

Screening of sibling species of *Anopheles culicifacies* Giles
in Sri Lanka, using DNA based methods and determining
population structure of sibling species E

by

Iresha Nilmini Harischandra



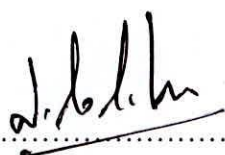
Thesis submitted to the University of Sri Jayewardenepura

for the award of the

Degree of Doctor of Philosophy in Zoology on 2015

CERTIFICATION OF SUPERVISORS

We certify that the candidate has incorporated all corrections, additions, and amendments recommended by the examiners.



24 / 03 / 2016

Prof. B. G. D. N. K. De Silva

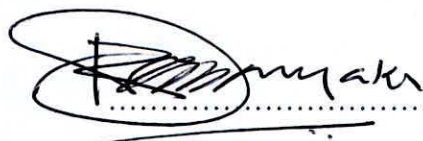
Date

Department of Zoology

Faculty of Applied Sciences

University of Sri Jayewardenepura

Gangodawila



24 / 03 / 2016

Prof. R.S. Dassanayake

Date

Department of Chemistry

Faculty of Science

University of Colombo

Colombo 07

DECLARATION

The work described in this thesis was carried out by me under the supervision of Prof. B. G. D. N. K. De Silva, Department of Zoology, Faculty of Applied Sciences, University of Sri Jayewardenepura and Prof. R.S. Dassanayake, Department of Chemistry, Faculty of Science, University of Colombo and a report on this has not been submitted in whole or in part to any university or any other institution for another Degree / Diploma

.....


Iresha Nilmini Harischandra

.....
24-03-2016

Date

We certify that the above declaration made by the candidate is true and the thesis is suitable for submission to the University for the purpose of evaluation.

.....


Prof. B. G. D. N. K. De Silva

.....


Prof. R.S. Dassanayake

TABLE OF CONTENTS

	Page
Table of contents.....	i
List of tables	x
List of figures	xiii
List of plates	xviii
Abbreviations	xix
Acknowledgements.....	xxiv
Abstract	xxvi
Chapter 1 – Introduction.....	1
1.1: Malaria	1
1.2: Mosquitoes as vectors of malaria.....	4
1.3: <i>Anopheles culicifacies</i> Giles, 1901.	7
1.4: Rationale of the current study.....	8
1.5: Specific aims of the study	10
Chapter 2 – Literature review.....	11
Part I	
2.1: Global malaria perspective	11
2.2: Malaria perspective in Sri Lanka	12
2.3: Malaria vector mosquitoes - <i>Genus Anopheles</i>	14
2.4: Burden of malaria – prevention and control.....	16

2.41: Vector control methods	17
2.5: Sibling species complexes of malaria vectors	19
2.6: Major malaria vector species in Sri Lanka - <i>Anopheles culicifacies sensu lato</i> Giles, 1901	21
2.6.1: Global distribution of <i>An. culicifacies</i>	22
2.6.2: Occurrence of <i>An. culicifacies</i> in Sri Lanka	23
2.7: Significance of precise identification of the <i>An. culicifacies</i> species complex	24
2.8: Methods to differentiate the sibling species in the complex	26
2.8.1: Cytotaxonomic techniques.....	26
2.8.1.I: Mitotic and meiotic karyotyping	27
2.8.1.II: Polytene chromosomes.....	29
2.8.2: Isozyme variations.....	29
2.8.3: Observation of cuticular hydrocarbon profiles	30
2.8.4: DNA based methods.....	30
2.8.4.I: Hybridization assay using DNA probes	30
2.8.4.II: Random Amplified Polymorphic DNA (RAPD).....	31
2.9: Ribosomal and mitochondrial genetic markers use for differentiation of sibling species	31
2.9.1: Ribosomal DNA (rDNA).....	31
2.9.1.I: Intergenic spacer region (IGS).....	36
2.9.1.II: Internal transcribed spacer regions	38
2.9.2: Mitochondrial DNA	40
2.9.2.I: Mitochondrial protein coding genes.....	43

2.10: Current status of the sibling species identification of <i>An. culicifacies</i> species complex	46
--	----

