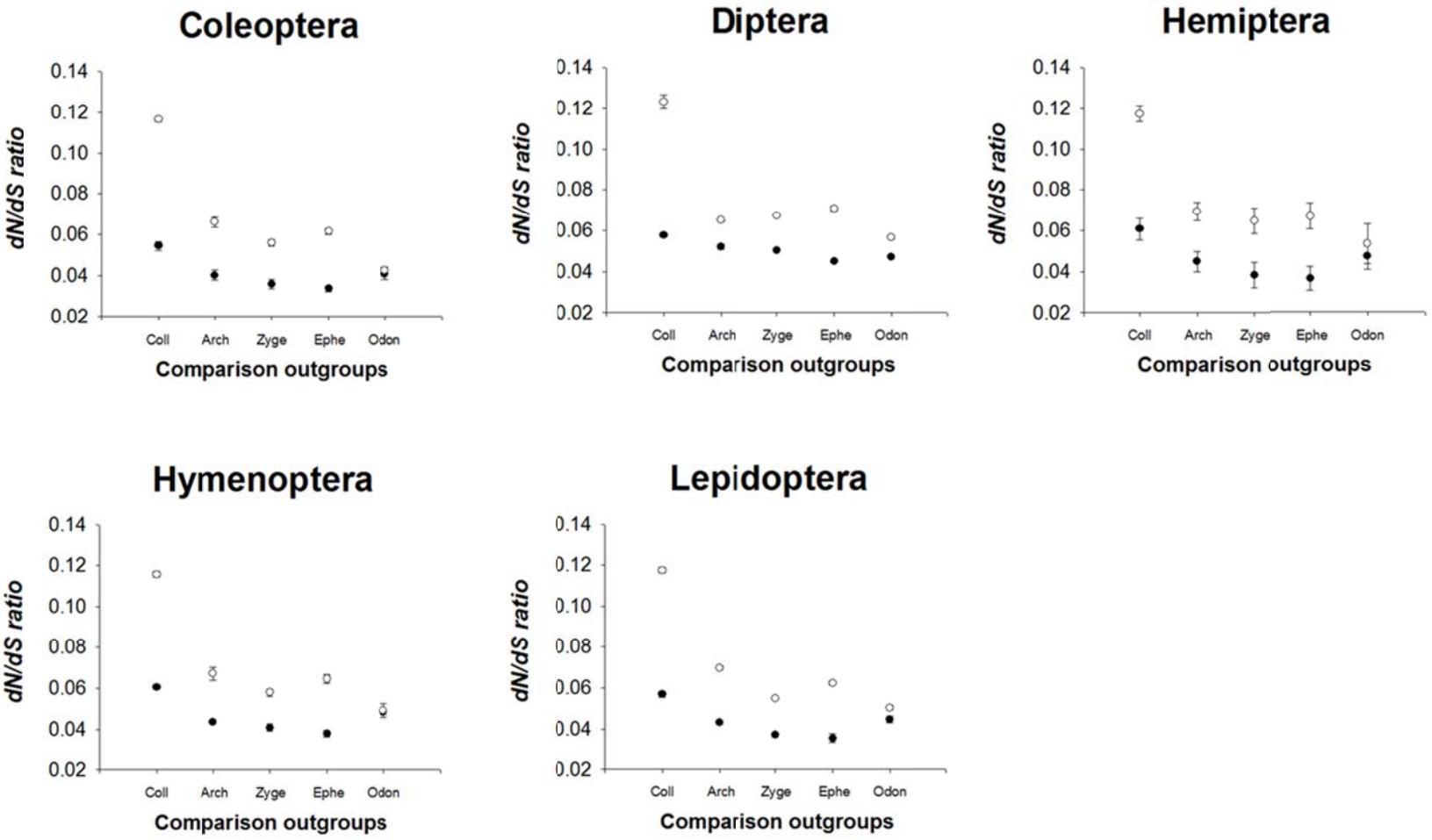

Coleoptera
Diptera
Hemiptera
Hymenoptera
Lepidoptera
dN/dS ratio
0.14
0.12
0.10
0.08
0.06
0.04
0.02
Coll
Arch
Zyge
Ephe
Odon
Comparison outgroups

Figure 3. Estimated selection pressure (*dN/dS* ratio) of insect two aces when compared with several common ancestors. Scatter plots of selection pressure in aces of several insect groups against that in aces of common ancestral outgroups. Black circles and white circles indicate *ace1* and *ace2*, respectively. Y-axis represents *dN/dS* ratios. X-axis indicates various insect common ancestors: Coll, Collembola (*Orchesella villosa*); Arch, Archaeognatha (*Haslundichilis viridis*); Zyge, Zygentoma (*Ctenolepisma longicaudata*); Ephe, Ephemeroptera (*Ecdyonurus levis*); Odon, Odonata (*Lyriothemis pachygastra*)

groups examined. In contrast, the *ace2* regression lines exhibited slopes of smaller than 1 (0.84-0.89), indicating the higher (compared to Coleoptera, Diptera and Lepidoptera) or identical (to Hemiptera) evolutionary rate of *ace2* compared to *ace1*.

Taken together, even in the comparison with the sister group insect species, *ace2* in the Hymenopteran insects still showed significantly higher evolutionary rate than *ace1* (Fig. 4). This finding further suggests that the specific event of putative functional transition from *ace1* to *ace2* observed in some Hymenopteran insects cannot be explained by the degree of evolutionary rate difference between two *aces* and is not due to the gradual accumulation of nucleotide substitution over even a relatively short evolutionary time (i.e., during the parallel evolution within Class Insecta). Therefore, the putative functional transition from *ace1* to *ace2* appears to have been driven more likely by a small number of non-synonymous mutations in *ace1* that result in the loss of AChE1 catalytic activity based on the finding that AChE1 possesses little catalytic activity in these Hymenopteran insects. In this study, because the evolutionary rate was estimated on the basis of *d* values (i.e., the nucleotide substitution number per site) or *d*N/*d*S ratios [i.e., ratio of nonsynonymous substitution rate per site (*d*N) to synonymous substitution rate per site (*d*S)] that were calculated from entire gene sequences without any weighting, a few but functionally significant changes in *ace*

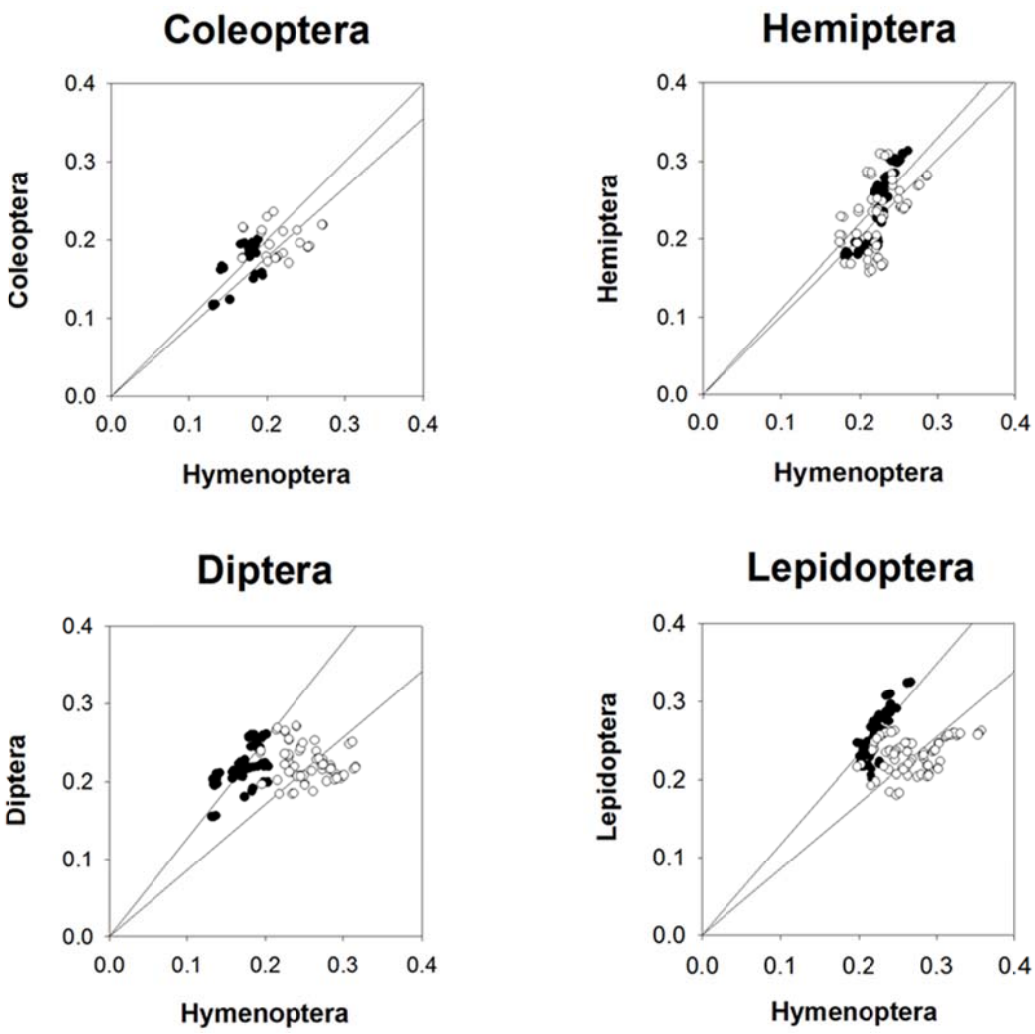


Figure 4. Scatter plots of evolutionary rates in *aces* of Hymenoptera against those of sister insect groups (Coleoptera, Hemiptera, Diptera and Lepidoptera). X-axis indicates evolutionary rates of aces in Hymenoptera whereas Y-axis indicate those in other sister insect groups. Black and white circles indicate *ace1* and *ace2*, respectively.

sequences could not be detected since they were counted as only a tiny part of non-synonymous nucleotide substitutions. With this in mind, the event of putative functional transition of *ace* in some Hymenopteran appears to have occurred in relatively recent time by only a few numbers of mutations and the similar but independent events also likely have occurred in several lineages of insects as many cases of putative functional transition of AChE1 to AChE2 were found in various insect species (Kim and Lee 2013). Thus, in order to investigate the nature of this event precisely, it would be necessary to identify and characterize in detail the non-synonymous substitutions causing dramatic alterations of AChE function.

In summary, *ace1* rather than *ace2* has been preserved to retain the relatively more essential function (i.e., synaptic transmission) following the split from their common ancestor when analyzed in a large time scale of evolution. Considering the origin of duplication, likely occurred after the split of Protostomian into its descendents with the well developed neural system, and the highly conserved nature of *ace1* during evolution, the ancestral *ace* of common ancestor with the neuronal function may have possessed properties that are not much different from those of present *ace1*. The presumably random event of functional transition from *ace1* to *ace2* appears to have occurred in relatively recent time by only a few numbers of mutations as it could not be detected by comparing their evolutionary rates at a relative large scale of evolutionary time.

CHAPTER 2.

Structural conservation and differentiation of two insect acetylcholinesterases during evolution

Structural conservation and differentiation of two insect acetylcholinesterases during evolution

Abstract

Insects possess two distinct acetylcholinesterases (AChE1 vs. AChE2), which are encoded by two paralogous loci originated from duplication. The deduced amino acid sequence identity between AChE1 and AChE2 is approximately less than 40% in average, indicating a long independent evolutionary history after duplication. However, kinetic analyses of several insect AChEs revealed that both AChE1 and AChE2 retain common catalytic properties of AChE but subtle kinetic differences also exist between these two AChEs. In this chapter, to understand how selection pressure has shaped the protein structure of AChEs and affected their functional differentiation during evolution, I measured and compared the nucleotide diversity (Pi) and amino acid site-specific selection pressure between AChE1 and AChE2 from various insects. The amino acid sequence comparison between AChE1 and AChE2 from a wide variety of insect taxa revealed a high degree of sequence conservancy in the functionally crucial domains, suggesting that presence of strong purifying selection pressure over these essential residues.

Interestingly, the EF-hand motif was mostly found in the AChE1 lineage but not in AChE2. In contrast, LRE-motif was partially conserved in the AChE2 lineage but not in AChE1. In addition, the AChE2-specific insertion domain appeared to have been introduced relatively more recently, perhaps during the radiation of insects, after the duplication. The comparison of five essential domains [i.e., the catalytic anionic site (CAS), peripheral anionic site (PAS), acyl binding pocket (ABP), oxyanion hole and catalytic triad] in the active-site gorge identified unique differences in amino acid residues of the PAS (Asp72 vs. Tyr72, Tyr121 vs. Met121) and the ABP (Cys288 vs. Leu288) between AChE1 and AChE2. Three-dimensional homology modeling of active-site gorge from insect AChEs with a particular focus on the PAS revealed that a subtle but consistent structural alteration in the active-site gorge topology was caused by the PAS amino acid substitution, likely resulting in the functional differentiation between two AChEs. Although *ace1* appears to have evolved at significantly slower rates to retain its essential function, it is likely that a few specific amino acid substitutions in *ace1* causing a dramatic reduction of enzyme activity may have occurred locally in more recent time and resulted in the functional transition from AChE1 to AChE2 as observed in some Hymenopteran insects.

1. Introduction

Acetylcholinesterase (AChE, EC 3.1.1.7) is the key enzyme for regulation of synaptic transmission by hydrolyzing the neurotransmitter acetylcholine at cholinergic synapses and neuromuscular junctions in the cholinergic nervous system (Rosenberry 1975). In addition to its typical neuronal function of acetylcholine hydrolysis, AChE is known to play other non-synaptic functions, including neurite outgrowth, synapse formation (Olivera et al., 2003), modulation of glial activation, tau phosphorylation (Ballard et al., 2005) and learning/memory (Gauthier et al., 1992; Shapira et al., 2001). AChE also shows high levels of sequence homology to other neuronal proteins, such as neurotactin and neuroligin (Johnson and Moore 2006).

In insects, two AChEs (AChE1 and AChE2) are encoded by two paralogous loci (*ace1* and *ace2*, respectively), which were originated from a duplication event that occurred long before the radiation of insects (Weill et al., 2002). The most recent phylogenetic analysis based a large scale data set, including several lower animal lineages, suggested that the duplication event have occurred after the divergence of Ecdysozoa and Lophotrochozoa from their Protostomian common ancestor but before the split of Ecdysozoa into their descendents (This study, Chapter 1). In addition, it was proposed that respective lineages of arthropod *ace1*

and *ace2* have undergone a divergent evolution along with those of Platyhelminthes. Comparison of evolutionary rate between the two *aces* in different insect groups revealed that *ace1* has evolved with a significantly slower rate compared to *ace2*, suggesting that the indispensable functions of synaptic transmission has been preserved in the *ace1* lineage (Chapter 1). In contrast, the *ace2* lineage showed a significantly higher evolutionary rate, implying its primary nature as the surplus copy that can be functionally diverged during evolution. Nevertheless, several cases of putative functional transition from AChE1 to AChE2 were reported in some insect species, particularly in several Hymenopteran insects (Kim et al., 2012; Kim and Lee 2013), where AChE2 were responsible for synaptic transmission whereas AChE1 retained little catalytic activity. Furthermore, the putative functional transition was presumed to be random event that occurred in relatively recent time by only a few numbers of mutations as it could not be detected by comparing their evolutionary rates at a relative large scale of evolutionary time. Therefore, it would worth investigating the *ace*-type specific non-synonymous nucleotide substitutions that are associated with dramatic functional alteration of AChE across a wide variety of insect species.

The deduced amino acid sequence identity between *ace1* and *ace2* is approximately less than 40% in average when the entire gene sequences were

compared, revealing that a substantial amount of non-synonymous substitutions has been accumulated in both lineages of *ace1* and *ace2* following duplication. Nevertheless, previous investigations on the kinetic properties of insect AChEs revealed that AChE2 as well as AChE1 generally retained high levels of catalytic efficiency although some differences in substrate and inhibitor kinetics were observed between AChE1 and AChE2 (Ilg et al., 2010; Kim et al., 2012; Kim et al., 2010; Lang et al., 2010). This finding suggests that the nucleotide sequences encoding amino acid residues of functional importance have been preserved in both *ace1* and *ace2,* and this notion is supported by the fact that the amino acid sequence identity between *ace1* and *ace2* increased to > 80% when only the functionally crucial domains were compared.

AChE belongs to the α/β hydrolase fold enzyme family, all of which possess a typical catalytic triad in the catalytic site that is composed of highly conserved residues of Ser (or Cys, Asp), Asp (or Glu) and His (Augustinsson 1949). In addition to this common catalytic triad, there are several functionally essential domains in the catalytic or acylation site of AChE, including the catalytic anionic site (CAS), peripheral anionic site (PAS), acyl binding pocket (ABP), and oxyanion hole. The catalytic site of AChE is located near the bottom of a deep and narrow cavity lined predominantly with 14 aromatic residues, which was named "active-site gorge" (Massoulíe et al., 1991). The PAS is located near the active-

site gorge entrance, approximately 20 Å apart from the catalytic triad (Sussman et al., 1991). It serves as a binding site for substrate and inhibitor at the first step of catalytic pathway (Mallender et al., 2000). The CAS, also known as the "anionic" subsite, is located near the bottom of the active-site gorge. It is composed of several aromatic residues, including Trp84 (in *Torpedo californica* AChE numbering hereafter), which has been known as choline binding site (Silman and Sussman 2005). The ABP is consisted of Trp233, Phe288, Phe290, Phe331 and Val400 residues in (Harel et al., 2000) and is known to play a role in determining the substrate/inhibitor specificity of AChE (Ariel et al., 1998; Kaplan et al., 2001). The oxyanion hole is located between the catalytic triad and choline binding site and is formed by three peptidic NH groups of Gly118, Gly119 and Ala201 residues (Harel et al., 1996), unlike the oxyanion hole of serine and cysteine proteases, which is formed by two peptidic NH groups (Menard and Storer 1992). It is required for stabilizing the tetrahedral transition state intermediate, by allowing the NH groups to form hydrogen bond with the carbonyl oxygen of substrate or inhibitor (Gerlt and Gassman 1993; Zhang et al., 2002). The catalytic triad of AChE, consisting of Ser200, Glu327, and His440 residues, is positioned near the bottom of the active-site gorge and involved in the hydrolysis of substrate (Sussman et al., 1993; Zhang et al., 2002).

