

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

**MONTAGEM DOS GENOMAS E ANOTAÇÃO DAS**  
**PROTEÍNAS DE *Leishmania (Viannia) guyanensis* E**  
*Leishmania (Viannia) shawi*

**Flávia Regina Florencio de Athayde**

Bióloga

ARAÇATUBA - SP  
2016

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

**MONTAGEM DOS GENOMAS E ANOTAÇÃO DAS**  
**PROTEÍNAS DE *Leishmania (Viannia) guyanensis* E**  
*Leishmania (Viannia) shawi*

**Flávia Regina Florencio de Athayde**

**Orientadora: Prof<sup>a</sup>. Adjunto Cáris Maroni Nunes**

Dissertação apresentada à Faculdade de Medicina Veterinária de Araçatuba – UNESP, Campus de Araçatuba, como parte das exigências para a obtenção do título de Mestra em Ciência Animal (Medicina Veterinária Preventiva e Produção Animal).

ARAÇATUBA – SP  
2016

Catálogo na Publicação(CIP)  
Serviço de Biblioteca e Documentação – FMVA/UNESP

Athayde, Flávia Regina Florencio de

At38m

Montagem dos genomas e anotação das proteínas de *Leishmania (Viannia) guyanensis* e *Leishmania (Viannia) shawi* / Flávia Regina Florencio de Athayde.

Araçatuba: [s.n], 2016.  
53f. il.

Dissertação (Mestrado) – Universidade Estadual Paulista “Júlio de Mesquita Filho”, Faculdade de Medicina Veterinária, 2016

Orientadora: Profa. Adj. Cárís Maroni Nunes

1. Genoma 2 .Genômica. 3. Leishmaniose. 4. Saúde pública.  
5. Trypanosoma. I. T.

CDD 572.8633



UNIVERSIDADE ESTADUAL PAULISTA

Câmpus de Araçatuba

CERTIFICADO DE APROVAÇÃO

TÍTULO: Montagem dos genomas e anotação das proteínas de *Leishmania (Viannia)*  
*guyanensis* e *Leishmania (Viannia) shawi*

**AUTORA: FLÁVIA REGINA FLORENCIO DE ATHAYDE**

**ORIENTADORA: CÁRIS MARONI NUNES**

Aprovada como parte das exigências para obtenção do Título de Mestra em CIÊNCIA ANIMAL,  
área: Medicina Veterinária Preventiva e Produção Animal pela Comissão Examinadora:

Prof. Dra. CÁRIS MARONI NUNES

Departamento de Apoio, Produção e Saúde Animal / Faculdade de Medicina Veterinária de Araçatuba - Unesp

Dra. FLÁVIA LOMBARDI LOPES

Departamento de Apoio, Produção e Saúde Animal / Faculdade de Medicina Veterinária de Araçatuba - Unesp

Dra. SILVANA DE CÁSSIA PAULAN

Doutora em Ciência Animal pela Faculdade de Medicina Veterinária de Araçatuba - Unesp

Araçatuba, 09 de dezembro de 2016.

## **DADOS CURRICULARES DO AUTOR**

**Flávia Regina Florencio de Athayde** - nascida em 26 de julho de 1991, na cidade de Birigui – SP. Possui graduação em Ciências Biológicas pela Faculdade de Ciências (UNESP), campus de Bauru, em dezembro de 2012. No ano de 2013 ingressou no programa de pós-graduação em Ciência Animal da Faculdade de Medicina Veterinária de Araçatuba (FMVA) - UNESP, como aluna especial do programa de mestrado, passando a aluna regular em 2014. Em agosto de 2016 integralizou os créditos e em 10 de novembro de 2016 foi aprovada no Exame Geral de Qualificação, com o trabalho intitulado “Montagem dos genomas e anotação das proteínas de *Leishmania (Viannia) guyanensis* e *Leishmania (Viannia) shawi*”.

*“É a dedicação ao trabalho que distingue um indivíduo do outro; não acredito em talentos.”*

**Euryclides de Jesus Zerbini**

À minha mãe Marta e minha irmã Rafaela, que me apoiaram em todos os momentos da minha caminhada acadêmica.

## **AGRADECIMENTOS**

À direção da Faculdade de Medicina Veterinária de Araçatuba, o programa de Pós-Graduação e a todos os funcionários da unidade. À CAPES por conceder a bolsa de mestrado.

À Professora Adj. Cárís Maroni Nunes que aceitou me orientar nessa jornada, sendo uma professora e pesquisadora muito competente, acima de tudo uma pessoa especial que me tratou com muito respeito, educação e carinho.

Aos amigos: Fernanda Müller, Dandara Fonseca, Luís Henrique Costa, Hení Falcão e Anirene Pereira, pela amizade que se estendeu além da faculdade.

Aos meus amigos e mentores: Pier Kenji Ito e Marco Milanesi, pelas ideias, colaboração e apoio desde o início acreditando em mim.

Ao amigo Leandro Costa do Nascimento, uma amizade especial que fiz graças a essa pesquisa que me ensinou muito e teve muita paciência, toda a minha gratidão.

Aos meus amigos e colegas do laboratório de Bioquímica e Biologia molecular, pelo convívio diário e por todo apoio prestado.

À minha mãe Marta, minha irmã Rafaela, meu namorado Fernando e sua família pelo carinho, paciência e apoio em todos os momentos dessa caminhada.

À todos que me apoiaram e ajudaram nessa jornada.



## SUMÁRIO

CAPÍTULO 1.....	12
CONSIDERAÇÕES GERAIS.....	13
OBJETIVO.....	23
REFERÊNCIAS .....	24
CAPÍTULO 2.....	28
INTRODUÇÃO .....	31
MATERIAL E MÉTODOS .....	32
Amostras biológicas e extração de DNA .....	33
Bibliotecas <i>Paired-End</i> .....	33
Sequenciamento e processamento dos dados .....	34
Montagem dos genomas.....	34
Predição dos genes .....	35
Análise de domínios das proteínas .....	36
RESULTADOS .....	37
DISCUSSÃO .....	37
CONCLUSÃO.....	46
AGRADECIMENTOS.....	46
REFERÊNCIAS .....	47
APÊNDICES.....	50

**Montagem dos genomas e anotação das proteínas de *Leishmania (Viannia) guyanensis* e *Leishmania (Viannia) shawi***

**RESUMO** – Os protozoários do gênero *Leishmania*, da família Trypanosomatidae, são responsáveis por causar as leishmanioses, zoonoses de distribuição mundial que atingem países de regiões tropicais e subtropicais, representando um importante problema de saúde pública. As espécies de *Leishmania* patogênicas à humanos são classificadas em dois subgêneros, *Leishmania* e *Viannia*, sendo responsáveis pelas três formas clínicas mais comuns da doença. O subgênero *Leishmania* tem representantes no Velho e no Novo Mundo, enquanto que as espécies do subgênero *Viannia* estão presentes somente nas regiões tropicais das Américas, sendo responsáveis pela manifestação tegumentar na forma cutânea e mucocutânea. O avanço nas tecnologias de sequenciamento e montagem dos genomas de espécies de *Leishmania* têm possibilitado melhor entendimento das leishmanioses, da relação parasita-hospedeiro e até mesmo a identificação diferencial das espécies, auxiliando no entendimento das diferenças nas manifestações clínicas. Com o intuito de ampliar as informações genômicas das espécies de *Leishmania* e do subgênero *Viannia*, trabalhamos com o sequenciamento de DNA genômico e ferramentas de bioinformática para a montagem do genoma e anotação de proteínas de duas espécies, *Leishmania (Viannia) guyanensis* e *Leishmania (Viannia) shawi*.

**Palavras-Chave:** genômica, leishmaniose, saúde pública, trypanosoma

**Genome assembly and protein annotation of *Leishmania (Viannia) guyanensis*  
and *Leishmania (Viannia) shawi***

**SUMMARY** – The protozoa of the Trypanosomatidae family *Leishmania* genus, are responsible for causing leishmaniosis, zoonosis of global distribution that affects countries in tropical and subtropical regions, representing an important public health problem. The *Leishmania* species pathogenic to humans are classified into two subgenera, *Leishmania* and *Viannia*, which are responsible for three of the most common clinical manifestations of this disease. The subgenus *Leishmania* has representatives in the Old and New World, while the *Viannia* species are present only in the tropical regions of the Americas. These are responsible for the tegumentary manifestation of the cutaneous and mucocutaneous forms. The advance in the technologies of sequencing and the assembly of the genomes of *Leishmania* species have allowed to better understand the leishmaniosis, the parasite-host relationship and even the differential identification of the species thus, helping to understand the differences in the clinical manifestations. In order to expand the genomic information of the *Leishmania* species and the subgenus *Viannia*, we worked with genomic DNA sequencing and bioinformatics tools for assembling the genome and annotating proteins of two species, *Leishmania (Viannia) guyanensis* and *Leishmania (Viannia) shawi*.

