BIOINFORMATION
Discovery at the interface of physical and biological sciences

open access www.bioinformation.net Hypothesis

Volume 8(9)

Identification and classification of detoxification enzymes from *Culex quinquefasciatus* (Diptera: Culicidae)

BP Niranjan Reddy[1, 2], B Prasad Rao[1, 2], GBKS Prasad[2] & K Raghavendra[1]*

[1]Vector Control Division, Insectary and Insecticide resistance lab, National Institute of Malaria Research (ICMR), Sector 8, Dwarka, New Delhi, India; [2]School of Studies in Biotechnology, Jiwaji University, Gwalior, MP, India; K Raghavendra – Email: kamarajur2000@yahoo.com; Phone: 91- 11- 25307 205; Fax: 91-11-25-307 177; *Corresponding author

Received April 12, 2012; Accepted April 28, 2012; Published May 15, 2012

Abstract:
Molecular characterization of the insecticide resistance has become a hot research topic ever since the first disease transmitting arthropod (*Anopheles gambiae*) genome sequence has unveiled in 2002. A recent publication of the *Culex quinquefasciatus* genome sequence has opened up new opportunities for molecular and comparative genomic analysis of multiple mosquito genomes to characterize the insecticide resistance. Here, we utilized a whole genome sequence of *Cx. quinquefasciatus* to identify putatively active members of the detoxification supergene families, namely cytochrome P450s (P450s), glutathione-*S*-transferases (GSTs), and choline/carboxylesterases (CCEs). The *Culex* genome analysis revealed 166 P450s, 40 GSTs, and 62 CCEs. Further, the comparative genomic analysis shows that these numbers are considerably higher than the other dipteran mosquitoes. These observed species-specific expansions of the detoxification super gene family members endorse the popular understanding of the involvement of these gene families in protecting the organism against multitudinous classes of toxic substances during its complex (aquatic and terrestrial) life cycle. Thus, the generated data set may provide an initial point to start with to characterize the insecticide resistance at a molecular level which could then lead the development of an easy to use molecular marker to monitor the incipient insecticide resistance in field environs.

Keywords: *Culex quinquefasciatus*, detoxification enzymes, cytochrome P450 (P450), glutathione-S-transferase (GST), choline/carboxylesterase (CCE)

Background:
The present day sustainable vector control activities are primarily dependent on use of chemical insecticides. Because of this reason, almost all of the mosquito-vectors around the globe have successfully learned to defend themselves from the existing insecticides that are being recommended by WHO. A decade ago, due to lack of genome sequence information from the disease transmitting mosquito species, it was challenging to understand the molecular aspects behind the evolution of insecticide resistance. The genome sequence of *Anopheles gambiae* (African malaria mosquito) was first published in 2002 **[1]**, followed by *Aedes aegypti* in 2007 **[2]** and very recently *Culex quinquefasciatus* **[3]** genome in 2010, has opened-up new possibilities to look into the insecticide resistance at a molecular level. Many conclusive reports on the candidate genes behind the molecular mechanisms of insecticide resistance from African malaria vector, *An. gambiae* have been published **[4-6]**; however, translating these studies for practical application is still a due. The present *Cx. quinquefasciatus* genome sequence further enhanced the capability to understand the molecular science of insecticide resistance through the comparative genomic studies. *Culex* species have acted as a model organism to study the population genetics and evolution of the insecticide resistance both in field and laboratory conditions **[6-8]**. Many of the long-term studies on monitoring the role of different processes that are important for insecticide resistance have been

