

DNA BARCODING OF THE PENINSULAR INDIAN MEMBERS OF THE BLACK FLIES (SIMULIUM: SIMULIIDAE: DIPTERA)

VEERANARAYANAN SURYA AATHMANATHAN¹, SANKARAPPAN ANBALAGAN² AND MUTHUKALINGAN KRISHNAN^{1*}

¹Department of Environmental Biotechnology, Bharathidasan University, Tiruchirappalli - 620 024, Tamil Nadu, INDIA

²Department of Zoology, Government Arts College, Melur - 625106, Madurai, Tamil Nadu, INDIA
e-mail: surya.aathmanathan@gmail.com

KEYWORDS

COI gene
Black fly
Simuliidae
phylogeny

Received on :
09.09.2020

Accepted on :
28.10.2020

***Corresponding author**

ABSTRACT

We examined the efficiency of DNA barcoding gene (cytochrome oxidase subunit I) for differentiating morpho-species and cytoforms of black fly species (subgenus, *Gomphostilbia*) complexes reported in India and analyzed their relationship with members found in South-East Asian countries. Results shows that the spatial heterogeneity is an important factor for determining the diversity of species. The phylogenetic trees derived from DNA barcodes grouped for selected taxa according to species recognized by morpho-taxonomic studies. Inter-specific sequence divergences within morphologically distinct species ranged from 0.66 to 0.99%, while higher divergences (73 to 78%) in species complexes suggest the presence of genetic diversity. DNA barcoding combined with a rigorous morpho-taxonomic context given an effective method for identification and phylogenetic relationship of black flies.

INTRODUCTION

The enzyme cytochrome oxidase subunit I (COI) gene is an effective tool to identify and detect unknown organisms, and are known as DNA barcoding gene. By the use of DNA barcoding gene, many organisms on taxonomically critical and evolutionary importance can be solved. It is widely applied in insect science. The utility of DNA barcoding for these purposes is subject to dispute in insects and these structure can have about 650 bp. Molecular identification and phylogeny using species identification genetic markers by using COI gene is regarded as efficient and their main advantage of DNA barcoding is the rapid acquisition of molecular data (Monaghan *et al.*, 2005). However, we used DNA barcoding gene to study the phylogenetic relationship of the black fly species.

Black flies represent a small and tiny insect of the family Simuliidae in nematoceros Diptera. Larvae play a vital role in stream ecosystem and they consider as good water quality indicators (Dinakaran and Anbalagan, 2007; Figuerio *et al.*, 2012). Most species of the adult female black flies bite humans, birds and other animals due to their need of blood for full egg development (Hernandez-Triana *et al.*, 2017) and other species are non-biting, they feed nectars. Simuliids are known vector for the parasitic disease of onchocerciasis (river blindness), which is responsible for transmitting parasitic nematode *Onchocerca volvulus* (Colebunders *et al.*, 2014; Anbalagan *et al.*, 2017). In India, black flies are represented by 88 species (74 named and 14 unnamed) in 6 subgenera

(*Eusimulium*, *Gomphostilbia*, *Montisimulium*, *Nevermannia*, *Simulium* and *Wilhelmia*). The subgenus *Gomphostilbia* had 22 species (20 named and 2 unnamed) in India (Adler 2020). Of these, 10 species are found in peninsular India.

The fauna of black fly is well known in India and identification of species is still difficult, need expertise for exact identification. To overcome this problem, DNA barcoding method is an excellent tool for identification of species (Ruiz-Arrondo *et al.*, 2018). The success of DNA barcoding using cytochrome oxidase subunit I (COI) gene for temperate organisms might not be reproduced in tropical areas because greater diversity and phylogeographic structure of the taxa (Moritz and Cicero 2004). Therefore, the efficiency of DNA barcoding gene for differentiating morphospecies and cytoforms of species complexes has been examined and also the phylogenetic relationships of seven species in the subgenus *Gomphostilbia* in Peninsular India by comparing phylogenetic trees with existing phylogenies have been investigated.

