

Increasing the digital repository of DNA barcoding sequences of sand flies (Psychodidae: Phlebotominae)

Magdalena Laurito^{1,2/+}, Iliana M Ontivero¹, Walter R Almirón^{1,2}

¹Universidad Nacional de Córdoba, Centro de Investigaciones Entomológicas de Córdoba, Facultad de Ciencias Exactas, Físicas y Naturales, Córdoba, Argentina

²Consejo Nacional de Investigaciones Científicas y Técnicas, Instituto de Investigaciones Biológicas y Tecnológicas, Córdoba, Argentina

Sand fly identification is complex because it depends on the expertise of the taxonomist. The females show subtle morphological differences and the occurrence of the species complexes are usual in this taxon. Therefore, a fragment of the cytochrome *c* oxidase subunit I (*COI*) gene is used for taxon barcoding to resolve this kind of problem. This study incorporates barcode sequences, for the first time, for *Evandromyia cortelezii* and *Migonemyia migonei* from Argentina. The nucleotide sequence divergences were estimated to generate a neighbour-joining (NJ) tree. The automatic barcode gap discovery (ABGD) approach was employed to find the barcode gaps and the operational taxonomic unit (OTU) delimitation. Other species of the subtribe were included. The frequency histogram of divergences showed a barcoding gap. The ABGD analysis identified 14 operational taxonomic units (OTUs) from 13 morphological species. Sequences of *Ev. cortelezii* and *Mg. migonei* formed well supported clusters and were diagnosed as primary species. These sequences are useful tools for molecular identification of the sand flies of the New World.

Key words: DNA barcoding - species delimitation - *Evandromyia cortelezii* - *Migonemyia migonei*

Phlebotominae is a subfamily of Psychodidae that is widely distributed with greater diversity in the tropical and subtropical regions.⁽¹⁾ Until 2004, the southernmost distribution limit of Phlebotominae was 31°35'S-60°17'W, Entre Ríos Province, Argentina.^(2,3) Since then, several new records began to provide evidences regarding the expanding range of distribution of Phlebotominae to higher latitudes in Córdoba Province, Argentina. Salomón et al.⁽³⁾ recorded the specimens of *Migonemyia migonei* at 30°54'S-62°18'W in 2004. Visintin et al.⁽⁴⁾ reported specimens of the *Cortelezii* complex at 30°51'S-62°56'W in 2012 and Ontivero et al.⁽⁵⁾ detected the southernmost record of both species in Córdoba City at 31°23'S-64°04'W during 2015-2016. The first autochthonous and the most southerly case of Tegumentary Leishmaniasis (TL) has been reported from the Córdoba Province in 2014.⁽⁵⁾

As not all sand fly species are vectors for pathogens, the accurate species identification during an entomological surveillance is needed to predict possible transmission foci. The use of morphological keys for the identification of Phlebotomine species depends on the expertise of the taxonomist. To accurately identify females of *Mg. migonei*,

e.g. dissection of the last abdominal segments must be done to see the stalk-shaped spermathecal.⁽⁶⁾ The subtle morphological differences between females or males and the occurrence of species complexes also make species identification difficult. In the *Cortelezii* complex⁽⁷⁾ females of *Evandromyia cortelezii* and *Ev. sallesi* are difficult to differentiate because they share diagnostic features of the length of the clypeus and the spermathecal duct. The species can only be distinguished using male characters: *Ev. cortelezii* has a hood-shaped paramere and a gonocoxite with a basal tuft of four long setae, and *Ev. sallesi* has the paramere without an apical hood-shaped structure and the gonocoxite has a basal tuft of five long setae.⁽⁶⁾ A fragment of the cytochrome *c* oxidase subunit I (*COI*) mitochondrial gene has been widely used in molecular taxonomy to resolve these kinds of issues.⁽⁸⁾ Even if several studies have employed *COI* barcode sequences to identify Phlebotominae species and revealed species complexes in America,^(9,10,11) (among others) its use can be problematic in closely related species because of incomplete lineage sorting and introgression events.⁽¹²⁾ The strategies designed to delimit species boundaries have been developed, among them is clustering, which does not require *a priori* definitions of species.⁽¹³⁾ Even if more than 220 sand fly species with barcode sequences are available in digital repositories, additional sequences are needed to enhance the efficiency of molecular identification. The present study incorporates, for the first time, barcode sequences of *Ev. cortelezii* and the first sequence for *Mg. migonei* from Argentina and validates *COI* as an effective marker for taxonomic identification in sandflies because it shows consistent differences between closely related species.

