Biota Neotropica 24(4): e20241680, 2024
www.scielo.br/bn

biotaneotropica
ISSN 1676-0611 (online edition)

Short Communication

Accessing the sand fly diversity of Tocantins, Northern Brazil: species delimitation using *COI* DNA barcoding

***Ocléia de Sousa Rodrigues Soares*[1*], *Bruno Leite Rodrigues*[2*] & *Paloma Helena Fernandes Shimabukuro*[3]**

[1]*Universidade Federal do Tocantins, Programa de Pós-Graduação em Biodiversidade, Ecologia e Conservação, Porto Nacional, TO, Brasil.*
[2]*Universidade de São Paulo, Faculdade de Saúde Pública, Programa de Pós-Graduação em Saúde Pública, São Paulo, SP, Brasil.*
[3]*Fiocruz Minas, Instituto René Rachou, Coleção de Flebotomíneos, Grupo de Estudos em Leishmanioses, Belo Horizonte, MG, Brasil.*
**Corresponding author: brunno2310@gmail.com*

SOARES, O.S.R., RODRIGUES, B.L., SHIMABUKURO, P.H.F. **Accessing the sand fly diversity of Tocantins, Northern Brazil: species delimitation using *COI* DNA barcoding**. Biota Neotropica 24(4): e20241680. https://doi.org/10.1590/1676-0611-BN-2024-1680

Abstract: Sand flies (Diptera, Psychodidae, Phlebotominae) are considered natural vectors of infectious agents, such as viruses, bacteria and protozoa. About 1,060 species are recognized, of which 73 have been recorded in the state of Tocantins, located in the transition of Amazon and Cerrado biomes, Northern Brazil. Here, we surveyed the sand fly fauna in different environments of the municipality of Palmas, including caves. Also, we evaluated a fragment of the cytochrome c oxidase subunit 1 (*COI*) as reliable for species delimitation in this region. The morphological identification of 163 sand flies revealed the presence of 26 species distributed in 13 genera. Of these, *Lutzomyia itambe*, *Deanemyia samueli*, *Pintomyia gruta*, *Psathyromyia barretti*, and *Sciopemyia servulolimai* had not yet been recorded in the state of Tocantins, thus, the sand fly fauna of the state is composed of 78 species. The first DNA sequence of *Edentomyia* sp., were generated, and probably correspond to an undescribed species, and new studies are needed to verify their real taxonomic identity. Also, eight nominal sand fly species were *COI*-sequenced for the first time, improving the DNA repositories for molecular species identification. The use of *COI* DNA barcodes proved to be efficient for identifying sand fly species in the state of Tocantins and revealed the existence of cryptic diversity for *Nyssomyia whitmani* and *Psathyromyia aragaoi* which need further investigations using an integrative taxonomy approach. However, *COI* sequences was ineffective in delimiting species of *Evandromyia* (*Aldamyia*) and *Psychodpygus* Chagasi series, and our limited sampling should be evaluated in more robust datasets to check the real usefulness of DNA sequences in identifying sand flies.
Keywords: *Cerrado; Amazon; Species delimitation; Molecular taxonomy; Phlebotominae; Caves.*

Acessando a diversidade de flebotomíneos do Tocantins, Norte do Brasil: delimitação de espécies usando o código de barras de DNA

Resumo: Os flebotomíneos (Diptera, Psychodidae, Phlebotominae) são considerados vetores biológicos de agentes infecciosos como vírus, bactérias e protozoários. São reconhecidas cerca de 1.060 espécies, das quais 73 foram registradas no estado do Tocantins, localizado na transição dos biomas Amazônia e Cerrado, Norte do Brasil. Neste estudo, foi realizado o levantamento da fauna de flebotomíneos em diferentes ambientes do município de Palmas, incluindo cavernas. Além disso, avaliamos um fragmento do gene citocromo c oxidase subunidade 1 (*COI*) para a delimitação de espécies nesta região. A identificação morfológica de 163 flebotomíneos revelou a presença de 26 espécies distribuídas em 13 gêneros. Destes, *Lutzomyia itambe, Deanemyia samueli, Pintomyia gruta, Psathyromyia barretti* e *Sciopemyia servulolimai* ainda não haviam sido registrados no estado do Tocantins, assim, a fauna de flebotomíneos do estado é composta por 78 espécies. A primeira sequência de DNA de *Edentomyia* sp. foi gerada e provavelmente corresponde a uma espécie não descrita, sendo novos estudos necessários para verificar sua real identidade taxonômica. Além disso, oito espécies foram sequenciadas para o *COI* pela primeira vez, melhorando os repositórios de DNA para identificação molecular de espécies. O uso de códigos de barras de DNA se mostrou eficiente para identificar espécies de flebotomíneos no estado do Tocantins e revelou a existência de diversidade críptica para *Nyssomyia whitmani* e *Psathyromyia aragaoi*, que precisam de mais investigações usando uma abordagem de taxonomia integrativa. No entanto, as sequências de *COI* foram ineficazes na delimitação de espécies

das séries *Evandromyia* (*Aldamyia*) e *Psychodpygus* série Chagasi, e nossa amostragem limitada deve ser avaliada em conjuntos de dados mais robustos para verificar a real utilidade das sequências de DNA na identificação de flebotomíneos.

Palavras-chave: *Cerrado; Amazônia; Delimitação de espécies; Taxonomia molecular; Phlebotominae; Cavernas.*

Introduction

Sand flies (Diptera, Psychodidae) are a group of dipteran insects that includes vector species of infectious agents such as viruses, bacteria and protozoa that cause diseases in humans and other animals (Maroli et al. 2013). Of these, the subfamily Leishmaniinae (Trypanosomatidae) stands out causing leishmaniases, a global public health concern due to the multiple forms of clinical manifestations (Espinosa et al. 2016, Akhoundi et al. 2016).

