

RESEARCH ARTICLE

DNA barcoding of sand flies (Diptera, Psychodidae, Phlebotominae) from the western Brazilian Amazon

Israel de Souza Pinto¹, Bruno Leite Rodrigues², Thais de Araujo-Pereira³, Paloma Helena Fernandes Shimabukuro⁴, Daniela de Pita-Pereira³, Constança Britto^{3*}, Reginaldo Peçanha Brazil⁵

1 Departamento de Ciências Biológicas, Instituto Federal de Educação, Ciência e Tecnologia do Pará, Itaituba, Pará, Brasil, **2** Programa de Pós-Graduação em Saúde Pública, Faculdade de Saúde Pública, Universidade de São Paulo, São Paulo, São Paulo, Brasil, **3** Laboratório de Biologia Molecular e Doenças Endêmicas, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Rio de Janeiro, Brasil, **4** Coleção de Flebotomíneos, Grupo de Estudos em Leishmanioses, Instituto René Rachou, Fundação Oswaldo Cruz, Belo Horizonte, Minas Gerais, Brasil, **5** Laboratório de Doenças Parasitárias, Instituto Oswaldo Cruz, Fundação Oswaldo Cruz, Rio de Janeiro, Rio de Janeiro, Brasil

* cbritto@ioc.fiocruz.br



OPEN ACCESS

Citation: Pinto IdS, Rodrigues BL, de Araujo-Pereira T, Shimabukuro PHF, de Pita-Pereira D, Britto C, et al. (2023) DNA barcoding of sand flies (Diptera, Psychodidae, Phlebotominae) from the western Brazilian Amazon. PLoS ONE 18(2): e0281289. <https://doi.org/10.1371/journal.pone.0281289>

Editor: Everton Falcão de Oliveira, Federal University of Mato Grosso do Sul, BRAZIL

Received: September 20, 2022

Accepted: January 19, 2023

Published: February 2, 2023

Copyright: © 2023 Pinto et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All relevant data are within the paper and its [Supporting Information file \(S1 Fig.\)](#)

Funding: This work was supported by grants from Fundação Carlos Chagas Filho de Amparo à Pesquisa do Estado do Rio de Janeiro/FAPERJ (E-26/202.782/2018, E-26/202.287/2019), Fundação de Amparo à Pesquisa do Estado de Minas Gerais – FAPEMIG (PPM-00676-18), and the Brazilian Research Council - Conselho Nacional de

Abstract

The subfamily Phlebotominae comprises important insects for public health. The use of complementary tools such as molecular taxonomy is necessary for interspecific delimitation and/or discovery of cryptic species. Here, we evaluated the DNA barcoding tool to identify different species in the southwestern Brazilian Amazon. For this, we collected sand flies in forest fragments along the highway BR-317, in the municipality of Brasiléia, state of Acre, Brazil. The specimens were DNA-barcoded using a fragment of the *cytochrome c oxidase subunit I (COI)* gene. The sequences were analyzed to generate K2P pairwise genetic distances and a Neighbour-joining tree. The sand fly barcodes were also clustered into Molecular Operation Taxonomic Units (MOTU) using Automatic Barcode Gap Discovery (ABGD) approach. A total of 59 *COI* sequences comprising 22 nominal species and ten genera were generated. Of these, 11 species had not been sequenced before, thus being new *COI* sequences to science. Intraspecific genetic distances ranged between 0 and 4.9%, with *Pin-tomyia serrana* presenting the highest values of genetic distance, in addition to having been partitioned into three MOTUs. Regarding the distances to the nearest neighbour, all species present higher values in relation to the maximum intraspecific distance, in addition to forming well supported clusters in the neighbour-joining analysis. The DNA barcoding approach is useful for the molecular identification of sand flies from Brasiléia, state of Acre, and was efficient in detecting cryptic diversity of five species which can be confirmed in future studies using an integrative approach. We also generated new *COI* barcodes for *Trichophoromyia auraensis*, *Nyssomyia shawi*, and *Psychodopygus paraensis*, which may play a role in the transmission of *Leishmania* spp. in the Brazilian Amazon.

Desenvolvimento Científico e Tecnológico/CNPq (304308/2019-3) – Financial Disclosure: CB is research fellow of CNPq (BPP) and FAPERJ (CNE); TAP is research fellow of FAPERJ; BLR and DPP (31010016001P0) are research fellows of Coordenação de Aperfeiçoamento de Pessoal de Nível Superior do Brasil (CAPES) - Finance Code 001. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

Introduction

Sand flies (Diptera, Psychodidae, Phlebotominae) are the main insects involved in the transmission of *Leishmania* spp. parasites to humans and other animals [1, 2], comprising about 500 species and 23 genera in Neotropical regions [3, 4]. The subfamily Phlebotominae is widely distributed in Brazil, and the Amazon biome presents the largest number of species recorded, including most of the genera *Bichromomyia*, *Nyssomyia*, *Psychodopygus*, and *Trichophoromyia* which includes incriminated vectors of the etiological agents of cutaneous leishmaniasis (CL) [1, 5]. The state of Acre, located in the southwest of the Brazilian Amazon, has a great diversity of sand fly species associated to different habitats, reservoirs, and protozoans [6–8], totaling 116 species of these insects [9], which increases the relevance of correct entomological monitoring in these localities.