conducted on *Culex* species **[8].** Some of important aspects of *Culex* species research are; origin of new adaptive mutations against the insecticide used in the vector control program, and their interaction with the existing insecticide resistance mutations, interaction with the environment, cost of a mutation in the presence and absence of an insecticide, establishment and migration of the mutations to a wide geographical areas, pleiotropic effect of a gene mutation on the fitness characteristics of the mosquito, etc **[7].** The *Culex* species being as an urban vector, many of its control efforts are focused on the usage of organophosphorous group based larvicides. Because of this the *Culex* species has been extensively investigated for the mechanisms behind the OP resistance. The established insecticide resistance mechanisms for OP compounds includes both target site mutations in acetylcholinesterase (*ace*) gene and over production of the detoxification enzymes, majorly esterases through gene amplification. The resistance mechanisms against insecticides in *Culex* species are similar to that of other disease causing mosquito-vectors (*An. gambiae* and *Ae. Aegypti)* and are grouped into two major groups; target site insensitivity and up-regulation of the detoxification enzymes. The detoxification enzymes consists of hundreds of genes from three supergene families, namely cytochrome P450s (P450s) or monoxygenases, glutathione-*S*-transferases (GSTs), and carboxyl/cholinesterases (CCEs). A plethora of information is available on the insecticide resistance in public domains that confirms the important role of detoxification enzymes (P450, GST, and CCE) in the evolution of insecticide resistance. These enzymatic groups possess a capability to virtually detoxify myriad classes of xenobiotics that are found in nature. In contrary, the target site mutations contribute resistance against a particular selected insecticide. Due to continuous vector control efforts using various strategies by placing the chemical insecticides at a center stage has created multiple insecticide selection pressure on the mosquito vectors. This particular situation has resulted in the appearance of mosquito isolates that are resistant to more than one insecticide. One such mosquito species with multiple insecticide resistance mechanisms is *Culex*. Nonetheless, this species has shown multiple resistances to all of the four major classes of insecticides, namely organochlorines, organophosphates, carbamates, and pyrethroids, especially in field situations **[9]**. These aspects coupled with the availability of a past history of chemical control activities and multiple-insecticide resistance information makes this species special to investigate the molecular insecticide resistance aspects in-depth. Giving a due importance to dissect the molecular basis of insecticide resistance through analysis of the detoxification supergene families, here, we utilized *Cx. quinquefasciatus* genome sequence to *in silico* fish-out the detoxification enzymes that belong to three major groups; P450s, GSTs, and CCEs. The aim of the present study was to investigate the detoxification enzymes from the *Culex* whole genome sequence and to classify them into respective gene families such that the information could be easily retrieved for further studies on delineation of the insecticide resistance processes. In addition, the comparative genomic analysis of *Culex* detoxification genes with *Drosophila*, *Aedes* and *Anopheles* was performed. Apart from common disease vectors (*Aedes* and *Anopheles*-model organisms for host-parasite interaction), the *Drosophila* was selected for comparative genomics due to its importance as a model organism.

Methodology:

Utilizing the published sequences of P450s, GSTs, and CCEs from *An. gambiae* and *D. melanogaster*, the whole genome sequence of *Cx. quinquefasciatus* (*Cx. quinquefasciatus* JHB CpipJ1.2, June 2008 data base) was scanned using the tBLASTn with default parameters (E-value-10, word size-3, similarity matrix-Blosum62, Gap penalties-opening: 11 and extension: 1) as a first step to find out the putatively active detoxification enzymes. Following which, the special characteristics of each of the enzyme group, namely cysteine heme-iron ligand signature i.e. conserved FXXGXXXCXG motif and ~ 500 amino acids (a. a.) protein length for P450s; SNAIL/TRAIL motif and ~ 200 a. a. protein length for GSTs; and catalytic triad sequence (Ser-His-Glu) and ~ 500 a. a. protein sequence length for CCEs, were applied to preliminarily confirm the status to their candidature. Following this, each of the putatively identified sequence was evaluated for having complete protein domains' structure and absence of multiple domains that are characteristics of a functional protein. Final list of detoxification enzymes were tabulated by removing the proteins with incomplete domain structure (possible pseudogene), and/or with multiple domains. For this, Conserved Domain (CD) search was performed against Conserved Domain Database using the protein sequence as a query. CD-search uses the RPS-BLAST to scan the pre-calculated PSSM. The result of CD search are graphical that identify and enlist the domain architecture present in the given protein sequence (the CD-search can be performed through NCBI and can be accessed at http://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml). Any partial domains in a given protein can be identified through this procedure.

Further to this, the confirmed enzymes were classified into various gene families based on their phylogenetic relationship with the classified gene family members' from *An. gambiae* and/or *Ae. aegypti*. The phylogenetic analysis of *Cx. quinquefasciatus* detoxification enzymes was performed by downloading the *An. gambiae* P450s and GSTs from VectorBase database AgamP3 build available at http://www.vectorbase.org/Anopheles_gambiae/Info/Index and the *D. melanogaster* esterase sequences from FB2012_02 available at http://www.flybase.org. The P450 sequences of the *Culex* and *Anopheles*; GST sequences of *Culex* and *Anopheles*; and CCE sequences of *Culex* and *Drosophila* were analyzed for drawing the evolutionary relationships among the genes. For all the phylogenetic analysis MEGA4.0 software was employed as described in Raghavendra et al. **[10]**. To construct phylogeny, the final protein multiple sequence alignment was used as an input with Jones-Taylor-Thornton (JTT) evolutionary model to assess the genetic distance between various taxa. Finally, the obtained phylogenies were statistically evaluated using the bootstrap test with 500 replicates.

