POPULATION GENETIC STRUCTURE AND BREEDING PATTERN OF THE YELLOW FEVER MOSQUITO *Aedes aegypti* SUBSPECIES WITH EMPHASIS ON THEIR DISEASE TRANSMISSION IN SUDAN

SARA ABDELRAHMAN ABUELMAALI ELAMIN

UNIVERSITI SAINS MALAYSIA

2022

POPULATION GENETIC STRUCTURE AND BREEDING PATTERN OF THE YELLOW FEVER MOSQUITO *Aedes Aegypti* SUBSPECIES WITH EMPHASIS ON THEIR DISEASE TRANSMISSION IN SUDAN

by

SARA ABDELRAHMAN ABUELMAALI ELAMIN

Thesis submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

October 2022

DECLARATION

I declare that the content which is presented in this thesis is my work which was done at Universiti Sains Malaysia unless informed otherwise. The thesis has not been previously submitted for any other degree.

____

Name of Student:
SARA ABDELRAHMAN ABUELMAALI ELAMIN
Matric number: P-BD0009/16(R)
Passport number: P04642026

____

Name of Supervisor:
Assoc. Prof. Dr. Nur Faeza Abu Kassim

ACKNOWLEDGEMENTS

I am eternally grateful to the Almighty God for his protection and guidance throughout my life and this research journey. Without him there is nothing I could ever possibly do. I am grateful to my supervisor, Assoc. Prof. Dr. Nur Faeza Abu Kassim for her counsel and encouragement. She stimulated deliberations that significantly shaped this work and was very understanding, supportive and patient throughout this journey of my study. I am also grateful to Prof. Dr. Zairi Jaal, Dr. Intan Hazelina and Assoc. Prof. Dr. Mustafa Fadzil Farid Wajidi who offer advice and guidance towards improving this work, as well as thought-provoking comments and suggestions greatly enhanced this study.

I would like to thank the WHO Collaborating Centre for Arbovirus Reference and Research at the Tropical Infectious Diseases Research and Education Centre (TIDREC), University Malaya, Malaysia for their hosting and facilitating the virus detection work, especially the director of TIDREC Prof. Dr. Sazaly Abu Bakar, Dr. Boon Teong Teoh and Dr. Sam Sing Sin for their guidance and support during my work in TIDREC. I am equally grateful to all the staff for their diverse assistance since this work would not have been possible without their help. I also acknowledge the comradeship of my family, friends, lab mates and anyone who has been on this journey with me.

Finally, I would like to acknowledge the supports of The Organization for Women in Science for Developing World (OWSD) for supporting my study with a Postgraduate Fellowship and fund reservation No.3240288516.

TABLE OF CONTENTS

DECLARATION .... **ii**

ACKNOWLEDGEMENTS .... **iii**

TABLE OF CONTENTS .... **iv**

LIST OF TABLES .... **ix**

LIST OF FIGURES .... **xi**

LIST OF ABBREVIATIONS .... **xiv**

LIST OF SYMBOLS .... **xvi**

LIST OF APPENDICES .... **xvii**

ABSTRAK .... **xviii**

ABSTRACT .... **xxi**

CHAPTER 1 INTRODUCTION **1**

1.1 General introduction 1

1.2 Justification of the study 5

1.3 Aim and objectives of the study 8

CHAPTER 2 LITERATURE REVIEW **9**

2.1 Arthropods transmitted diseases/arboviruses 9

2.1.1 Yellow fever 9

2.1.2 Yellow fever in Sudan 10

2.1.3 Dengue fever 10

2.1.4 Dengue fever in Sudan 11

2.1.5 Chikungunya virus 11

2.1.6 Chikungunya in Sudan 12

2.2 *Aedes aegypti* biology, ecology, classification, and distribution 12

2.2.1 *Aedes aegypti* biology and morphology 13

2.2.2 *Aedes aegypti* distribution and ecology 13

2.2.3 *Aedes aegypti* subspecies morphology 14

2.2.4 *Aedes* subspecies distribution and ecology 15

2.3 *Aedes aegypti* subspecies classification/debates 16

2.3.1 Morphological classification history /differentiation 16

2.3.2 Molecular classification of the two subspecies 17

2.3.3 *Aedes aegypti* subspecies vector competence 19

2.4 Genetic variations studies in *Aedes aegypti* subspecies 20

2.4.1 Mitochondrial markers.. 22

2.4.2 Cytochrome oxidase subunit 1 (*CO1*) 22

2.4.3 NADH dehydrogenase 4 (*ND4*) 23

2.4.4 Microsatellite markers in mosquitoes' population genetics............... 24

2.5 Virus isolation and Cell culture of *Aedes* mosquito..................................... 26

2.6 Virus detection techniques in mosquito vectors 27

2.7 Implications of population genetics studies on *Aedes aegypti* control 30

CHAPTER 3 THE DISTRIBUTION AND GENETIC VARIATION AMONG *Aedes aegypti* SUBSPECIES USING MITOCHONDRIAL DNA (*CO1* AND *ND4*) IN DIFFERENT AREAS OF SUDAN... 32

3.1 Introduction.. 32

3.2 Materials and Methods 34

3.2.1 Study sites 34

3.2.2 Mosquito collection and breeding sites survey 35

3.2.3 Morphological identification.. 35

3.2.4 DNA extraction 37

3.2.5 Molecular identification.. 38

3.2.6 PCR for Mitochondrial markers amplification................................ 39

3.2.7 Mitochondrial genes Sequences Analyses and molecular diversity indices.. 40

3.2.8 Phylogenetic Analyses 40

3.2.9 *Aedes aegypti* sub-species genetic variations and population genetic analyses 41

3.2.9(a) Genetic variations and isolation by distance test: 41

3.2.9(b) Population genetic analyses 41

3.3 Results 42

3.3.1 *Aedes aegypti* breeding sites, morphological and molecular identification at eight sites of Sudan 42

3.3.2 Distribution of the *Aedes aegypti* subspecies 43

3.3.3 Mitochondrial genes haplotype analysis 45

3.3.4 Molecular diversity indices in the *CO1* and *ND4* mitochondrial genes of the *Aedes aegypti* subspecies 46

3.3.5 Phylogenetic relationships in *Aedes aegypti* subspecies 47

3.3.6 Genetic Variance (Pairwise F_{ST}) and Mantel Test Results 51

3.3.7 Population genetic analyses: Different Study Sites Inferences on population groups and genetic barriers *Aedes aegypti* 56

3.3.8 Test of Neutrality and Natural Selection in Different Populations of *Ae. aegypti* 58

3.4 Discussion 61

3.4.1 Observations on breeding sites during the entomological survey of *Ae. aegypti subspecies* 61

3.4.2 Distribution of *Ae. aegypti* subspecies/forms 62

3.4.3 Genetic variation and population genetic of *Ae. aegypti* subspecies 64

CHAPTER 4 POPULATION STRUCTURE AND PHYLOGENETIC RELATIONSHIP OF *Aedes aegypti* SUBSPECIES POPULATIONS IN SELECTED GEOGRAPHICAL LOCATIONS IN SUDAN 68

4.1 Introduction 68

4.2 Materials and Methods 70

4.2.1 Mosquito Samples collection and classification 70

4.2.2 Genomic DNA extraction and molecular identification 71

4.2.3 Microsatellite DNA molecular techniques 71

4.2.4 Data analysis/ Microsatellite DNA 75

4.2.4(a) Standard genetic procedures and variability 75

4.2.4(b) Genetic structure 77

4.3 Results 79

4.3.1 Genetic Variability 79

4.3.2 Hardy-Weinberg equilibrium (HWE), Linkage disequilibrium, and F_{IS} 80

4.3.3 Genetic diversity, LD, HWE and FIS for the two subspecies of *Ae. aegypti* 81

4.3.4 Molecular variation and Differentiation in *Ae. aegypti* populations 92

4.3.5 Genetic structure of *Ae. aegypti* 96

4.4 Discussion 99

4.4.1 Genetic variations 100

4.4.2 Genetic Structure 102

CHAPTER 5 IDENTIFICATION OF FLAVIVIRUS AND ALPHAVIRUS VIRUSES CIRCULATING WITHIN *Aedes aegypti* FROM SUDAN COUPLED WITH METEOROLOGICAL DATA 106

5.1 Introduction 106

5.2 Materials and Methods 111

5.2.1 Study sites 111

5.2.2 *Aedes aegypti* mosquito entomological survey and meteorological records 113