MATERIALS AND METHODS

Taxa sampling

In the type of locality of each described species, larvae and pupae were collected manually from stream substrates (leaf litter, boulders and pebbles) by using a fine brush and forceps (Anbalagan *et al.*, 2012; Anbalagan *et al.*, 2018). The mature pupae were separated and placed on wet filter paper in a small plastic container for rearing purpose. Then, this same set up was maintained up to adult emergence (Anbalagan *et al.*

al., 2014). The collected specimens were preserved separately in the field by using 99% ethanol. After identification of species, the respective vial containing the larvae were considered for molecular studies. From the vial, two to ten mature larvae taken from the respective species and went for DNA barcoding studies.

PCR amplification

Total genomic DNA was extracted from individual larvae from each species according to the manufacturer's protocol for the QIAamp genomic DNA isolation kit (QIAGEN, Germany). The extracted genomic DNA was quantified by a spectrophotometer (Shimadzu, Japan) at 260/280 nm. The mitochondrial protein coding gene of cytochrome c oxidase subunit I (COI) was amplified by polymerase chain reaction according to Anbalagan *et al.* (2015). The amplified PCR products were identified by electrophoresis, using a 1% agarose gel followed by purified, and sequenced. Sequence alignments were performed using the Clustal W v.1.82 software (Thompson *et al.*, 1994). The sequences of seven species have been deposited in the GenBank database. In addition, sequences deposited in GenBank, which related members of *Gomphostilbia* were included as outgroups (Table 1).

Data analyses

A total of 28 COI sequences of *Simulium* (black flies) were selected for this study. Sequences of *Simulium* species of Indian sub-continent peninsula and some species from South Asian countries were selected for this study. Extensive and core sequence analysis tools of BOLD system was used to analyze and calculate both intraspecific and interspecific genetic divergence values based on the Kimura 2-parameter (K2P) model. All species barcodes option was used to test the frequency of successful identification of sequence for taxonomical study in the Java applet enabled program TaxonDNA v1.0 (Meier *et al.*, 2006). The unraveled diversity in *simulium* black flies were analyzed using methods, that designated operational taxonomic units based on the refined single-linkage algorithm called the Barcode Index Number System (BINs) (Ratnasingham and Hebert, 2013) in BOLD, this system automatically assigns COI sequences into the BIN numbers. Neighbor-joining (NJ), maximum likelihood (ML) and Bayesian (BA) methods were used to establish the close evolutionary relationship between close haplotypes. All phylogenetic analyses were implemented in MEGA X (Tamura *et al.*, 2013). Inter branch tree support was used to calculate the initial tree optimization using the bootstrapping method with 1,000 replicates. Node support was calculated using approximate likelihood ratio tests (Anisimova and Gascuel, 2006). Mega X (Kumar *et al.*, 2018) was used to select the best-fit DNA substitution model for the ML analysis based on the Akaike information criterion (AIC) and Bayesian Information Criterion (BIC). The best-fit model was the Tamura and Nei (TrN) model with a proportion of invariable sites of 0.1501 and with a gamma distribution of 0.2705 (Huelsenbeck and Ronquist, 2001).

RESULTS AND DISCUSSION

Genetic variation between Peninsular Indian members

Of 20 named species of *Gomphostilbia* in India, a total of

seven South Indian described species was taken for the present study. The intra-specific specific variation was found to be null as there was no same species duplication or strain repletion was taken for this study. The variation of this range is due to the haplotypic and genetic variation among members of *Gomphostilbia*. The interspecies species distance shared between species is 73%. This range of distancing serves as the middle ground for all the species similarity index. The less diverse species falls under the range of 0.5% to 1.0%, which consists of *S. (G.) takaokai* and *S. (G.) panagudiense* with their interspecific distance of 0.64%. The highest match occurred between *S. (G.) panagudiense*, *S. (G.) kottooreense* and *S. (G.) krishnani* with interspecific distance of 78.1%. The results prove the interspecific diverse distances of seven selected species of Indian subcontinent. Further, *S. (G.) takaokai* and *S. (G.) krishnani* are two diverse species which has diversified later among the species of *Gomphostilbia* members of the Peninsular India.