**Keywords:** Genomics, leishmaniosis, public health, trypanosoma

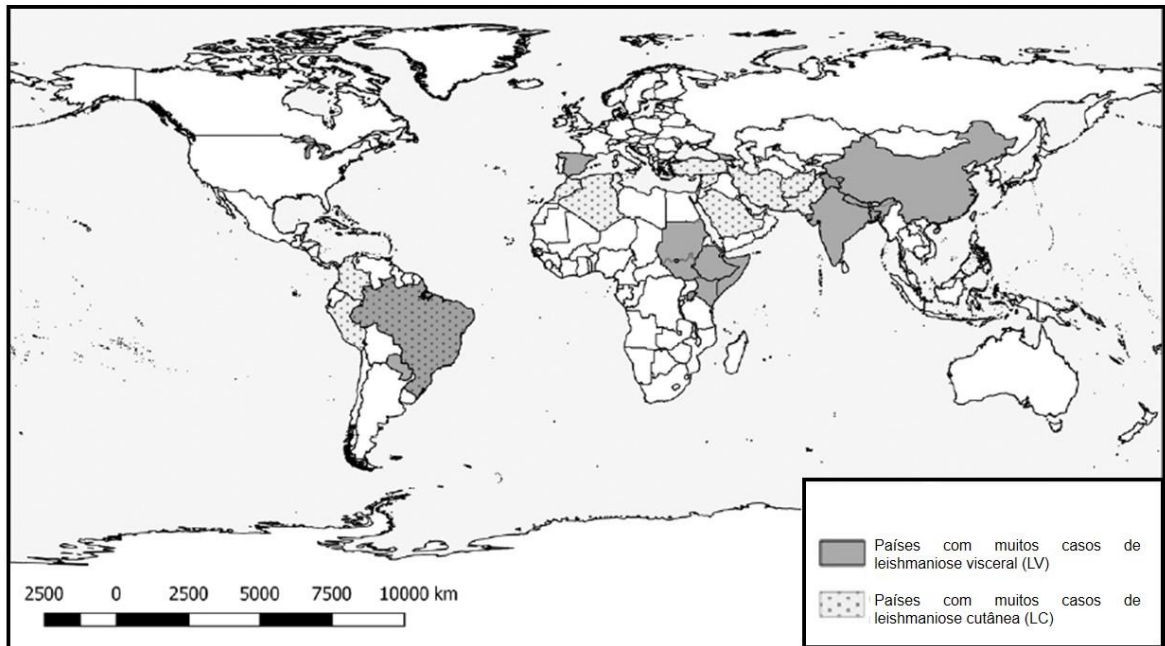
# **CAPÍTULO 1**

## CONSIDERAÇÕES GERAIS

As leishmanioses são doenças causadas por protozoários do gênero *Leishmania*, pertencentes à família Trypanosomatidae, da ordem dos Kinetoplastida. São doenças tropicais negligenciadas, que acometem humanos e várias espécies de animais, inclusive os domésticos. Apresentam distribuição mundial abrangendo, no Novo Mundo, as Américas Central e Sul e, no Velho Mundo, principalmente a África, parte da Europa e a Ásia ( BATES, 2007; DESJEUX, 2004).

As leishmanioses apresentam manifestações clínicas diversas, as quais podem ser categorizados em três tipos: (i) leishmaniose visceral (LV), na qual os parasitas se proliferam no fígado, baço e medula óssea, resultando em imunossupressão do hospedeiro; (ii) leishmaniose cutânea (LC), em que os parasitas permanecem no local de infecção causando ulcerações localizadas; e (iii) leishmaniose mucocutânea (LMC), com destruição crônica do tecido de mucosas (DESJEUX, 2004). Estimativas indicam que estas zoonoses estão presentes em 98 países, com incidência anual de 1,3 milhões de casos na forma cutânea e 0,2-0,4 milhões de casos na forma visceral, provocando cerca de 20.000 mortes por ano (WHO, 2013).

Dentre as formas clínicas, a leishmaniose visceral é a mais grave, podendo ser fatal se não tratada. Contudo a leishmaniose cutânea é a forma mais comum com um maior número de casos. A Organização Mundial da Saúde estimou que 90% dos casos globais de LV notificados até 2010 ocorreram em seis países: Bangladesh, Brasil, Etiópia, Índia, Sudão do Sul e Sudão. No caso da LC, mais de 70% dos casos globais ocorreram em dez países: Afeganistão, Argélia, Brasil, Colômbia, Costa Rica, Etiópia, República Islâmica do Irã, Peru, Sudão e Síria. O Brasil apresenta elevado número de casos de LV e LC, com 3.453 e 19.402 casos registrados, respectivamente (WHO, 2016) (Figura 1).

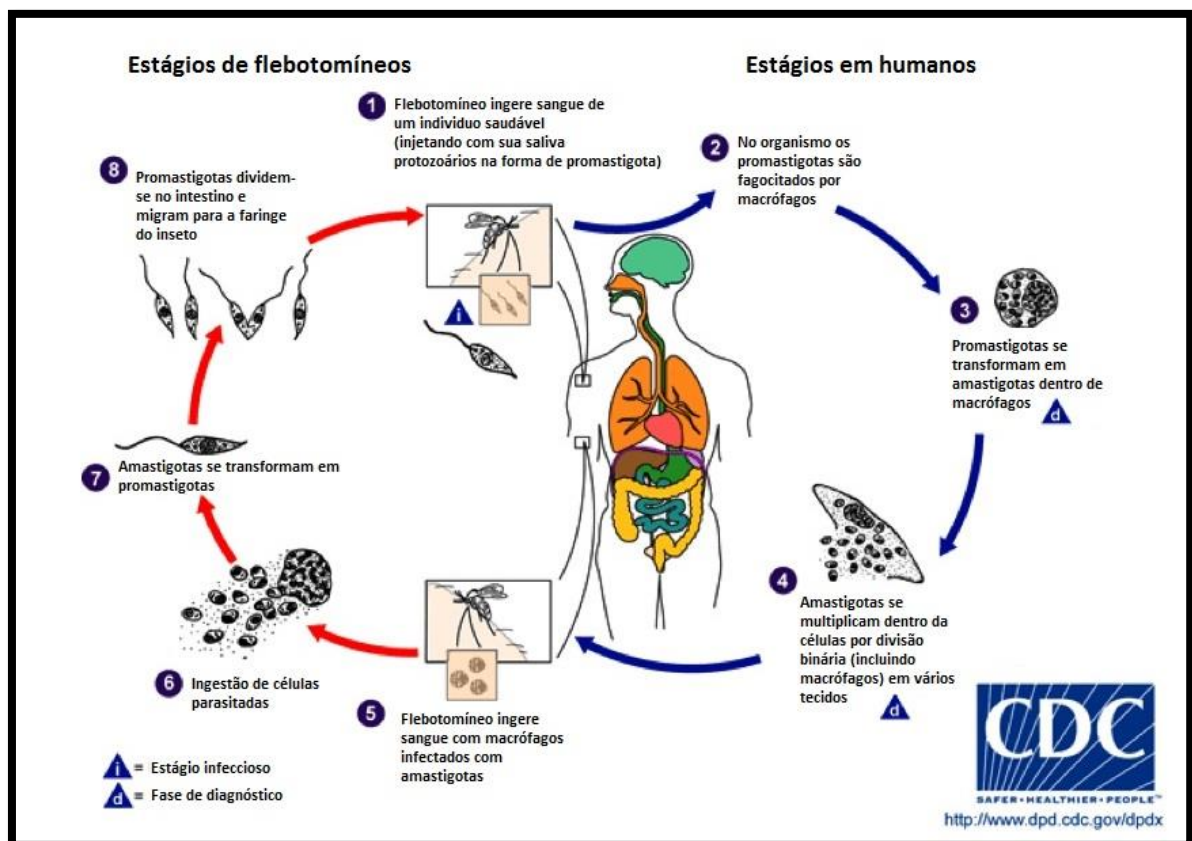


**FIGURA 1** – Distribuição mundial das manifestações clínicas de leishmaniose visceral e cutânea, mostrando que ambas estão presentes no Brasil. Fonte: adaptado de WHO (2016).

Embora existam diferentes espécies de *Leishmania* o ciclo de vida do protozoário é similar, sendo todos heteroxenos, ou seja, necessitam de um hospedeiro invertebrado e um vertebrado para completar o seu ciclo. Os parasitas apresentam duas formas distintas em seu ciclo biológico, a forma promastigota extracelular, flagelada e móvel, que se desenvolve no trato digestório do vetor, e a forma amastigota, intracelular, sem flagelo, imóvel, que são encontradas no interior das células do sistema mononuclear fagocitário, sobretudo, nos macrófagos dos hospedeiros mamíferos (FOREYT, 2005). Uma vez que invadem o macrófago, as formas promastigotas tornam-se amastigotas e se multiplicam; o aumento do número de parasitos acaba rompendo a célula, possibilitando a infecção de novas células (BATES, 2007).

O parasito *Leishmania* é transmitido para o vertebrado através da picada de fêmeas hematófagas da família Psychodidae, sub-família Phlebotominae, conhecidos como flebotomíneos, sendo os gêneros *Lutzomyia* e *Phlebotomus* distribuídos no Novo e no Velho Mundo, respectivamente os mais importantes. Os flebotomíneos se infectam a partir de hospedeiros infectados, quando são ingeridas

