See discussions, stats, and author profiles for this publication at: http://www.researchgate.net/publication/284437998

Genetic deviation in geographically close populations of the dengue vector Aedes aegypti (Diptera: Culicidae): influence of environmental barriers in South India

ARTICLE in PARASITOLOGY RESEARCH · JANUARY 2016
Impact Factor: 2.1 · DOI: 10.1007/s00436-015-4847-7

READS
55

14 AUTHORS, INCLUDING:

Dr. Chithravel VADIVALAGAN
Bharathiar University
10 PUBLICATIONS  16 CITATIONS
SEE PROFILE

Dr. Chellasamy Panneerselvam
University of Tabuk
43 PUBLICATIONS  456 CITATIONS
SEE PROFILE

Marcello Nicoletti
University of Cambridge
279 PUBLICATIONS  1,943 CITATIONS
SEE PROFILE

Paramasivan Rajaiah
Indian Council of Medical Research
48 PUBLICATIONS  251 CITATIONS
SEE PROFILE

All in-text references underlined in blue are linked to publications on ResearchGate, letting you access and read them immediately.
Genetic deviation in geographically close populations of the dengue vector Aedes aegypti (Diptera: Culicidae): influence of environmental barriers in South India

Chithravel Vadivalagan1 · Pushparaj Karthika2 · Kadarkarai Murugan1 · Chellasamy Panneeselvam1 · Manickam Paulpandi1 · Pari Madhiyazhagan1 · Hui Wei3 · Al Thabiani Aziz4 · Mohamad Saleh Alsalhi5 · Sandhanasamy Devanesan5 · Marcello Nicoletti6 · Rajaiah Paramasivan7 · Devakumar Dinesh1 · Giovanni Benelli8

Received: 15 September 2015 / Accepted: 20 November 2015 © Springer-Verlag Berlin Heidelberg 2015

Abstract Mosquitoes are vectors of devastating pathogens and parasites, causing millions of deaths every year. Dengue is a mosquito-borne viral infection found in tropical and subtropical regions around the world. Recently, dengue transmission has strongly increased in urban and semiurban areas, becoming a major international public health concern. Aedes aegypti (Diptera: Culicidae) is a primary vector of dengue. Shedding light on genetic deviation in A. aegypti populations is of crucial importance to fully understand their molecular ecology and evolution. In this research, haplotype and genetic analyses were conducted using individuals of A. aegypti from 31 localities in the north, southeast, northeast and central regions of Tamil Nadu (South India). The mitochondrial DNA region of cytochrome c oxidase 1 (CO1) gene was used as marker for the analyses. Thirty-one haplotypes sequences were submitted to GenBank and authenticated. The complete haplotype set included 64 haplotypes from various geographical regions clustered into three groups (lineages) separated by three fixed mutational steps, suggesting that the South Indian Ae. aegypti populations were pooled and are linked with West Africa, Columbian and Southeast Asian lineages. The genetic and haplotype diversity was low, indicating reduced gene flow among close populations of the vector, due to geographical barriers such as water bodies. Lastly, the negative values for neutrality tests indicated a bottle-neck effect and supported for low frequency of polymorphism among the haplotypes. Overall, our results add basic knowledge to molecular ecology of the dengue vector A. aegypti, providing the first evidence for multiple introductions of Ae. aegypti populations from Columbia and West Africa in South India.

Keywords Arbovirus · Mosquito-borne diseases · Yellow fever · CO1 · DNA barcoding · Phylogenetics · Ancestral lineage · Purifying selection · Molecular ecology

Electronic supplementary material The online version of this article (doi:10.1007/s00436-015-4847-7) contains supplementary material, which is available to authorized users.

