

UNIVERSIDADE ESTADUAL PAULISTA

“JÚLIO DE MESQUITA FILHO”

FACULDADE DE MEDICINA VETERINÁRIA

CAMPUS DE ARAÇATUBA

**SELEÇÃO *IN SILICO* DE SEQUÊNCIAS DE DNA
ESPÉCIE-ESPECÍFICAS DE *Leishmania***

Fernanda Müller de Oliveira

Médica Veterinária

ARAÇATUBA – SP

2015

UNIVERSIDADE ESTADUAL PAULISTA
“JÚLIO DE MESQUITA FILHO”
FACULDADE DE MEDICINA VETERINÁRIA
CAMPUS DE ARAÇATUBA

SELEÇÃO *IN SILICO* DE SEQUÊNCIAS DE DNA
ESPÉCIE-ESPECÍFICAS DE *Leishmania*

Fernanda Müller de Oliveira
Orientadora: Profa. Adjunto Cáris Maroni Nunes
Co-orientadora: Dra. Adriana Santana do Carmo

Tese apresentada a Faculdade Medicina Veterinária – Unesp, Campus Araçatuba, como parte das exigências para obtenção do título de Doutora em Ciência Animal (Medicina Veterinária Preventiva e Produção Animal).

ARAÇATUBA – SP

2015

Dados Internacionais de Catalogação na Publicação (CIP)
Ficha catalográfica elaborada pela Biblioteca da FMVA / UNESP

O48s Oliveira, Fernanda Müller de
Seleção *in silico* de sequências de DNA espécie-específicas
de *Leishmania* / Fernanda Müller de Oliveira. – Araçatuba,
2015.
88f. il. + 1 CD-ROM

Tese (Doutorado) – Universidade Estadual Paulista,
Faculdade de Medicina Veterinária de Araçatuba.

Orientadora: Profa. Cárís Maroni Nunes

Coorientadora: Profa. Adriana Santana do Carmo

1. Leishmaniose - diagnóstico. 2. Genoma. 3.
Oligonucleotídeos. 4. DNA. I. T.

CDD 636.089444

CERTIFICADO DE APROVAÇÃO

TÍTULO: Seleção "in silico" de sequências de DNA espécie-específicas de Leishmania

AUTORA: FERNANDA MÜLLER DE OLIVEIRA

ORIENTADORA: Dra. CÂRIS MARONI NUNES

Aprovada como parte das exigências para obtenção do Título de DOUTORA em CIÊNCIA ANIMAL (MEDICINA VETERINÁRIA PREVENTIVA E PRODUÇÃO ANIMAL) pela Comissão Examinadora.


Dra. FLÁVIA LOMBARDI LOPES

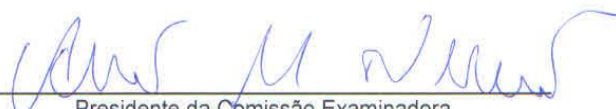

Dra. VÂNIA LÚCIA RIBEIRO DA MATTÁ


Dr. RODRIGO MARTINS SOARES


Dra. VALÉRIA MARÇAL FELIX DE LIMA


Dra. CÂRIS MARONI NUNES

DATA DA REALIZAÇÃO: 9 de dezembro de 2015.


Presidente da Comissão Examinadora
Dra. CÂRIS MARONI NUNES
- Orientadora -

DADOS CURRICULARES DO AUTOR

FERNANDA MÜLLER DE OLIVEIRA – nascida na cidade de Limeira– SP em 08 de abril de 1985. Possui graduação em Medicina Veterinária pela Faculdade de Medicina Veterinária de Araçatuba (UNESP), campus de Araçatuba, em dezembro de 2008. Possui MBA em gestão estratégica de negócios pela Faculdade Anhanguera de Limeira (2010), desenvolvendo como tema a gestão estratégica no SUS. Em 2012 obteve o título de Mestra em Ciência Animal, na área de Medicina Veterinária Preventiva e Saúde Animal pela Faculdade de Medicina Veterinária de Araçatuba (UNESP). No mesmo ano iniciou o Doutorado em Ciência Animal na mesma instituição, com período de Doutorado Sanduíche na University of Maryland, em College Park. Áreas de atuação biologia molecular, epidemiologia, genômica e parasitologia.

“A maior recompensa para o trabalho do homem não é o que ele ganha com isso, mas o que ele se torna com isso.”

John Ruskin

Aos meus pais, minha irmã e meu futuro marido.

AGRADECIMENTOS

À Universidade Estadual Paulista “Júlio de Mesquita Filho”, em especial ao Programa de Pós-Graduação em Ciência Animal.

À CAPES e CNPq, pelas bolsas de estudo durante o Doutorado, e pela oportunidade em realizar o Doutorado sanduíche na University of Maryland.

À Profa. Adj. Cáris Maroni Nunes e também orientadora, por ser uma professora e pesquisadora brilhante, competente, inspiradora tanto na vida pessoal como profissional, obrigada por todo suporte, orientação e amizade.

Ao Prof. Adj. José Fernando Garcia, pelo suporte, apoio e entusiasmo.

À Dra. Adriana Santana do Carmo pelo carinho, apoio e amizade.

Ao amigo e médico veterinário Pier Kenji Ito pelas ideias inspiradoras e por toda colaboração que teve neste trabalho.

Aos amigos: mestrande Flávia Athayde e ao doutorando Yuri Utsonomiya pelo apoio e colaboração empregados neste projeto.

Aos meus pais Marcos e Lindinalva, à minha irmã Flávia e ao meu noivo Kleber por todo amor, paciência e carinho proporcionados em toda essa caminhada, o apoio deles foi fundamental.

Aos amigos Priscilla Salvador, Hení Falcão, Luis Henrique Costa, Juliana Wagatsuma, Acácio Pacheco e Juliana Gamba por todos os momentos compartilhados.

Aos amigos e colegas do laboratório de Bioquímica e Biologia Molecular, Sarita Ferrari, Dandara Fonseca, Anirene Pereira, Silvana Paulan, Adam Utsunomiya, Tamiris Aguiar, Marco Milanesi, Rafaela Torrecilha, Ludmilla Zavarez, Kennya Martins e Beatriz Trigo pelo convívio diário e por todo apoio prestado.

À equipe da Deoxi Biotecnologia pela ajuda e serviços oferecidos para execução deste trabalho.

Ao Dr. Najib M. El-Sayed, da University of Maryland, pelo auxílio para o desenvolvimento do projeto e pelo tempo disponibilizado no laboratório de Cell Biology and Molecular Genetics, assim como toda sua equipe.

À todos que mesmo distantes me apoiaram essa jornada.

Sumário

LISTA DE ABREVIATURAS E TERMOS.....	11
CAPÍTULO 1 – CONSIDERAÇÕES GERAIS.....	16
<i>Leishmania</i> spp.	17
Ciclo biológico	19
Classificação das leishmanioses, etiologia das diversas manifestações clínicas	20
Leishmaniose visceral em cães e medidas de controle	21
Diferenciação de espécies de <i>Leishmania</i>	23
Genomas de <i>Leishmania</i> spp.	25
Objetivo	29
Objetivos específicos	29
REFERÊNCIAS	30
CAPÍTULO 2 – NOVA ABORDAGEM PARA IDENTIFICAÇÃO DE SEQUÊNCIAS DE DNA ESPECÍFICAS DAS DISTINTAS ESPÉCIES DE <i>Leishmania</i>.....	41
Resumo	41
Introdução	42
Métodos.....	44
Resultados	48
Discussão	58
Conclusão	63
Referências.....	64
APÊNDICE (S)	70

LISTA DE ABREVIATURAS E TERMOS

amplicon – termo utilizado para caracterizar a região amplificada, que é o produto da PCR

bootstrap – método estatístico de reamostragem com reposição, usada para construção de árvores filogenéticas

contigs – termo para montagem das primeiras sequências contíguas dos genomas

cross-dimer – quando ocorrem ligações entre os diferentes pares de oligonucleotídeos iniciadores numa reação de PCR

De Novo – montagem de genoma realizado sem base em genomas de referência, sem prévias montagens

dimer – regiões complementares iguais dos oligonucleotídeos iniciadores que se anelam entre si e não se ligam nas fitas de DNA

DPP – (do inglês) dual path platform/ teste rápido imunocromático

ELISA – (do inglês) enzyme-linked immunosorbent assay / teste sorológico

ePCR – (do inglês) eletronic polymerase chain reaction/ reação de alinhamento de sequências de DNA in silico

hairpin – região complementar ao oligonucleotídeo iniciador que se liga, formando um “grampo” no iniciador, prejudicando anelamento na fita de DNA

HRM – (do inglês) high resolution melting/ análise da temperatura de dissociação das fitas de DNA em alta resolução

Hsp70 - (do inglês) heat shock protein 70/ proteína de choque térmico de 70 kDa

IFAT – imunofluorescência indireta

ITS-1 - (do inglês) internal transcribed spacer 1 / espaço interno transcrito do DNA

L. aethiopica – *Leishmania aethiopica*

L. amazonensis – *Leishmania amazonensis*

L. arabica – *Leishmania arabica*

L. braziliensis – *Leishmania braziliensis*

L. donovani – *Leishmania donovani*

L. enrietti – *Leishmania enrietti*

L. gerbilli – *Leishmania gerbilli*

L. infantum – *Leishmania infantum*

L. major – *Leishmania major*

L. mexicana – *Leishmania mexicana*

L. panamensis – *Leishmania panamensis*

L. tarentolae – *Leishmania tarentolae*

L. tropica – *Leishmania tropica*

L. turanica – *Leishmania turanica*

LC – Leishmaniose cutânea

LMC – Leishmaniose mucocutânea

LV – Leishmaniose visceral

melting – dissociação de fitas duplas de DNA por temperatura

MLEE – (do inglês) multilocus enzyme electrophoresis / teste de múltiplos *loci* por meio da motilidade de enzimas na eletroforese

MLSA – (do inglês) multilocus sequence analysis/ tipagem de múltiplos loci de genes que são analisadas de forma conjunta

MLMT - (do inglês) multilocus microsatellite typing/ tipagem de múltiplos loci por meio de microssatélites do DNA

MLST - (do inglês) multilocus sequence typing/ tipagem de múltiplos loci que utiliza genes de manutenção

NGS – (do inglês) next generation sequence/ sequenciamento de DNA de nova geração

NJ – (do inglês) neighbor-joining – método de agrupamento entre vizinhos, de sequências por semelhanças, usado em filogenia para inferir similaridades entre as espécies e grupos

paired-end – biblioteca para sequenciamento em que as duas extremidades da fita de DNA são sequenciadas

PCR – (do inglês) polymerase chain reaction / reação em cadeia pela polimerase para amplificação de ácido desoxirribonucleico (DNA)

PCR-RFLP – (do inglês) polymerase chain reaction – restriction fragment polymorphisms/ PCR com restrição de fragmentos de polimorfismos baseadas em enzimas

qPCR – (do inglês) quantitative polymerase chain reaction/ real time PCR / técnica de reação em cadeia pela polimerase quantitativa ou em tempo real

scaffolds – termo para montagem de genomas, após a primeira fase de alinhamentos e montagens, formam-se grandes agrupamentos de nucleotídeos numa sequência contígua denominados *scaffolds*, é uma segunda etapa de montagem

script – escrita de termos para programação, linguagem de comandos

SNP – (do inglês) single nucleotide polymorphisms/ polimorfismos de nucleotídeos únicos

SSU – (do inglês) small subunit / subunidade menor do DNA ribossomal

supercontigs- agrupamentos de sequências contíguas de DNA realizadas na montagem, maiores que *contigs*

TriTrypDB – Banco de dados genômicos dos Tripanossomatídeos

SELEÇÃO *IN SILICO* DE SEQUÊNCIAS DE DNA ESPÉCIE- ESPECÍFICAS DE *Leishmania*

RESUMO - Leishmaniose, doença causada pelo protozoário *Leishmania spp.*, encontra-se em franca expansão no mundo e acomete milhões de pessoas anualmente, atingindo 98 países de regiões tropicais e subtropicais, representando um importante problema de Saúde Pública. As subespécies *Leishmania* e *Viannia* apresentam, no mínimo, 22 espécies, que são responsáveis pelas três manifestações clínicas mais comuns da doença para humanos. Rotineiramente, em áreas endêmicas, o diagnóstico laboratorial sorológico é empregado, porém, outras técnicas parasitológicas e/ou moleculares tem alta sensibilidade para identificação do gênero *Leishmania*. A publicação de genomas de *Leishmania* proporcionou um maior conhecimento de regiões e marcadores para identificação de espécies do protozoário e variações intra-específicas, todavia algumas espécies ainda tem sido subdiagnosticadas. Entretanto, a diferenciação das espécies de *Leishmania* em amostras clínicas ainda apresenta limitações. Assim, com o intuito de auxiliar na solução desta questão, buscamos sequências específicas para diferentes espécies de *Leishmania* depositadas em banco de dados genômicos de tripanossomatídeos e realizamos um ensaio *in silico* para verificação das nossas sequências. As informações geradas neste trabalho serão úteis numa plataforma diagnóstica para que sejam validadas com amostras clínicas.

Palavras-chave: Leishmaniose - diagnóstico, Oligonucleotídeos, Genoma, DNA.

IN SILICO SELECTION OF *Leishmania* SPECIE-SPECIFIC DNA SEQUENCES

ABSTRACT – Leishmaniasis, a disease caused by the protozoan *Leishmania* spp. is expanding in the world and affects millions of people annually, reaching 98 countries in tropical and subtropical regions, representing an important public health problem. Subspecies *Leishmania* and *Viannia*, represented by at least 22 species, are responsible for the three most common clinical manifestations of the disease in humans. Routinely, in endemic areas, serological diagnosis are employed. However, other parasitological and/or molecular techniques have high sensitivity to identify the genus *Leishmania*. The publication of *Leishmania* genomes have provided a better understanding of their composition and molecular markers to identify protozoan species and intra-specific variations, however some species are still underdiagnosed. However, the differentiation of *Leishmania* species in clinical samples still has limitations. Thus, in order to find the solution for this issue, we sought to design specific sequences for different species of *Leishmania* deposited in trypanosomatid genomic database, carrying out an in silico assay for verification of the selected sequences. The information generated in this study will constitute an interesting diagnostic platform, after proper validation with clinical samples.

Keywords: Leishmaniasis - diagnosis, *Primers*, Genome, DNA.

CAPÍTULO 1 – CONSIDERAÇÕES GERAIS

Leishmania spp. é o agente responsável pelas leishmanioses, umas das principais doenças tropicais negligenciadas, está distribuída mundialmente em 5 continentes, com grande número de espécies e manifestações clínicas, sendo que 90% dos casos de leishmaniose visceral, causados pelas espécies do complexo *Leishmania donovani*, ocorrem nos países Bangladesh, Brasil, Etiópia, Índia, Nepal e Sudão. Já a maior parte dos casos de leishmaniose tegumentar, representados principalmente pelas espécies *Leishmania major*, *Leishmania mexicana* e *Leishmania amazonensis* ocorrem no Afeganistão, Algéria, Brasil, Colômbia, Irã, Paquistão, Peru, Arábia Saudita, Síria e Tunísia. Outra importante manifestação clínica, que tem como agentes a *Leishmania braziliensis* e a *Leishmania peruviana*, é a forma mucutânea, ocorrendo principalmente no Brasil, Peru e Bolívia (WHO, 2013).

No Brasil, na década de 50 a leishmaniose visceral era encontrada apenas na região nordeste, mas a doença passou por um contínuo processo de urbanização e mudanças epidemiológicas que provocaram a chegada do vetor a novas regiões (WERNECK, 2010), além de um maior número de reservatórios capazes de albergar o parasita, e a doença se expandiu pelas regiões Norte, Centro-Oeste, Sul e Sudeste, sendo que o primeiro caso autóctone em humano do Estado de São Paulo ocorreu em 1.999. A *Leishmania* é capaz de infectar diversos hospedeiros como répteis, aves, mamíferos, como o homem, sendo que grande parte dos casos de leishmaniose ocorrem na população de baixa renda (ALVAR et al., 2006).

Os principais métodos diagnósticos utilizados são sorológicos, a observação do parasita e a detecção do DNA do protozoário por métodos moleculares. Embora a sensibilidade e especificidade destas técnicas sejam bastante variáveis, de modo geral, são capazes de identificar apenas o gênero

Leishmania, que somado ao diagnóstico epidemiológico, histórico clínico do paciente, observação de demais hospedeiros e reservatórios, bem como a distribuição geográfica, evidencia a ocorrência de algumas espécies. Contudo, o aprimoramento de técnicas diagnósticas permitiram maior acurácia e sensibilidade para identificação de algumas espécies do protozoário, diminuindo o subdiagnóstico. O advento da amplificação de DNA e a evolução das técnicas de sequenciamento resultaram em maior disponibilidade de genomas montados e anotados. Os primeiros genomas de tripanossomatídeos foram publicados em 2005 (EL-SAYED et al., 2005; IVENS et al., 2005) possibilitando o uso dessa informação para busca de sequências e novos marcadores que sejam espécie-específicos.

***Leishmania* spp.**

Leishmania spp. é um protozoário pertencente à família Trypanosomatidae com vinte e duas espécies isoladas de casos humanos, resultando em diferentes formas clínicas viscerais ou tegumentares (LAINSON; SHAW, 1987; WHO, 2013). A leishmaniose representa um importante problema de saúde pública sendo uma zoonose endêmica e em expansão no Brasil e no Mundo, ocorrendo principalmente nas Américas Central e Sul, parte da Europa e Ásia. Estimativas indicam que está presente em 98 países, com incidência anual de 1,3 milhões de casos da forma cutânea e 0,2-0,4 milhões de casos da forma visceral (DESJEUX, 2004; WHO, 2013). Ainda, as leishmanioses têm provocado cerca de 20.000 mortes por ano (WHO, 2013).

A classificação taxonômica dos protozoários do gênero *Leishmania* foi realizada de acordo com o comportamento das formas promastigotas no tubo digestivo do vetor, sendo divididos em dois subgêneros: *Viannia* para os parasitos que se desenvolvem na região do piloro e seção peripilária e *Leishmania* para os

com desenvolvimento na região suprapilária (LAISON; SHAW, 1987). Tal classificação embasava a antiga taxonomia, que aos poucos foi sendo atualizada com o conhecimento das sequências de DNA e as relações filogenéticas (Figura 1). A atual taxonomia separa algumas espécies de *Leishmania* não-patogênicas para humanos, encontradas em répteis, no subgênero *Sauroleishmania* (RAYMOND et al., 2012).

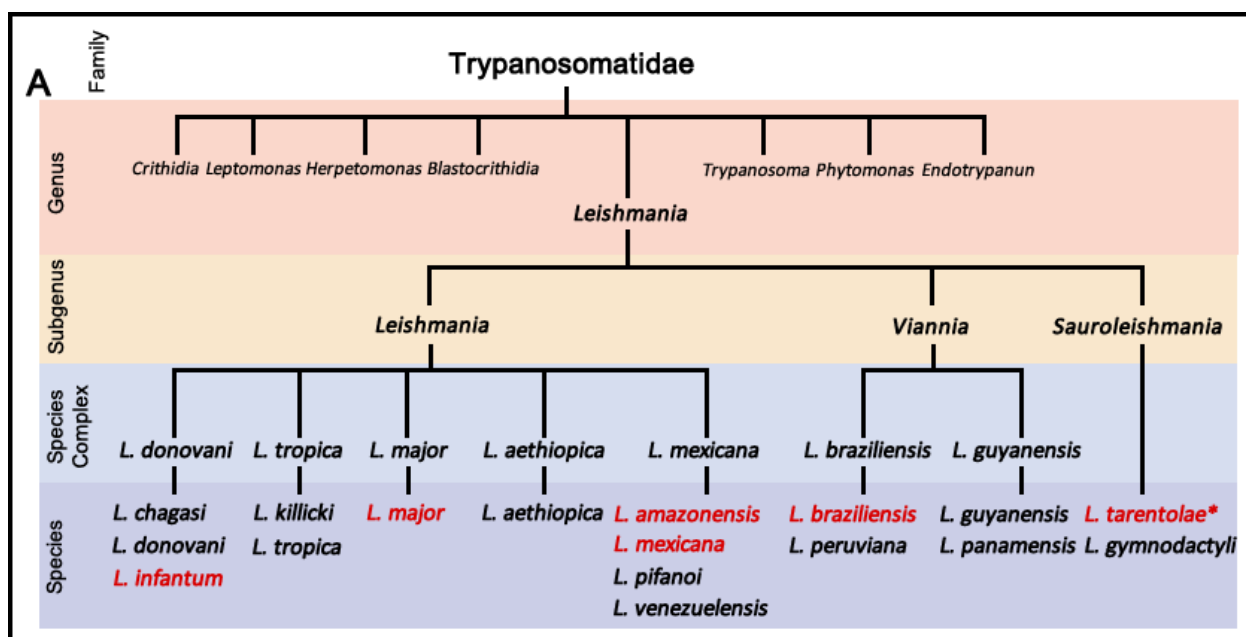


Figura 1-Classificação da *Leishmania* em gênero, subgênero, complexos e espécies. *L. infantum* a espécie mais importante da manifestação visceral da leishmaniose, *L. major*, *L. amazonensis* e *L. mexicana* das formas tegumentares da doença, *L. braziliensis* principal representante da forma mucocutânea e *L. tarentolae* (espécie não patogênica para humanos). Fonte: REAL et al.,(2013).

Ciclo biológico

Os principais vetores das Leishmânias são os flebotomíneos fêmeas, sendo os do gênero *Lutzomyia* associados à transmissão nas Américas. Durante o repasto sanguíneo, as formas promastigotas metacíclicas (formas infectantes) são depositadas na pele do hospedeiro e os parasitas são rapidamente capturados, principalmente por macrófagos, e em seguida diferenciados na forma amastigota, intracelular obrigatória, infectando o hospedeiro vertebrado (HANDMAN; BULLEN, 2002; KATO et al., 2010). O ciclo biológico do protozoário continua quando o hospedeiro infectado, contendo as formas amastigotas de *Leishmania* spp. é picado pelo flebotomíneo, que as ingere durante o repasto sanguíneo. No trato digestório do vetor, as formas amastigotas passam por um complexo processo e, após alguns dias (3-5 dias) transformam-se em promastigotas metacíclicas, sendo assim infectantes para outros mamíferos (BANETH, 2006; KILLICK-KENDRICK, 1990) (Figura 2).