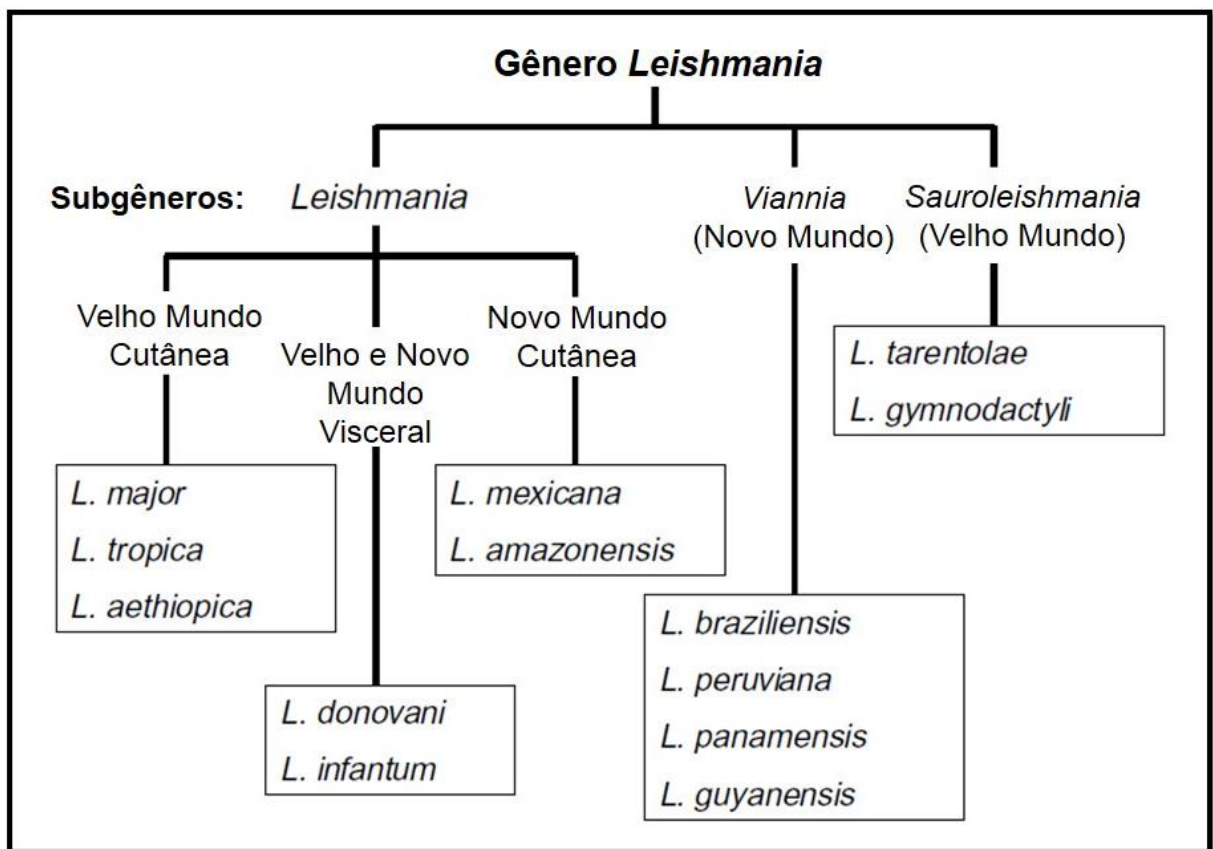
as formas amastigotas pelo vetor, as quais transformam-se em promastigotas dentro do sistema digestório, passando por várias mudanças, até se diferenciarem em formas metacíclicas infectantes para o hospedeiro vertebrado (BATES, 2007) (Figura 2).



**FIGURA 2-** Ciclo biológico de *Leishmania*. Fonte: "Center for Disease Control", EUA, 2016. Disponível em [www.cdc.gov/dpdx/](http://www.cdc.gov/dpdx/).

O comportamento das formas promastigotas no tubo digestório do vetor é usado como parâmetro de classificação taxonômica do gênero *Leishmania* em dois subgêneros: *Leishmania* para os parasitas que se desenvolvem na região suprapilária e *Viannia* para os com desenvolvimento na região do piloro e seção peripilária (LAISON; SHAW 1987). Essa classificação de subgênero em *Leishmania* e *Viannia* é confirmada por sequências de DNA e análises filogenéticas (NOYES et al., 2002).

O subgênero *Leishmania* tem espécies presentes no Velho e no Novo Mundo, porém o subgênero *Viannia* está distribuído exclusivamente na região tropical das Américas, causando infecções cutâneas e mucocutâneas (BATES, 2007). Além dos subgêneros *Leishmania* e *Viannia* há ainda o subgênero *Sauroleishmania*, responsável por causar leishmaniose principalmente em répteis (Figura 3).



**FIGURA 3** – Classificação das espécies de *Leishmania* em subgênero, incluindo informações das formas clínicas e distribuição geográfica. Parasitas nos subgêneros *Leishmania* e *Viannia* infectam mamíferos, enquanto que a *Sauroleishmania* infecta répteis como seus hospedeiros vertebrados. Fonte: adaptado Bates (2007).

Dentre as formas clínicas, a leishmaniose visceral é a mais grave, constituindo-se em um grave problema de saúde pública. Existem registros de surtos de LV muito antigos, com muitos casos de óbitos, como os grandes surtos na Índia entre 1890 e 1900, que despovoaram vilas inteiras; na África, nos anos de



1953-1954, quando ocorreu uma epidemia no Quênia, com mais de 3.000 casos. O Brasil sofreu um surto em 1935 e, nos anos seguintes, no estado do Ceará, com o diagnóstico de mais de 800 casos (REY, 2008).

No Brasil a LV ocorre nas cinco regiões e está em processo de expansão e urbanização pelo país desde a década de 80, devido aos constantes desmatamentos, crescimento urbano desordenado e aos processos migratórios, incluindo a migração de casos humanos e de cães infectados. Com isso, o ciclo de transmissão deixou de ser essencialmente rural, adaptando-se ao ambiente urbano (GONTIJO; MELO, 2004; MENDES et al., 2000). Esta urbanização da doença tornou os cães domésticos o principal reservatório urbano, com um importante papel na epidemiologia da doença (ALVAR et al., 2004). A prevalência de leishmaniose visceral canina está associada à urbanização da doença e à presença de leishmaniose em humanos. 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; HARHAY et al., 2011).

A leishmaniose tegumentar (LT) compreende a forma cutânea e mucocutânea, sendo prevalente em 82 países com incidência estimada em 1,5 milhões de casos por ano. Cerca de 90% dos casos são relatados na África, no Oriente Médio, e América Latina, incluindo o Brasil (DESJEUX, 2004; GOTO; LINDOSO, 2010; MURRAY et al., 2005).

Os agentes etiológicos da LT variam conforme a distribuição geográfica tendo, no Velho Mundo, somente espécies do subgênero *Leishmania*. Na Ásia, África e Europa os agentes são *Leishmania (L.) major*, *L. (L.) tropica*, *L. (L.) aethiopica* e algumas cepas de *L. (L.) infantum*. No Novo Mundo, além de algumas espécies do subgênero *Leishmania*, todas as espécies pertencentes ao subgênero *Viannia* são responsáveis por causar leishmaniose tegumentar, sendo a *Leishmania (Viannia) braziliensis* a espécie mais prevalente, seguida por *L. (L.) amazonensis*, *L. (V.) guyanensis* e *L. (V.) panamensis*. Existe ainda outras espécies como *L. (L.) mexicana*, *L. (L.) pifanoi*, *L. (L.) venezuelensis*, *L. (V.) peruviana*, *L. (L.) shawi* e *L. (V.) lainsoni*, as quais estão presentes principalmente na região amazônica e na América Central (GOTO; LINDOSO, 2010).

Há evidências da ocorrência de LT nas Américas antes da chegada dos colonizadores, retratadas em cerâmicas pré-colombianas feitas pelos incas do Peru e Equador, apresentando mutilações de lábios e narizes. Os conquistadores espanhóis também relataram a ocorrência da doença na população que habitava o continente no século XVII (REY, 2008). Apesar dessa zoonose apresentar ciclo natural nos animais silvestres, a constante urbanização tem levado o homem a frequentar esse ambientes de ciclo natural. Com isto, a doença já se encontra presente em grandes cidades, como resultado da adaptação de vetores e reservatórios ao ambiente urbano, processo este conhecido como a urbanização das leishmanioses no Novo Mundo (SILVA et al., 2005). Essa transmissão entre animais silvestre e domésticos, levou o homem a ser um reservatório acidental.

No Brasil a leishmaniose tegumentar tem ampla distribuição atingindo todos os estados do país com maior ocorrência de casos na região Norte, a qual possui muitas áreas em desmatamento, o que favorece o contato do vetor com o homem e outros animais silvestres (Quadro 1).

Há sete espécies no Brasil que causam a LT, sendo seis do subgênero *Viannia*: *L. (V) braziliensis*, *L. (V.) guyanensis*, *L. (V.) lainsoni*, *L. (V.) shawi*, *L. (V.) lindenbergi* e *L. (V.) naiffi*, e uma do subgênero *Leishmania*, a *L. (L.) amazonensis* (GRAMICCIA; GRADONI, 2005) (Quadro 2).

**Quadro 1** - Casos de leishmaniose tegumentar, de 2010 a 2014, no Brasil, segundo as grandes regiões e unidades federadas. Fonte: adaptado de Sinan/SVS.

Região e UF	2010	2011	2012	2013	2014
<b>Região Norte</b>	7.108	8.615	10.196	8.407	10.387
Rondônia	922	697	1.136	1.203	1.148
Acre	1.026	878	1.134	923	1.030
Amazonas	1.130	2.282	2.246	1.429	1.801
Roraima	641	2.331	444	524	463
Pará	2.331	3.595	4.076	3.058	4.356
Amapá	462	534	693	761	958
Tocantins	596	407	467	509	631
<b>Região Nordeste</b>	8.911	7.952	8.279	5.355	4.969
Maranhão	2.455	2.746	2.517	1.732	2.105
Piauí	146	173	120	69	85
Ceará	1.014	806	940	514	525
Rio Grande do Norte	82	7	6	6	8
Paraíba	83	26	67	35	28
Pernambuco	419	431	314	228	198
Alagoas	32	35	62	51	18
Sergipe	7	7	9	3	4
Bahia	4.673	3.721	4.244	2.717	1.998
<b>Região Sudeste</b>	2.428	2.179	1.388	1.15	1.406
Minas gerais	1.887	1.745	978	781	1033
Espírito Santo	124	122	107	73	89
Rio de Janeiro	86	65	41	25	19
São Paulo	331	247	262	271	319
<b>Região Sul</b>	253	317	439	296	373
Paraná	228	309	419	273	350
Santa Catarina	21	7	13	14	15
Rio Grande do Sul	4	1	7	9	8
<b>Região Centro-Oeste</b>	3.163	2.274	3.118	2.911	3.038
Mato Grosso do Sul	89	93	135	128	141
Mato Grosso do Sul	2.466	1.832	2.639	2.419	2.326
Goiás	553	314	310	355	535
Distrito Federal	55	35	34	20	36
UF ignorada	118	58	127	96	69
<b>Brasil</b>	21.981	21.395	23.547	18.226	20.296

**Quadro 2** – Espécies de *Leishmania* que causam leishmaniose tegumentar no Brasil. A leishmaniose cutânea difusa acontece tardiamente em pacientes que já se trataram de leishmaniose visceral. Fonte: adaptado de Marzochi, Schubach e Marzochi (1999).

