Commentary

(Special Topic)

Genomic and phylogenetic analysis of insect carboxyl/cholinesterase genes

Takuya Tsubota and Takahiro Shiotsuki*

*Invertebrate Gene Function Research Unit, National Institute of Agrobiological Sciences,
1–2 Owashi, Tsukuba, Ibaraki 305–8634, Japan*

(Received April 1, 2010; Accepted June 8, 2010)

The complete genome sequence data of insects in five orders are now available. In this commentary we provide recent genomic and phylogenetic analyses of insect carboxyl/cholinesterases (CCEs). Multiple CCEs are found in each insect species, and at present the silkworm *Bombyx mori* has the largest number and diversity of CCEs. Phylogenetic analyses of the 425 CCE sequences identified species or order specific clustering, but for some CCEs 1:1 orthologous relationships were apparent. Phylogenetic analyses will be useful for further functional analysis of CCEs and the development of species-specific insecticides and/or synergists. © Pesticide Science Society of Japan

Keywords: detoxification, phylogenetic analysis, carboxyl/cholinesterases (CCEs).

Introduction

The carboxyl/cholinesterase (CCE) superfamily is composed of functionally diverse proteins that hydrolyze carboxylic esters. CCEs are responsible not only for physiological regulation of particular endogenous compounds, such as hormones, pheromones, and acetylcholine, but also for the detoxification of exogenous compounds, such as those in the diet and in the environment. In some instances, CCEs are associated with insecticide resistance,¹⁾ and the gene identification of CCEs is significant not only for elucidation of the mechanisms of insecticide resistance but also for the development of novel insecticidal compounds pesticides. Recent completion of genome projects and the availability of expressed sequence tag (EST) clones has enabled the identification of multiple CCEs in each insect species.^{2–4)} Here we describe the genomic and phylogenetic analyses of 425 insect CCEs found in eight insect species whose genomes have been completely sequenced, and discuss the implications of the constructed phylogenetic tree.

Construction of the Phylogenetic Tree

ClustalW software was used to perform multiple sequence alignment prior to phylogenetic analysis. MEGA 4.0⁵⁾ was used to construct the phylogenetic tree using the neighbor-joining method with the JTT matrix. To evaluate branch strength in the phylogenetic tree, bootstrap analysis of 500 replicates was performed. The amino acid sequences of genes used for phylogenetic analysis were retrieved from the National Center for Biotechnological Information site (<http://www.ncbi.nih.gov/>). The accession number of each gene is listed in Supplemental Table S1. Potential signal peptides were predicted by the SignalP 3.0 program (<http://www.cbs.dtu.dk/services/SignalP>).

Phylogenetic Analysis of Insect CCEs

The database search revealed 425 insect CCEs, including those from insects whose genomes have been completely sequenced, namely the dipterans *Drosophila melanogaster*,¹⁾ *Anopheles gambiae*,⁶⁾ and *Aedes aegypti*;⁷⁾ the hymenopterans *Apis mellifera*⁸⁾ and *Nasonia vitripennis*;⁹⁾ lepidopteran *Bombyx mori*;^{10,11)} and the coleopteran *Tribolium castaneum*;⁹⁾ hemipteran *Acyrtosiphon pisum*.¹²⁾ The identified insect CCEs also included those of the lepidopterans *Helicoverpa armigera* and *Spodoptera littoralis*, in which EST clone analysis has been conducted,^{13,14)} and several other CCEs. Using these CCE amino acid sequences, a phylogenetic tree

* To whom correspondence should be addressed.

E-mail: shiotsuk@affrc.go.jp

Published online July 26, 2010

© Pesticide Science Society of Japan

Fig. 1. Phylogenetic tree of insect CCEs. MEGA 4.0⁵⁾ was used to construct the phylogenetic tree. Asterisks in the cladogram indicate bootstrap values greater than 50%. The nomenclature of the clades is basically according to previous articles.^{1,6)} The CCEs of *D. melanogaster* are colored light blue, *A. gambiae* dark blue, *A. aegypti* blue, *A. mellifera* orange, *N. vitripennis* red, *B. mori* purple, *H. armigera* reddish purple, *S. littoralis* dark purple, *T. castaneum* brown, *A. pisum* green and others black. The GenBank accession numbers of each gene are listed in Supplemental Table S1.