This subfamily is comprised of about 1,060 species widespread in both Eastern and Western hemispheres, being found in different ecotopes, predominantly in the tropics (Galati & Rodrigues 2023). Brazil holds a large diversity of these insects, where they can be found in preserved forest environments, degraded areas, caves, and rural and urban environments. The Brazilian state of Tocantins, located in the Northern region of the country, has a vegetation profile marked by the influence of the two largest Brazilian biomes, the Cerrado (Brazilian Savanah) and the Amazon rainforest, with a predominance of the Cerrado *stricto sensu* (Haidar et al. 2013). To date, 73 sand fly species were reported in the state of Tocantins (Aguiar & Vieira 2018), a high diversity which is justified by presenting transition areas, both vegetational and climatic, between the Cerrado and the Amazon biomes (Machado et al. 2017).

The entomological surveys of the sand fly fauna in Tocantins were carried out based on the morphological identification of these taxa (Lustosa et al. 1986, Andrade-Filho et al. 2001, 2004, Vilela et al. 2011, 2013, 2015, Machado et al. 2012, 2017). However, the use of integrative taxonomy tools can improve monitoring of the diversity of these insects, enabling the association of sexes, reducing ambiguous identifications, and detecting cryptic diversity within species (Depaquit 2014, Sousa Paula et al. 2021, Rodrigues & Galati 2023). One of these tools aims to sequence and analyze a fragment of the cytochrome c oxidase subunit 1 (*COI*) gene, the so-called DNA barcode (Hebert et al. 2003). This tool can help in the identification of sand flies, but only a quarter of the described species had this molecular marker deposited in DNA sequence repositories (Rodrigues & Galati 2023). Thus, evaluating the usefulness of this marker in identifying sand flies in Tocantins is of relevance for understanding local diversity and enabling integrative tools to be used to identify these vectors.

Material and Methods

1. Sample collection and processing

This study was performed in the Municipality of Palmas, capital of the state of Tocantins, which has an estimated population of 228,332 inhabitants and an area of 2,218,943 km² (IBGE, 2014). This region has a humid subhumid climate, average temperature of 28 °C, and the average annual precipitation varies from 1,600 to 1,700 mm. The landscape is characterized by Cerrado fields (Brazilian Savanah formations), typical for Central Brazil.

Entomological collections were carried out on private properties, with prior authorization from the owners, using HP light traps installed from 6:00 pm to 6:00 am, at 1.5 meters above ground level, for three consecutive nights. Captures began in January 2022 and ended in May 2022, totaling five months of captures. We selected two caves in a district of the Municipality of Palmas, Taquaruçu, named 'Boa Esperança' cave (10º3'06"S; 48º05'35"W) and Evilson's cave (10º2'31"S; 48º 12' 22"W). At each of the aforementioned collection sites, two traps were installed, one in the aphotic zone and the other in the photic zone of each cave. The 'Boa Esperança' cave can be characterized by the presence of ferruginous canga substrate with fragments of various dimensions and shapes, linked by iron hydroxide (Oliveira et al. 2008). The Evilson's cave is one of the main tourist attractions in Taquaruçu district due to its proximity to a waterfall of the same name, and its lithology is characterized as sandstone (Morais & Rocha 2011). In addition, two other traps were installed on farms in the North of Palmas, in the 'Serra do Carmo' farm (10º4'39"S; 48º20'8"W) and Iago's farm (10º7'17"S; 48º19'48"W), with one trap on each farm.

Sand flies were placed in 1.5 mL tubes containing 70% alcohol, and then the legs of each specimen were collected and frozen at −20°C for subsequent DNA amplification. For the morphological taxonomic study, the sand flies were subjected to clarification process using diluted commercial liquid detergent in distilled water (1:1), after 12 hours in this solution, the specimens were removed and placed for 15 minutes in 0.9% saline solution, and then in lactophenol for 24 hours. After removing lactophenol, females were dissected, separating the head and the final portion of the abdomen (last three segments), which were fixed in Berlese mounting medium between a slide and coverslip. The males were mounted without dissection. The morphological identification was carried out under a biological microscope following the identification keys and classification proposed by Galati (2018).

2. DNA extraction, PCR, and sequencing

The legs of the sand flies were directly added to the Polymerase Chain Reaction (PCR) mix for amplification of the DNA barcoding fragment of the cytochrome c oxidase subunit I (*COI*) gene. For this, we used the pair of primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 5'-TAAACTTCAGGGTGACCAAAAAATCA-3') (Folmer et al. 1994), and the DNA polymerase enzyme Green GoTaq™ mastermix (Promega). The amplification was carried out in an Applied Biosystems Proflex PCR System thermocycler, under the following conditions: 95 °C for two minutes, followed by 35 cycles at 95 °C for one minute, 52 °C for one minute, and 72 °C for one minute, finalizing

with a final extension of 72 °C for 10 minutes. PCR confirmation was performed through electrophoresis, using 1% agarose gel, stained with 5% ethidium bromide. Positive samples were sent to the DNA Sequencing Platform at the 'Instituto René Rachou', Fiocruz Minas (Brazil).

3. Sequence analysis

We manually checked the electropherograms to remove primer sequences and assemble contigs using SeqTrace v0.9. The *COI* barcodes were submitted to the Barcode of Life Data System – BOLD (Ratnasingham & Hebert 2007), and then to the NCBI GenBank (Sayers et al. 2022). Sequences were assigned to the Process IDs TOBAR001-23 – TOBAR132-23, and Accession Numbers OR773186 – OR773317 for BOLD and GenBank, respectively.

The alignment was performed with MUSCLE algorithm (Edgar 2004) implemented in the MEGA11 software (Tamura et al. 2021). We calculated the intra and interspecific genetic distances using the Barcoding Gap Analysis with uncorrected *p* distances in the BOLD environment. Then, we performed a maximum likelihood (ML) phylogenetic analysis to generate a *COI* gene tree to check the clustering pattern of sequences. For this, we used the IQ-TREE software (Nguyen et al. 2015) using automatic model selection, and 10,000 ultrafast bootstrap pseudoreplicates in the IQ-TREE web server (https://www.hiv.lanl.gov/content/sequence/IQTREE/iqtree.html, accessed on 05th March 2024). We included a sequence of *Sycorax konopiki* (KT946601) to root the ML tree, and edited the tree in the FigTree v.1.4.4 program (http://tree.bio.ed.ac.uk/software/figtree, accessed on 5th March 2024).

