Molecular and computational characterization of resistance genes of *Culex pipiens* complex mosquitoes

A Thesis
Submitted to the Department of Entomology,
Faculty of Science, Ain Shams University
For the award of the Ph. D. Degree
(Entomology)

By

Enas Hamdy Saad Mohamad Ghallab

M. Sc.

Supervisors

Prof. Dr. Magdi G. Shehata
Professor of Medical Entomology
Entomology Department, Faculty of Science
Ain Shams University

Prof. Dr. Hoda A. Farid
Professor of Medical Entomology
Entomology Department, Faculty of Science
Ain Shams University

Prof. Dr. Ragaa E. Hammad
Professor of Medical Entomology
Entomology Department, Faculty of Science
Ain Shams University

Dr. Doaa E. Soliman
Lecturer of Entomology
Entomology Department, Faculty of Science
Ain Shams University

Prof. Dr. Harry M. Savage
Chief of Ecology and Entomology Activity, Arbovirus Disease Branch
Centers for Disease Control and Prevention
Fort Collins, CO., USA

Entomology Department
Faculty of Science
Ain Shams University
2016

Biography

Name: Enas Hamdy Saad Mohamad Ghallab

Degrees Awarded: B.Sc. (Entomology) 2005
M.Sc. (Entomology) 2010

Department: Entomology

Faculty: Science

University: Ain Shams

Occupation: Assistant lecturer in Department of Entomology, Faculty of Science, Ain Shams University.

Date of Registration for Ph.D.: September, 2011

SUPERVISORS

Prof. Dr. Magdi G. Shehata
Professor of Medical Entomology
Department of Entomology, Faculty of Science
Ain Shams University

Prof. Dr. Hoda A. Farid
Professor of Medical Entomology
Department of Entomology, Faculty of Science
Ain Shams University

Prof. Dr. Ragaa E. Hammad
Professor of Medical Entomology
Department of Entomology, Faculty of Science
Ain Shams University

Dr. Doaa E. Soliman
Lecturer of Entomology, Department of Entomology
Faculty of Science, Ain Shams University

Prof. Dr. Harry M. Savage
Chief of Ecology and Entomology Activity, Arbovirus Disease Branch Centers for Disease Control and Prevention, Fort Collins, CO., USA

Acknowledgment

Thanks first and last to Allah for the utmost help and support during this work.

I wish to express my greatest gratitude, thanks and appreciation to ***Prof. Dr. Magdi G. Shehata***, Professor of Medical Entomology, Faculty of Science, Ain Shams University, for his supervision, his kind encouragement and valuable advice.

Heartfelt thankfulness goes to ***Prof. Dr. Hoda A. Farid***, Professor of Medical Entomology, Faculty of Science, Ain Shams University, for her supervision, her consistent scientific advice and guidance. Her invaluable effort made possible the achievement of this work.

Warmest thankfulness goes also to ***Prof. Dr***. ***Ragaa E. Hammad***, Professor of Medical Entomology, Faculty of Science, Ain Shams University, for her supervision, her unfailing guidance and support and her critical reviewing and guidance through finishing this thesis.

Deepest thanks to ***Dr. Doaa E. Soliman***, Lecturer of Entomology, Ain Shams University, for her supervision, her help, encouragement and profound supervision in different research aspects.

I am very grateful to ***Dr. Harry M. Savage,*** Chief Ecology and Entomology Activity, Arbovirus Disease Branch, Centers for Disease Control and Prevention, Fort Collins, Co.,

USA, for his professional advices, profound supervision, for the support and space to conduct this research project and for identifying *Culex pipiens* collected in Egypt.

I am very grateful to ***Dr***. ***Emad I. Khater***, Assistant professor of Entomology, Ain Shams University, for his continuous scientific follow up, assisting and directing my thinking to new research dimensions.

I would like to express my sincere appreciation to ***Dr. Linda Kothera,*** Associate Service Fellow at Centers for Disease Control and Prevention, Fort Collins, Co., USA, for suggesting the research topic, training on research methods used, continuous support and for her scientific guidance and directions through the research.