<b>Subgênero <i>Viannia</i></b>	<b>Acometimento clínico no homem</b>
<i>Leishmania braziliensis</i>	Lesões cutâneas e mucosas
<i>Leishmania guyanensis</i>	Predominantemente lesões cutâneas
<i>Leishmania panamensis</i>	Predominantemente lesões cutâneas
<i>Leishmania lainsoni</i>	Predominantemente lesões cutâneas
<i>Leishmania shawi</i>	Predominantemente lesões cutâneas
<i>Leishmania naiffi</i>	Predominantemente lesões cutâneas
<i>Leishmania lindenberg</i>	Lesões cutâneas
<b>Subgênero <i>Leishmania</i></b>	<b>Acometimento clínico no homem</b>
<i>Leishmania amazonensis</i>	Lesões cutâneas (eventualmente cutânea difusa)

A LT não recebe tanta atenção das autoridades públicas como a LV, mesmo com a sua crescente incidência mundial, provavelmente devido a sua baixa taxa de mortalidade. Assim, é uma doença negligenciada, com pouco interesse por doadores financeiros e profissionais para implementar atividade de pesquisa, que melhorem a prevenção e controle (REITHINGER et al., 2007).

O Ministério da Saúde tem dados que posicionam o Brasil em primeiro lugar entre os 12 países prioritários da América Latina, com maior número de casos de LT, principalmente na forma cutânea, com cerca de 20% evoluindo para forma mucocutânea. O estado do Amazonas registrou 9.000 casos no período de 2010 a 2015, dados divulgados na 1ª Conferência Internacional em Leishmaniose Cutânea na região amazônica, a qual ocorreu em agosto de 2015 (INPA – Instituto Nacional de Pesquisa da Amazônia).

Apesar da LT ter um amplo aspecto de manifestações clínicas, o desenvolvimento inicial da ferida no local da picada é semelhante, variando de uma única ulceração que pode sumir espontaneamente, para leishmaniose cutânea difusa (GOTO; LINDOSO, 2010; MACHADO et al., 2002). A relação da manifestação clínica com a espécie de *Leishmania* não é simples pois, muitas espécies podem coexistir no mesmo local e a identificação das espécies não é realizada devido à complexidade dos exames laboratoriais (GOTO; LINDOSO, 2010).

A leishmaniose cutânea difusa é uma forma clínica rara de leishmaniose humana, caracterizada pela presença de lesões nodulares que não ulceram, tendo sido descrita pela primeira vez no Brasil, com relatos posteriores em outros países das Américas e do Velho Mundo. As lesões são ricas em parasitas e as espécies frequentemente envolvidas são *L. (L.) mexicana* e *L. (L.) amazonensis* no Novo Mundo e *L. (L.) aethiopica* no Velho Mundo (BARRAL et al., 1995; GOTO; LINDOSO, 2010).

A LT também compreende a forma mucocutânea da leishmaniose (LM), conhecida como espúndia, e caracterizada por destruir tecido das mucosas, como a cavidade nasal e a faringe, causando lesões graves que podem levar a uma deformidade facial. A lesão na mucosa pode ser acompanhada de lesões cutâneas, e a espécie *L. (V.) braziliensis* está presente na maioria dos casos; outras espécies como *L. (V.) panamensis*, *L. (V.) guyanensis*, *L. (L.) amazonensis* e *L. (L.) major* também podem causar estas manifestações (GOTO; LINDOSO, 2010).

Pesquisas no campo genômico vêm gerando informações valiosas para estudos de patogenicidade e tratamento das leishmanioses, bem como para a elucidação da interação entre parasita e hospedeiro, melhorando o entendimento do mecanismo da infecção (EL-SAYED, 2005; PEACOCK et al., 2007; REAL et al., 2013). De fato, o conhecimento de genomas de alguns tripanosomatídeos possibilitará o desenvolvimento de vacinas, drogas e a busca de novos alvos para diagnóstico (MURRAY et al., 2005).

O primeiro genoma nuclear montado deste gênero foi o da *Leishmania (Leishmania) major*, responsável pela infecção cutânea no Velho Mundo. Este estudo demonstrou informações importantes como o fato do genoma estar organizado em 36 cromossomos e muitos dos genes serem ortólogos a outros membros da família Trypanosomatidae (IVENS et al., 2005).

A primeira espécie de *Leishmania* do subgênero *Viannia* que teve seu genoma montado e publicado foi a *L. (V.) braziliensis*, em 2007, responsável por causar leishmaniose mucocutânea. Concomitantemente, foi realizado o sequenciamento da *L. (L.) infantum* (PEACOCK et al., 2007), agente etiológico da leishmaniose visceral. Neste caso as duas espécies foram comparadas com *L. (L.)*

*major*, previamente sequenciada.

Ao comparar estas três espécies de *Leishmania* que causam manifestações clínicas diferentes, observa-se grande sintonia gênica (PEACOCK et al; 2007) e conteúdo de DNA muito próximo, mesmo estando geograficamente separadas. Contudo, existem diferenças que possibilitam a separação em subgêneros distintos: enquanto *L. (L.) major* e *L. (L.) infantum* do Velho Mundo apresentavam 36 cromossomos, *L. (V.) braziliensis* contém 35 cromossomos, devido a fusão do cromossomo 20 e 34 (SMITH, 2007).

Em 2011 foram publicados os genomas de *Leishmania (L.) donovani* (DOWNING et al., 2011) e *Leishmania (L.) mexicana* (ROGERS et al., 2011), demonstrando variação no número de cópias dos cromossomos entre espécies e cepas de *Leishmania*. Em 2012 foi publicado o primeiro genoma de espécie não patogênica para mamíferos, *Leishmania tarentolae*, encontrada em répteis, que ao ser comparada com as demais espécies, demonstrou ausência de 250 genes relacionados à fase de patogenicidade ao hospedeiro mamífero (RAYMOND et al., 2012).

No Brasil o primeiro genoma montado de *Leishmania* foi da *Leishmania (L.) amazonensis*, em 2013 (REAL et al., 2013), espécie está originalmente identificada no Brasil e agente etiológico de leishmaniose tegumentar, causando leishmaniose cutânea difusa. Apesar de ser semelhante à *Leishmania (L.) mexicana*, tendo em comum 23 famílias de proteínas compartilhadas somente entre elas, *L. (L.) amazonensis* apresenta 8 famílias de proteínas exclusivas que devem ter relação com a distribuição geográfica.

Recentemente, foi sequenciado e montado o genoma de outra espécie do subgênero *Viannia*, responsável por causar leishmaniose tegumentar, a *L. (V.) panamensis* (LLANES et al., 2015), contribuindo assim para melhor explorar características deste subgênero presente somente no Novo Mundo (Quadro 3).

**Quadro 3** – Resumo das informações dos principais genomas publicados de espécies de *Leishmania*. Adaptado de Real et al. (2013).

	<i>L. (L.) major</i>	<i>L. (V.) braziliensis</i>	<i>L. (L.) infantum</i>	<i>L. (L.) mexicana</i>	<i>L. (L.) donovani</i>	<i>L. (S.) tarentolae</i>	<i>L. (L.) amazonensis</i>	<i>L. (V.) panamensis</i>
<b>Contigs</b>	36	1041	37	35	2154	N/A	3199	-
<b>Genoma tamanho (Mb)</b>	32,8	33,7	32,1	32,1	32,4	30,4	29,6	30.69
<b>Cromossomos</b>	36	35	36	34	36	36	34	35
<b>Número de genes preditos</b>	8298	8153	8216	8007	8252	8201	8100	7.748
<b>C+G conteúdo (%)</b>	59,3	57,8	59,3	60,5	>60	57,2	58,5	57.56
<b>Referências</b>	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)	Llanes et al. (2015)

A informação genômica das espécies do subgênero *Viannia* ainda é escassa apesar de sua ampla distribuição no Brasil. A *Leishmania (Viannia) shawi*, inicialmente observada em animais silvestres, já foi isolada de infecções cutâneas em humanos no Estado do Pará (SHAW et al., 1992) e *Leishmania (Viannia) guyanensis*, importante causador de leishmaniose cutânea em humanos na região da Amazônia, no Brasil (GRAMICCIA; GRADONI, 2005).

## Objetivo

Com o intuito de ampliar as informações genômicas das espécies de *Leishmania* e do subgênero *Viannia*, trabalhamos com o sequenciamento de DNA genômico e com ferramentas de bioinformática para a montagem dos genomas e anotação de proteínas de duas espécies, *Leishmania (Viannia) guyanensis* e *Leishmania (Viannia) shawi*, agentes etiológicos de leishmaniose tegumentar no Brasil, comparando-as com outras espécies cujos genomas já foram montados e anotados.

## REFERÊNCIAS

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

BARRAL, A.; COSTA, J.M.; BITTENCOURT, A.L.; BARRAL-NETTO, M. C. E. Polar and subpolar diffuse cutaneous leishmaniasis in Brazil: clinical and immunopathologic aspects. **International journal of dermatology**, v. 34, n. 7, p. 474–479, 1995.