Sand fly sequences previously assigned to morphospecies were characterized at the molecular operational taxonomic unit (MOTU) level, which are groups of organisms based on their molecular similarity at a given genetic marker (Blaxter et al. 2005). We performed this analysis to associate morphologically-distinct species with MOTUs, thus enabling evaluation of the usefulness of *COI* DNA barcodes in the taxonomy of sand flies from Tocantins, Northern Brazil. For this, we implemented the partitions of the following algorithms: Automatic Barcode Gap Discovery – ABGD (Puillandre et al. 2012), Refined Single Linkage – RESL (Ratnasingham & Hebert 2007), TCS haplotype networks using statistical parsimony (Clement et al. 2000), and Poisson Tree Processes – PTP (Kapli et al. 2017). We performed two different ABGD analyses using both uncorrected *p* distances or K2P model in the web-server https://bioinfo.mnhn.fr/abi/public/abgd/ (accessed on 5th March 2024). The results of both distances were the same for our dataset, so we reported as a single analysis. For ABGD, we used the parameters Pmin = 0.005, Pmax = 0.1, and X = 1.0, considering the recursive partitions generated with the prior intraspecific divergence of 0.013. RESL sort barcode sequences into MOTUs according to their genetic distances, and then optimizes the clusters with a graphic analytical approach using a Markov Clustering (MCL). For this, we performed the analysis inside the BOLD environment using the 'cluster sequences' tool and default parameter. TCS infer haplotype networks using statistical parsimony, and we implemented this analysis in the software TCS version 1.21 using the default connection limit of 95%. PTP infer MOTUs using single-locus molecular phylogenies and seeks to differentiate stochastic population processes from speciation events. We performed the PTP analysis in the web-server https://mptp.h-its.org/#/tree (accessed on 5th March 2024), using the ML gene tree as input, single rate method, and p-value = 0.001.

Results

In total, 163 sand fly specimens were morphologically-identified at genus or species level. Of these, 26 nominal species were recorded, and some specimens of *Edentomyia*, *Evandromyia*, and *Psychodopygus* Chagasi series were identified only at genus level due to the lack of diagnostic characters for identification. *Lutzomyia itambe* Chaves Júnior, Lima, Paranhos & Andrade, 2023 (19,7%) and *Sciopemyia sordellii* (Shannon & Del Ponte, 1927) (19,1%) were the most abundant species, followed by *Martinsmyia oliveirai* (Martins, Silva & Falcão, 1970) (9,25%), *Edentomyia* sp. Galati, Andrade Filho, Silva & Falcão, 2003 (7,40%), and *Deanemyia samueli* (Deane, 1955) (6,79%) (Table 1). The occurrence of *Lu. itambe*, *Da. samueli*, *Pintomyia gruta* (Ryan, 1986), *Psathyromyia barretti* (Alves & Freitas, 2016), and *Sciopemyia servulolimai* (Damasceno & Causey, 1945) were observed for the first time in the State of Tocantins.

Of these, 132 specimens comprising all species, except for *Evandromyia evandroi* (Costa Lima & Antunes, 1936) and *Pa. barrettii*, were DNA barcoded (Table 1). We obtained the full length of this *COI* fragment (658 bp) with no visual indication of pseudogenes or NUMTs in the alignment. The maximum intraspecific *p* distances ranged from 0% to 3.8%, while the distances to the nearest neighbor varied from 0% to 14.29% (Table 1). *Psathyromyia aragaoi* (Costa Lima, 1932) shows the highest intraspecific divergences (3.8 %) followed by *Nyssomyia whitmani* (Antunes & Coutinho, 1939) (3.65%). In contrast, the species pairs *Evandromyia lenti* (Mangabeira, 1938)/*Evandromyia carmelinoi* (Ryan et al. 1986), and *Psychodopygus complexus* (Mangabeira, 1941)/*Psychodopygus wellcomei* Fraiha, Shaw & Lainson, 1971 have no differences in *COI* sequences, reaching 0% interspecific divergence (Table 1). Apart from these two comparisons, the lowest distance to the nearest neighbor was 5.47%, between *Ny. whitmani* and *Nyssomyia antunesi* (Coutinho, 1939), but were generally greater than 10% (Table 1).

The phylogenetic gene tree recovered well supported clades for nearest all species, except for the species pairs *Ev. lenti*/*Ev. carmelinoi* and *Ps. complexus*/*Ps. wellcomei* (Figure 1). The species delimitation algorithms sorted our *COI* sequences in accordance with morphological identification, also excluding the comparisons mentioned above. PTP, TCS, and RESL algorithms sorted our *COI* dataset into 26 MOTUs, while ABGD merged some of these forming 22 MOTUs (Figure 1). For both species delimitation scenarios, *Ev. lenti*/*Ev. carmelinoi* and *Ps. complexus*/*Ps. wellcomei* merged into the same MOTU. Also, species with great genetic divergences (*Pa. aragaoi* and *Ny. whitmani*) split into three and two MOTUs, respectively, regarding the delimitation of 26 MOTUs (Figure 1).

Table 1. Sand flies collected and processed for molecular analysis from the municipality of Palmas, state of Tocantis, Brazil, between January and May 2022. Number of *COI* sequences, maximum and mean intra-specific divergences, and the minimum interspecific distances (nearest neighbor, NN) for each species. Values of genetic distances are percentages. [a] new sand fly records for the state of Tocantins; [b] new *COI* DNA barcodes for sequence databases.