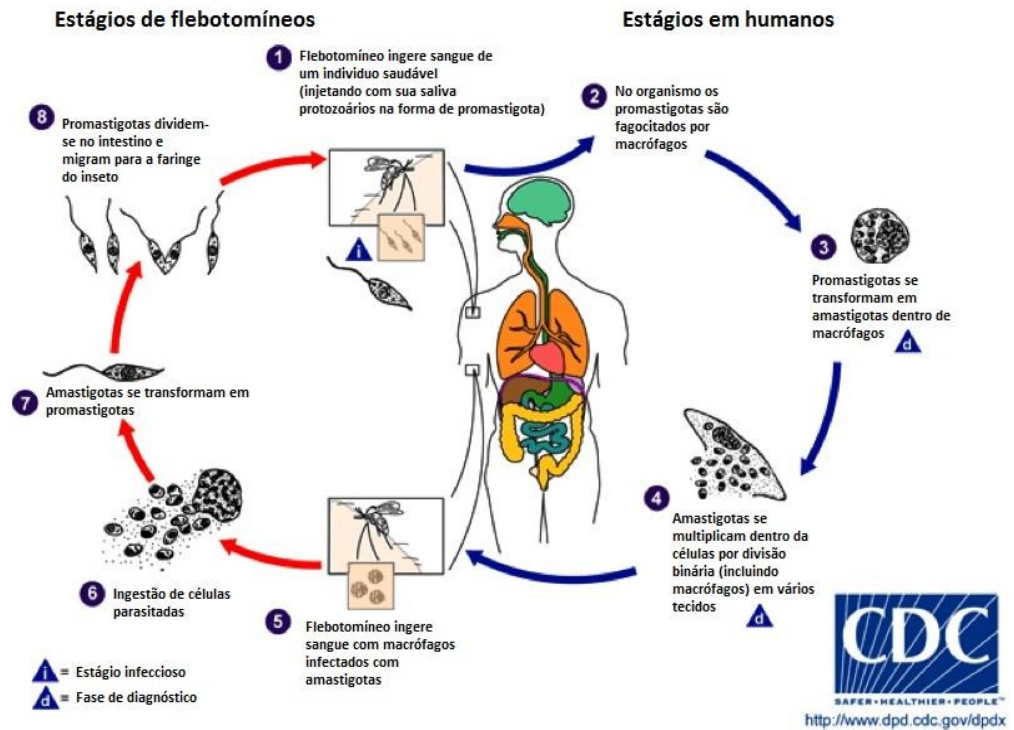


Figura 2- Ciclo biológico de *Leishmania*. Fonte: “Center for Disease Control”, EUA.

Classificação das leishmanioses, etiologia das diversas manifestações clínicas

As diferentes espécies de *Leishmania* causam amplo espectro de manifestações clínicas sendo classificadas como leishmanioses cutânea (LC), mucocutânea (LMC), e leishmaniose visceral (LV) (SCHÖNIAN et al., 2011), acometendo tanto crianças como adultos, além de várias espécies de mamíferos (DESJEUX et al., 2004). Ocorre variação clínica da doença de acordo com a região endêmica (MURRAY et al., 2005). Embora a variabilidade genética e a resposta imune do hospedeiro, em conjunto com a transmissão do flebotomíneo e

fatores ambientais sejam conhecidos por influenciar o resultado das infecções, o fator determinante é dependente da espécie de parasita infectante (PEACOCK et al., 2007).

A leishmaniose visceral (LV), a forma mais grave, tem como principal agente causador a *Leishmania infantum chagasi*, sendo o principal vetor a fêmea do flebotomíneo *Lutzomyia longipalpis*. Os cães (*Canis familiaris*) são considerados reservatórios domésticos da *Leishmania infantum chagasi* e caracterizados como principal fonte de infecção para o vetor (BRASIL, 2011; DEANE; DEANE, 1954; MARZOCHI et al., 1985; SHERLOCK, 1996).

Na América Latina há uma continua urbanização da leishmaniose visceral zoonótica, que no Brasil é causada principalmente por *Leishmania infantum chagasi* e leishmaniose cutânea, causada por *Leishmania braziliensis*, *Leishmania amazonensis* e *Leishmania guyanensis* (CAMARGO; LANGONI, 2006). Porém, há áreas que apresentam ocorrência concomitante de leishmaniose visceral e leishmaniose cutânea. A América Latina, a bacia do Mediterrâneo, central e o sudoeste da Ásia são áreas onde a LV é principalmente zoonótica (DESJEUX, 2001) assim, a redução da transmissão humana objetiva a redução do principal reservatório doméstico, o cão.

Leishmaniose visceral em cães e medidas de controle

A prevalência de leishmaniose visceral canina está associada à urbanização da doença e a presença de leishmaniose em humanos (HARHAY et al., 2011), o cão, principal reservatório, por ser fonte de infecção para o vetor e consequentemente aumentar o risco da infecção humana, favorece a manutenção do ciclo do parasito. O cão pode ou não apresentar sintomatologia clínica da doença, podendo ser classificados como doentes ou subclínicos sadios de acordo

com a presença ou ausência de sinais clínicos (ALVAR et al., 2004; SOLANO- GALEGO et al., 2009).

No Brasil, a porcentagem de cães infectados vivendo em áreas endêmicas para LV canina varia entre 1 a 67% (BERNADINA et al., 1997; COUTINHO et al., 1985; FRANÇA-SILVA et al., 2003).

Como medida de controle para LV, o Programa de Vigilância e Controle da Leishmaniose preconiza o diagnóstico e tratamento precoces do paciente humano, o controle vetorial por meio do manejo ambiental e uso de inseticidas e a identificação e eutanásia do reservatório canino, além da educação em saúde. Assim, nos municípios onde há registro da doença, inquéritos sorológicos são rotineiramente realizados na população canina (BRASIL, 2011). Os métodos sorológicos convencionais para diagnóstico de LV como ELISA e imunofluorescência indireta (IFAT), de modo geral, tem boa acurácia diagnóstica, porém são mais laboriosos. Já os testes imunocromatográficos, baseados na detecção de antígenos Rk26/rK39 são de rápida execução e foram desenvolvidos para áreas endêmicas (LAURENTI et al., 2014). Atualmente, os inquéritos sorológicos são feitos com DPP[®], que é um teste imunocromatográfico (teste rápido, do inglês *Dual Path Platform*), seguido do teste confirmatório ELISA (do inglês, *enzyme-linked immnosorbent assay*).

Uma das limitações dos testes sorológicos é que nenhum desses métodos permite a diferenciação das espécies de *Leishmania* envolvidas. Portanto, para a identificação da espécie faz-se necessário o isolamento do protozoário e/ou o diagnóstico molecular em laboratórios de referência, particularmente em áreas onde há ocorrência de múltiplas espécies (WHO, 2013). A identificação entre as espécies é crucial para execução de tratamentos de pacientes humanos, bem como em estudos epidemiológicos para se determinar as medidas de controle (SCHÖNIAN et al., 2003).

Diferenciação de espécies de *Leishmania*

O “padrão ouro” para identificação de cepas de *Leishmania* spp. em estudos epidemiológicos e taxonômicos é o MLEE (*multilocus enzyme electrophoresis*), baseado na motilidade de isoenzimas, que é também a base da atual taxonomia (ALONSO et al., 2010; RIOUX et al., 1990; SCHÖNIAN et al., 2010). Porém, há uma limitação no uso desta técnica, uma vez que se baseia em enzimas altamente conservadas e dependente de um padrão de mobilidade eletroforético e, se houver mudanças na composição de aminoácidos, estes podem não alterar a mobilidade das enzimas, levando a uma interpretação errônea, além da técnica ser laboriosa e só poder ser usada em amostras de culturas de *Leishmania* (KUHLS et al., 2007; OCHSERENTHEIR et al., 2006).

O diagnóstico molecular baseado na reação em cadeia pela polimerase (PCR) surgiu na década de 80 e tem sido amplamente utilizado para a identificação de gênero, podendo também apresentar sucesso quando da identificação de espécie de *Leishmania*. O kinetoplasto, composto por moléculas circulares de DNA situado na base do flagelo dos protozoários, apresenta regiões altamente conservadas, com grande número de cópias (10.000 cópias/parasita), possibilita um ensaio de alta sensibilidade, pode ser usado em amostras clínicas diversas (ALONSO et al., 2010; PRINA et al., 2007; RODGERS et al., 1990), sendo portanto, o principal alvo para diagnósticos moleculares (ARANSAY et al., 2000; CORTES et al., 2004; LAUCHAUD et al., 2002; MIMORI et al., 1998; RODGERS et al., 1990).

Sequências de DNA nuclear têm sido utilizadas como alvos para a distinção de espécies e cepas de *Leishmania* (SCHÖNIAN et al., 2011), como por exemplo sequências do gene da glicose-6-fosfato desidrogenase (CASTILHO et al., 2003, 2008) e da SSU (do inglês - *small subunit* do DNA) ribossomal (SAVANI et al., 2010; ULIANA et al., 1994).

Alvos de DNA ribossomal apresentam de 10 a 20 cópias e são muito utilizados nas técnicas de PCR-RFLP (do inglês *Restriction Length Polymorphisms*) ou seguidos de sequenciamento do fragmento, e podem ser utilizados em amostras clínicas para tipificação de espécies (BERZUNZA-CRUZ et al., 2015; KUHLS et al., 2005; SCHÖNIAN et al., 2003, 2011). Já a região espaçadora interna transcrita 1 (ITS-1) é muito utilizada para a identificação de variações intra-específicas por apresentar elevada conservação (EL TAI et al., 2000), assim como o gene que codifica a proteína de choque térmico de 70 kDa (Hsp70) que tem sido muito utilizado para estudos em filogenia e diversidade intra-específica (MONTALVO et al., 2010; REQUENA et al., 2012; SCHÖNIAN et al., 2010).

Seguido à técnica de qPCR, HRM (do inglês, *High resolution melting analysis*) baseada na análise molecular que mede a taxa de dissociação de fita dupla de DNA em fita simples conforme o aumento da temperatura da reação, que foi previamente amplificada pela PCR, têm sido muito utilizada, tal medida de dissociação é característica, ou seja, específica para cada produto de DNA (*amplicon*), sendo, portanto utilizado para identificação de espécies de *Leishmania* (TALMI-FRANK et al., 2010).

Diante das limitações da MLEE e, com o desenvolvimento da biologia molecular, surgiram novas técnicas como MLMT (do inglês *multilocus microsatellite typing*) que, por meio de marcadores de DNA microssatélites, é capaz de distinguir entre cepas e espécies de *Leishmania* (KUHLS et al., 2007; OCHSENREITHER et al., 2006; ODDONE et al., 2009; REALE et al., 2010). Outra variação das técnicas de *multilocus* é o MLST (*multilocus sequence typing*) em que enzimas padronizadas do MLEE são utilizadas e, em seguida, são desenhados oligonucleotídeos iniciadores para PCR e sequenciamento, seguidos de análise dos alinhamentos (MAURICIO et al., 2006; TSUKAYAMA et al., 2009; ZEMANOVÁ et al., 2007).

A combinação de diferentes alvos genômicos apresenta uma aparente vantagem, que é o uso da informação de vários *loci* para garantir a evidência de

certas espécies (VAN DER AUWERA; DUJARDIN, 2015). Utilizadas em estudos epidemiológicos para caracterização de cepas, para estudos da estrutura populacional e da evolução das espécies, a abordagem MLSA (*multilocus sequense analysis*) com genes de manutenção tem sido muito frequente (BOITÉ et al., 2012; MARLOW et al. 2014).

A possibilidade de se conhecer várias sequências moleculares forneceu informações adicionais para análises filogenéticas (HUGHES; PIOTINKVISKA, 2003). Com o passar dos anos, muitas informações de genes codificantes e não-codificantes usados como marcadores de DNA (BOITÉ et al., 2012; MONTALVO et al., 2010) permitiram o avanço do entendimento das relações filogenéticas assim como, a diferenciação das espécies (YANG et al., 2013). O uso de sequências do citocromo b por Yang et al. (2013) e do hsp70 por Montalvo et al. (2010) e Silva et al. (2010), auxiliam no estabelecimento de relações filogenéticas capazes de distinguir *Leishmania* de diferentes regiões.

Genomas de *Leishmania* spp.

Paralelamente aos avanços das técnicas de PCR, tecnologias de sequenciamento foram sendo desenvolvidas e aprimoradas. Atualmente, os sequenciadores de nova geração (*Next Generation Sequencing* – NGS – Figura 3) proporcionam um acelerado e eficiente conhecimento de genomas de diversos organismos, como o de alguns protozoários como *Leishmania* spp. e *Trypanosoma* spp. As pesquisas realizadas no campo genômico ajudaram a entender mais sobre a susceptibilidade, a expressão da doença, da relação hospedeiro-parasita, a imunidade anti-*Leishmania*, avançando ainda mais com o conhecimentos das sequências e/ou genomas de alguns tripanosomatídeos (EL-SAYED et al., 2005; IVENS et al., 2005) facilitando o entendimento e permitindo a

busca de novos alvos para diagnóstico, drogas e mesmo o desenvolvimento de vacinas (MURRAY et al., 2005).

O primeiro genoma nuclear publicado, e que serviu como base para os demais protozoários do gênero, foi o da *Leishmania major*, a principal representante da forma tegumentar da doença (IVENS et al., 2005).

Alta conservação e sintonia entre as espécies *Leishmania infantum*, *Leishmania braziliensis* e *Leishmania major* foi observada (PEACOCK et al., 2007), apresentando aproximadamente 200 diferenças no conteúdo de genes ou pseudogenes, incluindo 78 genes que são restritos a espécies individuais. Já o perfil de expressão gênica entre estas espécies, revelou que a regulação diferencial espécie-específico de genes conservados foi comum o que pode ter impacto sobre a sobrevivência do parasita no hospedeiro, pois os resultados suportam a hipótese de que as leishmanias são pré-adaptadas para submeter-se a sobrevivência intracelular, tendo pouca modulação na expressão do gene, ao nível de RNA (DEPLEDGE et al., 2009).

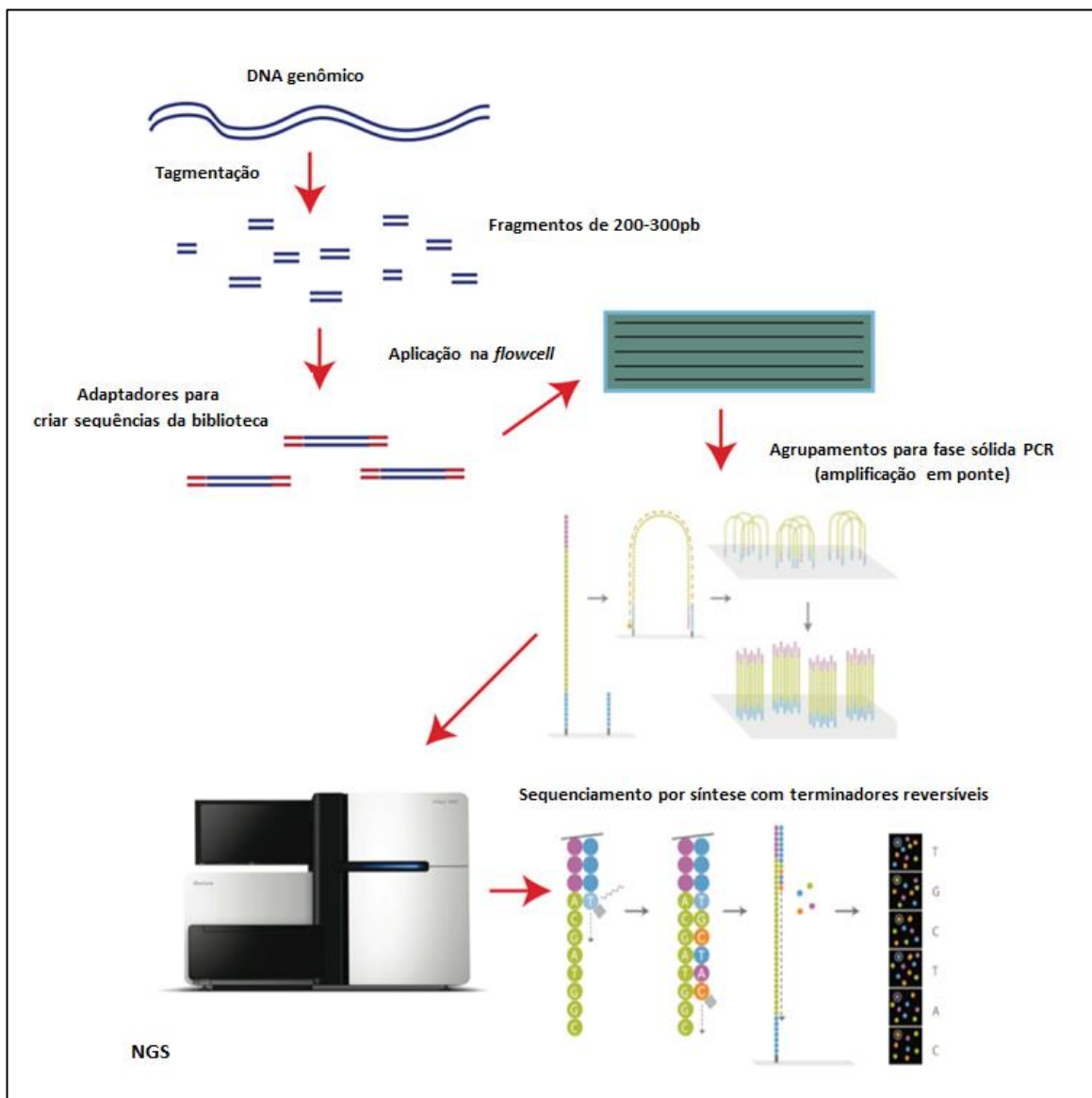


Figura 3-Esquema sequenciamento NGS. DNA genômico é tagmentado em fragmentos de 200-300pb, são inseridos adaptadores para formação da biblioteca, em seguida, aplicados numa lâmina de uma linha (*flowcell*) que vai para o sequenciador, são formados agrupamentos (*clusters*) numa amplificação em ponte e o sequenciamento é realizado por síntese complementar a fita, uma fluorescência para cada nucleotídeo é emitida, assim são reconhecidas as sequências para posterior análise, esquema adaptado de Illumina®.

Em 2011 os genomas de *Leishmania donovani* (DOWNING et al., 2011) e *Leishmania mexicana* (ROGERS et al., 2011) foram publicados, demonstrando a presença de cromossomos supranumerários e variado número de cópias entre as espécies e cepas de *Leishmania*. Em 2012 o primeiro genoma de espécie não patogênica, *Leishmania tarentolae*, encontrada em répteis, foi comparado às demais espécies, tendo sido observado um maior número de genes adaptados ao estágio promastigota do protozoário no vetor, com ausência de 250 genes em comparação à outras leishmânias, principalmente aqueles relacionados ao estágio amastigota do protozoário, que é a fase de patogenicidade relacionado ao hospedeiro (RAYMOND, et al., 2012).

O primeiro genoma montado no Brasil foi de *Leishmania amazonensis*, publicado por Real et al. (2013), espécie que foi inicialmente identificada no Brasil e é a principal representante da forma clínica mucocutânea (Tabela 1).

Tabela 1-Resumo das informações dos principais genomas publicados de espécies de *Leishmania* spp. Adaptado de Real et al. (2013).

	<i>L. (L.) major</i>	<i>L. (V.) braziliensis</i>	<i>L.(V.) infantum</i>	<i>L. (L.) mexicana</i>	<i>L. (V.) donovani</i>	<i>L. (S.) tarentolae</i>	<i>L. (L.) amazonensis</i>
Contigs	36	1041	37	35	2154	N/A	3199
Genoma tamanho (Mb)	32,8	33,7	32,1	32,1	32,4	30,4	29,6
Cromossomos	36	35	36	34	36	36	34
Número de genes preditos	8298	8153	8216	8007	8252	8201	8100
Gene densidade (genes/Mb)	260	228	252	256	254	270	273
C+G conteúdo (%)	59,3	57,8	59,3	60,5	>60	57,2	58,5
Referências	Ivens et al. (2005)	Peacock et al. (2007)	Peacock et al. (2007)	Rogers et al. (2011)	Downing et al. (2011)	Raymond et al. (2012)	Real et al. (2013)

Diante das informações dos genomas destes parasitas já disponíveis e da possibilidade de avançar mais no conhecimento genético de espécies de *Leishmania*, trabalhamos com a possibilidade de sequenciar novas espécies do protozoário e utilizar os genomas já publicados para buscar sequências espécie-específicas capazes de diferenciar as espécies de *Leishmania*, *in silico*.

Objetivo

Selecionar sequências-alvo de DNA de *Leishmania* específicas para diferenciação de espécies, baseado em informação genômica e validação *in silico* em culturas sequenciadas.

Objetivos específicos

- Selecionar sequências de DNA e oligonucleotídeos iniciadores espécie-específico para *Leishmania*
- Verificar a especificidade das sequências selecionadas utilizando a ferramenta de PCR eletrônico (ePCR)
- Validar, *in silico*, o painel de pares de oligonucleotídeos iniciadores selecionados, em amostras de referência de espécies de *Leishmania* sequenciadas
- Analisar filogeneticamente seis espécies de *Leishmania* com genomas completos, utilizando-se proteínas ortólogas de cópia única

REFERÊNCIAS

ALONSO, D.P.; COSTA, D.L.; MENDONÇA, I.L. de; COSTA, C.H.; RIBOLLA, P.E. Heterogeneity of *Leishmania infantum chagasi* kinetoplast DNA in Teresina (Brazil). **The American Journal and Tropical Medicine and Hygiene**, v.82, p. 819-821, 2010.

ALVAR, J.; CAÑAVATE, C.; MOLINA, R.; MORENO, J.; NIETO, J. Canine leishmaniasis. **Advances in parasitology**, v. 57, p. 1-88, 2004.

ALVAR, J.; YACTAYO, S.; BERN, C. Leishmaniasis and poverty. **Trends in Parasitology**, v. 22, n. 12, p. 552–557, 2006.

ARANSAY, A.M.; SCOULICA, E.; TSELENTIS, Y. Detection and identification of *Leishmania* DNA within naturally infected sand flies by seminested PCR on minicircle kinetoplastic DNA. **Applied Environment Microbiology**, v.66, p.1933–1938, 2000.

BANETH, G. Leishmaniasis. In: GREENE, C.E. **Infectious diseases of the dog and the cat**. 3.ed. Philadelphia: Elsevier, p.685-698, 2006.

BERNADINA, W.E.; DE LUNA, R.; OLIVA, G., CIARAMELLA; P. An immunodiffusion assay for the detection of canine leishmaniasis due to infection with *L. infantum*. **Veterinary Parasitology**, v.73, p.207–213, 1997.

BERZUNZA-CRUZ, M.; et al. *Leishmania (L.) mexicana* infected bats in Mexico: novel potential reservoirs. **PLoS Neglected Tropical Disease**, v.9; p.e0003438, 2015.

BOITÉ, M.C.; MAURICIO, I.L.; MILES, M.A.; CUPOLILLO, E. New insights on taxonomy, phylogeny and population genetics of *Leishmania (Viannia)* parasites based on multilocus sequence analysis. **PLoS Neglected Tropical Diseases**, v. 6; p.e1888, 2012.

BRASIL, Ministério da Saúde. Secretaria De Vigilância Em Saúde. Subcoordenação de Zoonoses Vetoriais e Raiva. **Nota Técnica N°/2011 – UVR/CGDT/DEVEP/SVS/MS**. Esclarecimentos sobre o diagnóstico sorológico da leishmaniose visceral canina utilizado na rede pública de saúde. Brasília: Ministério da Saúde, 2011.

CAMARGO, L.B.; LANGONI, H. Impact of leishmaniasis in public healthy. **The Journal of Venomous Animals and Toxins Including Tropical Diseases**, v.12, p.527-548, 2006.

CASTILHO, T.M.; CAMARGO, L.M.A., McMAHON-PRATT, D.; SHAW, J.J.; FLOETER-WINTER, L.M. A real-time polymerase chain reaction assay for the identification and quantification of American *Leishmania* species on the basis of glucose-6-phosphate-dehydrogenase. **American Journal of Tropical Medicine and Hygiene**, v.78, p.122-132, 2008.

CASTILHO, T.M.; SHAW, J.J.; FLOETER-WINTER, L.. New PCR assay using glucose-6-phosphate dehydrogenase for identification of *Leishmania* species. **Journal of Clinical Microbiology**, v.41, p.540–546, 2003.

