Ceylon Journal of Science 51(3) 2022: 319-323
DOI: http://doi.org/10.4038/cjs.v51i3.8039

SHORT COMMUNICATION

Genetic diversity of *Phlebotomus* (*Euphlebotomus*) *argentipes* species complex in Sri Lanka

P. Sharmintha, K. Gajapathy*, A.A.K.P. Amarasinghe and S.N. Surendran

Department of Zoology, University of Jaffna, Jaffna, Sri Lanka.

Received: 28/07/2021; Accepted: 15/08/2022

Abstract: Leishmaniasis has been recognized in Sri Lanka for many years and is becoming a serious public health threat. Sandflies, the known vector of leishmaniasis, are also found throughout the country. The presence of a species complex among the potential vector species; *Phlebotomus* (*Euphlebotomus*) *argentipes* sensu lato is one of the impediments to developing an effective control strategy. The current study assessed the genetic diversity of the species complex and compared the results to morphology-based taxonomy. Sandflies were collected from selected study sites in the Anuradhapura, Hambantota, Puttalam, and Jaffna districts between 2015 and 2018. They were identified morphologically using existing taxonomic keys. The PCR assay and subsequent DNA sequence analysis were performed for the cytochrome c oxidase subunit I (*COI*) gene. The morphological analysis revealed that the species complex contains more than two morphospecies. Nucleotide variation was observed in the *COI* sequences that were aligned. The uncorrected p distances within and between groups were 0.01%-0.08% and 2.1%, respectively. The *COI* sequence-based phylogenetic tree revealed the presence of two genetically distinct groups: one with a sensilla chaetica/second antennal flagellomere length ratio of less than 0.4 and another with a ratio of 0.45 and greater. This result contradicts the existing taxonomic keys. A PCR-based assay is proposed to differentiate the genetic groups based on the sequence variations.

***Keywords*:** *Argentipes* complex; morphospecies; molecular taxonomy; PCR.

INTRODUCTION

Leishmaniasis was made a notifiable disease in 2009 in Sri Lanka. Since then, the number of cases reported was more than a thousand in Hambantota and Anuradhapura districts (Amarasinghe and Wickramasinghe, 2020; Karunaweera *et al*., 2021). Matara, Kurunagala, and Polonnaruwa districts were reported with more than 500 cases. Few cases were notified in the Jaffna District as well (Galgamuwa *et al*., 2018). Several leishmaniasis outbreaks were reported in Matara, Anuradhapura, Hambantota, Kurunagala, Polonnaruwa and Mullaitivu districts (Semage, 2014; Karunaweera *et al*., 2021). According to the Epidemiology Unit of Sri Lanka, in 2018, 3271 cases were reported and 90% of them were from these districts. *Leishmania donovani* MON-37 was reported as the causative organism for both cutaneous leishmaniasis (CL) and visceral leishmaniasis (VL) in Sri Lanka. The suspected vector for *L. donovani* in Sri Lanka is *Phlebotomus* (*Euphlebotomus*) *argentipes* (Gajapathy *et al*., 2013) which exists as a species complex (Ilango, 2010; Gajapathy *et al*., 2011).

Insects have the most diverse species composition of any animal group. They may contain a large amount of previously unknown genetic diversity. With appropriate molecular markers, this genetic polymorphism can be identified precisely. Behura (2006) described how DNA markers such as mitochondrial DNA (mtDNA), microsatellites, Random Amplified Polymorphic DNA (RAPD), Expressed Sequence Tags (EST), Amplified Fragment Length Polymorphism (AFLP), transposon display, Sequence-Specific Amplification Polymorphism (S-SAP), and repeat-associated Polymerase Chain Reaction (PCR) have contributed significantly in understanding the genetic basis of insect diversity. As each marker has inherent advantages and disadvantages, the option of applying them depends on the objective of a study. Identification of the species, with limited morphological variations, is very hard. These morphologically obscure species are found in various insect vectors, including the sandfly species that transmit leishmaniasis. Distinguishing them solely based on their morphology is a difficult task (Struck and Cerca, 2019). Additionally, the extent to which these cryptic species differ in their DNA sequences is debatable (Muller, 2000), although numerous regions of the genome, including nuclear and mitochondrial DNA, can be used to discriminate between species. Due to its maternal origin, haploid status, and rapid evolution, mitochondrial DNA is primarily used as a molecular marker (Behura, 2006). Additionally, one advantage of using mitochondrial markers in insects is that the majority of the loci associated with the mtDNA can be amplified easily using universal primers (Behura, 2006).