Comparison of deduced amino acid sequences between AChE1 and AChE2,

with a particular focus on such functionally crucial domains, would identify any conserved changes that have been fixed in each lineage of AChE. In addition, comparison of three-dimensional (3D) structures of AChE1 and AChE2, if available, would enable to infer the functional differences between these two insect AChEs. The *D. melanogaster* AChE (i.e., AChE2) was crystallized and its 3D structure was determined (Harel et al., 2000) but no X-ray crystallographic analysis for insect AChE1 has been conducted yet. Nevertheless, recent advances in the 3D structural homology modeling enable to predict the tertiary structure of a protein from its amino acid sequences if one or more experimentally determined 3D structures of the homologous protein template are available. By using one of several crystal structures of AChE (i.e., *Torpedo* AChE, *Drosophila* AChE, human AChE and mouse AChE) as a template, therefore, the 3D structure of any insect AChE can be predicted by homology modeling (Arnold et al., 2006). Moreover, homology modeling can also be useful to investigate subtle structural differences between AChE1 and AChE2, whose typical structures have not been resolved experimentally.

In this chapter, to identify any conserved amino acid sequence motifs, domains, active-site gorge compositions and surface structures that are specific to each lineage of AChE1 and AChE2 and thereby to deduce any functional diversification occurred after duplication, amino acid sequence alignment was

conducted for the AChE1 and AChE2 from representative groups of insect (Coleoptera, Diptera, Hemiptera, Hymenoptera and Lepidoptera) and several species of common ancestral Hexapoda. In addition, 3D structural homology modeling of representative insect AChE1 and AChE2 was performed to infer the functional role of AChE type-specific amino acid substitutions, particularly in the active-site gorge.

2. Materials and methods

2.1. Alignment of amino acid sequence

AChE1 and AChE2 sequences were obtained from 65 species of insect and 2 species of hexapod and used for the alignment based on the method of MUSCLE in the Molecular Evolutionary Genetic Analysis (MEGA) v5.0 program (parameter default condition). After trimming out the signal peptide sequences and highly variable C-terminal amino acid sequences, only the core region (corresponding to amino acid residues 65–520 of *D. melanogaster* AChE; NM_057605.5) was aligned and used for finding sequence differentiation as well as for estimating the nucleotide diversity (Pi) and *dN/dS* ratio value per amino acid substitution of active site gorge between insect AChE1 and AChE2.

2.2. Identification of motif and domains

Pi and *dN/dS* ratio value per amino acid substitution in the active-site gorge of insect AChE1 and AChE2 were estimated by DnaSP v5.10.01 and selecton (http://selecton.tau.ac.il/index.html) programs, respectively. These values were used to determine any difference in the conserved position and other functional site between AChE1 and AChE2. Also, Sequence Identity and Similarity Service (SIAS) (http://imed.med.ucm.es/Tools/sias.html) and HHblit (http://toolkit.lmb.

uni-muenchen.de/hhblits) were used to detect the differences in the motifs, domains, conserved sequences and sequence similarity between two AChEs from several insects.

2.3. 3D structure modeling

The 3D structure of AChE created by the automated and aligned comparative protein modeling program of Swiss model server (http://swissmodel.expasy.org/) was used for the comparison of differences in the protein surface and active-site gorge topology between insect two AChEs. The homology models of both AChE1 and AChE2 were generated using the *Torpedo* AChE (PDB: 2ACEA) as templates. Then, the comparison of the protein surface and active-site gorge structure between insect AChEs was performed by UCSF Chimera an extensible molecular modeling system v. 1.8.1 (University of California, San Francisco, CA, USA).

3. Results and discussion

3.1. Identification and isolation of the AChEs from five hexapods

Partial cDNA encoding *ace1* and *ace2* were isolated from the head and thorax tissues of insect common ancestors (*Lyriothemis pachygastra*, *Lypace1* and *Lypace2*; *Ecdyonurus levis*, *Eclace1* and *Eclace2*; *Ctenolepisma longicaudata*, *Ctlace1* and *Ctlace2*; *Haslundichilis viridis*, *Havace1* and *Havace2*; *Tomocerus kinoshitai, Tokace1* and *Tokace2*). When the deduced amino acid sequences of *ace1* and *ace2* from the insect common ancestors were aligned with previously published insect two *ace* genes (Fig. 1), *Lypace1*, *Eclace1*, *Ctlace1*, *Havace1* and *Tokace1* exhibited high degrees of sequence identity to *ace1* of *Tribolium castaneum* (74.6%, 83.5%, 82.5%, 81.2% and 72.1%), *Aphis gossypii* (69%, 77.7%, 77.3%, 76.2% and 67.3%), *Bombyx mori* (74%, 82.9%, 81%, 79.4% and 71.7%), *Apis mellifera* (71.9%, 81%, 78.3%, 77.5% and 70.4%), *Aedes aegypti* (74.2%, 78.1%, 75.8%, 75% and 71.2%), respectively. Deduced amino acid sequences of *Lypace2*, *Eclace2*, *Ctlace2*, *Havace2* and *Tokace2* showed high degrees of identity to *ace2* of *T. castaneum* (70.4%, 66.9%, 66.5%, 64.6% and 47.1%), *A. gossypii* (62.7%, 61.5%, 60.4%, 59.8% and 44.6%), *B. mori* (66.9%, 66%, 67.5%, 62.3% and 45.8%), *A. mellifera* (67.1%, 64.8%, 65%, 62.9% and 46.2%), *A. aegypti* (61.9%, 60.2%, 60.4%, 60.8% and 44.6%). In summary, when

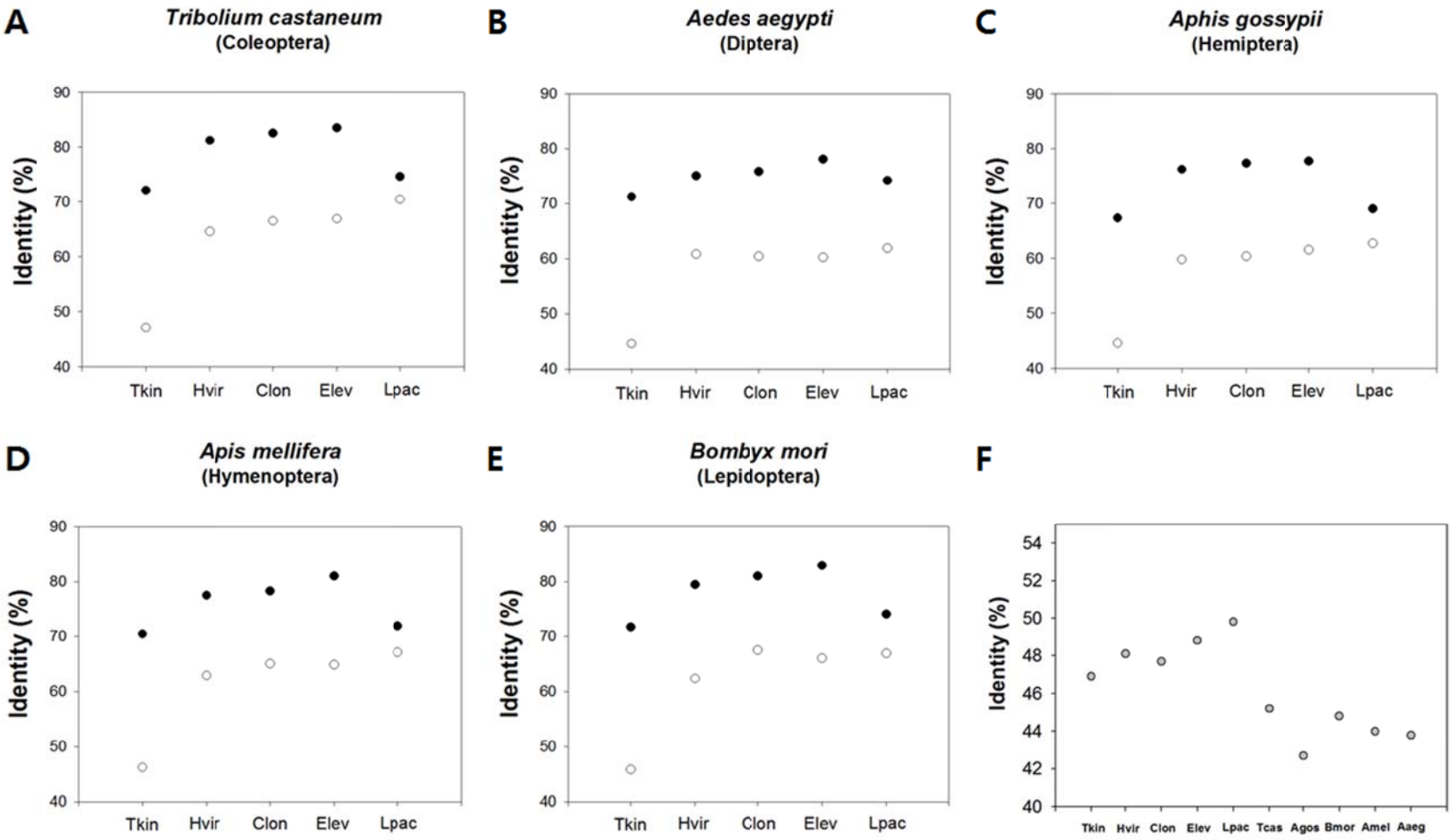

A
Tribolium castaneum
(Coleoptera)
B
Aedes aegypti
(Diptera)
C
Aphis gossypii
(Hemiptera)
D
Apis mellifera
(Hymenoptera)
E
Bombyx mori
(Lepidoptera)
F
Identity (%)
Tkin
Hvir
Clon
Elev
Lpac
Tcas
Agos
Bmor
Amel
Aaeg

Figure 1. Pairwise comparisons of deduced amino acid sequences of *aces* between five hexapod species and several insect species. Alignments were conducted by MUSCLE Method in the MEGA v5.0 and amino acid identities were calculated by SIAS server. (A) - (E) are the results of *ace1* and *ace2* from several insects when compared with those from five hexapods. (F) is the deduced amino acid sequence identities between *ace1* and *ace2* from several insects. Black and white bars indicated *ace1* and *ace2* gene**s**, respectively. Tcas, *Tribolium castaneum* (HQ260968.1 and HQ260969.1); Agos, *Aphis gossypii* (AB158637.1 and AF502081.1); Bmor, *Bombyx mori* (NM_001043915.1 and NM_001114641.1); Amel, *Apis mellifera* (XM_393751.4 and NM_001040230.1); Aaeg, *Aedes aegypti* (XM_001656927.1 and XM_001655818.1) Tkin, *Tomocerus kinoshitai*; Hvir, *Haslundichilis viridis*; Clon, *Ctenolepisma longicaudata*; Elev, *Ecdyonurus levis*; Lpac, *Lyriothemis pachygastra*

compared to the insect common ancestral species, insect *ace1* showed significantly higher deduced amino acid sequence identity than *ace2*, supporting that AChE1 is more conserved than AChE2 in insects

A moderate levels of deduced amino acid sequence identity (46.9~49.8%) were observed between *ace1* and *ace2* in the hexapod and insect species examined. However, these levels of sequence identity are higher than those between *ace1* and *ace2* in the insect species investigated (42.7~45.2%). The higher sequence identity between *ace1* and *ace2* in insect common ancestors apparently reflects their slow rate of evolution in spite of longer evolutionary history compared to insects. Therefore, the insect common ancestors *aces* likely retain more ancestral features, thereby serving as the reference as a relatively more primitive *ace*.