Genetic variation between members of South Asia

A total of 28 outgroup taxa including seven Indian species was taken for the phylogenetic analysis (Table 1.). The intra-specific specific variation was found to be null as there was no same species duplication or strain repletion was taken for this study. On the other hand, interspecific species variation was found to be 24.33% and 44.17% with a cumulative score of 76.45. The extreme pairwise sequence distance showed a result of *S. (G.) krishnani* and *S. (G.) johoreense* at highly distant with a score of 68.65 which indicates the close dissimilarity between the species. *S. (G.) panagudiense* and *S. (G.) takaokai* were the highly conserved and with least distant species in the cluster, with a score of 0.64. Results of the cluster with maximum number of specific hits per cluster using the Bayesian and

Table 1: Indian and out group taxa of the subgenus *Gomphostilbia*

Species	Accession no.	Country
<i>Simulium agasthyamalaiense</i>	MG757145	India
<i>Simulium angulistylum</i>	MF476247	Malaysia, Thailand
<i>Simulium asakoae</i>	MF101846	Malaysia, Burma, China, Thailand, Vietnam
<i>Simulium atratum</i>	MF476249	Indonesia
<i>Simulium cheongi</i>	MN514678	Malaysia, Indonesia, Thailand
<i>Simulium chiangdaoense</i>	LC472509	Thailand
<i>Simulium chiangraiense</i>	MF968964	Thailand
<i>Simulium decuplum</i>	MF968961	Malaysia, Burma, India
<i>Simulium dinakarani</i>	MG700551	India
<i>Simulium duolongum</i>	MK015718	Malaysia, Indonesia, Vietnam
<i>Simulium gombakense</i>	MG958564	Malaysia, Burma, Thailand
<i>Simulium huaikaoense</i>	MF968958	Thailand
<i>Simulium johoreense</i>	KY751930	Malaysia
<i>Simulium khaokhoense</i>	MF968969	Thailand
<i>Simulium kottooreense</i>	KP223707	India
<i>Simulium krishnani</i>	MG700552	India
<i>Simulium maeklangense</i>	MF968967	Thailand
<i>Simulium maleewongae</i>	MG958584	Thailand
<i>Simulium myanmarensis</i>	MF101842	Burma, Thailand
<i>Simulium paiense</i>	MG958580	Thailand
<i>Simulium pamiangense</i>	MF968953	Thailand
<i>Simulium panagudiense</i>	KP031495	India
<i>Simulium peteri</i>	KM977779	India
<i>Simulium rampae</i>	LC472508	Thailand
<i>Simulium takaokai</i>	KM985984	India
<i>Simulium thuathienense</i>	MG958582	Vietnam, Thailand
<i>Simulium whartoni</i>	MF476257	Malaysia, Indonesia
<i>Simulium yvonneae</i>	MH899077	Vietnam

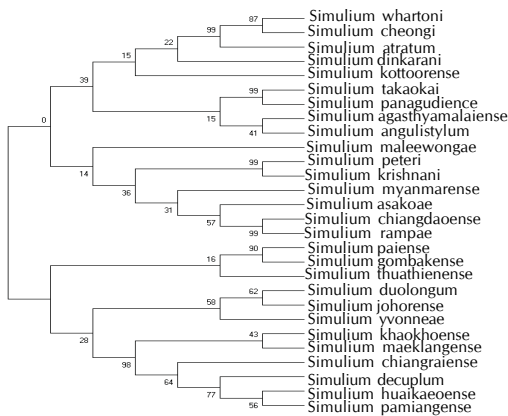


Figure 1: Neighbor joining tree for cytochrome oxidase I (COI) sequences of 28 species of *Simulium* black flies and members of their species-groups in Indian subcontinent and South Asian countries.