The presence of more than one morphospecies among the *Argentipes* complex has been reported in Sri Lanka (Surendran *et al*., 2005; Gajapathy *et al*., 2011). The current study was designed to characterize the species complex using DNA sequence variations in the cytochrome c oxidase subunit I (*COI*) region. Additionally, a PCR-based diagnostic assay based on sequence variations in the *COI* gene is proposed, which will be helpful for the identification of different genetic groups that exist in the country.

**Corresponding Author's Email:* gajapathyk@univ.jfn.ac.lk; gayan156@gmail.com
https://orcid.org/0000-0001-5728-6319

MATERIALS AND METHODS

Sample collection and morphological characterization

Phlebotomus (*Eup.*) *argentipes* flies were collected from selected study sites in Anuradhapura, Hambantota, Puttalam, Jaffna, and the associated islands (Delft and Pungudutivu) between 2015 and 2018 using cattle baited net and cattle baited hut. The sampling was done from 6 pm to 4 am every month. The flies were preserved in 70% ethanol. Flies were temporarily mounted on glass slides and were identified using established taxonomic keys (Lewis, 1978; Lane, 1992; Ilango, 2000; 2010). The morphometric and meristic features such as total length of the wings, the maximum width of the wing, length of radial vein 1 (R1), radial vein 2 (R2), radial vein 3 (R3), R2+R3, wing overlap, and the lengths and widths of the male genitalia were recorded. The morphologically identified flies were preserved in 70% alcohol individually for molecular analysis.

Designing the primers to amplify a part of the *COI* gene

The complete mitochondrial genome sequence of *Phlebotomus* (*Phlebotomus*) *papatasi* was obtained from the NCBI database (NCBI Reference Sequence: NC 028042.1). To design the primers, the *COI* gene was chosen from the complete mitochondrion genome sequence. The primer set for the *COI* sequence was designed to amplify a 715 base region. The primers were designed considering the characters such as the length (18 bp to 22 bp), melting temperature (55°C -60°C), and GC content (40% to 60%).

DNA extraction from the identified morphospecies

Both adult males and females of *Ph.* (*Eup.*) *argentipes* were used for DNA extraction using a commercially available DNA extraction kit (DNeasy® Blood and Tissue Kit, QIAGEN®), according to the manufacturer's instructions. The morphospecies was randomly chosen for the DNA extraction after they were classified based on the classification scheme proposed by Ilango (2000 and 2010). A total of 20 flies were selected from each sampling location. Extracted DNA samples were labelled and kept in the freezer at -20 °C. The purity and concentration of DNA were estimated using a Nanodrop spectrophotometer by measuring the absorbance of DNA at 260 nm and 280 nm. The purity of the extracted DNA was also determined by calculating the absorbance ratio at these wavelengths.

Polymerase Chain Reaction (PCR) assay to amplify *COI* gene

The designed *COI* forward primer: **F**: 5' GTTCTTGC GGGGGCTATTACT 3' and the reverse primer: **R**: 5' GA CGAGGTATTCCAGCTAGACC3' were used to amplify a part of the *COI* gene in mtDNA. The PCR reactions were performed in 25 µl volumes of PCR reaction mixture. Each reaction mixture contained 1 µl of extracted DNA, 0.5 µM of each primer, 2.5 mM of $MgCl_2$, 0.2 mM of dNTP mix, and 1.25 U of Taq DNA polymerase in 1x PCR buffer (Bioline, UK). The thermocycler was set to 94 °C for five minutes as an initial denaturation temperature. 35 cycles of amplification at 94 °C for 40 seconds, 45 °C for one minute, and 72 °C for one minute were performed, followed by a seven-minute final extension at 72 °C.