5.2.3 Tissue culture Media preparation for virus isolation using C6/36 114

5.2.4 Cell passage C6/36 115

5.2.5 Cell seeding 117

5.2.6 Samples pooling and filtration 118

5.2.7 Virus propagation (Infection of cell plates) 119

5.2.8 Viral RNA extraction 120

5.2.9 The RT-PCR for Alphaviruses and flaviviruses detection: 121

5.2.9(a) Pan-Flavivirus RT-PCR 121

5.2.9(b) Pan-*Alphavirus* RT-PCR 122

5.2.9(c) Gel electrophoresis 122

5.2.9(d) NS1 Ag rapid test detection 123

5.3 Results 124

5.3.1 Entomological survey and ecological parameters 124

5.3.2 Virus detection 126

5.4 Discussion 130

5.4.1 Entomological survey: 130

5.4.2 Meteorological parameters and their correlation with Adult *Aedes* abundance 131

5.4.3 The negative results possible explanations 132

5.4.3(a) RNA degradation 133

5.4.3(b) Virus detection method accuracy 133

5.4.3(c) Low virus prevalence in *Ae. aegypti* mosquito vector 134

5.4.3(d) Average sample size 135

CHAPTER 6 GENERAL SUMMARYAND RECOMMENDATIONS 137

6.1 General discussion and conclusion: 137

6.2 Recommendations 140

REFERENCES 143

APPENDICES

LIST OF PUBLICATIONS

LIST OF TABLES

Page

Table 3.1 Sites for collecting samples, sample numbers, and the distribution of the subspecies, as well as site coordinates and larval habitats. 44

Table 3.2 Molecular diversity indices and neutrality tests for *CO1* mitochondrial gene sequences of *Ae. aegypti* from eight sites in Sudan,...46

Table 3.3 Molecular diversity indices and neutrality tests for *ND4* mitochondrial gene sequences of *Ae. aegypti* from eight sites in Sudan...47

Table 3.4 Population divergence between *CO1* mitochondrial gene sequences samples (F_{ST} value) was performed in ARLEQUIN version 3.5 among eight populations of *Ae. aegypti* subspecies from Sudan...54

Table 3.5 Population divergence between *ND4* mitochondrial gene sequences samples (F_{ST} value) was performed in ARLEQUIN version 3.5 among eight populations of *Ae. aegypti* subspecies from Sudan...55

Table 3.6 Probabilities (P values) of the Pairwise divergence (F_{ST}) between *CO1* mitochondrial gene sequences of eight populations subspecies of *Ae. aegypti* in Sudan. 55

Table 3.7 Probabilities (P values) of the Pairwise divergence (F_{ST}) between *ND4* mitochondrial gene sequences of eight populations subspecies of *Ae. aegypti* in Sudan. 56

Table 3.8 Analysis of molecular variance (AMOVA) of *CO1* mitochondrial gene of two groups of *Ae. aegypti* collected from 8 sites in Sudan....61

Table 3.9 Analysis of molecular variance (AMOVA) of *ND4* mitochondrial gene of three Groups of *Ae. aegypti* collected from 8 sites in Sudan. 61

Table 4.1 Characteristics of seven polymorphic microsatellite loci *Ae. aegypti*. Primer sequence, accession number, fluorescent tag, repeat motif, allele size range (bp):..74

Table 4.2 Number of alleles, means and total number in eight populations within seven loci...82

Table 4.3 Summary statistics of number of individuals genotyped (N), allelic range and richness (A_R), number of alleles (N_A) and genetic diversity at each locus and each population of *Ae. aegypti*................83

Table 4.4 Summary statistics of microsatellite data study over eight populations of *Ae. aegypti* from Sudan. 86

Table 4.5 Linkage disequilibrium between pairs of microsatellite loci value obtained from Arlequin indicates number of linked loci at each population of *Ae. aegypti*....................91

Table 4.6 Hierarchical analysis of molecular variance (AMOVA) of the allele frequencies of seven microsatellite loci in two subspecies (groups) of eight *Ae. aegypti* populations 93

Table 4.7 Population pairwise F_{ST} of the allele frequencies of seven microsatellite loci in two subspecies (groups) of eight populations of *Aedes aegypti* subspecies from Sudan: 93

Table 4.8 Summary statistics of microsatellite data in the two subspecies of *Ae. aegypti* from Sudan. 94

Table 5.1 Preparation of three types of media used in tissue culture: 114

Table 5.2 Master mix for Pan-Flavivirus and Pan-Alphavirus RT-PCR:.........121

Table 5.3 Cycling conditions for Pan-*Flavivirus* and Pan-*Alphavirus* RT-PCR detection: 122

Table 5.4 *Aedes aegypti* adult females collected in different months of 2014: 125

Table 5.5 Coefficient correlation between the Abundance of adult *Aedes aegypti* and the mean of rainfall, humidity and temperature:...........126

LIST OF FIGURES

Page

Figure 2.1 *Aedes aegypti* subspecies (Huang, 2004). 19

Figure 3.1 Map of Sudan showing the location of two subspecies of *Ae. aegypti* collected in eight sites of Sudan. Note all *Ae. aegypti aegypti* were found in the East, whereas *Ae. aegypti formosus* was found in the West, with the main Nile and the White Nile separating their collection sites. 36

Figure 3.2 *Aedes aegypti* subspecies: At the right *Aedes aegypti aegypti* and at the left *Aedes aegypti formosus* (the photo by Carolyn McBride) 37

Figure 3.3 Median-joining haplotype network tree of *Ae. aegypti* subspecies/populations. The tree was constructed using 46 haplotype *CO1* mtDNA sequences from the eight study sites of Sudan. The size of each circle indicates the frequency of the incidence of each haplotype in the study populations. Between two haplotypes there is a minimum of one mutation. The network consists of two groups including haplotypes from both *Aaa* and *Aaf* populations. 50

Figure 3.4 Median-joining haplotype network tree of Ae. aegypti subspecies/populations. The tree was constructed using 19 haplotype ND4 mtDNA sequences from the eight study sites of Sudan. The size of each circle indicates the frequency of the incidence of each haplotype in the study populations. Between two haplotypes there is a minimum of one mutation...............................51

Figure 3.5 Results of the Mantel test on the correlation between genetic (F_{ST}) of cytochrome-oxidase 1 (*CO1*) mitochondrial gene and geographical distances (KM) of *Ae. aegypti* samples collected in 8 locations in Sudan...53

Figure 3.6 Results of the Mantel test on the correlation between genetic (F_{ST}) of NADH dehydrogenase subunit 4 (*ND4*) mitochondrial gene and geographical distances (KM) of *Ae. aegypti* samples collected in 8 locations in Sudan...54

Figure 3.7 Polar maximum likelihood phylogenetic tree from 140 *CO1* mitochondrial gene sequences of female *Ae. aegypti aegypti* (*Aaa*) and *Ae. aegypti formosus* (*Aaf*) form eight study sites in Sudan. Different colors indicate different subspecies/forms. Note most *Aaa* clustered in one group (blue colored), while *Aaf* clustered separately (red-colored). Exceptions to this were some *Aaa* from Tokar and Barakat clustered with *Aaf* populations...........................58

Figure 3.8 Polar maximum likelihood phylogenetic tree from 91 *ND4* mitochondrial gene sequences female *Ae. aegypti aegypti* (*Aaa*) and *Ae. aegypti formosus* (*Aaf*) form eight study sites in Sudan. Different colours indicate different subspecies/forms. Note most *Aaa* clustered in one group (blue coloured), while *Aaf* clustered separately (red coloured). Exceptions to this were some *Aaa* from Tokar and Barakat clustered with *Aaf* populations. 60

Figure 4.1 Allele frequency distribution for eight populations of *Ae. aegypti* across seven microsatellite loci. 90

Figure 4.2 Relationship between pairwise estimates of genetic distance (F_{ST}) and geographical distance (km) for *Ae. aegypti* microsatellite data. Trendline shows the general pattern of increasing genetic distance with greater geographic distance (IBD). Distances between sites were presented in Appendix 5. 96

Figure 4.3 Unrooted neighbour-joining tree based on DA genetic distance at seven nuclear microsatellites of *Ae. aegypti* from eight sites in Sudan, numbers at the nodes are percentage bootstrap support (only above 50% are shown) from 1000 replicates. The scale bar represents 5% sequence divergence. 98

Figure 4.4 Three-dimensional factorial correspondence analysis (FCA) showing multivariate relationships among eight *Ae. aegypti* populations based on seven microsatellite loci variation. Population 1 Port Sudan, 2 Tokar, 3 Kassala, 4 Barakat/ Gezira, 5 Kadugli, 6 Nyala, 7 Al-Fasher, 8 Al-Junaynah 98

Figure 4.5 Bayesian clustering analysis generated through STRUCTURE and STRUCTURE HARVESTER based on eight microsatellite loci of eight *Ae.aegypti* populations (reduce analysis of ‘pure’ Group 1) to determine the exact value of K. (a) Results of assignment tests for numbers of clusters K = 2 indicated along the x-axis. (b) Mean (±SD) log posterior probabilities (c) estimate of ΔK for each value of K (putative number of populations. Each vertical line represents one individual, and y-coordinates denote each individual's percentage assignment to each of the genetic clusters, represented by a different colour. Numbers from 1-8 are the study sites, 1 Port Sudan, 2 Tokar, 3 Kassala, 4 Gezira, 5 Kadugli, 6 Nyala, 7 Al Fasher and 8 Al Junaynah. 99

Figure 5.1 Map showing the collection sites of adult *Ae. aegypti*. 112

Figure 5.2 C3/36 cells confluent cell under light microscope 116

Figure 5.3 The cell suspension after transferred into a falcon tube and centrifuged. 116

