Egyptian Journal of Aquatic Biology & Fisheries
Zoology Department, Faculty of Science,
Ain Shams University, Cairo, Egypt.
ISSN 1110 – 6131
Vol. 29(4): 3459 – 3474 (2025)
www.ejabf.journals.ekb.eg

Mapping Tool of the Potential Restriction Sites for Field Identification of Major Mosquito Vectors in Aquatic Ecosystems

Aya Youssef Eissa Ahmed[1], Mona Gaber Shaalan[1]*, Amany Soliman Khaled[1], Enas Hamdy Ghallab[1]
Entomology Department, Faculty of Science, Ain Shams University, Cairo, Egypt. 11566

*Corresponding Author: mona.gaber@sci.asu.edu.eg

ARTICLE INFO
Article History:
Received: June 20, 2025
Accepted: Aug. 12, 2025
Online: Aug. 31, 2025

Keywords:
Mosquito vectors,
Restriction enzyme,
NEBcutter,
Molecular markers

ABSTRACT
Advances in genomic technologies have driven the development of molecular tools for mosquito species identification—an essential part of disease vector surveillance—especially during their aquatic life stages, where morphologically similar species can differ greatly in their role as vectors. Traditional genomic approaches, though accurate, are often labor-intensive, time-consuming, and costly, posing challenges for large-scale surveillance in resource-limited settings. This study proposes an automated, restriction enzyme–based method to distinguish three medically significant mosquito species: *Anopheles gambiae*, *Aedes aegypti*, and *Culex quinquefasciatus*. This focus is critically important as GST-mediated metabolic resistance to insecticides has been linked to increased *Plasmodium* infection in *Anopheles* mosquitoes, raising serious concerns for malaria epidemiology in sub-Saharan Africa. By targeting polymorphic regions within three glutathione S-transferase genes (*Gste1*, *Gste3*, *Gstt2*), species-specific restriction profiles are generated using three selected enzymes. Experimental validation demonstrated a classification accuracy exceeding 99%. The method enables simultaneous discrimination among species and improves reliability by minimizing false negatives that could arise from missing sequences. Unlike sequencing, this workflow is scalable, cost-effective, and suitable for rapid, field-friendly deployment. While this study focuses on three species, the method is adaptable to others with known GST polymorphisms. Overall, restriction enzyme mapping offers a robust, low-cost alternative to sequencing, with the potential to transform mosquito vector identification and enhance public health interventions.

INTRODUCTION

Distinguishing between these species in their aquatic form is a real challenge in the field, which makes it essential to use precise molecular methods for accurate identification before they emerge as adults.

Mosquitoes are responsible for the greatest number of human deaths globally each year. Mosquito vectors are capable of harboring or transmitting at least one type of mosquito-borne disease (**Harbach & Kitching, 1998**). Although we often associate

mosquitoes with their buzzing presence above ground, their life actually begins in water. These early stages are critical, especially because many mosquito species that look nearly identical in water can play very different roles when it comes to disease transmission. Anopheline mosquitoes transmit malaria, a parasitic infection that causes an estimated 219 million cases and over 400,000 deaths annually. Dengue is the most prevalent viral infection transmitted by Aedes mosquitoes. Other mosquito-borne viral diseases include Zika virus fever, yellow fever, West Nile fever, and Japanese encephalitis **(Gubler, 1996; Gubler, 2004)**.

Disease transmission is species-specific—each mosquito species transmits only particular pathogens it can acquire, support, and retransmit **(Thompson *et al*., 1996; Bogh *et al*., 1998)**. Distinguishing between these species in their aquatic form is a real challenge in the field, which makes it essential to use precise molecular methods for accurate identification before they emerge as adults.

Although mosquito monitoring programs have been widely developed worldwide, current procedures suffer from significant limitations. For example, detecting pathogens of mosquito-borne diseases may require several days. Moreover, one major limitation is the labor-intensive and time-consuming process of morphological classification, which relies on trained taxonomists using visual keys based on distinct morphological features **(Ree, 2003)**. However, the number of trained taxonomists is declining significantly **(Martineau *et al*., 2017)**, highlighting the urgent need for alternative automated methods with expert-level classification accuracy.

For a long time, mosquito identification has relied on morphological characteristics of both immature and adult stages **(Taylor & Provart, 2006).** However, variations in these traits, along with sample damage during preservation, often make accurate identification difficult. As a result, there have been considerable efforts to develop surveillance systems for early detection and diagnosis of mosquito-borne diseases. Among available methods, molecular analysis has proven to be the most reliable approach for identifying mosquitoes and other insects **(Hesson *et al*., 2010).**

Researchers have increasingly focused on molecular markers, which are based on nucleotide sequence mutations within the genome **(Aivazi and Vijayan, 2010).** These markers offer high reliability in species classification. By using molecular markers and restriction enzyme recognition sites, it is possible to conduct phylogenetic analyses and classify mosquitoes with high accuracy.

In this study, we aimed to demonstrate the effectiveness of restriction map analysis as a molecular marker for identifying major mosquito vectors. Genome mapping—characterizing a genome in terms of gene positions and DNA sequence organization—was used to validate the restriction mapping approach.

Molecular markers such as restriction fragment length polymorphism (RFLP), microsatellites or simple sequence repeats (SSRs), variable number tandem repeats (VNTRs), amplified fragment length polymorphism (AFLP), and random amplified polymorphic DNA (RAPD) have been widely used for genetic mapping **(Shu *et al.*, 2010)**.

To generate restriction maps, computational tools are employed to analyze gathered data. Among these tools, NEBcutter is considered the most comprehensive for restriction site analysis, offering features ranging from cloning analysis to gel electrophoresis predictions **(Gaudet *et al.*, 2007).** NEBcutter version 2.0 can process clusters smaller than 300 kb or sequence files under 1 MB in size using data from the REBASE database.

In this study, we focused on glutathione S-transferase (GST) genes as target domains. According to recent findings **(Tchouakui *et al.*, 2019)**, increased Plasmodium infection in insecticide-resistant *Anopheles* mosquitoes is linked to GST-mediated metabolic resistance. In *Anopheles gambiae*, six GST classes have been identified: Delta, Epsilon, Theta, Omega, Zeta, and Sigma **(Ranson & Hemingway, 2005).**

This study aimed to computationally investigate insecticide resistance genes across three major mosquito genera (*Anopheles*, *Aedes*, and *Culex*) using publicly available bioinformatics databases and tools. The goal was to compare sequence homology and restriction enzyme patterns to identify potential molecular markers for species differentiation.

MATERIALS AND METHODS

1. Gene selection via literature mining

Identification of glutathione S-transferases (GST) in mosquito vectors

We focused on metabolic resistance genes in three major genera: *Anopheles, Aedes*, and *Culex*. A curated list of 23 GST genes associated with metabolic resistance mechanisms was assembled. These genes represent six functional groups: delta (n=8), epsilon (n=8), theta (n=2), omega (n=1), zeta (n=1), and other miscellaneous detoxification-related genes (n=3). Gene selection was based on peer-reviewed literature obtained from:
Gene information was retrieved primarily from NCBI (https://www.ncbi.nlm.nih.gov/ and VECTORBASE (https://vectorbase.org/), which provide access to curated genomic and resistance-related data for vector species.