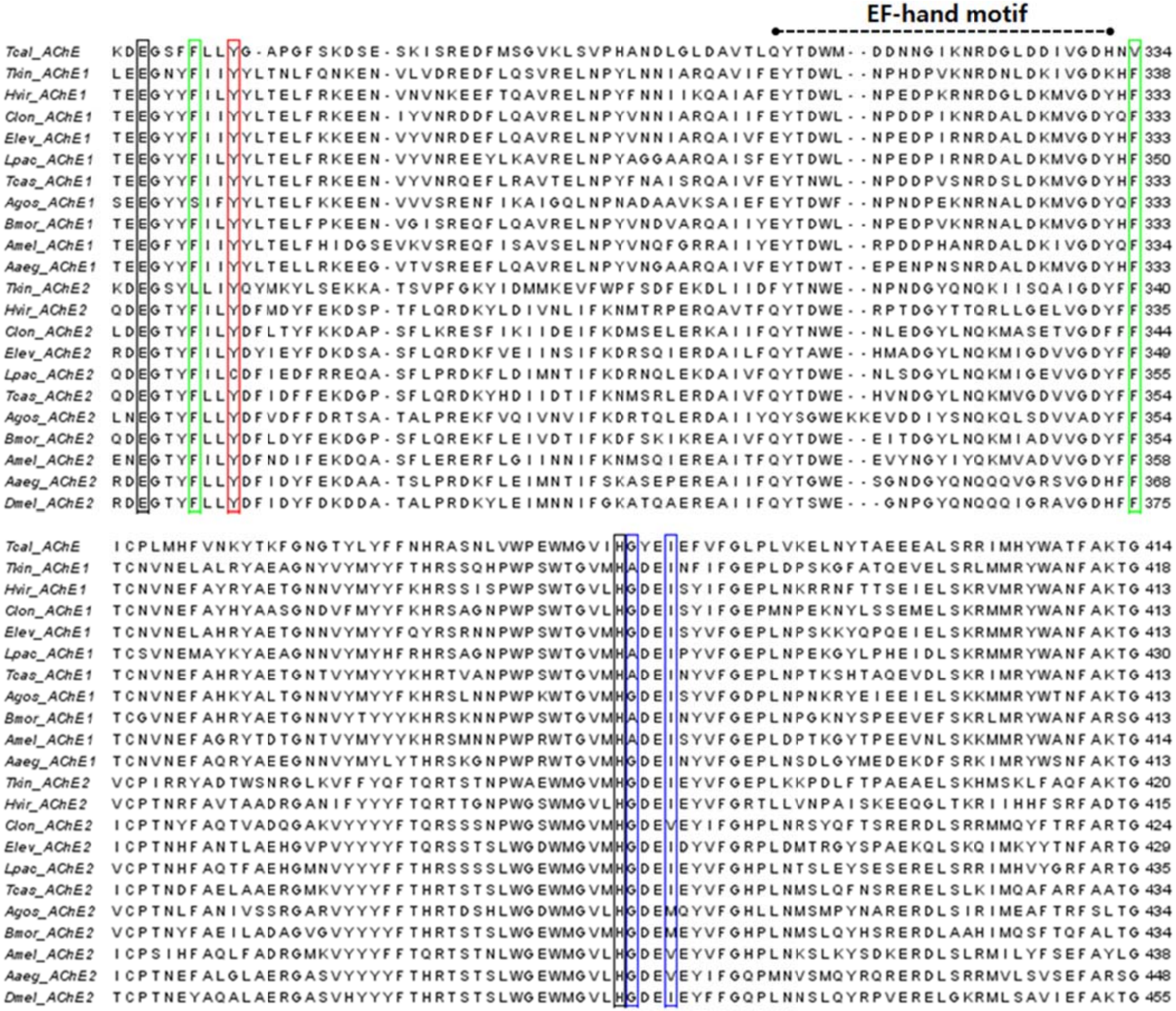
The alignment of deduced amino acid sequences of AChEs from the insect common ancestors with several insect AChEs and *Torpedo* AChE showed that insect common ancestors AChEs also have characteristics typical to cholinesterase (Fig. 2), including the omega-loop (Cys67-Cys94; the amino acid numberings were based on those of mature *Torpedo* AChE hereafter), choline-binding site (Trp84), three amino acid residues of catalytic triad (Ser200, Glu327 and His440) and the oxyanion hole (Gly118, Gly119 and Ala201). However, the amino acid residues of PAS (Tyr70, Asp72, Tyr121, Trp279 and Tyr334), CAS (Trp84, Tyr130, Glu199, Gly441 and Ile444) and ABP (Trp233, Phe288, Phe290, Phe331 and

```
                    Ω-loop                                  AChE2 specific insertion domain
Tcal_AChE     CQQYVDEQFPGFSGSEMWNPNREMSEDCLYLNIWVPSPRP-----------------------------------KSTTVM 46
Thin_AChE1    CSQIIDTLFPDFAGAAMWNPNTQVSEDCLFLNVVRPRAPYES-----------------------------DRPPKLAVM 51
Hvir_AChE1    CVQILDTVFGDFAGSAVWNPNTDLSEDCLYINVAVPHHRP-----------------------------------TNAAVM 46
Clon_AChE1    CMQIFDTVFGDFPGATMWNPNTQLNEDCLYINVVVPNPRP-----------------------------------KNSAVM 46
Elev_AChE1    CVQILDTVFGDFAGSAVWNPNTQLHEDCLYVNVVVPKPRP-----------------------------------KNAAVM 46
Lpac_AChE1    CVQIIDTVFGDFPGATMWNPNTPLSEDCLYLNVVAPVPRP-----------------------------APGTAAVM 48
Tcas_AChE1    CVQIIDTVFGDFPGATMWNPNTPLNEDCLYVNVVVPKPRP-----------------------------------TSAAVM 46
Agos_AChE1    CVQIFDTLFGDFPGATMWNPNSPVSEDCLYINVVVPRPRP-----------------------------------QNAAVM 46
Bmor_AChE1    CVQIVDTVFGDFPGAMMWNPNTDMQEDCLYINIVTPRPRP-----------------------------------KNAAVM 46
Amel_AChE1    CVQILDTVFGEFAGATMWNPNTPLSEDCLYVNVVVPRPRP-----------------------------------TNAAVM 46
Aaeg_AChE1    CVQIVDTVFGDFPGATMWNPNTPLSEDCLYINVVVPHPRP-----------------------------------KNSAVM 46
Thin_AChE2    CLQSDFVGFPDNRGEKVWTPNTNTSEDCLYLNIWIPASVFEE---------------------------KENKLATVM 51
Hvir_AChE2    CVQERSSYFPGFPGEEMWNPNTNVSEDCLYLNIWRPAPRE-----------------------------------GKAHLPIL 48
Clon_AChE2    CIQERCEYFPGFIGEEMWNPNTNISEDCLYLNLWVPKRVR-----------------------------HREVPKAPIL 50
Elev_AChE2    CYQERYEYFPGFEGEEMWNPNTNISEDCLYLNIWAPKKRK-----------------------------------DSPKAPLL 48
Lpac_AChE2    CYQERYEYFPGFEGEEMWNPNTNISEDCLYLNIWVPARFR---KGDG-----------------------------KERASLL 51
Tcas_AChE2    CYQERYEYFPGFEGEEMWNPNTNISEDCLYLNIWVPQRLRIRHHGEKLP---------------------QDRPKVPVL 58
Agos_AChE2    CYQERYEYFPGFVGEEMWSPNTKLSEDCLYLNIWIPKKQRTRHHSNN-----------------------AHHAKIPVL 56
Bmor_AChE2    CYQERYEYFPGFEGEEMWNPNTNISEDCLYLNIWVPQHLRVRHHQDKPL---------------------AERPKVPIL 58
Amel_AChE2    CYQERYEYFPGFPGEEMWNPNTNISEDCLYLNIWVPQKYRLRHKGDGSPGGNG----------------GPRNGLLPLL 63
Aaeg_AChE2    CIQERYEYFPGFAGEEMWNPNTNVSEDCLYLNIWVPTKTRLR-HGRGLNFGN--NDYFQDDDDFQRQ---HQSKGGLAML 74
Dmel_AChE2    CVQERYEYFPGFSGEEIWNPNTNVSEDCLYINVWAPAKARLR-HGRGANGGEHPNGKQADTDHLIHNGNPQNTTNGLPIL 79

Tcal_AChE     VWIYGGGFYSGSSTLDVYNGKYLAYTEEVVLVSLSYRVGAFGFLALH-----G-SQEAP--GNVGLLDQRMALQWVHDNI 118
Thin_AChE1    IWVYGGGFYSGTSTLDVYDHRTLAVTEQVLVVSMQYRVASLGFLYF------G-TPEVP--GNAGLFDQLMALQFIKDNI 122
Hvir_AChE1    VWIYGGGFYSGTSTLDVYDHRTLVSEENIILVSMQYRVASLGFLYF------D-TADVP--GNAGMFDQLMALQWVHDNI 117
Clon_AChE1    VWIYGGGFYSGTSTLDVYDHKTLVSEENIILVSMQYRVASLGFLFF------D-TSDVP--GNAGLFDQLMALQWVHDNI 117
Elev_AChE1    VWIYGGGFYSGTSTLDVYDHKTLVAEEKVILVSMQYRVASLGFLYF------D-SADVP--GNAGLFDQLMALQWVRDNI 117
Lpac_AChE1    VWVYGGGFYSGTSTLDVYDHRTLVAEEKVILVSMQYRVASLGFLFL------D-SAEAP--GNAGLYDQVMALRWIRDNI 119
Tcas_AChE1    VWVFGGGFYSGTNTLEVYDHNILVSEENIILVSMQYRVASLGFLYF------G-TPDVP--GNAGLFDQMMALQWVRDNI 117
Agos_AChE1    VWIFGGGFYSGSATLDIYDPKILVSEENVILVSMQYRVASLGFLYF------D-TEDVP--GNAGLFDQLMALQWVHENI 117
Bmor_AChE1    LWVFGGGFYSGTATLDVYDPKILVSEEKVVYVSMQYRVASLGFLFF------D-TADVP--GNAGLFDQLMALQWVKDNI 117
Amel_AChE1    VWIFGGGFYSGSATLDVYDHKTLVSEEKVILVSMQYRVASLGFLYF------G-TPDVP--GNAGLFDQVMALEWVRDNI 117
Aaeg_AChE1    LWIFGGGFYSGTATLDVYDHRTLASEENVIVVSLQYRVASLGFLFL------G-TPEAP--GNAGLFDQNLALRWVRDNI 117
Thin_AChE2    FWIYGGGFFSGSSALDVYNGAVLAAENEVVVVSTQYRVGPLGFLFF------G-TDDAP--GNVGLLDQVQALKWTNQHI 122
Hvir_AChE2    VWIYGGGYMSGTSTLDIYNADGLAAEGNVVVASMQYRIGAFGFLSL------G-TPEAP--GNVGLYDQALAIKWIKDNA 119
Clon_AChE2    VWIYGGGYMAGTSTLDVYYADIMAVANDVIIASMQYRVGAFGFLYF------D-HEEAP--GNVGLYDQISAIERLKNNA 121
Elev_AChE2    VWVYGGGYMSGTTTLDVYEASILAAENDVVVAAMQYRVGAFGFLYMPGEDGSR-DSEASVGGNMGLHDQVLALRWLRDNA 127
Lpac_AChE2    VWIYGGGYMSGTSTLDIYDADMMAAAEEVAVASMQYRVGAFGFLYL------G-VDEAP--GNAGLFDQALAIRWLRDNA 122
Tcas_AChE2    VWIYGGGYMSGTSTLDIYDADIIAATSDVIVASMQYRVGAFGFLYLSKYFPRG-SEEAP--GNMGMWDQALAIRWIKENA 135
Agos_AChE2    VWIYGGGYMSGTSTLDIYDGDLLAATFDVMIASMQYRLGAFGSLYLTPELPED-SDDAP--GNMGLWDQALAIKWIKENA 133
Bmor_AChE2    VWIYGGGYMSGTATLDLYKADIMASTSDVIVASMQYRVGAFGFLYLNKYFSPG-SEEAP--GNMGLWDQQLAIRWIKENA 135
Amel_AChE2    VWIYGGGFMSGTATLDVYNADIMAATSNVIIASMQYRVGAFGFLYLNKHFT-N-SEEAP--GNMGLWDQALALRWLRDNA 139
Aaeg_AChE2    VWIYGGGFMSGTSTLDVYNAEMLAAVGNVIVASMQYRVGSFGFFYLAPYL--N-DDDAP--GNVGLWDQALAIRWLKENA 149
Dmel_AChE2    IWIYGGGFMTGSATLDIYNADIMAAVGNVIVASFQYRVGAFGFLHLAPEMPSEFAEEAP--GNVGLWDQALAIRWLKDNA 157
```

```
Tcal_AChE     QFFGGDPKTVTIFGESAGGASVGMHILSPGSRDLFRRAILQSGSPNC-PWASVSVAEGRRRAVELGRNLNC------NLN 191
Thin_AChE1    ERFGGDPDNMTLFGESAGAVAVSLHLLSPLSRNLYNRAILQSGSATA-QWALITREESVLRGLRLAEAVNC-PHDKNRIN 200
Hvir_AChE1    QFFGGNPNNVTLFGESAGAVAVSLHLLSPLSRNLFTQAIMQSGSATA-PWAIISREESVLRGLRLAEAVGC-PKDRRNLP 195
Clon_AChE1    KFFGGNPNNVTLFGESAGAVAVSLHLLSPLSRNLFSQAIMQSGSATA-PWAIITREESILRGLRLAEAVGC-PRNRNNMR 195
Elev_AChE1    HAFGGNPNNITLFGESAGAVSVSLHLLSPLSRHLFNQAIMQSGSPTA-PWAIIGRDESVLRGLRLAEAVGC-PHSRSNLA 195
Lpac_AChE1    HAFGGNPENVTLFGESAGAVSVSFHLLSPLSRQLFSQAIMQSGGATA-PWAIISRTEAVLRALRLAEALGCSEVSAAGLN 198
Tcas_AChE1    AAFGGNPNNITLFGESAGAVSVSLHLLSPLSRNLFSQAIMESGSATA-PWAIISREESILRGLRLAEAVGC-PHERHELS 195
Agos_AChE1    KLFGGNPNNVTLFGESAGAVSVSLHLLSPLSRNLFNQAIMESGSSTA-PWAILSREESFSRGLKLAKAMGC-PDDRNEIH 195
Bmor_AChE1    GYFGGNPHNITLFGESAGAVSVSLHLLSPLSRNLFSQAIMQSGAATA-PWAIISREESILRGIRLAEAVHC-PHSRSDLA 195
Amel_AChE1    AAFGGNPDNVTLFGESAGAVSVSMHLLSPLSRHLFNQAIMQSGSPTA-PWAIISREESIVRGIRLAEAVGC-PHDRDNLQ 195
Aaeg_AChE1    HKFGGDPSRVTLFGESAGAVSVSLHLLSALSRDLFQRAILQSGSPTA-PWALVSREEATLRALRLAEAVNC-PHDATKLT 195
Thin_AChE2    NHFGGDSNSITVFGESAGGASTSYHSLSPLSRTLFKRSILMSGISTVDSWAVNHPKKAREVSYRLANLVGC------DTK 196
Hvir_AChE2    ISFGGDPHKITLFGESAGAGAVSLLLVSPVTRSLAHRGILQSGTLNA-PWSFMTAERARTVAEELAENCGC------GLN 192
Clon_AChE2    EVFGGDPDSITLFGESAGGGSVTIHLMSPETKHLVTRSIIQSSTLNA-PWSYMTGEMAEKVGVRLVSDVGC---DASKLH 197
Elev_AChE2    LAFGADPDTLTLFGESAGGGAVSAHILSPMSRGLVRRGILQSGTLNA-PWSYMTAENALQIGRTLADDVNC---NATTRM 203
Lpac_AChE2    AAFGADPESLTLFGESAGGGSVSIHLVSPVSKGLARRGIMQSGTVNA-PWSHMTAERAFAVARTLIKDCGCWNATTYPLP 201
Tcas_AChE2    AAFGGDPDLITLFGESAGGGSVSILLLSPVTKGLARRGILQSGTMNA-PWSYMSGERAPQIGKVLVEDCGC---NVSLLE 211
Agos_AChE2    AAFGADPETITLFGESAGGGSVSVHLISPETGGMVKRGIIQSGTVNA-PWSYMTGERAVEIAKKLLDDCNC---NSTSLD 209
Bmor_AChE2    RAFGGDPELITLFGESAGGGSVSLHMLSPEMKGLFKRGILQSGTLNA-PWSWMTGERAQDIGKVLIDDCNC---NSSLLA 211
Amel_AChE2    EAFGGDPELITIFGESAGGSSVSLHLISPVTRGLVRRGILQSGTLNA-PWSYMSGEKANEVATILVDDCGC---NSTMLN 215
Aaeg_AChE2    KAFGGDPDLITLFGESAGGSSVSLHLLSPVTRGLSRRGILQSGTLNA-PWSHMSAEKALSVAEALIDDCNC---NVTLLK 225
Dmel_AChE2    HAFGGNPEWMTLFGESAGSSSVNAQLMSPVTRGLVKRGMMQSGTMNA-PWSHMTSEKAVEIGKALINDCNC---NASMLK 233

Tcal_AChE     SDE-----------ELIHCLREKKPQELIDVEWNVLPFDS-IFRFSFVPVIDGEFFPTSLESMLNSGNFKKTQILLGVN 258
Thin_AChE1    --------------DVVACLRTVNSSELVSKEWG---TLG-ICEFPFVPVLDGSFLDEMPEVALKNGHFKQTPLLLGSN 261
Hvir_AChE1    --------------LVIDCLRHVNATILVSKEWG---TLG-ICEFPFVPVIDGAFLDESPKVSMIHQNFKKTNILMGSN 256
Clon_AChE1    --------------AVIDCLRKTNASDLVNNEWG---TLG-ICELPFVPVIDGSFLGEMPQKSLAERNFKKTSILMGSN 256
Elev_AChE1    --------------ATVECLRQVNASDLVANEWG---TLG-ICEFPFVPVIDGSFLDETPQRSLANKNFKKTNILMGSN 256
Lpac_AChE1    ESPD-GSVATPEMRAAAAKCLRKADAKKIVDSEWG---TLG-ICEFPFVPIVDGSFLDQTPARSLATKDFKKASILLGSN 273
Tcas_AChE1    --------------AVIDCLKKKDPIDLVNNEWG---TLG-ICESPFVPVIDGAFLDESPTRALANKNFKKTNILMGSN 256
Agos_AChE1    --------------KTVECLRKVNSSAMVEKEWD---HVA-ICFFPFVPVVDGAFLDDHPQKSLSTNNFKKTNILMGSN 256
Bmor_AChE1    --------------PMIECLRKKNADELVNNEWG---TLG-ICEFPFVPIIDGSFLDEMPVRSLAHQNFKKTNILMGSN 256
Amel_AChE1    --------------EVIDCLRVKDPVELVKNEWG---TLG-ICEFPFVPVIDGAFLDETPQRSLATSSFKKANIMMGSN 256
Aaeg_AChE1    --------------DTVECLRTKDPNVLVDNEWG---TLG-ICEFPFVPVVDGAFLDETPQRSLASGRFKKTDILTGSN 256
Thin_AChE2    SSPQ----------SIVTCLKEADPKKLNSIQWK--LTSGKIVSFPFKPTVDGYFLHDMPLNIVGTRDAKRHEIIIGTT 263
Hvir_AChE2    ITAG----------AIVQCLRNVDSKTISENQWN--SYTG-ILGFPSAPTVGGELFPALPFELLRETDLSGTEILIGSN 258
Clon_AChE2    DLSD-------DGIRDVMQCMRGIDAKNISVTQWN--SYWR-ILGSPSAPTIDGVLIPKDPIEMLKEADFNNTEILIGSN 267
Elev_AChE2    EREG--------GIAEVIQCLRAVDPKTLSVTQWN--SYSG-ILGFPSAPTIDGEFLTAHPLELLRRGDFDETEIMIGSN 272
Lpac_AChE2    EEPAQKSKENEKAIEEVMECMRFVEPKTISVQQWN--SYSG-ILGFPSAPTIDGVFLPKHPLEMLKEGDFGKTEILIGSN 278
Tcas_AChE2    TRPH----------EVIDCMRAVEAKTISLQQWN--SYSG-ILGFPSTPTVDGVFMPKHPMDMLAEGDYEDMEILVGSN 277
Agos_AChE2    SNPI----------GTMSCMRPVDASTISKKQWN--SYSG-ILGFPSAPTVDGILLPEHPLDMLAKANFSDIDILIGSN 275
Bmor_AChE2    KDPS----------LVMDCMRGVDAKTISVQQWN--SYTG-ILGFPSAPTVDGIFLPKDPDTMMKEGNFHNSEVLLGSN 277
Amel_AChE2    ENPA----------RVMACMRSVDAKTISVQQWN--SYWG-ILGFPSAPTIDGIFLPKHPLDLLREADFKDTEILIGNN 281
Aaeg_AChE2    DNPN----------YVMNCMRNVDAKTISVQQWN--SYSG-ILGFPSAPTIDGVFMTADPMTMLREANLEGVEILVGSN 291
Dmel_AChE2    TNPA----------HVMSCMRSVDAKTISVQQWN--SYSG-ILSFPSAPTIDGAFLPADPMTLMKTADLKDYDILMGNV 299
```



```
Tcal_AChE   KDEGSFFLLYG-APGFSKDSE-SKISREDFMSGVKLSVPHANDLGLDAVTLQYTDWM--DDNNGIKNRDGLDDIVGDHNV 334
Tkin_AChE1  LEEGNYFIIYYLTNLFQNKEN-VLVDREDFLQSVRELNPYLNNIARQAVIFEYTDWL--NPHDPVKNRDNLDKIVGDKHF 338
Hvir_AChE1  TEEGYYFILYYLTELFRKEEN-VNVNKEEFTQAVRELNPYFNNIIKQAIAFEYTDWL--NPEDPKRNRDGLDKMVGDYHF 333
Clon_AChE1  TEEGYYFIIYYLTELFRKEEN-IYVNRDDFLQAVRELNPYVNNIARQAIIFEYTDWL--NPDDPIKNRDALDKMVGDYQF 333
Elev_AChE1  TEEGYYFIIYYLTELFKKEEN-VYVNRDEFLQAVRELNPYVNNIARQAIVFEYTDWL--NPEDPIRNRDALDKMVGDYHF 333
Lpac_AChE1  TEEGYYFILYYLTELFRKEEN-VYVNREEYLKAVRELNPYAGGAARQAISFEYTDWL--NPEDPIRNRDALDKMVGDYHF 350
Tcas_AChE1  TEEGYYFIIYYLTELFRKEEN-VYVNRQEFLRAVTELNPYFNAISRQAIVFEYTNWL--NPDDPVSNRDSLDKMVGDYHF 333
Agos_AChE1  SEEGYYSIFYYLTELFKKEEN-VVVSRENFIKAIGQLNPNADAAVKSAIEFEYTDWF--NPNDPEKNRNALDKMVGDYQF 333
Bmor_AChE1  TEEGYYFILYYLTELFPKEEN-VGISREQFLQAVRELNPYVNDVARQAIIYEYTDWL--NPEDPVKNRNALDKMVGDYHF 333
Amel_AChE1  TEEGFYFIIYYLTELFHIDGSEVKVSREQFISAVSELNPYVNQFGRRAIIYEYTDWL--RPDDPHANRDALDKIVGDYQF 334
Aaeg_AChE1  TEEGYYFIIYYLTELLRKEEG-VTVSREEFLQAVRELNPYVNGAARQAIVFEYTDWT--EPENPNSNRDALDKMVGDYHF 333
Tkin_AChE2  KDEGSYLLIYQYMKYLSEKKA-TSVPFGKYIDMMKEVFWPFSDFEKDLIIDFYTNWE--NPNDGYQNQKIISQAIGDYFF 340
Hvir_AChE2  QDEGTYFILYDFMDYFEKDSP-TFLQRDKYLDIVNLIFKNMTRPERQAVTFQYTDWE--RPTDGYTTQRLLGELVGDYFF 335
Clon_AChE2  LDEGTYFILYDFLTYFKKDAP-SFLKRESFIKIIDEIFKDMSELERKAIIFQYTNWE--NLEDGYLNQKMASETVGDFFF 344
Elev_AChE2  RDEGTYFILYDYIEYFDKDSA-SFLQRDKFVEIINSIFKDRSQIERDAILFQYTAWE--HMADGYLNQKMIGDVVGDYFF 349
Lpac_AChE2  QDEGTYFILCDFIEDFRREQA-SFLPRDKFLDIMNTIFKDRNQLEKDAIVFQYTDWE--NLSDGYLNQKMIGEVVGDYFF 355
Tcas_AChE2  QDEGTYFLLYDFIDFFEKDGP-SFLQRDKYHDIIDTIFKNMSRLERDAIVFQYTDWE--HVNDGYLNQKMVGDVVGDYFF 354
Agos_AChE2  LNEGTYFLLYDFVDFFDRTSA-TALPREKFVQIVNVIFKDRTQLERDAIIYQYSGWEKKEVDDIYSNQKQLSDVVADYFF 354
Bmor_AChE2  QDEGTYFLLYDFLDYFEKDGP-SFLQREKFLEIVDTIFKDFSKIKREAIVFQYTDWE--EITDGYLNQKMIADVVGDYFF 354
Amel_AChE2  ENEGTYFILYDFNDIFEKDQA-SFLERERFLGIINNIFKNMSQIEREAITFQYTDWE--EVYNGYIYQKMVADVVGDYFF 358
Aaeg_AChE2  RDEGTYFLLYDFIDYFEKDAA-TSLPRDKFLEIMNTIFSKASEPEREAIIFQYTGWE--SGNDGYQNQQQVGRSVGDHFF 368
Dmel_AChE2  RDEGTYFLLYDFIDYFDKDDA-TALPRDKYLEIMNNIFGKATQAEREAIIFQYTSWE---GNPGYQNQQQIGRAVGDHFF 375

Tcal_AChE   ICPLMHFVNKYTKFGNGTYLYFFNHRASNLVWPEWMGVIHGYEIEFVFGLPLVKELNYTAEEEALSRRIMHYWATFAKTG 414
Tkin_AChE1  TCNVNELALRYAEAGNYVVYMYYFTHRSSQHPWPSWTGVMHADEINFIFGEPLDPSKGFATQEVELSRLMMRYWANFAKTG 418
Hvir_AChE1  TCNVNEFAYRYAETGNNVYMYYFKHRSSISPWPSWTGVLHGDEISYIFGEPLNKRRNFTTSEIELSKRVMRYWANFAKTG 413
Clon_AChE1  TCNVNEFAYHYAASGNDVFMYYFKHRSAGNPWPSWTGVLHGDEISYIFGEPMNPEKNYLSSEMELSKRMMRYWANFAKTG 413
Elev_AChE1  TCNVNELAHRYAETGNNVYMYYFQYRSRNNPWPSWTGVMHGDEISYVFGEPLNPSKKYQPQEIELSKRMMRYWANFAKTG 413
Lpac_AChE1  TCSVNEMAYKYAEAGNNVYMYHFRHRSAGNPWPSWTGVMHADEIPYVFGEPLNPEKGYLPHEIDLSKRMMRYWANFAKTG 430
Tcas_AChE1  TCNVNEFAHRYAETGNTVYMYYYKHRTVANPWPSWTGVMHADEINYVFGEPLNPTKSHTAQEVDLSKRIMRYWANFAKTG 413
Agos_AChE1  TCNVNEFAHKYALTGNNVYMYYFKHRSLNNPWPKWTGVMHGDEISYVFGDPLNPNKRYEIEEIELSKKMMRYWTNFAKTG 413
Bmor_AChE1  TCGVNEFAHRYAETGNNVYTYYYKHRSKNNPWPSWTGVMHADEINYVFGEPLNPGKNYSPEEVEFSKRLMRYWANFARSG 413
Amel_AChE1  TCNVNEFAGRYTDTGNTVYMYYYKHRSMNNPWPRWTGVMHADEISYVFGEPLDPTKGYTPEEVNLSKKMMRYWANFAKTG 414
Aaeg_AChE1  TCNVNEFAQRYAEEGNNVYMYLYTHRSKGNPWPRWTGVMHGDEINYVFGEPLNSDLGYMEDEKDFSRKIMRYWSNFAKTG 413
Tkin_AChE2  VCPIRRYADTWSNRGLKVFFYQFTQRTSTNPWAEWMGVMHGDEIEYVFGEPLKKPDLFTPAEAELSKHMSKLFAQFAKTG 420
Hvir_AChE2  VCPTNRFAVTAADRGANIFYYYFTQRTTGNPWGSWMGVLHGDEIEYVFGRTLLVNPAISKEEQGLTKRIIHHFSRFADTG 415
Clon_AChE2  ICPTNYFAQTVADQGAKVYYYYFTQRSSSNPWGSWMGVMHGDEVEYIFGHPLNRSYQFTSRERDLSRRMMQYFTRFARTG 424
Elev_AChE2  ICPTNHFANTLAEHGVPVYYYYFTQRSSTSLWGDWMGVMHGDEIDYVFGRPLDMTRGYSPAEKQLSKQIMKYYTNFARTG 429
Lpac_AChE2  VCPTNHFAQTFAEHGMNVYYYFFTHRSSSSLWGEWMGVMHGDEIEYVFGHPLNTSLEYSESERELSRRIMHVYGRFARTG 435
Tcas_AChE2  ICPTNDFAELAAERGMKVYYYFFTHRTSTSLWGEWMGVMHGDEIEYVFGHPLNMSLQFNSRERELSLKIMQAFARFAATG 434
Agos_AChE2  VCPTNLFANIVSSRGARVYYYFFTHRTDSHLWGDWMGVLHGDEMQYVFGHLLNMSMPYNARERDLSIRIMEAFTRFSLTG 434
Bmor_AChE2  VCPTNYFAEILADAGVGVYYYYFTHRTSTSLWGEWMGVMHGDEMEYVFGHPLNMSLQYHSRERDLAAHIMQSFTQFALTG 434
Amel_AChE2  ICPSIHFAQLFADRGMKVYYYFFTQRTSTNLWGEWMGVLHGDEVEYVFGHPLNKSLKYSDKERDLSLRMILYFSEFAYLG 438
Aaeg_AChE2  ICPTNEFALGLAERGASVYYYYFTHRTSTSLWGEWMGVLHGDEVEYIFGQPMNVSMQYRQRERDLSRRMVLSVSEFARSG 448
Dmel_AChE2  TCPTNEYAQALAERGASVHYYYFTHRTSTSLWGEWMGVLHGDEIEYFFGQPLNNSLQYRPVERELGKRMLSAVIEFAKTG 455
```