I would like to thank the ***Egyptian Ministry of Higher Education (MoHE), Cultural affairs and Mission sector***, for providing the financial assistance (Joint scholarship) for this research, as well as the ***Centers for Disease Control and Prevention***, Fort Collins, Co., USA, for offering the facilities and tools needed to conduct this work.

Many thanks goes to the American Mosquito Control Districts of: New Orleans (Louisiana), Dallas (Texas) and Florence (Mississippi) in the USA for providing us with *Culex quinquefasciatus* egg rafts. Also thanks goes to the Research Institute of Medial Entomology in Cairo, Egypt, for providing us with *Culex pipiens* larvae.

A special thanks to Centers for Disease Control and Prevention, Fort Collins, Co., USA, sequencing team for samples processing.

Thanks are also to ***Prof. Dr.*** **Adel Kamal** the head of Entomology Department, Faculty of Science, Ain Shams University.

My thanks are also passed to the staff members and colleagues of Entomology Department for the various help they offered me.

Dedication

I dedicate this thesis to my family for the support they have shown me and my sincere thanks are also due to who suffered a lot during the preparation of this work.

I truly appreciate my mother ***Somayah***, father ***Hamdy Ghallab***, sisters ***Eman*** and ***Aya*** for all their love and encouragement they have given me throughout my life.

Finally, I would like to express my sincere gratitude, love and appreciation to ***Mr. Jerry Arbogast*** and ***Mrs. Franci Calderone*** for their continuous help, encourage, support, for providing me with a lot of positive energy and self-confidence and finally for taking care of me while I was abroad.

Abstract

Enas Hamdy Saad Mohamad Ghallab. Molecular and computational characterization of resistance genes of *Culex pipiens* complex mosquitoes, Faculty of Science, Ain Shams University, 2016.

Mosquitoes in the *Culex pipiens* complex, the primary vectors of diseases, have been exposed to repeated applications of insecticides, particularly pyrethroids, which resulted in the development of resistance. Detoxification enzymes play a major role in the development of insecticide resistance. Two glutathione S-transferase (GST) genes responsible for insecticide resistance: CPIJ002663 "CpGSTD5/CqGSTD5" and CPIJ002681 "CqGSTD11", were presently used in a comprehensive molecular analysis for differentiating resistant and susceptible individuals of *Cx. pipiens* and *Cx. quinquefasciatus* collected from Egypt and the United States of America, respectively. Gene amplification, sequencing and cloning, were analyzed via bioinformatics databases. Data indicated 79.8% and 80.9% similarity of amplified CpGSTD5/CqGSTD5 respectively, to CPIJ002663 gene, whereas CqGSTD11 gene yielded 98% similarity to CPIJ002681 gene. The identity of translated amino acid sequences of these two genes was 99% for CpGSTD5 and CqGSTD5, and 96% for CqGSTD11. Although alignment of obtained amino acids sequences in NCBI conserved domains perceived polymorphic loci, the overall results revealed that no specific molecular marker for differentiating susceptible and resistant individuals was identified.

Keyword: *Cx. pipiens* complex, Detoxification enzyme, GSTs, Insecticide resistance.