Part II

2.11: Population genetic structure	50
2.12: Significance of understanding the population genetic structure of an organism.....	51
2.13: Population genetic models and processes	52
2.13.1: Genetic drift	53
2.13.2: Effective population size	54
2.13.3: Gene flow.....	54
2.13.4: Population expansions and constrictions – historic population demographics	56
2.14: Molecular techniques use in population genetic analysis	57
2.14.1: Microsatellites	58
2.14.2: Microsatellite mutation models.....	61
2.14.3: Microsatellite isolation	63
2.15: Population genetic studies for <i>An. culicifacies</i> using microsatellite markers	64
2.16: Methods of population genetic analysis	65
2.16.1: Statistical programs for population genetic analysis.....	66

Chapter 3 – Materials and methods..... 69

Part I

3.1: Mosquito sample collection.....	69
--------------------------------------	----

3.1.1: Identification of adult mosquitoes.....	71
3.2: Laboratory rearing of mosquitoes.....	71
3.2.1: Rearing for first filial (F ₁) generation.....	73
3.3: Preservation of adult mosquitoes.....	74
3.4: Cytogenetic identification of mosquitoes.....	74
3.5: DNA extraction of mosquitoes.....	75
3.6.: Identification of sibling species B and E based on mitochondrial DNA (mtDNA).....	76
3.6.1: BCE multiplex allele specific PCR method based on cytochrome oxidase II (<i>COII</i>) region of mitochondrial DNA.....	76
3.6.2: Amplification and sequencing of the Cytochrome oxydase II (<i>COII</i>) region of mitochondrial DNA.....	77
3.6.2.I: Analysis of nucleotide sequences of <i>COII</i> region.....	77
3.6.2.II: Analysis of amino acid sequences of <i>COII</i> region.....	78
3.6.2.III: Analysis of haplotype diversity of <i>An. culicifacies</i> sibling species based on the <i>COII</i> sequences.....	78
3.6.2.IV: Evolutionary relationship of <i>COII</i> nucleotide variants of <i>An. culicifacies</i>	79
3.6.3: Analysis of variations in the cytochrome oxidase 1 (<i>COI</i>) region of mitochondrial DNA as a molecular marker to distinguish sibling species B and E.....	81
3.6.4: Analysis of the Cytochrome b (<i>Cyt-b</i>) gene of mitochondrial DNA of sibling species B and E.....	81

3.7: Identification of <i>An. culicifacies</i> sibling species B and E based on ribosomal DNA (rDNA) regions.....	82
3.7.1: Analysis of Internal transcribed spacer I (ITS1) region of rDNA cistron for sibling species B and E	82
3.7.2: Analysis of internal transcribed spacer 2 (ITS2)region of <i>An. culicifacies</i> sibling species found in Sri Lanka.....	85
3.7.3: Intergenic spacer region (IGS) of rDNA cistron as a molecular marker to distinguish sibling species of <i>Anopheles culicifacies</i> in Sri Lanka.....	86
3.7.3.I: Expand long template PCR method	86
3.7.3.II: Reguar PCR method.....	86
3.7.4: D2 domain of 28S rDNA as a molecular marker to distinguish two sibling species found in Sri Lanka.....	87
3.7.5: D3 domain of 28S rDNA as a molecular marker to distinguish two sibling species found in Sri Lanka.....	87
3.8: Random amplified polymorphic DNA (RAPD) as a molecular marker to distinguish sibling species B and E.....	88

Part II

3.9: Analysis of population genetics structure of <i>An. culicifacies</i> sibling species E using microsatellite markers	90
3.9.1: Selection of microsatellite loci for the study	90
3.9.2: Optimization of microsatellite PCR	91
3.9.3: Visualization of PCR products.....	92
3.9.3.I: Visualization using agarose gel electrophoresis	92