2. Retrieval of gene sequences from GenBank

Full-length nucleotide sequences were retrieved in FASTA format from the NCBI GenBank database, Search terms combined gene symbols with genus names (e.g., GSTE3

AND *Anopheles*). Only complete coding sequences (CDS) were included for downstream analysis. The record was accessed through GenBank, then "FASTA" format option was selected. The complete nucleotide sequence was copied and saved as a .txt file for downstream analysis.

3. Homologue grouping and dataset preparation

For each gene, a triplet of homologous sequences—one from each mosquito genus—was compiled into individual FASTA files. This resulted in 23 unique sequence sets representing orthologous genes.

4. Multiple sequence alignment and variation assessment

Each triplet set was aligned using ClustalW (https://npsa-prabi.ibcp.fr/cgi-bin/align_clustalw.pl). Default parameters were used to detect:

Conserved regions: indicative of evolutionary constraint. Variable regions: suggestive of species-specific adaptations.

5. *In silico* restriction enzyme mapping

To visualize restriction enzyme cutting patterns, restriction maps were generated for each gene using NEBcutter v2.0 (https://www.labtools.us/nebcutter-v2-0/).
Steps:
- Each FASTA sequence was submitted to the tool.
- Maps were generated showing recognition sites for a variety of restriction enzymes.
- Sites affected by methylation (e.g., CpG sites) were excluded from further consideration..

6. Cross-species comparison of restriction patterns

The restriction maps were compared for each gene across the three mosquito species:
- Enzymes common to all three sequences were identified.
- Enzymes unique to one or two species were catalogued.
- Variations in restriction profiles were analyzed to assess sequence divergence and conserved motifs.

7. Identification of potential molecular markers

Based on alignment and restriction map differences:
- Genes exhibiting significant variation were prioritized as candidates for molecular diagnostics:

Differential restriction sites were evaluated for potential PCR-RFLP assay development.

The presence of conserved sites flanking variable regions was considered advantageous for assay design.

RESULTS

In this study, gene homologues related to insecticide resistance were analyzed across three major mosquito genera: *Anopheles, Culex*, and *Aedes*. Genes were categorized into functional groups, and homologues were compared among the species. Genes with no clear homologue or excessive divergence were excluded to focus on 20 representative genes with confirmed cross-species homology.

1. Homologue distribution across mosquito genera

1.1 Delta group (n = 8): *Culex* exhibited 14 delta genes, *Anopheles* 12, and *Aedes* 8 homologues.
1.2 Epsilon group (n = 8): *Culex* had 10, while *Anopheles* and *Aedes* each had 8 homologues.
1.3 Theta group (n = 2): *Culex* had 6, *Anopheles* 4, and *Aedes* 2 genes.
1.4 Omega group (n = 1): All three genera contained a single omega gene homologue.
1.5 Sigma group (n = 1): One gene was found in each of the three genera.
1.6 Zeta group (n = 0): *Anopheles* and *Aedes* contained one zeta gene each; no homologue was found in *Culex*.

2. Restriction mapping analysis

The selected 20 genes were further analyzed using three categories of *in silico* restriction enzyme markers:
- One-cutter enzymes
- Two-cutter enzymes
- Three-cutter enzymes

2.1 One-cutter enzyme analysis

Over 200 restriction sites were identified across the 20 gene sequences for all three mosquito species using one-cutter enzymes. Due to the high level of shared recognition patterns and lack of discriminatory potential, this category was deemed unsuitable for differential diagnostic assay development.

2.2 Two- and three-cutter enzyme analysis

In contrast, more than 90 distinct restriction sites were identified using two- and three-cutter enzymes. These maps revealed several genus-specific differences. For example:

- Gene: Gste6 (Epsilon group)
- NCBI Accession Numbers: *Anopheles* (AY070256), *Culex* (Gene ID: 6052516), *Aedes* (Gene ID: 5569858) (Fig. 1)

- Enzyme: **Bts1Mut1**, present in all three species (Fig. 2), were considered non-discriminatory and excluded from further diagnostic consideration.

- Recognition site: CAGTGNN, Yields three bands on 0.7% agarose gel (Fig. 3)

```
>AY070256.1 Anopheles gambiae glutathione S-transferase (GSTe6)
TCGAGCAAGCCGGTCCTGTACACGCACACGATTAGTCCCGCCGGCCGTGCGGTCGAGCTGACCGTGAAGGCGTTGAACCTTGACGTCGATGTTCGCGAGATGAACGTCTTCAAGGGTCAGCATATGAGCGAC
GAGTTCAAGAAGCTAAACCCCGTCCAAACGATTCCAACGCTGGACGACAACGGGTTCGTGCTGTGGGATAGCCACGCCATCATGATCTATCTGGCGCGCCGTTACGGTGCCGACTCCGGCCTCTACACGGAC
GAGTACGAGCAGCAGGCCCGCATCAATGCGGCCCTCTTCTTCGAGAGTTCGATCCTGTTCGCGCGGCTCCGCTTCTGCACGGACAATCTGACCGTGCTGGGCAAGAGTGCGATACCGGAGGAGAACCTGCAG
CGTGCGCTGGAAGGGCTGCAGCGGCTGGAGAGGATGCTACAGTCGGAGTATGTGGCCGGCGATCAGCTGACCATTGCGGATCTGAGCTGCGTGAGCAGTGTGGCCACACTGCACCTGATGCTGAAACCGTCG
GCGGAAGAGTTCCCCAAAACGTTCGCCTGGATGGACCGGTTGTCGAAGTTGCCGTACTACGGGGAGGTGATGGGACGGGGGCTTAAAGCGGCCGGAGAGCTGATGCAGACGCTTGGCAGCAAGAACAG

>CPIJ018632 Culex pipiens glutathione S-transferase (GSTe6)

ATGGGCAAAATCGTGCTGTACACCGCCAAGCTGAGTCCTCCGGGACGTTCCGTTGAGCTGACCGCCAAGGCTTTGGGCTTGGACCTGGACATCAAGCCGATCAACCTGATTGCTGGAGATCACCTGAAGCCGG
AATTCGTGAAGCTGAACCCCCAGCACACAATTCCGACGATCATCGATCAAGACGGAACGGTGGTGTACGACAGCCACGCGATCAACGTTTACTTGGCTACGAAGTACGCAAAGGACGACAGTCTGTATCCAGA
GGATGTGGCTACCCGGGCCAAGGTTAACGCTGCCCTGCACTTCGATTCTGGGGTTCTGTTCGCCCGACTGCGTTTCTACTTGGAACCCATTCTGTACTACGGATCAGGCGAAGCCCCCCAGGACAAGATCGAG
TACCTGTACAAGGCGTACCAACTGCTGAACGATACCCTGGTCAGCGATTACATAGTTGGCAACAAGCTCACCTTGGCAGATCTGAGCTGCATTGCCTCGATCTCCTCGATCCATGCCATTTTCCCGATCGATG
CGGCCAAGTACCCGCAACTGGCCGCTTGGGTGGCGCGACTTGCCAAGTTGCCCTACTACAAGGCCACCAACCAGGAGGGCGCCGAAGAGCTGGCTGCTCTCTATCGTGCCAAGCTGGCGGAGAATCG

>AAEL007946 aedes aegptyi glutathione S-transferase (GSTe6)
AGTCCCCGAATCAAGCTGTACACCAATCCCATCAGTCCTCCGGGACGATCGGTAGAACTGACGGCCAAAGCAATTGACCTTGACATTGAAGTCATCGCGATTGATCTGCTGGGAAATGATACCCTAAAACCGGA
CTATCTCCTCAAGAACCCACAGCACACGATTCCAATGATCGACGATGGCGGTAAGTTCATTTGGGACAGCCAGGCGATCAACGTGTACCTCACAACCGTCTACTCACGCAACGAAGATCTATATCCGAACGATC
CTTTCGTGCGCGCCAAGGTGAACGCCGGTCTGCATTTCAACTCCGGAGTGCTCTTCAGCCGTCTGAAACTGTTGATCTCCCCGGTGATTCGTGGATTCAAGCAGGACCTGGACCCGGAAAAGGTGGAATACTTT
AACATTGGGCTGCAGCTGCTCGAGGATACGCTGCACGCTGATTACTACATTGGCAATCGGATGACCTTGGCCGATTTGAGTTGCGTTTCGTCGGTGTCCAGCTTCGATGCGGTTCTGCCGATAAGTCATGAAAG
ATTTCCAAAAACCGTCGACTGGTTGAGACGGATGGAACAATTGCCGTACTATGGGGAAGCGAATGGAGAAGGTGCAAAGAAATTGGCGAAGGTTGTACAGAGCTTTCTGAAATAA
```

