



**Repositorio Institucional de la Universidad Autónoma de Madrid**

<https://repositorio.uam.es>

Esta es la **versión de autor** del artículo publicado en:

This is an **author produced version** of a paper published in:

Infection, Genetics and Evolution 18 (2013): 229-237

**DOI:** <http://dx.doi.org/10.1016/j.meegid.2013.05.020>

**Copyright:** © 2013 Elsevier

El acceso a la versión del editor puede requerir la suscripción del recurso

Access to the published version may require subscription

## Evolution and species discrimination according to the *Leishmania* heat-shock protein 20 gene

Jorge Fraga<sup>a</sup>, Ana Margarita Montalvo<sup>a</sup>, Gert Van der Auwera<sup>b</sup>, Ilse Maes<sup>b</sup>, Jean-Claude Dujardin<sup>b,c</sup>, José M. Requena<sup>d\*</sup>.

<sup>a</sup>Departamento de Parasitología, Instituto de Medicina Tropical “Pedro Kourí”, Habana, Cuba

<sup>b</sup>Department of BioMedical Sciences, Institute of Tropical Medicine, Antwerp, Belgium

<sup>c</sup>Department of BioMedical Sciences, University of Antwerp, Belgium

<sup>d</sup>Centro de Biología Molecular "Severo Ochoa" (CSIC-UAM), Madrid, Spain

\*Corresponding author. Tel.: +34 911964617; fax: +34 911964420. E-mail address:

[jmrequena@cbm.uam.es](mailto:jmrequena@cbm.uam.es) (J.M. Requena)

### Footnote:

Note: Nucleotide sequence data reported in this paper are available in GenBank, EMBL and DDBJ databases under the accession numbers JX630109-JX630148.

### Highlights

- *Hsp20* is a suitable molecular marker for *Leishmania* typing.
- The same pair of oligonucleotides allows *hsp20* amplification from different *Leishmania* species.
- Nine species and two subspecies are recognized among the *Leishmania* pathogenic species.
- A reliable taxonomy for *Leishmania* is possible based on molecular phylogenetic analyses.

## Abstract

The *Leishmania* genus comprises up to 35 species, of which 20 are responsible for human disease. However, the taxonomic status for many of them is under discussion. The small Heat Shock Proteins (sHSPs) have been extensively studied due to their importance in protecting cellular proteins from aggregation and maintaining cellular viability under intensive stress conditions. In *Leishmania*, a protein of this class was previously described, the 20-kDa heat-shock protein (HSP20), which is encoded by a single gene. In the present study, we used this target, alone or in combination with *hsp70* gene, to investigate the phylogenetic relationships among *Leishmania* species. Using a pair of degenerate primers it was possible amplifying a 370-bp fragment of the *hsp20* coding region in 39 strains of very different geographic origins, representing in total 16 *Leishmania* species. Nucleotide sequences were readily obtained by direct sequencing of the amplification products. Both phylogenetic trees and networks based on either *hsp20* sequences or combined datasets of *hsp20* and *hsp70* sequences were constructed. These phylogenic analyses allowed the division of the *Leishmania* genus into nine species: *L. (L.) donovani*, *L. (L.) major*, *L. (L.) tropica*, *L. (L.) aethiopica*, *L. (L.) mexicana*, *L. (V.) lainsoni*, *L. (V.) naiffi*, *L. (V.) guyanensis* and *L. (V.) braziliensis*. Additionally, by network analysis, the subspecies *L. (L.) donovani infantum* and *L. (V.) braziliensis peruviana* were recognized within the *L. (L.) donovani* and *L. (V.) braziliensis* species, respectively. Therefore, our phylogenetic analyses based on the *hsp20* sequence demonstrated that this gene is a suitable molecular marker for *Leishmania* typing and classification purposes. In addition, this study represents a solid contribution towards establishing a better and more reliable taxonomy for the genus *Leishmania*.

**Keywords:** HSP20; HSP70; Phylogeny; Molecular typing; Taxonomy; *Leishmania*

## 1. Introduction

The leishmaniasis are a complex of diseases caused by kinetoplastid flagellates of the genus *Leishmania*. Clinical manifestations range from simple self-healing cutaneous lesions (cutaneous leishmaniasis, CL) through metastasizing mucocutaneous forms (mucocutaneous leishmaniasis, MCL) to often lethal visceral disease (visceral leishmaniasis, VL). A recent epidemiological study recorded reports of endemic leishmaniasis transmission in a total of 98 countries and 3 territories on 5 continents. Furthermore, the global yearly incidence was estimated as 0.2 to 0.4 million cases of VL and 0.7 to 1.2 million of CL, while 12 million people worldwide are currently affected by the disease (Alvar et al., 2012). The genus *Leishmania* comprises some 35 species of morphologically similar kinetoplastid protozoa, among which 20 species are responsible for human disease. The severity of the clinical manifestation in immunocompetent persons depends on the particular infecting species. Therefore, distinguishing between species is crucial for the correct diagnosis and prognosis of the disease as well as for making decisions regarding treatment and control measures (Schönian et al., 2010).

In order to design reliable diagnostic tools, it is crucial to agree on a clear definition of the taxa to be identified (Bañuls et al., 2002). The classification of *Leishmania* was initially based on ecobiological criteria such as vector, geographical distribution, tropism, antigenic properties and clinical manifestations and, later, on immunological and biochemical data (Pratt and David, 1981; Lainson and Shaw, 1987; Schönian et al., 2010). Hierarchical taxonomic schemes have been proposed using the categories of subgenera, species complexes, species, and subspecies. The two subgenera *L.* (*Leishmania*) and *L.* (*Viannia*) are separated on the basis of their location in the vector's intestine (Lainson and Shaw, 1987), and the species within the subgenera were established using isoenzyme analysis (Rioux et al., 1990). However, the validity of these taxonomic classification schemes is still debated, and the status of some species are or have been questioned (Bañuls et al., 2007; Schönian et al.,

2010). It is therefore urgent to agree on the definition of species based on reliable and well-contrasted traits. The use of molecular sequence techniques is providing a large amount of additional data for phylogenetic analyses, but in spite of this, evolutionary relationships remain to be solved (Hughes and Piontkivska, 2003). Different molecular markers have been used for phylogenetic studies in *Leishmania*: the internal transcribed spacer (ITS) 1 and 2 of the ribosomal DNA array (Cupolillo et al., 1995; Dávila and Momen, 2000; Berzunza-Cruz et al., 2002; Orlando et al., 2002; Kuhls et al., 2005; Spanakos et al., 2008; Sukmee et al., 2008; Villinski et al., 2008), a repetitive DNA sequence (Piarroux et al., 1995), the gene for DNA polymerase  $\alpha$  catalytic subunit (POLA, Croan et al., 1997), the gene encoding the largest subunit of RNA polymerase II (RPOIILS, Croan and Ellis, 1996; Croan et al., 1997), the cytochrome oxidase II gene (COII, Ibrahim and Barker, 2001; Cao et al., 2011), the glycoprotein 63 gene (*gp63*, Mauricio et al., 2007), cysteine protease B genes (*cpb*, Hide et al., 2007), the mini-exon (Sukmee et al., 2008), 7SL RNA (Zelazny et al., 2005; Guan et al., 2012), the small subunit ribosomal RNA (SSU rRNA, Guan et al., 2012), and the cytochrome B gene (*cytB*, Luyo-Acero et al., 2004; Asato et al., 2009). Recently, Boité et al. (2012) evaluated the phylogeny of *Leishmania* (*Viannia*) parasites based on multilocus sequence analysis using different housekeeping genes. However, most of these studies were aimed to analyse isolates/strains from one particular geographic region, and only a few authors have investigated phylogenetic relationships within the entire genus (Croan and Ellis, 1996; Croan et al., 1997; Dávila and Momen 2000; Berzunza-Cruz et al., 2002; Luyo-Acero et al., 2004; Zelazny et al., 2005; Spanakos et al., 2008; Villinski et al., 2008; Asato et al., 2009; Guan et al., 2012).

In a recent work, we carried out a comprehensive phylogenetic study based on sequence analysis of a 1380 bp fragment of the *hsp70* coding region, including for the first time the 17 species commonly causing leishmaniasis and several strains per species (Fraga et al., 2010). As a result, eight monophyletic groups were clearly defined, four in each *Leishmania* subgenus (*Leishmania* and

*Viannia*). These groups correspond to the following species: *L. (L.) donovani*, *L. (L.) major*, *L. (L.) tropica*, *L. (L.) mexicana*, *L. (V.) lainsoni*, *L. (V.) naiffi*, *L. (V.) guyanensis* and *L. (V.) braziliensis*. Using phylogenetic network analysis, some subspecies were further recognized: *L. (L.) donovani infantum*, *L. (V.) guyanensis panamensis*, and *L. (V.) braziliensis peruviana*. More recently, phylogenetic data based on the 3'-untranslated region (UTR) of *hsp70-I* genes supported a separation between *L. (L.) tropica* and *L. (L.) aethiopica* (Requena et al., 2012), species that cannot be separated based on *hsp70* coding sequence evolutionary trees (Fraga et al., 2010).