Discussion:

The recent genomic sequences from *Anopheles* and *Aedes* species have enabled us to utilize the genomic sequence to develop and standardize the procedure to fish-out the detoxification enzymes. In the year 2010, Arensburger et al. have published the genome sequence of *Cx. quinquefasciatus* **[3],** and to date, to the best of our knowledge the *Culex* detoxification enzymes' related information is yet to be made available. Although the post-genomic era has brought simplifications in the way to analyze the genomic data, the manual screening and annotation

is necessary in order to obtain specific function related information from the genomes **[11].** Ever since the first disease causing mosquito genome has completed, the two mosquito biology research areas, namely insecticide resistance and understanding of the processes or basic genomic elements that are responsible for blockage of the pathogen growth inside the mosquito have flourished in comparison to the other scientific areas. The genomic and bioinformatic analysis of *Cx. quinquefasciatus* genome revealed 166 P450s, 40 GSTs, and 62 CCEs **Table 1 (see supplementary material) & (Figure 1)**. The total numbers of each of the supergene family are significantly expanded in *Culex* genome as compared to other sequenced dipteran mosquito species **Table 2a (see supplementary material)**. The differential expansion of the detoxification enzymes in *Culex* species is due to the expansion of the insecticide resistance causing gene families **Table 2b (see supplementary material).**