DNA sequencing and sequence analysis

Thirty four PCR products representing all the sampling locations and a range of SC2/AF2 ratios were sent for bi-directional Sanger sequencing (m/s Macrogen, South Korea). Good quality sequences were taken for the analysis based on the amount of baseline noise and base call quality. The ambiguous sequence patches were removed. MEGA 6.0 (Tamura *et al*., 2013) software was used to analyze the sequences. The edited sequences were aligned using the Clustal W alignment tool with default settings. Trimmed and aligned sequences were used to select the best tree building model using MEGA 6.0 (Tamura *et al*., 2013) based on the lower Bayesian Information Criterion (BIC) value. Aligned sequences were translated into amino acid sequences to identify that they are coding and there are no nuclear pseudo genes present. The translated sequences were checked in the NCBI protein blast programme (Blastp) to validate the sequences.

Phylogenetic tree construction

Edited sequences (n=34) were selected for the maximum likelihood phylogenetic tree construction using the Tamura-Nei model with 100 bootstrapping in MEGA 6.0 (Tamura et al, 2013). *Phlebotomus papatasi* (Gen Bank Acc. No: KR349298.1) and *Phlebotomus chinensis* (Gen Bank Acc. No: KR349297.1) (Ye *et al*., 2015) were used as outgroup.

Development of diagnostic PCR assay

The variable region from *COI* sequences was selected to design the diagnostic primers. A forward primer: F 5' GTG CGG GTT TTG CAT TAC TT 3' and a reverse primer: R 5' AAC ATC CTT GGC AAA TGT T 3' were designed using Primer 3 version 0.4.0. The expected size of the PCR product was ~450 bp. The PCR was performed in a 25 µl reaction mixture. Each reaction mixture contained 1.25 U of Taq DNA polymerase in 1x PCR buffer (QIAGEN®), 1 µl of DNA, 1.5 µM of each from the forward and reverse primers, 0.2 mM of dNTP mix, and 2.5 mM of $MgCl_2$. The samples were heated at 94°C for five minutes before 35 cycles of amplification at 94°C for 40 seconds, 50.4°C for one minute, and 72°C for one minute followed by a final extension at 72°C for seven minutes.

RESULTS AND DISCUSSION

A total of 950 sandflies were processed. Three morphospecies were distinguished based on the length ratio of sensilla chaetica (SC2) on antennal flagellomere 2 (AF2) and male genitalia characters as described by Ilango (2000 and 2010). All three morphospecies have been reported previously in Sri Lanka (Surendran *et al*., 2005; Gajapathy *et al*., 2011; Ranasinghe *et al*., 2012).

The *COI* amplicon was 715bp in length when designed primers were used. This is the first time the specified *COI* region has been amplified and studied. Since all previously studied regions of the *COI* demonstrated

extremely low variation, a new region was chosen for this study. Melting temperatures for forward and reverse primers are 60.13 °C and 59.71 °C, and the GC content for the forward and reverse primers are 52.38% and 54.55%, respectively.

COI sequences were found to vary among the individuals in the study (Genbank Acc.No. MZ452173-MZ452206). The uncorrected p distances within the groups were 0.01%-0.08% and between groups was 2.1%. The distanced value suggests that there are minimum variations found among the genetic groups. The low interspecific variations suggest the potential gene flow between the groups. The converted amino acid sequences were checked against the reference sequence, and all were coded, and no pseudo genes were found in the sequence. The constructed phylogenetic tree is presented in Figure 1.