Species	n (♀/♂)	%	*COI*-barcoded	Max. intra *p* distance	Mean intra *p* distance	Min. *p* distance to the NN
Bichromomyia flaviscutellata	2 (1/1)	1.23	1	0	N/A	11.7
Brumptomyia brumpti	1 (0/1)	0.61	1	0	N/A	12.77
Deanemyia samueli[ab]	11 (8/3)	6.79	10	0.3	0.06	14.29
Edentomyia sp.	12 (4/8)	7.4	9	0	0	12.61
Evandromyia begonae	1 (1/0)	0.61	1	0	N/A	13.53
Evandromyia carmelinoi	4 (2/2)	2.46	4	0.91	0.46	0
Evandromyia evandroi	1 (0/1)	0.61	–	–	–	–
Evandromyia walkeri	1 (1/0)	0.61	1	0	N/A	8.81
Evandromyia lenti	2 (2/0)	1.23	2	0.46	0.46	0
Evandromyia sp.	3 (1/2)	1.85	1	–	–	–
Lutzomyia itambe[ab]	32 (17/15)	19.75	27	0.91	0.15	11.7
Lutzomyia longipalpis	1 (1/0)	0.61	1	0	N/A	10.64
Martinsmyia oliveirai[b]	15 (8/7)	9.25	11	0.46	0.19	10.18
Micropygomyia acanthopharynx[b]	2 (2/0)	1.23	2	0.46	0.46	11.7
Micropygomyia echinatopharynx[b]	3 (1/2)	1.85	3	0.46	0.41	11.09
Micropygomyia longipennis[b]	3 (2/1)	1.85	2	0.76	0.76	12.92
Micropygomyia peresi	3 (2/1)	1.85	3	0.15	0.1	12.16
Nyssomyia antunesi	1 (0/1)	0.61	1	0	N/A	5.47
Nyssomyia whitmani	8 (5/3)	4.93	8	3.65	2.17	5.47
Pintomyia gruta[ab]	3 (3/0)	1.85	1	0	N/A	11.85
Psathyromyia aragaoi	5 (3/2)	3.08	5	3.8	2.37	11.85
Psathyromyia barretti[a]	1 (0/1)	0.61	–	–	–	–
Psathyromyia brasiliensis[b]	1 (1/0)	0.61	1	0	N/A	11.25
Psathyromyia hermalenti	6 (3/3)	3.7	5	0.61	0.43	12.01
Psychodopygus complexus	1 (0/1)	0.61	1	0	N/A	0
Psychodopygus wellcomei	1 (0/1)	0.61	1	0	N/A	0
Psycodophygus sp. *Chagasi Series*	7 (7/0)	4.32	6	–	–	–
Sciopemyia servulolimai[a]	1 (0/1	0.61	–	–	–	–
Sciopemyia sordellii	31 (20/11)	19.13	24	0.61	0.15	11.09
TOTAL	**163**	**100**	**132**			

Discussion

The morphological analysis of sand flies allowed us to identify 26 nominal species in the municipality of Palmas, five of them being recorded for the first time in the state of Tocantins (*Lu. itambe*, *De. samueli*, *Pi. gruta*, *Pa. barretti* and *Sc. servulolimai*), highlighting the rich biodiversity of the Brazilian Cerrado. Previous entomological surveys identified about 32 sand fly species in the district of Taquaruçu and Palmas (Machado et al. 2012, 2017), and surveys in other locations this diversity to 73 species (Aguiar & Vieira 2018). These findings indicate a great diversity of sand fly species in the State, justified by the fact that it presents transitional areas, both vegetational and climatic, between the Cerrado and the Amazon rainorest. Furthermore, most of the sampled specimens were found in cave environments, both inside and/or outside the photic zone, including also the new records for Tocantins state reported here, except for *Pa. barretti*, which was collected in the 'Serra do Carmo' Farm. In fact, many sand flies can inhabit cave environments, and caves have been an excellent source of diversity of these insects in Brazil (Dutra-Rêgo et al. 2022), and our results reinforce the great diversity of caves and the need to study them. Here, we increase the knowledge of the sand fly fauna in Tocantins, which now accounts for 78 nominal species.

The identification of sand flies is usually performed using morphological characters on dichotomous keys (Galati 2018), but the use of molecular tools can assist in the taxonomy and systematics of