Figure 5.4 Counting cells concentration using a hemocytometer under light microscope. 118

Figure 5.5 A plate monitored under microscope and any CPE change 120

Figure 5.6 Dengue NS1 Ag Rapid Test Cassette (Standard Diagnostic, INC.) . 123

Figure 5.7 Temperature, humidity, and rainfall average at Port Sudan from January to December 125

Figure 5.8 Temperature, humidity, and rainfall average at Kassala from January to December 126

Figure 5.9 Detrended Correspondence Analysis (DCA) in PAST 2.09 software reveal the influence of temperature, humidity, and rainfall on the abundance of adult *Ae. aegypti* mosquitoes at Port Sudan study 127

Figure 5.10 Gel electrophoresis picture showing positive control band (zika virus) for pan *Flavivirus* at approximate size 612 bp after RT-PCR amplification. Lane 1: Ladder 100 bp, Lane 6: *Ae. aegypti* lab strain, Lane 5: Positive control. Lane 2,3,4: the negative samples. 128

Figure 5.11 Gel electrophoresis picture showing positive control band (*chikungunya* virus) for pan *Alphavirus* at approximate size 682 bp after RT-PCR amplification. Lane 2: Ladder 100 bp, Lane 5: Positive control. Lane 3,4,5: the negative samples. 129

Figure 5.12 Test cassette is showing the negative result (coloured line appears at control line region) when tested with field strain mosquito homogenate 129

LIST OF ABBREVIATIONS

Ae.	*Aedes*
Aaa	*Aedes aegypti aegypti*
Aaf	*Aedes aegypti formosus*
Ag	Antigen
AIC	Akaike information criterion
AMOVA	Analysis of molecular variance
bp	Base pair
BI	Bayesian inference
BP	Bootstrap percentage
COI	Cytochrome c oxidase subunit
DENV	Dengue virus
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleotide triphosphate
DF	Dengue fever
DHF	Dengue haemorrhagic fever
DSS	Dengue Shock Syndrome
FCA	Factorial correspondence analysis
FDR	False discovery rate
gDNA	genomic DNA
GIS	Geographical Information System
HWE	Hardy Weinberg Equilibrium
IAM	Infinite allele mutation
IBD	Isolation by distance
LD	Linkage disequilibrium

MCMC	Markov chain monte Carlo
MgCl2	magnesium chloride
NS1	Non-structural 1
PCR	polymerase chain reaction
RNA	ribonucleic acid
RT-PCR	Reverse-transcriptase polymerase chain reaction
SAMOVA	Spatial analysis of molecular variance
SD	Standard deviation
SE	Standard error
SSM	Single stepwise mutation
ULV	Ultra low volume
USM	Universiti Sains Malaysia
WHO	World Health Organization

LIST OF SYMBOLS

AR	Allelic richness
FCT	Variance among group
FIS	Inbreeding coefficient
FSC	Variation among population within group
FST	Fixation index
h	Number of haplotypes
Gd	Gene diversity
h	Haplotype diversity
HE	Expected heterozygosity
HO	Observed heterozygosity
k	Number of groups
n	Samples size
Ne	Effective population size
NA	Number of alleles
P	P-value
Ta	Annealing temperature
π	Nucleotide diversity

LIST OF APPENDICES

Appendix 1 Gel electrophoresis of the ITS Aedes genus mosquito identification (700bp) and the species specific (Aedes aegypti) pattern after restriction enzyme

Appendix 2 The two mitochondrial genes gel electrophoresis photos on the right the ND4 (400bp) and on the left CO1 (860bp)

Appendix 3 Scoring diagrams

Appendix 4 Scoring information of seven nuclear microsatellites markers in eight *Aedes aegypti* populations from Sudan

Appendix 5 Distances in kilometres between each site and the others

STRUKTUR GENETIK POPULASI DAN CORAK PEMBIAKAN NYAMUK DEMAM KUNING SUBSPESIS, *Aedes aegypti* DAN PENEKANAN KE ATAS TRANSMISI PENYAKIT DI SUDAN

ABSTRAK

Kajian ini dijalankan bagi mengkaji taburan, kepelbagaian genetik, struktur dan penularan virus untuk lapan populasi daripada subspesies *Aedes aegypti aegypti* (*Aaa*) dan *Aedes aegypti formosus* (*Aaf*) di wilayah Sahel, Sudan. Taburan dan variasi genetik bagi kedua-dua subspesies di bahagian berlainan di Sudan telah diterokai dengan menggunakan dua penanda mitokondria (*CO1* & *ND4*). Struktur populasi dan hubungan filogenetik telah dikaji dalam pelbagai kawasan geografi di Sudan dengan menggunakan tujuh penanda mikrosatelit. Kemudian, virus Flavivirus dan Alphavirus telah diperiksa dalam *Ae. aegypti* dan hasil kajian menunjukkan bahawa kedua-dua subspesies *Ae. aegypti* boleh ditemui di Sahelian belt. Umumnya, taburan *Aaa* dapat ditemui di empat kawasan kajian di timur Sungai Nil iaitu Port Sudan, Kassala, Tokar dan Barakat, manakala taburan *Aaf* pula dapat ditemui di empat kawasan kajian di barat Sungai Nil iaitu Kadulgi, Al Junaynah, Al Fasher dan Nyala. Gen mitokondria *CO1* dan *ND4* menunjukkan aliran gen yang rendah dan kebolehubahan genetik yang besar untuk kedua-dua subspesies. Ujian SAMOVA untuk gen mitokondria menunjukkan bahawa lapan populasi dibahagikan kepada tiga kumpulan yang berbeza secara filogeografi dalam *CO1*. Data menunjukkan semua populasi *Aaf*, *Aaa* daripada Port Sudan dan Kassala, serta *Aaa* daripada Barakat/Gezira dan Tokar adalah berkelompok bersama. Kemudian, gen mitokondria *ND4* SAMOVA menunjukkan terdapat dua kumpulan/kelompok filogeografi yang berbeza dari dua populasi subspesies. *Hardy Weinberg Equilibrium* (HWE) untuk dua subspesies menunjukkan

penyimpangan subspesies/bentuk *Aaa* daripada HWE antara 7 lokus, manakala HWE untuk subspesies / bentuk *Aa*f sebaliknya menunjukkan pelepasan di semua lokus kecuali A10 dan M201. Plot Analisis Koresponden Faktor (FCA) menunjukkan bahawa populasi *Aaa* Port Sudan, Tokar dan Kassala dikelompokkan bersama (yang konsisten dengan filogenetik), manakala populasi *Aaf* Al-Fasher dan Nyala dikelompokkan bersama dan populasi Gezira, Kadugli, dan Al-Junaynah dinyatakan sebagai satu kumpulan. Analisis kelompok Bayesian menstrukturkan populasi kepada dua kumpulan, menyiratkan dua kumpulan berbeza secara genetik (subspesies). Pengasingan dengan ujian jarak menunjukkan hubungan korelasi sederhana hingga kuat yang signifikan antara jarak geografi dan variasi genetik ($p = 0.005$, $r = 0{,}586$) dan ($p = 0.003$, $r = 0{,}391$) di kedua-dua gen mitokondria dan mikrosatelit *CO1*, masing-masing. Walau bagaimanapun, ujian tersebut menunjukkan hubungan yang tidak signifikan ($p = 0.136$) tetapi hubungan kuat diperhatikan dalam gen *ND4* ($r = 0.710$). Prosedur kajian telah disahkan dengan menyatakan tidak adanya Flavivirus dan Alphavirus dalam sampel yang diperoleh dari kedua lokasi kajian. Tinjauan entomologi nyamuk betina dewasa *Aedes* yang diperolehi mendapati bahawa jumlah tertinggi adalah pada bulan April dan jumlah terendah adalah pada bulan September. Di dua tapak endemik arbovirus, nyamuk *Aedes* dewasa dikumpul sepanjang tahun dan untuk tujuan pengesanan Alphavirus dan Flavivirus, selepas prosedur kajian telah disahkan, ia mendedahkan ketiadaan Flavivirus dan Alphavirus dalam sampel yang diperolehi dari kedua-dua tapak kajian. Tinjauan entomologi ke atas nyamuk betina dewasa *Aedes* mendedahkan bahawa kelimpahan tertinggi kutipan betina dewasa pada bulan April dan terendah pada bulan September. Ujian korelasi antara kelimpahan dan data meteorologi menunjukkan korelasi negatif diantara suhu dan hujan manakala kelembapan menunjukkan positif korelasi. Dua subspesies vektor *Ae. aegypti* telah

dikenal pasti dalam kajian ini. Walaupun spesies tersebut mempunyai taburan geografi yang berbeza, kebolehubahan genetiknya mendedahkan perbezaan genetik yang ketara antara kedua-dua subspesies. Struktur genetik kajian membahagikan spesies tersebut kepada dua kumpulan walaupun ia tidak dipisahkan sepenuhnya. Tiada virus yang dikesan dalam kutipan nyamuk dewasa manakala semasa tinjauan entomologi, pengumpulan data meteorologi dewasa telah dikumpul dan dikaitkan dengan kepadatan nyamuk yang menunjukkan korelasi positif dengan kelembapan, manakala korelasi negatif dengan suhu dan kelembapan.