The phylogenetic tree for *COI* reveals that there were two genetically different groups (2 clades with strong bootstrap support (>70%) existing in the *Argentipes* complex in Sri Lanka. The SC2/AF2 ratios of the two genetically distinct groups were determined. One clade contained sandflies with an SC2/AF2 ratio of less than 0.4, while the other contained sandflies with an SC2/AF2 ratio of 0.45 or greater. It's interesting because previous classifications based on SC2/AF2 values (Ilango, 2000) classified morphospecies A as sandflies with an SC2/AF2 ratio greater than 0.5, whereas morphospecies B was classified as sandflies with an SC2/AF2 ratio less than 0.5, which is different from what was observed from the phylogenetic tree grouping in this study. Sandflies were randomly selected for molecular analysis in this study after they were classified according to Ilango's (2000) taxonomic classifications. As a result, sandflies with an SC2/AF2 ratio of 0.41–0.44 were excluded from the phylogenetic tree construction. Thus, using the measured morphometric characters, a principal component analysis (PCA) was performed in Minitab statistical software (Arend, 1993) to cross-check the morphometry-based classification based on the results of the phylogenetic tree analysis. The PCA analysis was performed based on the following two assumptions:

Assumption 1:

If the SC2/AF2 ratio is greater than 0.4, sandflies are classified as morphospecies A; if the ratio is less than 0.4, they are classified as morphospecies B.

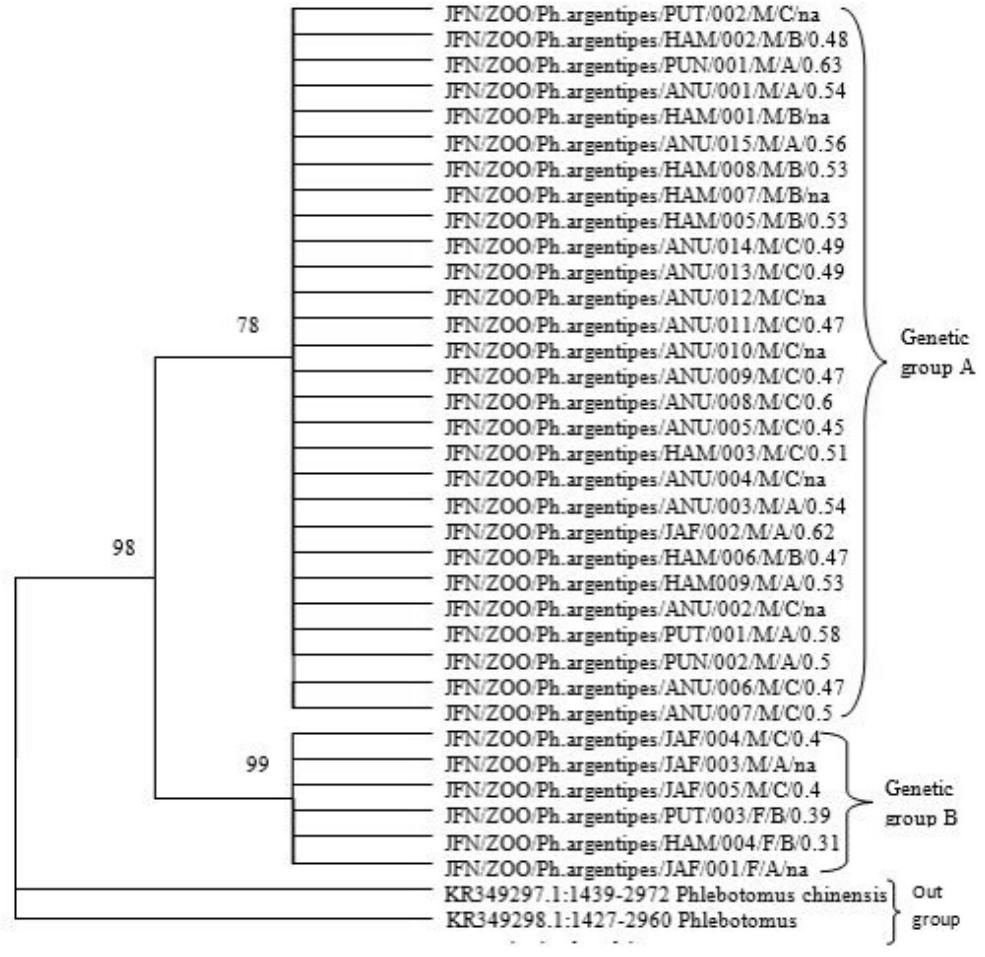

Figure 1: Molecular phylogenetic analysis by maximum likelihood method; Tamura-Nei model was selected based on the lowest BIC (Bayesian Information Criterion) scores. Non-parametric bootstrapping was performed (100 replicates) to validate the tree.