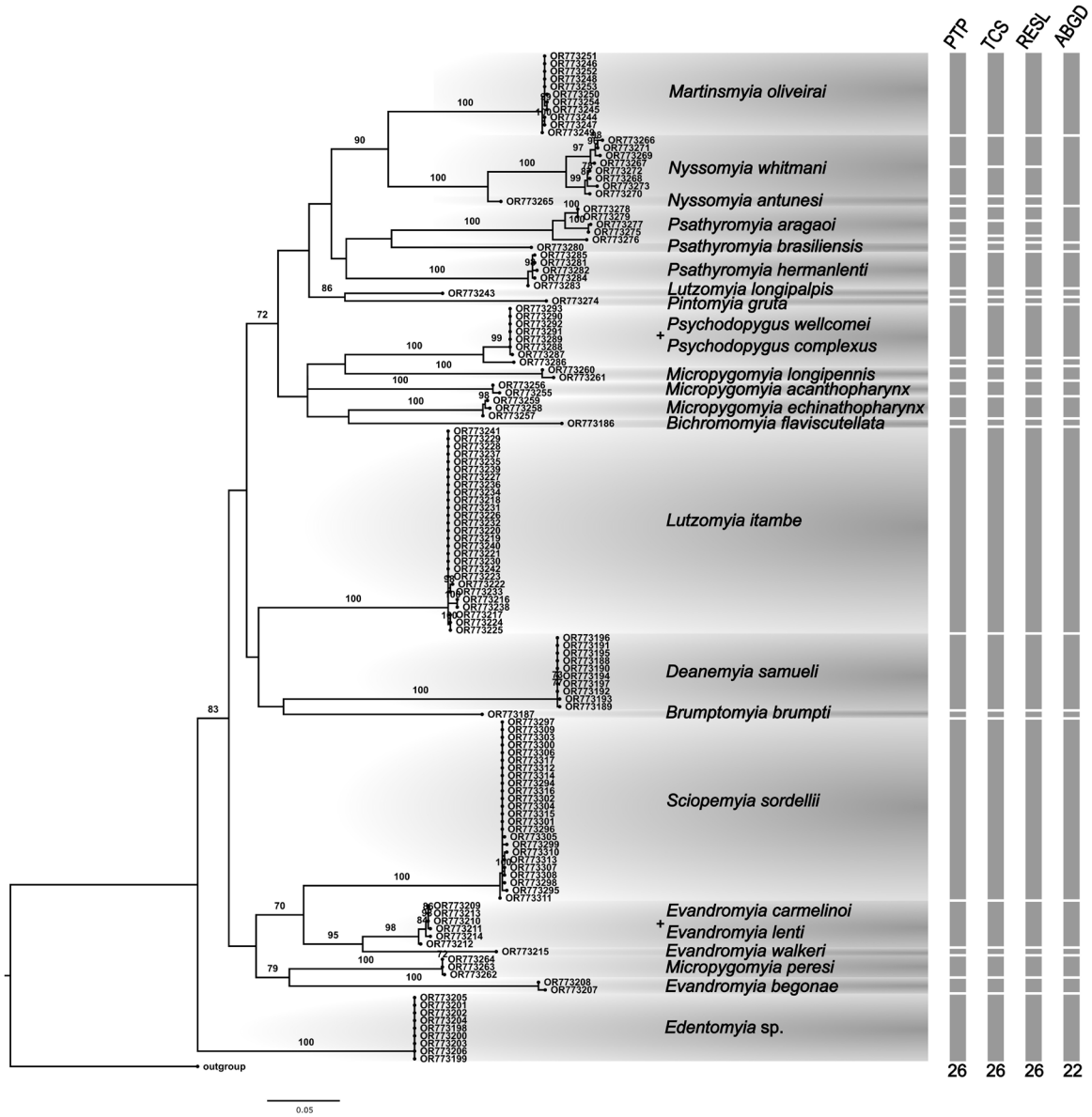

Figure 1. Maximum likelihood gene tree based on *COI* barcode sequences of sand flies from Palmas, Tocantins, Northern Brazil. Numbers near nodes indicate bootstrap values greater than 70. Lateral grey bars indicate the MOTU partitions by the species delimitation algorithms PTP, RESL, TCS and ABGD.

this group, especially in the case of the DNA barcoding using *COI* sequences for animals (Hebert et al. 2003). Further, the molecular taxonomy can be useful in discovering new taxa and elucidating taxonomic problems, despite requiring a robust reference library that is still under construction (Rodrigues & Galati 2023, Antil et al. 2023). To our knowledge, the sand fly *COI* sequences generated in this study were the first for the state of Tocantins. Besides that, our sequencing effort generated the first *COI* barcodes for *De. samueli*, *Lu. itambe*, *Mt. oliveirai*, *Mi. acanthopharynx*, *Mi. echinatopharynx*, *Mi. longipennis*, *Pi. gruta*, *Pa. brasiliensis*, and *Edentomyia* sp., which will be useful for future molecular identification of these taxa in other localities and possibly systematic studies of Phlebotominae.

The analysis of our *COI* sequences shows that intraspecific genetic distances are smaller than the interspecific ones. The 'barcoding gap' (Wiemers & Fiedler 2007) may be relevant to define *COI* sequences as reliable to species identification, although its absence in some datasets does not interfere with species delimitation. However, two species pairs – *Ev. lenti*/*Ev. carmelinoi* and *Ps. wellcomei*/*Ps. complexus* – were molecularly indistinguishable in this study. In the first case, these two species have substantial differences in both male and female characters and have been associated with other species of the subgenus *Evandromyia* (*Aldamyia*) in the 'evandroi complex' (Rodrigues et al. 2020), which to date comprises *Ev. lenti*, *Ev. carmelinoi*, *Ev. piperiformis* and *Ev. evandroi* of different localities of Brazil (Pinto et al. 2015, Rodrigues et al. 2018, 2020). In the second case, *Ps. wellcomei* and *Ps. complexus* are both part of the Chagasi series of the *Psychodopygus* genus, with most of its females being indistinguishable by morphology (Galati 2018). Most of the *Psychodopygus COI* sequences generated

in this study comprise females of Chagasi series and remained identified only at a generic level due to the ineffectiveness of barcodes in separating these two taxa. In fact, *COI* gene sequences seem to be ineffective in delimiting close-related species groups, merging different nominal species into the same clade or MOTU, as the case of the genera *Bichromomyia* (Melo et al. 2020), *Lutzomyia* (Pinto et al. 2015, Moya et al. 2020), *Pintomyia* (Cohnstaedt et al. 2011), and *Pressatia* (Rodrigues & Galati 2024). This may be a consequence of the recent natural history of these taxa, or hybridization events with introgression of mtDNA from different species. For the correct molecular delimitation of these taxa, *multilocus* tools, including molecular markers of the nuclear genome, must be evaluated (Dupuis et al. 2012, Liu et al. 2017).