Giovanni Benelli

g.benelli@sssup.it; benelli.giovanni@gmail.com

1 Division of Entomology, Department of Zoology, School of Life Sciences, Bharathiar University, Coimbatore 641 046, Tamil Nadu, India
2 Department of Zoology, Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore 641 043, Tamil Nadu, India
3 Institute of Plant Protection, Fujian Academy of Agricultural Sciences, Fuzhou 350013, Fujian, People’s Republic of China
4 Faculty of Science, Department of Biology, University of Tabuk, Tabuk 71491, Saudi Arabia
5 Department of Physics and Astronomy, Research Chair in Laser Diagnosis of Cancer, King Saud University, Riyadh, Kingdom of Saudi Arabia
6 Department of Environmental Biology, Sapienza University of Rome, Piazzale Aldo Moro 5, 00185 Rome, Italy
7 Centre for Research in Medical Entomology, Madurai, India
8 Department of Agriculture, Food and Environment, University of Pisa, Via del Borghetto 80, 56124 Pisa, Italy

Published online: 02 December 2015
Introduction

Mosquitoes are vectors of devastating pathogens and parasites, causing millions of deaths every year (Mehlhorn et al. 2012; Benelli 2015). Dengue is a mosquito-borne viral infection found in tropical and subtropical regions around the world. *Aedes aegypti* (L.) (Diptera: Culicidae) is a primary vector of dengue. Recently, dengue transmission has increased in urban and semiurban areas, becoming a major international public health concern (Chandran and Azeez 2015). The incidence of dengue has grown dramatically around the world in recent decades. The actual numbers of dengue cases are underreported, and many cases are misclassified. One recent estimate indicates 390 million dengue infections per year (95% credible interval 284–528 million), of which 96 million (67–136 million) manifest clinically (with any severity of disease); 3900 million people, in 128 countries, are now at risk of infection with dengue viruses (WHO 2015).

India alone accounted for 34% of global total infections. In 2001, a major outbreak was recorded in Chennai, 90% of the 861 cases recorded from the whole of Tamil Nadu (Ashok Kumar et al. 2010). Since then, the number of dengue incidences in Tamil Nadu have been raised considerably (Víctor et al. 2007), and the spread of the disease in rural (Paramasivan et al. 2006) and semiurban areas of Tamil Nadu has become a matter of concern for public health. Notably, during the past few years, the frequency of dengue haemorrhagic fever has increased remarkably in South India (Thangarahan et al. 2006). In 2010–2012, outbreaks of dengue and chikungunya-like illnesses with severe clinical manifestations were reported from several districts of Tamil Nadu (i.e., Tirunelveli, Virudunagar, Theni, Madurai, Thiruvallur, Vellore and Dharmapuri) (Wilson and Sevarkodyione 2014). Currently, there is no specific treatment for dengue, even if the development of a vaccine is in progress (Murrell et al. 2011), and anti-dengue nanodrugs also showed promising activity (Murugan et al. 2015, 2016a). Its prevention and control solely depends on effective vector control measures (Sujitha et al. 2015; Suresh et al. 2015; Benelli 2016).

The accurate identification of arthropod vectors is a crucial requirement to enhance effective control programmes. Traditional morphology-based taxonomic procedures are time-consuming and not always sufficient for identification to the species level. Therefore, a multidisciplinary approach to taxonomy that includes morphological, molecular and distributional data is of key importance (Munstermann and Conn 1997; Krzywinski and Besansky 2003). Moreover, the efficient monitoring of mosquito vector populations would be beneficial for understanding their dispersal capability and genetic diversity. After that Tautz and Arctander (2003) proposed to use DNA sequences as the main basis of biological classification, it has been suggested that sequencing the *COI* gene could allow DNA barcoding, facilitating such classification (Remigio and Hebert 2003). Later on, several studies have demonstrated that the *COI* gene is a valid molecular tool for identifying and monitoring mosquito species (Kumar et al. 2007; Werner and Kampen 2013; Huber et al. 2014; Deblauwe et al. 2015), and revealing cryptic species (Burns et al. 2007).

Population genetic studies have been conducted with *Ae. aegypti* from different regions across the world, using multi-markers (Herrera et al. 2006). These studies have provided information on population structure and dispersion rates at the micro- and macro-geographic levels, showing that environmental and social factors and human interventions (i.e., urbanisation, control activities) affect the population structure of this mosquito vector on medical and veterinary importance (Huber et al. 2002). Recent studies using the mitochondrial gene nicotinamide adenine dinucleotide dehydrogenase subunit 4 (*ND4*) were conducted with American (including Brazilian), African and Asian populations of the vector (Bracco et al. 2007), whereas in Brazil, utilisation of isozymes and random amplified polymorphism DNA (RAPD) markers indicated differentiation among populations (Costa-Ribeiro et al. 2006).