Table 1. Numbers of validated CCE groups annotated in the insect genomes.

Order	Diptera			Hymenoptera		Coleoptera	Lepidoptera	Hemiptera
Species	<i>Drosophila melanogaster</i>	<i>Anopheles gambiae</i>	<i>Aedes aegypti</i>	<i>Apis mellifera</i>	<i>Nasonia vitripennis</i>	<i>Tribolium castaneum</i>	<i>Bombyx mori</i>	<i>Acyrtosiphon pisum</i>
(A) α -Esterases	13	16	22	8	13	26	15	5
(B) Glutactin and uncharacterized (I)	4	4	8	0	0	0	0	0
(C) Lepidopteran JHEs and others	0	0	0	0	0	0	7	0
(D) Uncharacterized (I)	1	3	6	1	1	1	1	1
(E) Lepidopteran esterases	0	0	0	0	1	0	35	0
(F) Acetylcholinesterases	1	2	2	2	2	2	2	2
(G) Gliotactins	1	1	1	1	1	1	1	1
(H) Uncharacterized (II)	1	1	1	1	1	1	1	1
(I) Neuroligins	4	5	4	5	5	4	6	3
(J) β -Estrases (I)	1	1	1	2	1	1	1	9
(K) Integument esterases	3	0	1	1	4	2	2	0
(L) Nonlepidopteran JHEs and like enzymes	2	5	7	0	0	1	0	0
(M) β -Estrases (II)	2	4	1	2	11	7	1	6
(N) Neurotactins	2	2	2	1	1	2	1	0
Total	35	44	56	24	41	48	73	28

was constructed (Fig. 1). The phylogenetic tree segregated the CCEs into 14 groups (A to N), which were named basically as described by Oakeshott (Table 1 and Fig. 1).¹⁾ The number of CCEs found in each group of insect species analyzed is shown in Table 1. Detailed information (*i.e.*, GenBank accession number, sequence of the catalytic triad, and putative signal peptide for secretion) of each CCE used in phylogenetic analysis is given in Supplemental Table S1. A detailed explanation of the functional characteristics of each CCE group is given by Oakeshott.¹⁾ We provide below some of the key characteristics and members of each of the 14 major groups of insect CCEs.

(A) α -Esterases

The α -esterase group is composed of several subgroups of order-specific esterases; namely, dipteran α -esterase (I, II, III), hymenopteran α -esterase (I, II), coleopteran α -esterase (I, II), hemipteran α -esterase, and lepidopteran α -esterase (Fig. 1). Some of the dipteran α -esterases (III) are known to be involved in organophosphate (OP) resistance.^{9,15–19)} A hymenopteran CCE that is involved in OP resistance²⁰⁾ also belongs to this group. At least one of the lepidopteran CCEs in this group is hypothesized to function in the metabolism of toxins in the diet,²¹⁾ but whether this activity is directly related to insecticide resistance is unknown. Hemipteran whitefly CCE, which is associated with OP resistance,²²⁾ belongs to

this group, and CCEs of these insect species are in the hemipteran subgroup in the α -esterase group.

(B) Glutactin and uncharacterized CCEs

The only member of this clade that has been investigated functionally is glutactin of *D. melanogaster*. Most of the CCEs in this group are those of mosquitoes with catalytic residues (Fig. 1 and Supplemental Table S1).

(C) Lepidopteran JHEs

Juvenile hormone esterases (JHEs) form another group of CCEs with a clearly defined physiological substrate. JHEs have been extensively characterized from many insect orders. Previous analyses indicate that lepidopteran JHEs constitute an independent cluster that is distinct from JHEs from other insect orders,²⁾ and this is also evident in Fig. 1. Another sub-cluster close to lepidopteran JHEs was also found.

(D) Uncharacterized CCEs

It was found that one subgroup is in a clear 1 : 1 orthologous relationship without catalytic serine, and the other is a dipteran subgroup mostly with a GxSxG motif at the catalytic active site (Supplemental Table S1).

(E) Lepidopteran esterases

The lepidopteran esterase group is the largest among the

groups composed of just one insect order. Most CCEs in this group have a putative catalytic active center. Some CCEs in this group are known to be expressed in the antenna,²³⁾ while others (from *H. armigera*) are expressed in the gut.