In the search for additional molecular markers useful for improving the resolution of current phylogenetic classification, the *hsp20* gene was found to be a plausible candidate. Proteins belonging to heat-shock protein (HSP) family play an important role in folding, assembly, intracellular localization, secretion, regulation, stabilization and degradation of proteins (Young et al., 2004). Given their functional relevance, most HSPs are highly conserved in sequence and function (Folgueira and Requena, 2007). However, there is an exception, the small heat-shock proteins (sHSPs): although they are widespread in both eukaryotic and prokaryotic organisms, they are poorly conserved in sequence (Fu et al., 2006). Nevertheless, in spite of this low amino acid sequence conservation, the tridimensional structure of the  $\alpha$ -crystallin domain, which represents the signature motif of sHSPs, is well conserved. This domain is responsible for dimer formation, which is the basic functional unit of many sHSPs (Haslbeck et al., 2005). In addition, sHSPs contain amino- and carboxy-terminal extensions that are involved in modulating oligomerization, substrate binding and chaperone function (Sun and MacRae, 2005). In *Leishmania* and other trypanosomatids, only one member of the sHSPs exists, the 20-kDa heat-shock protein (HSP20) (Folgueira and Requena, 2007). Amino acid sequence alignment of HSP20s from *L. braziliensis*, *L. (L.) amazonensis*, *L. (L.) major* and *L. (L.) infantum* showed that this protein, even though conserved within the genus, has accumulated a significant number of amino acid substitutions in the different *Leishmania* species (Montalvo-Álvarez et al.,

2008). In this study, we investigate the discriminatory value of the *hsp20* coding region for phylogenetic analysis; the results indicate that this molecular marker is useful for species discrimination within the *Leishmania* genus, reinforcing and, even improving, previous phylogenetic classifications. In addition, we combined *hsp20* and *hsp70* sequence data in phylogenetic analyses as a strategy to further strengthen the species phylogeny within the *Leishmania* genus.

## **2. Material and Methods.**

### *2.1. PCR amplification of hsp20*

PCR primers were designed in conserved regions of the *hsp20* gene based on a multiple alignment of the previously characterized *hsp20* sequences: *L. (L.) amazonensis* (AM712297), *L. (V.) braziliensis* (XM\_001566535), *L. (L.) infantum* (XM\_001466741) and *L. (L.) major* (XM\_842817) (see Table 1). Finally, two degenerate primers, named *hsp20F* and *hsp20R*, we designed to amplify a fragment of around 370 bp (Table 2).

PCR products were amplified from the strains listed in Table 1. The reaction mix (50 µL) contained 1× standard PCR buffer including 1.5 mM MgCl<sub>2</sub>, 1× Q-buffer, 200 µM of each deoxynucleoside triphosphate, 0.5 U HotStarTaq Plus DNA polymerase (Qiagen, Hilden, Germany), 0.8 µM of each primer, and 10 ng of genomic DNA isolated from parasite culture. The thermal cycling parameters of the assay were: initial denaturation at 95 °C for 5 min; followed by 35 cycles consisting of 94 °C for 40 s, 58 °C for 1 min, and 72 °C for 1 min; and a final extension step of 8 min at 72 °C. Amplicons were analyzed on a 2% agarose gel, and sequenced directly without molecular cloning.

## 2.2. DNA sequencing

Both strands of the *hsp20* PCR products were sequenced using the oligonucleotides hsp20F and hsp20R as primers. DNA sequencing was carried out in the facilities of the Servicio de Genómica (Parque Científico de Madrid-UAM), using the Big Dye Terminators v3.1 kit (Applied Biosystem).

## 2.3. Phylogenetic analysis

For all analyses, the priming sites were trimmed from both ends of all sequences. In addition to the sequences generated in this study, previously published sequences (see Table 1 for a complete list of sequences) were also included in the alignments, which were carried out with the software package MEGA (Molecular Evolutionary Genetic Analysis Version 5.05, Tamura et al., 2011, [www.megasoftware.net](http://www.megasoftware.net)). The sequences alignment was screened for detect species-specific single nucleotide polymorphisms (SNPs) and indels. MEGA software was also used to build phylogenetic trees with both, distance and character-based methods, and to analyze synonymous versus non-synonymous nucleotide substitutions. The number of synonymous differences per synonymous site, and the number of nonsynonymous differences per non-synonymous site were averaged over all *Leishmania* sequence pairs, using the Nei–Gojobori method (Nei and Gojobori, 1986). Distances and characters from nucleotide sequences were estimated with the Kimura-2 parameter model (Kimura, 1980), and trees were built with the Neighbor-Joining (NJ) (Saitou and Nei, 1987), Minimum Evolution (ME) (Rzhetsky and Nei, 1992) and Maximum Parsimony (MP) (Eck and Dayhoff, 1966; Fitch, 1971) methods. Also we used the Maximum likelihood (ML) (Felsenstein, 1981) with appropriate substitution models (T92+I) chosen using the Akaike information criterion (AIC), as implemented in the MEGA 5.05 software. Distances from predicted amino acid sequences were determined with the *p*-distance model. As out-group, the *hsp20* sequences from two species of the genus *Trypanosoma*, evolutionarily related to *Leishmania*, were also used in our analysis. The support



of monophyletic groups was assessed by the bootstrap method (Felsenstein, 1985) with 2000 replicates. Additionally, phylogenetic networks were established with the SplitsTree4 software (Huson, 1998; Huson and Bryant, 2006), using the Kimura-2 parameter model (Kimura, 1980) and the Neighbor-Net method. Such networks depict alternative evolutionary paths supported by the data set.

Finally the *hsp20* and *hsp70* gene sequences (Table 1) were combined and analyzed using MEGA 5.05 and SplitsTree4 programs as specified above.

### 3. Results

A fragment covering 370-bp of the *hsp20* coding region (468-bp in length) was successfully PCR-amplified in 39 strains from 16 *Leishmania* species (Table 1), corresponding to the main *Leishmania* species and including isolates from different geographic origins. The PCR products were directly sequenced; Table 1 lists the accession numbers of the 39 newly determined sequences. Additionally, six *Leishmania hsp20* sequences present in the GenBank database were retrieved, amounting to a total of 45 sequences of different strains. For phylogenetic analyses, the *hsp20* sequences for *Trypanosoma brucei brucei* and *Trypanosoma cruzi* were retrieved from GenBank (Table 1).

The aligned sequence of *Leishmania hsp20* genes was 319-bp (primers were trimmed from the sequences), except for *L. (V.) lainsoni* strains that have 316-bp in the sequenced region. The *hsp20* sequences of *Leishmania* spp. were found to be GC rich (55.8 – 60.1%), with a sequence identity among *Leishmania* sequences between 82.2 and 100% (overall average 90.4%). The nucleotide sequence variation is sufficient to discriminate parasite species: 88 nucleotide positions (27.6%) are polymorphic and 82 positions (25.7%) are parsimony informative. Table 3 shows the 26 SNPs and indel positions that discriminate among *Leishmania* species studied. Species-specific SNPs were

detected for *L. (L.) infantum*, *L. (L.) tropica*, *L. (L.) major*, *L. (L.) aethiopica*, *L. (L.) amazonensis*, *L. (V.) naiffi*, *L. (V.) lainsoni*, *L. (V.) guyanensis* and *L. (V.) peruviana*. Also 3 positions represent deletions specific for the *L. (V.) lainsoni* species. No species-specific SNPs were detected for *L. (V.) braziliensis*, *L. (L.) donovani*, *L. (L.) mexicana* and *L. (L.) garnhami*.

The analysis of deduced amino acids for the sequenced *hsp20* region (105 amino acids) revealed substitutions at 29 positions (27.6%), of which 28 sites (26.6%) are parsimony informative. The number of synonymous substitutions per synonymous site (*dS*) is 32.6%, the number of non-synonymous substitutions per non-synonymous site (*dN*) is 5.6%. The ratio *dN* to *dS* substitutions in *hsp20* was  $dN/dS = 0.17$ , suggesting that this gene has been subject to purifying selection. The average dissimilarity between the *Leishmania* and *Trypanosoma* species is 47.2 %, which is three times as high as the highest intra-*Leishmania* values (17.2 %).

The Neighbor-Joining tree based on the *hsp20* sequence alignment (Fig. 1) shows that the subgenera *L. (Leishmania)* and *L. (Viannia)* are separated into distinct monophyletic clades. Within the *L. (Leishmania)* subgenus, the Old and New World *Leishmania* species are located in two different branches of the tree. Within the New World, 5 species/complexes can be reliably recognized: the *L. (L.) mexicana* complex, *L. (V.) lainsoni*, *L. (V.) naiffi*, the *L. (V.) guyanensis* complex, and the *L. (V.) braziliensis* complex. In the Old World, *hsp20* based trees support a separation of *L. (L.) major*, the *L. (L.) donovani* complex, *L. (L.) aethiopica* and *L. (L.) tropica*. All these clusters were also observed using ML, ME and MP phylogenies (data not shown), indicating that the derived groups are robust and not dependent of the evolutionary method used. Trees based upon amino acid sequences are in agreement with the nucleotide-based phylogenies (data not show).

Figure 2 displays a phylogenetic network obtained from the same sequences as in Fig. 1, excluding the *Trypanosoma* sequences. This network shows a separation of *L. (L.) infantum* (bootstrap

value 70 %) and *L. (V.) peruviana* (bootstrap value 87.4 %) as subgroups within the *L. (L.) donovani* and *L. (V.) braziliensis* complexes, respectively.