To the best of our knowledge, the population structure, lineage distribution and dispersal capabilities of *Ae. aegypti* from India have been scarcely investigated. Recently, Murugan et al. (2016b) reported high intraspecific divergence among Indian specimens belonging to the genus *Aedes* and emphasised that extensive sampling from wide study areas may provide clues for the high divergence among the same species. On the basis of the above postulated problem and on the alarming threat of dengue infection in South India, this research was carried out to enable rapid and accurate identification of the dengue vector *Ae. aegypti*, understanding its haplotype dispersal and association with ancestral lineages, and finally connect the results with the *Ae. aegypti* genetic differentiation among the haplotypes worldwide.

Materials and methods

Sample collection

Mosquito specimens used in the study were collected from 31 different sites in Tamil Nadu (South India). For each site, main temperature and rainfalls are reported in Supplementary Online Material Table S1. Adult collections were carried out in human dwelling areas. Individuals were identified based on the morphological characters (Gaffigan et al. 2015; Murugan et al. 2016b and references therein), then used for DNA extraction.
DNA extraction, PCR amplification and sequencing

Total genomic DNA was isolated from single whole mosquito sample using Qiagen DNeasy kit (Qiagen, Inc., Hilden, Germany) following manufacturer’s instructions. The amplified polymerase chain reaction (PCR) product was visualised in 1% agarose gel electrophoresis (AGE) (GENEI, Bangalore), and the image was documented with the gel documentation (Medic Care, India). For DNA barcode analysis, the 735-bp region of mitochondrial COI gene was targeted and amplified with the following primers: forward 5′-GGATTTGAAAATTTGATTTCCC-3′ and reverse 5′-AAAAATTTAATTCAGTTGGAAC-3′ (Kumar et al. 2007). The polymerase chain reaction (PCR) was carried out in ABI thermo-cycler. The 50-μl PCR reaction mixture consisted of 5 μl of extracted DNA, 1.5 mM MgCl2 (Sigma, Bangalore), 0.2 mM dNTPs (Sigma, Bangalore), 1× reaction buffer (Sigma, Bangalore), 1.5 U Taq DNA polymerase (Sigma, Bangalore) and 0.3 μM of each primer. PCR reaction conditions were as follows: an initial denaturation of 95 °C for 5 min, followed by 5 cycles of denaturation at 94 °C for 40 s, annealing at 51 °C for 1 min and 72 °C for 1 min, followed by a final extension at 72 °C for 10 min (Kumar et al. 2007). The final product was stored at −20 °C for further use. Sequencing was performed by using Sanger sequencing method in ABI 3500 XL Genetic Analyzer with manufacturer’s protocol of Chromos Biotech, Pvt. Ltd., Bangalore, India. The sequences were trimmed and edited using ClustalW and Bio Edit v.7.2.5 (Hall 1999) and submitted to NCBI-GenBank (Murugan et al. 2016b).

Data analysis

Species discrimination by DNA barcodes, nucleotide composition and genetic divergence of sequences The nucleotide sequence from each specimen was compared to barcode sequences on NCBI using BLASTn. Thirty-three sequences of Ae. aegypti haplotypes were retrieved from NCBI-GenBank and used for comparisons (Supplementary Online Material Tables S2 and S3). The nucleotide composition and AT bias was calculated using DnaSp v.5.1 (Librado and Rozas 2009). The evolutionary divergence and disparity index per site was calculated using the MEGA v.6 (Tamura et al. 2013).

Substitution rates, test for saturation and test of selection for the evolutionary divergence Other parameters of the sequences including the rate of transitions (Ts) and transversions (Tv) at the first, second and third codon positions were calculated and subsequently plotted against the F84 genetic distance for all the three codons to assess the saturation using DAMPE 5.3.10 (Xia 2013; see also Murugan et al. 2016b). Further test for substantial saturation of the sequences was checked following Xia et al. (2003) and Xia and Lemey (2009) (DAMBE). The probability of rejecting the null hypothesis of neutral evolution (p value) was calculated by Z-test of codon-based test of neutrality for analysis between sequences. Analyses were conducted using the Nei-Gojobori method (Nei and Gojobori 1986) for sequence pairs. All evolutionary analyses were carried out using MEGA v.6 (Tamura et al. 2013).