Despite the relevance to public health, the species-level identification of sand flies can be hampered due to the complexity of morphological characters, and the lack of differences between close-related taxa [4]. There are some complexes of cryptic species reported among sand flies [10], and taxa that are only known based on the morphological information from a single sex [4]. For this reason, integrative taxonomy approaches may use complementary tools to elucidate interspecific boundaries [11, 12]. One of these tools—the DNA barcoding—aims to assist in the species-level identification through the sequencing of a fragment of the *cytochrome c oxidase subunit I (COI)* gene, which allows the identification of closely related species [13]. Indeed, the sand fly systematics and taxonomy using DNA sequences are relatively understudied, and only about a third of the nominal species has been processed for any molecular marker, and a quarter of them for the *COI* gene [14]. In Brazil, some efforts have been made to evaluate DNA barcodes as effective in delimiting several species of different genera [15, 16], but information on Amazonian taxa is lacking, where some studies focused on a few numbers of close-related species [17–19].

Therefore, we aim to evaluate and make available new *COI*-barcodes for sand flies from a highly diverse region in the Brazilian Amazon, which may be of great relevance for future monitoring of these insects. Sand flies were collected in the municipality of Brasiléia, in the south of Acre state, in areas of current transmission of *Leishmania* spp. causing cutaneous leishmaniasis according to human cases registered in the recent years by the National Health Foundation (FUNASA) and the local Fernando Azevedo Correia Center of Health of Brasiléia [7]. In this study, we generated *COI* sequences for sand fly species already processed from other locations, which may help to elucidate the phylogeographic patterns and species discovery within these taxa in upcoming studies.

Materials and methods

Study area and collection

For the present study, sand flies were obtained from two collection sites situated in the rural zone of the municipality of Brasiléia, in areas transversely positioned along the BR-317 Trans Pacific highway connecting Brasiléia and the municipality of Assis Brasil, in the state of Acre, Brazil. The collection sites were identified as *Ramal* (or branch) based on the distance in kilometers of BR-317 Trans Pacific highway from the center of Brasiléia (Fig 1), and were selected due to the high incidence of human cutaneous leishmaniasis cases in forested areas close to residences in rural areas [7]. Collections were carried out in private areas after receiving permission from the landowners. Sand fly collections were performed under a permanent license of the Brazilian *Sistema de Autorização e Informação em Biodiversidade* (SISBIO) for the capture of zoological material (permit 32669–4).