3.9.3.II: Visualization using polyacrylamide gel electrophoresis and silver staining	92
3.9.3.III: Silver staining protocol	93
3.9.4: preparation of fluorescence labeled primers	93
3.9.5: PCR multiplexing	94
3.9.6: Genotyping of post PCR products.....	94
3.10: Data analysis	96
3.10.1: Preparation of genotypic data for population genetic software packages.....	97
3.10.2: Estimate the error rate of genotype allele scoring	98
3.10.3: Checking for null alleles and scoring errors	98
3.10.4: Observation of recent population bottleneck/ population expansion	99
3.10.5: Genetic diversity	100
3.10.6: Hardy Weinberg Equilibrium (HWE)	100
3.10.7: Pairwise linkage disequilibrium	101
3.10.8: Genetic differentiation among populations using F - statistics.....	101
3.10.9: Analysis of molecular variation (AMOVA)	101
3.10.10: Isolation by distance (IBD) analysis.....	101
3.10.11: Gene flow	102
3.10.12: Spatial analysis of molecular variance (SAMOVA) and detection of possible barriers to the gene flow among populations	102
3.10.13: Effective population size	103
3.10.14: Model based cluster analysis.....	104
3.10.15: Spatial genetic cluster analysis.....	104

3.11: Effect of monomorphic loci on genetic diversity and population genetic structure	106
Chapter 4 - Results.....	107
Part I	
4.1: Screening of <i>An. culicifacies</i> wild population using BCE allele specific PCR to distinguish the sibling species B and E	107
4.2: Analysis of <i>An. culicifacies</i> sibling species B and E using mitochondrial DNA	110
4.2.1: DNA Sequence polymorphism of Cytochrome Oxidase II region (<i>CO II</i>).....	110
4.2.2: Analysis of haplotype diversity in <i>COII</i> region of <i>An. culicifacies</i> in Sri Lanka.....	124
4.2.2.I: Minimum spanning algorithm of sibling species E haplotypes	126
4.2.3: Analysis of amino acid coding sequence of <i>COII</i> gene region.....	128
4.2.4: Evolutionary relationship of <i>An. culicifacies</i> in Sri Lanka using <i>COII</i> region	133
4.2.5: Characterization of sibling species B and E using the cytochrome oxidase I (<i>CO I</i>) region of mitochondrial DNA.....	137
4.2.6: Characterization of sibling species B and E using the Cytochrome b (<i>Cyt-b</i>) region of mitochondrial DNA	143
4.3: Analysis of <i>An. culicifacies</i> sibling species B and E using ribosomal DNA	148
4.3.1: Nucleotide sequence of Internal Transcribed Spacer I (ITS1) region of sibling species B and E.....	148

4.3.2: Internal Transcribed Spacer II (ITS2) region of sibling species B and E..	152
4.3.3: Nucleotide variations in the ribosomal intergenic spacer (IGS) region	156
4.3.4: Analysis of the Ribosomal D2 and D3 domain sequence variation among sibling species B and E.....	158
4.4: Random Amplified Polymorphic DNA (RAPD) analysis of sibling species B and E	164

Part II

4.5: Number of allele recorded for microsatellite loci.....	166
4.6: Error rate.....	170
4.7: Null alleles.....	170
4.8: Allele frequency of microsatellite loci.....	171
4.9: Allele richness	178
4.10: Gene diversity.....	178
4.11: Demographic stability of <i>An. culicifacies</i> sibling species E population in Sri Lanka.....	179
4.12: Intra population analysis	181
4.12.1: Hardy-Weinberg equilibrium.....	181
4.12.2: Pairwise linkage disequilibrium	189
4.13: Genetic differentiation.....	191
4.13.1: Inbreeding coefficient (F_{IS}).....	191
4.13.2: Fixation index in sub populations – Co-ancestry coefficient (F_{ST}).....	191
4.13.2.I: Fixation index in populations (F_{ST}) analyzed without Polymorphic markers	192
4.13.3: Gene Flow	194

4.13.4: Effective population size	196
4.13.5: Analysis of molecular variance (AMOVA)	197
4.13.6: Isolation by Distance (IBD)	200
4.14: Spatial analysis of molecular variance (SAMOVA) and barriers to the gene flow	204
4.15: Model based cluster analysis	206
4.15.1: Bayesian cluster analysis	206
4.15.1.I: Probability of assignment of individuals to each population cluster	215
4.15.2: Spatial genetic cluster analysis	217
Chapter 5 - Discussion	223
6.0 Conclusions.....	243
7.0 References.....	244
8.0 Appendices	316
Appendix I - Publications	316
Appendix II – Chemicals and Buffers.....	318
Appendix III – Instrumentation	323