Figure 2. Alignments of five hexapod acetylcholinesterases (AChEs) amino acid sequences with AChEs from some insects and *Torpedo californica*. The conserved amino acid residues corresponding to the Ω-loop, AChE2-specific insertion sequence and EF-hand motif are indicated by dotted lines. The residues forming the catalytic triad, acyl binding pocket, peripheral anionic site, catalytic anionic site and oxyanion hole are indicated by black, green, red, blue and pink boxes, respectively. Tcal, *Torpedo californica* (X03439.1); Tkin, *Tomocerus kinoshitai*; Hvir, *Haslundichilis viridis*; Clon, *Ctenolepisma longicaudata*; Elev, *Ecdyonurus levis*; Lpac, *Lyriothemis pachygastra*; Tcas, *Tribolium castaneum* (HQ260968.1 and HQ260969.1); Dmel, *Drosophila melanogaster* (NM057605.5); Agos, Aphis gossypii (AB158637.1 and AF502081.1); Bmor, *Bombyx mori* (NM_001043915.1 and NM_001114641.1); Amel, *Apis mellifera* (XM_393751.4 and NM_001040230.1); Aaeg, *Aedes aegypti* (XM_001656927.1 and XM_001655818.1).

Val400) were somewhat different between the AChE1 and AChE2. First, Ile444 and Gly441 of CAS in *Torpedo* AChE were partially replaced with Val444 or Met444 in insects AChE2 and Ala441 in insects AChE1. Second, Phe288 of ABP was replaced with Cys288 in insects AChE1 and Leu288 in insects AChE2. Third, Ile70, Asp72, and Tyr121 of PAS in AChE1 were replaced with Glu70, Tyr72, and Met121 in insects AChE2, respectively (Fig. 2).

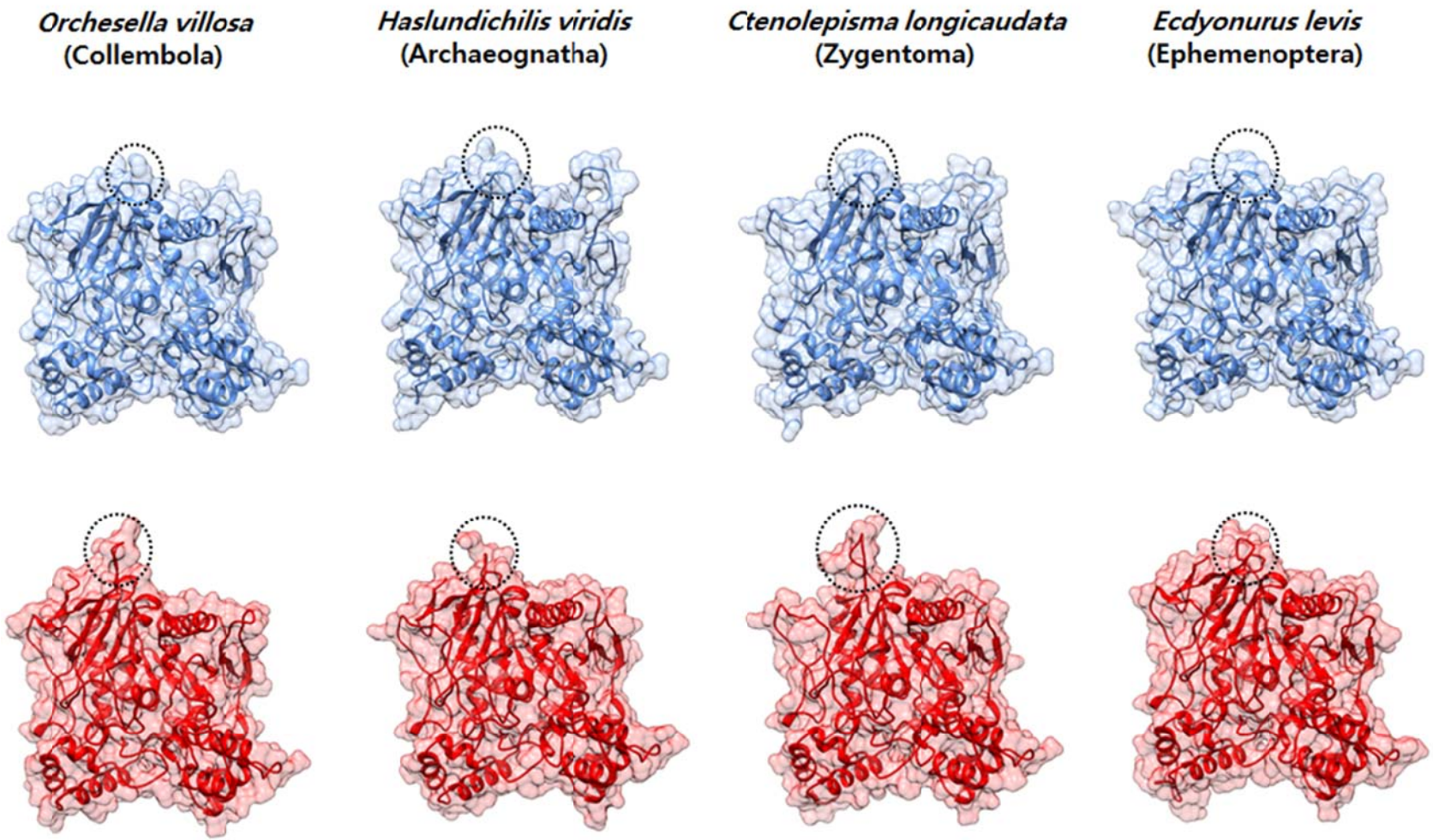
In the past several AChE related studies suggested that amino acids to compose active-site gorge from AChEs are changes may affect to their catalytic properties (Barak et al., 1994; Harel et al., 1992; Kaplan et al., 2001; Mallender et al., 2000; Saxena et al., 2003). Thus, these some replaced amino acid residues which are composed active-site gorge expect to affect to their catalytic properties of AChE1 and AChE2 in the insects. Also, these amino acids replacements seem to appear more frequently in insects AChE2 than AChE1. On the other hand, it is indicated that AChE1 is more conserved than AChE2 in during evolution after duplication.

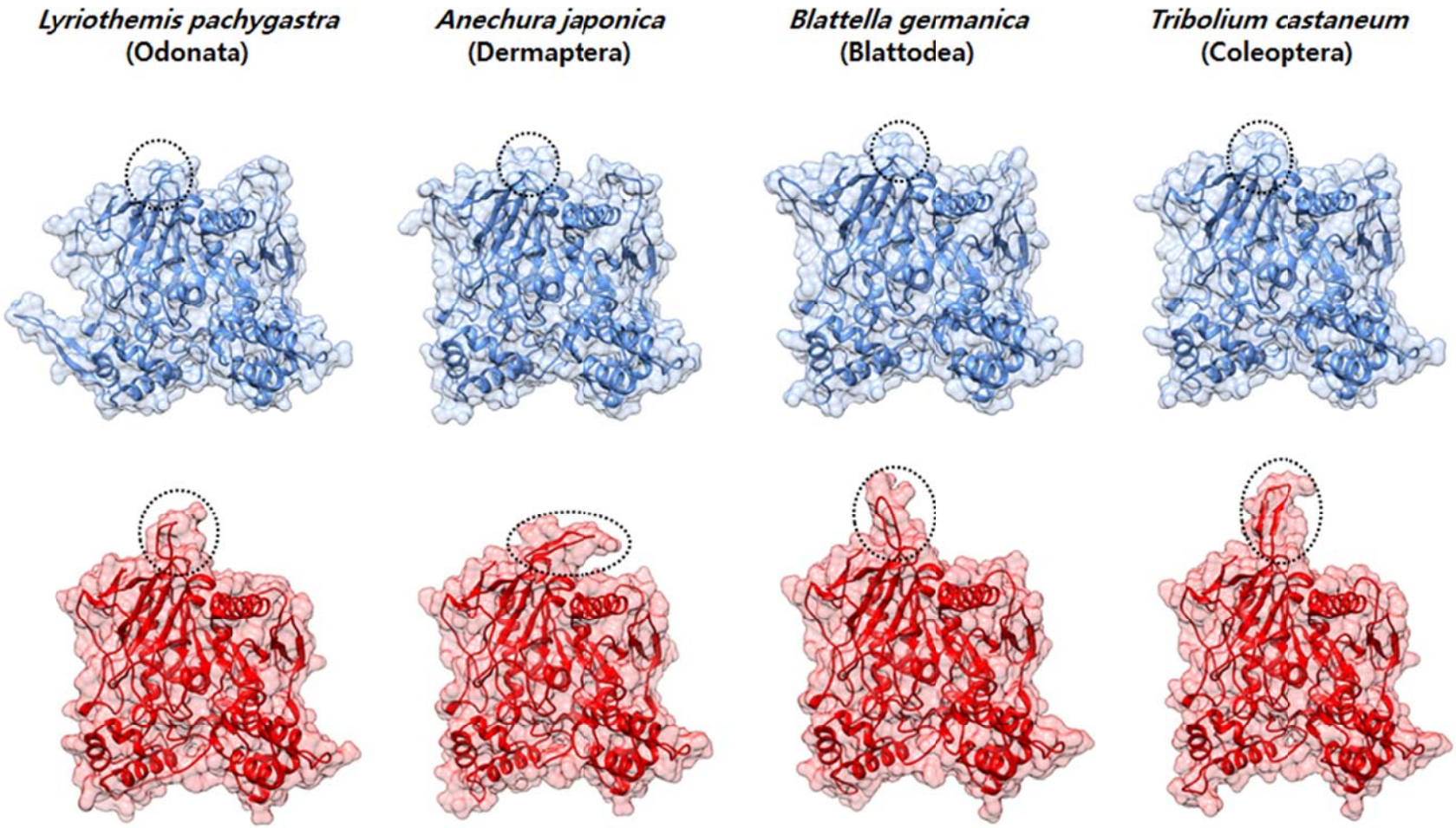
3.2. Identification of motifs and domains from insect two AChEs

Comparison of deduced amino acid sequences between AChE1 and AChE2 identified three motifs or domains that are specific to either AChE1 or AChE2 although the specificity of each domain is not absolutely conserved.

First, an insertion domain was found only in AChE2 (Ser103 to Thr109). This AChE2-specific insertion domain was not fixed in its amino acid sequence composition and length (Fig. 2). However, a tendency was found that the insertion domain has elongated over the evolutionary track: its presence was not apparent in the species of hexapod (*H. viridis*, AChE1: 7aa, AChE2: 9aa; *C. longicaudata*, AChE1: 7aa, AChE2: 11aa; *E. levis*, AChE1: 7aa, AChE2: 9aa; *L. pachygastra*, AChE1: 9aa, AChE2: 12aa) or its size was small (9-12 aa) in ancient insects (i.e., Odonata), whereas it has become generally longer (17-40 aa) in more evolved insect species such as *T. castaneum*, *B. mori* and *D. melanogaster*. When 3D structure was constructed based on homology modeling, the insertion domain was found to be located at the surface of enzyme (Fig. 3). As noticed by its primary amino acid sequence length, the volume of insertion domain became generally bulkier in more evolved species. Since presence of the AChE2-specific insertion domain became apparent in insects, particularly in more evolved species, its acquisition appears to be a relatively recent event, perhaps during the radiation of insect species. Despite such evolutionary aspect of the insertion domain, its function remains to be elucidated (Harel et al., 2000; Mutero and Fournier 1992; Seong et al., 2012; Wiesner et al., 2007).