Fig. 1. *Anopheles gambiae, Ae. aegypti, Cx. quinquefasciatus* glutathione S-transferases (GST) from epsilon group GSTe6 aligned sequence genes

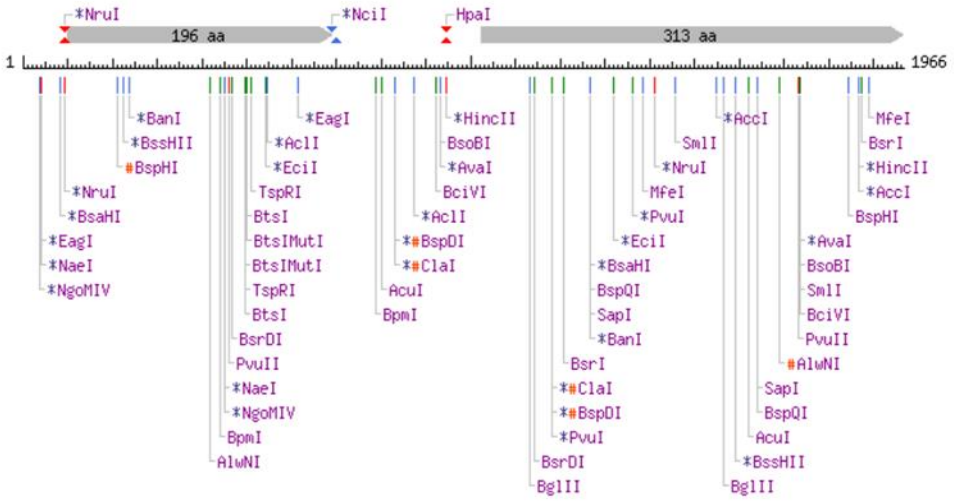

Fig. 2. Screenshot of the restriction map showing enzyme Bts1Mut1 from the two cutter group cuts GSTe6 genes twice

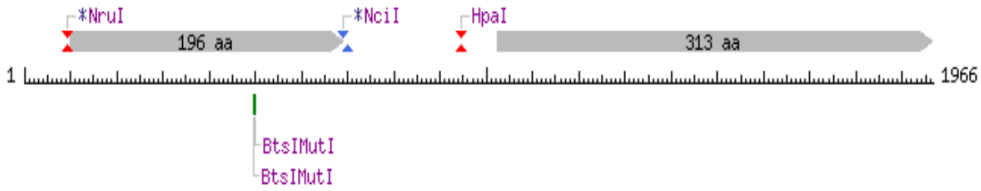

Fig. 3. Screenshot of the digested segment with BSPCN1 enzymes from the two cutter group

3. Identification of diagnostic markers

After filtering, three enzymes demonstrated genus-specific restriction patterns, suggesting their utility in species differentiation (Table 1):

Table 1. A list of glutathione S-transferase (GST) genes used as templates for diagnostic restriction enzyme marker development in mosquito genus differentiation

Gene Class	Gene name	NCBI ID		
		Anopheles gambiae	*Culex quinquefas..*	*Aedes aegypti*
Delta	GSTd1	**AF071160**	**AB443867**	**5568355**
	GSTd2	**Z71480**	**6033854**	**5568354**
	GSTd3	**AF513638**	**6033866**	**5568347**
	GSTd4	**AF513635**	**6033865**	**5568346**
	GSTd5	**AF513634**	**6033869**	**5568353**
	GSTd6	**AF513636**	**6033870**	**5573540**
	GSTd7	**AF071161**	**6033850**	**5568358**
	GSTd8	**AF071160**	**6031025**	**5810667**
Epsilon	GSTe1	**AF316635**	**6052511**	**5569844**
	GSTe2	**AF316636**	**6052510**	**11067855**
	GSTe3	**AY070234**	**6052512**	**5569859**
	GSTe4	**AY070254**	**6052513**	**5569847**

	GSTe5	**AY070255**	**6052515**	**5569856**
	GSTe6	**AY070256**	**6052516**	**5810639**
	GSTe7	**AF491816**	**6052508**	**5569858**
	GSTe8	**AY070257**	**6052514**	**5569853**
Theta	GSTt1	**AF515526**	**6047042**	**5571386**
	GSTt2	**AF515525**	**6047043**	**5571385**
Omega	GSTo1	**AY255856**	**6030837**	**23687505**
Sigma	GSTs1	**AF513639**	**6037723**	**5575278**