LIST OF TABLES

Table 2.1: Malaria vectors prevalent in Sri Lanka and neighboring countries.....	15, 16
Table 2.2: Anopheline species complexes	20, 21
Table 2.3: Summary of DNA based <i>An. culicifacies</i> sibling species identification methods	49
Table 2.4: Currently using population genetic analysis software.....	66, 67, 68
Table 3.1: Accession numbers of <i>COII</i> region sequences used for the phylogenetic reconstruction	79, 80
Table 3.2: Designed primer pairs and their sequences to amplify the ITS1 region	84
Table 3.3: RAPD primers and their sequences used in screening of <i>An. culicifacies</i> sibling species E	88, 89
Table 3.4: Forward and reverse primer sequences of selected 13 microsatellite loci	90, 91
Table 3.5: Fluorescence labels of forward primers	94
Table 3.6: Number of samples genotyped from each sampling site.....	95
Table 3.7: Pairwise matrix of geographical distance in km	106
Table 4.1: Number of wild and cytogenetically identified mosquitoes screened using BCE- PCR and their sibling species status.....	110
Table 4.2: Polymorphic nucleotide positions in the mitochondrial <i>COII</i> region sequences of sibling species E in India and Sri Lanka ...	116,117
Table 4.3: Polymorphic nucleotide positions in the mitochondrial <i>COII</i> region sequences of sibling species B India and Sri Lanka	123

Table 4.4: Mitochondrial <i>COII</i> haplotypes of <i>An. culicifacies</i> sibling species E	124
Table 4.5: Inter connection lengths for the minimum spanning tree between sequences	125
Table 4.6: Amino acid substitutions of sibling species B and E sequences of India and Sri Lanka	131,132
Table 4.7: NCBI GenBank accession numbers for <i>COI</i> , <i>COII</i> and <i>Cyt-b</i> sequences of sibling b species B and E	148
Table 4.8: NCBI BLAST search results for the IGS sequences... ..	158
Table 4.9: Total number of alleles amplified for each locus	166
Table 4.10: Number of alleles observed from each locus for six populations	167
Table 4.11: Number of private alleles in each population with respective loci	168
Table 4.12: Maximum frequency of alleles observed and the frequency in each population.....	169
Table 4.13: Loci showing errors while scoring in samples and replicates.....	170
Table 4.14: Null allele frequencies of eight loci for four analysis methods	171
Table 4.15: Allele richness values per locus and per population	178
Table 4.16: Gene diversity per locus in all populations.....	179
Table 4.17: Estimated of <i>p</i> -value for the heterozygosity tests of the <i>An. culicifacies</i> population	180,181
Table 4.18: HWE <i>p</i> - values for the all populations in each locus	183
Table 4.19: Summary statistics of genetic diversity at microsatellite loci among six populations of sibling species E	184,185,186,187, 188
Table 4.20: Significant linkage disequilibrium in pairs of loci	190

Table 4.21: F_{IS} values of the individuals at each population and in pooled population.....	191
Table 4.22: F_{ST} values and distance between (km) populations pairs	192
Table 4.23: F_{ST} values (using polymorphic loci) and distance between (km) populations pairs.....	193
Table 4.24: Significant comparison of F_{ST} values in two analyses ($p < 0.05$)	193, 194
Table 4.25: F_{ST} and Nm values for the population comparisons of <i>An. culicifacies</i> sibling species E populations	195
Table 4.26: F_{ST} and Nm values for the population comparisons using polymorphic loci ($*p < 0.0033$)	195
Table 4.27: Long-term N_e estimates based on genetic diversity (expected heterozygosity) in population, assuming SMM	197
Table 4.28: Global AMOVA results as a weighted average over loci	198
Table 4.29: AMOVA results within sampling populations	199
Table 4.30: Statistical values of Mantel test	200
Table 4.31: Results of SAMOVA.....	204, 205
Table 4.32: Table output of the Evanno method results	215
Table 4.33: The probability of samples to each cluster in sampling locations	216
Table 4.34: Pairwise F_{ST} values and the respective Nm values in three gene clusters.....	216