CORTES, S.; ROLÃO, N.; RAMADA, J.; CAMPINO, L. PCR as a rapid and sensitive tool in the diagnosis of human and canine leishmaniasis using *Leishmania donovani* s.l. - Specific kinetoplastid oligonucleotídeos iniciadores.

Transactions Royal Society of Tropical Medicine and Hygiene, v.98, p.12–17, 2004.

COUTINHO, S.G., NUNES, M.P., MARZOCHI, M.C.A., TRAMONTANO, N. A survey for American cutaneous and visceral leishmaniasis among 1342 dogs from areas in Rio de Janeiro (Brazil) where the human diseases occur. **Memórias do Instituto Oswaldo Cruz**, v.80, p.17–22, 1985.

DEANE, L. M.; DEANE, M. P. Dogs naturally infected by *Leishmania donovani* in Ceara. **Hospital**, v. 45, p.703-707, 1954.

DEPLEDGE, D.P.; KRYSTAL, J.E.; ALASDAIR, C.I.; AZIZ, N.; MAROOF, A.; KAYE, P.M.; SMITH, D.F. Comparative expression profiling of *Leishmania*: modulation in gene expression between species and in diferente host genetic backgrounds. **PLOS Neglected Tropical Diseases**, v.3, p. e476, 2009.

DESJEUX P. Leishmaniasis; current situation and new perspectives. **Comparative Immunology Microbiology & Infectious Diseases**, v.27, p.305-318, 2004.

DESJEUX P. The increase in risk factors for leishmaniasis worldwide. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v.95, p.239–243, 2001.

DOWNING, T.; et al. Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanisms of drug resistance. **Genome Research**, v.21, p.2143-2156, 2011.

EL TAI, N.O.; OSMAN, O.F.; EL FARI, M.; PRESBER, W.; SHONIAN, G. Genetic heterogeneity of ribosomal internal transcribed spacer (its) in clinical samples of

Leishmania donovani spotted on filter paper a revealed by single-strand conformation polymorphisms (sscp) and sequencing. **Transactions of the Royal Society of Tropical Medicine and Hygiene**. v.94, p.575-579, 2000.

EL-SAYED, N.M.; et al. The genome sequence of *Trypanosoma cruzi*, etiologic agent of Chagas disease. **Science**, v.309, p. 409–415, 2005.

FRANÇA-SILVA, J.C., COSTA, R.T., SIQUEIRA, A.M., MACHADO-COELHO, G.L.L., COSTA, C.A., MAYRINK, W., VIEIRA, E.P., COSTA, J.S., GENARO, O., NASCIMENTO, E. Epidemiology of canine visceral leishmaniasis in the endemic area of Montes Claros municipally, Minas Gerais State, Brazil. **Veterinary Parasitology**, v.111, p.161–173, 2003.

HANDMAN, E.; BULLEN, D.V.R. Interaction of *Leishmania* with the host macrophage. **Trends Parasitology**, v.18, p.332-334, 2002.

HARHAY, M. O.; OLLIARO, P. L.; COSTA, D. L.; COSTA, C. H. N. Urban parasitology: visceral leishmaniasis in Brazil. **Trends Parasitology**, v.27, p. 403-409, 2011.

HUGHES, A.L.; PIONTKIVSKA, H. Molecular phylogenetics of Trypanosomatidae: constrasting results from 18S rRNA and protein phylogenies. **Kinetoplastid Biology and Diseases**, v.1, p.1-10, 2003.

IVENS, A.C.; et al. The genome of the kinetoplastid parasite, *Leishmania major*. **Science**, v. 309, p. 436–442, 2005.

KATO, H.; GOMEZ, E.A.; CÁCERES, A.G.; UEZATO, H.; MIMORI, T.; HASHIGUCHI, Y. Molecular epidemiology for vector research on leishmaniasis. **International Journal of Environmental Research and Public Health**, v.7, p.814-826, 2010.

KILLICK-KENDRICK, R. The life cycle of *Leishmania* in the sandfly with special reference to the form infective to the vertebrate host. **Annual Parasitology Human Complete**, v.65, p.37-42, 1990.

KUHLS, K.; et al. Multilocus microsatellite typing (MLMT) reveals genetically isolated populations between and within the main endemic regions of visceral leishmaniasis. **Microbes and Infection**, v.9, p.334–343, 2007.

KUHLS, K.; MAURICIO, I.L., PRATLONG, F., PRESBER, W.; SCHÖNIAN, G. Analysis of ribosomal DNA internal transcribed spacer sequences of the *Leishmania donovani* complex. **Microbes and Infection**, v.7, p.1224–1234, 2005.

LACHAUD, L.; MARCHERGUI-HAMMAMI, S.; CHABBERT, E.; DEREURE, J.; DEDET, J.P.; BASTIEN, P. Comparison of six PCR methods using peripheral blood for detection of canine visceral leishmaniasis. **Journal of Clinical Microbiology**, v.40, 210–215, 2002.

LAISON, R.; SHAW, J. J. Evolution, classification and geographical distribution. In: PETERS, W.; KILLICK-DENDRICK, R. (Ed.) **The Leishmaniasis in Biology and Medicine**. London: Academic Press, 1987, p. 1-20.

LAURENTI, M.D.; DE SANTANA, L.M.V; TOMOKANE, T.Y.; DE LUCCA, H.R.L.; ASCHAR, M.; SOUZA, C.S.F.; SILVA, R.M.; MARCONDES,M.; DA MATTA, V.L.R. Comparative evaluation of the DPP(®) CVL rapid test for canine

serodiagnosis in area of visceral leishmaniasis. **Veterinary Parasitology**, v. 205, p.444-450, 2014.

MARLOW, M.A.; BOITÉ, M.C.; FERREIRA, G.E.M.; STEINDEL, M.; CUPOLILLO E. Multilocus sequence analysis for *Leishmania braziliensis* outbreak investigation. **Plos Neglected Tropical Diseases**, v.8, p. e2695, 2014.

MARZOCHI, M. C. A.; COUTINHO, S. G.; SOUZA, W. J. S.; TOLEDO, L. M.; GRIMALDI JR, G.; MOMEN, H.; PACHECO, R. S.; SABROZA, P. C.; SOUZA, M. A.; RANGEL JR, F. B.; TRAMONTANO, N. C. Canine visceral leishmaniasis in Rio de Janeiro, Brazil. Clinical, parasitological, therapeutical and epidemiological findings (1977-1983). **Memórias Instituto Oswaldo Cruz**, v.80, p.349-357, 1985.

MAURICIO, I.L.; et al. Towards multilocus sequence typing of the *Leishmania donovani* complex: Resolving genotypes and haplotypes for five polymorphic metabolic enzymes (ASAT, GPI, NH1, NH2, PGD). **International Journal of Parasitology**, v.36, p.757-769, 2006.

MIMORI, T.; et al. Rapid identification of *Leishmania* species from formalin-fixed biopsy samples by polymorphism-specific polymerase chain reaction. **Gene**, v.210, p.179–186, 1998.

MONTALVO, A.M.; et al. Heat-shock protein 70 PCR-RFLP: a universal simple tool for *Leishmania* species discrimination in the New and Old World. **Parasitology**, v.137, p.1159-1168, 2010.

MURRAY, H.W.; BERMAN, J.D.; DAVIES, C.R.; SARAVIA. N.G. Advances in leishmaniasis. **Lancet**, v. 366, p.1561-1577, 2005.

OCHSENREITHER, S.; KUHLS, K.; SCHAAR, M.; PRESBER, W.; SCHÖNIAN, G. Multilocus microsatellite typing as a new tool for discrimination of *Leishmania infantum* MON-1 strains. **Journal of Clinical Microbiology**, v.44: 495–503, 2006.

ODDONE, R.; et al. Development of a multilocus microsatellite typing approach for discriminating strains of *Leishmania (Viannia)* species. **Journal of Clinical Microbiology**, v.47, p.2818–2825, 2009.

PEACOCK, C.S., SEEGER, K., HARRIS, D., MURPHY, L., RUIZ, J.C., QUAIL, M.A., PETERS, N., ADLEM, E., TIVEY, A., ASLETT, M., KERHORNOU, A., IVENS, A., FRASER, A., RAJANDREAM, M.A., CARVER, T., NORBERTCZAK, H., CHILLINGWORTH, T., HANCE, Z., JAGELS, K., MOULE, S., ORMOND, D., RUTTER, S., SQUARES, R., WHITEHEAD, S., RABBINOWITSCH, E., ARROWSMITH, C., WHITE, B., THURSTON, S., BRINGAUD, F., BALDAUF, S.L., FAULCONBRIDGE, A., JEFFARES, D., DEPLEDGE, D.P., OYOLA, S.O., HILLEY, J.D., BRITO, L.O., TOSI, L.R., BARRELL, B., CRUZ, A.K., MOTTRAM, J.C., SMITH, D.F., BERRIMAN, M. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. **Nature Genetics**, v.39, p.839–847, 2007.

PRINA E.; ROUX E.; MATTEI D.; MILON G. Leishmania DNA is rapidly degraded following parasite death: an analysis by microscopy and real-time PCR. **Microbes and Infection**; v.9, p. 1307–1315, 2007.

RAYMOND, F.; BOISVERT, S.; ROY, G.; RITT, J.F.; LÉGARÉ, D.; ISNARD, A.; STANKE, M.; OLIVIER, M.; TREMBLAY, M.J.; PAPADOPOULOU, B.; OUELLETTE, M.; CORBEIL, J. Genome sequencing of the lizard parasite *Leishmania tarentolae* reveals loss of genes associated to the intracellular stage of human pathogenic species. **Nucleic Acids Research**, v.40, p.1131-1147, 2012.

REAL, F.; VIDAL, R.O.; CARAZZOLLE, M.F.; MONDEGO, J.M.C.; COSTA, G.G.L.; HERAI, R.H.; WÜRTELE, M.; DE CARVALHO, L.M.; RMONA E FERREIRA, R.; MORTARA, R.A.; BARBIERI, C.L.; MIECZKOWSKI, P.; DA SILVEIRA, J.F.; DA SILVA BRIONES, M.R.; PEREIRA, G.A.G.; BAHIA, D. The genome sequence of *Leishmania (Leishmania) amazonensis*: Functional annotation and extended analysis of gene models. **DNA Research.**, v.20, p.567-581, 2013.

REALE, S.; LUPO, T.; MIGLIAZZO, A.; DI MAURO, C.; CIPRI, V.; CALDERONE, S.; MANNA, L.; VITALE, F. Multilocus microsatellite polymorphism analysis to characterize *Leishmania infantum* strains isolated in Sicily. **Transbound Emerging Diseases**, v. 57, p. 37–41, 2010.

REQUENA, J.M.; CHICHARRO, C.; GARCÍA, L.; PARRADO, R.; PUERTA, C.J.; CAÑAVATE, C. 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**, v.5, p.87, 2012.

RIOUX, J.A.; LANOTTE, G.; SERRES, E.; PRATLONG, F.; BASTIEN, P.; PERIERES, J. Taxonomy of *Leishmania*. Use of isoenzymes. Suggestions for a new classification. **Annales de Parasitologie Humaine et Comparee**, v. 65, p.111-125, 1990.

RODGERS, M.R.; POPPER, S.J.; WIRTH, D.F. Amplification of kinetoplast DNA as a tool in the detection and diagnosis of *Leishmania*. **Experimental Parasitology**, v.71, p.267-275, 1990.

ROGERS, M.B.; HILLEY, J.D.; DICKENS, N.J.; WILKES, J.; BATES, P. A.; DEPLEDGE, D.P.; HARRIS, D.; HER, Y.; HERZYK, P.; IMAMURA, H.O.; THOMAS, D.; SANDERS, M.; SEEGER, K.; DUJARDIN, J.C.; BERRIMAN, M.;

SMITH, D.F.; HERTZ-FOWLER, C.; MOTTRAM, J.C. Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. **Genome Research**, v.21, p. 2129-2142, 2011.

SAVANI, E.S.M.M.; ALMEIDA, M.F.; CAMARGO, M.C.G.O.; DAURIA, S.R.N; SILVA, M.M.S.; OLIVEIRA, M.L.; SACRAMENTO, D.R.V. Detection of *Leishmania (Leishmania) amazonensis* and *Leishmania (Leishmania) infantum chagasi* in Brazilian bats. **Veterinary Parasitology**, v.168, p.5-10, 2010.

SCHÖNIAN, G.; KUHLS, K.; MAURICIO, I.L. Molecular approaches for a better understanding of the epidemiology and population genetics of *Leishmania*. **Parasitology**, v.138, p.405-425, 2011.

SCHÖNIAN, G.; MAURICIO, I.; CUPOLILLO, E. Is it time to revise the nomenclature of *Leishmania*?. **Trends in Parasitology**, v.26, p. 466-469, 2010.

SCHÖNIAN, G.; NASEREDDIN, A.; DINSE, N.; SHWEYNOCH, C.; SCHALLIG, H.D.F.H.; PRESBER, W.; JAFFE, C.L. PCR diagnosis and characterization of *Leishmania* in local and imported clinical samples. **Diagnostic Microbiology and Infectious Disease**, v.47, p.349-358, 2003.

SHERLOCK, I. A. Ecological interactions of visceral leishmaniasis in the state of Bahia, Brazil. **Memórias do Instituto Oswaldo Cruz**, v.91, p. 671-683, 1996.

SILVA, L.A.; DE SOUSA, C. D. S.; DA GRAÇA, G.C.; PORROZZI, R.; CUPOLILLO, E. Sequence analysis and PCR-RFLP profiling of the hsp70 gene as a valuable tool for identifying *Leishmania* species associated with human leishmaniasis in Brazil. **Infection, Genetics and Evolution**, v. 10, p. 77-83, 2010.

SOLANO-GALLEGO, L.; KOUTINAS, A.; MIRÓ, G.; CARDOSO, L.; PENNISI, M. G., FERRER, L.; BOURDEAU, P.; OLIVA, G.; BANETH, G. Directions for the diagnosis, clinical staging, treatment and prevention of canine leishmaniosis. **Veterinary Parasitology**, v. 165, p. 1-18, 2009.

TALMI-FRANK, D. et al. Detection and identification of old world leishmania by high resolution melt analysis. **PLoS Neglected Tropical Diseases**, v. 4, p. 4–8, 2010.

TSUKAYAMA, P.; LUCAS, C.; BACON, D.J. Typing of four genetic loci discriminates among closely related species of New World *Leishmania*. **International Journal for Parasitology**, v.39, p. 355-362, 2009.

ULIANA, S.R.; NELSON, K.; BEVERLY, S.M.; CAMARGO, E.P.; FLOETER-WINTER, L.M. Discrimination amongst *Leishmania* by polymerase chain reaction and hybridization with small subunit ribosomal DNA derived oligonucleotides. **Journal of Eukaryotic Microbiology**, v.41, p.324–330, 1994.

VAN DER AUWERA, G.; DUJARDIN, J-C. Species typing in dermal leishmaniasis. **Clinical Microbiology Reviews**, v. 28, p.265–294, 2015.

WERNECK, L.G. Geographic spread of visceral leishmaniasis in Brazil. **Cadernos de Saúde Pública**, v.26, p. 644-645, 2010.

WORLD HEALTH ORGANIZATION WHO. **Sustaining the drive to overcome the global impact of neglected tropical diseases: second WHO Report neglected tropical disease**. Geneva: WHO, 2013.

YANG, B.B; CHEN, D.L; CHEN, J.P; LIAO, L.; HU, X.S; XU, J.N. Analysis of kinetoplast cytochrome b gene of 16 *Leishmania* isolates from different foci of China: different species of *Leishmania* in China and their phylogenetic inference. **Parasites & Vectors**, v.6, p.32, 2013.

ZEMANOVÁ, E.; JIRKŮ, M.; MAURICIO, I.L.; HORÁK, A.; MILES, M.A.; LUKEŠ, J. The *Leishmania donovani* complex: genotypes of five metabolic enzymes (ICD, ME, MPI, G6PDH, and FH), new targets for multilocus sequence typing. **International Journal Parasitology**, v.37, p. 149–160, 2007.

CAPÍTULO 2 – NOVA ABORDAGEM PARA IDENTIFICAÇÃO DE SEQUÊNCIAS DE DNA ESPECÍFICAS DAS DISTINTAS ESPÉCIES DE *Leishmania*

Nova abordagem para identificação de sequências de DNA específicas das distintas espécies de *Leishmania**

Resumo

O gênero *Leishmania* apresenta pelo menos 22 espécies conhecidas por causarem diferentes patologias que afetam o ser humano, representando um importante problema de Saúde Pública. Mesmo com as técnicas de amplificação de DNA, diferenciar as espécies de *Leishmania* em amostras clínicas ainda apresenta limitações. Por outro lado, diversos tripanosomatídeos tiveram seus genomas disponibilizados em bancos de dados públicos possibilitando explorar novas sequências-alvo para melhoria do diagnóstico. Objetivou-se selecionar sequências de DNA espécie-específicas, num ensaio *in silico*, para desenvolver um painel para identificação das distintas espécies de *Leishmania*. Utilizando-se a base de dados disponível no banco genômico TriTrypDB, além de oligonucleotídeos iniciadores/ou sequências de DNA de *Leishmania* spp. comumente utilizados para diagnóstico molecular, foram selecionados 2.145 pares de oligonucleotídeos iniciadores que amplificam fragmentos de DNA do protozoário. Para validação dos mesmos por meio de ePCR, amostras de referência de seis espécies de *Leishmania* foram sequenciadas e seus genomas montados. Alvos específicos para 14 espécies de *Leishmania* (*L. amazonensis*, *L. aethiopica*, *L. arabica*, *L. braziliensis*, *L. donovani*, *L. enrietti*, *L. gerbilli*, *L. infantum*, *L. major*, *L. mexicana*, *L. panamensis*, *L. tarentolae*, *L. tropica* e *L. turanica*) foram identificados, totalizando 960 sequências de DNA candidatas potenciais à identificação de espécies de *Leishmania*. A seleção e validação *in*

silico de um painel de *oligonucleotídeos iniciadores* foi capaz de discriminar DNA de espécies de *Leishmania*, apresentando potencial para diagnóstico em amostras clínicas.

Palavras-chave: leishmaniose, ePCR, diagnóstico, genoma, sequências de DNA.

***Segundo normas da Revista PlosONE**

Introdução

O gênero *Leishmania* pertence à família Trypanosomatidae e compreende um grande número de espécies e subespécies, sendo pelo menos 22 delas infecciosas para humanos causando um amplo espectro de manifestações clínicas (DESJEUX, 2004; HAJJARAN et al., 2011). As formas clínicas das leishmanioses são classificadas em cutânea (LC), mucocutânea (LMC) e visceral (LV), acometendo tanto crianças quanto adultos, além de várias outras espécies de mamíferos (DESJEUX, 2004). A espécie *Leishmania major* que ocorre no Mediterrâneo, África, Oriente Médio e Índia é responsável, em grande parte, pela leishmaniose cutânea; a *L. braziliensis* aparece como principal agente da leishmaniose mucocutânea no Brasil e na América Latina, no Mediterrâneo e Brasil, a *Leishmania infantum chagasi* causa a leishmaniose visceral em humanos e cães; já a leishmaniose dérmica pós-calazar é ocasionada principalmente pela *Leishmania donovani*, sendo endêmica na Índia e Sri-Lanka (VAN DER AUWERA; DUJARDIN, 2015).

A identificação das espécies de *Leishmania* é importante para o tratamento adequado de pacientes humanos, bem como para estudos epidemiológicos (SCHÖNIAN et al., 2011), sendo geralmente realizada por meio do isolamento do protozoário e/ou diagnóstico molecular em laboratórios de referência (WHO, 2012). O “padrão ouro” para identificação de cepas de *Leishmania* spp. é o MLEE

(do inglês, *multilocus enzyme electrophoresis*) baseado na motilidade de isoenzimas, que serve como base para taxonomia vigente (ALONSO et al., 2010; MAURICIO et al., 2010; RIOUX et al., 1990). Entretanto, mudanças na composição de aminoácidos alteram a mobilidade das enzimas, o que pode levar a falhas na interpretação. Ademais, essa técnica é laboriosa e só pode ser realizada em amostras de cultivo de *Leishmania* (KUHLS et al., 2007; OCHSENREITHER et al., 2006).

Como alternativa, o diagnóstico molecular baseado na amplificação de DNA tem sido amplamente utilizado para a identificação de gênero, podendo também apresentar sucesso na identificação de espécie (CASTILHO et al., 2003; CASTILHO et al., 2008; EL TAI et al., 1996; KUHLS et al., 2005; MONTALVO et al., 2010; SAVANI et al., 2010; SCHÖNIAN et al., 2003; SCHÖNIAN et al., 2010; SCHÖNIAN et al., 2011; ULIANA et al., 1994). A combinação da amplificação de fragmentos de DNA, por PCR, com o sequenciamento tem permitido a diferenciação de espécies e cepas de *Leishmania* (BOITÉ et al., 2012; KUHLS et al., 2007; MARLOW et al., 2014; MAURICIO et al., 2006; OCHSENREITHER et al., 2006; ODDONE et al., 2009; REALE et al., 2010; TSUKAYAMA et al., 2009; ZEMANOVÁ et al., 2007). O volume de informação referente a genes envolvidos em diversos aspectos da leishmaniose tem aumentado e permitido o aprimoramento de meios diagnósticos (IVENS et al., 1998), o entendimento das relações filogenéticas, assim como a diferenciação das espécies (YANG et al., 2013).

A primeira espécie de *Leishmania* a ter o seu genoma nuclear sequenciado foi *Leishmania major*, tendo sido identificados 8.298 genes codificantes (IVENS et al., 2005). Posteriormente, os genomas de *L. braziliensis* e *L. infantum* foram montados e análises comparativas revelaram elevado grau de sintonia entre as espécies, com alta conservação na distribuição dos genes e de regiões repetitivas (PEACOCK et al., 2007). A adição do genoma de *L. mexicana* em estudos comparativos revelou significativa variação no número de cromossomos entre as

espécies (ROGERS et al., 2011) e, graças ao sequenciamento de *L. donovani*, novas análises demonstraram elevado número de polimorfismos de nucleotídeos únicos (SNP) (DOWNING et al., 2011). O tamanho dos genomas das espécies de *Leishmania* varia entre 29 e 34Mb (DOWNING et al., 2011; PEACOCK et al., 2007; ROGERS et al, 2011; S. DE TOLEDO, et al. 2010).

Semelhança significativa entre grupos de proteínas de superfície relacionadas à evolução de *Leishmania* e à interações com a célula hospedeira, pôde ser descrita com o sequenciamento da espécie *L. amazonensis* e análises comparativas com *L. mexicana* (REAL et al., 2013).

Como a identificação de genes que respondam questões biológicas não requer necessariamente genomas inteiros (GROVER et al., 2012), trabalhamos com a hipótese de que ao explorar todas as sequências de DNA dos tripanosomatídeos depositadas em um banco de dados (TriTrypDB), a seleção de um painel de sequências alvo de DNA específicas para diferenciação de espécies de *Leishmania* spp. seria possível.

Métodos

Seleção de sequências de DNA de *Leishmania* spp. para desenho pares de oligonucleotídeos iniciadores

Sequências-alvo de DNA para identificação de espécies de *Leishmania*, sequências de genes com função hipotética e com variação entre as espécies *Leishmania major*, *Leishmania infantum*, *Leishmania braziliensis* (PEACOCK et al., 2007) (Tabela suplementar 1) foram utilizadas para o desenho de pares de oligonucleotídeos iniciadores, totalizando nesta estratégia 541 pares.