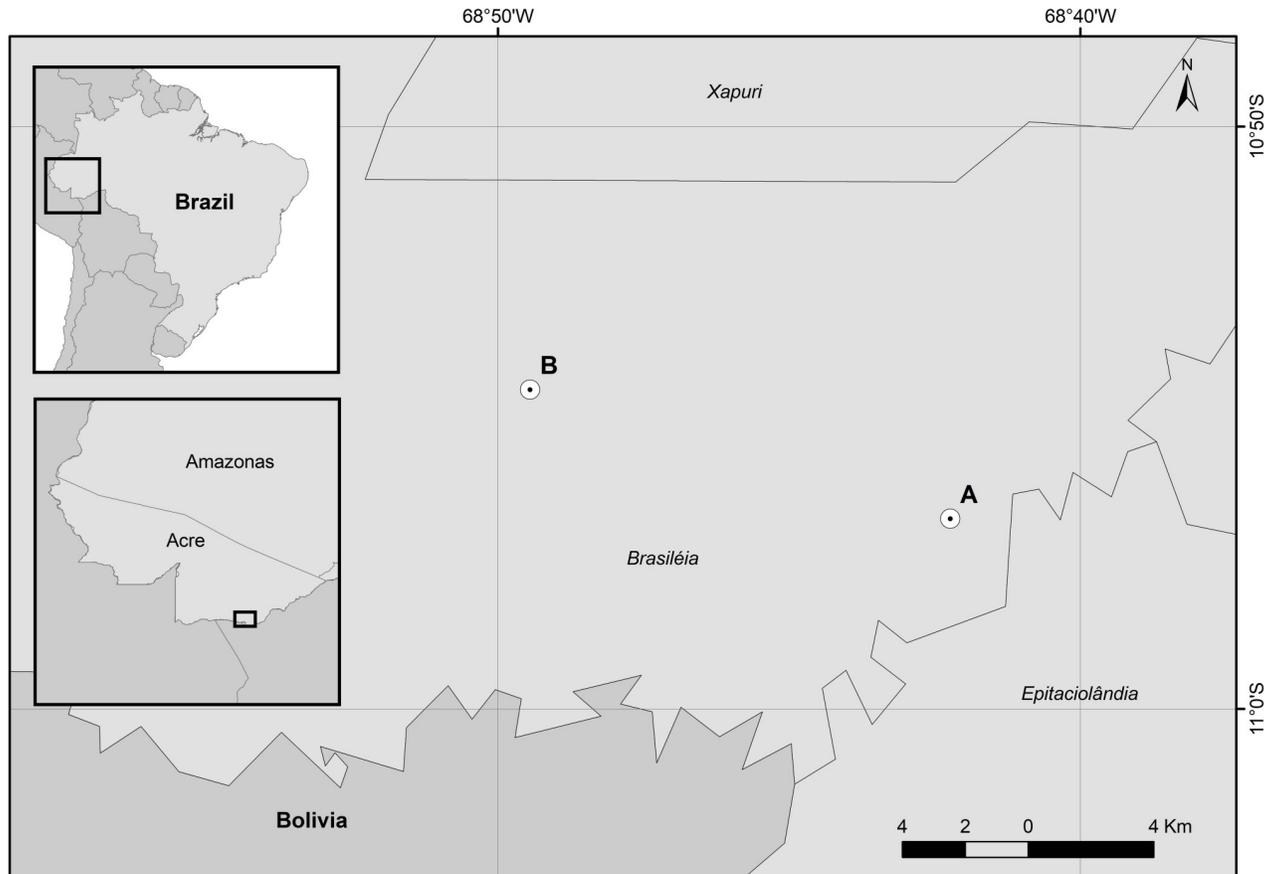


Fig 1. Map of the study area. Map of Brazil highlighting the state of Acre and the municipality of Brasiléia. The collection sites A and B correspond to Ramal 04 (km 04) and Ramal 13 (km 13) of the BR-317 Trans Pacific highway to Brasiléia, respectively.

<https://doi.org/10.1371/journal.pone.0281289.g001>

The municipality of Brasiléia, which has an estimated population of 27,000 inhabitants [20], is located in the south of Acre on the border with the municipality of Cobija–Bolivia and the Brazilian municipalities of Epitaciolândia, Assis Brasil, Sena Madureira and Xapuri. The insects were collected using HP light traps that operated overnight inside high-density forested areas, with the following coordinates: A—(S 10° 56′ 44″; W 68° 42′ 14″), and B—(S 10° 54′ 31″; W 68° 49′ 27″). A small container (220 mL) with 80% ethanol was connected to each trap to store the insects while they were being captured [21]. Male sand flies were screened and had their legs dissected for later DNA extraction, while the rest of the body was mounted between slides and coverslips for morphological identification according to the key and classification of [4]. The abbreviations of generic names follow the proposed by [22].

DNA barcoding

To generate the DNA barcodes, total DNA was extracted from the legs of each male sand fly specimen. All the legs available for each individual were placed in tubes (one for each specimen), and DNA extraction was performed. Genomic DNA was used to amplify/sequencing the DNA barcoding region of the *cytochrome c oxidase subunit I (COI)* gene, according to [15]. In summary, a rapid DNA extraction method for DNA amplification was performed as reported by [23]. To amplify a 658 bp fragment of the *COI* gene, PCR was carried out in a final volume of 50 μ L reaction mixture containing 2 μ L of genomic DNA template, 25 μ L 2X

Promega Go Taq Green™ Master Mix (Promega, Madison, WI, USA), and the primers LCO1490 (5′-GGTCAACAAATCATAAAGATATTGG-3′) and HCO2198 (5′-TAAACTT-CAGGGTGACCAAAAAATCA-3′) in final concentrations of 1.0 μM each [24]. The cycling conditions consisted of an initial denaturation step of 95°C for 3 min, followed by 37 cycles of 95°C for 1 min, 57°C for 1 min, 72°C for 1.5 min, and final extension at 72°C for 7 min. The amplified fragments were separated by agarose (2%) gel electrophoresis and purified using Illustra GFX PCR DNA™ and Gel Band Purification Kit (GE Healthcare, Pittsburgh, PA, USA). The purified fragments were sequenced bidirectionally using the same PCR primers for the *COI* gene. Reactions were performed using Big Dye Terminator v.3.1 Cycle Sequencing (Applied Biosystems Foster City, USA) following the manufacturer's specifications and assayed in the Sanger ABI 3730 DNA analyzer at Fundação Oswaldo Cruz (PDTIS/Fiocruz), Rio de Janeiro, Brazil. The generated sequences were submitted to GenBank nucleotide sequence database [25], and have been assigned the Accession Numbers OP346789-OP346847. In addition, our dataset was also submitted to the Barcode of Life Data System (BOLD) [26] for further analysis.

Sequences were initially visualized and edited in the software BioEdit v5.0.9.0 [27, 28], and then the alignment was performed using ClustalW [29] as implemented in the Mega v7 [30]. Visual inspection of the alignment indicated the absence of pseudogenes and/or nuclear copies of mitochondrial origin (NUMTs). The maximum intraspecific pairwise genetic distances, in addition of the minimum distance to the nearest-neighbour of each nominal species (interspecific), were generated using Kimura 2-parameters model (K2P) in the BOLD Systems environment.

In order to check the clustering pattern and the utility of *COI*-barcodes for the molecular identification of sand flies from the southwestern Brazilian Amazon, we combined our new sequences to those previously processed in GenBank for the sampled nominal species. Then, we reconstructed a neighbour-joining tree (NJ) for the combined dataset in the software MEGA v7 with 1,000 bootstraps pseudoreplicates and K2P model. The sequences of each specimen were also identified at the Molecular Operational Taxonomic Unit (MOTU) level [31], using the Automatic Barcode Gap Discovery (ABGD) algorithm [32] which is available at <https://bioinfo.mnhn.fr/abi/public/abgd/abgdweb.html> (last access at 13th Jun 2022). For this, we used the combined dataset with no missing data, and the parameters: $P_{min} = 0.001$, $P_{max} = 0.1$, $X = 1.0$, and K2P model. After the run, we considered the recursive partitions with prior maximal distances between 1% and 2.5%.