In order to appraise whether a combined analysis of *hsp20* and *hsp70* gene sequences would further improve the phylogenetic relationships within the *Leishmania* genus, *hsp20* and *hsp70* sequences from 42 strains of different geographic origin, accounting for 16 different *Leishmania* species (see Table 1), were combined and analyzed as before. The analysis, using NJ (Fig. 3), ML, MP and ME, showed a consistent separation of *Leishmania* genus into 9 monophyletic clusters, corresponding to the *L. (L.) mexicana* complex, *L. (V.) lainsoni*, *L. (V.) naiffi*, the *L. (V.) guyanensis* complex, the *L. (V.) braziliensis* complex, *L. (L.) major*, the *L. (L.) donovani* complex, *L. (L.) aethiopica* and *L. (L.) tropica*. In addition, this analysis allowed a distinction of *L. (V.) peruviana* as a subgroup within the *L. (V.) braziliensis* complex. The phylogenetic network obtained from the same sequences as in Fig. 3, excluding the *Trypanosoma* sequences, showed a separation of *L. (L.) infantum* (bootstrap value 74.5 %) and *L. (V.) peruviana* (bootstrap value 99 %) as subgroups within the *L. (L.) donovani* and *L. (V.) braziliensis* complexes, respectively (data not show).

#### 4. Discussion

Although sHSPs and their  $\alpha$ -crystallin domain have been used for phylogenetic purposes, being considered a good molecular marker for analyzing the evolution of species (Kriehuber et al., 2010), this is the first time that *hsp20* sequences have been employed for taxonomy of the *Leishmania* genus. In addition, the newly identified SNPs could be also used for *Leishmania* species typing, especially by PCR-RFLP and/or sequencing. Comparison of synonymous (silent, *dS*) and nonsynonymous (aminoacid changing, *dN*) nucleotide substitution rates in protein coding genes provides an important

means for understanding molecular evolution. The ratio of  $dN$  to  $dS$  obtained for the *hsp20* sequences was 0.17; this value confirms a purifying selection in this gene due to functional constriction of the protein. However, the value is large enough to support the use of *hsp20* coding sequences for analysis of the phylogenetic relationships within the *Leishmania* genus at the species, intra- and supra-species levels.

Our *hsp20* phylogenetic study included 45 strains from different geographic origin, belonging to 16 MLEE-defined *Leishmania* species that represent the most common causative agents of leishmaniasis in the New and Old World. As expected, *Leishmania hsp20* sequences were clearly distinct from the homologous sequences in *T. brucei* and *T. cruzi*. Also, our analyses agree with the division of mammalian species, proposed by Lainson and Shaw (1987), into the peripylarian *Viannia* and the suprapylarian *Leishmania* subgenera, both of which form distinct monophyletic clusters (Fig. 1 and 3). While *L. (Viannia)* is restricted to neotropical regions, *L. (Leishmania)* occurs both in the New (neotropical and southern neoarctic) and Old (palearctic, African and oriental) World (Kerr, 2000). This geographical dichotomy of the *Leishmania* subgenus (except for *L. (L.) infantum* / *L. (L.) chagasi*) is also reflected in the phylogenetic tree (Fig. 1 and 3), which shows a separation between the New (*L. (L.) mexicana* complex) and Old World groups. Similar conclusions were drawn with phylogenetic studies based on sequences from *polA* and *rpoHLS* (Croan et al., 1997), ITS rDNA (Dávila and Momen, 2000), 7SL RNA (Zelazny et al., 2005), *cytB* (Luyo-Acero et al., 2004; Asato et al., 2009) and *hsp70* (Fraga et al., 2010).

As shown in Fig. 1, the phylogenetic analysis based on the *hsp20* sequence allowed the separation of the mammalian-infecting *Leishmania* species, analyzed in this work, into 9 monophyletic clusters. Some of these clusters correspond to species as defined by standard MLEE (Table 1), but often they group several species. In recent times, the usefulness for taxonomic purposes of MLEE is under dispute, and the results of this work further support the idea that MLEE-species definition should

be left aside and definitely replaced for a classification based exclusively on genetic markers (da Silva et al., 2010; Van der Auwera et al., 2011; Requena et al., 2012). On the other hand, regarding the *Leishmania* phylogeny based on *hsp70* coding sequences (Fraga et al., 2010), the *hsp20* phylogeny and the combined analysis of *hsp20* and *hsp70* sequences reinforced previous conclusions, but, interestingly, allowed the discrimination of *L. (L.) tropica* and *L. (L.) aethiopica*, species which cannot be distinguished through evolutionary analysis based on *hsp70* coding sequences. Recently, using the 3'-UTR of *hsp70-I* genes as molecular marker, this separation was also strongly supported with a bootstrap value of 99 (Requena et al., 2012). Evidence of this separation was also obtained from phylogenetic studies based on non-coding sequences like ITS of the ribosomal DNA array and 7SL RNA gene (Dávila and Momen, 2000; Berzunza-Cruz et al., 2002; Zelazny et al., 2005; Spanakos et al., 2008; Villinski et al., 2008), but both species form a single cluster when coding sequences for *gp63* or *cytB* were used (Croan et al., 1997; Luyo-Acero et al., 2004; Mauricio et al., 2007; Asato et al., 2009). These two species are certainly very close, as the rock hyrax (*Procavia capensis*), the proven animal reservoir for *L. (L.) aethiopica*, is also the most likely original reservoir for *L. (L.) tropica* (Talmi-Frank et al., 2010). However, infection of *L. (L.) aethiopica* in man results in a spectrum of diseases ranging from simple self-healing CL to mucocutaneous forms, with lesions spreading into the nasal and/or oral mucosa, whereas *L. (L.) tropica* only produces CL (Bryceson et al., 1969; Hailu et al., 2005). All these elements suggest that *L. (L.) tropica* and *L. (L.) aethiopica* may be considered as two distinct species. Thus, on the basis of the molecular taxonomy, recently proposed by Fraga et al. (2010), and taking into account the results found in this work, we propose grouping the *Leishmania* strains analyzed in this study into the following nine species: *L. (L.) donovani*, *L. (L.) major*, *L. (L.) tropica*, *L. (L.) aethiopica*, *L. (L.) mexicana*, *L. (V.) lainsoni*, *L. (V.) naiffi*, *L. (V.) guyanensis* and *L. (V.) braziliensis*. Furthermore, within the *L. (V.) braziliensis* species (or complex), *L. (V.) peruviana* strains form a separate subgroup (Fig. 1 and 3), and, therefore, we suggest the status of subspecies for

this subgroup. The separation of *L. (V.) peruviana* is also strongly supported by the phylogenetic network analysis (Fig. 2).

In the literature, there is some controversy regarding whether *L. (V.) braziliensis* and *L. (V.) peruviana* should be considered either a sole species (Grimaldi et al., 1987; Arana et al., 1990), a heterogeneous species, with *L. (V.) peruviana* being a subspecies of *L. (V.) braziliensis*, or two distinct species (Dujardin et al., 1993; 1998; Victoir et al., 1998; Bañuls et al., 2000; García et al., 2005). Both species have been separated by MLEE (Bañuls et al., 2000), molecular karyotyping (Dujardin et al., 1993; 1995), RAPD (Bañuls et al., 1999; 2000), MLST (Tsukayama et al., 2009), *gp63* PCR-RFLP (Victoir et al., 1988) and *cytB* PCR-RFLP (García et al., 2005). The *hsp70* phylogeny supported a subspecies status for *L. (V.) peruviana* within *L. (V.) braziliensis* (Fraga et al., 2010). Odiwuor et al. (2012) studied the *L. (V.) braziliensis* complex genetically using amplified fragment length polymorphisms (AFLP) and also suggested that *L. (V.) peruviana* should be considered a subgroup of the *L. (V.) braziliensis* species, rather than a separated species. Recently Boité et al. (2012), using G6PD, 6PGD and MPI, constructed individual NJ trees and found that *L. (V.) peruviana* clustered within the *L. (V.) braziliensis* group. In agreement with these studies, our network analysis positioned *L. (V.) peruviana* as a subgroup separated from the *L. (V.) braziliensis* strains (Fig.2), even though they form a single node. Thus, we suggest considering *L. (V.) braziliensis* as a sole species, which is composed of two subspecies: *L. (V.) b. braziliensis* and *L. (V.) b. peruviana*.