Sand fly adults were captured with mini CDC light traps located at three sites: the Bajo Grande sewage treatment plant (31°23'38"S-64°04'36"W), the San Mar-

doi: 10.1590/0074-02760190208

Financial support: Agencia Nacional de Ciencia y Técnica, FONCYT (2013-1254) and Secretaría de Ciencia y Técnica, Universidad Nacional de Córdoba.

ML and WRA are members of the CONICET.

+ Corresponding author: mlaurito@conicet.gov.ar

ORCID: <https://orcid.org/0000-0002-6846-5029>

Received 18 June 2019

Accepted 05 November 2019



tín Urban Natural Reserve (31°21'44"S-64°15'47"W), and the Zoo Garden (31°25'34"S-64°10'33"W), once a week actively from 5 pm to 10 am, during March and April, 2018. The sand flies collected were killed by freezing and were stored at -20°C in 80% ethanol until processed. Species identification was based on the structures of the head and internal and external genitalia of females and males, respectively, which were mounted on the microscope slides using the taxonomic key of Galati.⁽⁶⁾ Male sand fly legs (between two and six, according to availability), thorax, and abdomen were separated before morphological preparations for subsequent molecular procedures. Voucher specimens were deposited in the collection of the Centro de Investigaciones Entomológicas de Córdoba, Argentina. The abbreviation of the generic names were according to Marcondes.⁽¹⁴⁾ DNA extractions were obtained from five male specimens of *Ev. cortelezzii* and one of *Mg. migonei*. The genomic DNA was extracted using the rapid method described by Pinto et al.⁽¹⁰⁾ The DNA obtained was used as the template to amplify the ~658 bp fragment of the *COI* barcode region and the amplification was performed as described in Folmer et al.⁽¹⁵⁾ The polymerase chain reaction (PCR) products were electrophoresed in 1% TAE agarose gel stained with GelRed (Biotium Inc, Hayward, USA). All sequencing reactions were carried out in both directions using an ABI3730XL automatic sequencer (Macrogen Inc., Korea) with the same set of PCR primers. The sequences were edited using BioEdit v. 7.2.⁽¹⁶⁾ The primer regions were removed from the sequences. Comparisons with available sequences were performed using Basic Local Alignment Search Tool (BLAST) to check for sequence homology and species identification of *Mg. migonei*. The *COI* gene sequences were aligned by nucleotides using the Muscle algorithm.⁽¹⁷⁾ The barcode sequences obtained in this study are deposited in GenBank under accession numbers: MN432906 (Cba1828M); MN432907 (Cba1825M); MN432908 (Cba1829F); MN432909 (Cba1837M); MN432910 (Cba1839M); MN432911 (Cba1856M). Pairwise nucleotide sequence divergences and the mean intra- and interspecific distances were estimated using Kimura two-parameter (K2P) distance⁽¹⁸⁾ implemented in Mega v7.⁽¹⁹⁾ Subsequently, the divergences obtained were used to generate a neighbour-joining (NJ) tree, using Mega v7 to evaluate the clustering pattern among species. The statistical support for the clusters was estimated using the Bootstrap support value (BSV) obtained from 1,000 replicates. Two mosquito sequences of *Culex bidens* (Diptera: Culicidae) were used as outgroup (GenBank numbers: KY581209 and KY581211). The automatic barcode gap discovery (ABGD) approach⁽¹³⁾ was employed to find barcode gaps for the primary species or OTUs delimitation, and to corroborate results obtained in the NJ topology. In this study, the default range of intraspecific divergence was limited between 0.001 and 0.1. Sequences from GenBank were obtained as references in the species delimitation analysis: *Ev. evandroi* (2), *Ev. termitophila* (4), *Expapillata firmatoii* (2), *Lutzomyia longipalpis* (6), *Mg. migonei* (5), *Pintomyia bianchigalatae* (2), *Pi. fischeri* (4), *Pi. misionensis*