Except for the two cases mentioned above, the *COI* DNA barcoding was effective in delimiting the remaining species, grouping conspecific sequences into well-supported clades and different MOTUs for each nominal species. In the pragmatic view of DNA barcoding, these sequences will be useful for identifying these taxa without the need for morphological evaluation. However, our dataset has a limited sample of species that occur in Tocantins, as well as in Brazil, and future studies should evaluate the real effectiveness of these sequences when analyzing them in global datasets with a greater diversity of species and populations of the same taxon. Despite limited spatial sampling, the species delimitation algorithms revealed the possibility of the presence of cryptic diversity in *Pa. aragaoi* and *Ny. whitmani*, splitting their *COI* sequences into three and two MOTUs each (regarding PTP, TCS, and RESL). This sympatric cryptic diversity is in accordance with previous investigation of *Pa. aragaoi* from the state of Acre where this species split in two MOTUs while comparing them with Colombian specimens (Pinto et al. 2023). Also, it has been questioned whether *Ny. whitmani* represents a complex of cryptic species (Rangel et al. 1996, Ishikawa et al. 1999, Margonari et al. 2004), and our results indicate that there is a sympatric cryptic diversity in cave environments in Tocantins for the aforementioned species: *Pa. aragaoi* and *Ny. whitmani*. These two species are widely distributed in South America and can occur in different ecosystems and varied environmental pressures, reinforcing the idea that they may represent cryptic species complexes. Therefore, integrative taxonomy studies must evaluate different populations and genetic lineages of these species in order to elucidate their real taxonomic status.

In conclusion, we improved the knowledge on the sand fly fauna in Tocantins, increasing the species list to 78. The sequencing and analysis of *COI* barcodes proved to be effective in delimiting most of the species studied, also detecting sympatric cryptic diversity within important sand fly vectors in this region. Other studies should evaluate these sequences in more robust data sets in terms of species and populations.

Supplementary Material

The following online material is available for this article:

Figure S1 – Expanded phylogenetic gene tree based on *COI* barcode sequences of sand flies from Palmas, Tocantins, Northern Brazil. Numbers near nodes indicate bootstrap values greater than 70.

Table S1 – Sample ID, GenBank Accession Numbers, Collection sites, and coordinates of Brazilian sand flies from Palmas, Tocantins, Central Brazil.

Acknowledgments

We thank Arquimedes Pereira Terêncio, Edivaldo Bessa Lessa, Eder Castro, Davi Cavalcante, Anivaldo Romeiro Vasco, Oziel Vieira Oliveira, Leonard Diniz and Anderson Soares Brito for their assistance in field work. We also thank the Leishmaniasis Study Group/Plebotominae Collection (Fiocruz/COLFLEB), Instituto René Rachou, Fiocruz Minas, Belo Horizonte, Minas Gerais, Brazil for their support in processing the samples for sequencing and to Renata de Barros Ramos Oliveira and Caroline Penido Rocha from the Sanger DNA Sequencing Platform at the René Institute Rachou, Fiocruz Minas.

This study was partially supported by scholarship grant of the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001. PHFS would like to thank additional financial support given by the Fundação de Amparo à Pesquisa de Minas Gerais (FAPEMIG) (Grant number PPM-00676-18).

Associate Editor

Gustavo Graciolli

Author Contributions

Ocléia de Sousa Rodrigues Soares: conceptualization; methodology; writing original draft; writing—review & editing.

Bruno Leite Rodrigues: methodology; formal analysis and visualization; writing original draft; writing—review & editing.

Paloma Helena Fernandes Shimabukuro: conceptualization; methodology; resources; supervision; writing—review & editing.

Conflicts of Interest

The authors declare no competing interests.

Data Availability

All sequences obtained from the study were deposited in the GenBank database under the accession numbers OR773186 – OR773317. The data supporting this article are published through the Integrated Publishing Toolkit (IPT) of SiBBr and is available under a CC BY-NC 4.0 license in SiBBr (Soares et al. 2024) and GBIF (https://doi.org/10.15468/pyy9r6).

References

AGUIAR, G.M. & VIEIRA, V.R. 2018. Regional distribution and habitats of Brazilian phlebotomine species. In: Rangel, E, Shaw, J (Eds.), Brazilian Sand Flies. Springer, Cham, pp. 251–298.

AKHOUNDI, M., KUHLS, K., CANNET, A., VOTÝPKA, J., MARTY, P., DELAUNAY, P. & SERENO, D. 2016. A historical overview of the classification, evolution, and dispersion of Leishmania parasites and sandflies. PLoS Neglected Tropical Diseases 10(3):e0004349.

ANDRADE FILHO, J.D., GALATI, E.A.B., ANDRADE, W.A. & FALCÃO, A.L. 2004. Description of *Micropygomyia* (*Silvamyia*) *echinatopharynx* sp. nov. (Diptera: Psychodidae) a new species of phlebotomine sand fly from the state of Tocantins, Brazil. Memórias do Instituto Oswaldo Cruz 99:609–15.

ANDRADE FILHO, J.D., VALENTE, M.B., ANDRADE, W.A., BRAZIL, R.P. & FALCÃO, A.L. 2001. Flebotomíneos do Estado de Tocantins, Brasil (Diptera: Psychodidae). Revista da Sociedade Brasileira de Medicina Tropical 34:323–91.

ANTIL, S., ABRAHAM, J.S., SRIPOORNA, S., MAURYA, S., DAGAR, J., MAKHIJA, S., BHAGAT, P., GUPTA, R., SOOD, U., LAL, R. & TOTEJA, R. 2023. DNA barcoding, an effective tool for species identification: a review. Molecular Biology Reports 50(1):761–75.

BLAXTER, M., MANN, J., CHAPMAN, T., THOMAS, F., WHITTON, C., FLOYD, R. & ABEBE, E. 2005. Defining operational taxonomic units using DNA barcode data. Philosophical Transactions of the Royal Society B: Biological Sciences 360(1462): 1935–1943.

CLEMENT, M., POSADA, D.C.K.A. & CRANDALL, K.A. 2000. TCS: a computer program to estimate gene genealogies. Molecular Ecology 9:1657–9.

COHNSTAEDT, L.W., BEATI, L., CACERES, A.G., FERRO, C. & MUNSTERMANN, L.E. 2011. Phylogenetics of the phlebotomine sand fly group Verrucarum (Diptera: Psychodidae: Lutzomyia). Am. J. Trop. Med. Hyg 84(6):913.

DEPAQUIT, J. 2014. Molecular systematics applied to Phlebotomine sandflies: Review and perspectives. Infection, Genetics and Evolution 28:744-56.