(F) Acetylcholinesterases

In addition to JHE, acetylcholinesterase (AChE) is one of the most extensively analyzed CCEs. The physiological substrate of this CCE is well-defined as acetylcholine. In Fig. 1, only some of the AChEs are shown and it is known to be identified in more abundant species. In general, at least two AChEs are found in each insect species, and they belong to distinct phylogenetic clusters (Fig. 1). An exception is the AChE of *D. melanogaster*, which has only one known AChE (*DmAChE*, *ace*). A commonly accepted nomenclature for AChE has not been established; thus, we proposed *ace*-orthologous-AChE (AO-AChE) for AChEs that belong to the same cluster as *DmAChE* (*ace*), and *ace* paralogous-AChE (AP-AChE) for AChEs located in a different cluster from *ace*.²⁴⁾

(G) Gliotactins

There is a clear 1:1 orthologous relationship for gliotactins. All have a GPGAG sequence motif at the active site, except *A. pisum* (Supplemental Table S1).

(H) Uncharacterized CCEs

A clear 1:1 orthologous relationship is found in this uncharacterized CCE group. The physiological function of CCEs in this group is not known.

(I) Neuroligins

The neuroligin group is composed of five subclusters, Nlg-1–5 (Fig. 1). In most cases, each insect species has five neuroligins. Only one of these genes is represented in each subcluster in the phylogenetic tree; however, there are exceptions: *D. melanogaster* and *A. aegypti* each lack Nlg-5 and, interestingly, *B. mori* has two Nlgs that are localized in the Nlg-4 cluster, as shown previously.^{10,11)} The chromosomal location of *nlg*s is conserved among insect species,²⁾ but we found this is not for *B. mori*.¹¹⁾

(J) β -Esterases (I)

CCEs in this group show a clear 1:1 relationship in one subgroup and include only *A. pisum* in the other subgroup (Fig. 1). The hemipteran *Nilaparvata lugens* CCE in this group is associated with resistance to OPs.²⁵⁾

(K) Integument esterases

The group is composed of small CCE clusters from each insect species except for *A. gambiae* (Fig. 1). The integument esterase of *Antheraea polyphemus*²³⁾ is hypothesized to clear pheromones from the external surface of the antenna. This is the only known CCE in this cluster.

(L) Non-lepidopteran JHEs

JHEs in this group show a clear 1:1 orthologous relationship. Interestingly, JHE of *Gryllus assimilis*, an orthopteran species, also belongs to this group (Fig. 1). Of these sequences, *DmJHE* is considered to function as JHE, whereas *DmCG8424* (JHEdup) putatively has a non-JHE function.²⁶⁾ There is a large subgroup for mosquitoes, and *A. aegypti* CCE 05200 protein is able to degrade JH *in vitro*.²⁷⁾

(M) β -Esterases (II)

The β -esterase group is composed of CCE clusters from multiple insect species. Of the CCEs in this cluster, the pheromone-degrading esterases of the lepidopteran *A. polyphemus*²³⁾ and the coleopteran *Popillia japonica* have been identified previously.²⁸⁾ JHE of *A. mellifera*²⁹⁾ is also included in this cluster, although its enzymatic activity is yet to be shown. This phylogeny suggests that the hymenopteran JHE may have evolved independently of JHEs from other insect orders. *D. melanogaster* CG6917, Est-6, is considered to have a role in reproduction, while CG17148, corresponding to Est-7, may be a functionless pseudogene.^{30–32)} This group also includes a subgroup with hemipteran CCEs, that is, five CCEs of *A. pisum* and *Myzus persicae* E4 and FE4, that are associated with resistance to OPs.^{33–35)}

(N) Neurotactins

This group is an outgroup with regard to the other CCEs (Fig. 1). A clear 1:1 orthologous relationship is evident in the neurotactin cluster, but there are several CCEs, which are less conserved with other CCEs, so their location in the phylogenetic tree may not be reliable, as evidenced by the low bootstrap value (Fig. 1 and data not shown). None possesses putative catalytic residues.

Conclusion

In this review we have analyzed the phylogeny of 425 insect CCEs primarily identified in insect genomes that have been completely sequenced. Some CCEs, in particular the noncatalytic CCEs, show a 1:1 orthologous relationship, while others are clustered in each insect species or order. CCEs in the former groups may have common functions in various insects, such as is the case for AChEs, while those in the latter groups may have order or species specific roles. The present phylogenetic tree will be improved by the incorporation of novel CCE genes identified from the genome sequences of insects from other orders. This analysis will be useful for functional analyses of CCEs and for the development of species-specific pesticides.