(2), *Pi. monticola* (4), and *Sciopemyia sordellii* (2) from Brazil; Colombia, and Perú. These species were selected because they were recorded from Argentina and belong to the subtribe *Lutzomyiina*, as *Ev. cortelezzii* and *Mg. migonei*. We attempted to assess the degree of divergence of the first barcode sequences for *Ev. cortelezzii* and *Mg. migonei* for the science and Argentina, respectively.

Sixty-two specimens were captured and were morphologically identified as follows (Table): *Ev. cortelezzii* s.s. (n = 12), *Cortelezzii* complex (n = 49), and *Mg. migonei* (n = 1). Five males of *Ev. cortelezzii* and one of *Mg. migonei* were sequenced. The mean K2P sequence distance within the 13 nominal species was 1.03%. The frequency histogram of mean *COI* K2P sequence intra- and interspecific divergences may be seen in Fig. 1,

TABLE
Information of specimens of Phlebotominae collected between March and April 2018 in Córdoba City

Species	Month	Site	Sex	N
Cortelezzii complex		BG	F	7
<i>Evandromyia cortelezzii</i> s.s.			M	6
Cortelezzii complex	March	SMUR	F	12
<i>Ev. cortelezzii</i> s.s.			M	1
<i>Migonemyia migonei</i>			M	1
Cortelezzii complex		ZG	F	1
Cortelezzii complex		BG	F	14
<i>Ev. cortelezzii</i> s.s.	M		4	
Cortelezzii complex	April	SMUR	F	14
<i>Ev. cortelezzii</i> s.s.			M	1
Cortelezzii complex		ZG	F	1

BG: bajo grande sewage treatment plant of Córdoba municipality; F: female; M: male; SMUR: San Martín Urban Reserve; ZG: zoo garden.

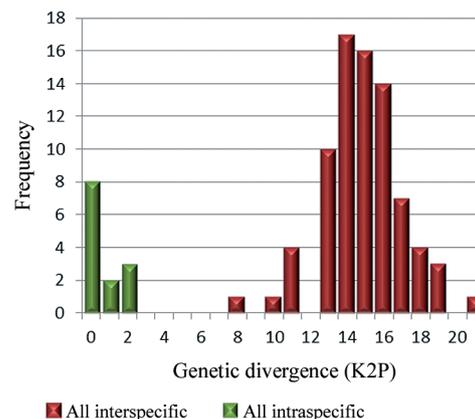


Fig. 1: frequency distribution of intraspecific and interspecific pairwise genetic divergence of the specimens belonging to the subtribe *Lutzomyiina* from Argentina, Brazil, Colombia, and Perú. Pairwise genetic distances were calculated using Kimura's two parameter (K2P) distance.