Selecionamos também regiões genômicas descritas na base de dados TriTrypDB que correspondessem aos seguintes critérios: genes codificadores de proteínas conhecidos ou conhecidos por projeção de outras espécies e que não

apresentassem ortologia com as espécies de *Leishmania* (*Leishmania* específicos), desenhando-se um par de oligonucleotídeo iniciador para cada gene, nesta busca foram desenhados 1.460 pares.

Adicionalmente, 144 referências de DNA ou oligonucleotídeos iniciadores descritos na literatura e comumente aplicados em técnicas diagnósticas (qPCR, PCR, *Multilocus enzyme electrophoresis* MLEE, *Multilocus sequence typing* MLST, *Multilocus microsatellite typing* MLMT, *Multilocus sequence analysis* MLSA) para identificação de *Leishmania* spp. foram utilizados (Tabela suplementar 2). Juntando-se as três estratégias de busca de sequências, 2.145 pares de oligonucleotídeos iniciadores foram selecionados.

Desenho de oligonucleotídeos iniciadores

O desenho de pares de oligonucleotídeos iniciadores foi realizado com auxílio do programa BatchPrimer3, com os parâmetros: tamanho entre 18-22pb, temperatura de *melting* entre 58-62°C e que resultassem em amplificação de fragmentos de 200-1000pb. A diferença de temperatura entre os pares de oligonucleotídeos esperada era inferior a 5°C e a porcentagem de CG, entre 40-60%, foram verificados se haveria possibilidade de formação de *hairpin*, *dimer* ou mesmo *cross-dimer*. Os pares de oligonucleotídeos que passaram por todas as verificações foram selecionados para o ePCR.

Sequenciamento de amostras de referência de *Leishmania* spp.

O sequenciamento de amostras de referência foi realizado com o objetivo de se obter dados que pudessem ser usados para validação do painel das sequências espécie-específicas selecionadas. A extração de DNA de amostras de cultura de *Leishmania* cedidas pela Fundação Oswaldo Cruz do Rio de Janeiro-FIOCRUZ (Tabela 1) foi realizada com kit comercial QIAmp DNA mini Kit (Qiagen®), segundo instruções do fabricante. Bibliotecas *paired-end* foram feitas

com o kit comercial DNA Nextera Library Preparation (Illumina®), e, após tagmentação, o tamanho médio dos fragmentos e a qualidade da biblioteca foram avaliados utilizando-se Bioanalyzer (Agilent Technologies®) (Figura suplementar 1). O sequenciamento foi realizado no sequenciador Miseq (Illumina®), com cobertura de 80 e 400 vezes (Tabela 1).

Tabela 1. Cepas de referência de espécies de *Leishmania* sequenciadas. Cepas de referência de espécies de *Leishmania* sequenciadas. *Espécies cujos genomas ainda não estão publicados.

Espécies	WHO código da cepa	Cobertura (vezes)
<i>L. (L.) mexicana</i>	(MHOM/BZ/1982/BEL21)	80
<i>L. (L.) guyanensis</i> *	IOC/L0565 (MHOM/BR/1975/M4147)	80
<i>L. (L.) amazonensis</i>	IOC/L0575 (MHOM/BR/1967/PH8)	80
<i>L. (L.) infantum</i>	(MHOM/BR/1974/PP75)	400
<i>L. (L.) shawi</i> *	(MHOM/BR/1984/M8408)	80

Os arquivos FASTQ contendo as sequências *forward* e *reverse* resultantes da metodologia utilizada (*paired-end*) foram avaliados quanto à qualidade total obtida e à qualidade individual por base sequenciada, efetuando-se a remoção daquelas que não atendessem os critérios de qualidade estabelecidos pelo programa Miseq *Reporter* (Illumina®). As sequências resultantes foram então submetidas a uma *pipeline* de montagem com uso do software de livre acesso ABySS 1.9.0 *Assembler*, de maneira a obter o maior número possível de bases que compõem uma sequência confiável (*scaffolds*) para cada espécie.

Amplificação de DNA por ePCR

Os 2.145 pares de oligonucleotídeos iniciadores selecionados foram validados, por ePCR (ferramenta do banco NCBI), utilizando-se os genomas de *Leishmania* do banco de dados TriTrypDB (Tabela suplementar 3) bem como as sequências (*scaffolds*) geradas a partir do sequenciamento das cepas de referência (Tabela 1). Para o alinhamento dos oligonucleotídeos iniciadores incorporados ao banco de dados, houve permissão de lacunas (*gaps*) de até dois nucleotídeos e permissão de erros de incorporação de nucleotídeos (*mismatched*) de até 2, assim a sensibilidade da técnica de ePCR foi aumentada, sem detrimento da especificidade. Contudo, mais de 706.000 resultados de anelamento dos oligonucleotídeos iniciadores com os genomas e com os *scaffolds* foram observados, os quais ocorreram em cromossomos, *contigs*, *supercontigs* ou em sequências aleatórias.

Na filtragem dos resultados foram descartadas todas as sequências que apresentaram alinhamento em espécies que não pertenciam ao gênero *Leishmania*. Os grupos identificados foram selecionados e separados de acordo com a espécie e o tipo de manifestação clínica da doença em humanos.

Filogenia

Para distinção das leishmanias com base molecular foi realizada a relação filogenética entre as espécies de *Leishmania* com genomas completos e anotados, sendo selecionados seis genomas de *Leishmania* (*Leishmania mexicana*, *Leishmania major*, *Leishmania braziliensis*, *Leishmania donovani*, *Leishmania infantum*, *Leishmania tarentolae*) e o genoma de *Trypanosoma brucei*, disponíveis no banco TriTrypDB (release 9.0.) A espécie *Trypanosoma brucei* foi utilizada como grupo externo para o estabelecimento da relação filogenética entre as espécies.

Dentre todas as cepas disponíveis (duas de *L. braziliensis*, uma de *L. donovani*, uma de *L. infantum*, três de *L. major*, uma de *L. mexicana*, uma cepa de *L. tarentolae* e três de *T. brucei*) foram encontrados 17.381 genes ortólogos. Dentre os genes ortólogos que codificam para proteínas foram filtrados os de cópia única, por *script* feito em Unix, tendo sido selecionadas apenas as proteínas comuns à todas as espécies investigadas, totalizando assim, 48 proteínas (Tabela suplementar 4). As funções das proteínas selecionadas foram definidas no banco de dados UniProt.

Os dados foram concatenados e organizados numa supermatriz envolvendo a combinação de todos os caracteres para que todas as proteínas pudessem ser analisadas simultaneamente. Os alinhamentos gerados pela supermatriz foram realizados com auxílio do programa Clustalx. A supermatrix com as proteínas alinhadas foi utilizada para construir uma árvore filogenética com *neighbor-joining* (NJ), utilizando o *p-distance* aminoácidos com auxílio do programa MEGA 6.0.

Resultados

Montagem das sequências de DNA de amostras de referência de *Leishmania spp.*

Os *scaffolds* obtidos a partir da montagem das sequências são compatíveis com os relatórios de qualidade gerados durante o processo de sequenciamento, e se correlacionam à cobertura esperada para cada espécie e respectivos números de bases montadas (Tabela 1 e 2). A intensidade dos dados de sequenciamento obtidos para a amostra de *L. mexicana* foi significativamente inferior às demais amostras, indicando redução proporcional à quantidade de DNA extraído da mesma, porém, não exercendo influência negativa no que diz respeito ao número total de bases montadas (Tabela 2). Em linhas gerais, os resultados obtidos

refletem 90% do total de bases montadas dos genomas de cada espécie quando comparados às suas respectivas referências, com uma média aproximada de 28Mb sequenciadas para cada amostra, quando comparada ao número total de bases contidos nos genomas de referência (31Mb) (Tabela 2).

Tabela 2. Montagem dos *scaffolds* de amostras de referência de *Leishmania* spp.. Resultados dos *scaffolds* dos genomas de seis espécies de *Leishmania* spp..

Espécies	N	n:500	L50	min	N80	N50	N20	E-size	max	Sum
<i>L. mexicana</i>	52258	15071	3529	500	1118	2166	4049	2949	32516	25.43e6
<i>L. amazonensis</i>	36024	6823	1179	500	3151	7245	14058	9044	43306	28.73e6
<i>L. guyanensis</i>	37691	8277	1502	500	2501	5405	10969	7454	56917	28.44e6
<i>L. shawi</i>	32246	6794	1206	500	3160	6974	13912	9031	51113	28.8e6
<i>L. infantum</i>	140488	2555	393	500	9848	22245	45820	29498	131675	30.65e6

n= número de *scaffolds* gerados na montagem; n:500 = número de *scaffolds* > 500bp; L50= número de *scaffolds* que compõem 50% do total de bases montadas; min= tamanho mínimo dos *scaffolds*; N80= tamanho mínimo dos *scaffolds* que cobrem 80% do total de bases montadas; N50= tamanho mínimo dos *scaffolds* que cobrem 50% do total de bases montadas; N20= tamanho mínimo dos *scaffolds* que cobrem 20% do total de bases montadas; E-size= tamanho esperado de um *scaffold* escolhendo uma base aleatória do conjunto; max= tamanho máximo dos *scaffolds* gerados; sum= total de bases montadas.

Seleção de sequências de DNA de *Leishmania* spp. e amplificação por ePCR

Dos 2.145 pares de *oligonucleotídeos iniciadores* utilizados para seleção de sequências alvo de DNA de *Leishmania* do banco de dados TriTrypDB, à ePCR, 2.126 identificaram o gênero, um grande número dos oligonucleotídeos iniciadores testados compartilharam sequências entre as espécies de *Leishmania* e um menor número, 960, revelou ser espécie específico (Tabela 3). Apenas 19 pares resultaram em amplificação de DNA de sequências que não pertenciam ao gênero *Leishmania*.

Quando se considerou somente os dados referentes aos genomas publicados e aos *scaffolds* das espécies de referência sequenciadas no presente estudo, observou-se que o número de oligonucleotídeos iniciadores compartilhando sequências também variou entre as espécies e entre cepas (Tabela 4).

Tabela 3. Número de *oligonucleotídeos iniciadores* específicos segundo a espécie de *Leishmania*. Número de *oligonucleotídeos iniciadores* que amplificaram regiões espécie-específicas de *Leishmania*.

Espécie	nPrimer	nPrimerspecific
<i>Leishmania_aethiopica</i>	259	44
<i>Leishmania_amazonensis</i>	72	1
<i>Leishmania_arabica</i>	255	38
<i>Leishmania_braziliensis</i>	560	156
<i>Leishmania_donovani</i>	257	11
<i>Leishmania_enriettii</i>	294	291
<i>Leishmania_gerbilli</i>	272	21
<i>Leishmania_guyanensis</i>	278	NA
<i>Leishmania_infantum</i>	284	19
<i>Leishmania_major</i>	367	103
<i>Leishmania_mexicana</i>	107	16
<i>Leishmania_panamensis</i>	398	28
<i>Leishmania_tarentolae</i>	174	153
<i>Leishmania_tropica</i>	265	57
<i>Leishmania_turanica</i>	274	22
<i>Leishmania_shawi</i>	311	NA
<i>Leishmania_sp</i>	240	235

Espécie= total de cepas para cada espécie utilizada neste trabalho; **nPrimer** = número total de oligonucleotídeos iniciadores que alinharam em determinada espécie; **nPrimerspecific** = número de oligonucleotídeos iniciadores específicos para determinada espécie de *Leishmania*; NA = não avaliado.

Com a montagem dos *scaffolds* das espécies de *Leishmania* foi possível verificar a concordância do alinhamento dos pares de oligonucleotídeos iniciadores selecionados nas espécies correspondentes no TriTrypDB com o da montagem *De Novo* realizada neste estudo (Tabela 5). Para as espécies *Leishmania shawi* e *Leishmania guyanensis* que foram sequenciadas pela primeira vez no presente estudo, não foi possível confrontar os resultados. Para as espécies *Leishmania infantum* e *Leishmania mexicana* todos os pares de oligonucleotídeos iniciadores que se anelaram aos *scaffolds* montados se anelaram aos genomas das espécies correspondentes no TriTrypDB.

Tabela 5. Validação do alinhamento dos pares de oligonucleotídeos iniciadores nas culturas de referência. Número de oligonucleotídeos iniciadores que se alinham aos *scaffolds* montados com as amostras de referência de *Leishmania* e que se alinham às espécies correspondentes depositadas no banco de dados TriTrypDB. **nPrimer scaffolds**= número de pares de oligonucleotídeos iniciadores que apresentaram alinhamento ao *scaffold* montado das culturas de referência de *Leishmania*. **nPrimer Total**= número de pares de oligonucleotídeos iniciadores com alinhamento nos genomas depositados do TriTrypDB e nos *scaffolds* montados.

Espécies	nPrimer <i>scaffolds</i>	nPrimer Total
<i>Leishmania amazonensis</i>	63	52
<i>Leishmania infantum</i>	207	207
<i>Leishmania mexicana</i>	58	58

Utilizando-se os pares de oligonucleotídeos iniciadores já descritos da literatura (Tabela suplementar 3), foi possível fazer a distinção entre seis espécies de *Leishmania*: *L. amazonensis*, *L. braziliensis*, *L. donovani*, *L. infantum*, *L. major*, e *L. panamensis*, pertencentes a dois subgêneros (*Viannia* e *Leishmania*). Destes, 28 pares de oligonucleotídeos iniciadores revelaram ser específicos para discriminação de espécies, de acordo com a especificidade apresentada à ePCR (Tabela 6).

Tabela 6. Sequências selecionadas a partir dos dados da literatura. As sequências estão divididas por subgênero *Viannia* e *Leishmania* e pela espécie que pôde ser discriminada *in silico*, com os pares de oligonucleotídeos iniciadores citados. As referências representam os artigos que publicaram o par de oligonucleotídeo iniciador ou a região de interesse e a concordância à ePCR concorda com a citação da referência.

Subgênero e espécies	Oligonucleotídeos iniciadores literatura	Referências	Concordância à ePCR
Viannia			
<i>Leishmania braziliensis</i>	argonaute/dicer-Primer4	Peacock et al., 2007/Toledo et al., 2010/Llanes et al., 2015	Sim
	putative-flagellar-calcium-binding-Primer1	Peacock et al., 2007	Sim
	ISVB-ISVC-Primer	Castilho et al., 2003	Sim
	b1-b2-Primer	Mimori et al., 1998	Sim
	mp1MPlintF/Rprimer	Tsukayama et al., 2009	Sim
	B6FPrimer	Oddone et al., 2009	Sim
	AC01RPrimer	Oddone et al., 2009	Sim
<i>Leishmania panamensis</i>	CSg48Primer	Oddone et al., 2009	Sim
Leishmania			
<i>Leishmania amazonensis</i>	a1/a2kDNAmnicircle-Primer	Mimori et al., 1998	Sim
<i>Leishmania infantum</i>	LEISH1-LEISH2-Primer	Piedra, 2002	Sim
<i>Leishmania donovani</i>	nh1F1/R3Primer	Mauricio et al., 2006	Sim
<i>Leishmania major</i>	metallo-peptidaseClanMG-Primer1	Peacock et al., 2007/Reyes-Urbe et al., 2012	Sim
	putative-carboxypeptidase-Primer1	Peacock et al., 2007/Isaza et al., 2008	Sim
	malic-enzyme-Primer1	Zemanová et al., 2007	Sim
	malic-enzyme-Primer2	Zemanová et al., 2007	Sim
	flavoprotein-subunit-like-Primer1	Oppedores e Szikora, 2006/Peacock et al., 2007	Sim
	flavoprotein-subunit-like-Primer2	Oppedores e Szikora, 2006/Peacock et al., 2007	Sim
	flavoprotein-subunit-like-Primer4	Oppedores e Szikora, 2006/Peacock et al., 2007	Sim
	flavoprotein-subunit-like-Primer5	Oppedores e Szikora, 2006/Peacock et al., 2007	Sim
	glucose-6-phosphate-1-dehydrogenase-G6PD-Primer1	Prina et al., 2007 /Zemanova et al., 2007/Boité et al., 2012/ Jara et al., 2013	Sim
	internaltranscribedspacer-1-Primer1	Ei Tai et al., 2001/Kuhls et al., 2005/Berzunza-Cruz et al., 2015	Não
	internaltranscribedspacer-1-Primer2	Ei Tai et al., 2001/Kuhls et al., 2005/Berzunza-Cruz et al., 2015	Não
	internaltranscribedspacer-1-Primer3	Ei Tai et al., 2001/Kuhls et al., 205/Berzunza-Cruz et al., 2015)	Não
	internaltranscribedspacer-1-Primer5	Ei Tai et al., 2001/Kuhls et al., 2005/ Marlow et al., 2014 /Berzunza-Cruz et al., 2015	Não
	mitochondrialmalatedehydrogenase-Primer1		Não
	RV1-RV2-Primer	Lauchaud et al., 2002/Ferroglio et al., 2006/Gomes et al., 2007	Não
	gpiF3/R4Primer	Mauricio et al., 2006	Não

Quando as espécies de *Leishmania* foram agrupadas segundo as formas clínicas da doença um conjunto de pares de oligonucleotídeos iniciadores partilhados entre as espécies de *Leishmania* pôde ser definido (Tabela 7). Do total de pares de oligonucleotídeos iniciadores (1.212) que se alinharam nas espécies envolvidas na forma cutânea, 550 pares foram exclusivos para as espécies que causam este tipo de manifestação clínica, enquanto que para as formas

mucocutânea (626) 263 foram exclusivos, e para a visceral (300), 132 pares de oligonucleotídeos iniciadores foram exclusivos.

A seleção de sequências-alvo de DNA não ortólogas com outros tripanosomatídeos e com variado número de cópias possibilitou a seleção de alvos específicos para 13 espécies de *Leishmania*: *L. aethiopica*, *L. arabica*, *L. braziliensis*, *L. donovani*, *L. enriettii*, *L. gerbilli*, *L. infantum*, *L. major*, *L. mexicana*, *L. panamensis*, *L. tarentolae*, *L. tropica* e *L. turanica*.

Tabela 7. Número de pares de oligonucleotídeos iniciadores partilhados entre as espécies de *Leishmania* segundo a forma clínica da doença. Número total de pares de oligonucleotídeos iniciadores que fazem alinhamento com sequências de DNA de espécies/cepas de *Leishmania* segundo a manifestação clínica da infecção humana. A diagonal (em cinza) representa o número total de oligonucleotídeos iniciadores que alinharam nas espécies agrupadas para cada conjunto de manifestação.

Manifestação clínica	Cutânea	Mucocutânea	Visceral	Sem classificação	<i>Lizard Leishmania</i>
Cutânea	1212 (550*)	389	160	279	20
Mucocutânea	389	626 (263*)	45	44	12
Visceral	160	45	300 (132*)	119	13
Sem classificação	279	44	119	650	13
<i>Lizard Leishmania</i>	20	12	13	13	174

(*) Número de pares de oligonucleotídeos iniciadores exclusivos para a manifestação clínica correspondente.

Filogenia

A filogenia baseada nas proteínas ortólogas de cópia única, resultou em uma árvore de perfil monofilético em que as espécies *Leishmania donovani* e *Leishmania infantum* (subgênero *Leishmania*), que são responsáveis pela forma visceral da doença e as espécies *Leishmania mexicana* e *Leishmania major* (subgênero *Leishmania*), com implicações na forma cutânea da doença ficaram, respectivamente, agrupadas em ramos com maior proximidade (Figura 1). *Leishmania braziliensis* (subgênero *Viannia*), envolvida na forma mucocutânea da leishmaniose e *Leishmania tarentolae* (subgênero *Sauroleishmania*), não-patogênica para humanos, foram posicionadas em ramos distintos, em comparação às outras espécies.

Adicionalmente, cada proteína foi também avaliada individualmente (sem concatenação das 48 proteínas selecionadas) em uma árvore construída com as espécies de *Leishmania* resultando em 78% de semelhança à árvore filogenética descrita pela concatenação das proteínas (dados não mostrados).

Comparando-se os resultados da tabela de sequências de *oligonucleotídeos iniciadores* selecionados com os ramos da árvore filogenética, observa-se complementariedade nos dados de distinção molecular entre as leishmânias, conforme taxonomia atual. *Leishmania infantum* e *Leishmania donovani* agrupadas no mesmo ramo partilham entre elas 241 pares de *oligonucleotídeos iniciadores*, enquanto que no ramo mais distante, *L. tarentolae* compartilha um baixo número de *oligonucleotídeos iniciadores* (no máximo 14) com outras espécies (Tabela 4).

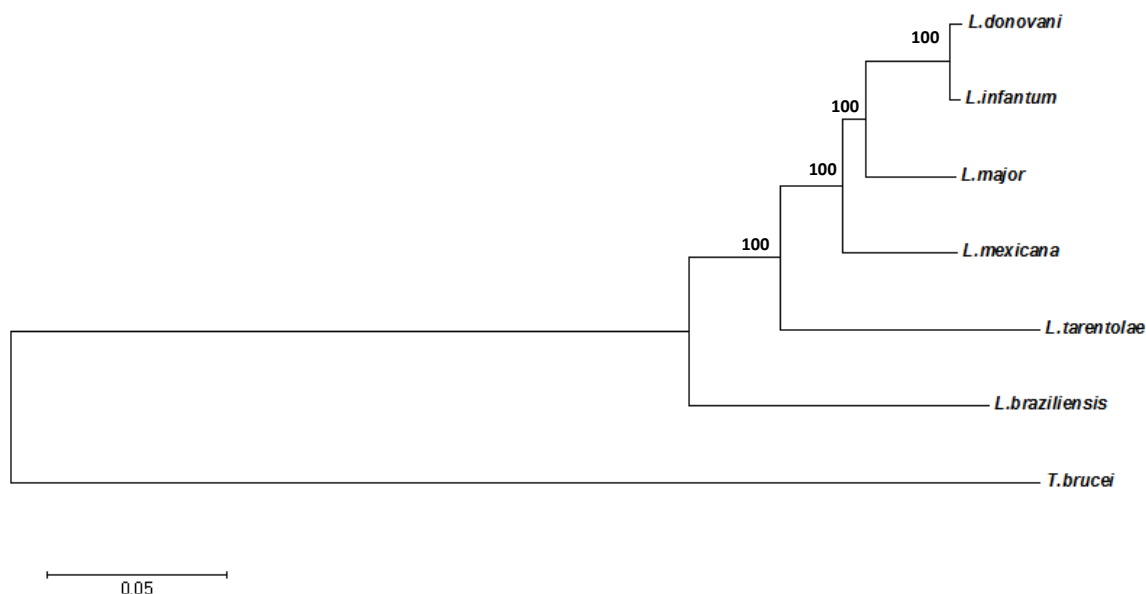


Figura 1. Árvore filogenética de *Leishmania* spp. Árvore filogenética usando *Neighbor-joining*, com modelo de substituição (p-distância de aminoácidos), *bootstrap* 100 vezes, utilizando-se 48 proteínas ortólogas e de cópia única, comuns às seis espécies de *Leishmania*: *L. donovani*, *L. infantum*, *L. major*, *L. mexicana*, *L. tarentolae* e *L. braziliensis*, além de *Trypanosoma brucei*, utilizada como grupo externo.

Discussão

A utilização da abordagem de seleção de sequências de DNA de *Leishmania* a partir de banco de dados genômicos associada àqueles já descritos na literatura resultou em alvos específicos para 14 espécies de *Leishmania*, totalizando 960 sequências-alvo de DNA específicas com potencial uso para a diferenciação de espécies de *Leishmania* spp. em amostras clínicas.