POPULATION GENETIC STRUCTURE AND BREEDING PATTERN OF THE YELLOW FEVER MOSQUITO *Aedes aegypti* SUBSPECIES WITH EMPHASIS ON THEIR DISEASE TRANSMISSION IN SUDAN

ABSTRACT

The current study investigated the distribution, genetic diversity, and structure, and virus transmission for eight populations of the two *Aedes aegypti* subspecies *Aedes aegypti aegypti* (*Aaa*) and *Aedes aegypti formosus* (*Aaf*) in Sudan's Sahelian region. The distribution and genetic variations of the two subspecies in different parts of Sudan were explored using two mitochondrial markers (*CO1* & *ND4*). Using seven microsatellite markers, the population structure, and phylogenetic relationships were investigated in their various geographical areas. *Flavivirus* and *Alphavirus* viruses were examined in *Ae. aegypti*. The results revealed that both *Ae. aegypti* subspecies are found in the Sahelian belt, *Aaa* and distribution was found in four study sites east of the Nile River named by Port Sudan, Kassala, Tokar and Barakat then *Aaf* distributed in the other four sites west of the Nile River named by Kadugli, Al Junaynah, Al Fasher and Nyala. The *CO1* and *ND4* mitochondrial genes demonstrated low gene flow and considerable genetic variability between the two subspecies. The SAMOVA test for mitochondrial genes revealed that the eight populations were divided into three phylogeographically distinct groups in *CO1*, all *Aaf* populations clustered together, while *Aaa* populations from Port Sudan and Kassala clustered together, and *Aaa* populations Barakat/Gezira and Tokar clustered together. In addition, the *ND4* SAMOVA revealed two phylogeographically distinct groups/clusters of the two subspecies populations. Hardy Weinberg Equilibrium (HWE) for the two subspecies revealed *Aaa* subspecies/form deviation from HWE

among the seven loci, while the HWE for the *Aaf* subspecies/form, on the other hand, exhibited departure in all loci except A10 and M201. The Factorial Correspondence Analysis (FCA) plots revealed that Port Sudan, Tokar, and Kassala *Aaa* populations were clustered together (which is consistent with the unrooted phylogenetic tree), while Al-Fasher and Nyala *Aaf* populations were clustered together, and Gezira, Kadugli, and Al-Junaynah populations were revealed to be one group. The Bayesian cluster analysis structured the populations into two groups, implying two genetically distinct groups (subspecies). Isolation by distance test revealed a moderate to strong relationship significant correlation between geographical distance and genetic variations ($p = 0.005$, $r = 0.586$) and ($P= 0.003$, $r = 0.391$) across both *CO1* mitochondrial and microsatellite genes, respectively. However, the test revealed a non-significant ($p = 0.136$) but strong relationship observed in the *ND4* gene ($r = 0.710$). In the two endemic sites of the arboviruses adult *Aedes* mosquitoes were collected throughout the year and for the purpose of *Alphaviruses* and *Flaviviruses* detection, after the study procedures had been validated, it revealed the absence of *Flavivirus* and *Alphavirus* in the obtained samples from both study sites. The entomological survey of *Aedes* adult female mosquitoes revealed that the highest abundance of collected adult females in April and the lowest in September and the correlation test between the abundance and the meteorological data revealed negative correlation with the both the temperature and rainfall while humidity showed a positive correlation. Two subspecies of the *Aedes aegypti* vector were identified in this study, and while they had different geographic distributions, their genetic variability revealed significant genetic differences between the two subspecies. The study's genetic structure divided them into two groups, although they were not completely separated. There are no viruses detected within the adult mosquito collection, while during the entomological

survey of adult collection meteorological data has been collected and correlated with the mosquito abundance which showed a positive correlation with the humidity and a negative correlation with the temperature and humidity.

CHAPTER 1

INTRODUCTION

1.1 General introduction

Arthropod-borne viruses (Arboviruses) continue to impose significant health and economic burdens on people living in endemic areas. Some of the most threatening viruses are spread primarily by *Aedes* mosquitoes (Ogunlade *et al.,* 2021). *Aedes aegypti* is a major vector known to transmit the most important arboviral diseases, including yellow fever, dengue, chikungunya, and zika viruses worldwide (Kamal *et al.,* 2018; Kraemer *et al.,* 2015). Despite its global importance, *Ae. aegypti* is almost certainly originated from Africa (Powell *et al.,* 2018). *Aedes aegypti* is a tropical and subtropical mosquito spreading all over the world. However, it is indigenous to the Sub-Saharan and African Sahelian regions, including Cameroon, Kenya, Nigeria, Morocco, Senegal, Western Sahara Tunisia, Algeria, Egypt, and Sudan (Kamal *et al.,* 2018; Kweka *et al.,* 2019). The exact burden of *Aedes*-borne arboviruses in Africa is unclear, even though tens of thousands die each year in Africa and the threat of outbreaks remains high, however, It is estimated that 70% (831 million) of the population is susceptible to arboviruses (Weetman *et al.,* 2018b).

Sudan has suffered from several arboviral diseases for decades since the 18th century (Ahmed *et al.,* 2020). Since 1955, Lewis highlighted the role of *Ae. aegypti* in yellow fever outbreaks in Sudan (Lewis, 1955). This notion was validated by some consequent reports since the species was found in places where outbreaks of arboviruses have happened (Ahmed *et al.,* 2019; Hajahmed *et al.,* 2020; Himatt *et al.,* 2015; Soghaier *et al.,* 2018). Yellow fever virus (YF) has been recorded since 1940 in Sudan and continued in recent decades with serious outbreaks (Soghaier *et al.,* 2013a;

Bushra *et al.,* 2019). Dengue fever viruses (DFV) have occurred in Sudan for many decades, mostly in eastern Sudan (Ahmed *et al.,* 2020; Himatt *et al.,* 2015; Seidahmed *et al.,* 2012), nevertheless lately in 2015, dengue fever first outbreak reported in Darfur states western Sudan (Ahmed *et al.,* 2019). Additionally, chikungunya viruses as well have been described in the country several times in various geographical localities (Soghaier *et al.,* 2018; Ahmed, 2020; Gould *et al.,* 2008; Adam *et al.,* 2016), in recent times, a significant chikungunya outbreak was stated 2018 - 2019 from several states of Sudan (Bushra *et al.,* 2019; Bracco *et al.,* 2007).

Aedes aegypti was first reported in Sudan in Khartoum by Balfour (1903) (Ahmed *et al.,* 2020b; Lewis, 1953). The vector was described broadly in different geographical areas in the country involving the eastern region (Port Sudan and Kassala), central region (Wad Madani and Khartoum), west part (Al Junaynah, and Al Fasher), and in the south (Nuba mountains) (Hajahmed *et al.,* 2020; Lewis, 1943, 1953; Mattingly, 1957; Seidahmed *et al.,* 2012). Recent studies have documented the presence of *Ae. aegypti* in endemic areas of dengue, chikungunya, and yellow fever. The absence of other potential vectors at these locations indicated the definitive role of *Ae. aegypti* on the transmission (Ahmed *et al.,* 2019; Elnour *et al.,* 2020). *Aedes aegypti* is known as the domestic mosquito, rest and breeds within and around the human environment mostly in human-made water containers (e.g. water jars, barrels, and plant vases) and feeds diurnal feeding on humans (anthropophilic) (Morrison *et al.,* 2008; Scott & Takken, 2012).

In Africa, contrary to other continents, two forms/subspecies are identified: *Ae. aegypti aegypti* (*Aaa*) and *Ae. aegypti formosus* (*Aaf*). The domestic form *Ae. aegypti aegypti* is pale brown having white abdominal scales, characterized by anthropophilic feeding and breed in man-made artificial containers. The wild form *Ae. aegypti*

formosus is dark in color and without the white abdominal scales, more sylvatic, breeds in natural habitats, such as tree holes, and more zoophilic in feeding (Mattingly, 1957; McBride *et al.,* 2014).

The dark form (*Aaf*) is confined completely to African content and distributed in the south of the Sahara until the northern borders in Sudan (Lewis, 1955). *Aedes aegypti formosus* has been suggested to be abundant all over Africa, however, it was identified in various parts of Africa including Liberia, Kenya, Senegal, Ghana, Uganda, and Sudan (Brown *et al.,* 2011; Captain-Esoah *et al.,* 2020; Gloria-Soria *et al.,* 2016; Lewis, 1944; Paupy *et al.,* 2010).

The *Ae. aegypti* subspecies are recognized to be different in their capacity to transmit arboviruses, *Aaa* populations are reported to have a significantly higher vector competence than *Aaf* for both yellow fever virus (YFV) and dengue virus (DENV-2) (Black *et al.,* 2002; Dickson *et al.,* 2014). Given that the two subspecies demonstrate dissimilar susceptibilities to arboviruses the understanding of their distribution and transmission role is essential. Consequently, a variety of genetic markers have been developed to understand the genetic differentiation and structure of *Ae. aegypti* vector populations. These markers consist of biochemical and molecular tools, including allozymes, nuclear DNA, microsatellite, and mitochondrial DNA (Tabachnick & Powell, 1979; Gloria-Soria *et al.,* 2016; Joyce *et al.,* 2018; Crawford *et al.,* 2017).