Second, the EF-hand motif was found in AChEs (Gln373 to His398 *Torpedo* amino acid numbering) (Fig. 2) and was located at the enzyme surface as judged


Orchesella villosa
(Collembola)
Haslundichilis viridis
(Archaeognatha)
Ctenolepisma longicaudata
(Zygentoma)
Ecdyonurus levis
(Ephemenoptera)


Lyriothemis pachygastra
(Odonata)
Anechura japonica
(Dermaptera)
Blattella germanica
(Blattodea)
Tribolium castaneum
(Coleoptera)

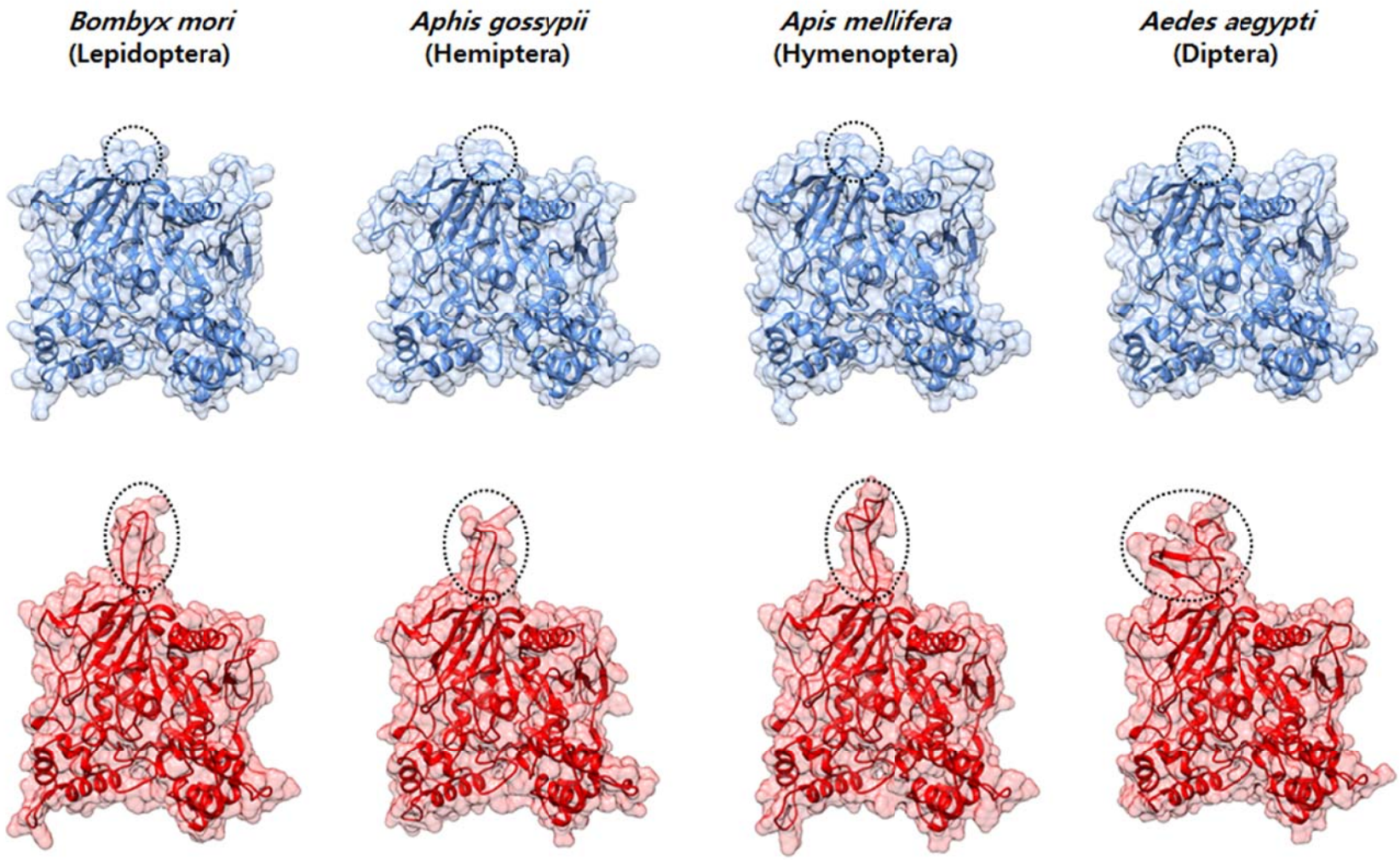

Bombyx mori
(Lepidoptera)
Aphis gossypii
(Hemiptera)
Apis mellifera
(Hymenoptera)
Aedes aegypti
(Diptera)

Figure 3. Three-dimensional structure of insect AChEs displayed by ribbon and accessible surface diagrams. The insect AChE1 and AChE2 are shown in blue and red, respectively. AChE2-specific insertion domains (from Ser106 to Thr109, amino acid numbering of *Torpedo californica*) are indicated by black dotted circles.

by 3D structure modeling (Fig. 4). The EF-hand motif is mainly composed of the E/F helices and a Ca^{2+} binding loop as first observed in the parvalbumin (Kretsing and Nockolds 1973). Several studies reported that once binding with calcium or another ion, enzymes having the EF-hand motif underwent conformational changes and their activities increased (Lee et al., 2001; Li et al., 2002; Schallreuter et al., 2007). The consensus sequences for EF-hand motif was well conserved in AChE1 whereas they were only rudimentary in AChE2 in the five insect orders (Coleoptera, Diptera, Hemiptera, Hymenoptera and Lepidoptera) examined. Since the EF-hand motif was also conserved in insect common ancestor (i.e., *H. viridis*, *C. longicaudata*, *E. levis* and *L. pachygastra*), it appears to have been present in ancient form of AChE1. Presence of EF-hand motif specifically in AChE1 implies the possible regulation of AChE1 activity via Ca^{2+} binding.

Third, the LRE-motif, which is composed of successive three amino acids of Leu, Arg and Glu, is known to have the function of adhesion with other proteins. In particular, it is found in many proteins that form the basal lamina of outer surface of cell membrane in the neuromuscular junction. (Hall and Sanes 1993; Hunter et al., 1991). In insects, no AChE1 is known to have LRE-motif, whereas it has been known to exist in AChE2 of some particular species from the endopterygote (Johnson and Moore 2013). The existence of LRE-motif in AChE2

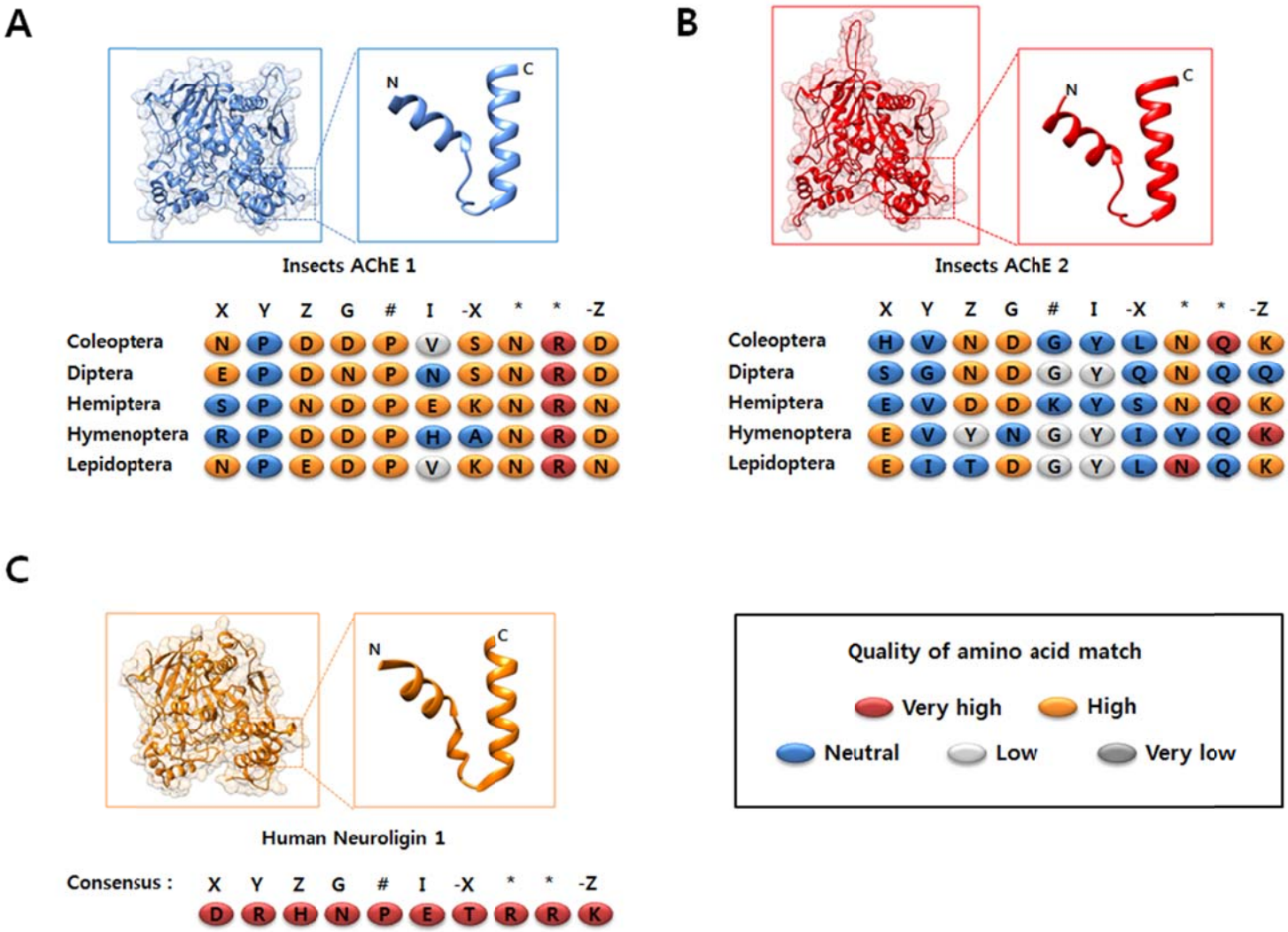

A
N
C
Insects AChE 1
X Y Z G # I -X * * -Z
Coleoptera N P D D P V S N R D
Diptera E P D N P N S N R D
Hemiptera S P N D P E K N R N
Hymenoptera R P D D P H A N R D
Lepidoptera N P E D P V K N R N
B
N
C
Insects AChE 2
X Y Z G # I -X * * -Z
Coleoptera H V N D G Y L N Q K
Diptera S G N D G Y Q N Q Q
Hemiptera E V D D K Y S N Q K
Hymenoptera E V Y N G Y I Y Q K
Lepidoptera E I T D G Y L N Q K
C
N
C
Human Neuroligin 1
Consensus : X Y Z G # I -X * * -Z
D R H N P E T R R K
Quality of amino acid match
Very high
High
Neutral
Low
Very low

Figure 4. Comparative analysis of the EF-hand motif homology in insect AChE1 (A) and AChE2 (B). The homology level of EF-hand motif was calculated by comparing insect AChEs with human neuroligin1 (C) using the HHblit homology detection server. The degrees of amino acid sequence homology were categorized from “very high” to “very low”, which were indicated by red l, yellow, blue, grey and dark grey colors, respectively.

was found in additional insect species (*H. viridis*, Archaeognatha; *S. infuscatum*, Odonata; *A. japonica*, Dermaptera; *S. avenae*, Hemiptera) (Table 1). Distribution of the LRE-motif across a wide variety of AChEs from invertebrates to vertebrates indicates the convergent evolution of this motif. In addition, the complete absence of the LRE-motif in AChE1 lineage implies the possibility of negative impacts of this motif on AChE1 function, thereby having been selected out However, the wide distribution of LRE motif in AChE2 lineage suggests its possible role in the adhesion with cell membrane or other proteins, which is likely required for the newly acquired function of AChE2.

In summary, the amino acid residues in functionally important domains are highly conserved in both AChE1 and AChE2. Nevertheless, there were some motifs or domains were identified to be specifically associated with either AChE1 (the EF-hand motif) or AChE2 (i.e., the insertion domain and LRE motif), which appear to play a role in functional diversification of insect AChEs.

3.3. Differences in the active-site gorge topology between insect two AChEs

In order to determine the structural conservation and any possible differentiation in the active-site gorge between insect two AChEs during the evolution after gene duplication, the degree of changes in the amino acid residues composing the active-site gorge was estimated by measuring the levels of the

Table 1. LRE (Leu-Arg-Glu) motif of AChE2 in insect and their location.

Order	Total No. of species	LRE-motif in AChE2	Family	No.	LRE-motif location*
Archaeognatha	1	1	Machilidae	1	309-311
Odonata	5	1	Libellulidae	1	528-533
Dermaptera	1	1	Forticulidae	1	256-258
Diptera	21	11	Culicidae	9	309-311
			Tephritidae	2	528-533
Hemiptera	9	1	Aphididae	1	528-533
Hymenoptera	10	9	Apidae	4	309-311
			Pteromalidae	3	309-311
			Megachilidae	1	309-311
			Formicidae	1	Close to 200
Other insects	22	-	-	-	-

* Amino acid numbering of AChE from *Torpedo californica*.

nucleotide diversity (Pi value) and site-specific selection (*dN/dS* ratio). The Pi values of the residues forming the five main domains (i.e., PAS, CAS, ABP, oxyanion hole and catalytic triad) were generally lower compared to the residues in other remaining region, indicating that these functionally crucial domains have been preserved during evolution (Fig. 5). This trend was true for both AChE1 and AChE2, and no significant difference in the Pi value levels was noticed between AChE1 and AChE2. Therefore, it appears that at least the minimum features required for the AChE functionality have been well retained in AChE2 as well as AChE1, which provides the fundamental basis for the observed AChE2 activity in most insects (Kim and Lee 2013).

The site-direct *dN/dS* of the residues forming the five main domains revealed that these residues of insect two AChEs have been undergone purifying selection. These results again mean that residues forming the active-site gorge have been highly conserved in both lineages of AChE1 and AChE2 (Fig. 6). Overall degrees of *dN*/*dS* of the resisdues in AChE2 were not significantly different from those of AChE1, suggesting that no apparent amino acid sequence differentiation of active-site gorge has occurred either in AChE2 except for the aforementioned sequence differences in the PAS and ABP.

To determine whether the differences in the amino acid composition in PAS and ABP result in any differences in the active-site gorge topology, 3D structures

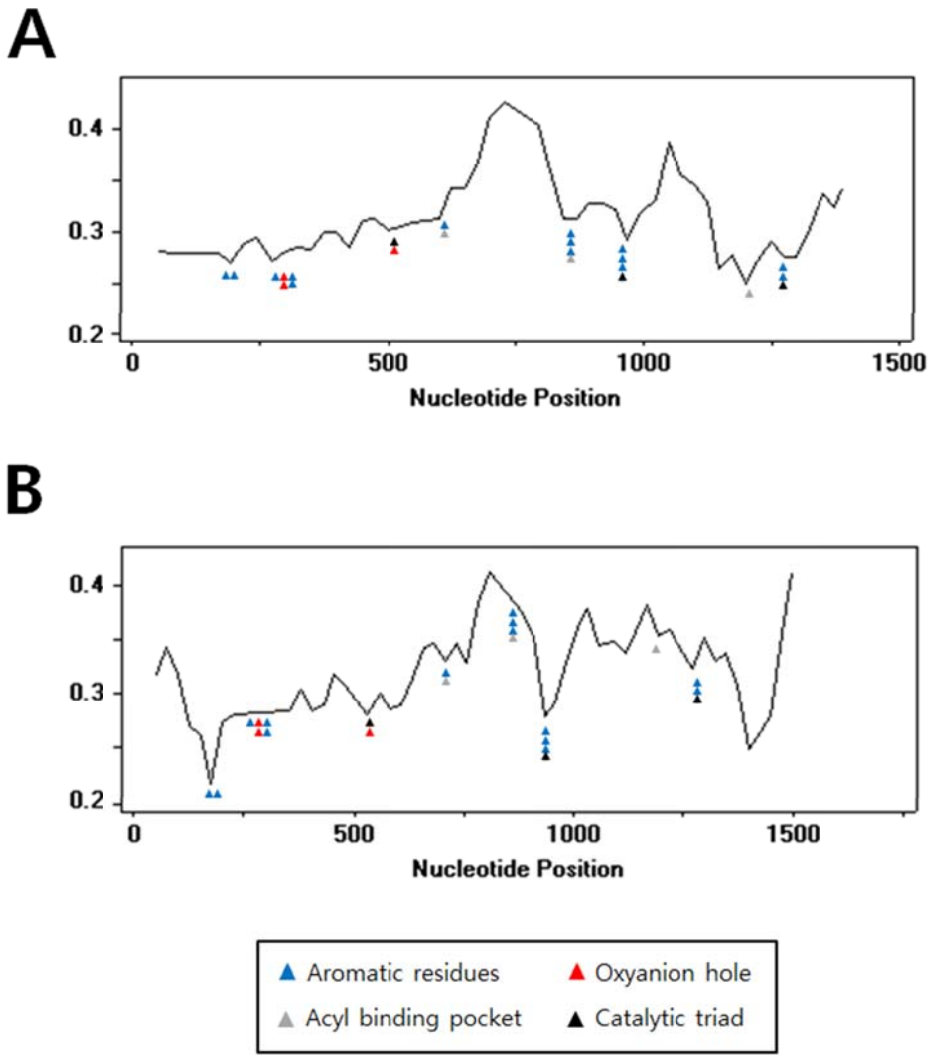


Figure 5. Estimation and comparison of nucleotide diversity (Pi value) between the core regions of insect AChE1 (A) and AChE2 (B). The 14 aromatic residues, residues of acyl binding pocket, oxyanion hole and catalytic triad are shown with blue, grey, red and black arrow heads, respectively.