Dos pares de oligonucleotídeos iniciadores que se alinham em sequências dos genomas de *Leishmania* depositados no banco de dados

TrypTrypDB, 100% também apresentaram alinhamento com os *scaffolds* das sequências de referência *Leishmania mexicana* (MHOM/BZ/1982/BEL21) e *Leishmania infantum* (MHOM/BR/1974/PP75) montadas em nosso estudo, validando a abordagem adotada para identificação de espécies específicas.

Apesar do alto número de pares de oligonucleotídeos iniciadores selecionados neste estudo apenas um par de oligonucleotídeos iniciadores foi específico para *Leishmania amazonensis* tanto nos genomas depositados no banco TriTrypDB como nos genomas de amostras de referência montados, sugerindo que este genoma possa ser menos diferenciado. *Leishmania enrietti* apresentou um elevado número de sequências específicas e compartilha poucos pares de oligonucleotídeos iniciadores com as demais leishmânias.

A falta de informação genômica depositada no banco de dados TriTrypDB com relação à *Leishmania shawi* e *Leishmania guyanensis* impossibilitou que fossem confrontados os resultados encontrados de alinhamentos que ocorreram nos *scaffolds* gerados em relação ao banco de dados utilizado, porém encontramos um elevado número de pares de oligonucleotídeos iniciadores que se alinharam nessas espécies e que são compartilhados entre as mesmas, assim como com as espécies *L. braziliensis* e *L. panamensis*, que pertencem ao subgênero *Viannia*.

Entretanto, uma das limitações foi o fato de que não pudemos confrontar o alinhamento dos oligonucleotídeos iniciadores para outras espécies de *Leishmania*, as quais encontramos alvos específicos nos genomas do TriTrypDB, visto que não dispunhamos de outras amostras de referência para serem sequenciadas em nosso estudo.

Na avaliação por ePCR do painel de pares de oligonucleotídeos iniciadores utilizando-se as sequências de DNA baseadas no TriTrypDB foi possível identificar 13 espécies de *Leishmania* por 845 pares de oligonucleotídeos iniciadores. Já nas sequências descritas na literatura como capazes de discriminar as espécies de *Leishmania*, por ensaio molecular, foi possível diferenciar seis espécies (*L.*

amazonensis, *L. braziliensis*, *L. donovani*, *L. infantum*, *L. major* e *L. panamensis*), utilizando-se 28 pares de oligonucleotídeos iniciadores. Tais sequências descritas em estudos de diferenciação de espécies de *Leishmania* foram incorporadas ao nosso conjunto de pares de oligonucleotídeos iniciadores selecionados para aumentar a chance de amplificação específica de algumas espécies. Nesta seleção, um único par de oligonucleotídeo iniciador (Tabela 6) foi específico para a identificação da espécie *L. amazonensis* (MIMORI et al., 1998). Da mesma forma, o par de oligonucleotídeo iniciador b1/b2 foi específico para *L. braziliensis*, conforme anteriormente descrito por Mimori et al. (1998). Por outro lado, os oligonucleotídeos iniciadores da região do kinetoplasto, descritos por Piedra (2002) e usados para identificação de *Leishmania* spp. revelaram-se específicos para identificação de *L. infantum*.

Quanto ao subgênero *Leishmania*, alguns oligonucleotídeos da região do DNA ribossomal foram específicos para *L. major*, embora ITS-1 tenha sido amplamente utilizado para a identificação de polimorfismos em *L. donovani* (EL TAI et al., 2001; SCHÖNIAN et al., 2003). As sequências do kinetoplasto que foram selecionadas pelos oligonucleotídeos iniciadores RV1/RV2 (FERROGLIO et al., 2006; GOMES et al., 2007; LACHAUD et al., 2002), anteriormente descrito como específico para o complexo *L. donovani*, alinharam-se especificamente em *L. major*.

O genoma de *Leishmania* é rico em sequências de microssatélites os quais são comumente utilizados nas técnicas de MLMT, por distinguirem os subgêneros *Viannia* e *Leishmania*, sendo um conjunto deles específicos para *L. major*, *L. donovani*, *L. infantum*, *L. tropica*, *L. guyanensis* e *L. braziliensis* (ALURU et al., 2015). Maurício et al. (2006) conseguiram discriminar algumas cepas de *Leishmania* do complexo *L. donovani*, enquanto Tsukayama et al. (2009) identificaram cepas relacionadas ao subgênero *Viannia*, principalmente envolvendo espécies do Novo Mundo como *L. braziliensis*, *L. panamensis*, *L. lansoni*, *L. peruviana* e *L. guyanensis*.

Enzimas de genes de manutenção (*housekeeping*), muito utilizadas em técnicas para tipificação de cepas, resultaram em alguns oligonucleotídeos iniciadores espécie-específicos; a enzima nucleosídeo fosforilase (Tabela 6 - *primer* nH1), sequência explorada por Mauricio et al. (2006), foi a única capaz de identificar *L. donovani* (MAURICIO et al., 2006). Os pares de oligonucleotídeos iniciadores que amplificam a região glucose-6-phosphate-1-dehydrogenase (G6PD) (BOITÉ et al., 2012; JARA et al., 2013; PRINA et al., 2007; ZEMANOVÁ et al., 2007) e enzima málica (ZEMANOVÁ et al., 2007) discriminaram a espécie *L. major* das demais espécies de *Leishmania*, em concordância com a literatura.

Do subgênero *Viannia*, as espécies *L. panamensis* e *L. braziliensis* puderam ser discriminadas quando se avaliaram os oligonucleotídeos iniciadores descritos na literatura frente ao banco de dados TriTrypDB. A sequência da enzima manose isomerase fosfato (oligonucleotídeo mp1Mplint) altamente conservada e utilizada para identificação de espécies do Novo Mundo, foi capaz de discriminar *L. braziliensis* também em nossa seleção (TSUKAYAMA et al., 2009). Já Oddone et al. (2009), com a abordagem MLMT, utilizaram alguns alvos para distinção do complexo *Viannia* porém, quando avaliados por ePCR, três dos pares de oligonucleotídeos iniciadores descritos (Csg48, B6F e ACO1) foram específicos para *L. braziliensis* e *L. panamensis*, respectivamente, sendo bons candidatos à diferenciação destas espécies. Com êxito, as sequências usadas nas técnicas MLEE, MLMT e MLSA, que buscam tipificação intra-específica para os estudos epidemiológicos e para as constantes revisões taxonômicas de *Leishmania* (BOITÉ et al., 2012; MARCILI et al., 2014; MAURICIO et al., 2006; ZHANG et al., 2013) identificam sequências-alvo em nosso estudo que possibilitaram a diferenciação de espécies, à ePCR.

O uso das proteínas descritas como diferenciadas entre as espécies *L. infantum*, *L. major* e *L. braziliensis* (PEACOCK et al., 2007) forneceu 84 pares de oligonucleotídeos iniciadores de sequências capazes de discriminar especificamente as três espécies, no presente estudo. Porém, a similaridade entre

L. braziliensis e *L. panamensis* fez com que algumas sequências alvo de DNA, anteriormente reconhecidos como específicos para *L. braziliensis*, fossem também observadas em *L. panamensis*, dado o fato de pertencerem ao mesmo subgênero *Viannia* e de seu genoma só ter sido recentemente publicado (LLANES et al., 2015).

Novos genomas de *Leishmania* e outros protozoários têm sido revelados, assim, informações nos bancos de dados públicos, como o TriTrypDB, são continuamente acrescentadas, tornando esse conhecimento cada vez mais exato devido ao aprimoramento de novas tecnologias aliadas ao uso da bioinformática para a correta montagem e anotação dos mesmos.

A possibilidade de avaliar diversas regiões e/ou sequências de DNA por ePCR, buscando não apenas variações entre cepas e sim regiões específicas por meio dos genomas publicados, permitiu a seleção de 960 sequências candidatas para distinção de 14 espécies de *Leishmania* sob a perspectiva inter-específica. Exemplo disto são as 156 possibilidades de identificação para a espécie *L. braziliensis*, principal representante da manifestação mucocutânea no Brasil e os 103 pares de oligonucleotídeos iniciadores para *L. major*, responsável pela leishmaniose tegumentar. Portanto, um vasto painel de sequências alvo de DNA poderá ser testado com fins diagnósticos.

A partir de dados de proteínas ortólogas de cópia única de seis espécies de *Leishmania* do banco TriTrypDB, construiu-se uma árvore filogenética que revela perfil monofilético, em concordância com os resultados de Marcili et al. (2014) (MARCILI et al., 2014). Este perfil monofilético também já foi demonstrado por meio da análise de sequências de Hsp70 de espécies e cepas de *Leishmania* (MONTALVO et al., 2010).

Análises filogenéticas com proteínas ortólogas de cópia única, altamente conservadas, sugerem lentas taxas de evolução entre espécies de tripanossomatídeos e, para o gênero *Leishmania*, as taxas de evolução são ainda menores em relação a outros tripanossomatídeos (STEVENS; RAMBAUT, 2001).

Apesar das marcantes diferenças na patogenicidade e nos ciclos biológicos, parasitas dos subgêneros *Leishmania*, *Viannia* e *Sauroleishmania* apresentam um perfil monofilético (MARCILI et al., 2014; MONTALVO et al., 2010). Contudo, Rougeron et al. (2015) descreveram uma alta taxa de heterogeneidade genética nas amostras de *Leishmania* spp., mesmo em isolados geograficamente próximos, dificultando a taxonomia das mesmas (ROUGERON et al., 2015). Estudos com cepas de *Leishmania* da África e da Eurásia questionam a atual classificação taxonomia do protozoário, principalmente para espécies que causam a forma visceral, como *L. infantum* e *L. donovani* (EL BAIDOURI et al., 2013). Entretanto, pudemos avaliar a filogenia de algumas espécies de *Leishmania* baseados em dados genômicos já publicados, confirmando as distâncias filogenéticas esperadas para este táxon.

Conclusão

A aplicação de nova abordagem, *in silico*, que compreendeu a seleção e análise do maior número de alvos de DNA de *Leishmania* depositados em um banco genômico já avaliados neste contexto, resultou em um conjunto de pares de oligonucleotídeos iniciadores específicos com potencial à diferenciação de 14 espécies distintas. A possibilidade de customização de um painel de sequências-alvo de DNA de *Leishmania* permitirá, por exemplo, o desenvolvimento de um teste que analise múltiplas amostras simultaneamente, com poder discriminatório entre distintas espécies de *Leishmania*.

Agradecimentos

À FIOCRUZ pelas amostras cedidas para o experimento, à CAPES pela bolsa de estudos do Programa de Doutorado Sanduíche no Exterior (PDSE); à equipe do laboratório de Biologia Celular do Prof. Najib M. El-Sayed, da

Universidade de Maryland, USA; à Deoxi Biotecnologia pela assessoria e realização do sequenciamento das amostras de referência.

Contribuição dos autores

Idealização e desenho do experimento: CMN, JFG, PKRKI, FMO. Realização do experimento: FMO, PKRKI, FFA, YTU, ASC, ATB. Contribuição com materiais/análises: NMES, CMN, JFG, PKRKI. Escrito por: FMO, CMN, JFG. Todos os autores aprovaram e revisaram a versão final do manuscrito.

Referências

ALURU, S. et al. Multilocus microsatellite typing of *Leishmania* and clinical applications: a review. **Parasite**, v. 22, p. 16, 2015.

BERZUNZA-CRUZ, M. et al. *Leishmania (L.) mexicana* Infected bats in Mexico: Novel potential reservoirs. **PLOS Neglected Tropical Diseases**, v. 9, p. e0003438, 2015.

BOITÉ, M. C. et al. New Insights on taxonomy, phylogeny and population genetics of *Leishmania* (Viannia) parasites based on multilocus sequence analysis. **PLoS Neglected Tropical Diseases**, v. 6; p. e1888, 2012.

CASTILHO, T. M.; SHAW, J. J.; FLOETER-WINTER, L. M. New PCR Assay using glucose-6-phosphate dehydrogenase for identification of *Leishmania* species. **Journal of Clinical Microbiology**, v.41, p.540–546, 2003.

CASTILHO, T.M.; CAMARGO, L.M.A., McMAHON-PRATT, D.; SHAW, J.J.; FLOETER-WINTER, L.M. A real-time polymerase chain reaction assay for the identification and quantification of American *Leishmania* species on the basis of glucose-6-phosphate-dehydrogenase. **American Journal of Tropical Medicine and Hygiene**, v.78, p.122-132, 2008.

DESJEUX, P. Leishmaniasis: Current situation and new perspectives. **Comparative Immunology, Microbiology and Infectious Diseases**, v. 27, p. 305–318, 2004.

DOWNING, T.; et al. Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanisms of drug resistance. **Genome Research**, v.21, p.2143-2156, 2011.

EL BAIDOURI, F. et al. Genetic structure and evolution of the *Leishmania* genus in Africa and Eurasia: what does MLSA tell us. **PLoS neglected tropical diseases**, v. 7, p. e2255, 2013.

EL TAI, N. O. et al. Genetic heterogeneity of ribosomal internal transcribed spacer in clinical samples of *Leishmania donovani* spotted on filter paper as revealed by single-strand conformation polymorphisms and sequencing. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 94, p. 575–579, 1996.

EL TAI, N. O. et al. *Leishmania donovani*: intraspecific polymorphisms of Sudanese isolates revealed by PCR-based analyses and DNA sequencing. **Experimental parasitology**, v. 97, p. 35–44, 2001.

FERROGLIO, E. et al. Characterization of *Leishmania infantum* strains in blood samples from infected dogs and humans by PCR-RFLP. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, v. 100, p. 636–641, 2006.

GOMES, A. H. S. et al. PCR identification of *Leishmania* in diagnosis and control of canine leishmaniasis. **Veterinary Parasitology**, v. 144, p. 234–241, 2007.

GROVER, C. E.; SALMON, A.; WENDEL, J. F. Targeted sequence capture as a powerful tool for evolutionary analysis. **American Journal of Botany**, v. 99, p. 312–319, 2012.

HAJJARAN, H. et al. Direct diagnosis of *Leishmania* species on serosity materials punctured from cutaneous leishmaniasis patients using PCR-RFLP. **Journal of clinical laboratory analysis**, v. 25, p. 20–4, 2011.

ISAZA, C. E. et al. A proposed role for *Leishmania major* carboxypeptidase in peptide catabolism. **Biochemical and Biophysical Research Communications**, v. 373, p. 25–29, 2008.

IVENS, A.C.; LEWIS, S.M.; BAGHERZADEH, A.; ZHANG, L.; CHAN, H.M.; SMITH, D.F. Genome Friedlin *Leishmania major* A physical map of the the *Leishmania major* Friedlin Genome. **Genome Research**, v. 8, p. 135-145, 1998.

IVENS, A.C.; et al. The genome of the kinetoplastid parasite, *Leishmania major*. **Science**, v. 309, p. 436–442, 2005.

JARA, M. et al. Real-time PCR assay for detection and quantification of *Leishmania* (Viannia) organisms in skin and mucosal lesions: Exploratory study of parasite load and clinical parameters. **Journal of Clinical Microbiology**, v. 51, p. 1826-1833, 2013.

KUHLS, K. et al. Analysis of ribosomal DNA internal transcribed spacer sequences of the *Leishmania donovani* complex. **Microbes and Infection**, v. 7, p. 1224–1234, 2005.

KUHLS, K. et al. Multilocus microsatellite typing (MLMT) reveals genetically isolated populations between and within the main endemic regions of visceral leishmaniasis. **Microbes and Infection**, v. 9, p. 334–343, 2007.

LACHAUD, L. et al. Comparison of six PCR methods using peripheral blood for detection of canine visceral leishmaniasis. **Journal of Clinical Microbiology**, v. 40, p. 210–5, 2002.

LLANES, A. et al. The genome of *Leishmania panamensis*: insights into genomics of the L. (Viannia) subgenus. **Scientific Reports**, v. 5, p. 8550, 2015.

MARCILI, A. et al. Phylogenetic relationships of *Leishmania* species based on trypanosomatid barcode (SSU rDNA) and gGAPDH genes: Taxonomic revision of *Leishmania* (*L.*) *infantum chagasi* in South America. **Infection, Genetics and Evolution**, v. 25, p. 44-51, 2014.

MARLOW, M. A. et al. Multilocus Sequence Analysis for *Leishmania braziliensis* Outbreak Investigation. **PLoS Neglected Tropical Diseases**, v. 8, p. e2695, 2014.

MAURICIO, I. L. et al. Towards multilocus sequence typing of the *Leishmania donovani* complex: Resolving genotypes and haplotypes for five polymorphic metabolic enzymes (ASAT, GPI, NH1, NH2, PGD). **International Journal for Parasitology**, v. 36, p. 757-769, 2006.

MIMORI, T. et al. Rapid identification of *Leishmania* species from formalin-fixed biopsy samples by polymorphism-specific polymerase chain reaction. **Gene**, v. 210, p. 179–186, 1998.

MONTALVO, A. M. et al. Heat-shock protein 70 PCR-RFLP: a universal simple tool for *Leishmania* species discrimination in the New and Old World. **Parasitology**, v. 137, p. 1159-1168, 2010.

OCHSENREITHER, S. et al. Multilocus microsatellite typing as a new tool for discrimination of *Leishmania infantum* MON-1 strains. **Journal of Clinical Microbiology**, v. 44, p. 495–503, 2006.

ODDONE, R. et al. Development of a multilocus microsatellite typing approach for discriminating strains of *Leishmania* (Viannia) species. **Journal of Clinical Microbiology**, v. 47, p. 2818–2825, 2009.

OPPERDOES, F. R.; SZIKORA, J. P. In silico prediction of the glycosomal enzymes of *Leishmania major* and trypanosomes. **Molecular and Biochemical Parasitology**, v. 147, p. 193–206, 2006.

PEACOCK, C. S. et al. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. **Nature genetics**, v.39, p.839-847, 2007.

PIEDRA, Y. M. Trabajos originales. v. 19, p. 109–116, 2002.

PRINA, E. et al. *Leishmania* DNA is rapidly degraded following parasite death: an analysis by microscopy and real-time PCR. **Microbes and Infection**, v. 9, p. 1307–1315, 2007.

REAL, F. et al. The genome sequence of *Leishmania (Leishmania) amazonensis*: Functional annotation and extended analysis of gene models. **DNA Research**, v. 20, p. 567-81, 2013.

REALE, S. et al. Multilocus microsatellite polymorphism analysis to characterize *Leishmania infantum* strains isolated in Sicily. **Transboundary and Emerging Diseases**, v. 57, p. 37–41, 2010.

REYES-URIBE, P. et al. Comparative zymographic analysis of metallopeptidase of *Leishmania* (Viannia) *peruviana* and *Leishmania* (Viannia) *braziliensis* isolates from Peru. **Parasitology International**, v. 61, p. 513–519, 2012.

RIOUX, J.A.; LANOTTE, G.; SERRES, E.; PRATLONG, F.; BASTIEN, P.; PERIERES, J. Taxonomy of *Leishmania*. Use of isoenzymes. Suggestions for a new classification. **Annales de Parasitologie Humaine et Comparee**, v. 65, p. 111–25, 1990.

ROGERS, M. B. et al. Chromosome and gene copy number variation allow major structural change between species and strains of *Leishmania*. **Genome Research**, v. 21, p. 2129–2142, 2011.

ROUGERON, V.; DE MEEÛS, T.; BAÑULS, A-L. A primer for *Leishmania* population genetic studies. **Trends in Parasitology**, v. 31, p. 52–59, 2015.

S. DE TOLEDO, J. Using Genomic Information to Understand *Leishmania* Biology. **The Open Parasitology Journal**, v. 4, p. 156–166, 2010.

SAVANI, E. S. M. M. et al. Detection of *Leishmania* (*Leishmania*) *amazonensis* and *Leishmania* (*Leishmania*) *infantum chagasi* in Brazilian bats. **Veterinary parasitology**, v. 168, p. 5–10, 2010.

SCHÖNIAN, G. et al. PCR diagnosis and characterization of *Leishmania* in local and imported clinical samples. **Diagnostic Microbiology and Infectious Disease**, v. 47, n. 1, p. 349–358, set. 2003.

SCHÖNIAN, G.; MAURICIO, I.; CUPOLILLO, E. Is it time to revise the nomenclature of *Leishmania*? **Trends in Parasitology**, v. 26, p. 466–469, 2010.

SCHÖNIAN, G.; KUHLS, K.; MAURICIO, I. L. Molecular approaches for a better understanding of the epidemiology and population genetics of *Leishmania*. **Parasitology**, p. 405–425, 2011.

STEVENS, J.; RAMBAUT, A. Evolutionary rate differences in trypanosomes. **Infection, Genetics and Evolution**, v. 1, p. 143–150, 2001.

TSUKAYAMA, P.; LUCAS, C.; BACON, D. J. Typing of four genetic loci discriminates among closely related species of New World *Leishmania*. **International Journal for Parasitology**, v. 39, p.355-362, 2009.

ULIANA, S.R.B.; NELSON, K.; BEVERLEY, S.M.; CAMARGO, E.P.; FLOETER-WINTER, L.M. Discrimination amongst *Leishmania* by polymerase reaction chain and hybridization with small subunit ribosomal DNA derived oligonucleotides. **Journal of Eukaryotes and Microbiology**, v. 41, p. 324-330, 1994.

VAN DER AUWERA, G.; DUJARDIN, J.-C. Species Typing in Dermal Leishmaniasis. **Clinical Microbiology Reviews**, v. 28, p. 265–294, 2015.

WORLD HEALTH ORGANIZATION. Post-kala-azar dermal leishmaniasis: a manual for case management and control. Geneva: **WHO**; 2012.

YANG, B.-B. et al. Analysis of kinetoplast cytochrome b gene of 16 *Leishmania* isolates from different foci of China: different species of *Leishmania* in China and their phylogenetic inference. **Parasites Vectors**, v. 6, p.32, 2013.

ZEMANOVÁ, E. et al. The *Leishmania donovani* complex: Genotypes of five metabolic enzymes (ICD, ME, MPI, G6PDH, and FH), new targets for multilocus sequence typing. **International Journal for Parasitology**, v. 37, p. 149–160, 2007.

ZHANG, C.-Y. et al. Phylogenetic and evolutionary analysis of Chinese *Leishmania* isolates based on multilocus sequence typing. **PloS one**, v. 8, p. e63124, 2013.

APÊNDICE (S)

Figura suplementar 1. Tamanho médio dos fragmentos gerados na biblioteca *Paired-end* Nextera® por Bioanalyzer para posterior sequenciamento de uma cepa de *Leishmania infantum*.