Within the *L. (L.) donovani* complex, four species have been described: *L. (L.) donovani*, *L. (L.) archibaldi*, *L. (L.) infantum*, and *L. (L.) chagasi* (Bañuls et al., 2007). All of them cause mainly visceral leishmaniasis in tropical and sub-tropical regions, even though asymptomatic infections are common (Rijal et al., 2010). However, the separation of these four species is not supported by most of the phylogenetic studies based on sequence analysis (Piarroux et al., 1995; Croan et al., 1997; Dávila and Momen, 2000; Berzunza-Cruz et al., 2002; Kuhls et al., 2005; Luyo-Acero et al., 2004; Mauricio et al.,

2007; Asato et al., 2009; Fraga et al., 2010; Cao et al., 2011). According to our data, based either on *hsp20* sequences or on a combination of *hsp20* and *hsp70* sequences, *L. (L.) archibaldi* is indistinguishable from *L. (L.) donovani*. In fact, many authors are proposing some time ago to synonymize *L. (L.) archibaldi* with *L. (L.) donovani* (Kuhls et al., 2005). Also, *L. (L.) infantum* and *L. (L.) chagasi* are indistinguishable each other (Fig. 1-3). Similar conclusion has been arisen after phylogenetic studies based on RAPD, microsatellites or DNA sequences; thus, there is a general agreement that *L. (L.) chagasi* and *L. (L.) infantum* are synonym names for the same *Leishmania* species. This is consistent with the theory of the introduction of *L. (L.) infantum* in the New World from a Portuguese population which had crossed the Atlantic Ocean most probably in the XVth century via infected dogs (Rioux et al., 1990; Cupolillo et al., 1994; Mauricio et al., 2000; Lukeš et al., 2007; Leblois et al., 2011). Thus, currently, within the *L. (L.) donovani* complex, the name of *L. (L.) donovani* is maintained for strains of this complex present in Asian and East Africa, whereas the Mediterranean basin and American strains are classified as *L. (L.) infantum* (Mauricio et al., 2000; Lukeš et al., 2007; Schönian et al., 2010). Support for this separation has been obtained by microsatellite and multilocus sequence typing (MLST) of the *L. (L.) donovani* complex (Mauricio et al., 2006; Kuhls et al., 2007). However, most of the phylogenetic trees based on sequence analysis do not show a clear separation between *L. (L.) donovani* and *L. (L.) infantum* strains (Fig. 1 and 3). Only phylogenetic network analysis either using *hsp70* sequences (Fraga et al., 2010) or *hsp20* (Fig. 2) allowed a distinction between *L. (L.) infantum* and *L. (L.) donovani* strains. Thus, we propose considering *L. (L.) donovani* as a species composed by two subspecies (*L. d. donovani* and *L. d. infantum*). Such subspecies are well justified as they have different vectors, reservoirs, and geographical distribution (Schönian et al., 2010).

Although we have maintained the term complex to define some phylogenetic groups (i.e. *L. (L.) donovani*, *L. (L.) mexicana*, *L. (V.) braziliensis* and *L. (V.) guyanensis*), this was for historical reasons.

Nevertheless, we propose to abandon the concept of species complexes, as none of them seems to consist of different monophyletic groups. Other species such as *L. (L.) pifanoi*, *L. (L.) aristidesi*, *L. (L.) venezuelensis*, *L. (L.) foralini* and *L. (V.) shawi* should also be studied in future phylogenetic analyses with this molecular marker in order to define their species or subspecies status inside of the *Leishmania* genus (Bañuls et al., 2007; da Silva et al., 2010; Schönian et al., 2010).

Unraveling the phylogenetic relationships that underlie the origin of contemporary taxa requires the study of independent genes that display different evolutionary constraints (Phillipe 1998), as the history of a gene might be different from that of the species in which it resides. In this regard, as it was demonstrated for the *hsp70* gene (Fraga et al., 2010; da Silva et al., 2010), *hsp20* represents a suitable molecular probe for *Leishmania* typing, complementing and reinforcing phylogenetic studies based on *hsp70* coding sequences. Thus, increasing the number of both suitable molecular markers and analyzed strains should be a forthcoming objective to establish a definitive and non-ambiguous taxonomy for the genus *Leishmania*.

## 5. Conclusions

The classification of *Leishmania* parasites in the nine species proposed in this work, which is supported by phylogenetic studies based on *hsp20* sequences, either alone or in combination with *hsp70* sequences, represents a framework to clarify the evolutionarily relationships existing among the different species and the status of some of them. Knowing the evolutionary history of the *Leishmania* genus not only helps to build a reliable and objective taxonomic classification system, but also allows extrapolation of parasite-linked biological features to closely related strains. Reduction, simplification



and delineation of the actual taxonomic *Leishmania* framework are crucial to ensure attaining the highest possible degree of practical relevance (Van der Auwera et al., 2010).

On the other hand, the use of *hsp20* as a molecular marker for both typing and phylogenetic purpose has clear advantages on many other molecular markers: (i) a single pair of oligonucleotides suffices to PCR amplify of *hsp20* from any of the *Leishmania* species; (ii) the relative low size of the amplicon (around 370-bp) allows obtaining the complete sequence in a single reaction.

## Acknowledgements

The authors would like to thank all colleagues and Institutes who kindly donated the *Leishmania* reference strains or DNA, among whom J. Arévalo (Instituto de Medicina Tropical Alexander von Humboldt, Lima, Peru); L. García (Centro Universitario de Medicina Tropical, Cochabamba, Bolivia); E. Cupolillo (Instituto Oswaldo Cruz, Rio de Janeiro, Brazil); G. Schönián (Institut für Mikrobiologie und Hygiene, Berlin, Germany); I. Mauricio and D. Evans (London School of Hygiene and Tropical Medicine, London, UK); P. Desjeux (Instituto Boliviano de Biología de Altura, La Paz, Bolivia); J. P. Dedet and J. A. Rioux (Centre National de Référence des Leishmania, Montpellier, France); J.J. Shaw (University of Sao Paulo, Sao Paulo, Brazil); G. Schoone and A. El Harith (Royal Tropical Institute, Amsterdam, The Netherlands); I. Felger (Swiss Tropical and Public Health Institute, Swiss); and the Leishpinet consortium (EU contract INCO-CT2005-015407). This work has been funded in part by the third framework program of the Belgian Directorate General for Development with ITM-A. The work at JMR's lab was supported by grants from the Ministerio de Ciencia y Tecnología (BFU2009-08986), Comunidad Autónoma de Madrid (S2010/BMD-2361) and Fondo de Investigaciones Sanitarias (ISCIII- RD12/0018/0009-FEDER). Also, an institutional grant from Fundación Ramón Areces is acknowledged.

## References

- Alvar, J., Vélez, I.D., Bern, C., Herrero, M., Desjeux, P., Cano, J., Jannin, J., den Boer, M., the WHO Leishmaniasis Control Team., 2012. Leishmaniasis Worldwide and Global Estimates of Its Incidence. PLoS ONE 7, e35671.
- Arana, M., Evans, D.A., Zolessi, A., Cuentas, A.L., Arevalo, J., 1990. Biochemical characterization of *Leishmania (Viannia) braziliensis* and *Leishmania (Viannia) peruviana* by isoenzyme electrophoresis. Trans. R. Soc. Trop. Med. Hyg. 84, 526-529.
- Asato, Y., Oshiro, M., Myint, C.K., Yamamoto, Y., Kato, H., Marco, J.D., Mimori, T., Gomez, E.A., Hashiguchi, Y., Uezato, H., 2009. Phylogenetic analysis of the genus *Leishmania* by cytochrome b gene sequencing. Exp. Parasitol. 121, 352-361.
- Bañuls, A.L., Dujardin, J.C., Guerrini, F., De Doncker, S., Jacquet, D., Arevalo, J., Noel, S., Le Ray, D., Tibayrenc, M., 2000. Is *Leishmania (Viannia) peruviana* a distinct species? A MLEE/RAPD evolutionary genetics answer. J. Eukaryot. Microbiol. 47, 197-207.
- Bañuls, A.L., Hide, M., Prugnotte, F., 2007. *Leishmania* and the leishmaniasis: a parasite genetic update and advances in taxonomy, epidemiology and pathogenicity in humans. Adv. Parasitol. 64, 1-109.
- Bañuls, A.L., Hide, M., Tibayrenc, M., 1999. Molecular epidemiology and evolutionary genetics of *Leishmania* parasites. Int. J. Parasitol. 29, 1137-1147.
- Bañuls, A.L., Hide, M., Tibayrenc, M., 2002. Evolutionary genetics and molecular diagnosis of *Leishmania* species. Trans. R. Soc. Trop. Med. Hyg. 96, S9-13.
- Berzunza-Cruz, M., Cabrera, N., Crippa-Rossi, M., Sosa-Cabrera, T., Pérez-Montfort R, Becker, I., 2002. Polymorphism analysis of the internal transcribed spacer and small subunit of ribosomal RNA genes of *Leishmania mexicana*. Parasitol. Res. 88, 918-925.

- Boité, M.C., Mauricio, I.L., Miles, M.A., Cupolillo, E. 2012. New insights on taxonomy, phylogeny and population genetics of *Leishmania* (*Viannia*) parasites based on multilocus sequence analysis. PLoS Negl. Trop. Dis. 6: e1888.
- Bryant, D., Moulton, V., 2004. Neighbor-Net: An agglomerative method for the construction of phylogenetic networks. Mol. Biol. Evol. 21, 255-265.
- Bryceson A, Foster WA, Lemma A, 1969. Clinical trial of CI-501 (Camolar) against cutaneous leishmaniasis in Ethiopia. Trans R Soc Trop Med Hyg 63: 152–153.
- Cao, D. P., Guo, X. G., Chen, D. L., Chen, J. P., 2011. Species delimitation and phylogenetic relationships of Chinese *Leishmania* isolates reexamined using kinetoplast cytochrome oxidase II gene sequences. Parasitol. Res. 109, 163-73.
- Croan, D., Ellis, J. 1996. Phylogenetic relationships between *Leishmania*, *Viannia* and *Sauroleishmania* inferred from comparison of a variable domain within the RNA polymerase II largest subunit gene. Mol. Biochem. Parasitol. 79, 97-102.
- Croan, D.G., Morrison, D.A., Ellis, J.T., 1997. Evolution of the genus *Leishmania* revealed by comparison of DNA and RNA polymerase gene sequences. Mol. Biochem. Parasitol. 89, 149-159.
- Cupolillo, E., Grimaldi, G. Jr., Momen, H., 1994. A general classification of New World *Leishmania* using numerical zymotaxonomy. Am. J. Trop. Med. Hyg. 50, 296-311.
- Cupolillo, E., Grimaldi, G., Momen, H., Beverley, S.M., 1995. Intergenic region typing (IRT): A rapid molecular approach to the characterization and evolution of *Leishmania*. Mol. Biochem. Parasitol. 73, 145-155.
- da Silva, L. A., de Sousa Cdos, S., da Graca, G. C., Porrozzi, R., Cupolillo, E., 2010. Sequence analysis and PCR-RFLP profiling of the *hsp70* gene as a valuable tool for identifying *Leishmania* species associated with human leishmaniasis in Brazil. Infect. Genet. Evol. 10, 77-83.