Genetic differentiation and haplotype network analysis Genetic variability for COI marker was evaluated by the number of haplotypes (h), number of variable sites (S), average number of nucleotide differences (k), haplotype diversity (Hd) and nucleotide diversity (π) using DnaSp v.5.0 (Librado and Rozas 2009). The Tajima’s D test (Tajima 1989) was performed to test the hypothesis that all mutations were selectively neutral (Kimura 1987) using DnaSp v.5.0. A median-joining haplotype network was built to examine the haplotype relationship among 31 individuals and TCS for the entire haplotype set (n=64) to visualise relationships among haplotypes with default parameters using PopART software (http://popart.otago.ac.nz/index.shtml).

Phylogenetic analysis The optimum substitution models were determined by the best fit model test for the selection of the model to be applied for sequences. The T92+G model was selected from 24 different nucleotide substitution models for the ten original sequences of the selected species in the present study based on the Akaike information criterion (AIC) and lowest Bayesian information criterion (BIC) values. The maximum likelihood (ML) tree was constructed, and the robustness of the clades of the tree was estimated using bootstrap analysis of 1000 replications with the elimination of all the codons containing gaps and missing data. The strength of the clades was assessed with the ten random addition replicates and bootstrap analysis. Branch lengths were calculated by average pathway method and were presented as the units of number of changes over the whole sequences. The evolutionary history was inferred using the neighbour-joining method (Saitou and Nei 1987). The optimal tree with the sum of branch length=4.36540551 is shown. The percentage of replicate trees in which the associated taxa clustered together in the bootstrap test (1000 replicates) was shown next to the branches (Felsenstein 1985). The evolutionary distances were computed using the Jukes-Cantor method (Jukes and Cantor 1969) and are in the units of the number of base substitutions per site. The rate variation among sites was modelled with a gamma distribution (shape parameter=1). The analysis involved 64 nucleotide sequences. Codon positions included were 1st+2nd+3rd+ noncoding. All positions containing gaps and missing data were eliminated. There were a total of 145 positions in the final dataset.
Results

In our analysis, the nucleotide sequences of the haplotypes of *Ae. aegypti* were generated from 31 sampling sites in Tamil Nadu, South India. Each of the haplotype was represented from one to twenty-nine individuals for sampling. The sequence alignments were straightforward and devoid of indels, pseudo genes and introns. Genomic DNA extracted from the whole insect produced good quality of >11 kb, and the PCR product size was approximately 750 bp. BLAST search showed 100 % identity, and the e-values were zero for almost all the sequences. To the best of our knowledge, little efforts have been done to shed light on genetic deviation in Indian populations of mosquito vectors. Here, we firstly recorded 31 haplotypes of *Ae. aegypti* from South India; all the sequences were submitted and authenticated in NCBI-GenBank, with respective accession numbers KT339653–KT339683.

The CO1 fragment of the haplotypes showed variations in the base composition, and it was expectedly AT biased. In agreement with the characteristic feature of mitochondrial genes, the sequence set was AT rich (65.50–69.70 %), whereas the GC content ranged from 30.22 to 34.44 %. The analysed haplotype set of sequences had 186 variable sites of which 179 were parsimony informative. The transition and transversion bias ($R$) was 4.75, 0.82 and 1.46 in the 1st, 2nd and 3rd codons, respectively. The number of transitions between A and T in all three codon positions (3.22, 14.24 and 10.86) was higher if compared to the transversions between G and C (0.36, 11.52 and 9.36). Similarly, the number of transitions between C and T and number of transition between A and G was similar (46.41, 24.24 and 19.78). Overall, Ts and Tv substitutions in all the 1st, 2nd and 3rd codons were (3.58, 92.82), (24.76, 48.48) and (22.22, 59.56), respectively. All the three codon positions of the CO1 fragment were tested individually for saturation by plot estimation of the number of substitutions versus the F84 genetic distance (Fig. 1). The plot of the 2nd codon showed that Ts and Tv increased along the F84 distance with considerable scattering. However, in the 1st and 3rd codons, the Ts and Tv exhibited a marked difference in plot patterns. The Ts showed a linear relationship in both the 1st and 3rd codons, whereas the Tv substitution exhibited a linear relation to the genetic distance prior and later experienced fall due to minimum substitution in certain pairs of taxa. Furthermore, the test of substantial saturation revealed that the value of Iss was lower than the Iss.c (Supplementary Online Material Table S4). It was evident that the selected sequences of the current study did not undergo substantial saturation and could be utilised for DNA barcoding and phylogenetic studies.