LIST OF FIGURES

Figure 1.1: The world map showing malaria affected regions.....	3
Figure 1.2: Global distribution of dominant or potentially important malaria vectors.....	6
Figure 2.1: Map representing areas of the world, endemic with malaria.....	12
Figure 2.2: Timeline of incidence and major events in Sri Lanka's malaria programme from 1911-2011	14
Figure 2.3: Distribution of the <i>An. culicifacies</i> complex (grey) and each of the five species A to E	23
Figure 2.4: Breeding places of <i>An. culicifacies</i>	24
Figure 2.5: Mitotic and meiotic karyotypes of male and female mosquitoes.....	28
Figure 2.6: General organization of rRNA gene of eukaryotes	32
Figure 2.7: Schematic representation of a transcription unit of eukaryotic rDNA with tandem repeats of 18S, ITS1, 5.8S, ITS2, and 28S.....	34
Figure 2.8: Gene arrangement of anopheline mitochondrial genome	41
Figure 2.9: Model of microsatellite mutation by replication slippage.....	61
Figure 3.1: A map of Sri Lanka showing the sample collecting sites	70
Figure 3.2: sequence of ITS1, flanking regions of 18s and 5.8s regions of <i>An. gambiae</i>	83
Figure 3.3: Allele 112bp of locus AcAVIII B40.....	96
Figure 3.4: Allele 200bp and 204bp of locus AcA59	96
Figure 3.5: Allele 200bp of locus AcA59	97

Figure 4.1a: Agarose gel electrophoresis image showing amplified BCE-PCR fragments.....	108
Figure 4.1b: Agarose gel electrophoresis image showing D2 PCR fragments (samples could not identify using BCE-PCR) and amplified BCE-PCR fragments.....	109
Figure 4.2: ClustalW multiple sequence alignment of <i>COII</i> sequences for sibling species E from current study	112,113,114,115
Figure 4.3: Nucleotide substitutions at BCE-PCR primer binding sites	118
Figure 4.4: ClustalW multiple alignment of <i>COII</i> sequences for sibling species B	120,121,122
Figure 4.5: Minimum spanning tree of sibling species E haplotypes.....	126
Figure 4.6: Minimum spanning network of sibling species E haplotypes	127
Figure 4.7: ClustalW multiple alignment of <i>COII</i> amino acid coding sequences	129,130
Figure: 4.8: Phylogenetic tree (1000 bootstrap) constructed from <i>COII</i> sequences using the neighbor-joining method.....	134
Figure 4.9: Phylogenetic tree (1000 bootstrap) constructed from <i>COII</i> sequences using the maximum likelihood method	135
Figure 4.10: Phylogenetic tree (1000 bootstrap) constructed from <i>COII</i> sequences using the maximum parsimony method.....	136
Figure 4.11: ClustalW multiple sequence alignment of the <i>COI</i> region sequences of sibling species B and E.....	138,139,140

Figure 4.12: Fragment of ClustalW multiple sequence alignment showing the variations in <i>COI</i> of sibling species B and E found in Sri Lanka and India.....	141
Figure 4.13: ClustalW multiple sequence alignment of the amino acid sequences <i>COI</i> region of sibling species B and E	142
Figure 4.14: ClustalW multiple alignment of the nucleotide sequences of <i>Cyt-b</i> region of sibling species B and E.....	144,145,146
Figure 4.15: ClustalW multiple alignment of the amino acid sequences of <i>Cyt-b</i> region of sibling species B and E.....	147
Figure 4.16: ClustalW multiple alignment of the nucleotide sequences of ITS1 region of sibling species B and E.	149,150,151
Figure 4.17: ClustalW alignment of ITS2 sequences	153,154,155
Figure 4.18: Agarose gel electrophoresis for the IGS region of <i>An. culicifacies</i> , <i>An. subpictus</i> and <i>Ae. aegypti</i> mosquitoes.	157
Figure 4.19: Nucleotide sequence pattern in the D2 domain of rDNA in sibling species B and E.....	159,160,161
Figure 4.20: Nucleotide sequence pattern in the D3 domain of rDNA in sibling species B and E.....	162,163
Figure 4.21: RAPD-PCR amplified fragments of <i>An. culicifacies</i> sibling species B and E mosquitoes.	165
Figure 4.22	
4.22a: Allele frequencies of AcAVIIB40 locus in all populations	172
4.22b: Allele frequencies of AcA59 locus in all populations	172
4.22c: Allele frequencies of AcAVIB213 locus in all populations.....	173