- Dávila, A.M., Momen, H., 2000. Internal-transcribed-spacer (ITS) sequences used to explore phylogenetic relationships within *Leishmania*. *Ann. Trop. Med. Parasitol.* 94, 651-654.
- Dujardin, J.C., Arevalo, J., Llanos-Cuentas, A., Càceres, A., Le Ray, D., 1993. From pulsed field to field: contribution of molecular karyotyping to epidemiological studies on New World leishmaniasis. *Arch. Inst. Pasteur Tunis* 70, 505–515.
- Dujardin, J.C., Bañuls, A.L., Dujardin, J.P., Arevalo, J., Tibayrenc, M., Le Ray, D., 1998. Comparison of chromosome and isoenzyme polymorphism in geographical populations of *Leishmania (Viannia) peruviana*. *Parasitology* 117, 547–554.
- Dujardin, J.C., Bañuls, A.L., Llanos-Cuentas, A., Alvarez, E., DeDoncker, S., Jacquet, D., Le Ray, D., Arevalo, J., Tibayrenc, M., 1995. Putative *Leishmania* hybrids in the Eastern Andean valley of Huanuco, Peru. *Acta Trop.* 59, 293-307.
- Eck, R.V., Dayhoff, M.O., 1966. Atlas of protein sequence and structure. National Biomedical Research Foundation, Silver Springs, Maryland.
- Felsenstein, J., 1981. Evolutionary trees from DNA sequences: A maximum likelihood approach. *J Mol Evol* 1981; 17: 368-76.
- Felsenstein, J., 1985. Confidence limits on phylogenies: An approach using the bootstrap. *Evolution.* 39, 783-791.
- Fitch, W.M., 1971. Towards defining the course of evolution: Minimum change for a specific tree topology. *Systematic Zoology.* 20, 406-416.
- Folgueira, C., Requena, J.M., 2007. A postgenomic view of the heat shock proteins in kinetoplastids. *FEMS Microbiol Rev.* 31:359-377.
- Fraga, J., Montalvo, A. M., De Doncker, S., Dujardin, J. C., Van der Auwera, G., 2010. Phylogeny of *Leishmania* species based on the heat-shock protein 70 gene. *Infect. Genet. Evol.* 10, 238-45.

- Fu, X., Jiao, W., Chang, Z., 2006. Phylogenetic and biochemical studies reveal a potential evolutionary origin of small heat shock proteins of animals from bacterial class A. *J Mol Evol* 62, 257–266.
- García, A.L., Kindt, A., Quispe-Tintaya, K.W., Bermudez, H., Llanos, A., Arévalo, J., Bañuls, A.L., De Doncker, S., Le Ray, D., Dujardin, J.C., 2005. American tegumentary leishmaniasis: antigen-gene polymorphism, taxonomy and clinical pleomorphism. *Infect. Genet. Evol.* 5, 109–116.
- Grimaldi, G. Jr., David, J.R., McMahon-Pratt, D., 1987. Identification and distribution of New World *Leishmania* species characterized by serodeme analysis using monoclonal antibodies. *Am. J. Trop. Med. Hyg.* 36, 270-287. Erratum in: *Am. J. Trop. Med. Hyg.* 37, 414.
- Guan, W., Cao, D. P., Sun, K., Xu, J. N., Zhang, J. R., Chen, D. L., Chen, J. P., 2012. Phylogenic analysis of Chinese *Leishmania* isolates based on small subunit ribosomal RNA (SSU rRNA) and 7 spliced leader RNA (7SL RNA). *Acta Parasitol.* 57, 101-13.
- Hailu A, Gebre-Michael T, Berhe N, Balkew M, 2005. Leishmaniasis. In: Yemane Berhane DHM, Helmut Kloos, editors. *Epidemiology and Ecology of health and disease in Ethiopia*. Shama Books.
- Haslbeck, M., Franzmann, T., Weinfurtner, D., Buchner, J., 2005. Some like it hot: the structure and function of small heat shock proteins. *Nat. Struct. Mol. Biol.* 12, 842–846.
- Hide, M., Bras-Gonçalves, R., Bañuls, A.L., 2007. Specific cpb copies within the *Leishmania donovani* complex: evolutionary interpretations and potential clinical implications in humans. *Parasitology* 134, 379-389.
- Hughes, A.L., Piontkivska, H., 2003. Phylogeny of *Trypanosomatidae* and *Bodonidae* (Kinetoplastida) based on 18S rRNA: evidence for paraphyly of *Trypanosoma* and six other genera. *Mol. Biol. Evol.* 20, 644-652.
- Huson, D.H., 1998. SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics.* 14, 68-73.
- Huson, D.H., Bryant, D., 2006. Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* 23, 254-267.

- Ibrahim. M.E., Barker, D.C., 2001. The origin and evolution of the *Leishmania donovani* complex as inferred from a mitochondrial cytochrome oxidase II gene sequence. *Infect. Genet. Evol.* 1, 61-68.
- Kerr, S.F., 2000. Palaearctic origin of *Leishmania*. *Mem. Inst. Oswaldo Cruz* 95, 75-80.
- Kimura, M., 1980. A simple method for estimating evolutionary rate of base substitutions through comparative studies of nucleotide sequences. *J. Mol. Evol.* 16, 111-120.
- Kriehuber, T., Rattei, T., Weinmaier, T., Bepperling, A., Haslbeck, M., Buchner, J., 2010. Independent evolution of the core domain and its flanking sequences in small heat shock proteins. *Faseb J* 24, 3633-3642.
- Kuhls, K., Keilonat, L., Ochsenreither, S., Schaar, M., Schweynoch, C., Presber, W., Schönián, G. 2007. Multilocus microsatellite typing (MLMT) reveals genetically isolated populations between and within the main endemic regions of visceral leishmaniasis. *Microbes Infect.* 9, 334-43.
- Kuhls, K., Mauricio, I.L., Pratlong, F., Presber, W., Schönián, G., 2005. Analysis of ribosomal DNA internal transcribed spacer sequences of the *Leishmania donovani* complex. *Microbes. Infect.* 7, 1224-1234.
- Lainson, R., Shaw, J.J., 1987. Evolution, classification and geographical distribution. In: Peters, W., Killick- Kendrick, R. (Eds). *The Leishmaniasis in Biological and Medicine*. Academic Press, London, pp. 1- 120.
- Leblois, R., Kuhls, K., Francois, O., Schonian, G., Wirth, T., 2011. Guns, germs and dogs: On the origin of *Leishmania chagasi*. *Infect Genet Evol* 11, 1091-5.
- Lukeš, J., Mauricio, I.L., Schönián, G., Dujardin, J.C., Soteriadou, K., Debet, J.P., Kuhls, K., Quispe Tintaya, K.W., Jirků, M., Chocholová, E., Haralambous, C., Pratlong, F., Oborník, M., Horák, A., Ayala, F.J., Miles, M.A., 2007. Evolutionary and geographical history of the *Leishmania donovani* complex with a revision of current taxonomy. *Proc. Natl. Acad. Sci.* 29, 9375-9380.

- Luyo-Acero, G.E., Uezato, H., Oshiro, M., Takei, K., Kariya, K., Katakura, K., Gomez-Landires, E., Hashiguchi, Y., Nonaka, S., 2004. Sequence variation of the Cytochrome b gene of various human infecting members of the genus *Leishmania* and their phylogeny. *Parasitology* 128, 483-491.
- Mauricio, I. L., Yeo, M., Baghaei, M., Doto, D., Pratlong, F., Zemanova, E., Dedet, J. P., Lukes, J., Miles, M. A., 2006. Towards multilocus sequence typing of the *Leishmania donovani* complex: resolving genotypes and haplotypes for five polymorphic metabolic enzymes (ASAT, GPI, NH1, NH2, PGD). *Int. J. Parasitol.* 36, 757-69.
- Mauricio, I.L., Gaunt, M.W., Stothard, J.R., Miles, M.A., 2007. Glycoprotein 63 (*gp63*) genes show gene conversion and reveal the evolution of Old World *Leishmania*. *Int. J. Parasitol.* 37, 565-576.
- Mauricio, I.L., Stothard, J.R., Miles, M.A., 2000. The strange case of *Leishmania chagasi*. *Parasitol. Today*. 16, 188-189.
- Montalvo-Álvarez, A.M., Folgueira, C., Carrión, J., Monzote-Fidalgo, L., Cañavate, C., Requena, J.M., 2008. The *Leishmania* HSP20 as antigenic during natural infections, but, as DNA vaccine, it does not protect BALB/c mice against experimental *L. amazonensis* infection. *J. Biomed. Biotechnol.* 695432.
- Nei, M., Gojobori, T., 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol. Biol. Evol.* 3, 418-426.
- Odiwuor, S., Veland, N., Maes, I., Arévalo, J., Dujardin, J.C., Van der Auwera, G., 2012. Evolution of the *Leishmania braziliensis* species complex from amplified fragment length polymorphisms, and clinical implications. *Infect. Genet. Evol.* 12, 1994-2002.
- Orlando, T.C., Rubio, M.A., Sturm, N.R., Campbell, D.A., Floeter-Winter, L.M., 2002. Intergenic and external transcribed spacers of ribosomal RNA genes in lizard-infecting *Leishmania*: Molecular structure and phylogenetic relationship to mammal-infecting *Leishmania* in the subgenus *Leishmania* (*Leishmania*). *Mem. Inst. Oswaldo Cruz* 97, 695-701.