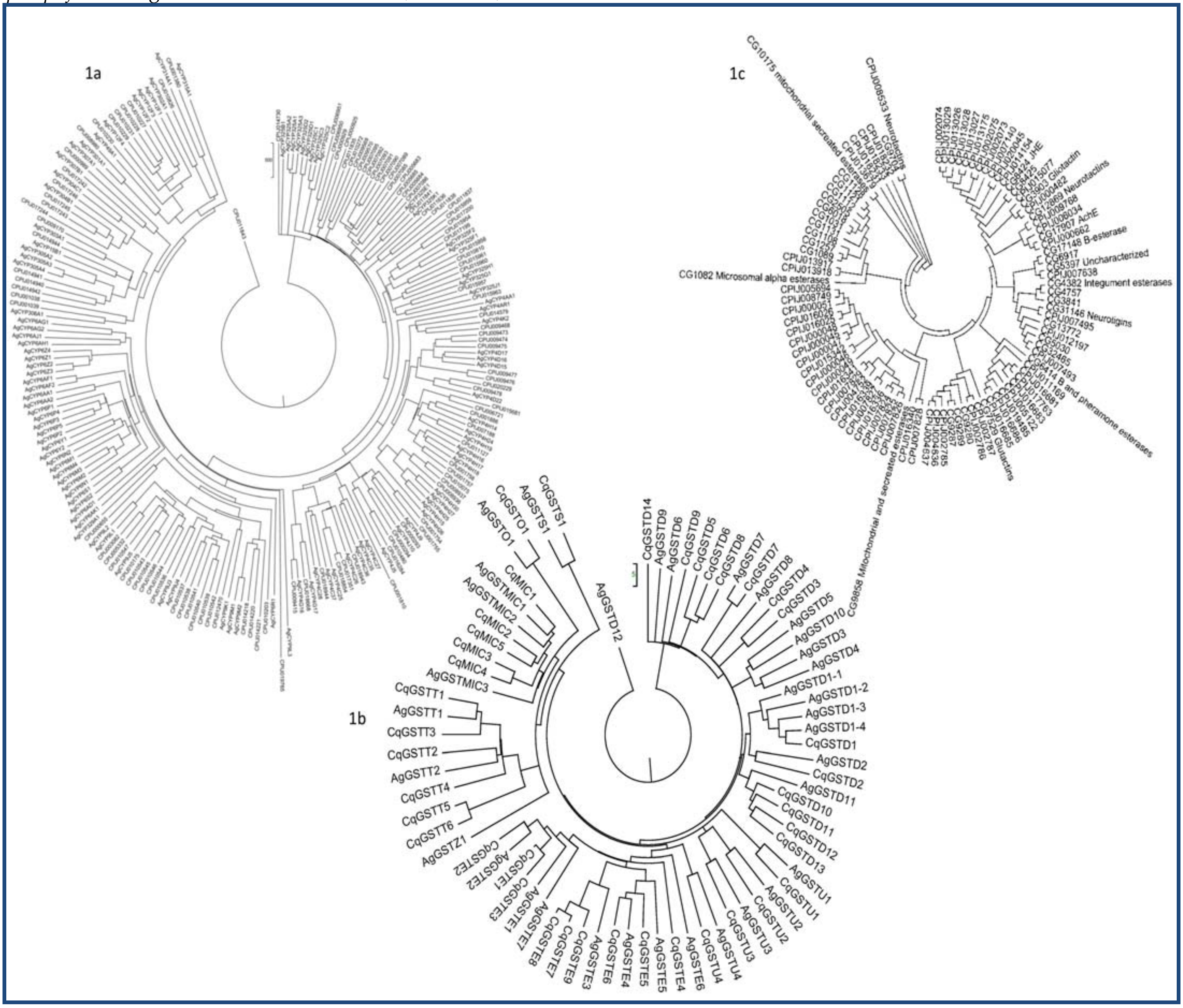

Figure 1a: The NJ based phylogenetic analysis of 166 CuP450s and 105 AgP450s; **Figure 1b**: The NJ based phylogenetic analysis of glutathione-S-transferases from *An. gambiae* (31 genes) and *Cx. quinquefasciatus* (40); **Figure 1c**: The NJ based carboxylesterase *Ae. aegypti* (49 genes) and *Cx. quinquefasciatus* (62).

Today it is known that CYP3 and CYP4 clan members from P450s, Delta-Epsilon gene family members from GSTs, and alpha-beta esterases from general esterases are primarily responsible for the insecticide resistance **[6].** The substantial expansion of detoxification enzymes in *Culex* might have occurred due to the species breeding preference to highly polluted water. Due to this, *Culex* mosquitoes might get exposed (some chemicals might have similar chemical structures as that of the insecticides that are being used in the vector control programs) to numerous kind of chemical molecules during the early stages of their development (aquatic phase of life cycle). David *et al.* **[12]** showed that larval breeding site has a significant influence over detoxification responses of the mosquitoes to various pesticides. According to Liska **[13]** the detoxification processes can be classified into two steps: (1) functionalization- where the foreign compound(s) get oxidized to create a reactive site (electrophilic site) by the phase I detoxification enzymes (P450s and esterases), (2) conjugation – utilizing the reactive site facilitated by the phase I system a water soluble compound will be added to the reactive site by

the GSTs. This particular action results in the bio-transformation of lipophilic xenobiotic compounds into a more water soluble byproducts and thus facilitates in easy excretion **[13, 14].**

The phase I detoxification enzymes (P450s) are categorized into four clans, viz. CYP2, CYP3, CYP4, and mitochondrial P450s **[15].** Of these, mitochondrial and CYP2 clans are important for performing the developmental regulations by facilitating in the production of juvenile hormone, while CYP3 and CYP4 clan members are important for detoxification of the xenobiotics. Furthermore, each of these four clans are divided into individual gene families based on the protein sequence identity (>55% and >40-55% protein identity is used to define a subfamily and a gene family, respectively). Of 16 P450 gene families identified in *An. gambiae*, the CYP4, CYP6, CYP9, and CYP325 gene families are important for insecticide resistance in insects **[16-22].** Due to the involvement of CYP2 and mitochondrial P450s in developmentally important functions, these gene families are least prone to gene duplications (Table 1a); in contrary CYP3 and CYP4 clan gene families' that are implicated in metabolizing and detoxification of foreign compounds are expanded grossly in the mosquito genomes (Table 2b). Of a total 166 *Culex* P450s, 77 and 66 genes belong to CYP3 and CYP4 clans, respectively that accounts for 86% of total P450s **Table 2b (see supplementary material) & (Figure 1a).** The comparative genomic analysis of P450s from the dipteran species revealed that CYP3 and CYP4 clans are alone contribute to 64-86% of the P450s **Table 2b (see supplementary material)**. The gene families (CYP2 and mitochondrial) that are responsible for developmental regulation are not expanded **Table 2a (see supplementary material)**. Furthermore, they have shown 1:1 secure orthologs in dipteran species (data not shown). The comparative genomic analysis shows that *Drosophila* has got least numbers of CYP3 and CYP4 clan members as compared with the other disease causing dipteran species. This may be due to restricted exposure of Drosophila to pesticides. In contrast, the rest of the three disease vectors are primary targets of human interventions to control the disease/s that are basically centered in using the insecticides to kill the vectors.