Molecular phylogenetic analysis by maximum likelihood method; Tamura-Nei model was selected based on the lowest BIC (Bayesian Information Criterion) scores. Non-parametric bootstrapping was performed (100 replicates) to validate the tree.

Assumption 2:

If the ratio SC2/AF2 is greater than 0.45, sandflies are classified as morphospecies A; if the ratio is less than 0.45, they are classified as morphospecies B.

The score plots obtained from the PCA analysis are given in Figure 2.

According to these score plots (Figures 2a and b) created based on the two assumptions, there was no clear grouping observed. Thus, it can be assumed that the classification scheme proposed based on these two assumptions are invalid. Hence, it can be confirmed that two genetic groups exist in the male *Ph*. (*Eup*.) *argentipes* species complex in the country, and this does not correspond with the current morphometry-based classification schemes (Ilango, 2000; 2010) practiced. Since the variation in the *COI* sequence was not high, diagnostic primers were designed to produce amplicons from genetic group B alone. The expected size of the amplicon was ~450 bp. The bands were obtained for the annealing temperatures of 44.0 °C, 44.3 °C, 45.1°C, 46.4 °C, 48.2 °C, and 50.4 °C in the gradient PCR. The bands obtained at 48.2 °C and 50.4 °C were considered ideal in quality and the latter was chosen as the annealing temperature. In the diagnostic assay performed on sandflies collected from various locations, no amplicons for genetic group A were observed, whereas an amplicon of 450bp was obtained for group B. (Figure 3). This is the first attempt to

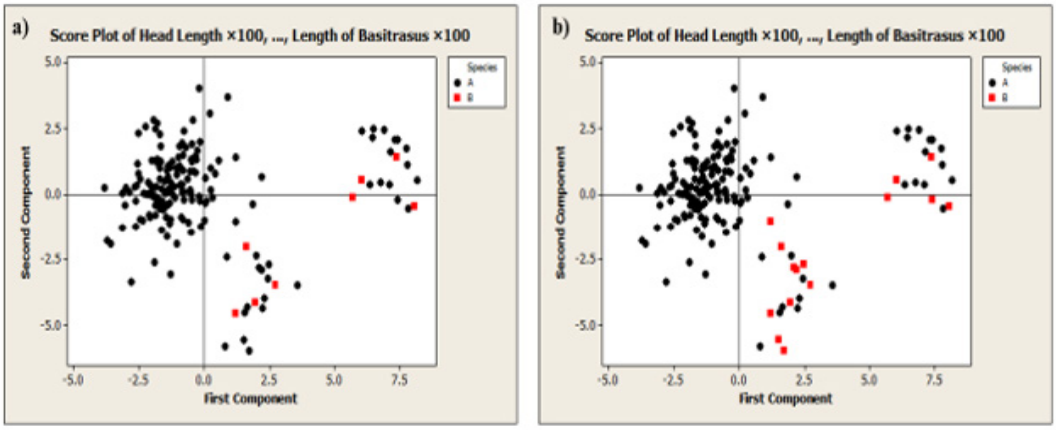

Figure 2: Score plots (**a)** Assumption 1*: **b)** Assumption 2*) from the PCA analysis for all measured morphological characters of male *Phlebotomus* (*Euphlebotomus*) *argentipes* in the studied sand fly population.

Score plots (**a)** Assumption 1*: **b)** Assumption 2*) from the PCA analysis for all measured morphological characters of male *Phlebotomus* (*Euphlebotomus*) *argentipes* in the studied sand fly population.