Since 1979, the genetic diversity of *Ae. aegypti* in different parts of the globe have been described (Powell & Tabachnick, 1979), genetic variation of thirty-four populations of *Ae. aegypti* around the globe were studied at 19 isozyme loci, their results revealed two primary groups may be distinguished as: the dark in colour,

typically sylvan African subspecies *formosus* and the light in colour, domestic subspecies *aegypti* (Tabachnick & Powell, 1979).

Recently, mitochondrial and microsatellite markers had been widely used in population genetic studies in disease vectors, particularly in mosquitoes (Ambrose *et al.,* 2014; Liu *et al.,* 2019). The mitochondrial DNA markers have been extensively and commonly used worldwide in phylogeny and population genetic studies of *Ae. aegypti*, their efficiency anticipated to their maternal inheritance and rapid divergence (Naim *et al.,* 2020; Joyce *et al.,* 2018; Moore *et al.,* 2013; Paupy *et al.,* 2012). Furthermore, microsatellites give more resolution for assessing genetic divergence between populations within taxa, allowing population biologists to draw more accurate predictions about population structure and, in some circumstances, the mobility between populations (Ellegren, 2004; Lanzaro & Tripet, 2005).

The genetic structure of mosquito populations and gene flow between them are important for maintaining effective control. The quantity of genetic exchange between groups not only reveals the type of mosquito movement or dispersal, but also provides an estimate of gene flow within the mosquito population (Ravel *et al.,* 2001). Furthermore, through genetic delineation of vector populations according to geographic locations, the potential threat of invading populations can be determined utilizing population genetic structure information (Failloux *et al.,* 2002; Halstead, 2019). Knowledge of population genetics could also be used to determine and improve the insecticide vector control, as well as to develop a model that can predict the rate at which insecticide resistance develops. The physiological, ecological, and behavioral heterogeneity among vector populations serves as a logical foundation for developing this strategy. This diversity among populations distinguishes each population as a distinct entity that can be used to develop strategies for long-term success. As well as

to research or develop conditions that allow transgenic mosquitoes to spread efficiently to replace wild type vectors, thereby blocking the transmission (Paupy *et al.,* 2004; Paupy *et al.,* 2008).

Detection of the viruses inside infected mosquitoes giving information on infection rates in such mosquito populations as well as predicting epidemics. In addition, the identification of these viruses is considered a key to outbreak control and reduces the burden and risk of arboviruses (Lequime *et al.,* 2016). Although *Ae. aegypti* as a vector of yellow fever, dengue fever, chikungunya and other arboviral diseases has a critical involvement in Sudan, its distribution, population dynamics, virus's transmission, vector density and behaviour, genetic structure, and genetic variation across various endemic regions in the country are little known.

1.2 Justification of the study

The subspecies remained distinct and had a different vectorial capacity in transmitting diseases, with *Aaa* having a greater capacity to transmit diseases than *Aaf* (Black *et al.,* 2002; Dickson *et al.,* 2014), consequently the study of their distribution and genetic differences essential towards improving the surveillance and control measures in the direction of arboviral diseases. However, there has been insufficient research on the genetic structure of the two subspecies in the African Sahelian region (Sylla et al, 2009; Crawford et al, 2017; Elnour et al 2020). Additionally, there is lack of reported information on virus infections within mosquito vectors in the area. Despite *Ae. aegypti* critical role in the transmission of important arboviral diseases in Sudan, such as yellow fever, dengue fever, and chikungunya, little is known about its distribution, population dynamics, genetic structure, and genetic variants across Sudan's endemic geographical regions.

Up to date there is no effective treatment for the dengue fever and other arboviral diseases. Insecticide-based vector control is less successful due to insecticide resistance (Huber *et al.,* 2001). Therefore, vector surveillance and elimination are considered the only effective method for dengue control, in the lack of a viable vaccine, specialized therapy, or curative medicine, the primary focus of dengue transmission prevention and control has switched to vector population reduction (Ravel *et al.,* 2001). Knowledge of population genetics can be useful in developing novel vector control strategies, scientists are now concentrating on genetically altering vectors to change their ability to transmit pathogens, these genetically modified vectors could then be introduced into natural populations to reduce vectorial capacity. However, in situations where population bottlenecks occur, allele frequencies may change and some alleles may disappear completely, resulting in the loss of the refractory genes. In such circumstances, the need for the reintroduction of these transgenic mosquitos grows, and a thorough understanding of population structuring as well as the temporal scale of gene frequency fluctuations becomes critical (Collins et al. 2000). The quantity of genetic exchange between groups not only estimates gene flow within the mosquito population, but it also shows the type of mosquito movement or dispersal. This understanding of gene flow, as well as their susceptibility to harbour and transmit diseases, is required for the establishment of effective control. Furthermore, by genetically describing vector populations according to geographic locations, the potential threat of invading populations can be determined using population genetic structure information (Failloux *et al.,* 2002). Additionally, a knowledge of population genetics could also be used to determine the best time to begin insecticide treatment and to create a model that can predict the rate of insecticide resistance development. Using population genetic structure information, the potential

threat of invading populations can be assessed by genetic delineation of vector populations according to geographic regions (Failloux et al. 2002), and to research or develop conditions that allow transgenic mosquitoes to spread efficiently to replace wild type vectors, thereby blocking pathogen transmission (Paupy *et al.,* 2004).

Virological monitoring of *Aedes* transmitted arboviruses in *Ae. aegypti* provides a non-human surveillance alternative for detecting viral circulation and following transmission dynamics during outbreaks (Madewell *et al.,* 2019). The emergence of human arboviral diseases is currently being used as a trigger for vector control measures. However, there will almost likely be a delay, and vector control may be insufficient to manage the pandemic (Lee & Rohani, 2005). Therefore, prevalence of dengue-infected mosquitoes must be monitored for a variety of reasons. It might be used as an early warning system for the spread of dengue fever in a given area (Kow *et al.,* 2001; Madewell *et al.,* 2019).

The influence of human activities and environmental conditions on mosquito population divergence is significant (Lequime *et al.,* 2016; Paupy *et al.,* 2004). Human activities such as water storage could cause significant extinction/recolonization of mosquito populations, which, combined with genetic drift, then increased genetic differentiation among mosquito populations (Bennett *et al.,* 2016; Paupy *et al.,* 2004).

1.3 Aim and objectives of the study

The current study aim to use strong molecular genetics tools to investigate the phylogeny, population structure, and population genetics of *Ae. aegypti* subspecies in Sudan. This study was aimed to investigate the distribution, genetic diversity, and virus transmission of the eight populations of the two subspecies of *Ae. aegypti* across the Sahelian belt in Sudan. Using two mitochondrial markers and seven microsatellite markers, the genetic diversity within and between the two subspecies in different areas of Sudan were compared. The specific objectives are as below:

Objective 1: To investigate the geographical distribution and genetic variations among *Ae. aegypti* subspecies/forms using mitochondrial DNA (*CO1* & *ND4*) in different study sites in Sudan.

Objective 2: To determine the population structure and phylogenetic relationships of *Ae. aegypti* populations in their different geographical locations in Sudan using microsatellite markers.

Objective 3: To identify the arboviruses (*Flavivirus* and *Alphavirus*) that are circulating within *Ae. aegypti* vector in Port Sudan and Kassala sites of Sudan and relate the vector abundance with meteorological data.

CHAPTER 2

LITERATURE REVIEW

2.1 Arthropods transmitted diseases/arboviruses

Arboviruses have spread rapidly around the world. They are disseminated largely due to international travel and the extensive presence of urban mosquito vectors among susceptible populations (Wilder-Smith *et al.,* 2017). The most prevalent arboviral infections affecting humans are mosquito-borne viruses divided into three families: Flaviviridae, which includes dengue and yellow fever viruses, Togaviridae, which includes chikungunya, and Bunyaviridae, which includes rift valley fever (Braack *et al.,* 2018).

However, the extent of arboviral diseases in Africa is largely unknown, and the population at risk of infection by *Aedes*-borne viral diseases is estimated at 271 million (63%) at risk of dengue infection, 34% for zika viruses, 23% for chikungunya, and 2% for yellow fever infection (Weetman *et al.,* 2018b).

2.1.1 Yellow fever

The Yellow fever (YF) was considered a global life-threatening arboviral disease that originated from Africa, however in the last decade unpreceded outbreak reports of yellow fever outside Africa (Tabachnick, 1991a). The YF killed about 78,000 deaths annually in African content, and approximately 51,000 deaths in 2018 (Litvoc *et al.,* 2018; Gaythorpe *et al.,* 2020a; Gaythorpe *et al.,* 2020b).