Ethics statement

Access to genetic heritage has been registered into the *Sistema Nacional de Gestão do Patrimônio Genético e do Conhecimento Tradicional Associado*—SisGen (A41DBDD). Sand fly collection was carried out under a permanent license provided by the *Sistema de Autorização e Informação em Biodiversidade*—SISBIO (permit 32669–4).

Results

In total, we generated 59 *COI* sequences for the sand flies sampling from the state of Acre ($n = 59$), southwestern of the Brazilian Amazon. These sequences are distributed in 22 species and ten genera (Table 1), and some of these *COI* sequences, representing 11 species, are new to science: *Evandromyia saulensis* (Floch & Abonnenc, 1944), *Lutzomyia flabellata* Martins & Silva, 1964, *Nyssomyia shawi* (Fraiha, Ward & Ready, 1981), *Psathyromyia dendrophyla* (Mangabeira, 1942), *Psychodopygus amazonensis* (Root, 1934), *Psychodopygus clautreii* (Abonnenc, Léger & Fauran, 1979), *Psychodopygus francoisleponti* Zapata, Depaquit & León 2012,

Table 1. List of nominal sand fly species collected in the western Brazilian Amazon which generate DNA barcode sequences. Number of sequences, intraspecific and interspecific (to the nearest neighbour) K2P genetic distances are presented.

Species	Number of sequences	Mean Intraspecific distance	Max Intraspecific distance	Distance to Nearest Neighbour
<i>Evandromyia bacula</i>	1	N/A	N/A	15.30
<i>Evandromyia saulensis</i> *	2	2.17	2.17	16.62
<i>Lutzomyia flabellata</i> *	1	N/A	N/A	11.83
<i>Nyssomyia antunesi</i>	2	0.46	0.46	5.23
<i>Nyssomyia shawi</i> *	1	N/A	N/A	11.48
<i>Nyssomyia whitmani</i>	3	0.51	0.61	5.23
<i>Pintomyia serrana</i>	6	3.12	4.9	11.13
<i>Pressatia choti</i>	3	0.2	0.3	14.39
<i>Psathyromyia abbonenci</i>	1	N/A	N/A	10.79
<i>Psathyromyia aragai</i>	1	N/A	N/A	12.19
<i>Psathyromyia dendrophyla</i> *	2	0.61	0.61	10.79
<i>Psychodopygus amazonensis</i> *	1	N/A	N/A	8.56
<i>Psychodopygus carrerai</i>	3	1.6	2.32	9.39
<i>Psychodopygus clautrei</i> *	2	0	0	7.24
<i>Psychodopygus davisi</i>	7	1.09	2.33	7.24
<i>Psychodopygus francoisleponti</i> *	3	0.1	0.15	11.32
<i>Psychodopygus hirsutus</i>	4	2.1	3.29	9.57
<i>Psychodopygus paraensis</i> *	4	0.54	1.07	9.39
<i>Trichophoromyia auraensis</i> *	5	0	0	2.49
<i>Trichophoromyia octavioi</i> *	3	0.61	0.77	2.49
<i>Trichopygomyia dasypodogeton</i> *	1	N/A	N/A	11.85
<i>Viannamyia furcata</i>	3	1.86	2.8	10.78

N/A, not applicable.

* species that were COI-DNA barcoded for the first time.

<https://doi.org/10.1371/journal.pone.0281289.t001>

Psychodopygus paraensis (Costa Lima, 1941), *Trichophoromyia auraensis* (Mangabeira, 1942), *Trichophoromyia octavioi* (Vargas, 1949), and *Trichopygomyia dasypodogeton* (Castro, 1939). In addition, we also sequenced previously published species for COI gene, but which are new to the state of Acre: *Evandromyia bacula* (Martins, Falcão & Silva, 1965), *Nyssomyia antunesi* (Coutinho, 1939), *Nyssomyia whitmani* (Antunes & Coutinho, 1939), *Pintomyia serrana* (Damasceno & Arouck, 1949), *Pressatia choti* (Floch & Abonnenc, 1941), *Psathyromyia abbonenci* (Floch & Chassignet, 1947), *Psathyromyia aragai* (Costa Lima, 1932), *Psychodopygus carrerai carrerai* (Barretto, 1946), *Psychodopygus davisi* (Root 1934), *Psychodopygus hirsutus* (Mangabeira, 1942), and *Viannamyia furcata* (Mangabeira, 1941).

Our new COI-barcode sequences have its full-length of 658 bp. For this dataset, the intra-specific mean and maximum K2P distances ranged from 0.0% to 3.1%, and 0.0% to 4.9%, respectively (Table 1). Most species did not obtain intraspecific distances greater than 3%, with the exception of *Pi. serrana* (maximum K2P = 4.9%). On the other hand, the minimum K2P distances to the nearest neighbour (interspecific) ranged from 2.49% to 16.62%, with a notable discrepancy in the species of the *Trichophoromyia* genus, which achieved interspecific values lower than the intraspecific comparisons of some species. Despite that, all species present higher values in relation to the maximum intraspecific distance (Table 1).