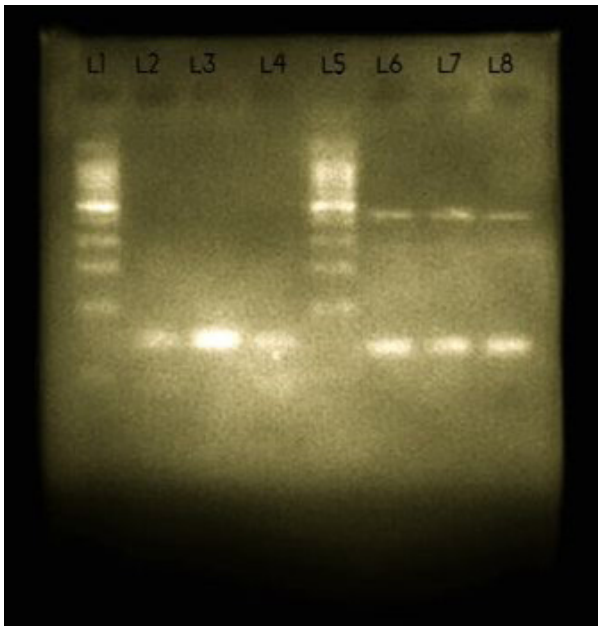

Figure 3: The 1.5% Agarose gel image stained with Ethidium bromide for the diagnostic fragments of morphospecies A and morphospecies B (Lane 1(L1) and L5 – 100bp ladder, L2-L4 PCR products for morphospecies A; L6-L8 PCR products for morphospecies B).

The 1.5% Agarose gel Image stained with Ethidium bromide for the diagnostic fragments of morphospecies A and morphospecies B (Lane 1(L1) and L5 – 100bp ladder, L2-L4 PCR products for morphospecies A; L6-L8 PCR products for morphospecies B).

develop a diagnostic PCR-based assay for separating the members from the *Argentipes* complex. When more genes other than *COI* are analyzed for sequence variations, it is possible to separate the genetic groups using a PCR-based diagnostic marker.

CONCLUSION

The phylogenetic tree constructed using the *COI* sequence revealed the existence of two genetically different groups among *Ph.* (*Eup.*) *argentipes* species complex. Diagnostic PCR assay which was developed based on the variations in the *COI* sequence was able to separate the genetic group B from the pool of sandflies belong to the *Argentipes* complex.

ACKNOWLEDGMENT

The financial support by the NSF through the grant (NSF/RG/2015/HS/04) is acknowledged

DECLARATION OF CONFLICT OF INTEREST

The authors declare that there is no conflict of interest

REFERENCES

Amarasignhe, A and Wickramasinghe, S. (2020). A Comprehensive Review of Cutaneous Leishmaniasis in Sri Lanka and Identification of Existing Knowledge Gaps. *Acta Parasitologica*, **65**:300–309. DOI: https://doi.org/10.2478/s11686-020-00174-6.

Arend, D.N. (1993). Choices (Version 4.0) [Computer software]. Champaign, IL: U.S. Army Corps of Engineers Research Laboratory. (CERL Report No.CH7-22510).

Behura, S. (2006). Molecular marker systems in insects: current trends and future avenues. *Molecular Ecology*, **15**: 3087-3113. DOI: https://doi.org/10.1111/j.1365-294X.2006.03014.x.

Gajapathy, K., Jude, P.J. and Surendran, S.N. (2011). Morphometric and meristic characterization of *Phlebotomus argentipes* species complex in northern Sri Lanka: evidence for the presence of potential leishmaniasis vectors in the country. *Tropical Biomedicine*, **28**(2): 259-268.

Gajapathy, K., Peiris, L.B., Goodacre, S.L., Silva, A., Jude, P.J., and Surendran, S. N. (2013). Molecular identification of potential leishmaniasis vector species within the *Phlebotomus* (*Euphlebotomus*) *argentipes* species complex in Sri Lanka. *Parasites & Vectors*, **6**(1), 1-9. DOI: https://doi.org/10.1186/1756-3305-6-302.