BATES, P. A. Transmission of *Leishmania* metacyclic promastigotes by phlebotomine sand flies. **International Journal for Parasitology**, v. 37, n. 10, p. 1097–1106, 2007.

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.

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.

DESJEUX, P. Leishmaniasis: current situation and new perspectives. **Comparative Immunology, Microbiology and Infectious Diseases**, v. 27, n. 5, p. 305–318, 2004.

EL-SAYED, N. M.; MYLER, P. J.; BLANDIN, G.; BERRIMAN, M.; CRABTREE, J.; AGGARWAL, G.; CALER, E.; RENAULD, H.; WORTHEY, E. A.; HERTZ-FOWLER, C.; GHEDIN, E.; PEACOCK, C.; BARTHOLOMEU, D. C.; HASS, B, J.; TRAN, A.; WORTMAN, J. R.; ALSMARK, U. C. M.; ANGIUOLI, S.; ANUPAMA, A.; BADGER, J.; BRINGAUD, F.; CADAG, E.; CARLTON, J. M.; CERQUEIRA, G. C.; CREASY, T.; DELCHER, A. L.; DJIKENG, A.; EMBLEY, T. M.; HAUSER, C.; IVENS, A. C.; KUMMERFELD, S. K.; PEREIRA-LEAL, J. B.; NILSSON, D.; PETERSON, J.; SALZBERG, S. L.; SHALLOM, J.; SILVA, J. C.; SUNDARAM, J.; WESTENBERGER, S.; WHITE, O.; MELVILLE, S. E.; DONELSON J. E.; ANDERSSON, B.; STUART, K. D.; HALL, N. Comparative genomics of trypanosomatid parasitic protozoa. **Science**, v. 309, n. 5733, p. 404–409, 2005.

FOREYT, W.J. **Parasitologia veterinária**. 5ª ed. São Paulo: Roca, 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.

GONTIJO, C.M.F.; MELO, M.N. Leishmaniose visceral no Brasil: quadro atual, desafios e perspectivas. **Revista Brasileira de Epidemiologia**, São Paulo, v.7, n.3, p.338-349, 2004.

GOTO H.; LINDOSO, J. A. L. Current diagnosis and treatment of cutaneous and mucocutaneous leishmaniasis. **Expert Review Anti-Infective Therapy**, v. 8, n. 4, p. 419–433, 2010.

GRAMICCIA, M.; GRADONI, L. The current status of zoonotic leishmaniasis and approaches to disease control. **International Journal for Parasitology**, v. 35, n. 11–12, p. 1169–1180, 2005.

HARHAY, M.O.; OLLIARO, P.L.; COSTA, D.L.; COSTA, C.H.N. Review: Urban parasitology: visceral leishmaniasis in Brazil. **Trends in Parasitology**, v.27, n.9, p.403-409, 2011.

IVENS, A. C.; PEACOCK, C. S.; WORTHEY, E. A.; MURPHY, L.; AGGARWAL, G.; BERRIMAN, M.; SISK, E.; RAJANDREAM, M.-A.; ADLEM, E.; AERT, R.; ANUPAMA, A.; APOSTOLOU, Z.; ATTIPOE, P.; BASON, N.; BAUSER, C.; BECK, A.; BEVERLEY, S. M.; BIANCHETTIN, G.; BORZYM, K.; BOTHE, G.; BRUSCHI, C. V.; COLLINS, M.; CADAG, E.; CIARLONI, L.; CLAYTON, C.; COULSON, R. M. R.; CRONIN, A.; CRUZ, A. K.; DAVIES, R. M.; DE GAUDENZI, J.; DOBSON, D. E.; DUESTERHOEFT, A.; FAZELINA, G.; FOSKER, N.; FRASCH, A. C.; FRASER, A.; FUCHS, M.; GABEL, C.; GOBLE, A.; GOFFEAU, A.; HARRIS, D.; HERTZ-FOWLER, C.; HILBERT, H.; HORN, D.; HUANG, Y.; KLAGES, S.; KNIGHTS, A.; KUBE, M.; LARKE, N.; LITVIN, L.; LORD, A.; LOUIE, T.; MARRA, M.; MASUY, D.; MATTHEWS, K.; MICHAELI, S.; MOTTRAM, J. C.; MÜLLER-AUER, S.; MUNDEN, H.; NELSON, S.; NORBERTCZAK, H.; OLIVER, K.; O'NEIL, S.; PENTONY, M.; POHL, T. M.; PRICE, C.; PURNELLE, B.; QUAIL, M. A.; RABBINOWITSCH, E.; REINHARDT, R.; RIEGER, M.; RINTA, J.; ROBBEN, J.; ROBERTSON, L.; RUIZ, J. C.; RUTTER, S.; SAUNDERS, D.; SCHÄFER, M.; SCHEIN, J.; SCHWARTZ, D. C.; SEEGER, K.; SEYLER, A.; SHARP, S.; SHIN, H.; SIVAM, D.; SQUARES, R.; SQUARES, S.; TOSATO, V.; VOGT, C.; VOLCKAERT, G.; WAMBUTT, R.; WARREN, T.; WEDLER, H.; WOODWARD, J.; ZHOU, S.; ZIMMERMANN, W.; SMITH, D. F.; BLACKWELL, J. M.; STUART, K. D.; BARRELL, B.; MYLER, P. J. The genome of the kinetoplastid parasite, *Leishmania major*. **Science**, v. 15, n. 3095733, p. 436–442, 2005.

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.

LLANES, A.; RESTREPO, C. M.; DEL VECCHIO, G.; ANGUIZOLA, F. J.; LLEONART, R. The genome of *Leishmania panamensis*: insights into genomics of

the *L. (Viannia)* subgenus. **Scientific reports**, v. 5, p. 8550, 2015.

MARZOCHI, M.C.A., SCHUBACH, A., MARZOCHI, K. B. F. **Leishmaniose Tegumentar Americana**, p. 39-64, São Paulo: Atheneu, 1999.

MACHADO, P.; ARAÚJO, C.; DA SILVA, A. T.; ALMEIDA, R. P.; D'OLIVEIRA JR, A.; BITTENCOURT, A.; CARVALHO, E. M. Failure of early treatment of cutaneous leishmaniasis in preventing the development of an ulcer. **Clinical infectious diseases : an official publication of the Infectious Diseases Society of America**, v. 34, n. 12, p. E69-73, 2002.

MENDES, W.S.; TROVÃO, J.R.; SILVA, A.A.M. Dinâmica da ocupação do espaço na cidade de São Luís e a leishmaniose visceral. **Caderno de Saúde Pública**, Rio de Janeiro, v.16, n.3, p.872, 2000.

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

NOYES, H.; PRATLONG, F.; CHANCE, M.; ELLIS, J.; LANOTTE, G.; DEDET, J. P. A previously unclassified trypanosomatid responsible for human cutaneous lesions in Martinique ( French West Indies ) is the most divergent member of the genus *Leishmania* ss. **Parasitology**, v.124, n. 1, p. 17–24, 2002.

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. O.; 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, n. 7, p. 839–47, 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.;ROMA 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.

REITHINGER, R.; DUJARDIN, J.-C.; LOUZIR, H.; PIRMEZ, C.; ALEXANDER, B.; BROOKER, S. Cutaneous leishmaniasis. **The Lancet. Infectious diseases**, v. 7, n. 9, p. 581–596, 2007.

REY, L. **Parasitologia**. 4ª ed, Rio de Janeiro: Guanabara, 2008

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.

SILVA E.S.; GONTIJO, C.M.F.; MELO, M.N. Contribution of molecular techniques to the epidemiology of neotropical *Leishmania* species. **Trends in Parasitology**, v.21, n. 12, p. 550-552, 2005.

SMITH, D. F.; PEACOCK, C. S.; CRUZ, A. K. Comparative genomics: From genotype to disease phenotype in the leishmaniasis. **International Journal for Parasitology**, v. 37, n. 11, p. 1173–1186, 2007.

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.

WORLD HEALTH ORGANIZATION WHO. **Weekly epidemiological record**. Geneva: WHO, 2016. Disponível em: <http://www.who.int/wer/2016/wer9122.pdf?ua=1> >

## **CAPÍTULO 2**

## **Montagem de novo dos genomas de *Leishmania (Viannia) guyanensis* e *Leishmania (Viannia) shawi***

**RESUMO** – Os protozoários do gênero *Leishmania* são responsáveis por causar diferentes formas clínicas da leishmaniose, zoonose com distribuição mundial. As espécies de *Leishmania* patogênicas aos seres humanos são classificadas em dois subgêneros, *Leishmania*, que ocorre no Velho e Novo Mundo, e *Viannia*, presente apenas nas regiões tropicais das Américas, sendo responsáveis pela manifestação tegumentar que faz referência às formas cutânea e mucocutânea. No presente estudo foi realizado o sequenciamento, montagem dos genomas e anotação de proteínas de duas espécies do subgênero *Viannia* presentes no Brasil: *Leishmania (V.) guyanensis* e *Leishmania (V.) shawi*. Na montagem obtivemos 3.999 *scaffolds* para *Leishmania (V.) guyanensis* e 5.032 *scaffolds* para *Leishmania (V.) shawi*, com 8.459 e 8.900 proteínas anotadas, respectivamente. Ao comparar a porcentagem de identidade das proteínas foi possível observar, quão próximas são estas espécies, apresentando mais de 95% de similaridade. Na análise dos domínios das proteínas, encontramos 11.069 domínios conservados entre as espécies *L. (V.) shawi*, *L. (V.) guyanensis*, *L. (V.) braziliensis* e *L. (V.) panamensis*, enquanto que domínios exclusivos de *L. (V.) shawi* e *L. (V.) guyanensis* são 43 e 23, respectivamente. A montagem dos genomas e a anotação de proteínas destas espécies de *Leishmania* aumentaram as informações genômicas destas espécies e do subgênero *Viannia* e contribuirão para o conhecimento deste subgênero.