The combined dataset comprised our 59 sequences plus 37 conspecific barcodes previously deposited and available in GenBank database, totaling 96 sequences. The neighbour-joining phenogram clustered the sequences according to the sampled nominal species with high values of

The analyzed species comprise specimens from the state of Acre and their conspecific sequences extracted from GenBank (when available). The number near nodes indicates bootstrap above 70. Tip labels marked with grey circles indicate clades with sequences generated in this study (state of Acre, Brazil), while black circles indicate those with Acre plus GenBank sequences. The unmarked clades have GenBank sequences only. The grey bar represents the species delimitation made by the ABGD algorithm.

The species delimitation made by ABGD algorithm, sorted the sequences into 28 and 23 MOTUs in the recursive partitions of $P = 1.29$ and $P = 2.15$, respectively (Fig 2 and S1 Fig). The most conservative scenario ($P = 2.15$) merged the species pairs *Ps. davisii*/*Ps. clautrei* and *Ny. antunesii*/*Ny. whitmani* into the same MOTU each, while the other nominal species formed at least one MOTU for each one (Fig 2 and S1 Fig). In contrast, the ABGD partition with $P = 1.29$ correctly delimited the above-mentioned species pairs, in addition to splitting specimens of *Ps. carrerai carrerai*, *Ps. davisii*, *Pa. abonnenci*, *Pa. aragaoi* into two different MOTUs, and *Pi. serrana* into three MOTUs (Fig 2 and S1 Fig).

Discussion

Here we provide the first *COI* DNA barcodes of sand flies from the state of Acre, located in the western Brazilian Amazon. These records comprise some sequences of species that were never analyzed for this marker, and can be used for future studies regarding the molecular identification and phylogeographic studies of this group. Further, we demonstrated that *COI*-barcodes have enough information for the species delimitation of several species from western Brazilian Amazon, and are useful in detecting cryptic diversity within these taxa, which may represent different cryptic species or structured populations. The relevance of these findings for entomological monitoring of sand flies, and for public health of areas with the transmission of leishmaniasis should be accessed in future studies.

The results of genetic divergences and the clustering pattern of neighbour-joining tree indicate that all species analyzed in this study can be correctly identified by DNA barcodes. Despite this, the species delimitation made by ABGD algorithm merged some close-related species, such as *Ps. davisii*/*Ps. clautrei* and *Ny. antunesii*/*Ny. whitmani*, when using more conservative parameters. Similarly, *Ny. whitmani* and *Ny. intermedia* from southeast Brazil are not correctly delimited depending on the parameters used in the ABGD analysis [15], and other species of this genus are merged into the same MOTU while analyzing with more diverged species [16]. These two genera, especially *Nyssomyia*, have little *COI* genetic differentiation between the nominal species compared to other sand fly genera, and may have different species merged into the same MOTU using distance-based methods such as ABGD. These results indicate that the best ABGD partition scheme for our dataset is the one with $P = 1.29$, which correctly delimited all the analyzed species, sometimes partitioning the same taxon into different MOTUs.

Studies that aim to provide sand fly DNA barcodes in the Brazilian Amazon are scarce, despite the great diversity of species found in the region. In most cases, the sequencing of *COI* fragments is performed for a few species from different locations, mainly because they are related to the transmission of pathogens, such as the species *Nyssomyia umbratilis* [17, 18] and *Bichromomyia* spp. [19]. Our study represented the first effort to sample different sand fly species and genera in the Brazilian Amazon, accounting for 22 analyzed species. Despite this, the state of Acre has 116 recorded nominal species of these insects [9], and our dataset comprises only about 19% of this diversity in a specific locality, demonstrating that future studies should assess the real usefulness of *COI* barcodes in the identification of different sand flies in Acre and Brazil, both in terms of different species and locations.

We generated for the first time the *COI* barcode sequences for 11 sand fly species, including some that are of relevance for the transmission of *Leishmania* pathogens in South America. Of these, the sequencing of *Th. auraensis* and *Th. octavioi* are especially important because the first is considered a putative vector of *Leishmania* (*Viannia*) in the Amazon region, including the state of Acre [33–35], in addition of their females being isomorphic [4], so the male-female association using barcodes may be implemented to correctly identify these insects in upcoming studies (e.g., [15, 36, 37]). We are providing the first *COI* barcodes for two vector species—*Ny. shawi* and *Ps. paraensis*; the former being associated to the transmission of *Leishmania guyanensis* in Bolivia [1], and the latter is considered the vector of *Leishmania naiffi* in Brazil [1, 38]. These and the other new barcode records are important for understanding the taxonomy and systematics of sand flies in the Americas, and could be used for future molecular identification of these taxa.

We revealed cryptic diversity in five sand fly species collected in the state of Acre: *Ps. carrerai carrerai*, *Ps. davisii*, *Pa. abonnenci*, *Pa. aragaoi*, and *Pi. serrana*. In the cases of *Ps. davisii*, *Pa. abonnenci*, and *Pa. aragaoi*, the barcode sequences were splitted into two MOTUs, one of them containing individuals from the state of Acre and the other with sequences from south-eastern Brazil (*Ps. davisii*; [15]) or Colombia (*Pa. abonnenci* and *Pa. aragaoi*; [39, 40]). This clustering pattern seems to be related to the geographic distances of the analyzed specimens in addition of the landscape heterogeneity within Brazil and South America, which may have promoted the isolation of these populations and interruption of the gene flow between them [41], a result previously seen when analyzing sand fly barcode sequences from broader spatial scales in Brazil [16]. On the other hand, *Ps. carrerai carrerai* presented an allopatric cryptic diversity, with sequences being splitted within the same locality (Brasiléia, state of Acre), while *Pi. serrana* splitted into three MOTUs and presents both allopatric and sympatric patterns of cryptic diversity. Whether or not these different MOTUs within nominal species are in fact different taxa, should be evaluated using integrative taxonomy approaches.