4.22d: Allele frequencies of AcA75 locus in all populations.....	173
4.22e: Allele frequencies of AcAIIB5 locus in all populations.....	174
4.22f: Allele frequencies of AcAVb93 locus in all populations.....	174
4.22g: Allele frequencies of AcA36 locus in all populations.....	175
4.22h: Allele frequencies of AcAVB93A locus in all populations.....	175
Figure 4.23: Percentage of polymorphic loci in each sampling site.....	176
Figure 4.24: Allelic pattern in sampling sites.....	177
Figure 4.25: The correlation between the distance between population pairs (km) and the Nm values.....	196
Figure 4.26: Percentages of molecular variance in pooled population.....	198
Figure 4.27: Percentages of molecular variations in sampled populations.....	199
Figure 4.28	
Figure 4.28a: Correlation between genetic differentiation (F_{ST}) of population pairs and geographical distances of <i>An. culicifacies</i> E.....	201
Figure 4.28b: Correlation between genetic distance ($F_{ST} / (1 - F_{ST})$) of population pairs and geographical distances of <i>An. culicifacies</i> E.....	202
Figure 4.28c: Correlation between genetic similarity (M) of population pairs and geographical distances of <i>An. culicifacies</i> E.....	203
Figure 4.29: Possible barrier to gene flow detected from the SAMOVA analysis	205
Figure 4.30: Bayesian cluster analysis using STRUCTURE	207
Figure 4.31: Graphical representation of the data set for the most likely K ($K=3$)	208
Figure 4.32: The graphical representation of proportionate gene clusters in each sampling sites	209
Figure 4.33: Triangle plot representation of 3 clusters ($K=3$).....	210

Figure 4.34: Plot of mean likelihood $L(K)$ and variance per K value. Mean $L(K)$ (ΔSD) over 20 runs for each K value	211
Figure 4.35: Graphic representation of the estimated probability of data for each K value.	212
Figure 4.36: The rate of change of the likelihood distribution for K values	213
Figure 4.37: The absolute value of the second order rate of change of the mean likelihood distribution.....	214
Figure 4.38: Veronoi diagram of the sampled populations and dummy points in a single plane.....	217
Figure 4.39: Graphical representation of TESS for K_{max} ranging from $K=2-9$ without admixture model	219
Figure 4.40: Analysis of TESS output data.....	220
Figure 4.41: Graphical representation of individuals assuming Conditionally Auto Regressive (CAR) admixture model (represents Isolation by distance). $K_{max} = K-2-9$	221
Figure 4.42: Graphical representation of individuals assuming BYM admixture model. $K_{max} = K-2-9$	222

LIST OF PLATES

Plate 3.1: Oviposition cups provided for each female mosquito	72
Plate 3.2: Larval rearing tray	73

ABBREVIATIONS

μ	Microsatellite mutation rate
A	Adenine
<i>A</i>	number of alleles
Anu.....	Anuradhapura
Mon.....	Monaragala
Tha	Thanamalwila
Kan.....	Kandy
Kat.....	Kataragama
Nik	Nikaweratiya
<i>Ae</i>	<i>Aedes</i>
AFLP	Amplified Fragment Length Polymorphism
AMC	Anti Malaria Campaign
AMOVA.....	Analysis of Molecular Variance
<i>An</i>	<i>Anopheles</i>
AT	Adenine + Thymine
BC	Before Christ
BPB	Bromo phenol blue
<i>Bti</i>	<i>Bacillus thurigiensis var israelensis</i>
BYM	Tess software specific admixture model
C.....	Cytocine
CAR	Conditionally Auto Regressive admixture model
cDNA	Complementary DNA