3.1 Marker 1 – BSPCN1 (for Anopheles gambiae)

- Target Gene: Gste1 (Epsilon group)
- NCBI Accession: AF316635, Alignment of the Gste1 gene from Epsilon group in *Anopheles*, *Culex*, *Aedes* mosquitos shown in Fig. (4).
- Enzyme: BSPCN1, from the three cutter group cuts (Fig. 5), Recognition site: CAGTG▲NN▼ Specific to *Anopheles*, produces 4 bands in 0.7% agarose gel (Fig. 6)

```
>AF316635.1 Anopheles gambiae glutathione S-transferase
ATGCCGAAACCTGTGCTATACACGGTTCATCTGAGCCCACCGTGCCGTGCCGTGGAGCTGACGGCAAAAGCGCTTGGATTGGAGTTAGAGCGCAAGCTTGTGAACTTACTTGCTGGGGAAAATTTAAC
ACCTGAGTTCTTGAAGCTCAATCCTAAGCATACGATCCCCGTGCTGGACGATAATGGGACGATCATCAGCGAAAGCCACGCGATCATGATCTATCTCGTGCGTAAGTACGGCCAGGGCGAAGGGAAGG
ATGCGCTGTACCCAACGGACATTGTCGAGCAGGCTCGGGTCAATGAGGCACTGCACTTCGAGTCCGGTGTGCTGTTTGCTCGGTTGCGATTCATTACCGAGCTGGTATTTTTCGCACGCAAACCAGAA
ATTCCGGAAGATCGCATCGAGTACGTTCGTACGGCGTACCGGTTGCTGGAAGACTCGCTGCAGAGCGATTACGTAGCCGGATCACGCATGACGATCGCTGACCTGAGCTGCATTTCCTCGGTCGCCTC
TATGGTGGGCTTTATCCCGATGGAAAGGTCCGAGTTCCCGCGGGTGCACGGGTGGATCGAGCGGATGAAGCAGTTGCCGTACTATGAGGAGATCAACGGTGCCGGTGCCACAGAGCTGGCCGAGTTCA
TTGTGGATATGTTGGCGAAGAATGCAAAACTGTAA

>CPIJ018628 culex quinqfitatus  glutathione S-transferase
ATGTCACGCCCGATTCTGTACACGCTCCACCTGAGCCCTCCCTGCCGGGCCGTCGAACTGACGGCCAAGGCGCTGGGCCTCGATTTGGAGCAGAAGGTGGTCAATCTGCTCACCGGGGACCACCTGAA
GCCGGATTATCTCAAGATGAACCCCCAGCATACGATTCCGGTGCTGGACGATAACGGGACCATCATTCCGGAGAGTCACGCCATCATGATCTACCTGGTGTCCAAGTACGGCAAGGACGACAGCCTGT
ACTCGAAGGAGATCGTCAAACAGGCCAAGCTGAACGCGGCGCTGCACTTTGAGTCGGGCGTTCTGTTTGCGCGGATGCGGTTCTTGTTTGAACCGATTCTGTTCCGAGGTAGCAGTGAAATGCCAGCC
GATCGCACGGAGTACATCGAAAAGGCTTATCAGCTGTTGGAAGACACCCTGGTTGATGACTACGTGGTTGGGAGCACCCCAACGATCGCCGATTTCAGCTGCATTTCTACGATCTCGTCGTTGATGGG
AGTTATCCCACTGGAGTCGACGAAGTATCCCAAAATTCTTGCCTGGGTGACCGTCTAAAGGCATTGCCGTACTACGAAGAAGCCAATGGAAGTGGTGCCATCCAGTTGTCATCAGCTGTTCTGGCTTC
CAAGGAGAAGAACGCGGCCAAGGCCTCGCTGTAA

>AAEL007954 Aedes aegypti  glutathione S-transferase
ATGGGCAAAATCAAACTCTACAGTTTCCCGCTGAGTCCACCTGGTCGAACGGTTCAGCTAACCGCGAAAGCCCTCGGCCTTGAGCTCGAATTTCATTCCGTCTCGGTGTTGGAAAAGGAGCACCTTAC
CGAGGAGTTCATCAAGATGAATCCGCAGCACACCATCCCGGTAATCGACGATGATGGGTTTGTGCTGTACGACAGTCATGCGATTGCCATCTACTTGGTATCGAAGTATGCCCCTGGAAATCGGCTCT
ATCCTACGAAGGACTTCAAGCAGCAGGCCCGTATCAACGCGATCTTGCACTTCGAGTCTGGAGTAATGTTCGCCAGGCTCCGATTCGTTGGAGATGCCATACAGAAGGCCAGTCACCAAGGTGAAGTT
CCACAGGATCGCGTGGAGTACGCACTGGAAGCCGTTGAGCTGCTGGAAGCCCTGCTGAGGGATGGACAGTATCTAGCTGGGGACCACGTGACATTGGGCGATATCAGTTGCGTGACTTCATTCTCGTT
TCTGGATGCTATGCTACCAGTGGAACGTGCCAAGTATCCGAAGGTTTACGCTTGGTATGAGCGGATGAAACACATCGAGGGATATGATGAGATCAACCAAAAAGCGGTGGATCAGTTGAACGGGGTCA
TTCAGGGGATTTTTGAGGGAAATAAGAGCAAATAG
```

Fig. 4. *Anopheles gambiae*, *Ae. aegypti*, *Cx. quinquefasciatus* glutathione S-transferases (GST) from Epsilon group GSTe1 aligned sequence genes

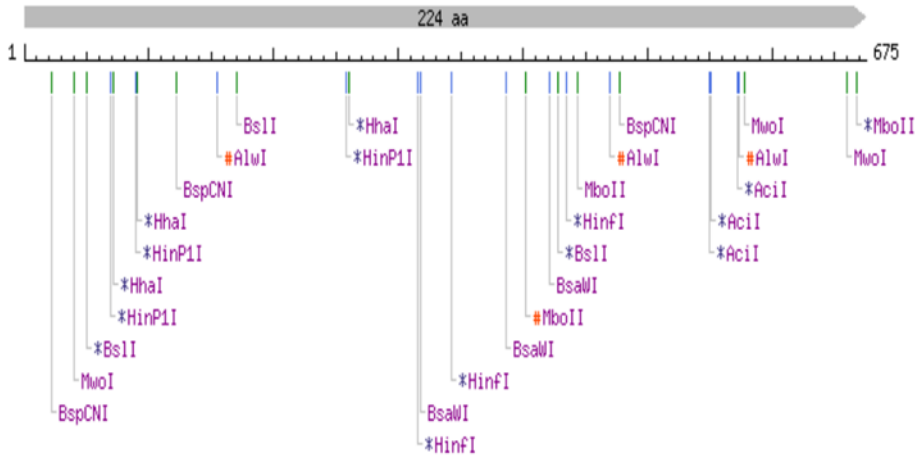

Fig. 5. Screenshot of the restriction map showing enzyme BSPCN1 from the three cutter group cuts GSTe1 gene in *An. gambiae* three times.