**Palavras-chave:** genômica, leishmaniose, saúde pública, trypanosoma

***De novo genome assembly of Leishmania (Viannia) guyanensis and Leishmania (Viannia) shawi***

**SUMMARY** – Leishmaniosis are zoonotic diseases caused by the *Leishmania* genus protozoa which are worldwide distributed and considered a public health problem. The genus *Leishmania* is divided into two subgenus (*Leishmania* and *Viannia*), with many species that are responsible for different clinical manifestations. The species from the subgenus *Viannia* are distributed in the tropical regions of the American continent and cause the cutaneous and mucocutaneous forms in human and animals. Genomic sequencing and protein annotation of *Leishmania* species has allowed better understanding the similarities and differences among species and subgenus. The genome sequencing of *Leishmania (Viannia) guyanensis* and *Leishmania (Viannia) shawi* resulted in 3,999 and 5,032 scaffolds with 8,459 and 8,900 annotated proteins respectively. Annotation of the proteins showed similarity higher than 95% among these *Viannia* species. Protein domain analysis and annotation revealed 11,069 conserved domains between *L. (V.) guyanensis*, *L. (V.) shawi*, *L. (V.) braziliensis* and *L. (V.) panamensis* whereas exclusive domains for *L. (V.) guyanensis* and *L. (V.) shawi* were 43 and 23, respectively. The assembly of *Leishmania (Viannia) guyanensis* and *Leishmania (Viannia) shawi* genomes, both belonging to the subgenus *Viannia*, besides to enrichment of the information about genomics of all *Leishmania* species contributes to improve the knowledge on this subgenus.

**Key - words:** Genomics, leishmaniosis, public health, trypanosoma

## INTRODUÇÃO

As leishmanioses são doenças negligenciadas que acometem humanos e várias espécies animais. Ocorrem no Novo Mundo, representado pelas Américas Central e Sul, e no Velho Mundo, principalmente na África, parte da Europa e na Ásia (BATES, 2007; DESJEUX, 2004; WHO 2013).

Os vetores responsáveis pela transmissão desta zoonose são fêmeas de flebotomíneos, sendo o gênero *Lutzomyia* o mais importante nas Américas. De acordo com o local de desenvolvimento das formas promastigostas no trato digestório dos flebotomíneos o gênero *Leishmania* é dividido em dois subgêneros: *Leishmania* para os parasitas que se desenvolvem na região suprapilária; e *Viannia* para os com desenvolvimento na região do piloro e seção peripilária (LAISON; SHAW 1987). Esta classificação de subgêneros é corroborada por sequências de DNA e análises filogenéticas (NOYES et al., 2002).

Estes dois subgêneros incluem cerca de 20 espécies patogênicas para humanos que diferem em sua distribuição geográfica e forma clínica. Na Ásia, África e Europa as espécies *Leishmania (Leishmania) major* e *Leishmania (Leishmania) tropica* são responsáveis por causarem a leishmaniose tegumentar (LT), que compreende as formas cutânea e mucocutânea. No Novo Mundo, além de alguns agentes etiológicos do subgênero *Leishmania*, todas as espécies pertencentes ao subgênero *Viannia* são responsáveis por causar leishmaniose tegumentar. A *Leishmania (Viannia) braziliensis* é a espécie mais prevalente, seguida por *L. (L.) amazonensis*, *L. (V.) guyanensis* e *L. (V.) panamensis*. Existem ainda outras espécies como *L. (L.) mexicana*, *L. (L.) pifanoi*, *L. (L.) venezuelensis*, *L. (V.) peruviana*, *L. (L.) shawi* e *L. (V.) lainsoni*, presentes principalmente na região amazônica e na América Central, causando a LT (GOTO; LINDOSO, 2010).

No Brasil, há sete espécies de *Leishmania* que causam leishmaniose tegumentar, sendo seis do subgênero *Viannia*: *L. (V) braziliensis*, *L. (V.) guyanensis*, *L. (V.) lainsoni*, *L. (V.) shawi*, *L. (V.) lindenbergi* e *L. (V.) naiffi*, e uma do subgênero *Leishmania*, a *L. (L.) amazonensis*.

A *Leishmania (Viannia) guyanensis* é importante causa de leishmaniose cutânea em humanos na região da Amazônia e tem como reservatório primário a preguiça e, como vetor principal, o *Lutzomyia umbratilis* (GRAMICCIA; GRADONI, 2005). A *L. (V.) shawi* também é causa da forma cutânea em humanos, parasita espécies de animais silvestres, incluindo vários mamíferos arbóreos na floresta tropical do Pará, e seu principal vetor é o *Lutzomyia whitmani* (GRAMICCIA; GRADONI, 2005; SHAW et al., 1991).

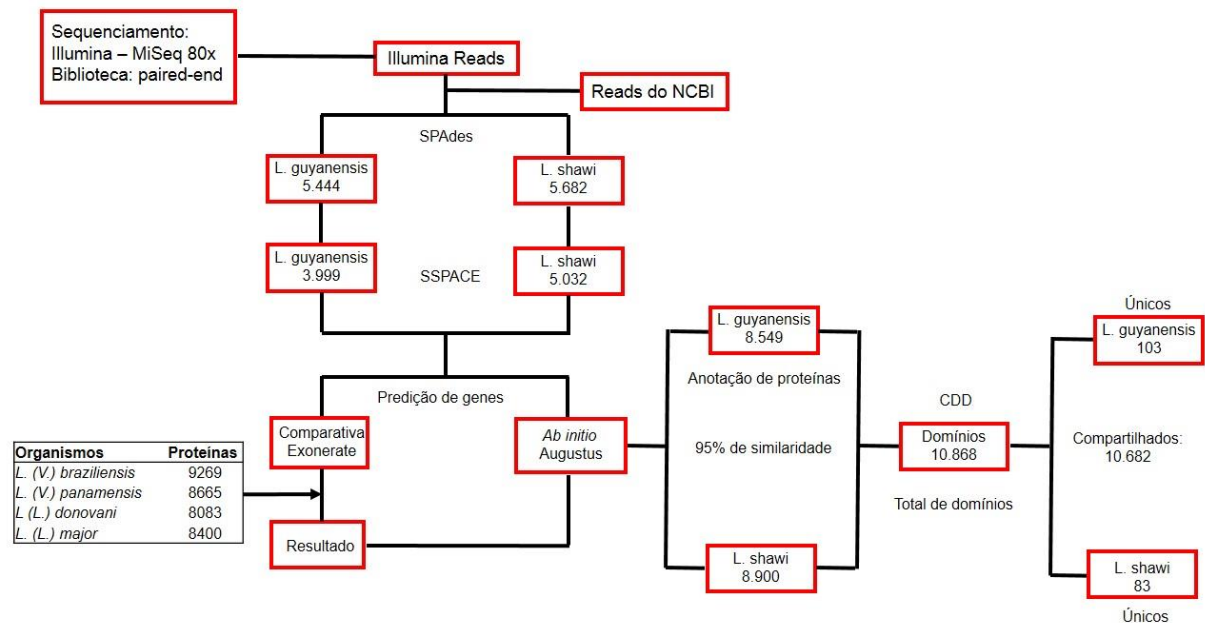
O conhecimento de genomas de alguns tripanosomatídeos tem possibilitado o desenvolvimento de vacinas, drogas e a busca de novos alvos para diagnóstico (MURRAY et al., 2005). O primeiro genoma nuclear montado foi o de *Leishmania (Leishmania) major*, responsável por infecção cutânea no Velho Mundo (IVENS et al., 2005). Em seguida foi realizada a montagem do genoma de *L. (V.) braziliensis*, agente causador de leishmaniose mucocutânea, juntamente com o de *L. (L.) infantum*, responsável pela forma visceral (PEACOCK et al., 2007). Posteriormente as espécies *L. (L.) amazonensis* (REAL et al., 2013) e a *L. (V.) panamensis* (LLANES et al., 2015), agentes de leishmaniose tegumentar, foram também sequenciadas.

Apesar do grande número de espécies e diferentes manifestações clínicas das leishmanioses, ainda há pouca informação sobre as espécies do subgênero *Viannia* como *L. (V.) guyanensis* e *L. (V.) shawi*. Deste modo, o objetivo do presente estudo foi sequenciar, montar os genomas e anotar as proteínas de *L. (V.) guyanensis* e *L. (V.) shawi*, comparando-as com outras espécies, no intuito de ampliar o conhecimento genômico das espécies do subgênero *Viannia*.

## **MATERIAIS E MÉTODOS**

Todos os procedimentos empregados neste estudo, exceto o detalhamento de processamento de DNA e comparação com as demais espécies de *Leishmania*, estão resumidos no fluxograma de trabalho (Figura 1), que apresenta as etapas de montagem e anotação do genoma, com alguns resultados, comparando as duas espécies em estudo.





**FIGURA 1** – Fluxograma de trabalho resumindo as análises bioinformáticas utilizadas no presente estudo, dando ênfase para *L. (V.) guyanensis* e *L. (V.) shawi*.

## Amostras biológicas e extração de DNA

O DNA genômico das amostras de *Leishmania (Viannia) guyanensis* (cepa MHOM/BR/1975/M4147) e *Leishmania (Viannia) shawi* (cepa MHOM/BR/1967/PH8), cedidas pela CLIOC da Fundação Oswaldo Cruz do Rio de Janeiro, foi extraído com o kit QIAmp DNA mini Kit (Qiagen®), seguindo-se protocolo do fabricante. A quantificação e a qualidade do DNA foram avaliados por espectrofotometria, sob luz ultravioleta (NanoDrop® ND1000 Thermo Scientific) e por fluorometria (Qubit®, Life Technologies™).

