Genetic characterization and molecular phylogeny of *Aedes albopictus* (Skuse) species from Sonitpur district of Assam, India based on COI and ITS1 genes

Momi Das1, Manuj K. Das2 & Prafulla Dutta3

1Department of Life Sciences, Dibrugarh University, Dibrugarh; 2Department of Bioengineering and Technology, Institute of Science and Technology, Gauhati University, Guwahati; 3Regional Medical Research Centre (Indian Council of Medical Research), Northeast Region, Dibrugarh, Assam

ABSTRACT

**Background & objectives:** *Aedes albopictus* (Skuse) is one of the major vectors of dengue which is an emerging threat in Northeast part of India. The morphological characterisation of mosquitoes is time consuming and lacks accuracy for distinguishing closely related species. Hence, molecular methods of mosquito identification, genetic diversity and molecular phylogeny have gained increased importance. This study was aimed to identify and characterize the most abundant species of *Aedes* vectors collected from different breeding spots in Assam, Northeast India employing molecular as well as bioinformatics tools.

**Methods:** *Ae. albopictus* species was genetically characterized with internal transcribed spacer1 (*ITS1*) and cytochrome *c* oxidase subunit I (*COI*) genes and sequence analysis was carried out following molecular methods like PCR amplification, DNA sequencing and multiple sequence analysis. Maximum likelihood molecular phylogeny was reconstructed to define the evolutionary relationship among studied isolates and isolates from other parts of Southeast Asia.

**Results:** Molecular study revealed that all five subject specimens belonged to *Ae. albopictus* species as per both *ITS1* and *COI* genes. Maximum likelihood tree based on *ITS1* and *COI* genes showed that isolates were distinctly grouped into separate clusters. Almost similar pattern of amino acid frequencies in *COI* gene was found amongst the five studied isolates. However, amino acid frequency in *ITS1* gene was found to be dissimilar, indicating polymorphisms in this gene, among the isolates.

**Interpretation & conclusion:** This is the first report among the Northeastern states of India describing the genetic make-up of *Ae. albopictus* species by virtue of highly conserved mitochondrial (mt) DNA and ribosomal (r) DNA gene sequences. This study also illustrates that the sequence diversity of these two genes in this mosquito species differs geographically which differentiate a population and brings unique identity.

**Key words** *Aedes albopictus*; COI; ITS1; molecular phylogeny; sequence analysis

INTRODUCTION

Mosquitoes are highly efficient vectors of diseases like dengue, malaria, yellow fever, chikungunya, etc. Presently, dengue is a serious arboviral disease in Asia, South and Central America, and Africa; transmitted mainly by *Aedes aegypti* and *Ae. albopictus* vectors. Several recent studies have detected the dengue virus in *Ae. albopictus* in rural and urban settings of India and identified this species as a potential dengue vector1–2. The first major widespread epidemic of dengue occurred in India during 1996 involving areas in Delhi3 and Lucknow4 which subsequently spread to all over the country5–6. Recently, dengue is becoming one of the fastest emerging vector-borne diseases in Assam. Very few cases were reported from this region in the last decade, but in 2012 and 2013, manifold increase in the number of dengue cases was recorded respectively7. In Sonitpur district of Assam, 13 cases were reported in the year 20127.

Correct identification of the *Aedes* species involved in arbovirus transmission is very important to design strategies for vector surveillance and control programme. Larval surveillance was conducted in dengue outbreak areas in Malaysia from 2008 until 20098. Moreover, many closely related species of mosquitoes with varying ecology and host preferences are nearly inseparable morphologically9. This poses difficulty in identification of mosquitoes to a species or even genus level10–12.

As a result, molecular methods of mosquito identification13–15, genetic characterization16–17, and molecular phylogeny18–19 have gained increasing importance. Beebe et al20 developed a molecular identification technique for
differentiating *Ae. aegypti* and *Ae. albopictus*. Molecular detection method is accurate, quicker and economical in identifying the species. While the use of nuclear genes is not remarkable\(^\text{21-24}\), mitochondrial (*mt*) genes have gained wide acceptance for analyzing genetic diversity in mosquitoes\(^\text{25-26}\); since mitochondrial genome has specific advantages such as maternal lineage, lack of recombination and lack of introns\(^\text{27}\). Among the *mt* genes, cytochrome *c* oxidase subunit I (*COI*) is reported to be the most conserved gene in terms of amino acid sequences, and hence has distinct advantage for taxonomic studies\(^\text{28}\). The *COI* gene is also used as molecular marker to distinguish unknown species, and enhance the identification and discovery of new species\(^\text{29}\), when morphological traits do not clearly discriminate species\(^\text{30}\). Besides *COI*, ribosomal DNA (rDNA) has been also used for molecular identification of species because it is one of the multigene families which are frequently distributed in the genome. Many researchers have used rDNA regions for molecular studies of *Aedes* mosquitoes\(^\text{31-32}\). Eukaryotic rDNA contains two internal transcribed spacer (ITS) regions, *i.e.* *ITS1* and *ITS2*, flanked by 18S and the 28S ribosomal region. Amongst them, the *ITS1* region is one of the most variable parts of the genome, which is able to affirm and align sequences according to the conserved parts of rRNAs. Hence, for phylogenetic analysis of closely related species it is the most suitable sequence\(^\text{33}\). Gene sequences such as *ITS1* and *ITS2* of rDNA, *COI* and *COII* are helpful in constructing phylogenetic relationship in between species\(^\text{34-49}\). Hence, molecular tools that could identify mosquito species even from a small piece of tissue from any developmental stage would be of great importance in mosquito taxonomy.