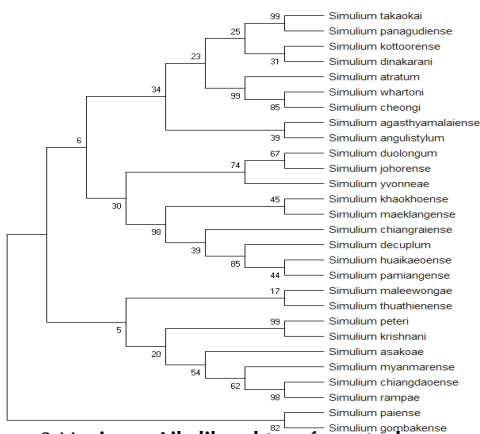


Figure 2: Maximum Likelihood tree for cytochrome oxidase I (COI) sequences of 28 species of *Simulium* black flies and members of their species-groups in Indian subcontinent and South Asian countries.

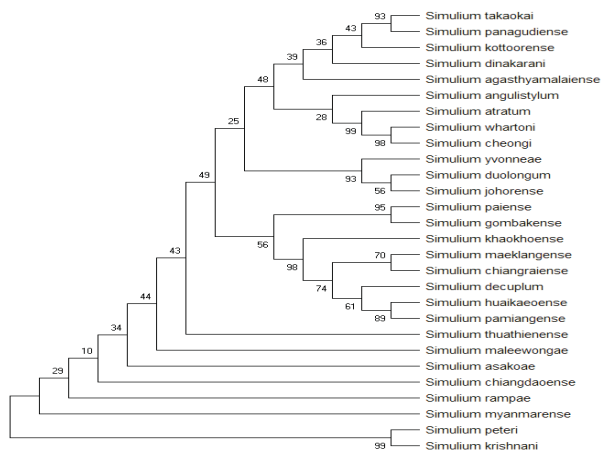


Figure 3: Bayesian tree for cytochrome oxidase I (COI) sequences of 28 species of *Simulium* black flies and members of their species-groups in Indian subcontinent and South Asian countries.

Kimkura -2- model algorithms. The maximum number of species fall under the distance score range of 75%, ensuring that all the barcodes rightly belong to the *Gomphostilbia*. The other clusters are two extremes categories of highly distant and least distant species of the categories.

The highly distant clusters consist of *S. (G.) cheongi* and *S. (G.) chiangdaoense* with an interspecific distance of 78.52%. *S. (G.) decuplum*, *S. (G.) huaikaeoense*, *S. (G.) pamiangense* and *S. (G.) peteri* with an interspecific distance of 78.71%. The highly distant cluster clearly showcases the region specific species variations. All the highly distant species represent, different prominent species of Thailand, Malaysia and India. *S. (G.) cheongi* and *S. (G.) chiangdaoense* represent Malaysian and Thailand species diversity. *S. (G.) decuplum*, *S. (G.) huaikaeoense*, *S. (G.) pamiangense* and *S. (G.) peteri* defines the species variation between Indian sub-continent and south east Asian countries. The closely related species cluster represented a total of four species with two species representing each from Thailand and India.

S. (G.) panagudiense and *S. (G.) takaokai* from India with an inter species distance of 0.6421%. *S. (G.) huaikaeoense* and *S. (G.) pamiangense* from Thailand with inter species distance of 0.9901%. These four species are highly conserved and appear at two distant ends of the clads representing their highly conserved nature and their geographical diversity. The pairwise sequence representation results also correlated with the phylogenetic analysis of all the 28 species from various regions of India and South East Asia. The phylogenetic trees which were constructed based on the different algorithms clearly indicated the species variation among the countries and its similarities between geographical locations.

The rooted Neighbor joining tree, ML tree and Bayesian tree (Figs. 1, 2 & 3) also represented the common clad to be arising from *Gomphostilbia* members of India and diverging to the other South-east Asian countries species. Comparing the species similarities among the Indian members of *Gomphostilbia* revealed the highly similar and highly conserved pattern of divergence. On the contrary, when compared with South-East Asian species the conservation was highly reduced and the geographic divergence was widely observed. Thus our study emphasizes the DNA barcode similarities and species divergence among Indian and South East Asian countries species.