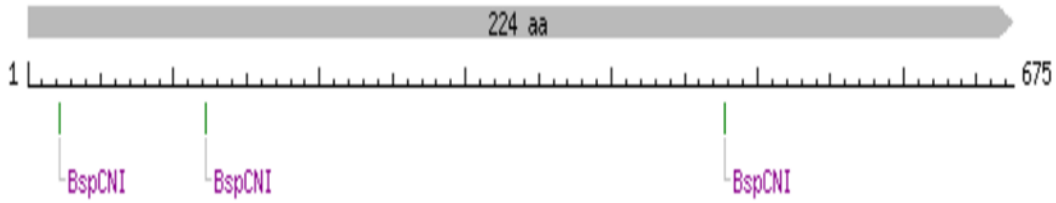

Fig. 6. Screenshot of the digested segment with BSPCN1 from the three cutter group cuts

3.2 Marker 2 – Bae1 (for Culex quinquefasciatus)

- Target Gene: Gste3 (Epsilon group)
- Gene ID: 6052512, Alignment of the Gste3 gene from Epsilon group in *Anopheles*, *Culex*, *Aedes* mosquitos showed in (Fig. 7).
- Enzyme: Bae1, from the three cutter group (Fig. 8), Recognition site: $_{▲}(N)_5{}^{▼}(N)_{10}ACNNNNGTAYC(N)_7{}_{▲}(N)_5$, specific to *Culex*, produces 4 bands in 0.7% agarose gel (Fig. 9)

```
>AY070234.1 Anopheles gambiae glutathione S-transferase (GSTe3) gene
ATGGCACCGATTGTGTTGTACAGTACCCGCCGCACGCCGGCCGGCCGTGCCGTGGAGCTGACGGCAAARATGATCGGTATCGAGCTGGATGTGCAGTACATTGATCTGGCGAAGAAGGAAAATA
TGACCGAAGAGTATTTGAAGGTAAGCGATCGCCTTCTTTTTTCGCGTCTTCTTCATCACTTTCACTGTGCATCTTCCATCCATCCGCAGATGAATCCGATGCACACCGTGCCGACGGTCAACGA
TAACGGTGTGCCGCTGTACGATAGCCATGCCATTATCAACTATTTGGTGCAGAAGTACGCAAAGGATGACACCCTTTACCCGGCGAAGGATCTAGTGAAGCAGGCCAACATCAACGCCCTGTTG
CACTTCGAGTCGGGCGTACTGTTCGCCCGGTTGCGCTGGATCCTGGAGCCGGTGTTCTACTGGGGCCAGACGGAGGTGCCACAGGAGAAGATCGACTCGGTACATAAAGCGTACGATCTGCTTG
AGGCTACGCTGAAGACATCTGGCACCGACTACCTCGTAGGCGGCACGATCACGCTTGCGGACATCTCCGTCAGCACATCGCTCTGCACGCTCAACGCGCTGTTCCCGGCCGATGCGAGCAAGTA
CCCGCTGGTGCTGGCCTACCTGAAGCGGCTGGAGCAAACGATGCCCCACTACCAGGAGATCAACACGGACCGGGCCAATGACGCGCTGCAGCTGTACAACCAGAAACTTGGCAAGGTTTAG

>CPIJ018628 culex quinquefasciatus glutathione S-transferase (GSTe3) gene
ATGGTGAACCCGATTCTCTACACCCTGCACCTGAGTCCTCCGTGCCGAGCGGTTGAACTTACCGCCAAGGCGTTGGGACTCCACCTTGATCAACGAGTGGTCAACCTGATGGCTGGAGATCACC
TGAAGCCGGAGTATTTGAAGATGAACCCTCGTCACACCGTCCCGGTTCTGGACGACAATGGTACCATTGTGTACGAAAGTCACGCCATTATGATCTATCTGGTGTCAAAGTACGGCAAGGACGA
CAGTCTGTACCCTAAGGATCTCGCCAAGCAGTCGAAGGTCAACGCAGCGTTGTTCTTCGAGTCGGGAGTTCTGTTCGCTCGCATGCGAACGATTACGTATCCGATCTTCCTGGAAGGTTGCCGC
GAGATTCCATCGGCGAAGGCAGAAGATGTTCGCAAGGCTTACCAACTGCTTGAGGATACCCTCGTGGAGAACTATGTTGCTGGAAGTTCGCTTACGATTGGTGACTTTAGCTGCATATCGACCA
TTTCTTCGGCTATGGGCTTCATTCCGATCGATCGGACCAAGTTTCCGAGGATCCACGGATGGATCAACCGGCTGAAGGAACTGCCGTATTACGAGGAGGCCAACGGTGGCGGTGCTACGCAGCT
ATCGACGGAAGTGCTGAAGTTGCTCGAGAAGAGG

>AAEL007947 Aedes aegypti glutathione S-transferase (GSTe3) gene
ATGTCCCCCATTACGTTGTACACTACTCGGCGCACACCGGCCGGCAGAGCTGTGGAAATCACCGCTAAGCTCATTGGCCTGGAACTGGACGTTAAGTTTATTGATTTGTCCAAGAAGGAACATC
TGACTGAGGAGTTTCTGAAGTTGAACCCACAACACACAGTGCCAACGATCGTTGATAATGGAGTCGCTCTGTACGATAGTCACGCAATCATCGTCTACCTGGTACAGAAATACGGCAAAGATGA
CGCCCTGTATCCGAAAGATCTGGTGACGCAAGCTCGTGTCAACGCATTGCTGCATTTCGAGTCGGGCATTTTGTTTGCTCGATTGAGAGGTACCTTGGAACCCATCTTCTACCATGGGTTCCCG
GAGGTTCCCCAGGAAAAGCTAACCGCCATCCACGGAGCGTACGATCTGCTAGAGGCCACCCTGAAATCAGACTATCTGGTTGGCGACTCGTTGACCCTTGCCGATGTCAGCTGCAGCACATCGC
TGTCCACTCTGAATGCCCTGTTCCCAATCGACGCCGAAAAATGCCCAAAATTGGTTGCGTATCTTCAGCGCTTAGAAGCGAACATGCCGAACTACAAGGAGCTGAACTCGGATCGTGCTGCCGA
AGCGGTAGCATTCTTCAAAGTCAAGCTGGAAGAAAACAAG
```

Fig. 7. *Anopheles gambiae*, *Ae. aegypti*, *Cx. quinquefasciatus* glutathione S-transferases (GST) from Epsilon group GSTe3 aligned sequence genes