Galgamuwa, L.S., Dharmaratne, S.D., and Iddawela, D. (2018). Leishmaniasis in Sri Lanka: spatial distribution and seasonal variations from 2009 to 2016. *Parasites & Vectors*, **11**(1), 1-10. DOI: https://doi.org/10.1186/s13071-018-2647-5.

Ilango, K. (2000). Morphological characteristics of the antennal flagellum and its sensilla chaetica with character displacement in the sandfly *Phlebotomus argentipes* Annandale and Brunetti sensu lato (Diptera: Psychodidae). *Journal of Biosciences*, **25**:163-172. DOI: https://doi.org/10.1007/BF03404911.

Ilango, K. (2010). A taxonomic reassessment of the *Phlebotomus argentipes* species complex (Diptera: Psychodidae: Phlebotominae). *Journal of Medical Entomology*, **47**:1-15. DOI: https://doi.org/10.1603/033.047.0101. PMID: 20180302.

Karunaweera, N.D., Senanayake, S., Ginige, S., Silva, H., Manamperi, N., Samaranayake, N., Dewasurendra, R., Karunanayake, P., Gamage, D., de Silva, N., Senarath, U. and Zhou, G. (2021). Spatiotemporal distribution of cutaneous leishmaniasis in Sri Lanka and future case burden estimates. *PLOS Neglected Tropical Diseases*, **15**(4): p.e0009346. DOI: https://doi.org/10.1371/journal.pntd.0009346.

Lane, R.P. (1992). Sandflies (Phlebotominae). Medical insects and arachnids. *Chapman and Hall*. pp. 78-119. DOI: https://doi.org/10.1007/978-94-011-1554-4_4.

Lewis, D.J. (1978). Phlebotamine sandflies (Diptera, Psychodidae) of the oriental region. *Bulletin of the British Museum* (*Natural History*), **37**: 218-331. DOI: https://doi.org/10.1111/j.1365-3113.1987.tb00194.x.

Muller, J. (2000). Mitochondrial DNA variation and the evolutionary history of cryptic *Gammarus fossarum* types. *Molecular Phylogenetics and Evolution*, **15**: 260-268. DOI: https://doi.org/10.1006/mpev.1999.0740.

Ranasinghe, S., Maingon, R.D., Bray, D.P., Ward, R.D., Udagedara, C., Dissanayake, M., .and de Silva, N. K. (2012). A morphologically distinct *Phlebotomus argentipes* population from active cutaneous leishmaniasis foci in central Sri Lanka. *Memórias do Instituto Oswaldo Cruz*, **107**(3), 402-409. DOI: https://doi.org/10.1590/s0074-02762012000300016.

Semage, S.N., Pathirana, K.P.N., and Agampodi, S.B. (2014). Cutaneous leishmaniasis in Mullaitivu, Sri Lanka: a missing endemic district in the leishmaniasis surveillance system. *International Journal of Infectious Diseases*, **25**, 53-55. DOI: https://doi.org/10.1016/j.ijid.2014.03.1382.

Surendran, S.N., Kajatheepan, A., Hawkes, N.J. and Ramasamy, R. (2005). First report on the presence of morphospecies A and B of *Phlebotomus argentipes* sensu lato (Diptera: Psychodidae) in Sri Lanka-implications for leishmaniasis transmission. *Journal of Vector-borne Diseases*, **42**: 155.

Struck, T. and Cerca, J. (2019). Cryptic species and their evolutionary significance. *Encyclopedia of Life Sciences*, pp.1-9. DOI: https://doi.org/10.1002/9780470015902.a0028292

Tamura, K., Stecher, G., Peterson, D., Filipski, A. and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Molecular Biology and Evolution*, **30**:.2725-2729. DOI: https://doi.org/https://doi.org/10.1093/molbev/mst197.

Ye, F., Liu, T., King, S.D. and You, P. (2015). Mitochondrial genomes of two phlebotomine sand flies, *Phlebotomus chinensis* and *Phlebotomus papatasi* (Diptera: Nematocera), the first representatives from the family Psychodidae. *Parasites and Vectors*, **8** (1), 472. DOI: https://doi.org/10.1186/s13071-015-1081-1.