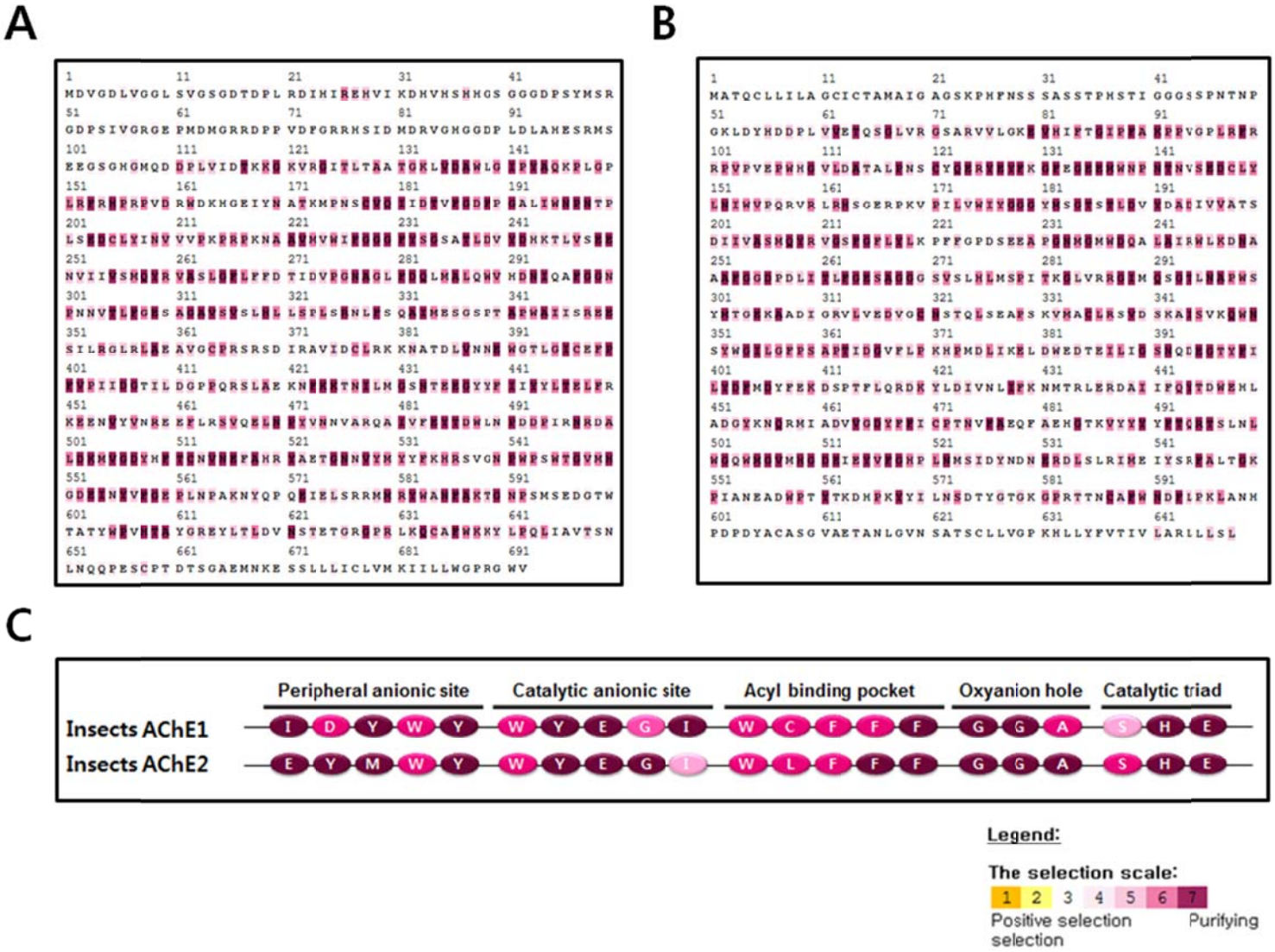

A
B
C
Peripheral anionic site
Catalytic anionic site
Acyl binding pocket
Oxyanion hole
Catalytic triad
Insects AChE1
Insects AChE2
Legend:
The selection scale:
1 2 3 4 5 6 7
Positive selection
Purifying selection

Figure 6. The degrees of site-direct selection posed to the amino acid residues of insect AChE1 (A) and AChE2 (B) as scored by *dN*/*dS* values. (C) Comparison of *dN*/*dS* values in the five main domains from active-site gorge between insect AChE1 and AChE2. A run on three sequences of AChEs from each five insect order (a total of 15 sequences) calculated by the MEC model. Positive selection is colored in shades of yellow, whereas purifying selection is colored in shades of magenta.

of insect two AChEs were generated and compared by homology modeling (Fig 7). The resulting 3D structures with a particular emphasis on the PAS showed clear differences in the active-site gorge topology that appears to be generated by different amino acid residues in the PAS (i.e., Asp72 and Tyr121 in AChE1 vs. Tyr121 and Met121 in AChE2). Generally, the active-site gorge entrance in insect AChE2 was narrower compared to that of AChE1 and the overall volume of gorge was smaller than that of AChE1. Similar observation was also reported in previous studies when the 3D crystal structure of *D. melanogaster* AChE2 was compared with that of *Torpedo* AChE (which is more similar to insect AChE1) (Harel et al., 2000; Wiesner et al., 2007). The other remaining difference in PAS residue composition (Ile vs. Glu) was determined to cause a noticeable alteration in the modeled structure of AChE. In addition, the structural difference caused by the different amino acid residues in the ABP (i.e., 288Cys in AChE1 vs. 288Leu in AChE2) was not visible in this model. More detailed investigation on their role in AChE structure and function remains to be studied.

Cross-comparison of amino acid residues between different hexapods and insects revealed that the Asp72 residue has been conserved in all insects and their common ancestors (i.e., hexapods), suggesting its critical role associated with the function of AChE1. In contrast, the amino acid residue at the corresponding site in AChE2 varied considerably (from Ser, Phe or Cys) and then the Tyr residue began

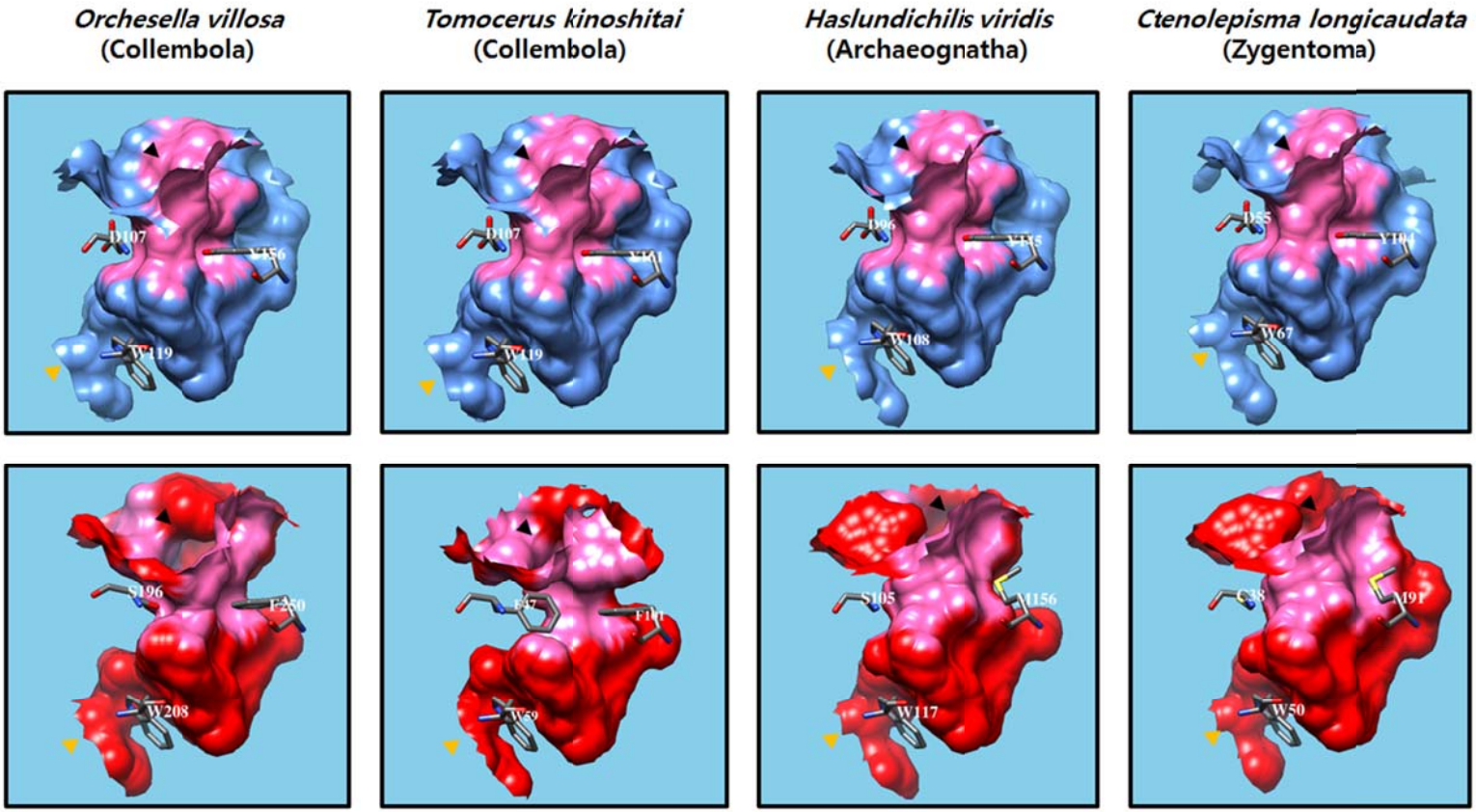


Figure 7. Comparison of the active-site gorge topology between AChE1 (top panels) and AChE2 (bottom panels) from several hexapods and insects. The Asp72, Tyr121 and Trp84 residues forming the peripheral anionic site (PAS) were shown with the stick model (amino acid numbering based on that of *Torpedo* AChE). The PAS region was shown in hot pink color. Black and orange arrows indicate the gorge entrance and backdoor, respectively.

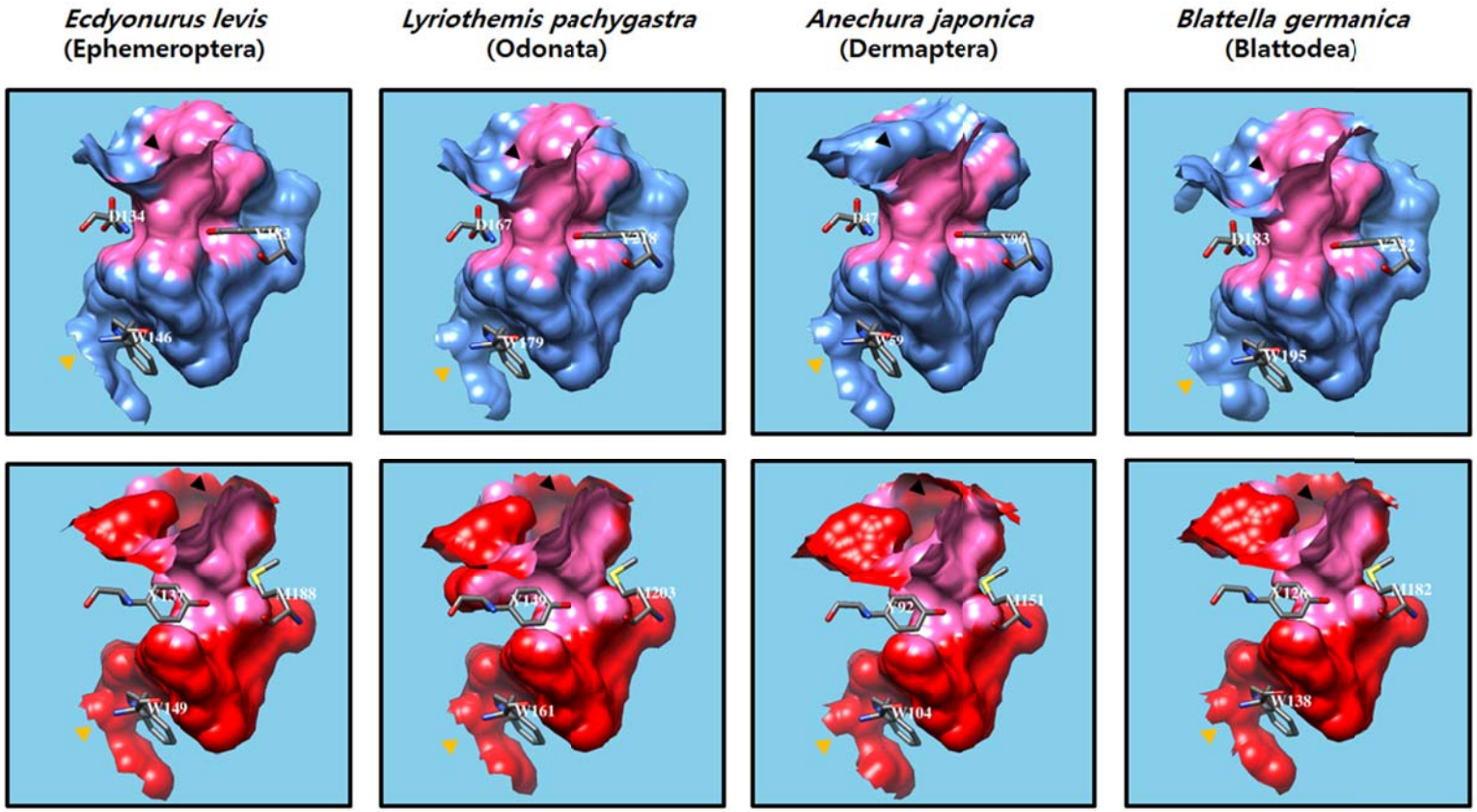


Figure 7. Continued

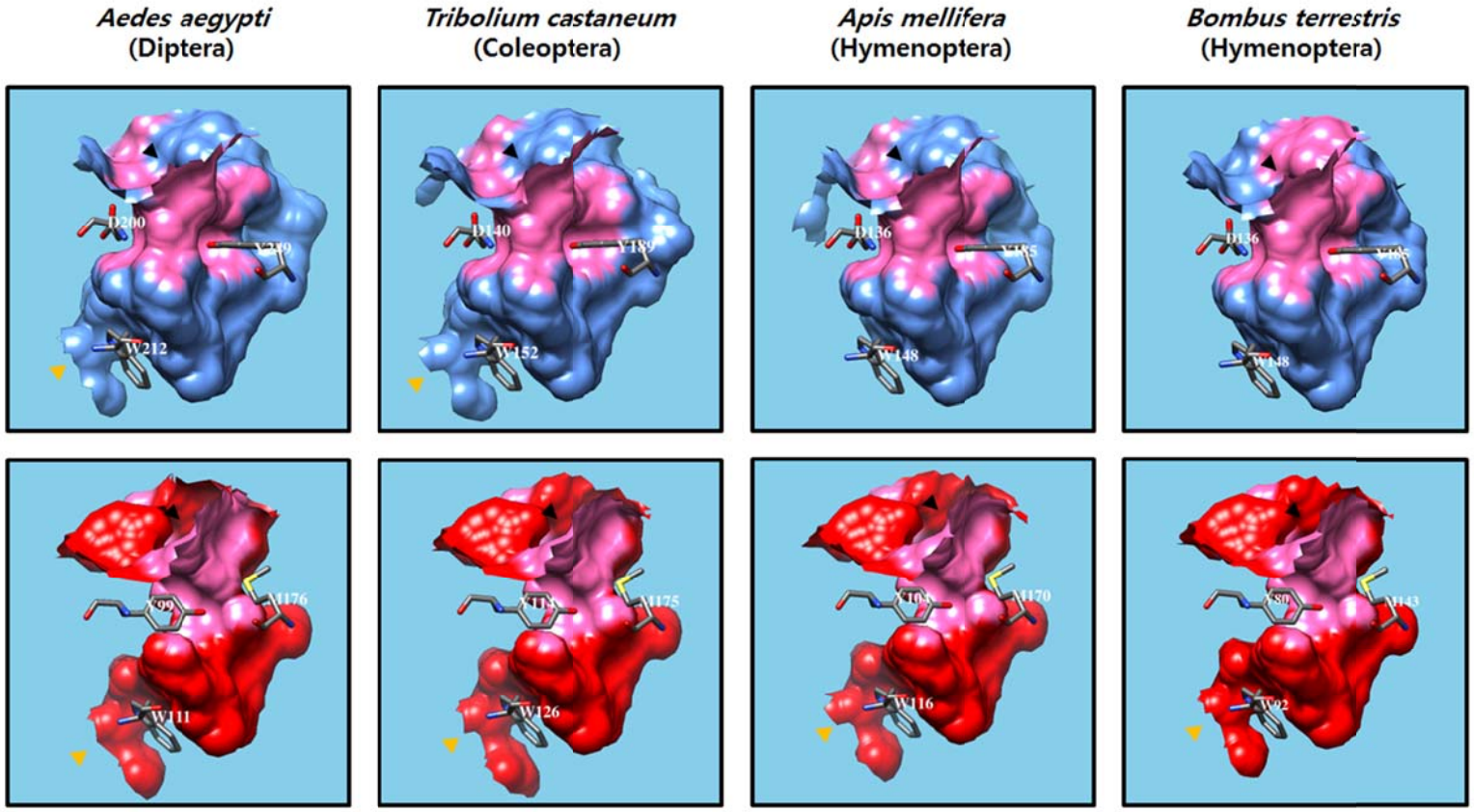


Figure 7. Continued

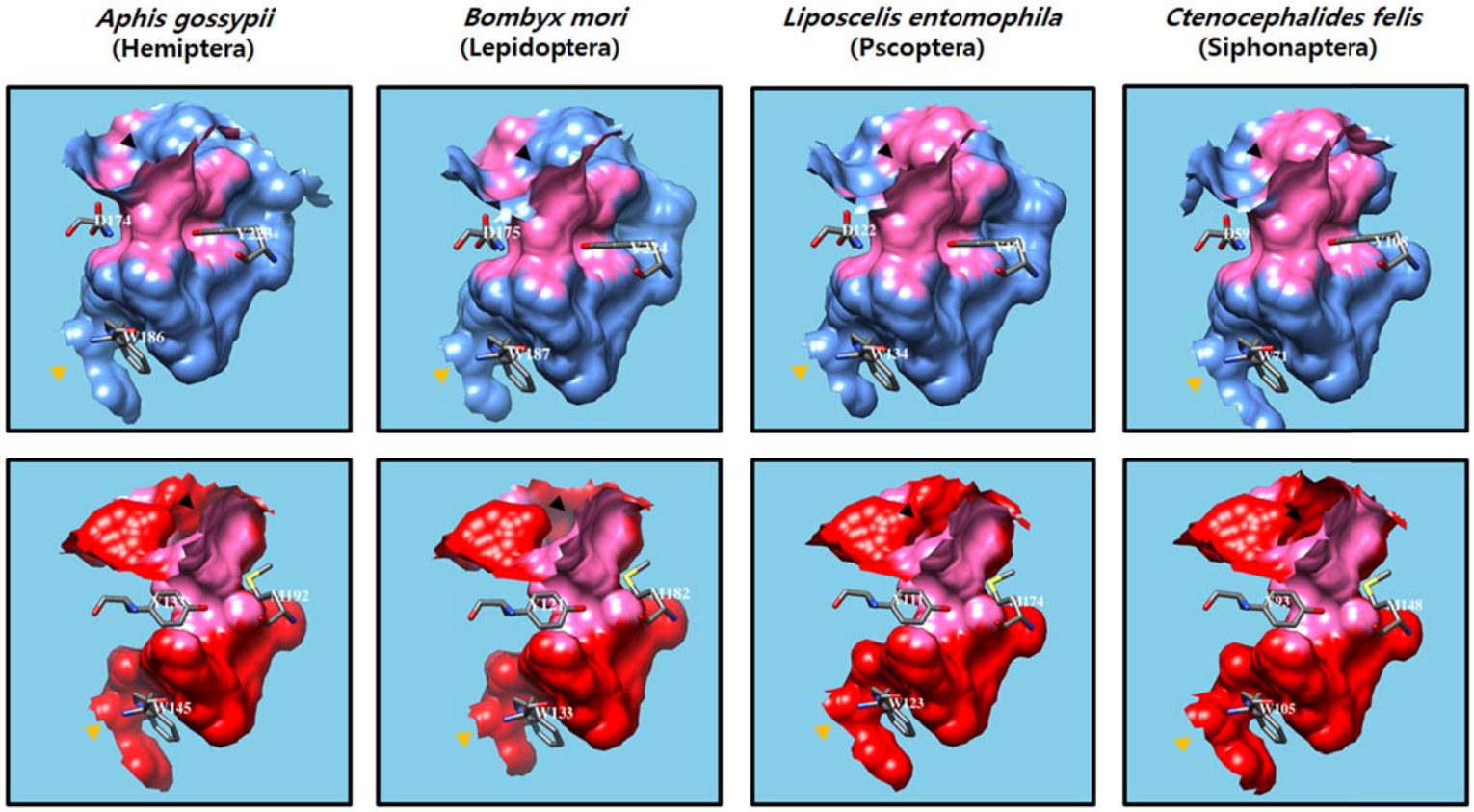


Figure 7. Continued

to appear from Ephemenoptera (*E. levis*) and was fixed in insects. The finding that the functional PAS has been well preserved in AChE1 indicates that PAS is a feature more typical to AChE1 lineage and is likely associated with the AChE1-specific kinetic property. Since PAS is known to be associated with the allosteric regulation of substrate binding and the phenomenon of substrate inhibition (Toutant 1989), the observed substrate inhibition pattern in AChE1 but not in AChE2 is likely due to the differentiation of PAS between AChE1 and AChE2.