<i>COI</i>	Cytochrome oxidase I
<i>COII</i>	Cytochrome oxidase II
<i>Cyt-b</i>	Cytochrome b
D2.....	Domain 2 of ribosomal DNA
D3.....	Domain 3 of ribosomal DNA
DDT.....	Dichlorodiphenyltrichloroethane
DIC.....	Deviance Information Criterion
DNA.....	Deoxyribonucleic acid
dNTP.....	Deoxy Nucleotide triphosphate
DVS.....	Dominant Vector Species / species complexes
EDTA.....	Ethylene-diamine-tetra-acetic acid
EMBL.....	European Molecular Biology Laboratory
EtBr.....	Ethidium bromide
ETS.....	External transcribed spacer
F ₁	First filial
F _{IS}	Inbreeding coefficient
F _{ST}	Fixation index
G.....	Guanine
GC.....	Guanine + Cytocine
G _{ST}	Analogue to the fixation index
H _e	Expected heterozygosity
He.....	Heterozygosity from allele frequency data
Heq.....	Heteroztgosity from number of alleles and sample sizes
H _o	Observed heterozygosity

HWE	Hardy-Weinberg Equilibrium
IAM.....	Infinite allele model
IBD.....	Isolation by distance
IGS.....	Intergenic Spacer
IRS	Indoor residual spraying
ITNs	Insecticide treated nets
ITS 1.....	Internal Transcribed Spacer 1
ITS 2.....	Internal Transcribed Spacer 2
<i>K</i>	gene pool clusters
<i>Ldh</i>	Lactate dehydrogenase
<i>M</i>	Genetic similarity
ML	Maximum likelihood
MP.....	Maximum parsimony
MCMC	Marcov chain monte carlo method
MDE.....	Mutation drift equilibrium
MgCl ₂	Magnesium Chloride
mRNA	Messenger RNA
mtDNA.....	Mitochondrial DNA
NCBI.....	National Center for Biotechnology Informaion
NJ.....	Neighbor joining
<i>ND</i>	<i>NADH dehydrogenase subunit</i>
<i>Ne</i>	Effective population size
NORs.....	Nucleolar Organizer Regions
<i>N_{ST}</i>	Analogue to the fixation index

NTS	Non transcribed spacer
ORF	Open Reading Frame
OTU	Operational taxonomic unit
PCD.....	Passive case detection
PCR.....	Polymerase Chain Reaction
<i>r</i>	null allele frequency
RAPD	Random Amplified Polymorphic DNA
rDNA.....	Ribosomal DNA
RFLP	Restriction fragment length polymorphism
RPM	Rounds per minute
rRNA.....	Ribosomal RNA
<i>Rs</i>	allele richness
R_{ST}	Analogue to the fixation index
SAMOVA.....	Spatial Analysis of Molecular Variance
SDS	Sodium dodecyl sulfate
SIT	Sterile insect technique
SMM	Stepwise mutation model
SNP	Single nucleotide polymorphism
SPR	Subtree pruning and redrafting
SSR	Simple sequence repeat
STR.....	Short tandem repeat
T.....	Thymine
TAE.....	Tris Acetate EDTA
TBE.....	Tris Borate EDTA

TE..... Tris EDTA (Ethylenediaminetetraacetic acid)
TEMED NNN'N' – tetramethyl-ethylenediamine
TPM Two phase mutation model
tRNA Transfer RNA
UV..... Ultra violet
VNTR..... Variable number of tandem repeat
WHO World Health Organization

ACKNOWLEDGEMENTS

I wish to extend my profound thanks to the Supervisor, Professor B.G.D.N.K. De Silva, Department of Zoology, University of Sri Jayewardenepura for providing invaluable guidance and precious encouragement throughout the research.

I extend my utmost gratitude and appreciation to my co-supervisor Professor R.S. Dassanayake, Department of Chemistry, University of Colombo for the constant guidance, advices and encouragement throughout the research.

I sincerely acknowledge the Head of the Department of Zoology during the period, Prof. B.G.D.N.K. De Silva and Prof. P.M. Manage for providing laboratory facilities and necessary guidance required to carry out the research project successfully.

I greatly acknowledge Dr. C.V. Chandrasekaran, Senior Lecturer, Department of Chemistry, University of Colombo, for providing laboratory facilities and chemical requirements during the first year of the study, prior to establish the molecular research laboratory facilities in the Department of Zoology, University of Sri Jayewardenepura.