The objective of the study was molecular identification of the most abundant species of *Aedes* vectors collected from different breeding spots of Sonitpur district, Assam and to establish the phylogenetic relationship with same species reported earlier from other parts of South-east Asia.

**MATERIAL & METHODS**

**Mosquito collection**

Wild type *Ae. albopictus* larvae were collected from different geographical areas of Sonitpur district of Assam in pre-monsoon, monsoon and post-monsoon seasons during the period 2010–12. Dark plastic containers (15 cm diam; 12 cm depth) were used as an ovitrap. Larvae were collected from the natural breeding habitats as well as from the ovitraps set in residential areas using Pasteur pipette. Samples were collected from different areas of Sonitpur district, namely Tezpur (26.63°N, 92.80°E), Bihaguri (26.11°N, 91.83°E), Dhekiajuli (26.70°N, 92.5°E), Balipara (26.82°N, 92.77°E), Rangapara (26.82°N, 92.65°E), Biswanath Chariali (26.73°N, 93.15°E), North Jamuguri (26.78°N, 92.91°E), Gogra T.E. (26.82°N, 92.68°E), Sirajuli (26.7°N, 92.54°E), Behali (26.85°N, 93.38°E) and Gohpur (26.88°N, 96.63°E) (Fig. 1). The collected larvae were reared to adult in the mosquito rearing room at the Defence Research Laboratory, Tezpur, Assam.

**Morphological identification**

The larvae collected from different sampling sites were identified using morphological characteristics such as comb scale and pecten teeth, and the adults reared from larvae were identified using standard keys\(^\text{50}\).

**Molecular identification**

Adult specimens were identified as *Ae. albopictus* by morphological characteristics. Individual adults from different areas were selected for molecular characterization.
Genomic and mitochondrial DNA extraction

Whole DNA was isolated from individual adult mosquito, reared from the collected larvae by DNeasy® Blood and Tissue Kit (QIAGEN, USA) according to manufacturer’s standard protocol. Isolated DNA was eluted with provided elution buffer and preserved at –20°C till further molecular analysis.

PCR amplification of ITS1 and COI partial DNA sequence

A 577 bp partial sequence of ITS1 gene from genomic DNA and a 520 bp partial sequence of COI gene from mitochondrial DNA was amplified from morphologically identified adult Ae. albopictus mosquitoes. The ITS1 partial sequence was amplified using AUF and AUR primer sequences, whereas COI partial fragment was amplified using the primers as described earlier. For both ITS1 and COI, amplification was carried out in total of 50 µl reaction volume containing PCR reaction buffer at 1.5x concentration, both forward and reverse primers at 20 pmol final concentrations (Bioserve Biotechnologies, Hyderabad), 3 mM MgCl₂ (Sigma-Aldrich, USA), 200 µM of each dNTP (Wraught, Germany) and 1 U/ml Taq DNA polymerase (Sigma-Aldrich, USA). About 2 µl of purified DNA was added as template to each amplification reaction. Amplification of these two gene fragments, two different thermal profiles were used. Table 1 illustrates about the oligonucleotide primers and PCR thermal profiles used for this study. All amplified PCR products were subjected to agarose gel electrophoresis (1.5%, prepared in 1x TAE buffer stained with ethidium bromide) with reference to a 50 bp DNA ladder. The DNA fragments were visualised in gel documentation system (Syngene, UK) under UV trans-illumination.

For DNA isolation, PCR master mix preparation and template addition, three different laminar air flow (LAF) hood (Labconco, Germany) were used. All LAF hoods were UV sterilised properly prior to use.