Yellow fever has three cycles of transmission, the sylvatic cycle occurs in the jungle including primate animals, and transmitted by wild species of *Aedes* mosquitoes, the intermediate cycle, transmitted to humans living or working near the forest, the transmission could be either animal (monkey) - human or human-human, and the urban cycle mainly transmitted by *Ae. aegypti*, is a human-human cycle usually initiated by an infected human from the jungle (Litvoc *et al.,* 2018; Gaythorpe *et al.,* 2020a; Gaythorpe *et al.,* 2020b).

2.1.2 Yellow fever in Sudan

The WHO classified Sudan as a country of high-risk transmission of yellow fever, since the 1800s, the yellow fever was believed to be endemic in Sudan, during the year 1940 Nuba mountains suffered from a huge outbreak with more than 15,000 cases and a 10% fatality rate (Ahmed *et al.,* 2020a; Soghaier *et al.,* 2013). Yellow fever outbreaks continued in South and West Sudan in the years 2005, 2012/13 respectively (Ahmed *et al.,* 2019; Ahmed *et al.,* 2019; Farnon *et al.,* 2005; Soghaier *et al.,* 2013b).

2.1.3 Dengue fever

Dengue viruses have become the most widespread, with the highest morbidity and mortality in the world, with nearly four million people at risk of infection (Weetman *et al.,* 2018; Fritzell *et al.,* 2018; Ahmed *et al.,* 2019).

Dengue is estimated to infect about 390 million people globally. In the last decades, dengue cases increased 30 folds and this may attribute to urbanization which helps the urban *Aedes* vectors to breed, human travels, virus development, and socioeconomic factors (Ferreira, 2012; Stanaway *et al.,* 2016; Zeng *et al.,* 2021). However, dengue was believed to originate from Africa, it was firstly reported and

isolated in Nagasaki, Japan 1943 during an outbreak. Dengue viruses belong to the family Flaviviridae is a single-strand positive-sense RNA virus (Wilder-Smith *et al.,* 2017).

2.1.4 Dengue fever in Sudan

Dengue fever (DF) had been reported in Sudan since 1907, in the last decades outbreaks of dengue were reported from different parts of the country (Ahmed *et al.,* 2019; Ahmed *et al.,* 2019; Seidahmed *et al.,* 2012). Port Sudan a dengue-endemic area with huge dengue cases recorded during outbreaks, Kassala also reported many outbreaks with high severity, Kordofan and Darfur states, El Gadarif, Gezira, and Khartoum States (Ahmed *et al.,* 2019; Ahmed *et al.,* 2020b; Elduma *et al.,* 2020; Himatt *et al.,* 2015; Seidahmed *et al.,* 2012; Soghaier *et al.,* 2013).

2.1.5 Chikungunya virus

Chikungunya is an infectious disease spread by mosquitos and caused by the Chikungunya virus (CHIKV), an Alphavirus belonging to the Togaviridae family. Chikungunya virus was originated in zoonotic cycles in Africa, the first case of CHIKV infection was discovered in 1952 in a feverish patient in Tanzania, Africa (Ahmed *et al.,* 2020a; Fritzell *et al.,* 2018; Khongwichit *et al.,* 2021; Weetman *et al.,* 2018b). CHIKV is an RNA virus with an envelope. The 11.8-kb CHIKV genome encodes non-structural proteins (NsP), capsid (C), and envelope (E) proteins (Khongwichit *et al.,* 2021). CHIKV nucleotide and amino acid sequence analyses from various isolates have revealed three primary CHIKV lineages: East/Central/South African (ECSA), West African (WA), and Asian. The virus mainly occurred in Africa

and Asia, there were different genotypes west/east/central and south African, and the Asian one (Sam & Bakar, 2006; Humphrey *et al.,* 2017).

2.1.6 Chikungunya in Sudan

The first report of the disease in Sudan in 1973 showed that the seroprevalence of CHIKV infection in Central Sudan (Ahmed *et al.,* 2020a). Chikungunya reported in many states: Kassala, Red Sea, Al Gadaref, River Nile, Northern State, South Darfur, and Khartoum (Adam *et al.,* 2016; Ahmed *et al.,* 2020a; Soghaier *et al.,* 2018). A significant chikungunya outbreak was recently reported in eastern Sudan, with more than 47,000 cases (Ahmed *et al.,* 2020a).

2.2 *Aedes aegypti* biology, ecology, classification, and distribution

Christopher mentioned that, *Ae. aegypti* had many nomenclatures in different areas and countries combined with the lack of control measures toward it during that time 1500–1900, only after 1900, after mosquito's transmission role become known by scientists and health workers as a public health concern (Powell *et al.,* 2018).

Today the *Ae. aegypti* "yellow fever mosquito" is known as an urban mosquito that is considered a major vector of many arboviral diseases worldwide, such as dengue, yellow fever, chikungunya, and zika viruses (Weetman *et al.,* 2018; Gloria-Soria *et al.,* 2016; Powell *et al.,* 2018).

Aedes aegypti is a major transmitter of arboviral diseases, it is normally acquiring the viruses while taking an infected blood meal, inside the mosquito midgut the virus propagates, and the mosquito becomes infected after 8-12 days (Black *et al.,* 2002; Dickson *et al.,* 2014). It is effective at harboring these viruses since it is a

domesticated species that is bred, rests, and feeds in the human environment (Jansen & Beebe, 2010; Souza-Neto *et al.,* 2019).

2.2.1 *Aedes aegypti* biology and morphology

The *Aedes* mosquito eggs are normally laid on a moist surface one at a time, the eggs are submerged in water within 48 hours, they hatch into larvae. Eggs of *Ae. aegypti* are resistant to desiccation and can survive for up to a year in the absence of water (Faull & Williams, 2015; Sota & Mogi, 1992). Before becoming pupae, the larvae grow and develop through four instars. After around 48 hours, adult mosquitoes emerge from the pupae (Kweka *et al.,* 2019; Amoa-Bosompem *et al.,* 2021; Muktar *et al.,* 2016). Depending on the temperature, the entire process from egg to adult can take as little as 6-9 days.

The adult females are a diurnal active mosquito has a preference to feed on human blood (anthropophilic) while males do not feed on blood. The females bite during the daytime especially dusk and dawn, they have intermittent feeding and their gonotrophic cycle predominantly includes multiple hosts (Muktar *et al.,* 2016; Powell, 2018; Wilder-Smith *et al.,* 2017). The females lay their eggs on a suitable moist surface after developing eggs for 2-3 days, completing the life cycle. Their favorite breeding sites are man-made artificial containers including barrels, discarded tires, plant/flower pots, and other homes- water storage containers (Jansen & Beebe, 2010; Soares-Pinheiro *et al.,* 2016; Wilder-Smith *et al.,* 2017).

2.2.2 *Aedes aegypti* distribution and ecology

The *Aedes aegypti* is a globally distributed in tropical and subtropical native to Sub-Saharan Africa, involving Kenya, Senegal, Cameroon, Nigeria, Morocco,

Western Sahara, Algeria, Tunisia, Egypt, and Sudan (Kweka *et al.,* 2019; Weetman *et al.,* 2018; Gloria-Soria *et al.,* 2016). *Aedes aegypti* is found globally between 35° N and 35° S, at latitudes that roughly correspond to a 10°C winter isotherm, which appears to be the species' limitation temperature for overwintering. The species are highly adapted to urban environments, breeding in stagnant water found in manufactured containers, garbage heaps, and tyres (Kweka *et al.,* 2019; Kamal *et al.,* 2018; Brady *et al.,* 2014).

Aedes aegypti has environmental adaptations that make it highly robust, or capable of quickly returning to pre-disaster levels after interruptions caused by natural events such as drought or human control methods (Faull & Williams, 2015). One of the remarkable adaptations is the capacity of *Ae. aegypti* eggs to endure drought. *Aedes aegypti* eggs can survive without water for several months, and when water is available, larvae can hatch from the eggs (Kamal *et al.,* 2018; Kweka *et al.,* 2019; Pinheiro *et al.,* 2016).

2.2.3 *Aedes aegypti* subspecies morphology

Aedes aegypti have two subspecies/forms, this is known to occur only in the African continent (Figure 2.1) (Huang, 2004; Mattingly, 1957; Powell & Tabachnick, 2013a). The two forms/subspecies were first described by Mattingly named; *Ae. aegypti aegypti* (*Aaa*) and *Ae. aegypti formosus* (*Aaf*). The pale to brownish black domestic variant *Ae. aegypti aegypti* is found throughout Africa's coastal plains as well as in the tropics and subtropics outside of Africa.

Aedes aegypti breeds exclusively in human-made containers and is highly anthropophilic (preferring to bite humans) (Gaythorpe *et al.,* 2020; Ngugi *et al.,* 2017). Moreover, The darkly coloured and often sylvan *Ae. aegypti formosus*, breeds in forest

settings, and is mostly zoophilic (preferring to bite nonhuman animals) , it is darker in colour and lacks of white abdominal scales (Captain-Esoah *et al.,* 2020; Huang, 2004; Mattingly, 1957; McBride *et al.,* 2014). The dark form (*Aaf*) is restricted to Africa, south of the Sahara (Lewis, 1955; Powell *et al.,* 2018).