## Bibliotecas *paired-end*

Para o preparo das bibliotecas de DNA *paired-end*, utilizou-se o kit *Nextera® DNA Library Preparation* (Illumina Inc., San Diego, CA), seguindo-se as instruções do fabricante. As amostras de DNA genômico de cada espécie de *Leishmania* foram

fragmentadas por ação enzimática (tagmentação), e ligadas à adaptadores. Em seguida, os fragmentos foram denaturados e amplificados por reação em cadeia da polimerase (PCR), utilizando-se oligonucleotídeos iniciadores (*primers*) randômicos. A qualidade e a integridade das bibliotecas foram avaliadas utilizando-se eletroforese capilar baseada em chip (Bioanalyzer® Agilent); o tamanho do inserto das bibliotecas foi de aproximadamente 400 pares de bases (pb).

### **Sequenciamento e processamento dos dados**

O sequenciamento de DNA foi realizado por síntese, utilizando-se o kit *MiSeq Reagent v3* e a plataforma de sequenciamento Illumina MiSeq®, com a cobertura de 80 vezes. Após o preparo das bibliotecas de DNA, os fragmentos ligados aos adaptadores foram hibridizados à célula de fluxo (*flow cell*), gerando aglomerados (*clusters*).

A identificação e remoção de adaptadores em ambas as extremidades de cada *read*, bem como das bases com qualidade inferior a 20 (probabilidade de erro de 0,01) foram realizados por meio do programa Trimmomatic v0.36 (BOLGER; LOHSE; USADEL, 2014).

### **Montagem dos genomas**

Com o objetivo de obter melhor resultado na montagem dos genomas, seis arquivos contendo *reads paired-end* de *Leishmania (V.) guyanensis* e um da espécie *Leishmania (V.) shawi*, disponíveis no arquivo de sequências *read Sequence Read Archive* (SRA) do banco de dados do *National Center for Biotechnology Information* (NCBI), foram também utilizados. Estes arquivos foram gerados por sequenciamento em plataforma Illumina HiSeq®; as sequências apresentavam tamanho de 100 pares de bases (pb), e os tamanhos médios dos insertos foram calculados utilizando-se script em Perl (Quadro 1).

**Quadro1** - Características dos arquivos de sequências *reads* dos genomas de *Leishmania (V.) guyanensis* e *Leishmania (V.) shawi*

Espécie	Fonte*	Total de "reads"	% GC	Tamanho (pb)	Tamanho inserto (pb)
<i>Leishmania (V.) guyanensis</i>	SRX767379	12.450.070	53	100	321
	ERX442765	24.128.282	57	101	310
	ERX442766	35.415.718	57	101	366
	ERX442767	24.642.388	57	101	283
	ERX442768	19.618.618	57	101	313
	ERX442769	27.809.294	57	101	350
<i>Leishmania (V.) shawi</i>	SRX764331	10.220.958	51	100	350

\*Disponíveis em: <https://www.ncbi.nlm.nih.gov/sra>

A montagem *de novo* dos genomas de *Leishmania* foi realizada utilizando-se o programa SPAdes (BANKEVICH, 2012), o qual realiza a correção dos *reads* sequenciados e, em seguida, faz a construção de bibliotecas de *k-mer* (gerados pela fragmentação dos *reads*) usadas para a construção de grafos *de Bruijn*. Nos estágios subsequentes o programa realiza operações de grafos teóricos (sem a geração de *k-mers*), baseados na topologia dos grafos, na cobertura e no tamanho das sequências. Por fim, as sequências consenso de DNA (*scaffolds*) são restauradas.

Posteriormente os *scaffolds* construídos pelo SPAdes foram submetidos ao SSPACE v2.0 (BOETZER et al., 2010), pacote de programas que possibilita a extensão dos *scaffolds* montados através do mapeamento dos *reads*, utilizando o programa *Bowtie* (LANGMEAD et al., 2009), seguindo-se à extensão dos *scaffolds*, pelo programa SSAKE (WARREN et al., 2007). Cada extensão é computada apenas se o posicionamento dos *reads* mapeados satisfizer a distância que compreende o tamanho do inserto inicial.

Para a análise da qualidade das montagens dos genomas foi utilizado o programa QUAST v4.2 (GUREVICH, 2013), o qual monta tabelas e gráficos a partir dos dados de montagem, possibilitando a comparação entre diferentes métodos com diferentes parâmetros, havendo ou não um genoma de referência.

## Predição de genes

A predição de genes foi realizada utilizando-se os programas Exonerate (SLATER; BIRNEY, 2005) e Augustus (STANKE; MORGENSTERN, 2005). O Exonerate determina estruturas gênicas através do alinhamento de uma única proteína à sequência genômica alvo. Para tanto, o programa utiliza o método heurístico de alinhamento exaustivo das sequências contra um banco de dados de referência, criado com sequências de *L. (V.) braziliensis*, *L. (V.) panamensis*, *L. (L.) donovani* e *L. (L.) major*, disponíveis no banco de dados TriTrypDB (disponível em: <http://tritrypdb.org/common/downloads/release-9.0/>).

Os resultados obtidos pelo Exonerate foram utilizados como arquivo de sugestões (*hints*) para a segunda etapa da predição, com o programa Augustus. Este programa, além de realizar o processo de forma *ab initio*, ou seja, baseando-se puramente nas sequências, utiliza também arquivo de sugestões que podem ser compostos, por exemplo, de proteínas e marcadores de sequências genéticas (ESTs).

O arquivo de saída do Augustus, além de apresentar as sequências gênicas anotadas, também apresentou as mesmas transcritas em proteínas. Com um *script em Perl*, as sequências de nucleotídeos referentes aos genes e às proteínas foram extraídas em arquivos separados, no formato fasta, para serem usadas em análises posteriores.

A avaliação do grau de similaridade entre as espécies de *Leishmania* foi realizada por meio de alinhamento das proteínas de *L. (V.) guyanensis* e de *L. (V.) shawi*, pelo algoritmo BLASTp (ALTSCHUL et al., 1997), com *e-value* de  $1 \times 10^{-5}$ , contra sequências proteicas de um conjunto de referências retiradas do banco de dados TriTrypDB (disponíveis em: <http://tritrypdb.org/common/downloads/release-9.0/>). As espécies usadas foram *L. (V.) braziliensis* (9.269 proteínas), *L. (V.) panamensis* (8.665 proteínas), *L. (L.) donovani* (8.083 proteínas) e *L. (L.) major* (8.400 proteínas). As proteínas de *L. (V.) guyanensis* e *L. (V.) shawi* foram também alinhadas entre si.

## **Análise de domínios das proteínas**

As proteínas foram também analisadas conforme seus domínios conservados, os quais são responsáveis por caracterizar a estrutura tridimensional das proteínas, cuja análise pode explicar a evolução molecular. Para tal, foi utilizado o *Conserved Domain Database* (CDD), que contém a localização dos domínios preservados nas proteínas (MARCHLER-BAUER et al., 2013), é formatado para o RPS-BLAST, o qual usa o algoritmo de “*reverse position-specific BLAST*”. As sequências de proteínas são comparadas a uma matriz de escores posição-específica (PSSM) preparada a partir de alinhamentos de domínios conservados (MARCHLER-BAUER et al., 2013).

Com RPS-BLAST, sequências proteicas de *L. (V.) guyanensis* e *L. (V.) shawi*, e das referências *L. (V.) braziliensis*, *L. (V.) panamensis*, *L. (L.) donovani* e *L. (L.) major*, retiradas do banco de dados TriTrypDB (disponíveis em: <http://tritrypdb.org/common/downloads/release-9.0/>), foram confrontadas por PSSM, usando *e-value* de  $1 \times 10^{-5}$  e o máximo de 20 domínios para cada descrição. Os resultados das diferentes espécies foram então comparados entre si.

## **RESULTADOS**

O sequenciamento dos genomas resultou em *reads* com tamanhos variando entre 35 e 300 pares de base (pb), sendo o tamanho médio de 171 pb para *L. (V.) guyanensis* e 248 pb em *L. (V.) shawi*. Após a remoção de adaptadores, bem como de bases com baixa qualidade utilizando-se o programa Trimmomatic v0.36, foi possível utilizar 81,8% do total de *reads* gerados para *L. (V.) guyanensis*, com tamanho médio de 150 pb, e 72,4% para *L. (V.) shawi*, com tamanho médio de 233 pb. A montagem resultou num total de 5.444 e 5.682 *scaffolds*, com N50 de 70.610 e 31.650 bases para *L. (V.) guyanensis* e *L. (V.) shawi*, respectivamente (Tabela 1). Em adição, com o intuito de aperfeiçoar os resultados da montagem, foi utilizado o programa SSPACE v.2.0, o qual realiza a extensão dos *scaffolds*, de acordo com

informações do tamanho do inserto e do resultado do mapeamento dos *reads* contra os *scaffolds*. Estes procedimentos resultaram em um total de 3.999 e 5.032 *scaffolds*, com N50 de 118.499 e de 35.801 bases para as espécies *L. (V.) guyanensis* e *L. (V.) shawi*, respectivamente (Tabela 1).