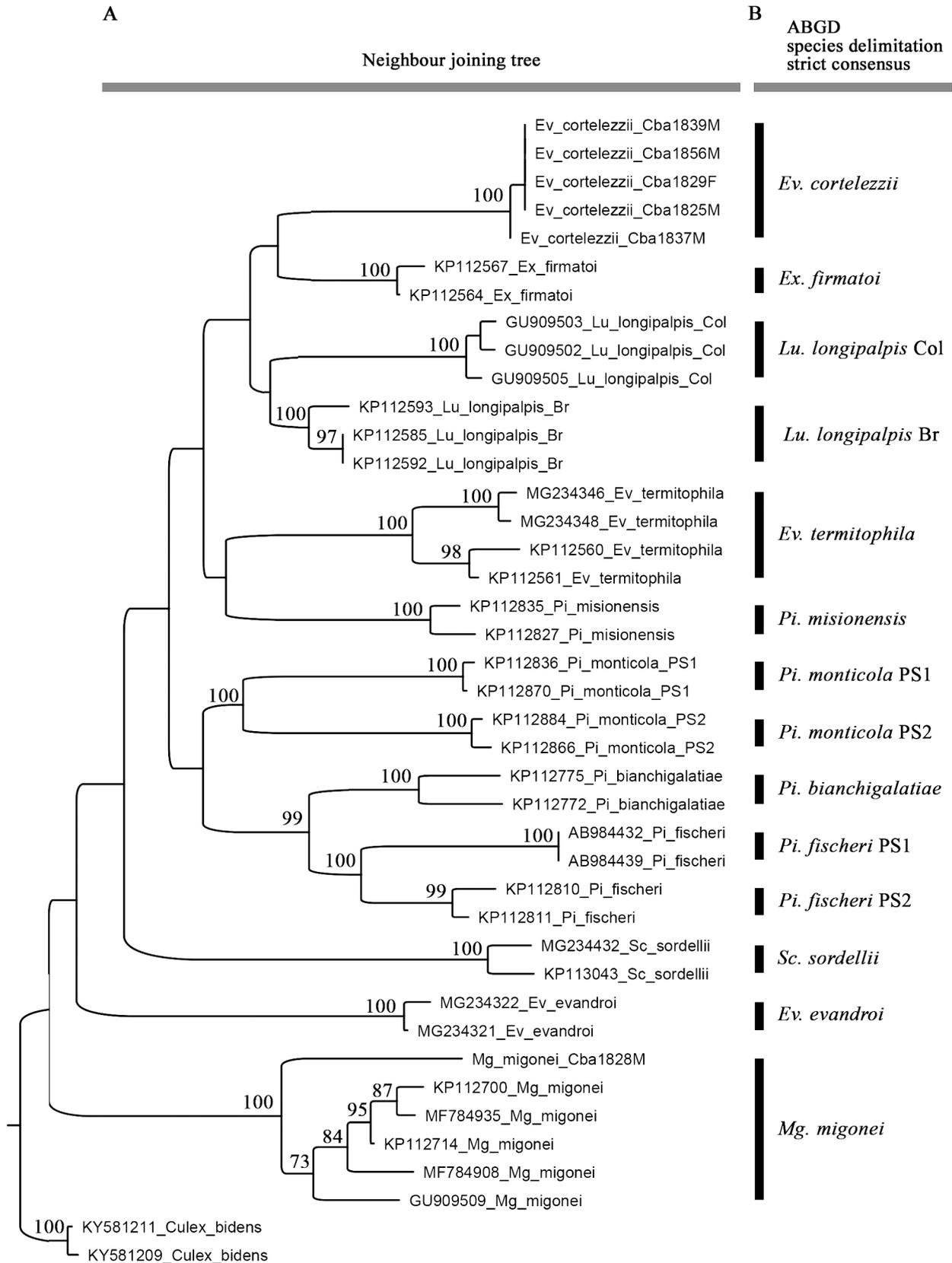


Fig. 2: neighbour-joining tree and species delimitation analysis of 41 cytochrome *c* oxidase subunit I (*COI*) sequences generated from specimens belonging to the subtribe *Lutzomyiina* from Argentina, Brazil, Colombia, and Perú. (A) neighbour-joining tree. Numbers at branches indicate bootstrap support value ($\geq 70\%$); (B) automatic barcode gap discovery (ABGD) analysis.