DNA sequencing and multiple sequence alignment

For each gene, a total of five PCR products were selected for DNA sequencing. Amplified PCR products were purified using the QIAquick PCR purification kit (QIAGEN, USA) according to the manufacturer’s standard protocol. Purified PCR products were outsourced to Biolinke, New Delhi, India for capillary sequencing involving Sanger’s chain termination reaction. Double pass sequencing for both forward and reverse strand was performed for each of selected samples. Nucleotide sequences, thus obtained were edited in BioEdit software version 6.0.7 and consensus sequence was created isolate wise by aligning the forward and reverse nucleotide strand. Consensus sequences thus obtained were searched in basic local alignment search tool (BLAST) for similarity and deposited in GenBank, NCBI under accession number KC527656-KC527660 (ITS1) and KT260197-KT260201 (COI). Multiple sequence alignment (MSA) for both ITS1 and COI genes sequence was performed separately with sequences available in GenBank from other parts of Southeast Asia. The MSA was carried out in molecular evolutionary genetics analysis (MEGA) software version 6.06 using ClustalW tool following default parameters. Distance amongst multiple aligned sequences was calculated using Tamura 3-parameter model.

Molecular phylogenetic tree reconstruction and relationship elucidation

Morphological data can fetch the related species, but molecular phylogeny can help in revealing closely related as well as sibling species under a genus. In molecular phylogenetic analysis, rDNA ITS1 and mt COI genes sequences of Ae. albopictus were used for construction of phylogenetic tree using Tamura 3-parameter model in MEGA 6.0 software. Also, each sequence was subjected to 1000 bootstrap replications for accuracy of the constructed phylogenetic tree. A cluster was considered significant if it contain >50% permuted tree.

RESULTS

The morphological characteristics of both larvae and adult mosquitoes elucidated the subject specimens as Ae. albopictus. Figure 2 shows morphological characters of larvae used for identification of species. The 577 bp partial coding fragment of rDNA ITS1 gene and 520 bp of

<table>
<thead>
<tr>
<th>Target sequence</th>
<th>Primer name</th>
<th>Primer sequence; 5'–3'</th>
<th>Size</th>
<th>PCR thermal profile</th>
</tr>
</thead>
<tbody>
<tr>
<td>ITS1</td>
<td>AUF</td>
<td>TCAAAAATTAAGGGTAGTGT</td>
<td>577 bp</td>
<td>94°C for 5 min; 30 cycles of 94°C for 30 sec; 55°C for 30 sec; 72°C for 2 min; 72°C for 4 min</td>
</tr>
<tr>
<td></td>
<td>AUR</td>
<td>GACTTCAACTCGTCTAAAG</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COI</td>
<td>For-COI</td>
<td>GAGGATTGAGAATTGATTAGT</td>
<td>520 bp</td>
<td>94°C for 10 min; 30 cycles of 94°C for 30 sec; 55°C for 30 sec; 72°C for 2 min; 72°C for 4 min</td>
</tr>
<tr>
<td></td>
<td>Rev-COI</td>
<td>CCGGTATAAACTAAAAATATAT</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Das et al: Genetic characterization and molecular phylogeny of *Aedes albopictus*

Partial coding fragment of *mt COI* gene were successfully amplified by the PCR protocol (Figs. 3a and b). The DNA sequencing result revealed that only five isolates were of high quality and the rest were not found suitable. BLAST results of the subject specimens indicated identity with *Ae. albopictus* species for both *ITS1* gene (99% maximum identity) and *COI* gene (100% maximum identity). In MSA analysis, a single base mutation (Adenine instead of Guanine) in *COI* gene was observed at nucleotide position 2118 in two of the collected isolates, viz. isolate KT260199 and KT260201 and rest three isolates were found to have similar nucleotide sequences to that of other isolates from India and Southeast Asia. On the contrary, MSA of *ITS1* region revealed noticeable information about deletion, insertion and mutation in the collected samples. In BLAST analysis, maximum six base deletion was recorded in two isolates (KC527656 and KC527660) followed by a five base deletion in one isolate (KC527660). Number of other deletions present in the isolates has been given in Table 2. For reconstruction of the phylogeny, the Southeast Asian isolates which showed similarity and identity in BLAST result were included. Phylogenetic tree was built in two different ways, i.e. one in between the isolates investigated in present study; and the other in between isolates studied and the isolates reported from Southeast Asia, which provided a better picture about the diversity and evolutionary relationships. To the best of our knowledge this is the first report from Northeast region of India depicting the molecular phylogenetic relationship of *Ae. albopictus* mosquito species based on *ITS1* and *COI* genes.