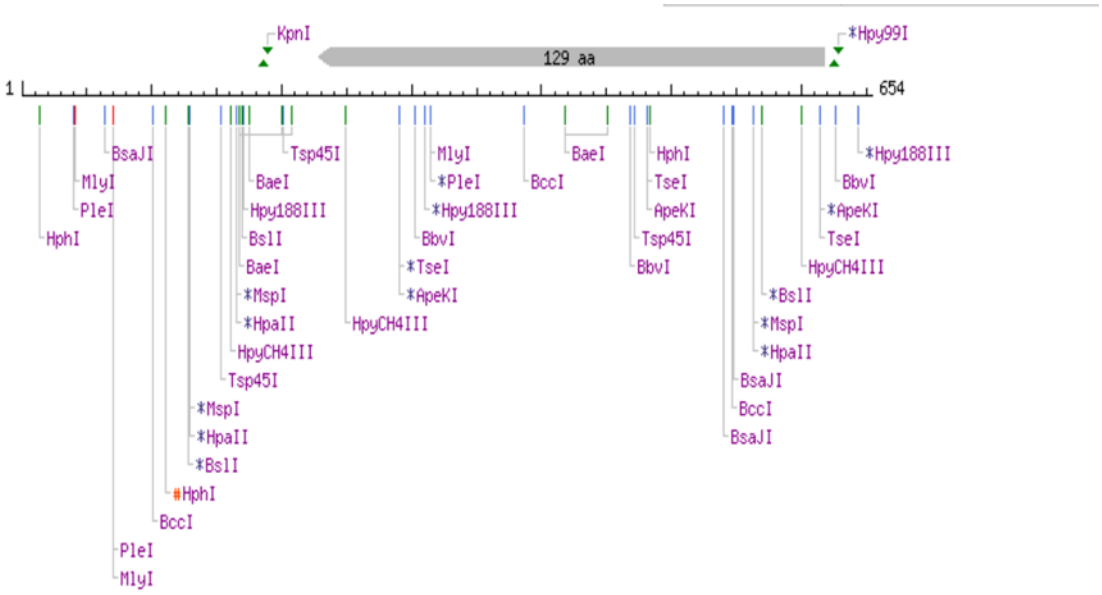

Fig. 8. Screenshot of the restriction map showing enzyme Bae1 from the three cutter group cuts GSTe3 genes three times

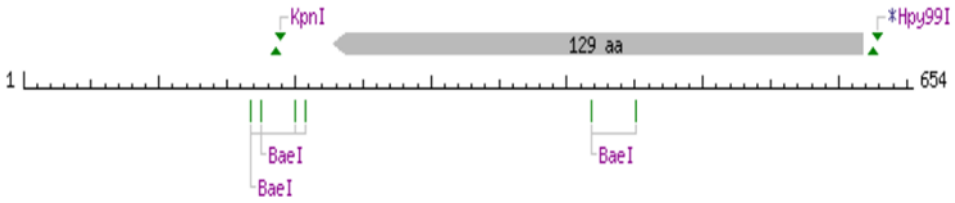

Fig. 9. Screenshot of the digested segment with Bae1 from the three cutter group cuts

3.3 Marker 3 – AF1111 (for *Aedes aegypti*)

- Target Gene: Gstt2 (Theta group)
- Gene ID: 5571385, Alignment of the Gstt2 gene from Epsilon group in *Anopheles*, *Culex*, *Aedes* mosquitos showed in (Fig. 10).
- Enzyme: AF1111, from the two cutter group (Fig. 11), Recognition site: A˅CRYG˄T, specific to *Aedes*, produces 4 bands on agarose gel (Fig. 12).

```
>AF515525.1 Anopheles gambiae glutathione S-transferase (gstT2)
ATGTCCCGTAGTGTTAAACTTTACTATGACCTCATGTCGCAACCGTCGAGGGCGCTTTACATATTCTTATCGACCAACAAGATCCCCTTTGACCGGTGCCCGATAGCG
CTGCGAAAAATGCAGCACAAGACGGACGAATACCGCCGGCAGGTGAATCGGTACGGCAAGGTGCCCTGCATCGTGGACGGCAGCTTCCGGCTAGCCGAGAGCGTCGCC
ATCTACCGGTACCTGTGCCGAGAGTTTCCGACCGACGGCCACTGGTACCCGAGCGACACGGTGCGCCAGGCACGCGTCGACGAGTACCTTTCCTGGCAGCACCTAAAC
CTGCGCGCCGACGTGTCGCTATACTTTTTCCACGTGTGGCTAAACCCGCTGCTGGGCAAGGAGCCGGACGCCGGCAAGACGGAACGGTTGCGCCGGCGGCTGGACGGT
GTGCTGAACTTTTTCGACCAGGAGCTGCTGTCGGCCGGTAGCGGGCAGGCGTTTCTGGCGGGCGATCGCATCAGCATTGCCGACCTATCGGCCGCTTGCGAGATTGAG
CAGGCGAAAATAGCCGGGTACGATCCGTGTGAGGGCCGGCCGGCACTGGCCAGCTGGCTGACGGCGGTGCGGGAGCGAACCAATCCGTACTACGACGAGGCGCACAAA
TACGTTTATCGGCTTTCGCCGGACCACATCGTAA

>CPIJ014053 Culex quinqfitatus glutathione S-transferase (gstT2)
GCCTCCCGGGCACTAAAATACTACTACGACCTGCTGTCCCAGCCGTCCCGCGCCCTCTACATCCTGCTGGAGCAGAGCAAGATTCCGTACGAAAAGTGTCCGGTGGCG
CTTCGAAAGTTGGAAAACCGCACCACCGAGTTCAAGGAGACCATCAACCGGTTCGGCAAGGTCCCGTGCATCGTGCACGGGGACTTCAAGCTGGCGGAGAATGTGGCC
ATTCTGCGATATTTGTCGCGGGAGTGTCCCGATCTGCCGGACTTTTGGTATCCGCGGGAAAATCTGGCTCGGGCCCGGGTTGACGAATATTTAGAGTGGCAGCACACG
ACGATTCGAGCCCAGTGCGCGCAGTACTTCCTGTACGTTTGGATGACGCCGCTGCTGGGGATGGAGCTGGACCCGGCGAAGGTCGAGCGGCTGCGGAAGAACATGGCC
GACGGGTTGGACCTGTTTCAGCGGGTGTGGTTGGAGGACGGGAAGAAGGCGTTTGTGGCGGGGGAGAGATTGAGCTACGCGGACGTGCTGGCAGCTTGCGAGATTGAA
CAGCCCAAGTGTGCGGGGTACGATCCGAGGGAGGGCCGGCCACAGTTGGCCGCGTGGATGGAACGGGTGCGGCAGGGGACGAATCCGTACTATGACGAGGCGCACAAG
ATTTTGTACAAGTTTACGCCGAAGGAAATCGCGA

>AAEL009016 Aedes aegypti glutathione S-transferase (gstT2)
AACGGTCGCAGCATTCGCTTCTACTACGACCTGATATCGCAACCTTGCCGAGCTTTGTACATTTTCCTTGAGCAGAACAAAATCCACTATCAAAAGTGTCCCATTGCG
CTGCGGAAATGGGAGCACACAACGCCCGAATATCTGCAAAATGTGAACCGTTTCGGTAAGGTGCCCGCCATTGTGGACGGAAAGAACTTCAAATTGGCGGAAAGTATC
GCCATTCTGCGATATTTGGCACGAGAGTTCACCGTACCGGACCATTGGTATCCGAGGGATAGCCGTAGGAGAGCCAGGGTTGACGAGTACTTGGAGTGGCAGCATTCG
AACACCCGGTTGCACTGCGCTGGGTATGTGAGGTACGTGTGGCGAGGACCTCTGCGAGGGGAAACCATGGACCCGAGGGTTGCCAAACGGTTGAAAGCGGAGATGGTC
GGTTGCTTGGATTTCATCGAGACGAACGTTTTACAGCGGGACGTGCATTTCATAGCAGGCGATGAAATCTCGATCGCTGATCTGGAGGCGGCTTGCGAGATCGAACAG
CCAAAACTAGCCGGGTACGACGCACGTGTCGGTCGACCGAAGCTGACAGCGTGGATGCAAAGGGTAAAGGAAACCACCCAACCAGACTACGACGAGGCGCATAAGGTT
TTGAATAAATTTGCGCCTACTGCGACGTGA
```

Fig. 10 *Anopheles gambiae, Ae. aegypti, Cx. quinquefasciatus* Glutathione S-transferases (GST) from Theta group GSTt2 aligned sequence genes.