- Phillipe, H., 1998. Molecular phylogenetic in kinetoplasts. In: Coomb, G.H., Vickerman, K., Sleight, M.A., Warren, A. (Eds.), Evolutionary relationships among protozoa. Systematics association, London, pp. 195-212.
- Piarroux, R., Fontes, M., Perasso, R., Gambarelli, F., Joblet, C., Dumon, H., Quilici, M., 1995. Phylogenetic relationships between Old World *Leishmania* Strains revealed by analysis of a repetitive DNA sequence. Mol. Biochem. Parasitol. 73, 249-252.
- Pratt, D.M., David, J.R., 1981. Monoclonal antibodies that distinguish between New World species of *Leishmania*. Nature. 291, 581-583.
- Requena, J.M., Chicharro, C., Garcia, L., Parrado, R., Puerta, C.J., Cañavate, C., 2012. Sequence analysis of the 3'-untranslated region of HSP70 (type I) genes in the genus *Leishmania*: its usefulness as a molecular marker for species identification. Parasites & vectors 5, 87.
- Rijal, S., Uranw, S., Chappuis, F., Picado, A., Khanal, B., Paudel, I. S., Andersen, E. W., Meheus, F., Ostry, B., Das, M. L., Davies, C., Boelaert, M., 2010. Epidemiology of *Leishmania donovani* infection in high-transmission foci in Nepal. Trop. Med. Int. Health. 15, 21-8.
- Rioux, J.A., Lanotte, G., Serres, E., Pratlong, F., Bastien, P., Perieres, J., 1990. Taxonomy of *Leishmania*. Use of Isoenzymes. Suggestion for a new classification. Ann. Parasitol. Hum. Comp. 65, 111-125.
- Rzhetsky, A., Nei, M., 1992. A simple method for estimating and testing minimum evolution trees. Mol. Biol. Evol. 9, 945-967.
- Saitou, N., Nei, M., 1987. The neighbor-joining method: A new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406-425.
- Schönian, G., Mauricio, I., Cupolillo, E., 2010. Is it time to revise the nomenclature of *Leishmania*? Trends Parasitol. 26, 466-9.



- Spanakos, G., Piperaki, E.T, Menounos, P.G., Tegos, N., Flemetakis, A., Vakalis, N.C., 2008. Detection and species identification of Old World *Leishmania* in clinical samples using a PCR-based method. Trans. R. Soc. Trop. Med. Hyg. 102, 46-53.
- Sukmee, T., Sriripattanapipong, S., Mungthin, M., Worapong, J., Rangsin, R., Samung, Y., Kongkaew, W., Bumrungsana, K., Chanachai, K., Apiwathanasorn, C., Rujirojindakul, P., Wattanasri, S., Ungchusak, K., Leelayoova, S., 2008. A suspected new species of *Leishmania*, the causative agent of visceral leishmaniasis in a Thai patient. Int. J. Parasitol. 38, 617-622.
- Sun, Y., MacRae, T.H., 2005. Small heat shock proteins: molecular structure and chaperone function. Cell Mol Life Sci 62, 2460–2476.
- Talmi-Frank, D., Nasereddin, A., Schnur, L. F., Schonian, G., Toz, S. O., Jaffe, C. L., Baneth, G., 2010. Detection and identification of old world *Leishmania* by high resolution melt analysis. PLoS. Negl. Trop. Dis. 4, e581.
- Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M., Kumar, S., 2011. MEGA5: molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Mol. Biol. Evol. 28, 2731-9.
- Tsukayama, P., Lucas, C., Bacon, D.J., 2009. Typing of four genetic loci discriminates among closely related species of New World *Leishmania*. Int. J. Parasitol. 39, 355–362.
- Van der Auwera, G., Fraga, J., Montalvo, A. M., Dujardin, J. C., 2011. *Leishmania* taxonomy up for promotion? Trends Parasitol. 27, 49-50.
- Victoir, K., Bañuls, A.L., Arevalo, J., Llanos-Cuentas, A., Hamers, R., Noel, S., De Doncker, S., Le Ray, D., Tibayrenc, M., Dujardin, J.C., 1998. The gp63 gene locus, a target for genetic characterization of *Leishmania* belonging to subgenus *Viannia*. Parasitology 117, 1–13.

- Villinski, J.T., Klena, J.D., Abbassy, M., Hoel, D.F., Puplampu, N., Mechta, S., Boakye, D., Racznik, G., 2008. Evidence for a new species of *Leishmania* associated with focal disease outbreak in Ghana. *Diagnost. Microbiol. Infect. Dis.* 60, 323-327.
- Young, J.C., Agashe, V.R., Siegers, K., Hartl, F.U., 2004. Pathways of chaperone-mediated protein folding in the cytosol. *Nat. Rev. Mol. Cell. Biol.* 5, 781-91.
- Zelazny, A.M., Fedorko, D.P., Li, L., Neva, F.A., Fisher, S.H., 2005. Evaluation of 7SL RNA gene sequences for the identification of *Leishmania* spp. *Am. J. Trop. Med. Hyg.* 72, 415-420.

**Table 1.** Sequences used in this study.

Species <sup>a</sup>	Strain name	Country	Accession	
			Number sequences	hsp20 Number sequences <sup>b</sup>
<i>L. aethiopica</i>	MHOM/ET/89/GERE	Ethiopia	JX630109	FN395018
	NLB 107-08	Kenya	JX630110	FN395019
	MHOM/ET/83/169-83	Ethiopia	JX630111	FN395020
	MHOM/ET/72/L100	Ethiopia	JX630112	FN395021
<i>L. tropica</i>	MHOM/IN/79/DD7	India	JX630113	FN395025
	MHOM/KE/81/NLB_030B	Kenya	JX630114	FN395026
<i>L. major</i>	UQ_8	Sudan	JX630115	FN395022
	L137	Spain	JX630116	FN395023
	GITHURE	Kenya	JX630117	FN395024
	MHOM/IL/80/Friedlin	Israel	XM_842817 <sup>c</sup>	XM_001684512 <sup>c</sup>
<i>L. donovani</i>	MHOM/SD/--/1-S	Sudan	JX630118	FN395027
	MHOM/SD/82/GILANI	Sudan	JX630119	FN395029
	MHOM/NP/03/BPK282	Nepal	XM_003862626 <sup>bc</sup>	NA
<i>L. archibaldi</i>	MHOM/SD/97/LEM3463	Sudan	JX630120	FN395030

<i>L. infantum</i>	MHOM/MT/85/BUCK	Malta	JX630121	FN395031
	MHOM/ES/1988/LLM175	Spain	JX630122	NA
	MCAN/ES/98/LLM-877	Spain	XM_001466741 <sup>c</sup>	XM_001470287 <sup>c</sup>
<i>L. chagasi</i>	MCAN/BR/06/MAIKE	Brazil	JX630123	FN395031
	MHOM/BR/07/WC	Brazil	JX630124	FN395036
	MHOM/BR/07/ARL	Brazil	JX630125	FN395037
<i>L. mexicana</i>	MHOM/GT/01/U1103	Guatemala	XM_003872264 <sup>c</sup>	XM_003877072 <sup>c</sup>
<i>L. amazonensis</i>	MHOM/PE/02/LH2312	Peru	JX630126	FN395038
	MHOM/BR/73/M2269	Brazil	JX630127	EU599091
	IFLA/BR/67/PH8	Brazil	AM712297 <sup>b</sup>	NA
<i>L. garnhami</i>	MHOM/VE/76/JAP78	Venezuela	JX630129	EU599092
<i>L. braziliensis</i>	MHOM/BO/--/CUM180	Bolivia	JX630130	FN395039
	MHOM/PE/02/LH2182	Peru	JX630131	FN395040
	MHOM/BO/94/CUM29	Bolivia	JX630132	FN395041
	MHOM/PE/91/LC2177	Peru	JX630133	FN395042
	MHOM/BR/06/ICA	Brazil	JX630134	FN395043
	MHOM/BR/75/M2903	Brazil	JX630135	M87878

	MHOM/BR/75/M2904	Brazil	XM_001566535 <sup>c</sup>	XM_001566275 <sup>c</sup>
<i>L. peruviana</i>	MHOM/PE/03/LH2864	Peru	JX630136	FN395044
	MHOM/PE/03/LH2439	Peru	JX630137	FN395045
	MHOM/PE/90/LC468	Peru	JX630138	FN395046
	MHOM/PE/90/LCA08	Peru	JX630139	EU599089
<i>L. guyanensis</i>	MHOM/PE/02/LH2372	Peru	JX630140	FN395051
	MHOM/GF/85/LEM699	French Guiana	JX630141	FN395052
	MHOM/BR/07/029-ZAV	Brazil	JX630142	FN395053
<i>L. panamensis</i>	MCHO/PA/00/M4039	Panama	JX630143	FN395055
<i>L. naiffi</i>	MDAS/BR/78/M5210	Brazil	JX630144	FN395056
	MDAS/BR/79/M5533	Brazil	JX630145	FR872767
<i>L. lainsoni</i>	MHOM/BO/95/CUM71	Bolivia	JX630146	FN395047
	MHOM/PE/91/LC1581	Peru	JX630147	FN395048
	MHOM/PE/02/LH2344	Peru	JX630148	FN395049
<i>T. cruzi</i>	TINF/BR/63/CL Brener	Brazil	XM_816981 <sup>c</sup>	XM_812645 <sup>c</sup>
<i>T. b. brucei</i>	927/4 GUT at 10.1	Kenya	XM_838833 <sup>c</sup>	XM_824101 <sup>c</sup>

---

<sup>a</sup>Species as defined by MLEE typing. Whenever available, the full WHO code of the strain is provided.