Simulium (Gomphostilbia) Enderlein, the third largest in the genus *Simulium* Latreille s. l., is one of the two most abundant and diverse subgenera in the Oriental Region (Takaoka, 2012), where about 10% of total species have been recorded in India. The subgenus *Gomphostilbia* was first defined morphotaxonomically by Crosskey (1967), and its definition has been modified by various black fly taxonomists (Datta, 1973; Davies and Gyorkos 1987; Takaoka and Davies, 1996). The subgenus *Gomphostilbia* is characterized by the combination of the haired katepisternum and the bare pleural membrane of the adult thorax, by which it was easily distinguished from the related subgenera *Hebridosimulium*, *Inseliellum*, *Morops*, *Nevermannia* and *Simulium*.

The highest similarity occurred between *S. (G.) panagudiense*, *S. (G.) kottoorensae* and *S. (G.) krishnani* with interspecific

distance of 78.1%. Of seven species selected, the above three species are related each other. The morphological characters of these three species revealed the similar outcome that characters like stalk of ventral paired filaments in pupae and abdominal segments without dorsal pair of conical protuberances shared the common characters. The less similarity found between *S. (G.) takaokai* and *S. (G.) krishnani*, it may have spatial heterogeneity.

However, inter-specific variation between species is related with spatial factor, which may determine the diversity of genes of *Gomphostilbia* species

The fauna of tropical regions have generally greater biodiversity, compared with temperate regions (Moritz and Cicero 2004). Numerous hypotheses have been projected to elucidate this latitudinal gradient (Willig *et al.*, 2003) at the genetic level, as faster rates of molecular evolution in the tropics (Wright *et al.*, 2006). The present study shows that latitudinal gradient play an important role to determine the diversity of species. The phylogenetic trees emphasize the lineage and origin of species, reflects that Indian species is primitive and their origin wander to south-east Asian countries and this finding agree with the results of previous molecular and morphological phylogenies (Anbalagan *et al.*, 2017).

ACKNOWLEDGEMENTS

We thank Science and Engineering Research Board, New Delhi (Ref. No: ECR/2016/000191/LS) for financial assistance.