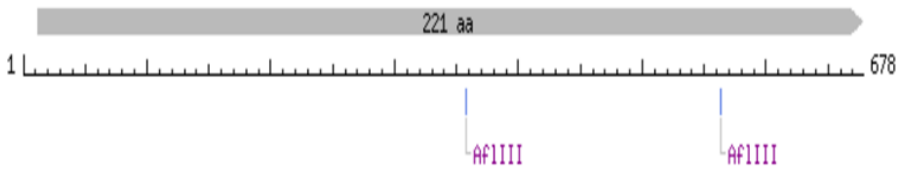

Fig. 11. Screenshot of the restriction map showing enzyme AF1 111 from the two cutter group cuts GSTt2 gene two times

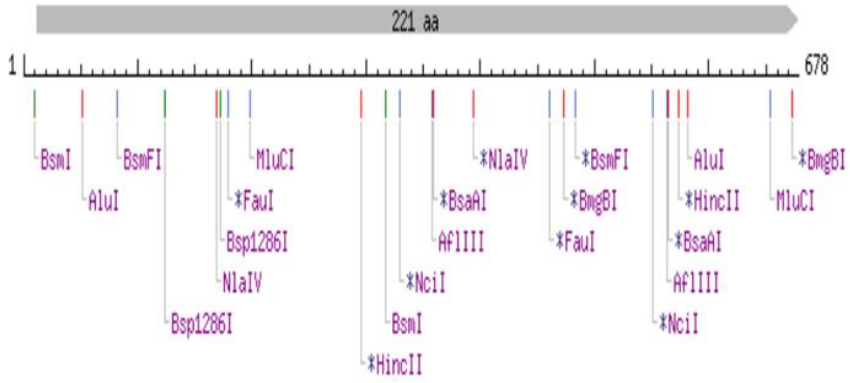

Fig. 12. *Anopheles gambiae, Ae. aegypti, Cx. quinquefasciatus* GSTt2 sequence digested with AF1 111 two- cutter marker

DISCUSSION

This study presents a molecular approach for the identification of three medically important mosquito species *Anopheles gambiae, Culex quinquefasciatus,* and *Aedes aegypti* based on species-specific variations in the glutathione S-transferase (GST) gene families, specifically **GSTe1**, **GSTe3**, and **GSTt2**, respectively. The application of restriction fragment length polymorphism (RFLP) using distinct restriction enzymes **(BSPCN1, Bae1, and AF1 111)** allows for the differentiation of these species through easily interpretable banding patterns in agarose gel electrophoresis.

The enzyme **BSPCN1**, categorized within the 3-cutter group, targets the sequence CTCAG(N)7NN, which occurs three times in *Anopheles gambiae,* yielding four distinct bands on a 0.7% agarose gel. Similarly, the enzyme **Bae1**, which recognizes the complex sequence (N)5(N)10ACNNNNGTAYC(N)7(N)5, is specific to *Culex quinquefasciatus* and generates four bands on a 0.07% agarose gel. In the case of *Aedes aegypti*, the 2-cutter enzyme **AF1 111**, which recognizes the sequence A$^{\blacktriangledown}$CRYG$_{\blacktriangle}$T, also produces four bands, corresponding to three restriction sites within the targeted sequence.

The presence or absence of these enzyme recognition sites serves as a reliable diagnostic feature for species identification. While the use of a single restriction enzyme may be sufficient due to the specificity of the target sequence, employing a combination of two or more enzymes enhances the accuracy and robustness of the identification process. This is particularly important in field settings, where degraded DNA or sequence mutations may obscure expected restriction patterns.

Previous literature has emphasized the necessity for accurate mosquito species identification as a cornerstone of effective vector control strategies **(Otranto *et al.*, 2009; Tipayamongkholgul *et al.*, 2009)**. Morphological methods, although traditional and widely used, often face limitations in distinguishing between closely related species,

particularly when morphological damage occurs during sample collection **(Apote, 1947; Batovska *et al.*, 2016)**. Moreover, conventional taxonomic keys require significant entomological expertise and may not be practical in large-scale epidemiological studies.

While several restriction analysis tools exist, many are limited in their capacity to process large sequences or handle high-throughput datasets **(Vincze, 2003; Gaudet *et al.* 2007)**. Tools such as NEBcutter, although useful, can process only short DNA sequences (<300 kb) and may not be suited for comprehensive analyses of multiple genetic targets. The approach proposed in this study leverages species-specific GST gene signatures and overcomes these limitations, providing a practical alternative for rapid, large-scale mosquito identification.

Our restriction enzyme mapping method offers a powerful upgrade to current mosquito surveillance systems by solving some of their most pressing challenges. First, it removes the heavy reliance on rare taxonomic experts. Using a simple, standardized molecular protocol with over 99% classification accuracy, field technicians—even without advanced entomology training—can confidently identify species. This drastically shortens response times: what once took days using sequencing or pathogen screening can now be done in just hours when paired with rapid DNA extraction kits. Faster identification means we can map vector hotspots in real time during outbreaks and act immediately.

It's also far more cost-effective—cutting expenses by over 80% compared to next-generation sequencing—because it avoids expensive equipment, specialized reagents, or advanced lab facilities. Instead, it uses only basic PCR and gel electrophoresis tools, which most district-level laboratories in resource-limited regions already have. Continuous development tools such as RFLP-kenzy **(Laref *et al.*, 2024)** demonstrated the value of *in silico* approaches for rapidly screening multiple enzymes and pinpointing those that cut at least one sequence but not all, thereby enabling clear separation of closely related taxa. This perspective supports our findings in highlighting the growing importance of *in silico* digestion tools for accurate and cost-effective mosquito vector identification.

One of its biggest strengths is the focus on aquatic life stages (larvae and pupae). This lets us target mosquitoes before they become adults, adding a proactive layer to control efforts and complementing traditional adult surveillance. Our multi-enzyme profiling system (BSPCN1 + Bae1 + AF1111) also improves reliability by reducing errors from incomplete digestion or natural DNA variations—critical for working with degraded samples from the field.

Finally, the method fits easily into existing surveillance networks. Results can be uploaded to platforms like VectorBase for resistance gene monitoring. In short, this

approach combines speed, affordability, and early intervention—making it a practical, scalable solution for mosquito-borne disease control.