Electropherogram Summary

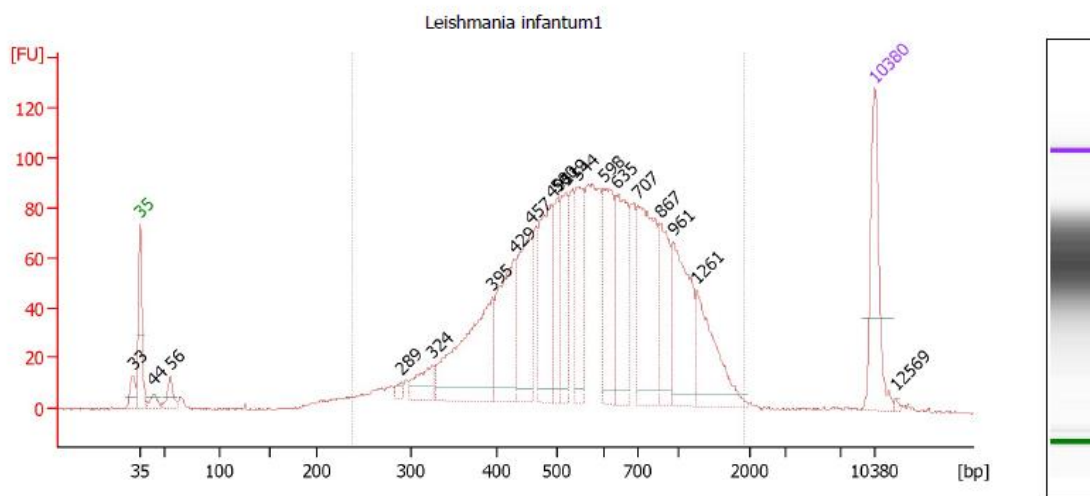


Tabela suplementar 1. Análise comparativa publicada por Peacock et al. (2007) das espécies *L. braziliensis*, *L. major* e *L. infantum*.

Supplementary Table 2. <i>Leishmania</i> loci with species-specific differences and					
conserved pseudogenes					
Product	<i>L. major</i>	<i>L. infantum</i>	<i>L. braziliensis</i>	<i>T. brucei</i>	<i>T. cruzi</i> (Tc00.1047053...)
<i>L. major</i> specific					
Hypothetical protein	LmjF10.0185	LinJ10.0475	-	-	-
Hypothetical protein	LmjF10.1225	LinJ10.1585	LbrM10_V2.1350	Tb927.8.4690	5,061,536
EF hand protein	LmjF13.1450	LinJ13.1380	-	-	-
Hypothetical protein	LmjF31.0260	LinJ31.0270	LbrM31_V2.0370	-	-
Cysteine peptidase	LmjF35.3910	LinJ35.4000	LbrM34_V2.3890	-	-
<i>L. infantum</i> specific					
Hypothetical protein	-	LinJ02.0670	-	-	-
Hypothetical protein	LmjF08.0135	LinJ08.0140	LbrM08_V2.0140	Tb05.6E7.480	51,165,360
Tuzin	LmjF08.0795	LinJ08.0750	LbrM08_V2.0690	-	-
Hypothetical protein	LmjF10.1085	LinJ10.1430	LbrM10_V2.1190	-	-
Phosphatidylinositol	LmjF14.0020	LinJ14.0020	LbrM14_V2.0020	-	50,885,990
3-kinase					
Hypothetical protein	LmjF15.0840	LinJ15.0890	-	-	50,650,940
Hypothetical protein	LmjF20.1175	LinJ20.1210	LbrM20_V2.5335	-	-
Methylenetetrahydrofo	LmjF22.0340	LinJ22.0330	-	-	51,180,920
late dehydrogenase					
Hypothetical protein	-	LinJ22.0410	-	-	-
repeat					
Hypothetical protein	-	LinJ22.0680	-	-	-
Hypothetical protein	LmjF24.2045	LinJ24.1430	-	-	-
Glutathionyl-	LmjF25.2380	LinJ25.2500	LbrM25_V2.1980		508,479,110
spermidine synthase					
Hypothetical protein	LmjF28.0420	LinJ28.0330	LbrM28_V2.0430	Tb11.01.0230	50,881,930
Multidrug resistance	-	LinJ30.1840	LbrM24_V2.1400	Tb927.8.2160	-
like protein					
n-acyl-l-amino acid	-	LinJ31.1490	-	-	-
amidohydrolase					
Hypothetical protein	LmjF31.2195	LinJ31.2630	-	-	-
p-nitrophenyl-	-	LinJ31.3030*	-	-	-
phosphatase					
Hypothetical protein	LmjF31.2465	LinJ31.3160	-	-	-
Hypothetical protein	LmjF32.1505	LinJ32.1900	LbrM32_V2.1680	-	-
Hypothetical protein	LmjF33.2905	LinJ33.2710	LbrM33_V2.3195	Tb927.2.5530	5,034,134
Hypothetical protein	LmjF34.3645	LinJ34.3170	-	-	-
phosphatidylinositol/	-	LinJ36.2050	LbrM35_V2.0690	-	51,029,320
phosphatidylcholine/					
sec14 cytosolic factor					
Hypothetical protein	LmjF36.0580	LinJ36.2060	LbrM35_V2.0700	-	-
Glyceraldehydes 3-	LmjF36.2350	LinJ36.4900	-	-	51,018,760
phosphate					
dehydrogenase					
Hypothetical protein	LmjF36.3995	LinJ36.2750	LbrM35_V2.4235	-	-
<i>L. braziliensis</i> specific					

SLACS	-	-	LbrM02_V2.0550	Tb09.211.5015	-
SLACS	-	-	LbrM02_V2.0690	Tb09.211.5015	-
SLACS	-	-	LbrM02_V2.0720	Tb09.211.5015	-
Hypothetical repeat protein	-	-	LbrM03_V2.0960	-	-
31-0-demethyl-FK506 methyltransferase	LmjF04.1165	LinJ04.1185	LbrM04_V2.1180	-	-
Hypothetical protein	-	-	LbrM05_V2.0960	-	-
Hypothetical protein	LmjF06.1195	LinJ06.1235	LbrM06_V2.1180	-	-
Hypothetical protein	LmjF07.0155	-	LbrM07_V2.0155	Tb927.8.1790	503565.29
Hypothetical protein	LmjF10.0930	LinJ10.1280	LbrM10_V2.0930	Tb927.8.4360	50,686,570
Hypothetical protein	LmjF11.0280	LinJ11.0280	LbrM19_V2.0040	-	-
Argonaute	LmjF11.0570	LinJ11.0500	LbrM11_V2.0360	Tb10.406.0020	-
Hypothetical protein	-	-	LbrM11_V2.1100	-	-
Hypothetical protein	LmjF11.1365	LinJ11.1345	LbrM11_V2.1130	-	-
Hypothetical protein	-	-	LbrM13_V2.1570	-	-
Hypothetical protein	LmjF14.0470	LinJ14.0470	LbrM14_V2.0480	-	30,885,740
Hypothetical protein repeat	-	-	LbrM14_V2.0940	-	-
Hypothetical protein	LmjF22.0105	LmjF22.0110	LbrM22_V2.0105	-	-
Hypothetical protein	LmjF22.0130	-	LbrM22_V2.0130	-	511,807,130
Hypothetical protein	-	-	LbrM22_V2.0810	-	-
Hypothetical protein	-	-	LbrM22_V2.1380	-	-
RNase III domain gene	-	-	LbrM23_V2.0390	Tb927.8.2370	-
Hypothetical protein	-	-	LbrM23_V2.0550	Tb927.8.2470	51,442,764
Hypothetical protein	LmjF23.0845	LinJ23.0995	LbrM23_V2.0950	-	-
Hypothetical protein	-	-	LbrM24_V2.1430	-	-
Hypothetical protein	LmjF22.0150	-	LbrM25_V2.0540	-	-
Rnase III gene	-	-	LbrM25_V2.1020	Tb927.3.1230	-
Hypothetical protein	-	-	LbrM26_V2.0010	-	-
Adenine	-	-	LbrM26_V2.0120*	Tb927.7.1790	507,519,150
phosphoribosyl- transferas	LmjF26.0520	LinJ26.0500	LbrM26_V2.0530	Tb927.7.1420	50,357,910
Hypothetical protein	-	-	LbrM26_V2.0810*	-	-
Glutathione peroxidase Hypothetical protein	LmjF27.1125	LinJ27.0720	LbrM27_V2.1220	-	-
Hypothetical protein	-	-	LbrM27_V2.2790	-	-
Hypothetical protein	LmjF28.0915	LinJ28.0955	LbrM28_V2.0980	Tb11.02.5680	50,881,380
Hypothetical protein	-	-	LbrM28_V2.1240	-	-
Tyrosine/dopa decarboxylase	LmjF30.2500	LinJ30.2830	LbrM30_V2.2460	-	-
Hypothetical protein	-	-	LbrM30_V2.2580	-	-
Hypothetical protein	-	-	LbrM31_V2.2370	-	-
Hypothetical protein	-	-	LbrM32_V2.0080	-	-
Hypothetical protein	-	-	LbrM32_V2.3800	Tb11.01.8270	50,408,570
Hypothetical protein	LmjF33.0180	LinJ33.0190	LbrM33_V2.0190	-	-
Hypothetical protein	-	-	LbrM33_V2.3180	-	-
Hypothetical protein	LmjF33.3215	LinJ33.3005	LbrM33_V2.3490	--	-

Beta galactofuranosyl transferase	-	-	LbrM20_V2.0480		504115
Hypothetical protein	-	-	LbrM34_V2.2010	-	-
Galactokinase	-	-	LbrM34_V2.3650	-	-
Amino acid transporter	-	-	LbrM35_V2.1500	-	-
Hypothetical protein	LmjF36.1395	LinJ36.1255	LbrM35_V2.1530	Tb10.70.1140	-
Hypothetical protein	-	-	LbrM35_V2.6620	Tb10.6k15.1610	-
Hypothetical protein	LmjF36.6785	LinJ36.0455	LbrM35_V2.7140	Tb10.6k15.2470	506,247,180
Absent in <i>L. braziliensis</i>					
Hypothetical protein	LmjF01.0730	LinJ01.0740	-	Tb09.160.3240	510,155,140
Protein kinase	LmjF01.0750	LinJ01.0760	LbrM01_V2.0720	-	511,585,160
Tagatose-6- phosphate kinase	LmjF02.0030	LinJ02.0030	LbrM02_V2.0030	-	-
Aminopeptidase P1	LmjF02.0040	LinJ02.0010	LbrM02_V2.0060	Tb927.3.2090	50,399,120
Hypothetical protein	LmjF04.0430	LinJ04.0420	LmjF04_V2.0470	-	50,385,550
Hypothetical protein	LmjF04.0540	LinJ04.0540	LbrM04_V2.0590	-	504,033,190
Hypothetical protein	LmjF04.0600	LinJ04.0600	LbrM04_V2.0690	-	-
Hypothetical protein	LmjF04.0900	LinJ04.0920	LbrM04_V2.0690	-	-
Viscerotropic gene	LmjF05.0240	LinJ05.0340	LbrM05_V2.0230	-	503,583,100
Pteridine transporter	LmjF06.0310	LinJ06.0310	LmjF06_V2.0290	Tb927.8.3620	50,663,390
Hypothetical protein	LmjF04.0430	LinJ04.0420	LmjF04_V2.0470	-	50,385,550
Hypothetical protein	LmjF06.0630	LinJ06.0650	-	-	-
Flavoprotein subunit	LmjF07.0800	LinJ07.0870	LbrM07_V2.0880	-	-
protein	LmjF07.0910	LinJ07.0960	LbrM07_V2.0965	Tb08.12016.440	50,373,530
Hypothetical protein	LmjF12.0160	LinJ12.0140	-	Tb927.6.1140	507,603,150
Hypothetical protein	LmjF13.1390	LinJ13.1320	LbrM13_V2.1160	-	-
carboxypeptidase	LmjF14.0180	LinJ14.0180	LbrM14_V2.0180	-	-
Hypothetical protein	LmjF14.0270	LinJ14.0270	LbrM14_V2.0270	-	50,660,530
Fatty acid elongase	LmjF14.0700	LinJ14.0730	LbrM14_V2.0680	Tb927.7.4160	50,666,110
Guanine nucleotide-binding protein Kinesin	LmjF14.0760 LmjF14.0810	LinJ14.0800 LinJ14.0850	LbrM14_V2.0740 LbrM14_V2.0785	- Tb927.7.4110	510989.30 -
Hypothetical protein	LmjF14.1095	LinJ14.1150	LbrM14_V2.1090	-	-
Hypothetical protein	LmjF16.0880	LinJ16.0910	LbrM16_V2.0885	Tb08.26E13.110	50,990,130
Hypothetical protein	LmjF16.1640	LinJ16.1730	LbrM16_V2.1700	-	-
Transporter (sugar)	LmjF18.0040	LinJ18.0040	LbrM18_V2.0050	Tb10.61.2747	507,993,310
Hypothetical protein	LmjF18.1650	LinJ18.1660	LbrM18_V2.1690	Tb10.389.1350	50,778,790
Glycerol uptake	LmjF19.1347	LinJ19.1260	LbrM19_V2.1570	Tb10.61.0380	51,135,540
protein	LmjF20.0460	LinJ20.0760	LbrM20_V2.4700	Tb927.1.1140	
Hypothetical protein	LmjF20.1480	LinJ20.1530	LbrM20_V2.5730	Tb927.1.3300	51,063,650
Hypothetical protein	LmjF21.0450	LinJ21.0350	LbrM20_V2.0510	-	-
Hypothetical protein	LmjF21.0891	LinJ30.4480	LbrM21_V2.1120	-	-
Hypothetical protein	LmjF21.1230	LinJ21.0900	LbrM21_V2.1430	Tb10.70.7300	506855
Hypothetical protein	LmjF21.1570	LinJ21.1250	LbrM21_V2.1800	-	-
Hypothetical protein	LmjF22.0190	LinJ22.0180	LbrM22_V2.0190	Tb927.7.2180	511,807,210

Hypothetical protein	LmjF22.0250	LinJ22.0230	LbrM22_V2.0240	-	-
Hypothetical protein	LmjF22.1510	LinJ22.1330	LbrM22_V2.1410	-	-
Argininosuccinate synthase Oxoreductase	LmjF23.0260 LmjF23.0670	LinJ23.0300 LinJ23.0810	LbrM23_V2.0290 LbrM23_V2.0770	- -	- 503831.20
Hypothetical protein	LmjF23.0790	LinJ23.0930	LbrM23_V2.0880	Tb927.8.2710	50,415,139
HASPA	LmjF23.1040	LinJ23.1160	-	-	-
SHERP	LmjF23.1082 LmjF23.1088 LmjF23.1050	LinJ23.1200 LinJ23.1170	-	-	-
HASPB	LmjF23.1080 LmjF23.1086 LmjF23.1060	LinJ23.1190 LinJ23.1180	-	-	-
Transcription elongation factor	LmjF23.1070 LmjF24.0200	LinJ24. (not assembled)**	LbrM24_V2.0190	-	507,669,104
Hypothetical protein Multi-pass transmembrane protein	LmjF24.0220 LmjF24.0700	LinJ24. (not assembled)** LinJ24.0350	LbrM24_V2.0210 LbrM24_V2.0710	- Tb11.02.3050	- 503789.20
Hypothetical protein	LmjF26.0990	LinJ26.0950	LbrM26_V2.1000	Tb927.7.950	50,346,320
Hypothetical protein	LmjF27.0020	LinJ27.0020	LbrM27_V2.0030	Tb927.5.3540	511,649,110
Hypothetical protein	LmjF27.0960*	LinJ27.0470	LbrM27_V2.1040	Tb11.55.0022	50,636,750
Hypothetical protein	LmjF27.1740	LinJ27.1245	-	-	-
Hypothetical protein	LmjF27.1770	LinJ27.1270	LbrM27_V2.1890	-	504,153,120
Adenylate cyclase	LmjF28.0090	LinJ28.0080	LbrM28_V2.0100	-	-
Hypothetical protein	LmjF28.0150	LinJ28.0150	LbrM28_V2.0160	Tb11.02.5200	50,677,990
Hypothetical protein	LmjF28.0430	LinJ28.0320	LbrM28_V2.0440	-	-
Hypothetical protein	LmjF28.0840	LinJ28.0880	LbrM28_V2.0910	Tb11.02.5740	50,693,343
Hypothetical protein	LmjF28.2450	LinJ28.2570	LbrM28_V2.2660	Tb11.01.3590	50,602,520
D-lactate dehydrogenase Lipase	LmjF29.0280 LmjF29.1260	LinJ28.0440 LinJ29.1500	LbrM29_V2.0540 LmjF29_V2.1340	- Tb927.3.3860	- 50,402,921
Hypothetical protein	LmjF29.2880	LinJ29.2570	LbrM29_V2.2910	-	-
Hypothetical protein	LmjF30.1180	LinJ30.1550	LbrM30_V2.1300	Tb927.6.2680	50,996,550
Hypothetical protein	LmjF30.1190	LinJ30.1560	LbrM30_V2.1310	-	-
Hypothetical protein	LmjF30.1985	LinJ30.2320	LbrM30_V2.1930	Tb927.6.3270	51,150,724
Hypothetical protein	LmjF30.2010	LinJ30.2350	LbrM30_V2.1960	-	-
Phospholipase c	LmjF30.2950	LinJ30.3320	LbrM30_V2.2930	-	-
Triacylglycerol lipase	LmjF31.0830	LinJ31.0860	LbrM31_V2.1010	-	-
Hypothetical protein	LmjF31.1040	LinJ31.1390	LbrM31_V2.1240	Tb927.4.4230	51,000,730
Hypothetical protein	LmjF31.1330	LinJ31.1700	LbrM31_V2.1090	-	-
Hypothetical protein	LmjF31.2240	LinJ31.2680	-	-	-
3' nucleotidase/ nuclease Hypothetical protein	LmjF31.2300 LmjF31.3000	LinJ31.2980 LinJ31.3720	LbrM31_V2.2570 LbrM31_V2.3370	- -	- -
Hypothetical protein	LmjF32.0280	LinJ32.0650	LbrM32_V2.0350	Tb10.61.2290	-
Hypothetical protein	LmjF32.1540	LinJ32.1940	-	-	-
Hypothetical protein	LmjF32.1880	LinJ32.2300	LbrM32_V2.2070	-	-
Hypothetical protein	LmjF32.3410	LinJ32.3970	LbrM32_V2.3710	Tb11.01.8390	50,672,530
Hypothetical protein	LmjF32.3420	LinJ32.3980	-	Tb11.01.8380	50,672,540

Hypothetical protein	LmjF32.3690	LinJ32.4200	LbrM32_V2.3930	Tb11.01.8190	511,529,150
Hypothetical protein	LmjF33.1000	LinJ33.1020	LbrM33_V2.1190	-	-
Hypothetical protein	LmjF33.2420	LinJ30.0580	LbrM33_V2.2695	-	-
Hypothetical protein	LmjF33.2840	LinJ33.2640	LbrM33_V2.3130	-	-
Hypothetical protein	LmjF34.0705	LinJ34.4270	LbrM20_V2.0635	Tb927.4.3710,	50,412,910
aminophospholipid translocase Hypothetical protein	LmjF34.3220 LmjF35.0450	LinJ34.2740 LinJ35.0540	LbrM20_V2.2800 LbrM34_V2.0470	Tb927.4.1510 -	511003.10 -
Hypothetical protein	LmjF35.0780	LinJ35.0880	LbrM34_V2.0770	-	-
Hypothetical protein	LmjF35.0920	LinJ35.1020	LbrM34_V2.0910	Tb927.5.710	51,145,920
Ribosomal protein L32 like protein Hypothetical protein	LmjF35.1960 LmjF35.3580	LinJ35.2110 LinJ35.3680	LbrM34_V2.1870 LbrM34_V2.3540	- Tb927.5.2430	- 507087.60
Hypothetical protein	LmjF35.3650	LinJ35.3750	LbrM34_V2.3620	Tb09.211.2800	50,700,970
Hypothetical protein	LmjF35.4490	LinJ35.4610	LbrM34_V2.4470	Tb09.211.1670	50,955,190
L-ribulokinase	LmjF36.0060	LinJ36.2610	LbrM35_V2.0100	-	50,975,749
Trypanothione	LmjF36.4300	LinJ36.6840	-	-	504,147,280
synthetase-like protein Hypothetical protein	LmjF36.4310	LinJ36.6830	LbrM35_V2.4550	Tb106k15.0660	504,147,260
Hypothetical protein Carboxypeptidase Absent in <i>L. major</i> Hypothetical protein Hypothetical protein Hypothetical protein cyclopropane-fatty-	LmjF36.4600 LmjF36.6260 LmjF02.0100 LmjF04.0350 LmjF04.0465 -	LinJ36.5210 LinJ36.0120 LinJ02.0080 LinJ04.0350 LinJ04.0460 LinJ08.0560	LbrM35_V2.4850 LbrM35_V2.6580 LbrM02_V2.0110 LbrM04_V2.0380 LbrM04_V2.0510 LbrM08_V2.0590	Tb10.6k15.0340 - - - - -	503847.50 - 411427.10 - - -
acyl-phospholipid synthase, putative					
Pteridine transporter ft5 Hypothetical protein	LmjF10.0350 LmjF10.0425	LinJ10.0660 LinJ10.0720	LbrM10_V2.0370 LbrM10_V2.0430	- -	- -
Hypothetical protein	LmjF11.0475	LinJ11.0480	LbrM11_V2.0160	-	-
Hypothetical protein	LmjF15.0595	LinJ15.0620	LbrM15_V2.0630	-	
Hypothetical protein	-	LinJ15.1280	LbrM15_V2.1300	-	
Flagellar calcium binding protein	LmjF16.0910	LinJ16.0950	LbrM16_V2.0920	Tb08.5H5.30	
Flagellar calcium binding protein Hypothetical protein	LMJF16.0920 LmjF16.1385	LinJ16.0960 LinJ16.1450	LbrM16_V2.0930 LbrM16_V2.1440	Tb08.5H5.50 Tb927.8.5050	
Phosphate repressible	-	LinJ20.0040	LbrM10_V2.0990	-	
phosphate permease Hypothetical protein	LmjF20.0355	LinJ20.0470	LbrM20_V2.4590	-	
Hypothetical protein	LmjF20.1475	LinJ20.1520	LbrM20_V2.5720	-	-
phosphoinositide-	LmjF22.1680	LinJ22.1500	LbrM22_V2.1590	Tb11.02.3780	504,149,160
specific phosphatase C Hypothetical protein	LmjF23.0400	LinJ23.0440	LbrM23_V2.0440	-	-
Hypothetical protein	-	LinJ23.0980	LbrM23_V2.0930	-	-
Hypothetical protein	LmjF27.0745	LinJ31.1010	LbrM27_V2.0830	Tb11.47.0016	