Unlike the PAS showing consistent difference in topology between AChE1 and AChE2, the backdoor that was visualized by 3D modeling was highly conserved (see orange arrow heads in Fig. 7) not only across a wide variety of hexapods and insects examined but also between AChE1 and AChE2, suggesting that this region has been conserved after duplication during evolution. Interestingly, only exception was found in some Hymenopteran insects (*Apis mellifera* and *Bombus terrestris*), in which the formation of backdoor was not obvious in AChE1. Since the AChE backdoor is known to play an important role in cleaning the reaction products (acetate and choline), thereby enhancing the catalytic efficiency of AChE (Colletier et al., 2007; Nachon et al., 2008; Sanson et al., 2011). A detailed comparison of the deduced amino residues forming the backdoor (Trp84, Trp432, and Tyr442, *Torpedo* AChE numbering) or adjacent to it (Ser329 and Ile439, *Torpedo* AChE numbering) among a wide array of insects

revealed that the backdoor-forming residues of insects AChEs were highly conserved with those of *Torpedo* AChE (i.e., Trp84, Trp432, and Asp442) and completely conserved in all insects examined. The residues adjacent to the backdoor in insect AChE1 (Tyr329 and Met439), however, were different from those corresponding to *Torpedo* AChE (i.e., Ser329 and Ile439). Interestingly, the Tyr329 residue was substituted with Phe in AChE1 of *A. mellifera* and *B. terrestris,* suggesting that this residue change likely causes the defect of backdoor in AChE1. To confirm whether this Tyr-to-Phe substitution actually alters the backdoor topology in AChE1, the reverse Phe–to-Tyr substitution was introduced into the *A. mellifera* AChE1 (AmAChE1) sequences and its 3D structure was predicted by homology modeling (Fig. 8). The substitution of Phe with Tyr clearly restored the backdoor topology that is similar to those of other insect AChE1. In the 3D structures of AmAChE1 with the Phe–to-Tyr substitution as well as other insect AChE1, the presence of Tyr391 (Ser329 in *Torpedo* AChE) appears to affect the spatial arrangement of Met503 (Ile439 in *Torpedo* AChE) in favor of opening the backdoor. In contrast, the substitution of Tyr391 with Phe results in the alteration of angles of Met503, blocking the backdoor formation. Therefore, the Tyr-to-Phe substitution occurred in AChE1 from some Hymenopteran insects is likely the main determining factor associated with alteration of backdoor topology. A 3D structure modeling of AmAChE1 and AmAChE2 in a previous study also

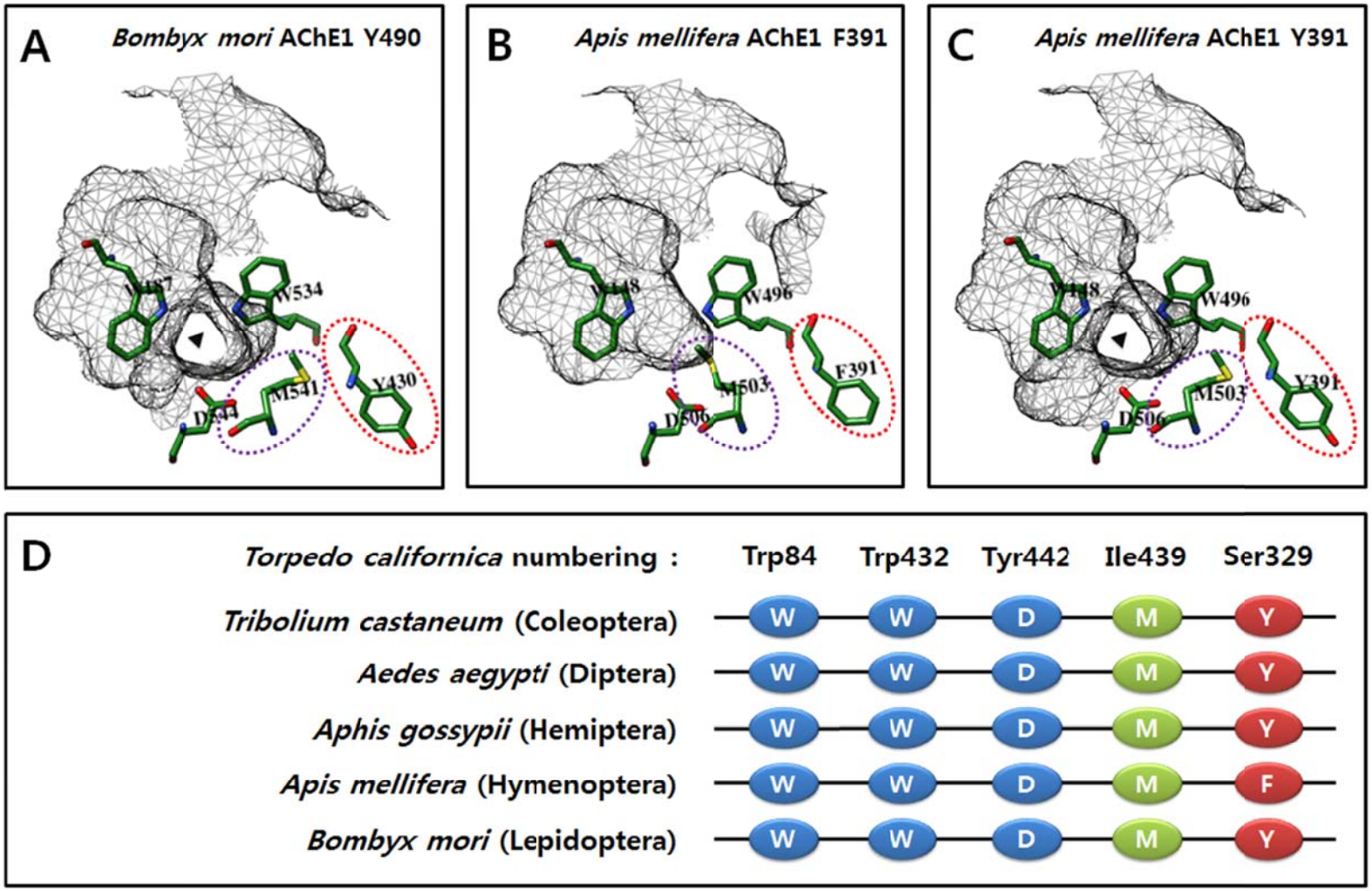

A
Bombyx mori AChE1 Y490
W534
M541
Y430
B
Apis mellifera AChE1 F391
W496
M503
F391
C
Apis mellifera AChE1 Y391
W496
M503
Y391
D
Torpedo californica numbering :
Trp84
Trp432
Tyr442
Ile439
Ser329
Tribolium castaneum (Coleoptera)
W W D M Y
Aedes aegypti (Diptera)
W W D M Y
Aphis gossypii (Hemiptera)
W W D M Y
Apis mellifera (Hymenoptera)
W W D M F
Bombyx mori (Lepidoptera)
W W D M Y

Figure 8. Comparison of the backdoor region topology between AChE1 and AChE2 from *Apis mellifera* and other insects. The residues forming the backdoor or ones adjacent to it were shown in green. Blue dot circles indicated Met residue (Ile439, amino acid numbering of *Torpedo* AChE) that appear to affect backdoor opening. Red dot circles indicated Tyr or Phe residues (Ser329) that differ between *A. mellifera* AChE1 and AChE2 and likely affect the spatial arrangement of the Met residue. Backdoor topologies of AChE1 from *Bombyx mori* and *A. mellifera* are shown in (A) and (B), respectively. Backdoor topology of *A. mellifera* AChE1 with the reverse Phe-to-Tyr substitution is shown in (C). The consensus residues forming backdoor in five insect species are represented in (D).

reported that the side chain angle of a choline-binding site residue (Trp344 in AmAChE1 vs. Trp336 in AmAChE2) at the active-site gorge entrance was different by approximately 90°, likely resulting in the alteration of PAS conformation (Kim et al., 2012). Taken together, in addition to such alteration in the PAS topology from insect AChE1 and AChE2, the blocked backdoor topology are likely responsible for the significantly reduced catalytic efficiency of AmAChE1. Besides these AmAChE1-specific differences, the consensus differences in the PAS residues found between all insect AChE1 and AChE2 (i.e., Ile70, Asp72, and Tyr121 AChE1 vs. Glu70, Tyr72, and Met121 in AChE2) may function as additive factors associated with the significant reduction in AmAChE1 activity. Introduction of the Tyr-to-Phe substitution to other insect AChE1 (or vice versa; i.e., Phe-to-Tyr substitution to AmAChE1) singly or in combination with other substitutions at different PAS sites by *in vitro* mutagenesis in conjunction with functional analysis would be necessary to confirm the effects of these AmAChE1-specific substitutions on the catalytic property.

In summary, differentiation of active-site gorge, particularly at PAS and ABP, between insect AChE1 and AChE2 seems to have occurred early in the evolution, probably before the divergence of hexapods. Such differentiation in active-site gorge structure appears to be responsible for the observed differences in catalytic properties between AChE1 and AChE2. The introduction of Tyr-to-Phe

substitution to AChE1, causing a significant reduction of enzyme activity, appears to have occurred relatively recently and locally in some Hymenopteran insects, thereby functioning as a driving force for the functional transition from AChE1 and AChE2. Since similar cases (i.e., use of AChE2 as the main catalytic enzyme) have been reported in many other insect species, such functional transition of replacement was proposed to have been occurred with multiple origins across a wide variety of insects. A larger scale survey of mutations that are functionally similar the Tyr-to-Phe substitution in AChE1 would be required to understand the putatively multiple origins of such event.

APPENDIX. Information data table of *aces* and *ace-like* sequences from metazoans

Class	Order	Species	ace1	ace2
Insecta	Diptera	*Aedes aegypti*	XM_001656927.1	XM_001655818.1
		Aedes albopictus	AB218421.1	AB218420.1
		Anopheles gambiae	Genome data	BN000067.1
		Anopheles darlingi		Genome data
		Anopheles funestus	EZ976793.1	Genome data
		Anopheles stephensi	Genome data	Genome data
		Culex tritaeniorhynchus	AB122152.1	AB122151.1
		Culex pipiens	AJ489456.1	AM159193.1
		Culex pipiens pallens	AY762905.1	
		Culex quinquefasciatus	XM_001847396.1	XM_001842175.1
		Phlebotomus papatasi	JQ922267.1	Genome data
		Musca domestica		AF281162.1
		Stomoxys calcitrans		HM125963.1
		Haematobia irritans		AY466160.1
		Cochliomyia hominivorax		FJ830868.1
		Lucilia cuprina		JF776371.1
		Ceratitis capitata		EU130781.1
		Exorista sorbillans		HM028669.1
		Bactrocera dorsalis		AY155500.1
		Drosophila melanogaster		NM_057605.5
		Drosophila albomicans		Genome data
		Drosophila ananassae		Genome data

APPENDIX. Continued

Class	Order	Species	ace1	ace2
Insecta	Diptera	*Drosophila biarmipes*		Genome data
		Drosophila bipectinata		Genome data
		Drosophila elegans		Genome data
		Drosophila erecta		Genome data
		Drosophila eugracilis		Genome data
		Drosophila ficusphila		Genome data
		Drosophila kikkawai		Genome data
		Drosophila miranda		Genome data
		Drosophila sechellia		Genome data
		Drosophila simulans		Genome data
		Drosophila yakuba		Genome data
		Drosophila grimshawi		Genome data
	Lepidoptera	*Helicoverpa assulta*	DQ001323.1	AY817736.1
		Helicoverpa armigera	JF894118.1	JF894119.1
		Plutella xylostella	AY773014.2	AY061975.1
		Chilo suppressalis	EF453724.1	EF470245.1
		Cnaphalocrocis medinalis		FN538987.1
		Bombyx mandarina	EU262633.2	EF166089.1
		Bombyx mori	NM_001043915.1	NM_001114641.1
		Melitaea cinxia	GQ489250.1	GQ489251.2
		Heliconius melpomene	Genome data	
		Danaus Plexippus	Genome data	EHJ76382

APPENDIX. Continued

Class	Order	Species	ace1	ace2
Insecta	Lepidoptera	*Cydia pomonella*	DQ267977.1	DQ267976.1
		Manduca sexta	Genome data	Genome data
	Hymenoptera	*Nasonia vitripennis*	XM_001600408.2	XM_001605518.2
		Nasonia giraulti	Genome data	Genome data
		Nasonia longicornis	Genome data	Genome data
		Apis mellifera	XM_393751.4	NM_001040230.1
		Apis florea	XM_003693203.1	XM_003694182.1
		Bombus impatiens	XM_003488192.1	Genome data
		Bombus terrestris	XM_003399342.1	XM_003401864.1
		Oomyzus sokolowskii		HM212643.1
		Megachile rotundata	XM_003699292.1	XM_003701146.1
		Atta cephalotes	Genome data	Genome data
		Polyrhachis vicina		JF742990.1
		Acromyrmex echinatior	Genome data	
		Camponotus floridanus	Genome data	
		Linepithema humile	Genome data	
		Pogonomyrmex barbatus	Genome data	
		Solenopsis invicta	Genome data	
		Harpegnathos saltator	Genome data	
	Siphonaptera	*Ctenocephalides felis*	FN645950.1	FN645951.1
	Coleoptera	*Leptinotarsa decemlineata*	JF343436.1	JF343437.1
		Tribolium castaneum	HQ260968.1	HQ260969.1

APPENDIX. Continued

Class	Order	Species	ace1	ace2
Insecta	Coleoptera	*Alphitobius diaperinus*	EU086057.1	EU086056.1
		Lasioderma serricorne	GU211888.1	
	Hemiptera	*Aphis gossypii*	AB158637.1	AF502081.1
		Aphis glycines	JQ349160.1	
		Rhopalosiphum padi	AY667435.1	AY707318.1
		Sitobion avenae	AY819704.1	AY707319.1
		Myzus persicae	AY147797.1	AF287291.1
		Schizaphis graminum	AF321574.1	
		Acyrthosiphon pisum	XM_001948618.2	XM_001948953.2
		Nilaparvata lugens	HQ605041.1	JN688930.1
		Nephotettix cincticeps	AB264392.1	AF145235.1
		Cimex lectularius	JN563927.1	GU597839.1
		Bemisia tabaci	EF675188.1	EF675190.1
	Psocoptera	*Liposcelis bostrychophila*	FJ647185.1	EF362950.1
		Liposcelis paeta	GU214754.1	
		Liposcelis decolor	FJ647186.1	FJ647187.1
		Liposcelis entomophila	EU854149.2	EU854150.1
	Phthiraptera	*Pediculus humanus corporis*	AB266606.1	AB266605.1
		Pediculus humanus capitis	AB266615.1	AB266614.1
	Blattodea	*Blattella germanica*	DQ288249.1	DQ288847.1
	Orthoptera	*Locusta migratoria manilensis*	EU231603.1	

APPENDIX. Continued

Class	Order	Species	ace1	ace2
Insecta	Dermaptera	*Anechura japonica*	Partial cloning	Partial cloning
	Odonata	*Lyriothemis pachygastra*	Partial cloning	Partial cloning
		Orthetrum albistylum speciosum	Partial cloning	Partial cloning
		Sympetrum infuscatum	Partial cloning	Partial cloning
		Coenagrion lanceolatum	Partial cloning	Partial cloning
		Calopteryx atrata	Partial cloning	Partial cloning
		Calopteryx japonica	Partial cloning	Partial cloning
	Ephemeroptera	*Ecdyonurus levis*	Partial cloning	Partial cloning
	Archaeognatha	*Haslundichilis viridis*	Partial cloning	Partial cloning
	Zygentoma	*Ctenolepima longicandata*	Partial cloning	Partial cloning
Entognatha	Entomobryomorpha	*Orchesella villosa*	FJ228227.1	FJ228228.1
		Tomocerus kinoshitai	Partial cloning	Partial cloning
Branchiopoda	Cladocera	*Daphnia pulex*	Uniprot: E9FRY6	Uniprot: E9G1H3
Chilopoda	Geophilomorpha	*Strigamia maritima*	Genome data	Genome data
Arachnida	Ixodida	*Rhipicephalus microplus*	AJ223965.1	
		Rhipicephalus appendiculatus	AJ006338.1	
		Rhipicephalus decoloratus	AJ006337.1	
		Dermacentor variabilis	AY212906.1	
	Trombidiformes	*Tetranychus urticae*	GQ461353.1	
		Tetranychus evansi	JQ779843.1	

APPENDIX. Continued

Phylum	Class	Species	ace1	ace2	ace3	ace4
Nematoda	Chromadorea	*Caenorhabditis elegans*	NM_078259.6	AF025378.3	AF039650.4	NM_064561.4
		Caenorhabditis brenneri	uniprot:G0MGN3	uniprot: G0NVV5	uniprot:G0MVQ6	uniprot:G0MVQ7
		Caenorhabditis briggsae	U41846	AF030037	AF159504	AF159505.1
		Caenorhabditis remanei	uniprot: E3MKC8	uniprot: E3LM27	uniprot: E3LF28	uniprot: E3LF29
		Heligmosomoides polygyrus	JF439067	JF439068		
	Secernentea	*Bursaphelenchus xylophilus*	GU166345.1	GU166346.1	GU166347.1	
		Dictyocaulus viviparus	DQ375489.1	AY218789.1		
Phylum	**Class**	**Species**	**ace-like 1**	**ace-like 2**	**ace-like 3**	**ace-like 4**
Choanoflagellata		*Monosiga brevicollis*	XM_001750752.1	XM_001743828.1	XM_001749298.1	
Porifera	Demospongiae	*Amphimedon queenslandica*	XM_003387193.1			
Placozoa	Tricoplacia	*Trichoplax adhaerens*	XM_002117107.1	XM_002117094.1	XP_002117136.1	XP_002117133.1
Cnidaria	Anthozoa	*Nematostella vectensis*	XM_001629623.1	XM_001627440.1		
Ctenophora	Tentaculata	*Mnemiopsis leidyi*	ML073237a			
Platyhelminthes	Digenea	*Schistosoma mansoni*	AF279461.1			
		Schistosoma haematobium	AY228511.1			
		Schistosoma bovis	AF279463.1			
	Trematoda	*Clonorchis sinensis*	GAA53463.1	DF144213.1		
Mollusca	Gastropoda	*Lottia gigantea*	Genome data	Genome data		
	Cephalopoda	*Doryteuthis opalescens*	AF065384.1			
	Bivalvia	*Crassostrea gigas*	EKC37862.1			
Annelida	Polychaeta	*Capitella teleta*	ELU02276.1	ELU12508.1	ELU18693.1	ELT99434.1
	Clitellata	*Helobdella robusta*	jgi|Helro1|76935	jgi|Helro1|120724		