The thirty-one haplotypes detected from area-wide sampling in South India differentiated among themselves, forming unique clusters in the median-joining haplotype network (Fig. 2a) and represented in geo map (Fig. 2b). G1 cluster comprised of largest set of identical haplotype sequences which were distinguished from the several other clusters like G2, G3 and G4 by four and one mutational step, respectively. The H10, H2, H3, H11 and H14 haplotypes were ideal and did not from any clusters; they varied with G1 by one (H10, H2), six (H3), one hundred and seventy-six (H11) and one hundred and seventy-five (H14) mutational steps. The highest number of variations was observed from the haplotypes from Namakkal (H11) and Krishnagiri (H14). The haplotypes of the present study were compared with the existing haplotypes worldwide, and the nucleotide sequences were retrieved from NCBI (Supplementary Online Material Tables S2 and S3). The TCS haplotype network of the entire set of 64 haplotypes of the CO1 fragment formed six main groups based on the identity and low genetic divergences (Fig. 3). Accordingly, the major cluster G11 comprised of mixed combination of South Indian haplotypes (H15, H16, H19, H20–H26, H28, H31, H36–H38, H41–H45, and H47) together with the West African and Southeast Asian lineages. Most of the haplotype of the present study was unique and was connected with each other, except those from Ramanathapuram (H18) and Villupuram (H30), which form distinct group with the haplotype from Singapore. The G8 cluster comprised of haplotypes from Dharmapuri (H5) and Kanyakumari (H9) associated with the Columbian lineage. The other haplotypes such as H2, H3, H10, H11 and H14 were isolated and did not from any clusters with the other haplotypes.

The genetic diversity indices of *Ae. aegypti* haplotypes are shown in Table 1. The highest haplotype diversity (0.933 and 0.8971) and variable sites (6 and 4) are found in the cluster G8 (Colombian lineage) and G11 (mixed lineage of West Africa and Southeast Asia). The neutrality indices of Tajima’s D and Fu’s S were found to be negative values (−1.909 and −1.701) which indicated low frequency of polymorphism and highlighted recent population expansion of this species, respectively.

The Z-test of selection was performed to trace the evolutionary lineage among the haplotypes (Fig. 4). The estimation of positive selection revealed that the test statistic nonsynonymous and synonymous substitutions ratio (dN-dS) possessed maximum negative values and p-values >0.01. On the other hand, the estimation of purifying selection had maximum positive values for the test statistic (dS-dN), and the maximum p values were significant ($p < 0.01$). Hence, the probability of rejecting null hypothesis in strict neutrality in favour of the alternate hypothesis of purifying selection was true for the evolutionary selection of the haplotype sequences of the *Ae. aegypti*. Significantly, the haplotypes H11 and H14 exhibited wide variations in the entire selection. The genetic divergence of the entire haplotype population of *Ae. aegypti* was represented in Fig. 5. The conspecific divergence ranged from 0 to 0.034, and the threshold level of the nucleotide
divergence was 0.09%. According to the 10× rule of Hebert et al. (2003), 100% of the cases were resolved up to the species level in the entire set of haplotypes because the average genetic divergence was 0.009 (0.09%) which was significantly lower than the postulated 3% level (Hebert et al. 2013) and can be utilised for discrimination up to species level. In the
case of intraspecies divergence of the haplotypes representing Tamil Nadu (South India), it is noteworthy that only few haplotype comparisons revealed low divergence. The highest evolutionary divergence between
the sequences was 1.88, whereas the lowest was 0.22 (Fig. 6). The high divergence among the species of the same geographical region (Tamil Nadu) indicates larger difference in base composition than the expected, perhaps based on the evolutionary divergence of the lineage.