Supplemental Table S1 is available in the online publication at <http://www.jstage.jst.go.jp/browse/jpestics/>.

References

- 1) J. G. Oakeshott, C. Claudianos, P. M. Campbell, R. D. New-

- comb and R.J. Russell: "Comprehensive Molecular Insect Science-Pharmacology," Vol. 5, ed. by L. I. Gilbert, K. Iatrou and S. S. Gill, Elsevier, Oxford, pp. 309–381, 2005.
- 2) C. Claudianos, H. Ranson, R. M. Johnson, S. Biswas, M. A. Schuler, M. R. Berenbaum, R. Feyereisen and J. G. Oakeshott: *Insect Mol. Biol.* **15**, 615–636 (2003).
- 3) K. Mita, M. Morimyo, K. Okano, Y. Koike, J. Nohata, H. Kawasaki, K. Kadono-Okuda, K. Yamamoto, M. G. Suzuki, T. Shimada, M. R. Goldsmith and S. Maeda: *Proc. Natl. Acad. Sci. USA* **100**, 14121–14126 (2003).
- 4) Q. Xia, Z. Zhou, C. Lu, D. Cheng, F. Dai, B. Li, P. Zhao, X. Zha, T. Cheng, C. Chai, G. Pan, J. Xu, C. Liu, Y. Lin, J. Qian, Y. Hou, Z. Wu, G. Li, M. Pan, C. Li, Y. Shen, X. Lan, L. Yuan, T. Li, H. Xu, G. Yang, Y. Wan, Y. Zhu, M. Yu, W. Shen, D. Wu, Z. Xiang, J. Yu, J. Wang, R. Li, J. Shi, H. Li, G. Li, J. Su, X. Wang, G. Li, Z. Zhang, Q. Wu, J. Li, Q. Zhang, N. Wei, J. Xu, H. Sun, L. Dong, D. Liu, S. Zhao, X. Zhao, Q. Meng, F. Lan, X. Huang, Y. Li, L. Fang, C. Li, D. Li, Y. Sun, Z. Zhang, Z. Yang, Y. Huang, Y. Xi, Q. Qi, D. He, H. Huang, X. Zhang, Z. Wang, W. Li, Y. Cao, Y. Yu, H. Yu, J. Li, J. Ye, H. Chen, Y. Zhou, B. Liu, J. Wang, J. Ye, H. Ji, S. Li, P. Ni, J. Zhang, Y. Zhang, H. Zheng, B. Mao, W. Wang, C. Ye, S. Li, J. Wang, G. K.-S. Wong and H. Yang: *Science* **306**, 1937–1940 (2004).
- 5) K. Tamura, J. Dudley, M. Nei and S. Kumar: *Mol. Biol. Evol.* **24**, 1596–1599 (2007).
- 6) H. Ranson, C. Claudianos, F. Orrelli, C. Abgrall, J. Emingway, M. V. Sharaknova, M. F. Unger, F. H. Collins and R. Feyereisen: *Science* **298**, 179–181 (2002).
- 7) C. Strode, C. S. Wondji, J. P. David, N. J. Hawkes, N. Lumjuan, D. R. Nelson, D. R. Drane, S. H. P. P. Karunaratne, J. Hemingway, W. C. Black IV and H. Ranson: *Insect Biochem. Mol. Biol.* **38**, 113–123 (2008).
- 8) C. Claudianos, H. Ranson, R. M. Johnson, S. Biswas, M. A. Schuler, M. R. Berenbaum, R. Feyereisen and J. G. Oakeshott: *Insect Mol. Biol.* **15**, 615–636 (2006).
- 9) J. G. Oakeshott, R. M. Johnson, M. R. Berenbaum, H. Ranson, A. S. Cristino and C. Claudianos: *Insect. Mol. Biol.* **19** (Suppl. 1): 147–163 (2010).
- 10) Q.-Y. Yu, C. Lu, W.-L. Li, Z.-H. Xiang and Z. Zhang: *BMC Genomics* **10**, 553 (2009).
- 11) T. Tsubota and T. Shiotsuki: *BMC Genomics* **11**, 377 (2010).
- 12) J. S. Ramsey, D. S. Rider, T. K. Walsh, M. De Vos, K. H. J. Gordon, L. Ponnala, S. L. Macmil, B.A. Roe and G. Jander: *Insect Mol. Biol.* **19** (Suppl. 2): 155–164 (2010).