The GST supergene family of insects is divided into eight (that include one unclassified GST class) classes, namely Delta, Epsilon, Theta, Sigma, Omega, Iota, Unclassified, and microsomal GSTs. Of these, Delta-and Epsilon-classes are important for the detoxification of xenobiotics **[11, 23].** The GST supergene family forms the phase II detoxification system where the conjugation reactions occur to render the xenobiotics more soluble or to make them sequestered so that xenobiotics or insecticides will become inactive in the cell. In *Culex* 57% (23/40) of total GSTs belong to the Delta-Epsilon class **Table 2b (see supplementary material) & (Figure 1b).** The comparative GST supergene family analysis suggests that 55 to 66% of total GSTs belong to the Delta-Epsilon class. Delta-Epsilon classes are primarily responsible for detoxification process while the function of other GST classes is yet to be elucidated **[23]**. The comparative genomic analysis show that classes Iota, Unclassified, and Microsomal GSTs are absent from the *Drosophila,* while class Zeta is absent from the *Culex*. The maximum number of variations in the gene copy numbers is observed in Delta- and Epsilon-classes **Table 1 (see supplementary material)**.

Esterases are classified into two major groups based on their cellular functions; (a) metabolic enzymes (dietary detoxification, hormone and pheromone processing esterases) (b) neuro/developmental functions **[24, 25].** These two groups are further classified into gene families, namely alpha, beta, acetylcholinesterases, neurotactin, neuroligin, gliotactin, glutactin, juvenile, and unknown (still to classify) gene families **[24, 25].** Of which alpha esterases (phase I detoxification enzymes) are majorly involved in xenobiotic-detoxification processes. Of a total 62 CCEs identified in *Culex* genome 50% (28 genes) of which are belonging to the alpha-esterases **Table 2 (see supplementary material) & (Figure 1c).** The comparative analysis of dipteran CCEs revealed that 30 to 50% of total CCEs are alpha-esterases. The rest of the gene families identified in the *Culex* genome and their respective copy numbers are given in **Table 1 (see supplementary material).** As described in the methods section, the classification of the *Culex* esterases were preformed based on the phylogenetic relationship with the reference *Drosophila* esterases **(Figure 1c).** Similar to the case with P450s and GSTs, the gene families that are responsible for the detoxification of the xenobiotics are expanded in esterases, i.e., alpha esterases. The highest number of alpha esterases is reported from *Culex* (28 genes). The comparative genomic analysis show that except *Drosophila* (3 genes) rest of the dipteran mosquitoes have lost the integument esterases during the evolution. Furthermore, there is considerable expansion of the juvenile hormone gene copies observed in the mosquitoes (10-13) while only three were reported from *Drosophila*. Interestingly, in all the analyzed species, a single ortholog gene copy that is classified under uncharacterized esterases is reported **Table 2a (see supplementary material).** Finally, the *Cx. quinquefasciatus* detoxification enzyme's data further corroborate the popular understanding that detoxification enzymes undergo adaptive evolution to satisfy the need of an organism for its broad environmental adaptability **[26]**. Furthermore, it is evident from the analysis that the strong Darwinian selection will favor the organism to evolve new functions through the extensive duplication of genes. Such a mechanism is evident from the significant expansion (locally and globally in the genome) of CYP4, CYP6, CYP9, and CYP325 cytochrome P450 gene families, Delta and Epsilon GSTs, and alpha esterases **Table 2b (see supplementary material)** that are implicated in causing insecticide resistance in the class Insecta.

In conclusion, the present study identified 268 detoxification genes that belong to P450, GST, and CCE supergene families. This is the first report on the full information about these genes in *Cx. quinquefasciatus.* These data may act as a raw material for further studies on insecticide resistance. Molecular characterization of the detoxification enzymes involves retrieval, identification, confirmation, and transcriptional profiling of the genes. This needs an expert curated detoxification gene's data set, and the process involved herein is not straightforward **[3, 24, 25, 27]**. The comprehensive listing of the detoxification enzymes along with their groupings may helps in easy *in silico* retrieval of the enzyme related information for molecular characterization of insecticide resistance. The present information may also help in understanding evolution of the detoxification supergene families that are directly and/or indirectly responsible for insecticide resistance in insects. However, as the generated information on *Culex* detoxification genes is based on *in silico* analyses and thus further studies are needed to confirm the

exact number of active genes and their functional roles in various biological processes in *Culex* mosquitoes.

Acknowledgement:
BPNR was supported by Council of Scientific and Industrial Research (CSIR)-SRF fellowship. BPR was supported by Indian Council of Medical Research (ICMR-SRF) fellowship. We are highly grateful to the continuous support provided by the Director, National Institute of Malaria Research, New Delhi.