Hypothetical protein	LmjF27.0875	LinJ27.0490	LbrM27_V2.0950	Tb11.47.0005	51,079,720
Hypothetical protein	LmjF27.1470	LinJ27.2350	LbrM27_V2.1610	Tb11.18.0012	50,343,310
eukaryotic translation release factor	LmjF27.1710	LinJ27.1220	LbrM27_V2.1850	Tb11.22.0012	506,127,110
Dynein heavy chain	LmjF27.2590	LinJ27.1940	LbrM27_V2.2710	Tb927.2.5270	50,523,110
Hypothetical protein	LmjF28.1165	LinJ28.1220	LbrM28_V2.1260	-	-
Hypothetical protein	LmjF30.1220	LinJ30.1590	LbrM30_V2.1340	-	50,996,520
Hypothetical protein	LmjF31.1165	LinJ31.1530	LbrM31_V2.1370	-	-
helicase	LmjF32.1590	LinJ32.1990	LbrM32_V2.1760	Tb11.01.6420	50,367,720
Hypothetical protein	LmjF32.1630	LinJ32.2020	LbrM32_V2.1790	Tb11.01.6450	506,221,129
Hypothetical protein	LmjF33.0115	LinJ33.0120	LbrM33_V2.0120	-	-
Hypothetical protein	LmjF33.3110	LinJ33.2910	LbrM33_V2.3400	Tb927.2.5820	51,118,990
Histone H4	LmjF35.0015	LinJ35.0020	LbrM34_V2.0020	-	-
Hypothetical protein	LmjF35.2725	LinJ35.2560	LbrM34_V2.2620	-	-
Absent in <i>L. infantum</i>					
lysophospholipase	LmjF24.1840	LinJ10.0030	LbrM24_V2.1910	Tb927.8.6390	50,679,770
Hypothetical protein	LmjF27.0870	LinJ27.0480	LbrM27_V2.0940	Tb11.47.0006	51,079,730
Hypothetical protein	LmjF29.2120	-	LbrM29_V2.2100	Tb927.3.4640	510,329,120
Polynucleotide kinase 3'-phosphatase	LmjF30.0190	-	LbrM30_V2.0190	Tb927.6.1580	505,807,190
Hypothetical protein	LmjF32.1790	LinJ32.2190	LbrM32_V2.1960	Tb11.01.6630	50,844,560
Hypothetical protein	LmjF32.1950	LinJ32.2430	LbrM32_V2.2140	Tb11.01.6790	50,844,310
Pseudogene or absent in all three species					
Zinc binding	LmjF03.0290	LinJ03.0140	-	-	-
dehydrogenase	LmjF06.0400	LinJ06.0400	LbrM06_V2.0380	-	-
Fructose biphosphate					
aldolase Hypothetical protein*	LmjF07.0745	LinJ07.0810	LbrM07_V2.0835	-	-
ATPase subunit 9	LmjF07.1105	LinJ07.1180	LbrM07_V2.1180	-	-
Fatty acid elongase	LmjF14.0665	LinJ14.0665	LbrM14_V2.0655	-	-
Hypothetical protein	LmjF14.0870	LinJ14.0910	LbrM14_V2.0840	Tb927.7.4050	50,698,330
CMP-sialic acid	LmjF20.1055	LinF20.1080	-	-	-
transporter Hypothetical protein	LmjF21.0795	LinJ21.0715	LbrM21_V2.0880	-	-
carboxypeptidase	LmjF27.1245	LinJ27.0835	LbrM27_V2.1350	-	-
Phosphoglycerate	LmjF27.1720	LinJ27.1230	-	Tb11.02.0010	50,612,520
kinase	-	-	LbrM28_V2.0580	-	-
Hypothetical protein	LmjF28.2965	LinJ28.3115	LbrM28_V2.3180	Tb11.01.2930	-
Hypothetical protein	-	-	LbrM29_V2.0740	-	506,989,150
Hypothetical protein	-	-	LbrM29_V2.2060	see below	-
Epsilon-adaptin	LmjF30.1545	LinJ30.1900	LbrM30_V2.1600	-	-
Glutamyl carboxypeptidase	LmjF31.2660	LinJ31.3350	LbrM31_V2.3010	Tb927.8.1910	5107657.10
Glutamyl	LmjF31.2785	LinJ31.3490	LbrM31_V2.3140	-	-
carboxypeptidase Mu-adaptin 4	LmjF31.3035	LinJ31.3760	LbrM31_V2.3420	Tb927.4.7050	50,507,330
Clathrin assembly sigma-adaptin	LmjF32.0095	LinJ32.0455	LbrM32_V2.0150	Tb10.61.2530	506,525,104

MAP Kinase	LmjF32.0795	LinJ32.1195	-	Tb11.01.5650	511,727,230
Hypothetical protein	LmjF32.2405	LinJ32.2915	LbrM32_V2.2640	-	-
Hypothetical protein	LmjF32.2998	LinJ32.3540	LbrM32_V2.3280	-	508,707,300
Oxidoreductase	LmjF33.1770	LinJ33.2120	LbrM33_V2.2030	-	-
Hypothetical protein	LmjF33.3060	LinJ33.2860	LbrM33_V2.3340	-	50,661,930
Hypothetical protein	LmjF34.3952	LinJ34.3510	LbrM20_V2.3570	-	-
Hypothetical protein	LmjF35.0480	LinJ35.0570	LbrM34_V2.0500	-	-
Hypothetical protein	LmjF35.4730	LinJ35.4850	LbrM34_V2.4700	-	-
Hypothetical protein	LmjF36.0325	LinJ36.2335	LbrM35_V2.0410	-	-
Beta adaptin	LmjF36.5595	LinJ36.6085	LbrM35_V2.5870	Tb10.6k15.1180	50,979,910
* duplication followed by differentiation.					
** Not assembled means that this sequence is not in the assembled contigs in GeneDB but is present in unassembled reads and small contigs that remain.					
a putative orthologues not found in <i>T. brucei</i> or <i>T. cruzi</i> but sequence simialrity to a					
shotgun sequence of <i>T. vivax</i> (Tviv1634d12.p1k_2) available from http://www.sanger.ac.uk/Projects/T_vivax					
Identifiers in red represent pseudogenes, those in black are considered as complete coding genes. "Degenerate" indicates pseudogenes that were too degenerate in nature to reconstitute conceptual protein coding regions. The table also contains a number of genes that appear to be pseudogenes in all three species.					

Tabela suplementar 2. *Oligonucleotídeos iniciadores e regiões genômicas baseadas na literatura.*

	Alvos	Oligonucleotídeos iniciadores		pB	Referências
1	kinetoplasto*	GGGGTTGGTGAAAATAGGG	TTTGAACGGGATTTCTG	720	Aransay et al., 2000
2	heat shock protein 70*	GGACGCCGGCACGATTKCT	CCTGGTTGTTGTTTCAGCCACTC	1245	Auwerda et al., 2013
3	heat shock protein 70*	GGACGCCGGCACGATTKCT	CGAAGAAGTCCGATACGAGGGA	552	Auwerda et al., 2013
4	heat shock protein 70*	GTGCACGACGTGGTGCTGGTG	CCTGGTTGTTGTTTCAGCCACTC	723	Auwerda et al., 2013
5	kinetoplasto*	CTRGGGGTTGGTGAAAATAG	TWTGAACGGGRTTCTG	900 (variável 1 região)	Berzuna-Cruz et al., 2015
6	18S ribossomal*	GCTGTAGGTGAACCTGCAGCAGCTGGA TCATT	CCCCTCTCCTCCTCCCC	1550 (tamanho do gene)	Berzuna-Cruz et al., 2015
7	isocitrate dehydrogenase	GTGCAACGTTGGCATCAAGT	ATGGGATTGTGCCACTGAGG	186	Boite et al., 2012/Marlow et al., 2014
8	isocitrate dehydrogenase	TGCAACGTTGGCATCAAGTG	TACCGTTTGGGCTCTTCCAC	94	Boite et al., 2012/Marlow et al., 2014
9	isocitrate dehydrogenase	GGAAAAGCTGATCCTGCCCT	GTACCGTTTGGGCTCTTCCA	213	Boite et al., 2012/Marlow et al., 2014
10	isocitrate dehydrogenase	AGGAAAAGCTGATCCTGCCCT	TGGTACCGTTTGGGCTCTTC	216	Boite et al., 2012/Marlow et al., 2014
11	isocitrate dehydrogenase	CCTCAGTGGCACAATCCCAT	AGTGGGTACTTGCGCATCAA	251	Boite et al., 2012/Marlow et al., 2014
12	G6PD*	GTCGGTTATCCTATTCGGGTC	ATCACAATGATGGTCAACGCAC	333	Castilho et al., 2003
13	G6PD*	TACTCGCCATGTCGGAGG	ATCACAATGATGGTCAACGCAC	333	Castilho et al., 2003
14	kinetoplasto*	GTTAGCCGATGGTGGTCTTG	CACCAATTTTCCGATTTTG	447	Cortes et al., 2004
15	ITS-5.8S	TATAGCGCACAACCGCGTAT	ATGCACGGGATGACACAAT	257	El Tai et al., 2000
16	ITS-5.8S	TCCTCAATGTTCTGGGCGAC	ATGTGATTCCGGGGTTTCCC	108	El Tai et al., 2000
17	ITS-5.8S	ATAGACCCACTTGGGACCGA	GGTGAACTTTCGGGCGGATA	186	El Tai et al., 2000
18	ITS-5.8S	TCAACGGACGAGATCCAAGC	GACATTGTAGTGCGCGTGTC	207	El Tai et al., 2000
19	ITS-5.8S	TATCCGCCCGAAAGTTCACC	TCCATCGCGACACGTTATGT	356	El Tai et al., 2000
20	ITS-1*	CTGGATCATTTTCCGATG	TGATACCACTTATCGCACTT	350	El Tai et al., 2000

21	ITS-1	AACGCACCGCCTATACAA	TCATCCATCGGACACGTTA	138	El Tai et al., 2000/Kuhls et al., 2005/Berzunza-Cruz et al., 2015
22	ITS-1	AAAGGCCGATCGACGTTGTA	GAAGCCAAGTCATCCATCGC	168	El Tai et al., 2000/Kuhls et al., 2005/Berzunza-Cruz et al., 2015
23	ITS-1	GTTTCGGTTTTTGGCGGGAG	CGACACGTTATGTGAGCCGT	71	El Tai et al., 2000/Kuhls et al., 2005/Berzunza-Cruz et al., 2015
24	ITS-1	GGCTTTCCACATACACAGC	TTGTGTATAGGCGGTGCGTT	91	El Tai et al., 2000/Kuhls et al., 2005/Berzunza-Cruz et al., 2015
25	ITS-1	ACAACCTCGGGGAGGCTTATTC	TTGTGTATAGGCGGTGCGT	130	El Tai et al., 2000/Kuhls et al., 2005/Berzunza-Cruz et al., 2015
26	miniexon gene*	GGGAATTCAATATAGTACAGAACTG	GGGAAGCTTCTGTACTTTATTGGTA	400-500/var iável	Fernandes et al., 1994
27	kinetoplasto*	AACTTTTCTGGTCTCCGGGTAG	AAACCGGCTTAGCGGGACGC	139	Francino et al., 2006
28	kinetoplasto*	CTTTTCTGGTCCCGGGTAGG	CACCTGGCCTATTTACACCA	145	Lauchaud et al., 2002/Ferroglio et al., 2006/Gomes et al., 2007/Motazediam et al., 2008/Lima Junior et al., 2009
29	aspartate-aminotransferase-asat-	CAGGGACAAGCTCCCGATG	TCATGTCCAAAAGAAGCCGTT	160	Le Blancq e Lanham, 1984/Mauricio et al., 2006/Marco et al., 2006
30	mitochondrial malate dehydrogenase	ATGCCGAAGCTGAATGCCTA	GGTGCAACTTCACAGCAAGGT	103	Marlow et al., 2014
31	mitochondrial malate dehydrogenase	GGCCTCGGAATGGTCAATCT	TGCAACTTCACAGCAAGGTCG	258	Marlow et al., 2014
32	aspartate-aminotransferase-asat*	ACGAGCGCCGTCCGYAA	TTCCYMCATCCACCAAGC	1200 (tamanho gene)	Mauricio et al., 2006
33	glucose-6-phosphateisomerase-GPI*	GAATCCCTTTTCAAGATGAGCGATTAT	CCCCTGAGAGGCAATCACAG	1800 (tamanho gene)	Mauricio et al., 2006
34	nucleoside phosphorylase 1-NH1*	CTTGCTTACGCCGAGATAC	GAAAAAAAAGACGCTTCACACAAGC	1020 (tamanho gene)	Mauricio et al., 2006
35	nucleoside phosphorylase 2-NH2*	ACGTGGGCGAACGCAC	GCCATCTACACCTTCAGTGCCTCGGTC	1020 (tamanho gene)	Mauricio et al., 2006
36	6-phosphogluconate dehydrogenase-PGD*	GAACGAATCCCTTATTCTCYAGT	GGAACCGGTTGAGCGGC	1440 (tamanho gene)	Mauricio et al., 2006
37	aspartate-aminotransferase-asat	CTGGCAGAAGATTCAGGGACA	TCATGTCCAAAAGAAGCCGTT	173	Mauricio et al., 2006
38	aspartate-aminotransferase-asat	CAACCTCGTCATTGGTGCCTA	CATGTCCAAAAGAAGCCGTT	91	Mauricio et al., 2006
39	aspartate-aminotransferase-asat	AACCTCGTCATTGGTGCCTAC	TTCATGTCCAAAAGAAGCCGTT	92	Mauricio et al., 2006
40	aspartate-aminotransferase-asat	TGGCAGAAGATTCAGGGACA	TTCATGTCCAAAAGAAGCCGTT	173	Mauricio et al., 2006
41	glucose-6-phosphateisomerase-PGI	GTTGGGAACCTCAGACAGCA	ACGCATGGAGAATGGCTTCA	547	Mauricio et al., 2006 /Tsukayama et al., 2009

42	glucose-6-phosphateisomerase-PGI	AAATACGGCAGCGAGTCCAT	TCACGCATGGAGAATGGCTT	482	Mauricio et al., 2006 /Tsukayama et al., 2009
43	glucose-6-phosphateisomerase-PGI	AGAAATACGGCAGCGAGTCC	ACGTTGATAACCCGGTTGCT	289	Mauricio et al., 2006 /Tsukayama et al., 2009
44	glucose-6-phosphateisomerase-PGI	GGCCACTGCGACTTTCAATG	TGCTGTCTGAGGTTCCCAAC	71	Mauricio et al., 2006 /Tsukayama et al., 2009
45	glucose-6-phosphateisomerase-PGI	CTCTTGAAGTTGCCTGCGTG	TAGACTTGCCCGTTTGACCC	439	Mauricio et al., 2006 /Tsukayama et al., 2009
46	kinetoplasto*	GTGGGCGTATCTGCTGATGAC	CAAAAAGCGAGGGACTGCGGA	103	Mimori et al., 1998/Pereira et al., 2008/Lima Junior et al., 2009
47	kinetoplasto*	TGCGAGGATAAAGGGAAGAA	GTGCCCTGACTTGCATGTCTA	62	Mimori et al., 1998/Pereira et al., 2008/Lima Junior et al., 2009
48	heat shock protein 70	CAAGGGTAAGAACCTGGCGT	CGCGTCGATCTCCTTGTTCA	985	Montalvo et al., 2010/Requena et al., 2012/Marlow et al., 2014/Auwera et al., 2014
49	heat shock protein 70	ACAAGGGTAAGAACCTGGCG	ATCTCCTTGTTGAGCGTGGC	979	Montalvo et al., 2010/Requena et al., 2012/Marlow et al., 2014/Auwera et al., 2014
50	heat shock protein 70	GCCACGCTGAACAAGGAGAT	ATGATCGGGTGCATACGCT	119	Montalvo et al., 2010/Requena et al., 2012/Marlow et al., 2014/Auwera et al., 2014
51	heat shock protein 70	TTGACTTCCAGGCCACCATC	TCGATCTCCTTGTTGAGCGT	853	Montalvo et al., 2010/Requena et al., 2012/Marlow et al., 2014/Auwera et al., 2014
52	heat shock protein 70	CCCTCGTGTCGGACTTCTTC	CGATCTCCTTGTTGAGCGTG	681	Montalvo et al., 2010/Requena et al., 2012/Marlow et al., 2014/Auwera et al., 2014
53	microsatélite*	AAACGTGCAAAGGCACATC	TCTATTACCGCGCTCATGCT	79	Oddone et al., 2009
54	microsatélite*	GTGTTGCGTAAACGTCGAAA	AAAAGGCCGGTTTCAAATTC	97	Oddone et al., 2009
55	microsatélite*	TTGACGTGTACACCGCTCTT	TTCTGAGAAAGGCAACCGATA	76	Oddone et al., 2009
56	microsatélite*	CATGTAGGCATGCGGTTGTA	GCTCCTTTTCTCGTTTGAAC	100	Oddone et al., 2009
57	microsatélite*	GCTTTGCTTGACTGGAGAG	GGAGGGAAAAGGAAGTCCAG	97	Oddone et al., 2009
58	microsatélite*	CATTTGAGCTGCACGTGTCT	AGGGGAGAGTGTGTTGGTG	94	Oddone et al., 2009
59	microsatélite*	CTCCTGGAACGGCTAACAC	TGATATGAGGCACATTGAGC	115	Oddone et al., 2009
60	microsatélite*	CACACCTGCTACTGGTCTC	TCTGTTTCATGACATGCCTTT	92	Oddone et al., 2009
61	microsatélite*	GTGGGTATGCGTGTGTCTCT	ATTAAAGTTGCCACCCTCAC	100	Oddone et al., 2009

62	microsatellite*	CAACAGCAAAGCACAAAGAA	CAGCAATGCCGATAAGAAAC	91	Oddone et al., 2009
63	microsatellite*	TGCGAGTATACCTCTCCTTCA	CAACAACAACAACCGAGAGG	99	Oddone et al., 2009
64	microsatellite*	CACCTCTTGCCTGCACTT	TTTAAACGTCGGTCTGTGTG	95	Oddone et al., 2009
65	microsatellite*	GGTATGCGTGGATATGAAGC	CTCGGCATCGCAGTTTC	77	Oddone et al., 2009
66	microsatellite*	ACGTCAGCACACAAACGTC	CTTCTTCCTGCTTGCCTCT	99	Oddone et al., 2009
67	microsatellite*	GGGTGTCGAGGATGAGGT	TAGTGCCATTAGGGGCTCA	93	Oddone et al., 2009
68	flavoprotein	AAGAAGGATGCCGCTGTCAA	CCGAAGAGAGAGAAGCCTCG	86	Opperdore e Szikora, 2006/Peacock et al., 2007
69	flavoprotein	ACTACACGATGGGTGGCTTG	TCAAGCCCTTGTGCTTCTCA	399	Opperdore e Szikora, 2006/Peacock et al., 2007
70	flavoprotein	CACTACACGATGGGTGGCTT	AAGCCCTTGTGCTTCTCACC	398	Opperdore e Szikora, 2006/Peacock et al., 2007
71	flavoprotein	AGCGAGGCTTCTCTCTTTC	AAGCCACCCATCGTGTAGTG	193	Opperdore e Szikora, 2006/Peacock et al., 2007
72	flavoprotein	GAAGCGAGGCTTCTCTCTCT	CAAGCCACCCATCGTGTAGT	196	Peacock et al., 2007
73	viscerotropic-leishmaniasis-antigen	GTGCACTGATGCGCTAATGG	TCGTCGTGAGTGTGAGTGTG	534	Peacock et al., 2007
74	actin	GGAGATGACGAACAGTGCCA	CCGCGAGATGTCGGAACATA	257	
75	actin	ACGGCAACATTGTCTGTCT	TCAGCGACGACAGAATGGAG	168	Peacock et al., 2007
76	actin	CGCCAGAGCGAAAATACAGC	CAGACTCGTCGTACTIONGCTC	102	Peacock et al., 2007
77	actin	TCCATCGTCTGCGATAACGG	GCTGTATTTTCGCTCTGGCG	996	Peacock et al., 2007
78	putative-carboxypeptidase	CAACCTGCGGCACTTTCTTG	TCCAGCTTCGCAACTGTCAT	465	Peacock et al., 2007/Isaza et al., 2008
79	putative-carboxypeptidase	TCTACAGCTGCACAACCTGC	AGCGCAATCTGCTTCTCCTT	606	Peacock et al., 2007/Isaza et al., 2008
80	putative-carboxypeptidase	ACAACCTGCGGCACTTTCTT	AAGCGCAATCTGCTTCTCCT	596	Peacock et al., 2007/Isaza et al., 2008
81	putative-carboxypeptidase	CACAACCTGCGGCACTTTC	CCAGGAGGCACTTGGTGTAG	745	Peacock et al., 2007/Isaza et al., 2008
82	putative-carboxypeptidase	ATGACAGTTGCGAAGCTGGA	GGGAAAGCCTGCCAGTAG	771	Peacock et al., 2007/Isaza et al., 2008
83	argonaute-dicer	GTTCTCAGTACCCAGGTCGC	GGTATCAGCGAGACCGAAGG	644	Peacock et al., 2007/Toledo et al., 2010/Llanes et al., 2015
84	argonaute-dicer	GGTAACAGGGTCCACAAGCA	CCACATAGTCGGACCCGATG	628	Peacock et al., 2007/Toledo et al., 2010/Llanes et al., 2015

					2015
85	argonaute-dicer	TGGTAACAGGGTCCACAAGC	CTCCTCCGGGTCGTTTGA	970	Peacock et al., 2007/Toledo et al., 2010/Lanes et al., 2015
86	argonaute-dicer	GCAGGTTCTCAGTACCCAGG	TAGTTGTGGAACCGGCCTTC	284	Peacock et al., 2007/Toledo et al., 2010/Lanes et al., 2015
87	argonaute-dicer	TCCACAAGCATGACGGGATT	GTCTCTCCGGGTCGTTTT	961	Peacock et al., 2007/Toledo et al., 2010/Lanes et al., 2015
88	small subunit RNA*	TACTGGGCGTCAGAG	GGGTGTCATCGTTTGC	153	Prina et al., 2007
89	DNA polar	TGATTGCGACTGACCAAGC	TCAGGGGTAATGAACTTGAG	133	Prina et al., 2007
90	G6PD*	GACGAGAAGGTGTGCGG	GCGGTCGTTGTTGATG	180	Prina et al., 2007
91	hexokinase*	TCATCCCGAAGAGTGC	AGAAGGTGAACCCAG	132	Prina et al., 2007
92	kinetoplasto*	CCTATTTTACACCAACCCCCAGT	GGGTAGGGGCGTTCTGCGAAA	120	Prina et al., 2007
93	glucose-6-phosphate-1-dehydrogenase-G6PD	TCGGCTATGCTCGCACTAAG	GTAGGGGTGGACCTCATCCT	956	Prina et al., 2007 /Zemanova et al., 2007/Boité et al., 2012/ Jara et al., 2013
94	glucose-6-phosphate-1-dehydrogenase-G6PD	CGACAAGGTGGACGACTTCA	CGAATACCGATGGCGGAAGA	122	Prina et al., 2007 /Zemanova et al., 2007/Boité et al., 2012/ Jara et al., 2013
95	glucose-6-phosphate-1-dehydrogenase-G6PD	AGGATGAGGTCCACCCCTAC	AGCTCGTCCTTACGCACAAA	256	Prina et al., 2007 /Zemanova et al., 2007/Boité et al., 2012/ Jara et al., 2013
96	glucose-6-phosphate-1-dehydrogenase-G6PD	TCTTCCGCCATCGGTATTCG	TCTCCTCCCGAGGTAGTGG	203	Prina et al., 2007 /Zemanova et al., 2007/Boité et al., 2012/ Jara et al., 2013
97	glucose-6-phosphate-1-dehydrogenase-G6PD	ACGACAAGGTGGACGACTTC	CTTCAGCACCGACACCTTCT	566	Prina et al., 2007 /Zemanova et al., 2007/Boité et al., 2012/ Jara et al., 2013
98	kinetoplasto*	AAC TTTCTGGTCCTCCGGGTAG	ACCCCCAGTTCCCGCC	123	Rey et al., 2002
99	metallo-peptidase	AGAAGGCACTTCTGTGGACG	CCAACAAC TGTCTTGGAGCC	151	Reyes-Uribe et al., 2012/Peacock et al., 2007
100	metallo-peptidase	GGAGGAGAAGTACCCGAGT	CTGAACGATGTTTGCCACCG	190	Reyes-Uribe et al., 2012/Peacock et al., 2007
101	metallo-peptidase	CGGTGGCAAACATCGTTCAG	GGTGCCGACATCCTTCAAGA	449	Reyes-Uribe et al., 2012/Peacock et al., 2007
102	metallo-peptidase	AGAAGATGCAAGAGGCCACC	CCACAGAAGTGCCCTTCTCCA	161	Reyes-Uribe et al., 2012/Peacock et al., 2007
103	metallo-peptidase	TGGAGAAGGCACTTCTGTGG	GACAGCGTACGCGTAGAAGA	458	Reyes-Uribe et al., 2012/Peacock et al., 2007
104	chaperonin HSP60-mitochondrialprecursor	GGTCGCTCGCTTAATACGGA	CGCAGGGAGGCTATAAGCAA	743	Rey-Ladino et al., 1997