showing that a clear barcoding gap was established. The NJ tree was built using the K2P distances between the specimens, and the primary species delimitation from the strict partition obtained from the ABGD analysis, as shown in Fig. 2. The ABGD analysis provided ten possible partitions for these specimens (Fig. 3). The most conservative partitioning (strict partitioning) identified 14 OTUs (Fig. 2) with values of prior intraspecific divergence between 0.17% and 3.5% because values below 0.17% or above 3.5% can over- or underestimate the number of species, respectively. The 14 OTUs mentioned were identified from 13 morphological species because one of the original groups, *Pi. fischeri*, was split into two barcode groups (Fig. 2), showing a deep intraspecific variation. The clusters enclosing *Pi. fischeri*, provisional species (PS1 and PS2), showed group sequences from Brazil and Perú, separately. A plausible explanation is related to the historical events of vicariance, which can reduce gene flow between subpopulations of the same species, leading to speciation.⁽²⁰⁾ Over time, morphological differentiation tends to take longer because changes in the morphological traits require changes in the multiple genes.⁽²¹⁾ The sequences of *Lutzomyia longipalpis* from Brazil and Colombia were considered *a priori* as the two provisional species because the species in Brazil showed a deep intraspecific variation between the regions of the country, as evidenced by Pinto et al.⁽¹⁰⁾ The previously mentioned concept has been supported by the outcomes. The sequences of *Pi. monticola* were also treated as the two provisional species (PS1 and PS2) because they belong to a cryptic species complex.⁽¹⁰⁾ The sequences of the nominal species *Mg. migonei* formed a well-supported cluster (100% BSV), with a mean intraspecific K2P distance of 2.24% and was recognised as an OTU by the ABGD partitioning (Fig. 2). It is important to note that *Mg. migonei* sequences belonging to Argentina differ from those found in Brazil and Colombia (Fig. 2). Novel sequences of *Ev. cortelezzii* were grouped in a well-supported cluster (100% BSV) and diagnosed as the primary species by the ABGD algorithm (Fig. 2).

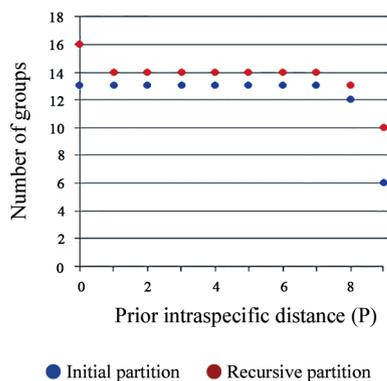


Fig. 3: number of groups among the 39 specimens of sand flies from Argentina, Brazil, Colombia, and Perú based on the values of prior intraspecific divergence found by the automatic barcode gap discovery (ABGD) software as potential barcode gaps using a range of 0.001 to 0.1 for prior intraspecific divergence.

Even though all the *Ev. cortelezzii* specimens that were sequenced had been collected in Córdoba Province in Argentina, it is worth noting that they were not clustered with other congeneric species that were also considered, as shown by Pinto et al.⁽¹⁰⁾ In the latter, sequences of *Ev. carmelioni* were grouped with *Ev. lenti*, from specimens collected in the distant states of Espírito Santo and Mato Grosso in Brazil and were not recognised as species by the ABGD partitioning. In our analyses, three species of the genus *Evandromyia* from Argentina and Brazil were diagnosed as different OTUs.

Mg. migonei and *Ev. cortelezzii* are widely distributed in Argentina,⁽³⁾ both species being important in public health. *Mg. migonei* is associated with the agents of visceral Leishmaniasis (VL) and TL, *Leishmania braziliensis*, and *L. infantum*, respectively,^(22,23) and was considered the putative vector of VL in Santiago del Estero Province.⁽²²⁾ Female specimens of the *Cortelezzii* complex were found naturally infected with *L. braziliensis* in Chaco Province.⁽²⁴⁾

The first sequences of *COI* barcode region are now available for sand flies for Argentina. The sequence of *Mg. migonei* in its southernmost known distribution and the sequences of *Ev. cortelezzii* are the first for the species and allow the molecular identification of the females of the complex. Further studies should be carried out to obtain *COI* sequences from the males of *Ev. sallesi* and to solve identification problems of females of the complex. These sequences are useful tools for the future molecular identification of these species in other locations in the New World.

AUTHORS' CONTRIBUTION

ML, IMO and WRA - Conceptualisation, investigation, resources and writing - original draft; ML and IMO - formal analysis and methodology; WRA - funding acquisition and supervision.