I wish to express my thanks to all the academic and non academic staff members of the Department of Zoology, University of Jayewardenepura for their generous support provided in various ways to complete my research project.

The financial assistance by the National Research Council (Grant No: 09-21) and University of Sri Jayewardenepura (Grant No: ASP/06/RS/2009/006) are greatly acknowledged for their support provided through research grants. Also the genetic society of Department of Zoology, University of Sri Jayewardenepura is greatly appreciated for the financial assistance.

Anti Malaria Campaign (AMC) of Sri Lanka, Regional Malaria Officers and entomological assistants attached to the AMC are kindly acknowledged for providing mosquito samples for the study.

Prof. Nilmini Gunawardena, Senior Lecturer, Dr. Menaka Hapugoda, Senior Lecturer, Mr. M.Y.D. Dayanath, research officer and the research team in Molecular Medicine Unit of University of Kelaniya are greatly acknowledged for the support given in the optimizations of silver staining procedure.

My friends, Dr. Sajani Dias, Dr. Prasad Premaratne, Dr. Chamila Adhikaram are acknowledged for the support throughout the project. My fellow graduate research students in the molecular genetics laboratory, Ms. Pamoda Jayathunga, Ms. Dilakshini Dayananda and Ms. Sachini Fernando are kindly appreciated.

Most of all, I would like to thank to my loving parents, brother and sister for their support and for always believing in my ability to succeed during my PhD program.

Screening of sibling species of *Anopheles culicifacies* Giles in Sri Lanka, using DNA based methods and determining population structure of sibling species E

Iresha Nilmini Harischandra

ABSTRACT

Anopheles culicifacies, major vector of malaria in Sri Lanka is a complex of five sibling species designated as A, B, C, D and E. Sri Lanka is reported to have B and E, among them sibling species E is the major vector. Precise identification and characterization of sibling species and the understanding of the genetic structure of vector population is prerequisite in malaria vector control and prevention of re-occurrence programs. Therefore, the study was designed to investigate the potential of mitochondrial and ribosomal DNA regions to distinguish the *An. culicifacies* sibling species in Sri Lanka and to determine the population genetic structure of sibling species E in Sri Lanka. Mosquito samples were collected from six localities of Sri Lanka during 2009-2012 and the genomic DNA was extracted. BCE-PCR assay developed in India was used to test its utility in sibling species identification in Sri Lanka. Sequencing of mitochondrial and ribosomal DNA regions were carried out to investigate the genetic diversity in these regions of Sri Lankan sibling species. Then the cytogenetically identified sibling species E mosquitoes (n=193) were genotyped using microsatellite markers to analyze the population genetic structure using standard analyses, genetic differentiation, Hardy-Weinberg equilibrium, linkage disequilibrium, demographic stability, clustering analysis, AMOVA, SAMOVA, gene flow and isolation by distance. BCE-PCR profiles

of some sibling species of B and E were different from the profile reported in India limiting its utility in Sri Lanka suggesting existence of genetic variants. None of the DNA region was able to distinguish two sibling species because of the heterogeneity in mitochondrial regions and the similarity in ribosomal DNA regions in two sibling species. Further, non-synonymous substitutions were detected in *COII* gene amongst sibling species E populations in Sri Lanka (04%). In the genetic structure analysis, five microsatellite loci were highly polymorphic across populations with high allelic richness (4.963 – 9.830). Deviation from Hardy-Weinberg Equilibrium was observed in four loci in pooled population and in five subpopulations ($p < 0.006$) with heterozygosity deficits. The population was observed in mutation drift equilibrium ($p < 0.0062$) with a trend towards a recent population expansion creating a risk of spreading disease if there is an indigenous *Plasmodium* parasite population. Genetic differentiation was observed (F_{ST} 0.03331 - 0.23184) in some population pairs with reduced gene flow which were not supported by the isolation by distance model ($r^2 = 0.3057$, $p < 0.0180$). Bayesian clustering analysis identified the presence of three sympatric clusters in the studied population. Further, the central hill region of the country was found to be act as a possible barrier to the gene flow among population.