2.2.4 *Aedes* subspecies distribution and ecology

Aedes aegypti has been identified in the majority of Eastern Mediterranean region countries, including Djibouti, Egypt, Oman, Pakistan, Saudi Arabia, Somalia, Sudan, and Yemen (Kweka *et al.,* 2019). The domestic form subspecies *Ae. aegypti aegypti* (*Aaa*) distributed in tropical and subtropical countries, while *Ae. aegypti formosus (Aaf)* was reported exclusively in Sub-Saharan Africa except in the coastal areas (Gloria-Soria *et al.,* 2016; Moore *et al.,* 2013; Mattingly, 1957). It has been recorded in numerous African countries, including Liberia, Kenya, Senegal, Ghana, Uganda, and Sudan and has also been hypothesized that the dark variety is widespread throughout Africa (Brown *et al.,* 2011; Captain-Esoah *et al.,* 2020; Gloria-Soria *et al.,* 2016; Lewis, 2009; Paupy *et al.,* 2010).

The *Aaa* is a light anthropophilic species that breeds in an urban environment, while *Aaf* is a dark wild form that prefers natural breeding places and feeds on animal blood (Moore *et al.,* 2013; Brown *et al.,* 2011; Powell & Tabachnick, 2013; Captain-Esoah *et al.,* 2020). *Aedes aegypti* thrives in human habitats, whereas *Ae. aegypti formosus*, the sylvatic form, breeds in forest tree holes (Sylla *et al.,* 2009). The two subspecies were discovered to be sympatric along Kenya's coast (Ngugi *et al.,* 2017).

Mattingly (1957) reported that the *Ae. aegypti* was discovered in natural breeding grounds in Kenya and Uganda, away from human settlements. Lewis (1953) also suggested that under urbanization conditions, *Aaf* can become exceedingly

domesticated and contribute significantly to the transmission of human arboviruses (Gloria-Soria *et al.,* 2016; Lewis, 1953; Powell & Tabachnick, 2013a). Similarly, (Futami *et al.,* 2020) discovered recently that the two subspecies were sympatric in both artificial and natural containers, implying that *Aaf* had adapted to different habitats (Sylla *et al.,* 2009).

Human activities including water storage and use, and disposal of water-holding containers have a considerable impact on *Ae. aegypti* breeding in individual houses (Paupy *et al.,* 2004). The most important primary elements that may influence *Ae. aegypti* development in different container types include the frequency of water replenishment, the availability of food for the larvae, the degree of solar exposure, and container covering (Ngugi *et al.,* 2017).

2.3 *Aedes aegypti* subspecies classification/debates

2.3.1 Morphological classification history /differentiation

Mattingly (Mattingly, 1957) was first categorize the two *Ae. aegypti* subspecies as distinctive and geographically representative forms. He described the domestic pale form variety has a range of colour depths but is always paler and browner (at least in the female) and has pale scaling on the first abdominal tergite, while the other wild dark form varies from the pale form in that the strongly blackish dark portions of the thorax and abdomen, and there is no bleaching or extension of pale scaling on any part of the body. The first abdominal tergite never has any pale scales (Captain-Esoah *et al.;* Paupy *et al.,* 2010).

While a broad discussion on *Ae. aegypti* types or subspecies has occurred after Mattingly (1957) considered them as subspecies; McClelland (1874) wrote despite the

colour, behaviour, and scaling pattern differences in population, *Ae. aegypti* cannot be divided into specific infrastructure entities. In addition, Moore (1979) indicated that they have not considered a distinct species once they can mate and produce viable generation, they tested in the laboratory and found to freely mate with each other. The two subspecies were found in sympatry in four collections north of the Gambia in the middle Savannah region (Sylla *et al.,* 2009).

(Huang, 2004) documented that the diagnostic features to identify the *Aaa* subspecies are the large, median patch of pale scales in the first abdominal tergite; and a row of small white scales along posterior border female terga II–VII, and in the *Aaf* form the first abdominal tergite lack of the median patch of pale scales as well as the row of small pale scales along posterior terga II–VII. (Tabachnick, 1991a) suggested that the ancestor of the domestic form of *Ae. aegypti* breed in wild habitat (ex: tree holes) and fed on non-human blood meals.

2.3.2 Molecular classification of the two subspecies

Previously genetic studies using isoenzyme loci supported up Mattingly's classification, showing that the sylvan and domestic variants of *Ae. aegypti*, at least in East Africa, clearly reflect two sympatric gene pools (Tabchnick & Powell, 1979; Fallioux *et al.,* 2002). *Aedes aegypti* subspecies/forms have stayed genetically distinct for more than 30 years and even more (Brown *et al.,* 2011; Mattingly, 1957; Powell & Tabachnick, 2013a).

Despite the sympatric distribution of the forms in Rabai district of Kenya, which appears to provide ample potential for gene flow, the dimorphism is persists, these findings support the theory that *Ae. aegypti* is a single polytypic species (Moore, 1979). A study by Tabachnick (1985) used isozyme analysis to establish genetic relatedness discovered that west African sylvan is genetically distinct from other domestic forms from around the world. Powell & Tabachnick (2013) indicated that morphological similarities/differentiation seemed not related to genetic variations specified by common molecular markers such as allozymes, microsatellite, and SNPs. The scale pattern of these two subspecies/forms is considered a great genetic variable character between and within various populations of *Ae. aegypti* in different habitats, genetic variation behind the scaling pattern simply a single Mendelian genes variation while no other interacting genes evolved (Powell & Tabachnick, 2013a).

Brown and his group (Brown *et al.,* 2011) demonstrated that *Aaa* and *Aaf* subspecies/forms other parts of Africa did not show any significant genetic differentiation. A study in Senegal found that subspecies accounted for only a minor, non-significant percentage of the variance, implying that West African *Aaa* and *Aaf* are monophyletic and the dark form *Aaf*, is the ancestor of the light form *Aaa* (Sylla *et al.,* 2009). Global research by (Gloria-Soria *et al.,* 2016) revealed that in Kenya which the two subspecies existed, they showed different patterns, they founded to be mixing freely in a district, while they show the opposite in another geographically near the place, they remain genetically distinct while they are together in the same habitat, the different ecology of the geographically closer sites may play a role whether the subspecies inbreed or not (Gloria-Soria *et al.,* 2016). Outside of Africa, *Ae. aegypti* populations are descended from ancestral African populations and are monophyletic (Brown *et al.,* 2011; Futami *et al.,* 2020; Gloria-Soria *et al.,* 2016).

Recently, a group of researchers from Ghana identified the *Ae. aegypti* subspecies molecularly using a random amplified polymorphic DNA polymerase chain reaction. The molecular approach they use (Random amplification of polymorphic DNA-PCR) is an effective tool for genetically isolating and differentiating *Aedes* populations. The two morphologically identified subspecies (n =110) were identified as *Ae. aegypti aegypti* (69%) and *Ae. aegypti formosus* (31%) (Captain-Esoah *et al.,* 2020).

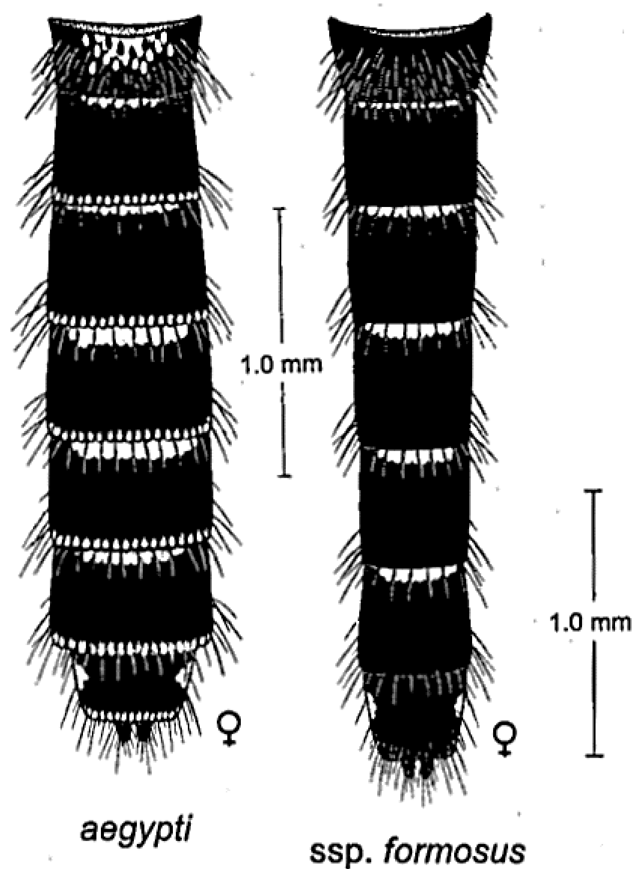

Figure 2.1 *Aedes aegypti* subspecies (Huang, 2004).