DUTRA-RÊGO, F., FREIRE, M.L., CARVALHO, G.M. & ANDRADE-FILHO, J.D. 2022. Revisiting the cave-dwelling sand flies (Diptera, Psychodidae, Phlebotominae) from Brazil: Diversity and potential role in the transmission of *Leishmania* Ross, 1903 (Kinetoplastida: Trypanosomatidae). Medical and Veterinary Entomology 36(4):408–23.

DUPUIS, J.R., ROE, A.D. & SPERLING, F.A. 2012. Multi-locus species delimitation in closely related animals and fungi: one marker is not enough. Molecular Ecology 21(18):4422–36.

EDGAR, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res 32(5):1792–7.

ESPINOSA, O.A., SERRANO, M.G., CAMARGO, E.P., TEIXEIRA, M.M. & SHAW, J.J. 2016. An appraisal of the taxonomy and nomenclature of trypanosomatids presently classified as *Leishmania* and *Endotrypanum*. Parasitology 145(4):430–42.

FOLMER, O., BLACK, M., HOEH, W., LUTZ, R. & VRIJENHOEK, R. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3:294–299.

GALATI, E.A.B. & RODRIGUES, B.L. 2023. A review of historical Phlebotominae taxonomy (Diptera: Psychodidae). Neotropical Entomology 52(4):539–59.

GALATI, E.A.B. 2018. Phlebotominae (Diptera, Psychodidae): classification, morphology and terminology of adults and identification of American taxa. In: Rangel EF, Shaw JJ (eds) Brazilian Sand Flies. Springer Internacional Publishing AG, Cham pp. 9–212.

HAIDAR, R.F., FAGG, J.M., PINTO, J.R., DIAS, R.R., DAMASCO, G., SILVA, L.D. & FAGG, C.W. 2013. Seasonal forests and ecotone areas in the state of Tocantins, Brazil: structure, classification and guidelines for conservation. Acta Amazonica 43:261–90.

HEBERT, P.D., RATNASINGHAM, S. & DE WAARD, J.R. 2003. Barcoding animal life: cytochrome c oxidase subunit 1 divergences among closely related species. Proceedings of the Royal Society of London. Series B: Biological Sciences 270(1):S96–9.

IBGE-Instituto Brasileiro de Geografia e Estatística. Censo Demográfico. 2014. [Accessed in September 28, 2024]. Available in: https://cidades.ibge.gov.br/xtras/perfil.php.

ISHIKAWA, E.A., READY, P.D., DE SOUZA, A.A., DAY, J.C., RANGEL, E.F., DAVIES, C.R. & SHAW, J.J. 1999. A mitochondrial DNA phylogeny indicates close relationships between populations of *Lutzomyia whitmani* (Diptera: Psychodidae, Phlebotominae) from the rain-forest regions of Amazonia and northeast Brazil. Memórias do Instituto Oswaldo Cruz 94:339–45.

KAPLI, P., LUTTEROPP, S., ZHANG, J., KOBERT, K., PAVLIDIS, P., STAMATAKIS, A. & FLOURI, T. 2017. Multi-rate Poisson tree processes for single-locus species delimitation under maximum likelihood and Markov chain Monte Carlo. Bioinformatics 33(11):1630–8.

LIU, J., JIANG, J., SONG, S., TORNABENE, L., CHABARRIA, R., NAYLOR, G.J. & LI, C. 2017. Multilocus DNA barcoding–species identification with multilocus data. Scientific Reports 7(1):16601.

LUSTOSA, E.S., NAVES, H.A., CARVALHO, M.E. & BARBOSA, W. 1986. Contribuição ao Conhecimento da Fauna Flebotominica do Estado de Goiás-1984-1985. NOTA PRÉVIA I. Revista de Patologia Tropical/Journal of Tropical Pathology 15(1).

MACHADO, T.D., MINUZZI-SOUZA, T.T., FERREIRA, T.D., FREIRE, L.P., TIMBÓ, R.V., VITAL, T.E., NITZ, N., SILVA, M.N., SANTOS, A.D., SALES, N.M. & OBARA, M.T. 2017. The role of gallery forests in maintaining Phlebotominae populations: potential *Leishmania* spp. vectors in the Brazilian savanna. Memórias do Instituto Oswaldo Cruz 112:681–91.

MACHADO, T.O., BRAGANÇA, M.A., CARVALHO, M.L. & ANDRADE FILHO, J.D. 2012. Species diversity of sandflies (Diptera: Psychodidae) during different seasons and in different environments in the district of Taquaruçú, state of Tocantins, Brazil. Memórias do Instituto Oswaldo Cruz 107:955–9.

MARGONARI, C.S., FORTES-DIAS, C.L. & DIAS, E.S. 2004. Genetic variability in geographical populations of *Lutzomyia whitmani* elucidated by RAPD-PCR. Journal of medical entomology 41(2):187–92.

MAROLI, M., FELICIANGELI, M.D., BICHAUD, L., CHARREL, R.N. & GRADONI, L. 2013. Phlebotomine sandflies and the spreading of leishmaniases and other diseases of public health concern. Medical and Veterinary Entomology 27(2):123–47.

MELO, L.B., ALENCAR, R.B. & SCARPASSA, V.M. 2020. Molecular taxonomy and phylogenetic inferences of *Bichromomyia flaviscutellata* complex based on the *COI* gene DNA barcode region. Infect. Genet. Evol. 81:104256.

MORAIS, F. & ROCHA, S.D. 2011. Cavernas em arenito no planalto residual do Tocantins (Sandstone Caves in The Planalto Residual do Tocantins). Espeleo-Tema 22(1):127–37.

MOYA, S.L., PECH-MAY, A., QUINTANA, M.G., MANTECA-ACOSTA, M. & SALOMON, O.D. 2020. Phylogenetic relationships of closely-related phlebotomine sand flies (Diptera: Psychodidae) of *Nyssomyia* genus and *Lutzomyia* subgenus. Memorias do Instituto Oswaldo Cruz 115:e200220.