Contents

Title	Pages
I. Introduction……	1
Aim of the work……	6
II. Literature Review……..	7
1. Insecticide resistance problem in mosquito control……..	7
2. Insecticide resistance mechanisms…….……..	10
3. Detoxification enzymes……	11
4. Surveillance and management of insecticide resistance……	13
4.1. Tracking mosquito baseline susceptibility to pyrethroids	13
4.2. Molecular mechanisms of insecticide detoxification in mosquitoes……..	18
4.3. Biological implications of insecticide resistance……	27
III. Materials and Methods……..	30
1. Origin of tested mosquitoes……..	30
1.1. *Culex quinquefasciatus*…….	30
1.2. *Culex pipiens*……..	33
2. Insecticide Bioassay……..	34
3. Molecular studies……..	37
3.1. DNA extraction……..	37
3.1.1. Robotic DNA extraction……..	37
3.1.2. Manual DNA extraction……..	43
3.2. Molecular identification and confirmation of mosquito species using the rDNA-ITS region……..	44
3.3. Detection of insecticide resistance genes……	46
3.3.1. Amplification of Glutathione S-transferases (GST) genes..	46
3.3.1.1. Primer design……	46
3.3.1.2. PCR Purification……...	50
3.3.1.3. Direct Sequencing…….	51
3.3.1.4. Cloning……..	51
3.4. Data Analysis……..	57
IV. Results……..	59
1. Collection of *Culex* mosquitoes……	59
2. Determination of insecticide resistance…….....	60
3. Molecular studies…….	65
3.1. Molecular identification and confirmation of mosquito species using the rDNA-ITS region……	65

3.2. Detection of insecticide resistance genes……..	67
3.2.1.Primer design via database homology search for potential GST genes…….....	67
3.2.2. Amplification, sequencing and cloning of selected GST genes……	70
3.2.3. Bioinformatics analysis of gene sequences…….....	75
3.2.3.1. Sequence identity of the sequenced genes, CpGSTD5, CqGSTD5 and CqGSTD11…….	76
3.2.3.2. Protein orthologues of the sequenced genes, CpGSTD5, CqGSTD5 and CqGSTD11…….	77
3.2.3.3. Polymorphism analysis…….	82
3.2.3.4. Phylogenetic analysis……...	85
V. Discussion and Conclusions…….	89
1. Insecticide susceptibility/resistance status…….	89
2. Molecular characterization of CpGSTD5, CqGSTD5 and CqGSTD11 in members of the *Cx. pipiens* complex…….	94
2.1. Amplification, sequencing and cloning of selected GST genes……..	97
2.2. Identity of sequenced genes……	100
2.3. Phylogenetic analysis……..	101
VI. Summary……..	104
VII. References……..	109
Arabic Summary	

List of Tables and Figures

I- Tables

Table	Title	Pages
Table 1	Primers used to amplify rDNA-ITS region (Crabtree *et al*., 1995)……...……..	45
Table 2	Designed primers for six GST genes……..	48
Table 3	*Culex pipiens* adults collected from Egypt….....	59
Table4	*Culex quinquefasciatus* populations collected from the USA…...……...……...	60
Table5	Mortality of *Cx. quinquefasciatus* populations from treated and untreated areas in the USA after exposure to diagnostic doses of permethrin……	63
Table 6	Mortality of *Cx. quinquefasciatus* populations from treated and untreated areas in the USA after exposure to diagnostic doses of malathion…….	64
Table 7	Primer pairs used for the amplification of the selected genes…...……...……	71
Table 8	Polymorphic positions in amino acids of studied genes and their status…….	83