References:

[1] Holt R A *et al. Anopheles gambiae Science.* 2002 **298**: 129 [PMID: 12364791]
[2] Nene V *et al. Aedes aegypti Science.* 2007 **316**: 1718 [PMID: 17510324]
[3] Arensburger P *et al. Science.* 2010 **330**: 86 [PMID: 20929810]
[4] Awolola T S *et al. Trans R Soc Trop Med Hyg.* 2008 **103**: 1139 [PMID: 18829056]
[5] Donnelly M J *et al. Trends Parasitol.* 2009 **25**: 213 [PMID: 19369117]
[6] Reddy B P N *et al. Parasit Vectors.* 2012 **5**: 63 [PMID: 22463777]
[7] Labbé P *et al.* Massachusetts, USA: Elsevier; 2011
[8] Labbé P *et al. Culex pipiens PLoS Genet.* 2007 **3**: e205 [PMID: 18020711]
[9] Raghavendra K *et al. Acta Trop.* 2011 **118**: 50 [PMID: 21315680]
[10] Raghavendra K *et al. A dthali Curr Sci.* 2009 **97**: 923
[11] Reddy B P N *et al. Comput Biol Chem.* 2011 **35**: 114 [PMID: 21514893]
[12] David J *et al. BMC Genomics.* 2010 **11**: 216 [PMID: 20356352]
[13] Liska DJ. *Altern Med Rev.* 1998 **3**: 187 [PMID: 9630736]
[14] Nebert DW. *Annu Rev Biochem.* 1987 **56**: 945 [PMID: 3304150]
[15] Feyereisen R. *Biochem Soc Trans.* 2006 **34**: 1252 [PMID: 17073796]
[16] McLaughlin LA *et al. Insect Mol Biol.* 2008 **17**: 125 [PMID: 18353102]
[17] Chiu T L *et al.Proc Natl Acad Sci U S A.* 2008 **105**: 8855 [PMID: 18577597]
[18] Christmas P *et al. J Biol Chem.* 2001 **276**: 38166 [PMID: 11461919]
[19] Djouaka R F *et al. BMC Genomics.* 2008 **9**: 538 [PMID: 19014539]
[20] Tares S *et al. Biochem Biophys Res Commun.* 2000 **268**: 677 [PMID: 10679264]
[21] Wen Z *et al. Gene.* 2001 **272**: 257 [PMID: 11470532]
[22] Bogwitz M R *et al.Proc Natl Acad Sci U S A.* 2005 **102**: 12807 [PMID: 16120680]
[23] Reddy BPN *et al. Parasitol Res.* 2011 **109**: 1165 [PMID: 21494844]
[24] Ranson H *et al. Science.* 2002 **298**: 179 [PMID: 12364796]
[25] Claudianos C *et al. Insect Mol Biol.* 2006 **15**: 615 [PMID: 12364796]
[26] Thomas JH. *PLoS Genet.* 2007 **3**: e67 [PMID: 17500592]
[27] Tijet N *et al. Gene.* 2001 **262**: 189 [PMID: 11179683]

Edited by P Kangueane
Citation: **Reddy *et al.*** Bioinformation 8(9): 430-436 (2012)

BIOINFORMATION

open access

Supplementary material:

Table 1: The comprehensive list of the detoxification enzymes -cytochrome P450s (P450s), glutathione-S-transferases (GSTs), and carboxylesterases (CCEs) - of *Culex quinquefasciatus*. The table also shows the total number of genes of each of the gene family that are identified in the present analysis of the *Culex* genome sequence. The P450 enzyme supergene family is divided into four clans, viz. mitochondrial, CYP2, CYP3, and CYP4. Each of the clan is further divided into gene families, namely mitochondrial gene families (CYP12 gene family, 301, 302, 314, 315), CYP2 clan (CYP15, 303, 304, 305, 306, 307), CYP3 clan (CYP6, 9, 329), and CYP4 clan (CYP4, 325 gene families). In the similar way, GST supergene family has been classified into Delta, Epsilon, Sigma, Theta, Omega, Unclassified, and microsomal GST classes. Likewise, the CCE supergene family divided into alpha, beta, Juvenile hormone processing, glutactin, gliotactins, neuroligins, and neurotactins (pl. see ref [24-25] for the basis of classification of these supergene families).