In summary, the DNA barcoding strategy is useful for the molecular identification of sand flies from western Brazilian Amazon, and was efficient in detecting cryptic diversity of *Pi. serrana*, which can be confirmed in future studies using an integrative approach. Further, it was generated remarkable *COI*-barcode records, like *Th. auraensis*, *Ny. shawi*, and *Ps. paraensis* which play a role in the transmission of *Leishmania* spp. in the Brazilian Amazon.

Supporting information

S1 Fig. Expanded neighbour-joining gene tree of sand fly *COI* barcode sequences. The analyzed species comprise specimens from the state of Acre and their conspecific sequences extracted from GenBank (when available). The number near nodes indicates bootstrap values greater than 70. Tip labels marked with grey circles indicate clades with sequences generated in this study (state of Acre, Brazil). The unmarked clades have GenBank sequences only. The grey bar represents the species delimitation made by the ABGD algorithm. (TIF)

Acknowledgments

To FUNASA, for the support on sand fly collections, specially to Mr. Francisco Firmino and Mr. Ademar, and to the Municipality Department of Health of Brasiléia, for the logistical support and operational assistance in the conduct of field work. To Gustavo Rocha Leite for the preparation of the map. To the *Instituto Chico Mendes de Conservação da Biodiversidade*

(ICMBio) for providing the collection permissions via SISBIO, and to the support of the sequencing facilities of Fundação Oswaldo Cruz (Fiocruz) in Rio de Janeiro.

Author Contributions

Conceptualization: Israel de Souza Pinto.

Formal analysis: Israel de Souza Pinto, Bruno Leite Rodrigues, Paloma Helena Fernandes Shimabukuro.

Funding acquisition: Constança Britto, Reginaldo Peçanha Brazil.

Investigation: Israel de Souza Pinto, Thais de Araujo-Pereira, Paloma Helena Fernandes Shimabukuro, Daniela de Pita-Pereira.

Methodology: Israel de Souza Pinto, Bruno Leite Rodrigues, Paloma Helena Fernandes Shimabukuro.

Project administration: Constança Britto, Reginaldo Peçanha Brazil.

Resources: Constança Britto, Reginaldo Peçanha Brazil.

Supervision: Constança Britto, Reginaldo Peçanha Brazil.

Writing – original draft: Bruno Leite Rodrigues.

Writing – review & editing: Israel de Souza Pinto, Thais de Araujo-Pereira, Paloma Helena Fernandes Shimabukuro, Daniela de Pita-Pereira, Constança Britto, Reginaldo Peçanha Brazil.

References

1. Maroli M, Feliciangeli MD, Bichaud L, Charrel RN, Gradoni L. Phlebotomine sandflies and the spreading of leishmaniasis and other diseases of public health concern. *Med Vet Entomol.* 2013; 27(2): 123–147. <https://doi.org/10.1111/j.1365-2915.2012.01034.x> PMID: 22924419
2. Cecílio P, Cordeiro-da-Silva A, Oliveira F. Sand flies: Basic information on the vectors of leishmaniasis and their interactions with *Leishmania* parasites. *Commun Biol.* 2022; 5(1): 1–2305. <https://doi.org/10.1038/s42003-022-03240-z> PMID: 35379881
3. Shimabukuro PHF, de Andrade AJ, Galati EA. Checklist of American sand flies (Diptera, Psychodidae, Phlebotominae): genera, species, and their distribution. *ZooKeys.* 2017;(660): 67–106. <https://doi.org/10.3897/zookeys.660.10508> PMID: 28794674
4. Galati EAB. Morphology and Terminology of Adults and Identification of the American Taxa. In: Rangel EF, Shaw JJ, editors. *Brazilian sand flies.*, Springer: Cham; 2018. pp. 251–298.
5. Brazil RP, Rodrigues AA, Andrade Filho JD. Sand fly vectors of *Leishmania* in the Americas- a mini review. *Entomol Ornithol Herpetol.* 2015; 4(2): 144. <https://doi.org/10.4172/2161-0983.1000144>
6. Ávila MM, Brilhante AF, de Souza CF, Bevilacqua PD, Galati EA, Brazil RP. Ecology, feeding and natural infection by *Leishmania* spp. of phlebotomine sand flies in an area of high incidence of American tegumentary leishmaniasis in the municipality of Rio Branco, Acre, Brazil. *Parasit Vectors.* 2018; 11(1): 64. <https://doi.org/10.1186/s13071-018-2641-y> PMID: 29373995
7. Araujo-Pereira T, Pita-Pereira D, Baia-Gomes SM, Boité M, Silva F, Pinto IS, et al. An overview of the sandfly fauna (Diptera: Psychodidae) followed by the detection of *Leishmania* DNA and blood meal identification in the state of Acre, Amazonian Brazil. *Mem Inst Oswaldo Cruz.* 2020; 115: e200157. <https://doi.org/10.1590/0074-02760200157> PMID: 33206821
8. Brilhante AF, Lima L, Ávila MM, Medeiros-Sousa AR, Souza JF, Santos NP, et al. Remarkable diversity, new records and *Leishmania* detection in the sand fly fauna of an area of high endemicity for cutaneous leishmaniasis in Acre state, Brazilian Amazonian Forest. *Acta Trop.* 2021; 223: 106103. <https://doi.org/10.1016/j.actatropica.2021.106103> PMID: 34416187
9. Godoy RE, Andrade AJ, Shimabukuro PHF, Brilhante AF. Sand fly (Diptera Psychodidae, Phlebotominae) records in Acre, Brazil: a dataset. *SciELO [Preprint].* 2022. [cited 2022 March 18]. Available from: <https://doi.org/10.1590/SciELOPreprints.3792>