LITERATURE CITED

Ariel, N., Ordentlich A., Barak D., Bino T., Velan B., Shafferman A., 1998. The 'aromatic patch' of three proximal residues in the human acetylcholinesterase active centre allows for versatile interaction modes with inhibitors. Biochem J. 335: 95-102

Arnold, K., Bordoli L., Kopp J., Schwede T., 2006. The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. Bioinformatics. 22: 195-201

Augustinsson, K.B., 1949. Substrate Concentration and Specificity of Choline Ester-Splitting Enzymes. Arch Biochem. 23: 111-126

Baek, J.H., Kim J.I., Lee D.W., Chung B.K., Miyata T., Lee S.H., 2005. Identification and characterization of ace1-type acetylcholine sterase likely associated with organophosphate resistance in Plutella xylostella. Pestic Biochem Phys. 81: 164-175

Ballard, C.G., Greig N.H., Guillozet-Bongaarts A.L., Enz A., Darvesh S., 2005. Cholinesterases: roles in the brain during health and disease. Curr Alzheimer Res. 2: 307-318

Barak, D., Kronman C., Ordentlich A., Ariel N., Bromberg A., Marcus D., Lazar A., Velan B., Shafferman A., 1994. Acetylcholinesterase Peripheral Anionic Site Degeneracy Conferred by Amino-Acid Arrays Sharing a Common Core. J Biol Chem. 269: 6296-6305

Colletier, J.P., Royant A., Specht A., Sanson B., Nachon F., Masson P., Zaccai G., Sussman J.L., Goeldner M., Silman I., Bourgeois D., Weik M., 2007. Use of a 'caged' analogue to study the traffic of choline within acetylcholinesterase by kinetic crystallography. Acta Crystallogr D Biol Crystallogr. 63: 1115-1128

Eyre-Walker, A., Gaut B.S., 1997. Correlated rates of synonymous site evolution across plant genomes. Mol Biol Evol. 14: 455-460

Fournier, D., Berge J.B., Dealmeida M.L.C., Bordier C., 1988. Acetylcholinesterases from Musca-Domestica and Drosophila-Melanogaster Brain Are Linked to

Membranes by a Glycophospholipid Anchor Sensitive to an Endogenous Phospholipase. J Neurochem. 50: 1158-1163

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 Molec. 32: 765-775

Gauthier, M., Belzunces L.P., Zaoujal A., Colin M.E., Richard D., 1992. Modulatory Effect of Learning and Memory on Honey-Bee Brain Acetylcholinesterase Activity. Comp Biochem Phys C. 103: 91-95

Gerlt, J.A., Gassman P.G., 1993. Understanding the Rates of Certain Enzyme-Catalyzed Reactions - Proton Abstraction from Carbon Acids, Acyl-Transfer Reactions, and Displacement-Reactions of Phosphodiesters. Biochemistry. 32: 11943-11952

Hall, Z.W., Sanes J.R., 1993. Synaptic structure and development: the neuromuscular junction. Cell. 72 Suppl: 99-121

Harel, M., Kryger G., Rosenberry T.L., Mallender W.D., Lewis T., Fletcher R.J., Guss J.M., Silman I., Sussman J.L., 2000. Three-dimensional structures of Drosophila melanogaster acetylcholinesterase and of its complexes with two potent inhibitors. Protein Sci. 9: 1063-1072

Harel, M., Quinn D.M., Nair H.K., Silman I., Sussman J.L., 1996. The X-ray structure of a transition state analog complex reveals the molecular origins of the catalytic power and substrate specificity of acetylcholinesterase. J Am Chem Soc. 118: 2340-2346

Harel, M., Sussman J.L., Krejci E., Bon S., Chanal P., Massoulie J., Silman I., 1992. Conversion of Acetylcholinesterase to Butyrylcholinesterase - Modeling and Mutagenesis. P Natl Acad Sci USA. 89: 10827-10831

Huchard, E., Martinez M., Alout H., Douzery E.J.P., Lutfalla G., Berthomieu A., Berticat C., Raymond M., Weill M., 2006. Acetylcholinesterase genes within the Diptera: takeover and loss in true flies. P Roy Soc B-Biol Sci. 273: 2595-2604

Hunter, D.D., Cashman N., Morris-Valero R., Bulock J.W., Adams S.P., Sanes J.R., 1991. An LRE (leucine-arginine-glutamate)-dependent mechanism for adhesion of neurons to S-laminin. J Neurosci. 11: 3960-3971

Ilg, T., Schmalz S., Werr M., Cramer J., 2010. Acetylcholinesterases of the cat flea Ctenocephalides felis: identification of two distinct genes and biochemical characterization of recombinant and in vivo enzyme activities. Insect Biochem Mol Biol. 40: 153-164

Johnson, G., Moore S.W., 2006. The peripheral anionic site of acetylcholinesterase: Structure, functions and potential role in rational drug design. Curr Pharm Design. 12: 217-225

Johnson, G., Moore S.W., 2013. The Leu-Arg-Glu (LRE) adhesion motif in proteins of the neuromuscular junction with special reference to proteins of the carboxylesterase/cholinesterase family. Comp Biochem Physiol Part D Genomics Proteomics. 8: 231-243

Jordan, I.K., Rogozin I.B., Wolf Y.I., Koonin E.V., 2002. Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. Genome Res. 12: 962-968

Kaplan, D., Ordentlich A., Barak D., Ariel N., Kronman C., Velan B., Shafferman A., 2001. Does "butyrylization" of acetylcholinesterase through substitution of the six divergent aromatic amino acids in the active center gorge generate an enzyme mimic of butyrylcholinesterase? Biochemistry. 40: 7433-7445

Kim, J.I., Jung C.S., Koh Y.H., Lee S.H., 2006. Molecular, biochemical and histochemical characterization of two acetylcholinesterase cDNAs from the German cockroach Blattella germanica. Insect Mol Biol. 15: 513-522

Kim, Y.H., Cha D.J., Jung J.W., Kwon H.W., Lee S.H., 2012. Molecular and Kinetic Properties of Two Acetylcholinesterases from the Western Honey Bee, Apis mellifera. Plos One. 7

Kim, Y.H., Choi J.Y., Je Y.H., Koh Y.H., Lee S.H., 2010. Functional analysis and molecular characterization of two acetylcholinesterases from the German cockroach, Blattella germanica. Insect Mol Biol. 19: 765-776

Kim, Y.H., Lee S.H., 2013. Which acetylcholinesterase functions as the main catalytic enzyme in the Class Insecta? Insect Biochem Molec. 43: 47-53

Kretsing, R.H., Nockolds C.E., 1973. Carp Muscle Calcium-Binding Protein: II. Structure Determination and General Description. J Biol Chem. 248: 3313-3326

Lang, G.J., Zhang X.H., Zhang M.Y., Zhang C.X., 2010. Comparison of catalytic properties and inhibition kinetics of two acetylcholinesterases from a lepidopteran insect. Pestic Biochem Phys. 98: 175-182

Lee, D.W., Kim S.S., Shin S.W., Kim W.T., Boo K.S., 2006. Molecular characterization of two acetylcholinesterase genes from the oriental tobacco budworm, Helicoverpa assulta (Guenee). Bba-Gen Subjects. 1760: 125-133

Lee, J., Jung I.D., Nam S.W., Clair T., Jeong E.M., Hong S.Y., Han J.W., Lee H.W., Stracke M.L., Lee H.Y., 2001. Enzymatic activation of autotaxin by divalent cations without EF-hand loop region involvement. Biochem Pharmacol. 62: 219-224

Li, F., Han Z.J., 2002. Purification and characterization of acetylcholinesterase from cotton aphid (Aphis gossypii Glover). Arch Insect Biochem. 51: 37-45

Li, Q., Liu Y., Zhao W., Chen X.Z., 2002. The calcium-binding EF-hand in polycystin-L is not a domain for channel activation and ensuing inactivation. FEBS Lett. 516: 270-278

Lu, Y.H., Park Y., Gao X.W., Zhang X., Yao J.X., Pang Y.P., Jiang H.B., Zhu K.Y., 2012. Cholinergic and non-cholinergic functions of two acetylcholinesterase genes revealed by gene-silencing in Tribolium castaneum. Sci Rep. 2

Mallender, W.D., Szegletes T., Rosenberry T.L., 2000. Acetylthiocholine binds to asp74 at the peripheral site of human acetylcholinesterase as the first step in the catalytic pathway. Biochemistry. 39: 7753-7763

Massoulíe, J., Barnard E., Chatonnet A., Bacou F., Doctor B.P., Quinn D.M., 1991. Cholinesterases: Structure, Function, Mechanism, Genetics, and Cell Biology. Am Chem Soc. 394-398

Menard, R., Storer A.C., 1992. Oxyanion Hole Interactions in Serine and Cysteine Proteases. Biol Chem H-S. 373: 393-400

Mutero, A., Fournier D., 1992. Posttranslational Modifications of Drosophila Acetylcholinesterase - Invitro Mutagenesis and Expression in Xenopus Oocytes. J

Biol Chem. 267: 1695-1700

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 Bioph Res Co. 307: 15-22

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

Olivera, S., Rodriguez-Ithurralde D., Henley J.M., 2003. Acetylcholinesterase promotes neurite elongation, synapse formation, and surface expression of AMPA receptors in hippocampal neurones. Mol Cell Neurosci. 23: 96-106

Pezzementi, L., Chatonnet A., 2010. Evolution of cholinesterases in the animal kingdom. Chem-Biol Interact. 187: 27-33

Revuelta, L., Piulachs M.D., Belles X., Castanera P., Ortego F., Diaz-Ruiz J.R., Hernandez-Crespo P., Tenllado F., 2009. RNAi of ace1 and ace2 in Blattella germanica reveals their differential contribution to acetylcholinesterase activity and sensitivity to insecticides. Insect Biochem Molec. 39: 913-919

Rosenberry, T.L., 1975. Acetylcholinesterase. Adv Enzymol Ramb. 43: 103-218

Sanson, B., Colletier J.P., Xu Y., Lang P.T., Jiang H., Silman I., Sussman J.L., Weik M., 2011. Backdoor opening mechanism in acetylcholinesterase based on X-ray crystallography and molecular dynamics simulations. Protein Sci. 20: 1114-1118

Saxena, A., Fedorko J.M., Vinayaka C.R., Medhekar R., Radic Z., Taylor P., Lockridge O., Doctor B.P., 2003. Aromatic amino-acid residues at the active and peripheral anionic sites control the binding of E2020 (Aricept (R)) to cholinesterases. Eur J Biochem. 270: 4447-4458

Schallreuter, K.U., Gibbons N.C.J., Elwary S.M., Parkin S.M., Wood J.M., 2007. Calcium-activated butyrylcholinesterase in human skin protects acetylcholinesterase against suicide inhibition by neurotoxic organophosphates. Biochem Bioph Res Co. 355: 1069-1074

Seong, K.M., Kim Y.H., Kwon D.H., Lee S.H., 2012. Identification and characterization

of three cholinesterases from the common bed bug, Cimex lectularius. Insect Mol Biol. 21: 149-159

Shapira, M., Thompson C.K., Soreq H., Robinson G.E., 2001. Changes in neuronal acetylcholinesterase gene expression and division of labor in honey bee colonies. J Mol Neurosci. 17: 1-12

Silman, I., Sussman J.L., 2005. Acetylcholinesterase: 'classical' and 'non-classical' functions and pharmacology. Curr Opin Pharmacol. 5: 293-302

Sussman, J.L., Harel M., Frolow F., Oefner C., Goldman A., Toker L., Silman I., 1991. Atomic-Structure of Acetylcholinesterase from Torpedo-Californica - a Prototypic Acetylcholine-Binding Protein. Science. 253: 872-879

Sussman, J.L., Harel M., Silman I., 1993. Three-Dimensional Structure of Acetylcholinesterase and of Its Complexes with Anticholinesterase Drugs. Chem-Biol Interact. 87: 187-197

Tamura, K., Peterson D., Peterson N., Stecher G., Nei M., Kumar S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol Biol Evol. 28: 2731-2739

Toutant, J.P., 1989. Insect Acetylcholinesterase - Catalytic Properties, Tissue Distribution and Molecular-Forms. Prog Neurobiol. 32: 423-446

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. P Roy Soc B-Biol Sci. 269: 2007-2016

Wiesner, J., Kriz Z., Kuca K., Jun D., Koca J., 2007. Acetylcholinesterases - the structural similarities and differences. J Enzym Inhib Med Ch. 22: 417-424

Wilson, A.C., Carlson S.S., White T.J., 1977. Biochemical Evolution. Annu Rev Biochem. 46: 573-639

Yue, J.X., Li J., Wang D., Araki H., Tian D., Yang S., 2010. Genome-wide investigation reveals high evolutionary rates in annual model plants. BMC Plant Biol. 10: 242

Zhang, Y.K., Kua J., McCammon J.A., 2002. Role of the catalytic triad and oxyanion hole in acetylcholinesterase catalysis: An ab initio QM/MM study. J Am Chem Soc.

124:10572-10577

Zhang, Z., Li J., Zhao X.Q., Wang J., Wong G.K., Yu J., 2006. KaKs_Calculator: calculating Ka and Ks through model selection and model averaging. Genomics Proteomics Bioinformatics. 4: 259-263

KOREAN ABSTRACT

두 곤충 아세틸콜린에스터레이즈의 진화적 기원과 양상 및 구조적 보전과 분화 분석

서울대학교 대학원

농생명공학부 곤충학 전공

차덕재

초록

아세틸콜린에스터레이즈는 (AChE) 곤충을 포함한 대부분의 동물의 콜린성 신경계 내에서 신경 전달의 중추적인 역할을 한다. 곤충은 오래 전 곤충 분화 이전에 유래되었을 것이라 생각되는 서로 다른 유전자좌 (*ace1* 및 *ace2*)로부터 발현되는 상이한 두 종의 아세틸콜린에스터레이즈 (AChE1 과 AChE2)를 가진다. 본 연구에서는, 두 *ace* 의 복제 기원과 복제 후 AChE 의 분화와 관련된 구조적 특성을 이해하기 위해 계통발생학적 분석 및 구조적 모델링 분석을 수행하였다. 알려진 절지동물의 AChE1 과 AChE2 유전자뿐만 아니라 본 실험에서 클로닝한 곤충 ace 유전자 및 유전체 정보가 알려진 하등동물에서 확보한 아세틸콜린에스터레이즈-유사 (AChE-like)

유전자를 대상으로 계통발생학적 분석을 수행하였다. 분석 결과 곤충 AChE 1 과 2 의 최종 공통조상은 편형동물과 공통조상을 공유하는 것으로 보이며, 더 나아가 절지동물의 AChE 1 과 2 계통은 편형동물의 AChE 와 함께 평행진화하고 있는 것으로 추정된다. AChE1 과 AChE2 유전자의 분화가 초래된 AChE 유전자의 복제시점은 전구동물 공통조상으로부터 탈피동물상문과 연체동물상문이 분화된 이후, 그러나 탈피동물상문이 후대로 분화되기 전인 것으로 추정된다. AChE1 계통은 AChE2 계통에 비해 현저히 낮은 진화율을 가지며 높은 도태(purifying selection)압이 주어지는 것으로 나타났는데, 이러한 결과는 복제된 후 AChE1 계통이 상대적으로 신경전달 기능에 있어 더 필수적인 기능을 유지하고 있음을 시사한다. 일부 벌목 곤충에서 관찰되는 AChE 의 기능적 전환이 이와 같은 진화율 차이로 설명될 수 없는 점을 감안하면, 이러한 AChE 의 기능적 전환은 소수의 돌연변이에 따른 AChE 기능의 극적인 변화에 의해 비교적 최근에 발생한 것으로 보여진다.

AChE1 과 AChE2 간에는 전반적인 아미노산 서열의 유사도는 낮지만 기능적으로 중요한 영역에서는 고도의 유사도를 보여주었는데, 이는 이러한 필수 아미노산 잔기에 강한 도태압이 주어진다는 것을 시사한다. 흥미롭게도, EF-hand motif 는 주로 AChE1 계통에서 발견되지만 LRE-motif 는 AChE2 계통에서 부분적으로 보전되는 것으로 나타났다. AChE2 특이적인 삽입 영역은 AChE2 계통에서만 관찰되었는데, 이는 비교적 매우 최근에, 아마도, 곤충 종의 급격한 팽창기에 도입된 것으로 추정된다. Active-site gorge 의 주요 다섯 영역[i.e., the catalytic anionic site (CAS), peripheral anionic site (PAS), acyl binding pocket (ABP), oxyanion hole 및 catalytic triad]의 아미노산 서열비교를 통해 PAS (Asp72 대 Tyr72, Tyr121 대 Met121)와 ABP(Cys288

대 Leu288) 의 아미노산 잔기에서 AChE1 과 2 간에 차이가 있음을 확인하였다. 특히 PAS 에 초점을 둔 곤충 AChE 의 active-site gorge 의 3 차원 모델링을 바탕으로 PAS 의 아미노산 변화에 의해 active-site gorge topology 에서 두 AChE 간 기능적 분화가 예상되는 사소하지만 일관된 구조 변경이 있음을 알 수 있었다. 비록 AChE1 이 상당히 느린 속도의 진화를 보임으로써 신경기능에서 필수적인 기능을 유지하는 것으로 보여지지만, 몇몇 벌목 곤충들에서 발견되는 것과 같은 AChE 의 기능전환 현상은 AChE1 의 효소활성 감소를 현저히 유발할 수 있는 소수의 돌연변이에 의해 유발된 것으로 사료된다. 이러한 현상은 비교적 최근에 여러 곤충 계통에서 지엽적으로 나타났을 것으로 추정된다.

검색어: 곤충 아세틸콜린에스터레이즈, 유전자 복제, 진화적 거리, 도태압, active-site gorge, 계통발생학적 분석, 3 차원 모델링

학번: 2011-21321