CONCLUSION

This study introduces a reliable, cost-effective, and scalable molecular tool for the differentiation of three major mosquito vector *Anopheles gambiae, Culex quinquefasciatus,* and *Aedes aegypti* through the use of restriction enzymes **BSPCN1**, **Bae1**, and **AF1 111** targeting the GST gene family. The method is based on diagnostic restriction sites that produce clear and reproducible banding patterns in agarose gel electrophoresis. Although a single enzyme may be sufficient for identification, the combined use of two or more enzymes is recommended to increase diagnostic confidence and minimize errors due to sequence variability or incomplete digestion.

Given the global expansion of mosquito vectors and their critical role in the transmission of diseases such as malaria, dengue, and West Nile virus, the molecular technique described here offers a valuable tool for entomological surveillance, vector mapping, and integrated vector management programs.

This approach lays the groundwork for a scalable system that can be adapted to a wide range of mosquito and vector surveillance needs. To broaden its scope, future development could include other high-priority disease vectors such as *Anopheles stephensi*—an emerging urban malaria threat—and *Aedes albopictus*, a secondary carrier of several arboviruses. By pinpointing unique DNA variations in their glutathione S-transferase (GST) genes and drawing on curated genomic databases like VectorBase, we can design precise, species-specific detection tools. New restriction enzymes—such as BsmAI for the *An. stephensi* GSTe2 gene—could be identified and tested using our established computational workflow for sequence alignment and restriction site mapping.

ETHICAL APPROVAL

This research paper was approved by the research ethics committee from Faculty of Science, Ain Shams University (ASU-SCI/ENTO/2025/3/3).

REFERENCES

Aivazi, A. A. and Vijayan, V. (2010) Efficacy of Ruta graveolens extract and its synergistic effect with cypermethrin against *Anopheles stephensi* Liston larvae. Toxicol Environ Chem 92:893–901, DOI: 10.1080/02772240903092306

Apote, N. (1947) Pharmaceutical abstracts-section I. Bacteriology. (1948). Journal of the American Pharmaceutical Association (Scientific Ed.), 37, 2–70. . New Engl J Med 237:149, DOI: 10.1056/NEJM194707312370502

Batovska, J.; Blacket, M. J.; Brown, K. and Lynch SE (2016) Molecular identification of mosquitoes (Diptera: Culicidae) in southeastern Australia. Ecol Evol 6:3001–3011, DOI: 10.1002/ece3.2095

Bogh, C.; Pedersen, E. M.; Mukoko, D. A. and Ouma, J. H. (1998). Permethrin-impregnated bednet effects on resting and feeding behaviour of lymphatic filariasis vector mosquitoes in Kenya. Medical and Veterinary Entomology, 12(1), 52–59, DOI: 10.1046/j.1365-2915.1998.00091.x

Gaudet, M.; Fara, A. G.; Sabatti, M.; Kuzminsky, E. and Mugnozza, G. S. (2007). Single-reaction for SNP Genotyping on Agarose Gel by Allele-specific PCR in Black Poplar (*Populus nigra* L.). Plant Molecular Biology Reporter, 25(1-2), 1–9, DOI:10.1007/s11105-007-0003-6

Gubler, D. J . (2004). The changing epidemiology of yellow fever and dengue, 1900 to 2003: full circle? Comparative Immunology, Microbiology and Infectious Diseases, 27(5), 319–330, DOI: 10.1016/j.cimid.2004.03.013

Gubler, D. J. (1996). The global resurgence of arboviral diseases. Transactions of the Royal Society of Tropical Medicine and Hygiene, 90(5), 449–451, DOI: 10.1016/s0035-9203(96)90286-2

Harbach, R. E. and Kitching, I. J. (1998). Phylogeny and classification of the Culicidae (Diptera). Systematic Entomology, 23, 327e370, DOI: 10.1046/j.1365-3113.1998.00072.x

Hesson, J. C.; Lundström, J. O.; Halvarsson, P.; Erixon, P. and Collado, A. (2010). A sensitive and reliable restriction enzyme assay to distinguish between the mosquitoes *Culex torrentium* and *Culex pipiens*. Medical and Veterinary Entomology, 24(2), 142–149, DOI: 10.1111/j.1365-2915.2010.00871.x

Laref, N.; Belkheir, K.; Belazreg, M. and Hireche, A. (2024). RFLP-kenzy: a new bioinformatics tool for in silico detection of key restriction enzyme in RFLP technique, Beni-Suef University Journal of Basic and Applied Sciences (2024) 13:83, DOI: https://doi.org/10.1186/s43088-024-00531-8

Martineau, M.; Conte, D.; Raveaux, R.; Arnault, I.; Munier, D. and Venturini, G. (2017). A survey on image-based insect classification. Pattern Recognition, 65, 273–284, DOI:10.1016/j.patcog.2016.12.020

Otranto, D.; Capelli, G. and Genchi, C. (2009). Changing distribution patterns of canine vector borne diseases in Italy: leishmaniosis vs. dirofilariosis. Parasites & Vectors, 2(Suppl 1), S2, DOI: 10.1186/1756-3305-2-S1-S2

Ranson, H. and Hemingway, J. (2005). Mosquito Glutathione Transferases. Gluthione Transferases and Gamma-Glutamyl Transpeptidases, 226–241, DOI: 10.1016/S0076-6879(05)01014-1

Ree, H. I. (2003). Taxonomic Review and Revised Keys of the Korean Mosquitoes (Diptera: Culicidae). Entomological Research, 33(1), 39–52, DOI:10.1111/j.1748-5967.2003.tb00047.x

Shu, Y.; Li, Y.; Zhu, Z.; Bai, X.; Cai, H.; Ji, W. and Zhu, Y. (2010). SNPs discovery and CAPS marker conversion in soybean. Molecular Biology Reports, 38(3), 1841–1846, DOI: 10.1007/s11033-010-0300-2

Taylor, J. and Provart, N. J. (2006): CapsID: a web-based tool for developing parsimonious sets of CAPS molecular markers for genotyping. – BMC Genetics. 7 (1), 27, DOI: 10.1186/1471-2156-7-27

Tchouakui, M.; Chiang, M. C.; Ndo, C.; Kuicheu, C. K.; Amvongo-Adjia, N.; Wondji, M. J. and Wondji, C. S. (2019). A marker of glutathione S-transferase-mediated resistance to insecticides is associated with higher Plasmodium infection in the African malaria vector *Anopheles funestus*. Scientific Reports, 9(1), DOI: 10.1038/s41598-019-42015-1

Thompson, D.; Malone, J.; Harb, M.; Faris, R.; Huh, O.; Buck, A. and Kline, B. (1996) Bancroftian filariasis distribution and diurnal temperature differences in the southern Nile delta. Emerg Infect Dis 2:234–235, DOI: 10.3201/eid0203.960313

Tipayamongkholgul, M.; Fang, C. T.; Klinchan, S.; Liu, C. M. and King, C. C. (2009) Effects of the El Niño-southern oscillation on dengue epidemics in Thailand, 1996-2005. BMC Public Health 9:422, DOI: 10.1186/1471-2458-9-422

Vincze, T. (2003). NEBcutter: a program to cleave DNA with restriction enzymes. Nucleic Acids Research, 31(13), 3688–3691, DOI: 10.1093/nar/gkg526