The ML tree constructed from the haplotypes of the present study formed monophyletic clades with five distinct clusters (Fig. 7). The major clusters one and four comprised of fifteen and seven haplotypes, respectively. However, the haplotypes from Coimbatore and Viruthunagar exhibited wide divergence and formed separate clades. In the ML analysis, the best fit model of the nucleotide substitution was evaluated based on the AIC corrected value of 2154.76 and BIC score of 2599.54. The T92+G model was selected for ML analysis of the 31 sequences representing the partial CO1 gene sequence. Bootstrapping of the ML analysis (1000 replications) was implemented in NNI ML heuristic method with highest log likelihood (−1015.6396).

<table>
<thead>
<tr>
<th>Cluster</th>
<th>Parameter</th>
<th>h</th>
<th>s</th>
<th>k</th>
<th>Hd</th>
<th>π</th>
<th>Tajima’s D</th>
<th>Fu’s FS</th>
</tr>
</thead>
<tbody>
<tr>
<td>G3</td>
<td></td>
<td>2</td>
<td>2</td>
<td>0.441</td>
<td>0.221</td>
<td>0.00093</td>
<td>1.069</td>
<td>−0.6347</td>
</tr>
<tr>
<td>G7</td>
<td></td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>0.833</td>
<td>0.01745</td>
<td>−0.887</td>
<td>95.59</td>
</tr>
<tr>
<td>G8</td>
<td></td>
<td>5</td>
<td>6</td>
<td>2.2</td>
<td>0.933</td>
<td>0.00455</td>
<td>−1.909</td>
<td>−0.9317</td>
</tr>
<tr>
<td>G11</td>
<td></td>
<td>2</td>
<td>4</td>
<td>1.572</td>
<td>0.891</td>
<td>0.0031</td>
<td>−1.701</td>
<td>−0.1325</td>
</tr>
<tr>
<td>G12</td>
<td></td>
<td>3</td>
<td>2</td>
<td>0.933</td>
<td>0.733</td>
<td>0.933</td>
<td>−0.304</td>
<td>4.6384</td>
</tr>
</tbody>
</table>

*Tajima’s D*—a negative Tajima’s D signifies an excess of low frequency polymorphisms relative to expectation. Not significant, *P*=0.10
*Fu’s FS*—a negative value evidences an excess number of alleles and recent population expansion. Significant, *P*=0.02
h number of haplotypes, s number of variable sites, k average number of nucleotide differences, Hd haplotype diversity, π nucleotide diversity

Fig. 2  a Median-joining network and b geo map for *Aedes aegypti* haplotypes from South India; four distinctive clusters represents the identical haplotypes in the South Indian region of Tamil Nadu. *H1* Salem; *H4* Cuddalore; *H5* Dharmapuri; *H6* Dindugal; *H7* Erode; *H8* Kancheepuram; *H9* Kanyakumari; *H12* Madurai; *H13* Nagapattinam; *H15–H17* Niligiris, Perambalur, Pudukottai; *H18* Ramanathapuram; *H19–H22* Sivagangai, Thanjavur, Theni, Tiruchirapalli, Thiruvallur, Thiruvanamalai, Thirunelveli, Tiruppur, Tiruvur, Tuticorin, Vellore; *H30* Villupuram; *H31* Viruthunagar (see Supplementary Online Material Table S2 and S3 for full list of haplotypes)
Discussion

Our study was focused on geographically close populations of the primary dengue vector *Ae. aegypti*. Due to its role as an invasive species and globally important disease vector, it is critical to understand the patterns by which *Ae. aegypti* populations have associated with humans. Recently, Brown et al. (2014) utilised the genetic markers to show the impact of anthropogenic forces in influencing sub-speciation of invasive and epidemiologically important subspecies of *Ae. aegypti* from the ancestral clade *Ae. aegypti* formosus. DNA analysis provides an accurate way to identify vector species, and the use of molecular data, in combination to morphological methods, has resolved some long-standing taxonomic problems.