II- Figures

Figure	Legend	Pages
Fig. 1	Illustrated map showing *Cx. quinquefasciatus* egg rafts collection areas in USA. Arrows point to each of three collection sites: Texas, Louisiana and Mississippi……..……	32
Fig. 2	Adult mosquitoes in "modified ice cream carton" rearing cages……..……..	32
Fig. 3	Illustrated map showing *Cx. pipiens* larval collection areas (stars) in Egypt…….	34
Fig. 4	Wheaton bottles (250 ml) used in CDC Bottles Bioassay…….....……	35
Fig. 5	Replicates of three sets of Wheaton bottles used in the bioassay and ready for mosquito testing……..……	35
Fig. 6	(a) The BioRobot Universal System® machine. (b) Composition of the internal Worktable setup of the robotic machine, where: (1) is the tip-tray holder, (2) cooling and heating system, (3) Vacuum base, and (4) multi-well plate holder……..……..……..…….….	39
Fig. 7	BB beads. ……..……..…….....……	40
Fig. 8	96 wells' S block. ……..……..……	40
Fig. 9	Mixer mill, with the white racks fixed in position. ……..……	40
Fig. 10	96-well plates centrifuge. ……..……...……..	42
Fig. 11	Screenshot of the sample tracking system through the BioRobot Universal System® software…….....	42
Fig. 12	Diagram illustrating the composition of rDNA region, and the positions of the primers used in the amplification of this region in *Cx. quinquefasciatus* and *Cx. pipiens* (Crabtree *et al*., 1995), and differentiating it from *Cx. restuans*, where "18S, ITS1, 5.8S, ITS2 and 28S" are the rDNA region parts and "PQ10, R6 and CP16" are the primers used in the amplification process. ……..……..……..……	44
Fig. 13	Illustration of the composition of the pCR™4-TOPO® TA vector, (TOPO® TA Cloning® Kit for Sequencing: User guide. ……..……..……..……..……..……..……..……...	52
Fig. 14	Mortality of adult *Cx. quinquefasciatus* in bottles treated with permethrin (400 µg/bottle)…….... ……..…….	63
Fig. 15	Mortality of adult *Cx. quinquefasciatus* in bottles treated with malathion (43 µg/bottle)… ……..……...……..……..	64

Fig. 16 Amplification of the rDNA-ITS region (≈700bp) from *Cx. pipiens* field-collected in Egypt. QL: Qaliubiya, Ca: Cairo, G: Giza, Bs: Beni-Suef and Fa: Faiyoum populations. (+ve) is the positive control ITS2 of *Cx. restuans* (≈500 bp) and (-ve) is the blank control.…….……... 66

Fig. 17 Amplification of the rDNA-ITS region from *Cx. quinquefasciatus* collected in the USA (≈700bp). (+ve) is the positive control ITS2 of *Cx. restuans* (≈500 bp) and (-ve) is the blank control.…….……..…….. 66

Fig. 18 Alignment of the amino acids of the six annotated selected genes by CLUSTALW showing the extent of similarity among these genes. ……..……...……..……..… 68

Fig. 19 Diagram of the positions of the two protein conserved domains of the GST family in the six selected genes (CPIJ018629, CPIJ002678, CPIJ010814, CPIJ002681, CPIJ002680 and CPIJ002663) using Pfam database. ……...……..……..……..……..… 70

Fig. 20 Diagram showing the structural compositions of CPIJ002663 and CPIJ002681 genes. ……..…… 72

Fig. 21 Amplification of CPIJ002663 gene (≈ 556bp) from populations of *Cx. pipiens* and *Cx. quinquefasciatus* collected from Egypt and the USA. QL's: Qaliubiya population, Ca's: Cairo population, G's: Giza population, Bs's: Beni suef population and Fa's: Faiyum population. (-ve) is the blank control. FLO's: Florence treated families, NOT's: New Orleans treated families, TXU's: Texas untreated families, TXT's: Texas treated families, NOU's: New Orleans untreated families and MSU's: Mississippi untreated families, all from USA……..……... 74

Fig. 22 Amplification of CPIJ002681 gene (≈ 828bp) from different populations of *Cx. quinquefasciatus* collected from the USA. FLO's: Florence treated families, NOT's: New Orleans treated families, TXU's: Texas untreated families, TXT's: Texas treated families, NOU's: New Orleans untreated families and MSU's: Mississippi untreated families. ……. 74

Fig. 23 Alignment of the translated amino acids of CpGSTD5 and CqGSTD5 with the predicted amino acid sequences of CPIJ002663 and their orthologues in *Ae. aegypti* "AAGSTD5" and *An. gambiae* "AgGSTD2" by CLUSTALW.. ……..……..……..……..……..……. 78

Fig. 24 Schematic diagram for the structure of CqGTD5 and CpGSTD5 proteins and its orthologues in *Ae. aegypti* (AaGSTD5) and *An. gambiae* (AgGSTD2), based on the analysis of NCBI conserved domain database compared to the annotated CPIJ002663 protein structure..... 79