10. Sousa-Paula LD, Pessoa FA, Otranto D, Dantas-Torres F. Beyond taxonomy: species complexes in New World phlebotomine sand flies. *Med Vet Entomol.* 2021; 35(3): 267–283. <https://doi.org/10.1111/mve.12510> PMID: 33480064
11. Dayrat B. Towards integrative taxonomy. *Biol J Linn Soc Lond.* 2005; 85(3): 407–417. <https://doi.org/10.1111/j.1095-8312.2005.00503.x>
12. Padiál JM, Miralles A, De la Riva I, Vences M. The integrative future of taxonomy. *Front Zool.* 2010; 7(1): 16. <https://doi.org/10.1186/1742-9994-7-16> PMID: 20500846
13. Hebert PD, Cywinska A, Ball SL, DeWaard JR. Biological identifications through DNA barcodes. *Proc R Soc Lond B Biol Sci.* 2003; 270(1512): 313–321. <https://doi.org/10.1098/rspb.2002.2218> PMID: 12614582
14. Rodrigues BL, Galati EA. Molecular taxonomy of phlebotomine sand flies (Diptera, Psychodidae) with emphasis on DNA barcoding: a review. *Acta Trop.* 2023; 238: 106778. <https://doi.org/10.1016/j.actatropica.2022.106778> PMID: 36435214
15. Pinto IS, Chagas BD, Rodrigues AA, Ferreira AL, Rezende HR, Bruno RV, et al. DNA barcoding of neotropical sand flies (Diptera, Psychodidae, Phlebotominae): species identification and discovery within Brazil. *PLoS One.* 2015; 10(10):e0140636. <https://doi.org/10.1371/journal.pone.0140636> PMID: 26506007
16. Rodrigues BL, Carvalho-Costa LF, Pinto IS, Rebêlo JM. DNA barcoding reveals hidden diversity of sand flies (Diptera: Psychodidae) at fine and broad spatial scales in Brazilian endemic regions for leishmaniasis. *J Med Entomol.* 2018; 55(4): 893–901. <https://doi.org/10.1093/jme/tjy032> PMID: 29562268
17. Scarpassa VM, Alencar RB. *Lutzomyia umbratilis*, the main vector of *Leishmania guyanensis*, represents a novel species complex? *PLoS One.* 2012; 7(5): e37341. <https://doi.org/10.1371/journal.pone.0037341> PMID: 22662146
18. Scarpassa VM, Cunha-Machado AS, Alencar RB. Multiple evolutionary lineages for the main vector of *Leishmania guyanensis*, *Lutzomyia umbratilis* (Diptera: Psychodidae), in the Brazilian Amazon. *Sci Rep.* 2021; 11(1): 15323. <https://doi.org/10.1038/s41598-021-93072-4> PMID: 34321495
19. Melo LB, Alencar RB, Scarpassa VM. Molecular taxonomy and phylogenetic inferences of *Bichromomyia flaviscutellata* complex based on the COI gene DNA barcode region. *Infect Genet Evol.* 2020; 81: 104256. <https://doi.org/10.1016/j.meegid.2020.104256> PMID: 32084625
20. IBGE, Diretoria de Pesquisas, Coordenação de População e Indicadores Sociais, Estimativas da população residente com data de referência 10 de julho de 2021. Available from: <https://cidades.ibge.gov.br/brasil/ac/brasileia/panorama>, accessed in 10th July 2022.
21. Pinto IS, Tonini JFR, Ferreira AL, Falqueto A. A brief inventory of sand flies (Diptera, Psychodidae) from the National Forest of the Rio Preto, state of the Espírito Santo, southeastern Brazil. *Biota Neotrop.* 2012; 12: 323–326. <https://doi.org/10.1590/S1676-06032012000100025>
22. Marcondes C.B. 2007. A proposal of generic and subgeneric abbreviations for Phlebotomine sandflies (Diptera: Psychodidae: Phlebotominae) of the world. *Entomol News.* 2007; 118: 351–356. <https://doi.org/10.3157/0013-872X>
23. Jowett T. Preparation of nucleic acids. In: Roberts DB, editor. *Drosophila: a practical approach.* Oxford: IRL Press; 1998. pp. 347–371.
24. Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Mol Mar Biol Biotechnol.* 1994; 3: 294–299. PMID: 7881515.
25. Sayers EW, Cavanaugh M, Clark K, Pruitt KD, Schoch CL, Sherry ST, et al. GenBank. *Nucleic Acids Res.* 2021; 49(D1): D92–96. <https://doi.org/10.1093/nar/gkaa1023> PMID: 33196830
26. Ratnasingham S, Hebert PD. BOLD: The Barcode of Life Data System (<http://www.Barcodinglife.org>). *Mol Ecol Notes.* 2007; 7(3): 355–364. <https://doi.org/10.1111/j.1471-8286.2007.01678.x> PMID: 18784790
27. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser.* 1999; 41: 95–98.
28. Hall T, Biosciences I, Carlsbad C. BioEdit: an important software for molecular biology. *GERF Bull Biosci.* 2011 Jun; 2(1): 60–61.
29. Thompson JD, Gibson TJ, Higgins DG. Multiple sequence alignment using ClustalW and ClustalX. *Curr Protoc Bioinformatics.* 2003; 2: 2–3. <https://doi.org/10.1002/0471250953.bi0203s00> PMID: 18792934
30. Kumar S., Stecher G., & Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 2016; 33(7): 1870–1874. <https://doi.org/10.1093/molbev/msw054> PMID: 27004904