Fig. 4 Codon-based test to determine the type of selection inferred by sequence pairs. Scatter plots show the test statistics (dN-dS) (light blue plot) and (dS-dN) (green plot) for positive and purifying selection, respectively; *p* values (dark blue plot) are marked for each selection separately (*p*<0.05 is considered significant).

Fig. 5 Distribution of genetic distance in haplotypes of *Aedes aegypti*. Entries for intraspecific divergence were approximately <3%. The dotted line represents the threshold level of the divergence among the haplotypes.

Fig. 6 Estimation of evolutionary divergence among the *Aedes aegypti* haplotypes. Dot plot shows the nucleotide divergence among the haplotypes of Tamil Nadu (South India); deep intraspecies nucleotide divergence was observed in certain pairs of taxa indicating the difference in base composition.
questions (de la Herran et al. 2001). As a consequence, DNA-based approaches to mosquito identification (Kang and Sim 2013; Werner and Kampen 2013; Deblauwe et al. 2015). Genetic diversity (Krida et al. 1998; Pfeiler et al. 2013) and molecular phylogeny (Sharma et al. 2013) have growing attention. Although use of nuclear genes is not uncommon (Surendran et al. 2011), mitochondrial genes have gained primary adoption for analysing genetic diversity in mosquitoes (Galtier et al. 2009). Construction of DNA barcodes for each species of mosquito would provide an important tool for identification of mosquito species and may enable description of the species biodiversity of this important group of insects (Murugan et al. 2016b). Here, we present the first report about construction of DNA barcodes for 31 haplotypes of *Ae. aegypti* from South India.

The milestones of the DNA-based identification depend on the effective recovery of target genomic DNA, ease in analysing the sequence information and utilising the sequence data to discriminate species and subspecies level identification. In coherence to it, all the sequences enabled accurate identification in BLAST search, and no indels or pseudo genes were present. The composition of the targeted CO1 gene was AT rich which was supported by Cywinska et al. (2006). The effective application of DNA sequence data to molecular diagnostics depends on patterns of nucleotide substitution and the rate of variation among sites (Blouin et al. 1998).

Accordingly, the CO1 region of the haplotypes was characterised by the accumulation of Ts and Tv in linear fashion in the 2nd codon, whereas the 1st and 3rd codons experienced low rate of transversions and high transitions. Similar case was observed in the population genetics of *Ae. aegypti* in Brazil, in which all substitutions found were transitions (Scarpassa et al. 2008). On contrary, Wang et al. (2012) showed high rates of transversions in the main mosquito species of China. This could be possibly due to the wide genetic divergence caused by substitutions among the congenerics and conspecifics. Ideology of DNA barcoding depends on the potential to discriminate inter and intra species, in which the congeneric divergence is higher than the conspecific divergence (Ruiz-Lopez et al. 2012). Since the haplotypes of the present study represented a single conspecific cluster from a closed geographical location, the rate of low frequency polymorphisms and high synonymous substitutions (dS) would have influenced the high rate of transitions than transversions. This could be effectively declared by increasing the haplotype sampling in the study location in the future research. Deep intraspecific divergences were noted in the H11 and H14 haplotypes, in agreement with recent studies on *Collessius pseudotaeniatus* mosquitoes in Pakistan (Ashfaq et al. 2014).