<sup>b</sup>Sequences from strains retrieved from GenBank, the remaining ones were determined in this study.

<sup>c</sup>Accession numbers starting with XM are derived from a contemporary annotation of full genome sequences, and were also retrieved from GenBank database.

NA: not available in GenBank.

**Table 2.** Primers used for *hsp20* gene PCR and sequencing.

Primer	Primer Sequence (5'-3') <sup>a</sup>	Nucleotide position <sup>b</sup>
hsp20F	RGRGACTCGCTCAKCAACAGCG	21-43
hsp20R	CGTTGAAGSTGGCCTTGATTTTGCTG	365-391

<sup>a</sup> R=A or G; K=G or T; S=G or C

<sup>b</sup> The annealing position of the primers is given relative to GenBank accession entry AM712297 (*L. (L.) amazonensis* strain PH8).

**Table 3.** Single nucleotide polymorphisms and indel positions among *Leishmania* species.

Species <sup>b</sup>	Nucleotide at <i>hsp20</i> alignment positions <sup>a</sup>																										
	49	52	63	83	84	108	113	114	115	117	129	135	144	159	168	179	181	186	231	238	245	252	258	291	292	296	
	G	C	T	T	A	C	T	C	T	A	C	G	A	G	C	C	C	T	G	T	T	C	T	G	C	A	
<i>L. donovani</i> (4)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
<i>L. infantum</i> (6)	.	.	.	.	.	.	.	.	.	.	.	<u>T</u>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
<i>L. tropica</i> (2)	.	.	.	.	.	.	.	.	.	<u>G</u>	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	
<i>L. major</i> (4)	A	<u>T</u>	.	<u>C</u>	.	<u>T</u>	.	.	G	.	.	.	<u>G</u>	.	<u>T</u>	<u>T</u>	.	.	<u>A</u>	.	.	.	.	.	.	.	
<i>L. aethiopica</i> (4)	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	<u>T</u>	<u>G</u>	.	<u>C</u>	.	.	.	.	.	.	
<i>L. mexicana</i> (1)	.	.	.	.	.	.	.	.	G	.	.	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	
<i>L. amazonensis</i> (3)	.	.	.	.	.	.	.	T	G	<u>T</u>	T	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	
<i>L. garnhami</i> (1)	.	.	.	.	.	.	.	T	G	.	T	.	.	A	.	.	.	.	.	.	.	.	.	.	.	.	
<i>L. naiffi</i> (2)	.	.	.	.	.	.	.	.	.	C	.	<u>A</u>	.	.	.	.	.	.	<u>C</u>	G	.	.	.	.	.	.	
<i>L. lainsoni</i> (3)	.	.	<u>C</u>	.	G	.	-	-	-	C	.	.	.	.	.	.	.	.	.	.	.	.	<u>G</u>	.	.	<u>T</u>	
<i>L. guyanensis</i> (3)	.	.	.	.	G	.	.	.	.	C	.	.	.	<u>C</u>	.	.	.	.	.	G	<u>C</u>	.	.	.	<u>A</u>	.	
<i>L. panamensis</i> (1)	.	.	.	.	G	.	.	.	.	C	.	.	.	<u>C</u>	.	.	.	.	.	G	<u>C</u>	.	.	.	<u>A</u>	.	
<i>L. peruviana</i> (4)	A	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G	.	<u>T</u>	.	<u>A</u>	.	.	
<i>L. braziliensis</i> (7)	.	.	.	.	.	.	.	.	.	C	.	.	.	.	.	.	.	.	.	G	.	.	.	G	.	.	

<sup>a</sup> Relative to GenBank accession entry AM712297 (*L. (L.) amazonensis* strain PH8). Species-specific SNPs are underlined, conserved positions are indicated by a point, and the deleted positions in the *L. lainsoni hsp20* are indicated by a dash

<sup>b</sup> In brackets: number of sequences corresponding to each *Leishmania* species



**Fig. 1.** Neighbor-Joining phylogeny of the *hsp20* sequences listed in Table 1, based on an alignment of 319 nucleotides. Distances were estimated using the Kimura-2 parameter model. Bootstrap support of the branches was inferred from 2000 replicates, and is given in percentages at the internodes when exceeding 70%. The tree is drawn to the scale at the bottom, expressed as distance per nucleotide. Supported monophyletic species and (sub)genera are depicted at the right, irrespective of the species classification presented in Table 1. Old World clusters are indicated by a dot on the branch leading to the cluster, while a block is used for New World groups. The tree was rooted with the corresponding sequences from *T. cruzi* and *T. brucei*.

**Fig. 2.** Complete phylogenetic network of the *Leishmania* sequences of Table 1 and Fig. 1, excluding the *Trypanosoma* outgroup. It was constructed with the Neighbor-Net algorithm (Bryant and Moulton, 2004), excluding all conserved sites. The Kimura 2-parameter model for nucleotides was used, calculating the fraction of differences between each pair of sequences. Each of the three panels (A)–(C) is drawn to the scale indicated, expressed as distance per nucleotide (Fig. 1) in the *hsp20* alignment. (A) *L. (Viannia)* subgenus sequences separate into four species. Squares are used to indicate the position of subspecies. (B) Old World sequences of the *L. (Leishmania)* subgenus are separated into four species. Within *L.(L.) donovani*, the square shows the location of *L. donovani infantum* strains. (C) Phylogenetic relationships into the New World *L. (Leishmania)* subgenus. The species *L. garnhami*, *L. mexicana*, and *L. amazonensis*, as listed in Table 1, are denoted by gar, mex, and ama, respectively.

**Fig.3.** Neighbor-Joining phylogeny based on a combination of the *hsp20* and *hsp70* sequences listed in Table 1. Distances were estimated using the Kimura-2 parameter model. Bootstrap support of the branches was inferred from 2000 replicates, and is given in percentages at the internodes when exceeding 70%. The tree is drawn to the scale at the bottom, expressed as distance per nucleotide. Supported monophyletic species and (sub)genera are depicted at the right, irrespective of the species classification presented in Table 1. Old World clusters are indicated by a dot on the branch leading to the cluster, while a block is used for New World groups.