105	chaperonin HSP60-mitochondrialprecursor	TTGCTTATAGCCTCCCTGCG	TCATCCTCGATGCGCTTCTC	365	Rey-Ladino et al., 1997
106	chaperonin HSP60-mitochondrialprecursor	AGGGTCGCTCGCTTAATACG	GGGAGGCTATAAGCAACCCG	741	Rey-Ladino et al., 1997
107	chaperonin HSP60-mitochondrialprecursor	GTGTCGAAAGGTGACCGTCT	ATAAGCAACCCGGTACCACC	306	Rey-Ladino et al., 1997
108	chaperonin HSP60-mitochondrialprecursor	GGTGGTACCGGGTTGCTTAT	CATCCTCGATGCGCTTCTCT	376	Rey-Ladino et al., 1997
109	kinetoplasto	TACACCAACCCCAAGTTTCC	ATGGGTGCAGAAATCCCGTT	80	Rogers et al., 1990
110	kinetoplasto	TTACACCAACCCCAAGTTTCC	AATGGGTGCAGAAATCCCGT	82	Rogers et al., 1990
111	kinetoplasto	TCCGAGCCCAAAAATGGCA	GCGTTCTGCGAAAAGCGAAA	77	Rogers et al., 1990
112	kinetoplasto	CACCAACCCCAAGTTTCCC	TGGGTGCAGAAATCCCGTTC	77	Rogers et al., 1990
113	kinetoplasto	CTCCGAGCCCAAAAATGGC	GCGTTCTGCGAAAAGCGAA	78	Rogers et al., 1990
114	kinetoplasto*	GTGGGGGAGGGGCGTTCT	ATTTTACACCAACCCCAAGTT	120	Rogers et al., 1990
115	mannose phosphate isomerase*	CCCTTTGGTTGTTCGGT	TCATACGCATAGGAGCA	1464	Tsukayama et al., 2009
116	mannose phosphate isomerase*	TTTACAAGCAGTACCCGATGA	GGCTGCATACCAGGTTCTCATC	1464	Tsukayama et al., 2009
117	malate dehydrogenase*	TCACAACCGCAACTACGA	CTACTCACGATAACGGCAGA	1156	Tsukayama et al., 2009
118	malate dehydrogenase*	GCCGCCAACGACAGGGAC	GCCGCCAACGACAGGGAC	1156	Tsukayama et al., 2009
119	glucose-6-phosphate isomerase*	AATGTTCCTCATACCCCTCT	TTCCGTCCGTCTCCTG	2084	Tsukayama et al., 2009
120	glucose-6-phosphate isomerase*	TGGGATTGGCGGCAGCGACCTT	CGCCACAGGTACTGGTCGT	2084	Tsukayama et al., 2009
121	6-phosphogluconate dehydrogenase-PGD*	CACCAGTCCGCTCCC T	GCCTCTGTATTTACGCTTC	1728	Tsukayama et al., 2009
122	6-phosphogluconate dehydrogenase-PGD*	TTCAAGGCGTTCCTACATT	ACTCGCCAGCGTTGTGATA	1728	Tsukayama et al., 2009
123	putative-flagellar-calcium-binding	CCTGCCGAGAATTGGTGAGT	GTTGCCTGGGTCACCCCTAG	154	Tull et al., 2004/Peacock et al., 2007
124	putative-flagellar-calcium-binding	CAGGGCTGCGTGGATATTCT	TCTTCTGTGTCTGGGTCA	415	Tull et al., 2004/Peacock et al., 2007
125	putative-flagellar-calcium-binding	GACATCGTGAGAGAGCCTT	CCTGGGTACCCCTTAGTGTTT	353	Tull et al., 2004/Peacock et al., 2007
126	putative-flagellar-calcium-binding	GACACCTCCGGCAATATGCT	CTGGGTACCCCTTAGTGTTCA	196	Tull et al., 2004/Peacock et al., 2007
127	putative-flagellar-calcium-binding	TCCGGCAATATGCTGGTGAG	GTTGCCTGGGTCACCCCTTA	195	Tull et al., 2004/Peacock et al., 2007

128	SSU rDNA*	GATCCAGCTGCAGGTTACCC	GGTGTATCCGTCAACGGAC	540	Uliana et al., 1994
129	SSU rDNA*	CCAAGCTGCCCAGTAGAAT	TGCGGCGGATAAACACC	439	Uliana et al., 1994
130	kinetoplasto*	GGGGAGGGGCGTTCTGCGAA	CCGCCCTATTTTACACCAACCCC	120	Volpini et al., 2004/Andrade et al., 2006
131	malic enzyme	CAACCATCGCAAGGCAACAA	CCAGGTAGTAGTTCTCGGGC	334	Zemanová et al, 2007
132	malic enzyme	AACCATCGCAAGGCAACAAC	CCAGGTAGTAGTTCTCGGGC	333	Zemanová et al, 2007
133	malic enzyme	AAGTACCGCTGCTTCAACGA	CCAGGTAGTAGTTCTCGGGC	910	Zemanová et al, 2007
134	malic enzyme	ATTCCCGCTGTCAAATCCGT	CCAGGTAGTAGTTCTCGGGC	464	Zemanová et al, 2007
135	malic enzyme	TGCAAGCCTGAAGACTCTCG	CCAGGTAGTAGTTCTCGGGC	599	Zemanová et al, 2007
136	fumarate hydratase	GACCCGTTTGTGACCTCTT	CGTACTTGTGGCACGCTTTC	100	Zemanová et al, 2007
137	fumarate hydratase	CTGCCACTACTACGACGACC	AGATACTGGGACGGCTCCAT	294	Zemanová et al, 2007
138	fumarate hydratase	CCCTAGTGTCGAAGGTGTCG	CTTGCTGGCGTAGAGGTCAA	503	Zemanová et al, 2007
139	fumarate hydratase	CGTTCCCTAGTGTCGAAGG	GTCGTCGTAGTAGTGGCAGG	751	Zemanová et al, 2007
140	fumarate hydratase	GATTGACCTCTACGCCAGCA	GGTCGTCGTAGTAGTGGCAG	266	Zemanová et al, 2007
141	isocitrate dehydrogenase*	ATGTTCCGCCATGTTTCGGC	TTACGCGCTCATCGCCTT	1300 (tamanho gene)	Zemanová et al, 2007
142	malic enzyme*	CGCAACCGCTTCACCAATAAGGGC	CAACTCCTTCTCCAGGTAGTAGTT	1700 (tamanho gene)	Zemanová et al, 2007
143	mannose phosphate isomerase*	ATGTCTGAGCTCGTAAAGCT	CTACCTGTCGCTCAAGTC	1220 (tamanho gene)	Zemanová et al, 2007
144	glucose-6-phosphate dehydrogenase*	ATGTCGGAAGAGCAGTCT	TCACAGCTTATTCGAGGGAA	1690 (tamanho gene)	Zemanová et al, 2007

Tabela suplementar 3. Banco de dados TriTrypDB. Genomas disponíveis no banco de dados TriTrypDB dos tripanosomatídeos usados no estudo.

Organismo	NCBI taxon ID	Fonte de dados	Versão do genoma	Tamanho do genoma (Mbp)	Total genes (contagem)
<i>Crithidia fasciculata</i> strain Cf-CI		BeverleyLab	6/1/2015	32,63	9619
<i>Endotrypanum monterogeii</i> strain LV88		GeneDB	1/16/2015	32,52	8414
<i>Leishmania aethiopica</i> L147	1206056	GeneDB	12/16/2014	31,63	8826
<i>Leishmania amazonensis</i> MHOM/BR/71973/M2269		GenBank	7/25/2013	29,03	
<i>Leishmania arabica</i> strain LEM1108		GeneDB	12/16/2014	31,27	8755
<i>Leishmania braziliensis</i> MHOM/BR/75/M2903	1295825	GeneDB	12/16/2014	35,21	9379
<i>Leishmania braziliensis</i> MHOM/BR/75/M2904	420245	GeneDB	1/16/2013	32,09	8505
<i>Leishmania donovani</i> BPK282A1	981087	GeneDB	1/16/2013	32,44	8195
<i>Leishmania donovani</i> strain BHU 1220		GenBank	11/6/2013	32,41	
<i>Leishmania enriettii</i> strain LEM3045		GeneDB	12/16/2014	30,76	8831
<i>Leishmania gerbilli</i> strain LEM452		GeneDB	12/16/2014	31,29	8700
<i>Leishmania infantum</i> JPCM5	435258	GeneDB	1/16/2013	32,13	8381
<i>Leishmania major</i> strain Friedlin	347515	GeneDB	1/16/2013	32,86	9378
<i>Leishmania major</i> strain LV39c5	860569	GeneDB	1/16/2015	32,33	9101
<i>Leishmania major</i> strain SD 75,1	860570	GeneDB	1/16/2015	31,24	8931
<i>Leishmania mexicana</i> MHOM/GT/2001/U1103	929439	GeneDB	1/16/2013	32,11	9063
<i>Leishmania panamensis</i> MHOM/COL/81/L13	1295824	GeneDB	12/16/2014	31,26	8778
<i>Leishmania</i> sp, MAR LEM2494	1303197	GeneDB	12/16/2014	30,81	8581
<i>Leishmania tarentolae</i> Parrot-TarII		ULAVAL	6/22/2011	31,63	8530
<i>Leishmania tropica</i> L590	1206058	GeneDB	12/16/2014	33,02	8938
<i>Leishmania turanica</i> strain LEM423		GeneDB	12/16/2014	31,51	8717
<i>Leptomonas pyrrhocoris</i> H10		LukesLab	5/5/2015	30,38	10148
<i>Leptomonas seymouri</i> ATCC 30220		YurchenkoLab	5/6/2015	27,8	8595
<i>Trypanosoma brucei</i> Lister strain 427		GeneDB	10/20/2010	26,75	9302
<i>Trypanosoma brucei</i> TREU927	185431	GeneDB	8/28/2013	35,83	12094
<i>Trypanosoma brucei</i> gambiense DAL972	679716	GeneDB	1/16/2013	22,15	10000
<i>Trypanosoma congolense</i> IL3000	1068625	GeneDB	1/16/2013	41,37	13358
<i>Trypanosoma cruzi</i> CL Brener Esmeraldo-like		GeneDB	9/16/2014	32,53	10597
<i>Trypanosoma cruzi</i> CL Brener Non-Esmeraldo-like		GeneDB	9/16/2014	32,53	11106
<i>Trypanosoma cruzi</i> Dm28c	1416333	GenBank	11/13/2013	27,35	11398
<i>Trypanosoma cruzi</i> JR cl, 4	914063	GenBank	1/17/2013	41,48	
<i>Trypanosoma cruzi</i> Sylvio X10/1		GenBank	10/2/2012	38,59	10947
<i>Trypanosoma cruzi</i> Tula cl2	1206070	GenBank	6/26/2013	83,51	
<i>Trypanosoma cruzi</i> marinkellei strain B7		Franzen	1	38,65	10282
<i>Trypanosoma cruzi</i> strain CL Brener	353153	GeneDB	9/16/2014	36,03	3397
<i>Trypanosoma cruzi</i> strain Esmeraldo	366581	GenBank	1/17/2013	38,08	
<i>Trypanosoma evansi</i> strain STIB 805		Schnauffer	6/3/2014	25,43	10174
<i>Trypanosoma grayi</i> ANR4		FieldLab	3/21/2014	20,95	10686
<i>Trypanosoma rangeli</i> SC58	429131	GenBank	10/30/2013	14,02	7479
<i>Trypanosoma vivax</i> Y486	1055687	GeneDB	1/16/2013	47,5	12581

Tabela suplementar 4. Proteínas ortológicas de cópia única comuns à 6 espécies de *Leishmania* usadas para filogenia.

N	Protein name	Function	Size (amino acids) by species					
			L. br	L. dn	L. in	L. mj	L. mx	L. ta
1	2,3-bisphosphoglycerate-independent phosphoglycerate mutase-like	Catalysis of the reaction: 2-phospho-D-glycerate = 3-phospho-D-glycerate; this reaction does not require the cofactor 2,3-bisphosphoglycerate	834	833	833	833	833	833
2	25 Kd elongation factor 1-beta, putative	Functions in chain elongation during polypeptide synthesis at the ribosome.	77	77	77	77	77	77
3	2-aminoethylphosphonate:pyruvateaminotransferase e-like protein	Catalysis of the reaction: (2-aminoethyl)phosphonate + pyruvate = L-alanine + phosphonoacetaldehyde.	381	381	381	435	366	381
4	2-hydroxy-3-oxopropionate reductase, putative	Catalysis of the reaction: (R)-glycerate + NADP+ = 2-hydroxy-3-oxopropionate + NADPH + H+.	299	299	299	299	299	299
5	2-oxoglutarate dehydrogenase, E2 component, dihydrolipoamide succinyltransferase, putative	Catalysis of the reaction: succinyl-CoA + dihydrolipoamide = CoA + S-succinyl-dihydrolipoamide	391	391	389	389	389	389
6	2-oxoglutarate dehydrogenase, e3 component, lipoamidedehydrogenase-like protein	Interacting selectively and non-covalently with FAD, flavin-adenine dinucleotide, the coenzyme or the prosthetic group of various flavoprotein oxidoreductase enzymes, in either the oxidized form, FAD, or the reduced form, FADH2.	518	478	478	622	542	516
7	2-oxoisovalerate dehydrogenase alpha subunit, putative	Catalysis of the reaction: 3-methyl-2-oxobutanoate + lipoamide = S-(2-methylpropanoyl)dihydrolipoamide + CO2.	479	479	479	479	479	479
8	2-oxoisovalerate dehydrogenase beta subunit, mitochondrial precursor, putative	*The branched-chain alpha-keto dehydrogenase complex catalyzes the overall conversion of alpha-keto acids to acyl-CoA and CO2. It contains multiple copies of three enzymatic components: branched-chain alpha-keto acid decarboxylase (E1), lipoamide acyltransferase (E2) and lipoamide dehydrogenase (E3).	366	366	366	366	366	366
9	3,2-trans-enoyl-CoA isomerase, mitochondrial precursor-like protein	*Able to isomerize both 3-cis and 3-trans double bonds into the 2-trans form in a range of enoyl-CoA species.	296	296	296	296	296	296
10	33 kDa inner dynein arm light chain, axonemal, putative	-	425	247	247	247	247	440
11	3-demethylubiquinone-9 3-methyltransferase, putative	Catalysis of the reaction: S-adenosyl-L-methionine + 3-demethylubiquinone-9 = S-adenosyl-L-homocysteine + ubiquinone-9.	291	321	321	321	291	295
12	3-methylcrotonoyl-CoA carboxylase beta subunit, putative	*Catalysis of the joining of two substances, or two groups within a single molecule, with the concomitant hydrolysis of the diphosphate bond in ATP or a similar triphosphate.	579	536	536	534	534	329
13	6-phosphogluconate dehydrogenase, decarboxylating, putative	Catalysis of the reaction: 6-phospho-D-gluconate + NADP+ = D-ribulose 5-phosphate + CO2 + NADPH + H+.	479	479	479	479	479	479
14	6-phosphogluconolactonase	Catalysis of the reaction: 6-O-phosphono-D-glucono-1,5-lactone + H(2)O = 6-phospho-D-gluconate + H(+).	267	267	267	267	267	267
15	8-oxoguanine DNA glycosylase, putative	Catalysis of the removal of oxidized purine bases by cleaving the N-C1' glycosidic bond between the oxidized purine and the deoxyribose sugar. The reaction involves the formation of a covalent enzyme-substrate intermediate. Release of the enzyme and free base by a beta-elimination or a beta, gamma-elimination mechanism results in the cleavage of the DNA backbone 3' of the apurinic (AP) site.	569	565	565	565	565	899
16	A-1 protein, putative	-	174	175	175	175	175	175
17	A/G-specific adenine glycosylase, putative	Any molecular function by which a gene product interacts selectively and non-covalently with DNA (deoxyribonucleic acid)	491	501	501	501	501	587
18	C-14 sterol reductase, putative	Catalysis of the reaction: NADP+ + 4,4-dimethyl-5-alpha-cholesta-8,24-dien-3-beta-ol = NADPH + H+ + 4,4-dimethyl-5-alpha-cholesta-8,14,24-trien-3-beta-ol.	437	437	437	437	437	437
19	C-5 sterol desaturase	Catalysis of an oxidation-reduction (redox) reaction, a reversible chemical reaction in which the oxidation state of an atom or atoms within a molecule is altered. One substrate acts as a hydrogen or electron donor and becomes oxidized, while the other acts as hydrogen or electron acceptor and becomes reduced.	284	274	274	274	274	262
20	C-8 sterol isomerase-like protein	Catalysis of the geometric or structural changes within one molecule. Isomerase is the systematic name for any enzyme of EC class 5.	223	223	223	223	223	223

21	Csl4p homologue	Interacting selectively and non-covalently with an RNA molecule or a portion thereof.	281	285	285	285	285	280
22	D-3-phosphoglycerate dehydrogenase-like protein	Interacting selectively and non-covalently with nicotinamide adenine dinucleotide, a coenzyme involved in many redox and biosynthetic reactions; binding may be to either the oxidized form, NAD ⁺ , or the reduced form, NADH.	408	407	407	511	407	439
23	D-tyrosyl-tRNA deacylase	Catalysis of the hydrolysis of any ester bond.	152	152	152	152	152	152
24	Dynein light chain 1, cytoplasmic, putative	*Any cellular process that depends upon or alters the microtubule cytoskeleton, that part of the cytoskeleton comprising microtubules and their associated proteins.	89	89	89	89	89	89
25	Golgi/lysosome glycoprotein	-	622	624	624	623	623	624
26	L-gulonolactone oxidase	Catalysis of the reaction: L-gulono-1,4-lactone + O ₂ = L-xylulo-hex-3-ulonolactone + H ₂ O ₂ .	498	495	495	502	502	502
27	L-ribulokinase	Catalysis of the reaction: ATP + L(or D)-ribulose = ADP + L(or D)-ribulose 5-phosphate.	600	563	563	563	563	567
28	Lsm5p	-	128	128	128	128	128	128
29	Lsm7p protein	The process of removing sections of the primary RNA transcript to remove sequences not present in the mature form of the RNA and joining the remaining sections to form the mature form of the RNA.	98	91	91	91	91	91
30	N-acetylglucosamine-6-phosphate deacetylase-like protein	Catalysis of the hydrolysis of any carbon-nitrogen bond, C-N, with the exception of peptide bonds.	432	432	432	432	432	432
31	N-acetyltransferase subunit Nat1, putative	Catalysis of the transfer of a group, e.g. a methyl group, glycosyl group, acyl group, phosphorus-containing, or other groups, from one compound (generally regarded as the donor) to another compound (generally regarded as the acceptor). Transferase is the systematic name for any enzyme of EC class 2.	711	711	711	711	711	711
32	Na/H antiporter-like protein	Catalysis of the transfer of a solute or solutes from one side of a membrane to the other according to the reaction: solute(out) + H ⁺ (in) = solute(in) + H ⁺ (out).	146 2	150 0	150 0	150 0	1499	149 9
33	Pyridoxal kinase	Catalysis of the reaction: ATP + pyridoxal = ADP + pyridoxal 5'-phosphate	302	302	302	302	302	302
34	Ran-binding protein 1	The directed movement of substances within a cell.	158	158	158	158	158	158
35	Ran-binding protein	*May act as an adapter protein to couple membrane receptors to intracellular signaling pathways. May be involved in signaling of ITGB2/LFA-1 and other integrins. Enhances HGF-MET signaling by recruiting Sos and activating the Ras pathway. Enhances dihydrotestosterone-induced transactivation activity of AR, as well as dexamethasone-induced transactivation activity of NR3C1, but not affect estrogen-induced transactivation. Stabilizes TP73 isoform Alpha, probably by inhibiting its ubiquitination, and increases its proapoptotic activity. Inhibits the kinase activity of DYRK1A and DYRK1B. Inhibits FMR1 binding to RNA (By similarity).	200	213	213	213	213	214
36	S-adenosyl-methyltransferase mraW-like protein	Catalysis of the transfer of a methyl group to an acceptor molecule.	491	489	489	489	489	490
37	T-complex protein 1, beta subunit, putative	Interacting selectively and non-covalently with ATP, adenosine 5'-triphosphate, a universally important coenzyme and enzyme regulator.	529	529	529	529	529	529
38	T-complex protein 1, eta subunit, putative	Interacting selectively and non-covalently with ATP, adenosine 5'-triphosphate, a universally important coenzyme and enzyme regulator	571	570	570	570	570	570
39	T-complex protein 1, gamma subunit, putative	*Interacting selectively and non-covalently with ATP, adenosine 5'-triphosphate, a universally important coenzyme and enzyme regulator.	551	551	551	551	551	551
40	U1 small nuclear ribonucleoprotein, putative	Interacting selectively and non-covalently with any nucleic acid	281	171	375	283	282	282
41	U2 small nuclear ribonucleoprotein 40K, putative	*Interacting selectively and non-covalently with an RNA molecule or a portion thereof.	308	308	308	308	308	307
42	U2 splicing auxiliary factor, putative	Interacting selectively and non-covalently with an RNA molecule or a portion thereof.	210	210	210	210	210	210
43	U3 small nucleolar ribonucleoprotein protein MPP10, putative	Any process involved in the conversion of a primary ribosomal RNA (rRNA) transcript into one or more mature rRNA molecules	682	675	675	677	679	401
44	U3 snoRNA-associated protein UTP11, putative	*Any process involved in the conversion of a primary ribosomal RNA (rRNA) transcript into one or more mature rRNA molecules.	360	186	186	361	361	363
45	U6 snRNA-associated Sm-like protein LSm4p	Any process involved in the conversion of one or more primary RNA transcripts into one or more mature RNA	116	197	197	189	190	116

		molecules						
46	V-type ATPase, C subunit, putative	Enables the transfer of hydrogen ions from one side of a membrane to the other./Catalysis of the hydrolysis of various bonds, e.g. C-O, C-N, C-C, phosphoric anhydride bonds, etc. Hydrolase is the systematic name for any enzyme of EC class 3.	224	224	224	224	224	224
47	X-pro, dipeptidyl-peptidase,serine peptidase, Clan SC, family S15, putative	Catalysis of the hydrolysis of N-terminal dipeptides from a polypeptide chain	686	686	686	686	686	686
48	Zn-finger domain protein, putative	Interacting selectively and non-covalently with zinc (Zn) ions.	1219	1217	1217	1221	1217	1222

N = Número das proteínas selecionadas. Nome da proteína (TritypDB) e função descrita (Uniprot), tamanho em quantidade de aminoácidos por espécie. - = ausência de informação sobre a função da proteína. L.br= *Leishmania braziliensis*; L.dn=*Leishmania donovani*; L.in=*Leishmania infantum*; L.mj= *Leishmania major*; L.mx= *Leishmania mexicana*; L.ta= *Leishmania tarentolae*.