P450s				GSTs	COEs	
CYP 6	**CYP4**	**CYP325**	**CYP 304**	**Delta GSTs**	**Mitochondrial & secreted**	**Juvenile**
CPIJ000298	CPIJ014579	CPIJ006951	CPIJ017242	AB443867	CPIJ000046	CPIJ002073
CPIJ000299	CPIJ009468	CPIJ006950	CPIJ017243	AB443867	CPIJ000047	CPIJ002075
CPIJ002535	CPIJ009473	CPIJ015959	CPIJ017244	AB443867	CPIJ000045	CPIJ002074
CPIJ002536	CPIJ009474	CPIJ017200	CPIJ017245	AB443867	CPIJ000048	CPIJ013026
CPIJ002537	CPIJ009475	CPIJ015954	CPIJ017246	CPIJ002663	CPIJ008749	CPIJ013027
CPIJ002538	CPIJ009476	CPIJ017199	*5*	CPIJ002675	CPIJ005694	CPIJ013028
CPIJ003361	CPIJ009477	CPIJ015958	**CYP 307**	CPIJ002674	CPIJ000051	CPIJ013029
CPIJ003375	CPIJ020229	CPIJ010810	CPIJ000989	CPIJ002678	CPIJ000049	CPIJ013175
CPIJ003376	CPIJ009478	CPIJ015961	*1*	CPIJ002679	CPIJ016025	CPIJ014154
CPIJ003377	CPIJ015681	CPIJ015960	**CYP 301**	CPIJ002660	CPIJ016026	CPIJ020045
CPIJ003378	CPIJ006721	CPIJ015957	CPIJ008980	CPIJ000304	CPIJ015342	CPIJ007140
CPIJ003389	CPIJ001886	CPIJ015963	*1*	CPIJ002683	CPIJ004752	CPIJ009768
CPIJ004410	CPIJ007188	CPIJ014730	**CYP 12**	CPIJ002682	CPIJ007829	CPIJ015077
CPIJ004411	CPIJ011127	*34*	CPIJ010227	CPIJ010814	CPIJ016341	*11*
CPIJ005899	CPIJ001758	**CYP9**	CPIJ010228	CPIJ002681	CPIJ007825	**Uncharacterized**
CPIJ005900	CPIJ001757	CPIJ019765	CPIJ010229	CPIJ002680	CPIJ007826	CPIJ007638
CPIJ005952	CPIJ010075	CPIJ010203	CPIJ010230	CPIJ002676	CPIJ016336	*1*
CPIJ005953	CPIJ008937	CPIJ014220	CPIJ010231	*14*	CPIJ007824	**neuroligin**
CPIJ005954	CPIJ008936	CPIJ014221	*5*	**Epsilon**	CPIJ007827	CPIJ007493
CPIJ005955	CPIJ001754	CPIJ014218	**CYP302**	CPIJ018628	CPIJ016339	CPIJ007495
CPIJ005956	CPIJ001755	CPIJ012470	CPIJ010826	CPIJ018627	CPIJ007828	CPIJ012197
CPIJ005957	CPIJ000293	CPIJ010542	*1*	CPIJ018629	CPIJ016340	*3*
CPIJ005958	CPIJ000294	CPIJ010539	**CYP 314**	CPIJ018630	CPIJ018232	**neurotactin**
CPIJ005959	CPIJ010480	CPIJ010537	CPIJ001380	CPIJ018631	CPIJ018233	CPIJ008533
CPIJ008566	CPIJ016284	CPIJ010538	*1*	CPIJ018632	CPIJ018231	*1*
CPIJ009085	CPIJ001810	CPIJ010540	**CYP 315**	CPIJ018625	CPIJ013679	**gliotactin**
CPIJ010858	CPIJ018943	CPIJ010541	CPIJ011843	CPIJ018624	*26*	CPIJ000482
CPIJ011129	CPIJ017351	CPIJ010536	*1*	CPIJ018626	**Microsomal alpha**	*1*
CPIJ015223	CPIJ018854	CPIJ010544		*9*	CPIJ013917	
CPIJ015428	CPIJ018944	CPIJ010543		**Omega**	CPIJ013918	
CPIJ016356	CPIJ018668	CPIJ010545		CPIJ000031	*2*	
CPIJ016846	CPIJ009415	CPIJ010546		*1*	**Glutactin**	
CPIJ016847	*32*	CPIJ010175		**Sigma**	CPIJ002786	
CPIJ016848	**CYP325**	CPIJ010547		CPIJ006160	CPIJ002787	
CPIJ016849	CPIJ011837	CPIJ005332		*1*	CPIJ002785	
CPIJ016850	CPIJ011838	CPIJ003082		**Theta**	CPIJ004636	
CPIJ016851	CPIJ011636	*21*		CPIJ014052	CPIJ004637	
CPIJ016852	CPIJ011841	**CYP329**		CPIJ014053	*5*	
CPIJ016853	CPIJ005683	CPIJ000655		CPIJ014051	**Beta & pheromone**	
CPIJ016854	CPIJ005684	1		CPIJ014054	CPIJ011169	
CPIJ016855	CPIJ005685	**CYP 306**		CPIJ019572	CPIJ005122	
CPIJ016856	CPIJ007086	CPIJ001038		CPIJ020053	CPIJ017763	
CPIJ016857	CPIJ007085	CPIJ001039		*6*	CPIJ016681	
CPIJ017014	CPIJ007089	*2*		**unclassified**	CPIJ016683	
CPIJ017462	CPIJ007090	**CYP 305**		CPIJ009434	CPIJ019485	
CPIJ017609	CPIJ007091	CPIJ014940		CPIJ016212	CPIJ016686	
CPIJ018494	CPIJ007093	CPIJ014941		CPIJ014694	CPIJ016685	
CPIJ019586	CPIJ007092	CPIJ014942		CPIJ018633	*8*	
CPIJ019587	CPIJ007095	*3*		*4*	**Aceltylcholines**	
CPIJ019673	CPIJ009570	**CYP 15**		**microsomal GSTs**	CPIJ000662	
CPIJ019700	CPIJ009569	CPIJ014944		CPIJ012756	CPIJ006034	
CPIJ019702	CPIJ010272	*1*		CPIJ018241	*2*	
CPIJ019703	CPIJ011835	**CYP 303**		CPIJ012754		
CPIJ019704	CPIJ000925	CPIJ009170		CPIJ012755		
CPIJ019751	CPIJ000929	*1*		CPIJ015233		
55				*5*		