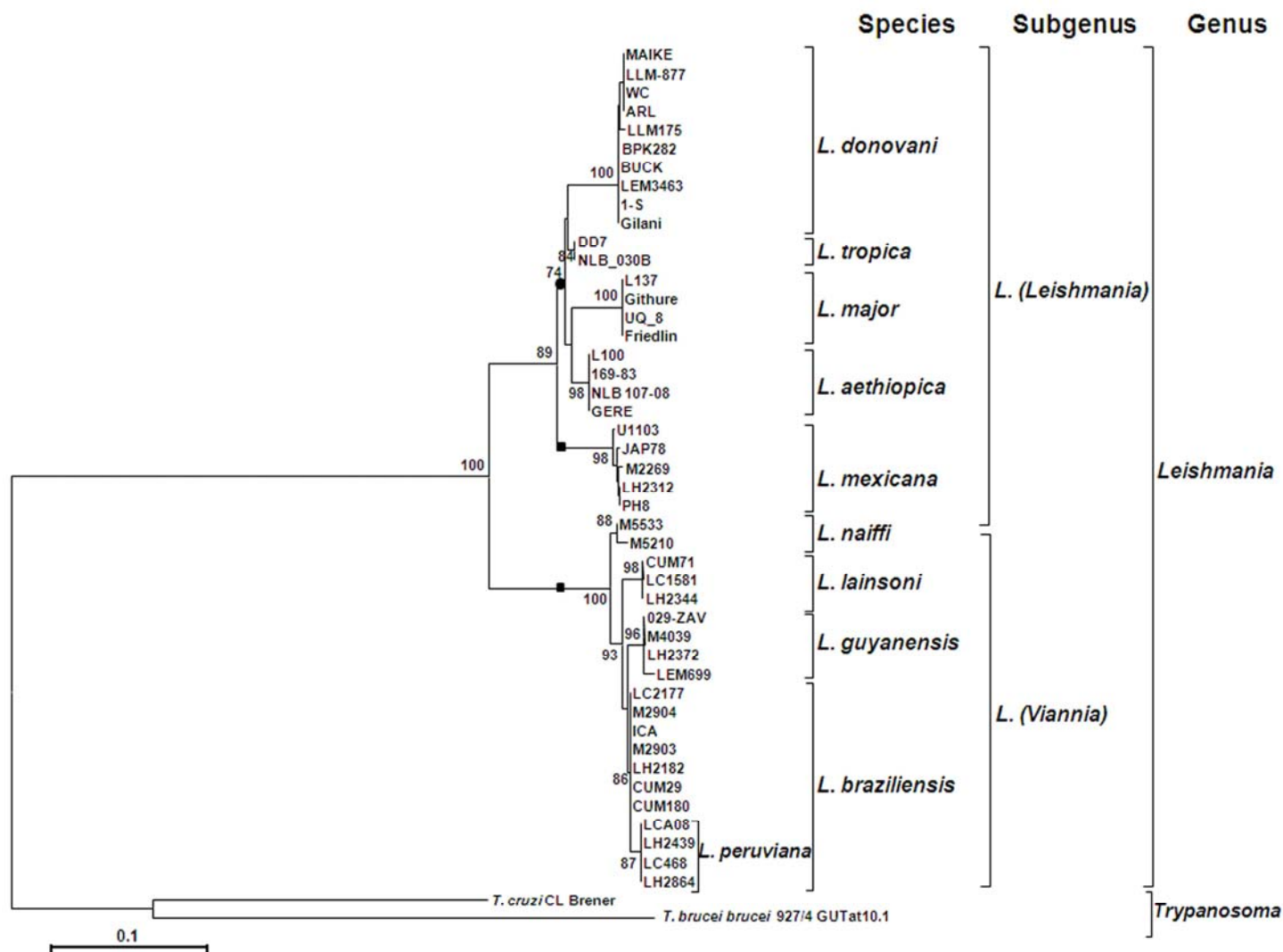


Fig 1

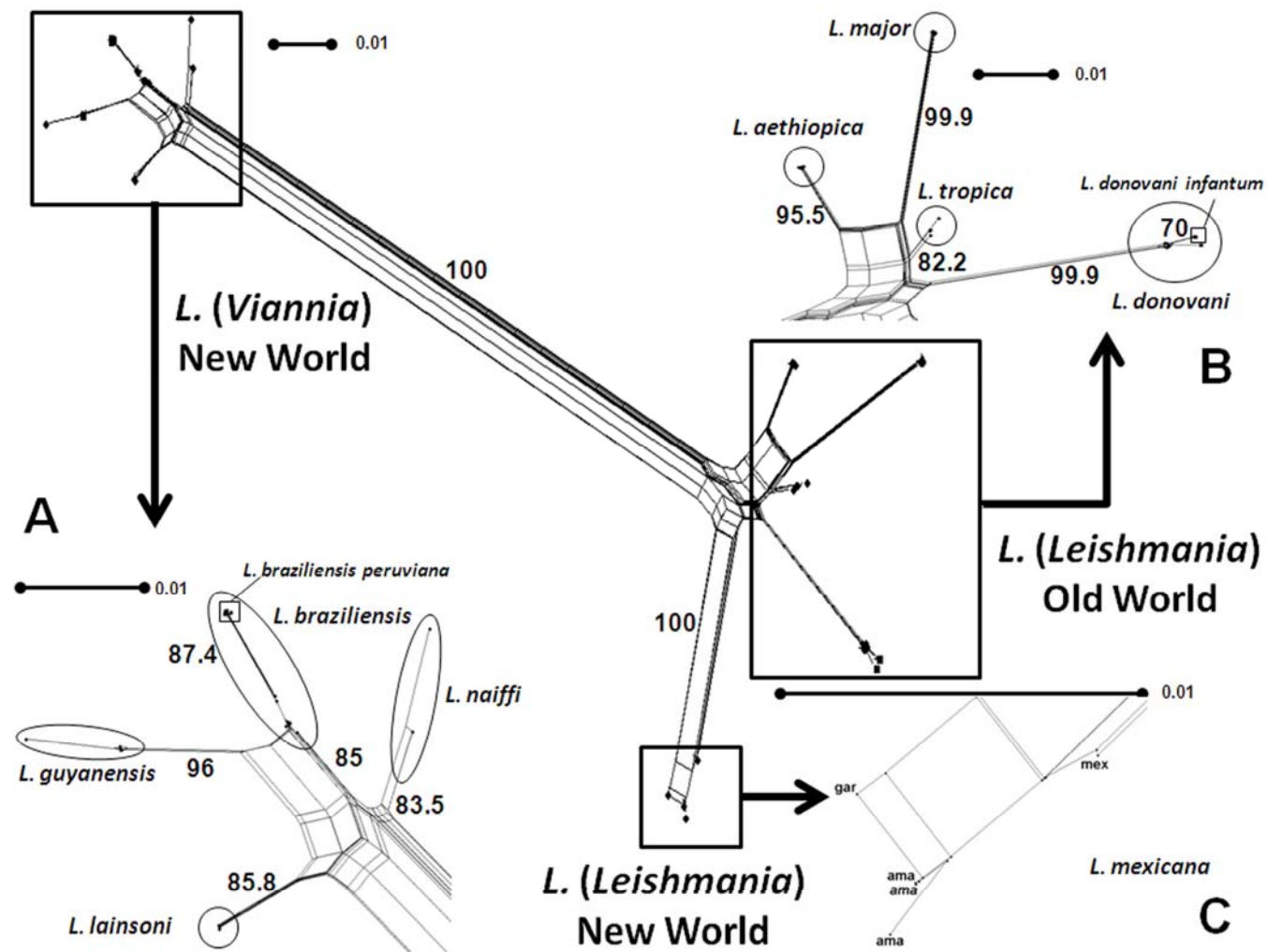


Fig. 2

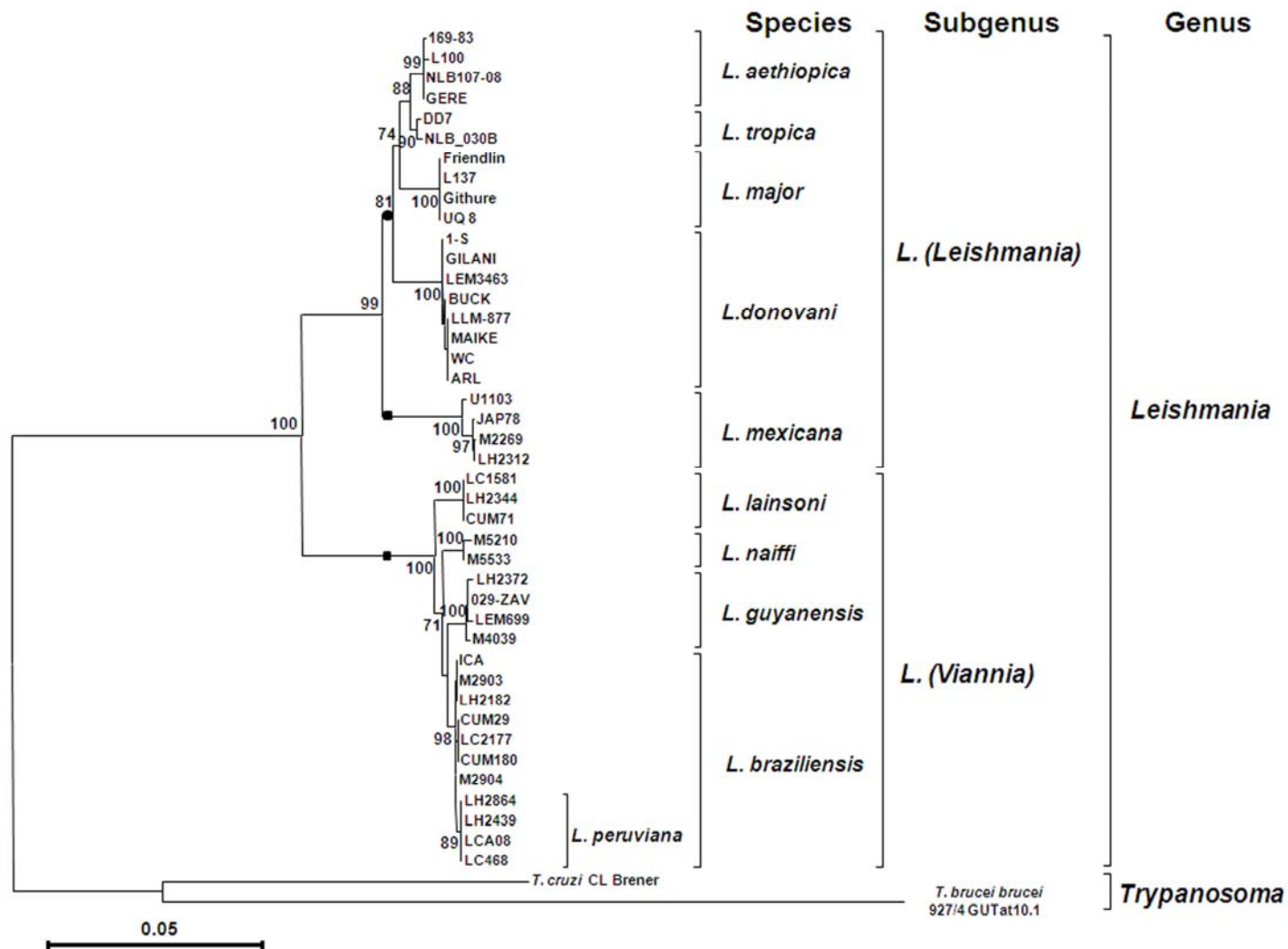


Fig. 3