REFERENCES

- Adler, P.H. 2020. World blackflies (Diptera: Simuliidae): a comprehensive revision of the taxonomic and geographical inventory [2020].
- Anbalagan, S., Arun Prasanna, V., Dinakaran, S. and Krishnan, M. 2014. Two new species of *Simulium* (*Gomphostilbia*) (Diptera: Simuliidae) from Peninsular India with keys to Peninsular Indian members of the genus *Simulium*. *Zootaxa*. **3861**(5): 451-465.
- Anbalagan, S., Kannan, M., Dinakaran, S. and Krishnan, M. 2017. Species diversity of black flies (Diptera: Simuliidae) in Oriental region and molecular phylogeny of the subgenus *Gomphostilbia* members. *J. Vector Borne diseases*. **54**: 80-86.
- Anbalagan, S., Kannan, M., Dinakaran, S., Balasubramanian, C. and Krishnan, M. 2018. Hierarchical dynamics influence the distribution of immature black flies (Diptera: Simuliidae). *Acta Tropica* **177**: 105-115.
- Anbalagan, S., Pratheep, T., Dinakaran, S. and Krishnan, M. 2012. Effects of two leaf litter species on the colonization of macroinvertebrates in a tropical stream of India. *The Bioscan*. **7**(3): 533-538.
- Anisimova, M. and Gascuel, O. 2006. Approximate likelihood-ratio test for branches: a fast, accurate, and powerful alternative. *Systematic Biology*. **55**: 539-552.
- Colebunders, R., Hendy, A., Nanyunja, M., Wamala, J.F. and Van Oijen M. 2014. Nodding syndrome - a new hypothesis and new direction for research. *International J. Infectious Diseases*. **27**: 74-7.
- Crosskey, R.W. 1967. The classification of *Simulium* Latreille (Diptera: Simuliidae) from Australia, New Guinea and the Western Pacific. *J. Natural History*. **1**: 23-51.
- Datta, M. 1973. New species of black flies (Diptera: Simuliidae) of the subgenera *Eusimulium* Roubaud and *Gomphostilbia* Enderlein from the Darjeeling area, India. *Oriental Insects*. **7**: 363-402.
- Davies, D.M. and Györkös, H. 1987. The Simuliidae (Diptera) of Sri Lanka. Descriptions of species in the subgenera *Eusimulium* and *Gomphostilbia* of the genus *Simulium*. *Canadian J. Zoology*. **65**: 1483-1502.
- Dinakaran, S. and Anbalagan, S. 2007. Modern trends for assessment of forest streams and rivers of southern Western Ghats using Caddisflies. *The Bioscan*. **2**(2): 109-112.
- Figueiró, R., Gil-Azevedo, L.H., Maia-Herzog, M. and Monteiro, R.F. 2012. Diversity and microdistribution of black fly (Diptera: Simuliidae) assemblages in the tropical savanna streams of the Brazilian cerrado. *Mem. Inst. Oswaldo Cruz*. **107**(3): 362-369.
- Hernández-Triana, L.M., Montes de Oca, F., Prosser, S.W., Hebert, P.D., Gregory, R.T. and McMurtrie, S. 2017. DNA barcoding as an aid for species identification in austral black flies (Insecta: Diptera: Simuliidae). *Genome*. **60**: 348-57.
- Huelsenbeck, J.P. and Ronquist, F.R. 2001. MrBayes: Bayesian inference of phylogeny. *Bioinformatics*. **17**: 754-755.
- Kumar, S., Stecher, G., Li, M., Knyaz, C. and Tamura, K. 2018. MEGA X: Molecular Evolutionary Genetics Analysis across computing platforms. *Molecular Biology and Evolution*. **35**: 1547-1549.
- Meier, R., Shiyang, K., Vaidya, G. and Ng, P.K.L. 2006. DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. *Systematics Biology*. **55**: 715-728.
- Monaghan, M.T., Balke, M., Gregory, T.R. and Vogler, A.P. 2005. DNA-based species delineation in tropical beetles using mitochondrial and nuclear markers. *Philos Trans R Soc Lond B Biol Sci*. **360**(1462): 1925-33.
- Moritz, C. and Cicero, C. 2004. DNA barcoding: promise and pitfalls. *PLoS Biology*, **2**: e354.
- Ratnasingham, S. and Hebert, P.D. 2013. A DNA-based registry for all animal species: The Barcode Index Number (BIN) System. *PLoS One*. **8**(7): e66213.
- Ruiz-Arondo, I., Hernández-Triana, L.M., Ignjatović-Āepina, A. et al. 2018. DNA barcoding of blackflies (Diptera: Simuliidae) as a tool for species identification and detection of hidden diversity in the eastern regions of Spain. *Parasites Vectors*. **11**: 463 pp.
- Takaoka, H. 1996. The geographical distribution of the genus *Simulium* Latreille in the Oriental and Australasian Regions. *Japanese J. Tropical Medicine and Hygiene*. **24**: 113-124.
- Takaoka, H. 2012. Morphotaxonomic revision of *Simulium* (*Gomphostilbia*) (Diptera: Simuliidae) in the Oriental Region. *Zootaxa*. **3577**: 1-42.
- Tamura, K., Stecher, G., Peterson, D., Filipiński, A. and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. *Molecular Biology and Evolution*. **30**: 2725-2729.
- Willig, M.R., Kaufman, D.M. and Stevens, R.D. 2003. Latitudinal gradients of biodiversity: pattern, process, scale, and synthesis. *Annual Review of Ecology, Evolution, and Systematics*. **34**: 273-309.
- Wright, S., Keeling, J. and Gillman, L. 2006. The road from Santa Rosalia: a faster tempo of evolution in tropical climates. *Proceedings of the National Academy of Sciences, USA*, **103**: 7718-7722.