Based on *ITS1* gene, maximum likelihood (ML) tree showed that these five isolates have been distinctly separated into two clusters, i.e. two isolates (KC527656 and KC527660) lie in a same cluster but three isolates (KC527657, KC527658, and KC527659) fall into two sub-clusters under another cluster. Whereas, in ML tree with two other Southeast Asian isolates, it was observed that isolate KC527656 and KC527659 were grouped into a single cluster with two isolates from Malaysia, within which only one isolate from Assam (KC527656) was
found sub-clustered together with isolates from Malaysia and others (KC527657-KC527659) were grouped into a separate sub-cluster. But isolate KC527660 individually assorted into a single cluster showing that it is distinct from other isolates (Figs. 4 a and b). For COI gene, these studied isolates showed different relationship scenario with other isolates from Southeast Asia (Figs. 4 c and d). The two isolates, i.e. KT260199 and KT260201 sub-clustered with isolates from Singapore as they are closely related. Other three isolates were found to have sub-clustered separately into two different group, viz. isolate KT260198 and KT260200 with Karnataka and Maharashtra, and isolate KT260197 with another isolate from Singapore. Constructing relationship among these studied isolates, it was found that these isolates have grouped into two separate clusters; KT260199 and KT260201 in one cluster and isolate KT260197, KT260198 and KT260200 in another cluster.

Almost similar pattern of amino acid frequencies in COI gene was found amongst the five studied isolates. One of the isolate showed low lysine content in this coding gene, which is supplemented with increased number of other amino acids (Fig. 5a). On the contrary, amino acid frequency in ITS1 gene was found to be dissimilar indicating that polymorphisms in this gene differentiate each of these isolates. However, arginine was found highest in frequencies in one isolate (KC527657) and isoleucine, methionine and tryptophan were recorded as nil in isolates KC527657, KC527656 and KC527660 respectively (Fig. 5b).

**DISCUSSION**

Morphological identification methods have limited capacity in differentiating sibling and closely related species of mosquitoes and this could be overcome by DNA-based identification methods such as rDNA region and mt gene COI. DNA-based identification methods use molecular markers such as nuclear ribosomal ITS, cytochrome b (Cyt-b) and cytochrome c oxidase subunits, COI, COII. Thus, this method can be used for the identification of adult as well as aquatic stages of the Aedes vectors during vector surveillance in arboviral outbreaks like dengue. In this present study, though all individual specimens were identified as Ae. albopictus both by morphological and molecular determination, but variations in nucleotide compositions were observed both in COI and ITS1 genes. Nucleotide variation was observed high in ITS1 gene from all the isolates in comparison to that of the COI implying that ITS1 is more polymorphic than COI gene33. However, the COI gene is highly conserved that has distinct advantage for taxonomic studies but it is species specific. Although, diversity in nucleotide sequence may be notable for geographically discriminating sibling species from a species under a genus. No significant change was seen in COI encoded amino acid composition, but decrease in lysine content was recorded in
Das et al: Genetic characterization and molecular phylogeny of Aedes albopictus

one of the isolates. On the contrary, considerable fluctuations in amino acid content in between isolates were noticed in the ITS1 region.

During the retrieval of sequences from gene bank data base, only 11 representative COI gene sequences could be retrieved from Southeast Asia. In contrast to this, only two ITS1 gene sequence from Malaysia were obtained from the data base as a representative of Southeast Asia for molecular phylogenetic analysis. In case of ITS1 gene though we found only two other isolates from Malaysia, but phylogenetic analysis showed that genetic diversity between these isolates was less and hence they are in close proximity in the phylogenetic tree. However, in case of COI gene, studied isolates were found distantly grouped into separate clusters. Though, the two selected marker genes are useful in genetic characterization of this mosquito species, but mt COI gene reflects a better picture for genetic characterization and molecular phylogenetic analysis and the data obtained in present study are consistent in case of COI gene. However, the phylogenetic analysis of the isolates showed close relation among the isolates in both ITS1 and COI genes (Figs. 4 b and d for ITS1 and COI genes respectively). But, they are distantly related when compared with Southeast Asian isolates (Figs. 4 a and c for ITS1 and COI genes respectively).

CONCLUSION

Mosquito identification using molecular tools and phylogeny related information are lacking from North-east region of India. The molecular data of conserved gene not only identify the species, subspecies but also help in describing the relationship among species and sibling species. Proper information of biogeography, diversity, and
population genetic structure of the prevailing mosquito species might decipher their origin, spread and distance wise genetic relationship resulting in proper taxonomic information. Keeping in view, the increased incidence of dengue in India, the molecular information of dengue vector need to be explored extensively.

Conflict of interest
The authors declare that they have no conflict of interest.

ACKNOWLEDGEMENTS
The authors would like to sincerely thank the Director, Defence Research Laboratory, Tezpur, Assam for providing research facility and constant encouragement. Ms. Momi Das was supported by research fellowship from the Defence Research and Development Organisation, New Delhi.

REFERENCES