2.3.3 *Aedes aegypti* subspecies vector competence

Aedes aegypti subspecies are considered to differ in their capacity to transmit diseases, *Aaa* populations have more capacity to harbour and transmit both dengue

fever and yellow fever viruses than *Aaf* populations (Black *et al.,* 2002; Dickson *et al.,* 2014). However reports indicate that *Aaf* populations have less competence than *Aaa*, the vector competence is a result of several aspects including vector abundance, distribution, longevity, biting behaviour, and adaptations to different environmental factors (Black *et al.,* 2002; Captain-Esoah *et al.,* 2020; Dickson *et al.,* 2014).

Interestingly, in west Africa, a group of researchers found that populations of *Aaf* can be competently transmitters of Flaviviruses (Futami *et al.,* 2020). Since the two subspecies show different susceptibilities to dengue viruses it is important to understand their distribution and role in disease transmission.

2.4 Genetic variations studies in *Aedes aegypti* subspecies

Various genetic population studies in *Ae. aegypti* have demonstrated significant genetic differences in global populations during the past 50 years. *Aedes aegypti* populations exhibited a wide variation in morphology, vector competence, and genetic differences (Tabachnick, 1991b; Tabachnick & Powell,1979; Wallis *et al.,* 1985). The population genetic structure of *Ae. aegypti* has been studied using a variety of genetic markers, such as allozymes (Tabachnick & Powell, 1979; Tabachnick *et al.,* 1985; Tabachnick, 1991b), single nucleotide polymorphism (Herrera et al., 2008; Kotsakiozi *et al.,* 2018; Lee *et al.,* 2019), nuclear genes (Brown *et al.,* 2014; Fernando *et al.,* 2020a), mitochondrial DNA including the cytochrome oxidase subunit 1 (COI) gene (Paupy *et al.,* 2008; Naim *et al.,* 2020; Paupy *et al.,* 2012; Bennett *et al.,* 2016; Cook *et al.,* 2005; Calvez *et al.*, 2016) and nicotinamide adenine dinucleotide (NAD) dehydrogenase subunit 4 (Calvez *et al.,* 2016; Fernando *et al.,* 2020a; Moore *et al.,* 2013b; Paupy *et al.,* 2008, 2012), and microsatellite DNA (Calvez *et al.,* 2016; Lee *et al.,* 2019; Paupy *et al.,* 2008, 2012; Shi *et al.,* 2017).

Aedes aegypti genetic variations have been characterized since 1979 (Tabachnick & Powell, 1979), the authors indicated that *Ae. aegypti* mosquito is a widely distributed mosquito associated with human demonstrating variations in their morphology and behaviour, this makes evolutionary genetic studies in *Ae. aegypti* needed to complete the picture. Another genetic variation study using 11 isozyme loci in *Ae. aegypti* from eight locations from Africa and other continents showed that the sylvan west African populations is genetically divergent from the other different geographical locations (Tabachnick *et al.,* 1985).

Genetic diversity study surveyed the genetic variation of sylvian and domestic populations of *Ae. aegypti* in Cameroon using both mitochondrial and microsatellite markers revealed that reduced gene flow between the sylvan and domestic populations (Paupy *et al.,* 2008). A study using mosquitoes collected from 13 countries across five continents and clustering them using the Bayesian method using microsatellite loci also supports the existence of two divergent groups in *Ae. aegypti*, and the high genetic difference between tropical non-African populations and African populations corresponds to previous descriptions of subspecies *Ae. aegypti* (Brown *et al.,* 2011).

A global assessment of the genetic structure of seventy-nine populations from thirty countries across six continents using twelve microsatellite markers, the study indicated that globally *Ae. aegypti* divided genetically into two units (for most of the globe) matching with the old, divided subspecies of *Ae. aegypti*, the *Ae. aegypti aegypti* and *Ae. aegypti formosus* (Gloria-Soria *et al.,* 2016). Another global study in mitochondrial genes *ND4* revealed that, *Ae. aegypti* around the globe had dual origins in African (Moore *et al.,* 2013).

2.4.1 Mitochondrial markers

Mitochondrial DNA, the maternally inherited DNA with no introns, which found in many copies made it easy to amplify, with no recombination characterized resulted in only one demographic history rapid divergence, and high polymorphism rate and divergence levels due to relatively increased mutation rate. This is a reason for its power as genetic markers for evolutionary relationships (Cook *et al.,* 2005; Morlais & Severson, 2002; Naima *et al.,* 2020; Fernando *et al.,* 2020a).

Mitochondrial markers had been expansively used through periods in mosquito's phylogeny and genetic diversity studies (Costa-da-Silva *et al.,* 2005; Fernando *et al.,* 2020b; Gloria-Soria *et al.,* 2016; Joyce *et al.,* 2018; Moore *et al.,* 2013; Paupy *et al.,* 2008, 2012; Sousa *et al.,* 2017). Mitochondrial DNA has been proved as a genetic marker evaluating species gene flow, commonly has been used in population genetic studies, particularly in *Ae. aegypti* in different geographical populations (Twerdochlib *et al.,* 2012; Joyce *et al.,* 2018; Naim *et al.,* 2020; Elnour *et al.,* 2020; Moore *et al.,* 2013).

2.4.2 Cytochrome oxidase subunit 1 (*CO1*)

Cytochrome oxidase subunit 1 (*CO1*) is a mitochondrial gene in the respiratory chain terminal catalyst contributory across the cell membrane in proton translocation and electron transport. The *CO1* is the biggest among the three mitochondrial-encoded cytochrome oxidase subunits. *CO1* protein sequence contains highly conserved functional domains and variable regions, which makes it preferably used mitochondrial marker in evolutionary studies. *Aedes aegypti* cytochrome oxidase subunit 1 is 1537 base-pair length coding for 512 amino acids (Morlais & Severson, 2002).

The DNA barcoding technique that uses the cytochrome oxidase subunit 1 mitochondrial gene (*CO1*) is extremely effective in distinguishing between vertebrate and invertebrate species (Morlais & Severson, 2002). The *CO1* was widely and commonly used as a mitochondrial marker in *Aedes* mosquito genetic variations studies, many recent studies have been used in *Ae. aegypti* in different countries (Elnour *et al.,* 2020; Naim *et al.,* 2020; Salgueiro *et al.,* 2019; Joyce *et al.,* 2018). A recent study analysed mitochondrial genetic diversity and phylogeographic relationships of *Ae. aegypti* mosquitoes collected from seven districts in Sri Lanka. The study revealed a high genetic diversity for *CO1* marker which might be due to high levels of gene flow observed in the study between Sri Lanka and Southeast Asia (Fernando *et al.,* 2020a).

Research study of genetic diversity of *Ae. aegypti* mosquito vector in El Salvador, Central America used (mtDNA) cytochrome c oxidase subunit 1, the mitochondrial DNA sequences were also discovered to have a high level of genetic variation (Joyce *et al.,* 2018). Another study in Penang, Malaysia characterized the *Ae. aegypti* and *Aedes albopictus* population structure and genetic diversity using cytochrome oxidase I (Naim *et al.,* 2020). Elnour and his research group (2020) studied the population genetic diversity of *Ae. aegypti* from Red Sea State, Sudan, using *CO1* mitochondrial gene, they found only four haplotypes.

2.4.3 NADH dehydrogenase 4 (*ND4*)

The mitochondrial gene nicotinamide adenine dinucleotide dehydrogenase (NADH) subunit 4 (*ND4*) gene controlled the production of NADH dehydrogenase 4 protein, the active mitochondrial protein including in complex I, a large enzyme complex (Kühlbrandt, 2015). The electrons went from ubiquinone through many other

enzyme complexes to provide energy for ATP production (Kühlbrandt, 2015). The *ND4* (NADH dehydrogenase subunit 4) mitochondrial gene marker have been widely used for genetic variations studies in *Ae. aegypti* around the globe in Brazil (Fraga *et al.,* 2013; Sousa *et al.,* 2017), Peru (Costa-da-Silva *et al.,* 2005), Bolivia (Paupy *et al.,* 2012), Senegal, west Africa and Kenya (Moore *et al.,* 2013), Saudi Arabia (Ali *et al.,* 2016) and a global study by (Fernando *et al.,* 2020b). Researchers studied the genetic polymorphism between sylvan and domestic *Ae. aegypti* populations from in Cameroon using mtDNA-*ND4*, revealed that the two populations are structured by low gene flow between them (Paupy *et al.,* 2008).

2.4.4 Microsatellite markers in mosquitoes' population genetics

Microsatellites are repeated tandem copies of simple sequence repeats (2-6 bp) that dispersed the whole genome, the repetitive sequences were conserved but their length differs. The microsatellites are characterized by a copiousness in genomes, high variableness degree in the repeat sequence, and reproducibility (Vieira *et al.,* 2016). Microsatellites are universally used in both conservation genetics and population genetic studies with the unique developed fluorescent microsatellite primers for PCR amplification. Microsatellites are frequently greatly polymorphic, with many alleles separating in different populations. Microsatellites are widely distributed in the chromosomes and could have multiple loci or single in different individuals which can be specific to an individual insect (Ellegren, 2004).

Microsatellite can be interpreted simply like Mendelian variation; their variation is known to be caused by mismatch errors directing to crossover within repeats. In that case, when dissimilar crossover happens, one allele increases the repeat units while the other allele decreased repeat units. Mostly alleles modify into same