REFERENCES

1. WHO - World Health Organization. Control of the leishmaniasis: report of a meeting of the WHO Expert Committee on the Control of Leishmaniasis. World Health Organization Technical Report Series. 2010; 949: 22-6.
2. Salomón OD, Rosa JR, Stein M, Quintana MG, Fernández MS, Visintín AM, et al. Phlebotominae (Diptera: Psychodidae) fauna in the Chaco region and Cutaneous Leishmaniasis transmission patterns in Argentina. Mem Inst Oswaldo Cruz. 2008; 103(6): 578-84.
3. Salomón OD, Quintana MG, Rosa JR. Ecoepidemiología de la Leishmaniasis cutánea en la Argentina. Salud(i)Ciencia. 2008; 16(5): 514-20.
4. Visintín AM, Beranek MD, Amieva MJ, Rosa JR, Almirón WR, Salomón OD. Spread of Phlebotominae in temperate climates: province of Córdoba, Argentina. Mem Inst Oswaldo Cruz. 2016; 111(1): 75-8.
5. Ontivero IM, Beranek MD, Rosa JR, Ludueña-Almeida FF, Almirón WR. Seasonal distribution of phlebotomine sandfly in a vulnerable area for tegumentary leishmaniasis transmission in Córdoba, Argentina. Acta Trop. 2018; 178: 81-5.
6. Galati EAB. Morfologia, terminologia de adultos e identificação dos táxons da América. In: Rangel EF, Lainson R, orgs. Flebotomíneos do Brasil. Rio de Janeiro: Fiocruz; 2003. 53-175.

7. Szelag EA, Rosa JR, Galati EA, Andrade Filho JD, Salomón OD. Considerations on the species complex of the *Corteleezzii* series (Diptera: Psychodidae) and description of *Evandromyia chacuensis* sp. nov., a new Phlebotomine species of the Chaco Region, Argentina. *J Med Entomol.* 2018; 55(4): 902-9.
8. Hebert PDN, Cywinska A, Ball SL, de Waard JR. Biological identifications through DNA barcodes. *Proc R Soc Lond B Biol Sci.* 2003; 270(1512): 313-21.
9. Hoyos-López R, Uribe S, Vélez I. [Typification of Colombian specimens of *Lutzomyia longipalpis* (Diptera: Psychodidae) by "Barcoding"]. *Rev Colomb Entomol.* 2012; 38(1): 134-40.
10. Pinto IDS, Chagas BD, Rodrigues AAF, 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.
11. Rodrigues BL, Carvalho-Costa LF, Pinto IDS, Rebêlo JMM. 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.
12. Beeble NW. DNA barcoding mosquitoes: advice for potential prospectors. *Parasitology.* 2018; 145(5): 622-33.
13. Puillandre N, Lambert A, Brouillet S, Achaz G. ABGD, automated barcode gap discovery for primary species delimitation. *Mol Ecol.* 2012; 21: 1864-77.
14. Marcondes CB. A proposal of generic and subgeneric abbreviations for phlebotomine sandflies (Diptera: Psychodidae: Phlebotominae) of the world. *Entomol News.* 2007; 118(4): 351-7.
15. 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: 294-9.
16. Hall TA. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series.* 1999; 41(41): 95-8.
17. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004; 32(5): 1792-7.
18. Kimura M. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol.* 1980; 16(2): 111-20.
19. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol Biol Evol.* 2016; 33(7): 1870-4.
20. Avise JC. *Phylogeography: the history and formation of species.* Cambridge: Harvard University Press; 2000.
21. Coyne JA, Orr HA. *Speciation.* Sunderland: Sinauer Associates, Inc; 2004.
22. Salomón OD, Quintana MG, Bezzi G, Morán ML, Betbeder E, Valdéz DV. *Lutzomyia migonei* as putative vector of visceral leishmaniasis in La Banda, Argentina. *Acta Trop.* 2010; 113(1): 84-7.
23. Moya SL, Giuliani MG, Acosta MM, Salomón OD, Liotta DJ. First description of *Migonemyia migonei* (França) and *Nyssomyia whitmani* (Antunes & Coutinho) (Psychodidae: Phlebotominae) natural infected by *Leishmania infantum* in Argentina. *Acta Trop.* 2015; 152: 181-4.
24. Rosa JR, Pereira P, Brazil RP, Andrade Filho JD, Salomón OD, Szelag EA. Natural infection of *Corteleezzii* complex (Diptera: Psychodidae: Phlebotominae) with *Leishmania braziliensis* in Chaco, Argentina. *Acta Trop.* 2012; 123(2): 128-31.