- 13) M. G. Teese, P. M. Campbell, C. Scott, K. H. J. Gordon, A. Southon, D. Hovan, C. Robin, R. J. Russell and J. G. Oakeshott: *Insect Biochem. Mol. Biol.* **40**, 1–16 (2010).
- 14) N. Durand, G. Carot-Sans, T. Chertemps, N. Montagne, E. Jacquin-Joly, S. Debernard and M. Maibeche-Coisne: *Insect Mol. Biol.* **19**, 87–97 (2010).
- 15) A. Vaughan, N. Hawkes and J. Hemingway: *Biochem. J.* **325**, 359–365 (1997).
- 16) C. Tittiger and V. K. Walker: *Biochem. Genet.* **35**, 119–138 (1997).
- 17) R. D. Newcomb, P. M. Campbell, D. L. Ollis, E. Cheah, R. J. Russell and J. G. Oakeshott: *Proc. Natl. Acad. Sci. USA* **94**, 7464–7468, (1997).
- 18) C. Claudianos, R. J. Russell and J. G. Oakeshott: *Insect Biochem. Mol. Biol.* **29**, 675–686 (1999).
- 19) F. D. Guerrero: *Insect Biochem. Mol. Biol.* **30**, 1107–1115 (2000).
- 20) Y. C. Zhu, A. K. Dowdy and J. E. Baker: *Insect Biochem. Mol. Biol.* **29**, 417–425 (1999).
- 21) T. Tsubota, M. Shimomura, T. Ogura, A. Seino, T. Nakakura, K. Mita, T. Shinoda and T. Shiotsuki: *Insect Biochem. Mol. Biol.* **40**, 100–112 (2010).
- 22) M. Alon, F. Alon, R. Nauen and S. Morin: *Insect Biochem. Mol. Biol.* **38**, 940–949 (2008).
- 23) Y. Ishida and W. S. Leal: *Insect Biochem. Mol. Biol.* **32**, 1775–1780 (2002).
- 24) A. Seino, T. Kazuma, A.-J. Tan, H. Tanaka, Y. Kono, K. Mita and T. Shiotsuki: *Pestic. Biochem. Physiol.* **88**, 92–101 (2007).
- 25) G. J. Small and J. Hemingway: *Insect Mol. Biol.* **9**, 647–653 (2000).
- 26) E. J. Crone, T. D. Sutherland, P. M. Campbell, C. W. Coppin, R. J. Russell and J. G. Oakeshott: *Insect Biochem. Mol. Biol.* **37**, 540–549 (2007).
- 27) H. Bai, P. Ramaseshadri and S. R. Palli: *Insect Biochem. Mol. Biol.* **37**, 829–837 (2007).
- 28) Y. Ishida, and W. S. Leal and *Proc. Natl. Acad. Sci. USA* **102**, 14075–14079 (2005).
- 29) A. Mackert, A. M. do Nascimento, M. M. G. Bitondi, K. Hartfelder and Z. L. P. Simões: *Comp. Biochem. Physiol., Part B: Biochem. Mol. Biol.* **150**, 33–44 (2008).
- 30) D. B. Meikle, K. B. Sheehan, D. M. Phillis and R. C. Richmond: *J. Insect Physiol.* **36**, 93–101 (1990).
- 31) R. C. Richmond, K. M. Nielsen, J. P. Brady and E. M. Snella: "Ecological and Evolutionary Genetics of *Drosophila*," ed. by J. S. F. Barker, W. T. Starmer and R. J. MacIntyre, Plenum, New York, pp. 273–293, 1990.
- 32) E. S. Balakirev, V. R. Chechetkin, V. V. Lobzin and F. J. Ayala: *Genetics* **164**, 533–544 (2003).
- 33) A. L. Devonshire and R. M. Sawicki: *Nature* **280**, 140–141 (1979).
- 34) L. M. Field, A. L. Devonshire and B. G. Forde: *Biochem. J.* **251**, 309–312 (1988).
- 35) L. M. Field, M. S. Williamson, G. D. Moores and A. L. Devonshire: *Biochem. J.* **294**, 569–574 (1993).