The primary presumed ancient form is *Ae. aegypti formosus* sylvan mosquito (Aaf) restricted to the sub-Saharan region of Africa, while *Ae. aegypti aegypti* (Aaa)
was widespread across most of the tropical and subtropical regions, associated with humans (Tabachnick and Powell 1978). In this study, the haplotypes of *Ae. aegypti* in South Indian region were investigated to check the association of them with the ancestral lineages. HaplotypeS formed three clusters which were found associated with West Africa, Colombia and South East Asian lineages. This grouping of haplotypes could be due to the factors of vector competence and insecticide tolerance from different origins (Sim et al. 2013). Bracco et al. (2007) suggested that missing of one lineage or incomplete colonisation is perhaps due to micro-evolutionary forces acting against one another. According to the hypothesis of Tabachnick and Powell (1979), our results indicated that the G8 cluster could be formed from the ancestor clades from Columbia. However, G11, which was associated with the populations of Western Africa and South East Asia, may be considered as a recently introduced group. As supporting evidence, the G8, G11 and G13 clusters yielded negative values of Tajima’s D which was consistent for purifying selection which also suggested for the high number of low polymorphism frequency in the haplotypes. Recently, Navarro et al. (2013) focused on the bottle-neck effect in the haplotypes of *Aedes albopictus* in South America populations. The combined colonisation of different lineages might have occurred due to multiple introductions of populations by the active adult flight and passive transport of immature and adult stages via trade and air travels. The haplotype lineage association with the West Africa and Columbia was in close congruence with that of the studies by Jaimes-Dueñez et al. (2015), who reported the widespread of Aaf and Aae lineages in Colombia.

In the present study, the average genetic variability was moderate if compared to Scarpassa et al. (2008). The nucleotide diversity of the observed clusters was 0.009, 0.004 and 0.051 which was similar to those of earlier studies in Columbia (Jaimes-Dueñez et al. 2015), Brazil (Bracco et al. 2007), Venezuela (Urdaneta et al. 2005) and Mexico (Gorrochotegui-Escalante et al. 2002). The low genetic diversity may be due to the hypothesis that the South Indian region has poor geographical isolation by the water bodies on the all the three directions [Arabian Sea (west), Indian Ocean (south) and Bay of Bengal (east)], which reduces the process of reinvasion from distant countries, thus interrupting the gene flow occurring between neighbouring lineages. However, our results supported the fact that the *CO1* of *Ae. aegypti* gene was more variable than the nicotinamide adenine dinucleotide dehydrogenase subunit 5 (*ND5*) and cytochrome b (*Cytb*) genes 43 (Mousseon et al. 2005). In studies on the *ND4* gene (Herrera et al. 2006), the nucleotide diversity values were higher than those obtained here, possibly because in *Ae. aegypti*, the *CO1* gene is more conservative than the *ND4* gene.

*Ae. aegypti* is a poikilothermic species; therefore, changes in response to environmental factors such as humidity, temperature and rainfall can affect the genetic structure of its populations (Polson et al. 2012). Considering this characteristic feature of *Ae. aegypti*, a pilot work of the meteorological parameters has been recorded in the 31 sampling sites during the collection period (Supplementary Online Material Table S1), and future research will emphasise the connection between the presence of two ancestral lineages (Columbia and West Africa) among South Indian haplotypes, as well as the influence of climatic factors on their establishment.

**Conclusions**

Overall, this research provides the first evidence for multiple introductions of *Ae. aegypti* populations from Columbia and West Africa into South India. Factors leading to the genetic dispersal of *Ae. aegypti* and their connection with the potential relevance of different climatic factors deserve further research. To our mind, the current study provides evidence for the prominent role of DNA barcoding in mosquito vector identification (Werner and Kampen 2013; Huber et al. 2014; Deblauwe et al. 2015) and sheds light on the presence of genetically distinct dengue vector lineages in Tamil Nadu. These findings contribute to a better understanding of the epidemiological aspects of dengue and aid in improving mosquito control strategies to prevent or reduce the epidemic impacts in India.

**Acknowledgments**  Heinz Mehlihorn and two anonymous reviewers improved an earlier version of our research. The authors would like to thank the SAARP Trust (Coimbatore, India) and the financial support rendered by King Saud University, Saudi Arabia through the Vice Deanship of Research Chairs.

**Compliance with ethical standards**  All applicable international and national guidelines for the care and use of animals were followed. All procedures performed in studies involving animals were in accordance with the ethical standards of the institution or practice at which the studies were conducted.

**Conflict of interest**  The authors declare no conflicts of interest. G. Benelli is an Editorial Board Member of *Parasitology Research*. This does not alter the author’s adherence to all the *Parasitology Research* policies on sharing data and materials.

**Informed consent**  Informed consent was obtained from all individual participants included in the study.

**References**


community knowledge and behaviour following a dengue epidemic in Chennai city, Tamil Nadu, India. Trop Biomed 27:330–336