Table 2a: Comparative distribution of various gene families or classes of cytochrome p450s, glutathione-S-transferases, and carboxylesterases from sequenced dipteran species. The enzyme data for the *Drosophila, Aedes*, and *Anopheles* were retrieved from Claudianos *et. al.* [25] And for *Culex* freshly analyzed data is used.

Species	*D. melanogaster*	*A. gambiae*	*A. aegypti*	*Cx. quinquefasciatus*
Cytosolic GSTs				
Delta	11	12	9	14
Epsilon	14	8	8	9
Omega	4	1	1	1
Sigma	1	1	1	1
Theta	4	2	4	6
Zeta	2	1	1	**0**
Iota	**0**	1	1	1
Others (unclassified)	**0**	3	2	3
Microsomal	**0**	4	3	4
Total GSTs	***36***	***33***	***30***	***39***
Cytochrome P450s				
CYP4	32	45	59	66
CYP3	36	42	84	77
CYP2	6	10	11	13
Mitochondrial CYPs	11	9	10	10
P450 total	***85***	***106***	***164***	***166***
CCEs				
Mitochondrial & secreted&microsomal	13	16	22	28
Hormone/semiochemical processing				
D clade/ integument esterases	3	0	0	0
E clade/ beta esterases	2	4	2	8
F clade/ juvenile hormone esterases	3	10	12	13
Neuro/developmental				
H clade/glutactin	5	10	7	5
I clade/ uncharacterized	1	1	1	1
J clade/ acetylcholinesterases	1	2	2	2
K clade/ gliotactins	1	1	1	1
L clade/ neuroligins	4	5	5	3
M clade/ neurotactins	2	2	2	1
CCEs total	***35***	***51***	***54***	***62***

Table 2b: Comparative distribution of various IR gene families or classes of cytochrome P450s, glutathione-S-transferases, and carboxylesterases from the sequenced dipteran species

	Culex quinquefasciatus	***Anopheles gambiae***	***Aedes aegypti***	***Drosophila melanogaster***
P450s	**166**	**105**	**160**	**85**
CYP3 Clan	*77*	36	82	36
CYP4 Clan	66	32	57	32
Percent contribution	*143 (86%)*	*68 (64%)*	*139 (87%)*	*68 (80%)*
GSTs	**40**	**31**	**29**	**38**
Delta	14	12	8	11
Epsilon	9	8	8	14
Percent contribution	*23 (57%)*	*20 (66%)*	*16 (55%)*	*25 (66%)*
Esterases	**62**	**51**	**49**	**35**
Alpha class	*28 (50%)*	*16 (30%)*	*22 (40%)*	*13 (40%)*