Fig. 25 Alignment of the translated amino acids of CqGSTD11 with the predicted amino acid sequences of CPIJ002681 and their orthologues in *Ae. aegypti "*AAGSTD11*"* and *An. gambiae "*AgGSTD11*"* by CLUSTALW……..…….. 80

Fig. 26 Schematic diagram for the structure of CqGSTD11 proteins and its orthologues in *Ae. aegypti* (AaGSTD11) and *An. gambiae* (AgGSTD11), based on the analysis of NCBI conserved domain database, compared to the annotated CPIJ002681 protein structure. ……...……..…… 81

Fig. 27 NJ tree based on aligned 528 nucleotides of CpGSTD5 and CqGSTD5 generated from 90 representative samples of *Cx. pipiens* from Egypt (EM1-3) and six *Cx. quinquefasciatus* populations from the USA. The accession name of *Cx. quinquefasciatus* sequences retrieved from VectorBase is: CPIJ002663. Bootstrap values obtained by 1000 replicates are indicated on the tree branches…… 86

Fig. 28 NJ tree based on aligned 694 nucleotides of CqGSTD11 generated from 80 representative samples of *Cx. quinquefasciatus* populations. The accession name of *Cx. quinquefasciatus* sequences retrieved from VectorBase is: CPIJ002681. Bootstrap values obtained by 1000 replicates are indicated on the tree branches……..…….. 87

Fig. 29 NJ tree based on aligned 176 aa of CpGSTD5, CqGSTD5 and CqGSTD11 generated from 90 representative samples of *Cx. pipiens* from Egypt and of *Cx. quinquefasciatus* from the USA, with their orthologous in *Cx. quinquefasciatus*, *Ae. aegypti* and *An. gambiae* The accession name of *Cx. quinquefasciatus* sequences retrieved from VectorBase are: CPIJ002663 and CPIJ002681. *Ae. aegypti* and *An. gambiae* accession no. are (AAEL001071, AGAP004173) and (AAEL010582, AGAP004378), were used as outgroup taxon. Bootstrap values obtained by 1000 replicates are indicated on the tree branches. ……... 88

List of Abbreviations

18S	18S ribosomal RNA
28S	28S ribosomal RNA
5.8S	5.8S ribosomal RNA
aa	amino acid
BA-1 grinding buffer	bovine serum albumin grinding buffer
BB	Beaded Beads
Bti	*Bacillus thuringiensis israelensis*
CCEs	Carboxyl/cholinesterases
CDC	Centers of Disease Control and prevention
CDs	Conserved domains
Cyt P450/ P450	Cytochrome P450
DDE	dichlorodiphenyldichloroethylene
DDT	dichlorodiphenyltrichloroethane
E. coli	*Escherichia coli*
EM	Egyptian mosquitoes
FISH	Fluorescent *in situ* hybridization
GSH	reduced Glutathione
G-site	GSH binding site
GSTs	Glutathione S-transferases
H-site	hydrophobic substrate binding
ITN	insecticide-treated nets
ITS	Internal transcribed spacer
IUPAC	International Union of Pure and Applied Chemistry
kdr	knockdown resistance
LB broth	*Luria Bertani* growth medium
NADPH reductase	nicotinamide adenine dinucleotide phosphate reductase
NJ trees	neighbor-joining tree
OPs	Organophosphates
ORF	open reading frame
PBS	Phosphate-buffered saline
PSI blast	Position specific Iterated blast
rDNA	Ribosomal DNA
rpm	Rotation per minute
RT-PCR	Reverse Transcription Polymerase Chain Reaction
S.O.C.	Super Optimal Catabolite Medium
SDS-PAGE	Sodium Dodecyl Sulfate Polyacrylamide Gel Electrophoresis
S-Lab	laboratory susceptible strains
SNP	Single nucleotide polymorphism
tBLASTn	Translated, Basic Local Alignment Search Tool is an algorithm, nucleotides
x g	times gravity