NGUYEN, L.T., SCHMIDT, H.A., VON HAESELER, A. & MINH, B.Q. 2015. IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. Molecular Biology and Evolution 32(1):268–74.

OLIVEIRA, M.A., FERREIRA, R.L., CARNEIRO, M.A. & DIOTAIUTI L. 2008. Ecology of Cavernicola pilosa Barber, 1937 (Hemiptera: Reduviidae: Triatominae) in the Boa Esperança Cave, Tocantins, Brazil. Ecotropica 14:63–8.

PINTO, I.S., CHAGAS, B.D., RODRIGUES, A.A., FERREIRA, A.L., REZENDE, H.R., BRUNO, R.V., FALQUETO, A., ANDRADE-FILHO, J.D., GALATI, E.A.B., SHIMABUKURO, P.H.F., BRAZIL, R.P. & PEIXOTO, A.A. 2015. DNA barcoding of neotropical sand flies (Diptera, Psychodidae, Phlebotominae): species identification and discovery within Brazil. PLoS One 10(10):e0140636.

PINTO, I.S., RODRIGUES, B.L., DE ARAUJO-PEREIRA, T., SHIMABUKURO, P.H.F., DE PITA-PEREIRA, D., BRITTO, C. & BRAZIL, R.P. 2023. DNA barcoding of sand flies (Diptera, Psychodidae, Phlebotominae) from the western Brazilian Amazon. Plos One 18(2):e0281289.

PUILLANDRE, N., LAMBERT, A., BROUILLET, S. & ACHAZ, G.J. 2012. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. Molecular ecology 21(8):1864–77.

RANGEL, E.F., LAINSON, R., SOUZA, A.A., READY, P. & AZEVEDO, A.C. 1996. Variation between geographical populations of *Lutzomyia* (*Nyssomyia*) *whitmani* (Antunes & Coutinho, 1939) sensu lato (Diptera: Psychodidae: Phlebotominae) in Brazil. Memórias do Instituto Oswaldo Cruz 91:43–50.

RATNASINGHAM, S. & HEBERT, P.D. 2007. BOLD: The Barcode of Life Data System (http://www. Barcodinglife. org). Molecular Ecology Notes 7(3):355–364.

RODRIGUES, B.L., BATON, L.A. & SHIMABUKURO, P.H. 2020. Single-locus DNA barcoding and species delimitation of the sandfly subgenus *Evandromyia* (*Aldamyia*). Medical and Veterinary Entomology 34(4):420–31.

RODRIGUES, B.L., CARVALHO-COSTA, L.F., PINTO, I.S. & REBÊLO, J.M.M. 2018. DNA barcoding reveals hidden diversity of sand flies (Diptera: Psychodidae) at fine and broad spatial scales in Brazilian endemic regions for leishmaniasis. Journal of Medical Entomology. 55(4):893–901.

RODRIGUES, B.L. & GALATI, E.A. 2023. Molecular taxonomy of phlebotomine sand flies (Diptera, Psychodidae) with emphasis on DNA barcoding: a review. Acta Tropica 238:106778.

RODRIGUES, B.L., & GALATI, E.A.B. 2024. New sand fly (Diptera, Psychodidae) records and *COI* DNA barcodes in the state of Maranhão, Eastern Amazon, Brazil. Acta Tropica 250:107095.

SAYERS, E.W., CAVANAUGH, M., CLARK, K., PRUITT, K.D., SCHOCH, C.L., SHERRY, S.T. & KARSCH-MIZRACHI, I. 2022. GenBank. Nucleic Acids Res, 50(D1):D161–D164.

SOARES, O.D.S.R., RODRIGUES, B.L. & SHIMABUKURO, P.H.F. 2024. Accessing the sand fly diversity in caves of Tocantins, Northern Brazil: species delimitation using COI DNA barcoding. Version 1.0. 2024. FIOCRUZ - Oswaldo Cruz Foundation. Occurrence dataset. https://ipt.sibbr.gov.br/sibbr/resource?r=sf_tocantins&v=1.0 and https://doi.org/10.15468/pyy9r6.

SOUSA-PAULA, L.D., PESSOA, F.A., OTRANTO, D. & DANTAS-TORRES, F. 2021. Beyond taxonomy: species complexes in New World phlebotomine sand flies. Medical and Veterinary Entomology 35(3):267–83.

TAMURA, K., STECHER, G. & KUMAR, S. 2021. MEGA11: molecular evolutionary genetics analysis version 11. Molecular Biology and Evolution 38(7):3022–7.

VILELA, M.L., AZEVEDO, A.C. & GODOY, R.E. 2015. Description of a new phlebotomine species of the Brazilian Cerrado from sandstone caves in Tocantins State, Brazil: *Lutzomyia* (*Lutzomyia*) *elizabethrangelae* sp. nov. (Diptera: Psychodidae). Journal of Medical Entomology 52(4):596–603.

VILELA, M.L., AZEVEDO, C.G., CARVALHO, B.M. & RANGEL, E.F. 2011. Phlebotomine fauna (Diptera: Psychodidae) and putative vectors of leishmaniases in impacted area by hydroelectric plant, state of Tocantins, Brazil. PLoS One 6(12):e27721.

VILELA, M.L., PITA-PEREIRA, D.D., AZEVEDO, C.G., GODOY, R.E., BRITTO, C. & RANGEL EF. 2013. The phlebotomine fauna (Diptera: Psychodidae) of Guaraí, state of Tocantins, with an emphasis on the putative vectors of American cutaneous leishmaniasis in rural settlement and periurban areas. Memórias do Instituto Oswaldo Cruz 108:578–85.

WIEMERS, M. & FIEDLER, K. 2007. Does the DNA barcoding gap exist?–a case study in blue butterflies (Lepidoptera: Lycaenidae). Frontiers in Zoology 4:1–16.

Received: 10/07/2024
Accepted: 23/10/2024
Published online: 02/12/2024