31. Blaxter M, Mann J, Chapman T, Thomas F, Whitton C, Floyd R, et al. Defining operational taxonomic units using DNA barcode data. *Philos Trans R Soc Lond B Biol Sci.* 2005; 360(1462): 1935–1943. <https://doi.org/10.1098/rstb.2005.1725> PMID: 16214751
32. Puillandre N, Lambert A, Brouillet S, Achaz G. ABGD, Automatic Barcode Gap Discovery for primary species delimitation. *Mol Ecol.* 2012; 21(8): 1864–1877. <https://doi.org/10.1111/j.1365-294X.2011.05239.x> PMID: 21883587
33. Araujo-Pereira T, Pita-Pereira D, Boité MC, Melo M, Costa-Rego TA, Fuzari AA, et al. First description of *Leishmania (Viannia)* infection in *Evandromyia saulensis*, *Pressatia* sp. and *Trichophoromyia aur-aensis* (Psychodidae: Phlebotominae) in a transmission area of cutaneous leishmaniasis in Acre state, Amazon Basin, Brazil. *Mem Inst Oswaldo Cruz.* 2017; 112: 75–78. <https://doi.org/10.1590/0074-02760160283> PMID: 28076470
34. Teles CB, Pessoa FA, Medeiros JF, Camargo LM. Comments about the finding of *Trichophoromyia aur-aensis* (Mangabeira 1942) sandfly specimens infected by *Leishmania (Viannia)* parasites in Acre state. *Mem Inst Oswaldo Cruz.* 2017; 112: 517–519. <https://doi.org/10.1590/0074-02760170024> PMID: 28591315
35. Santos TV, Silveira FT. Increasing putative vector importance of *Trichophoromyia* phlebotomines (Diptera: Psychodidae). *Mem Inst Oswaldo Cruz.* 2020; 115: e190284. <https://doi.org/10.1590/0074-02760190284> PMID: 32049097
36. Kasap OE, Linton YM, Karakus M, Ozbel Y, Alten B. Revision of the species composition and distribution of Turkish sand flies using DNA barcodes. *Parasit Vectors.* 2019; 12(1): 410. <https://doi.org/10.1186/s13071-019-3669-3> PMID: 31439012
37. Lozano-Sardaneta YN, Paternina LE, Sanchez-Montes S, Quintero A, Ibanez-Bernal S, Sanchez-Cordero V, et al. DNA barcoding and fauna of phlebotomine sand flies (Diptera: Psychodidae: Phlebotominae) from Los Tuxtlas, Veracruz, Mexico. *Acta Trop.* 2020; 201: 105220. <https://doi.org/10.1016/j.actatropica.2019.105220> PMID: 31618609
38. Rangel EF, Lainson R, Carvalho BM, Costa SM, Shaw JJ. Sand fly vectors of American cutaneous leishmaniasis in Brazil. In: Rangel EF, Shaw JJ, editors. *Brazilian sand flies*. Springer: Cham; 2018. pp. 341–80.
39. Contreras- Gutierrez MA, Vivero RJ, Velez ID, Porter CH, Uribe S. DNA barcoding for the identification of sand fly species (Diptera, Psychodidae, Phlebotominae) in Colombia. *PloS One.* 2014; 9(1): e85496. <https://doi.org/10.1371/journal.pone.0085496> PMID: 24454877
40. Romero-Ricardo L, Lastre-Meza N, Perez-Doria A, Bejarano EE. DNA barcoding to identify species of phlebotomine sand fly (Diptera: Psychodidae) in the mixed leishmaniasis focus of the Colombian Caribbean. *Acta Trop.* 2016; 159: 125–131. <https://doi.org/10.1016/j.actatropica.2016.03.017> PMID: 26992298
41. Van Strien MJ, Holderegger R, Van Heck HJ. Isolation-by-distance in landscapes: considerations for landscape genetics. *Heredity.* 2015; 114(1): 27–37. <https://doi.org/10.1038/hdy.2014.62> PMID: 25052412