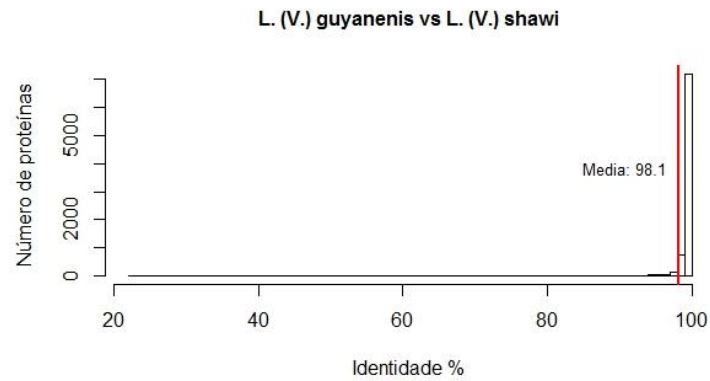
**Tabela 1** - Resumo da montagem de novo dos genomas de *Leishmania (V.) guyanensis* e *Leishmania (V.) shawi* de acordo com os programas SPAdes e SSPACE v.2.0

<b>Espécie</b> <b>Assembly</b>	<i>L. (V.) guyanensis</i>		<i>L. (V.) shawi</i>	
	<b>SPAdes</b>	<b>SSPACE</b>	<b>SPAdes</b>	<b>SSPACE</b>
<i>scaffolds</i> (>= 0 bp)	5.444	3.999	5.682	5.032
GC (%)	57,46	57,5	57,5	57,53
N50	70.610	118.499	31.65	35.801
N75	36.850	59.536	16.506	17.794
L50	116	75	277	250
L75	262	170	608	552
Maior <i>scaffold</i> (bp)	503.403	674.613	196.256	196.298
Total de bases montadas em Mb	30,7	31,1	30,5	30,7

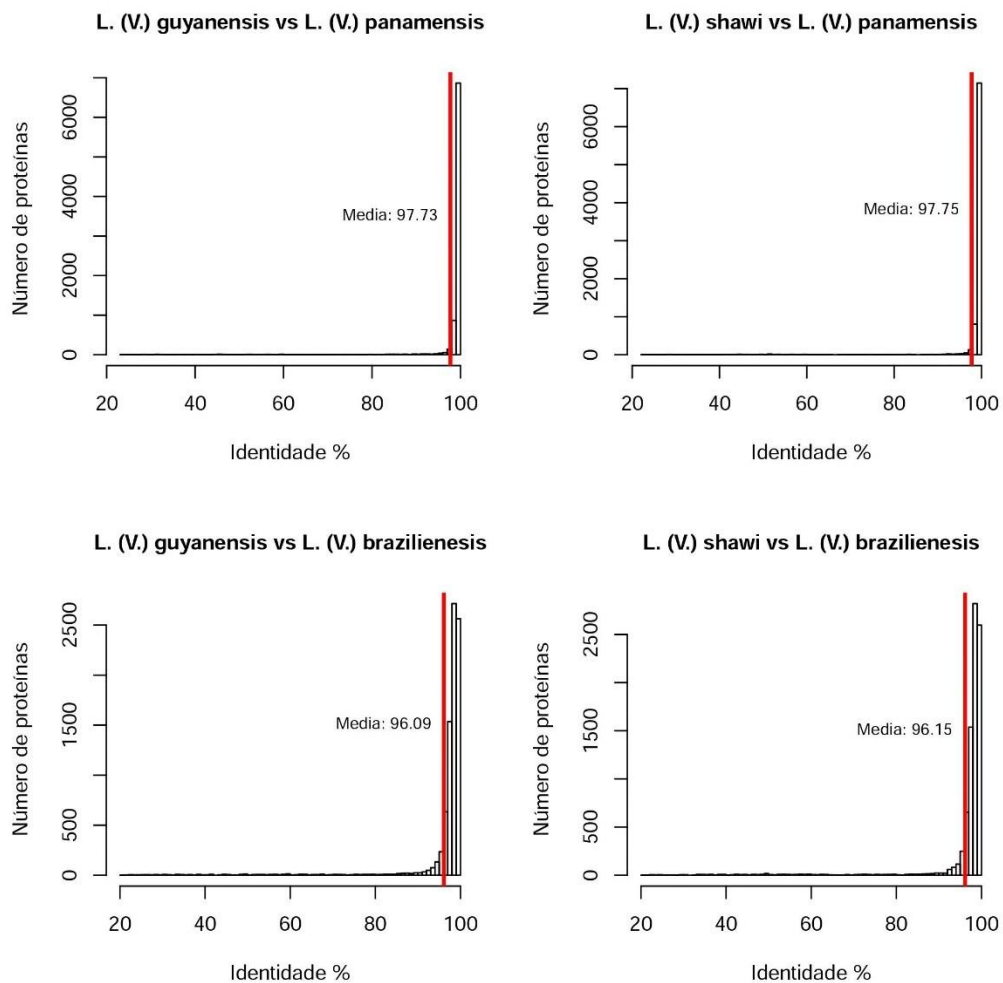
pb: pares de bases; N50: tamanho mínimo dos *scaffolds* que cobrem 50% do total de bases montadas; N75: tamanho mínimo dos *scaffolds* que cobrem 75% do total de bases montadas; L50: número de *scaffolds* que compõem 50% do total de bases montadas; L75: número de *scaffolds* que compõem 75% do total de bases montadas. Mb significa um milhão pares de bases de nucleotídeos montados.

Na anotação de *L. (V.) guyanensis* obteve-se 8.459 proteínas traduzidas das sequências preditas e 8.900 proteínas de *L. (V.) shawi*. O alinhamento destas proteínas por meio da ferramenta BLASTp, demonstrou ampla similaridade das sequências proteicas entre as espécies de *Leishmania* avaliadas, mantendo identidade de aminoácidos superior a 75%. A comparação de *L. (V.) guyanensis* e *L. (V.) shawi* demonstrou identidade média de 98% entre elas (Figura 2). Quando estas espécies foram comparadas com *L. (V.) panamensis* e *L. (V.) braziliensis*, a identidade foi de 97% e 96%, respectivamente (Figura 3). A comparação com uma espécie que causa a leishmaniose visceral, *L. (L.) donovani* revelou média em ambas de 77% de identidade entre os aminoácidos das proteínas (Figura 4A e 4B). Já com *L. (L.) major*, que também causa leishmaniose cutânea mas pertence a outro subgênero, a média de identidade foi de 76% com *L. (V.) guyanensis* (Figura 4C) e

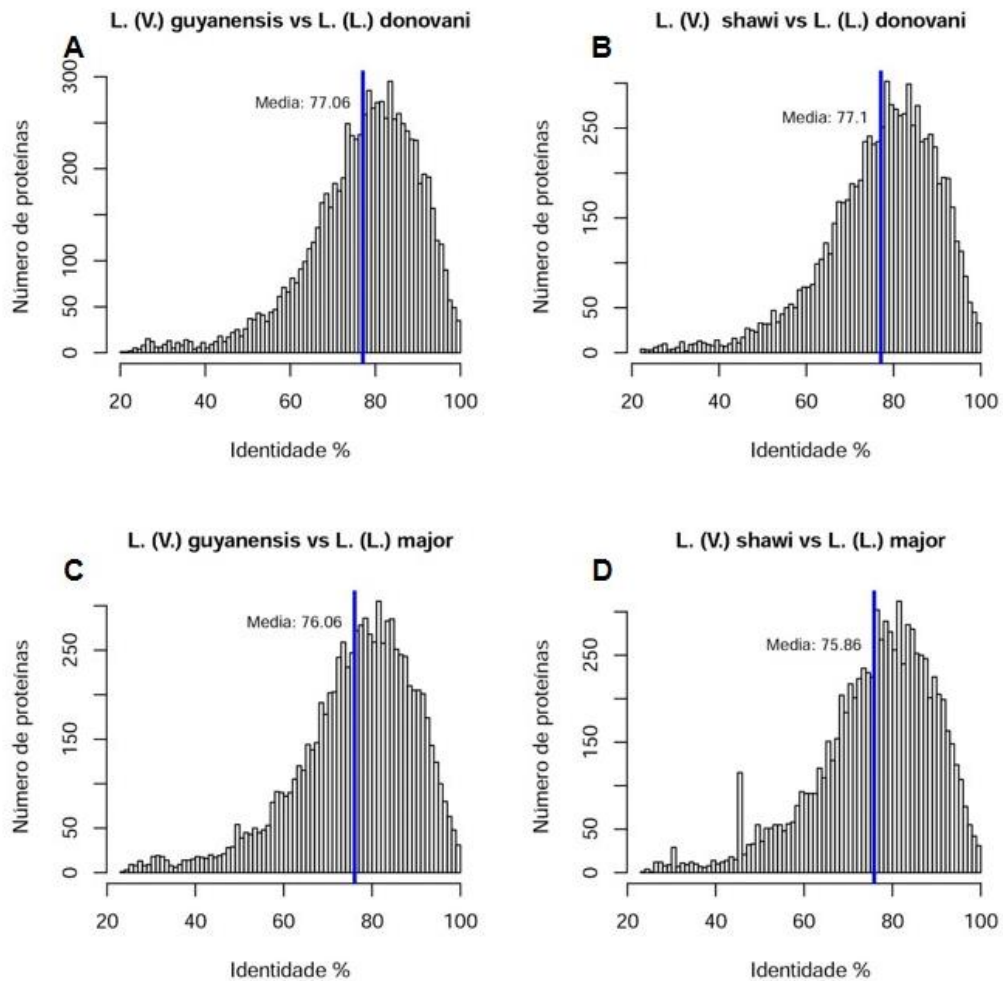
75% com *L. (V.) shawi* (Figura 4D).



**FIGURA 2** - Identidade entre as proteínas de *L. (V.) guyanensis* e *L. (V.) shawi*.



**FIGURA 3** - Identidade entre as proteínas de *L. (V.) guyanensis* e *L. (V.) shawi* em comparação com *L. (V.) panamensis* e *L. (V.) braziliensis*.

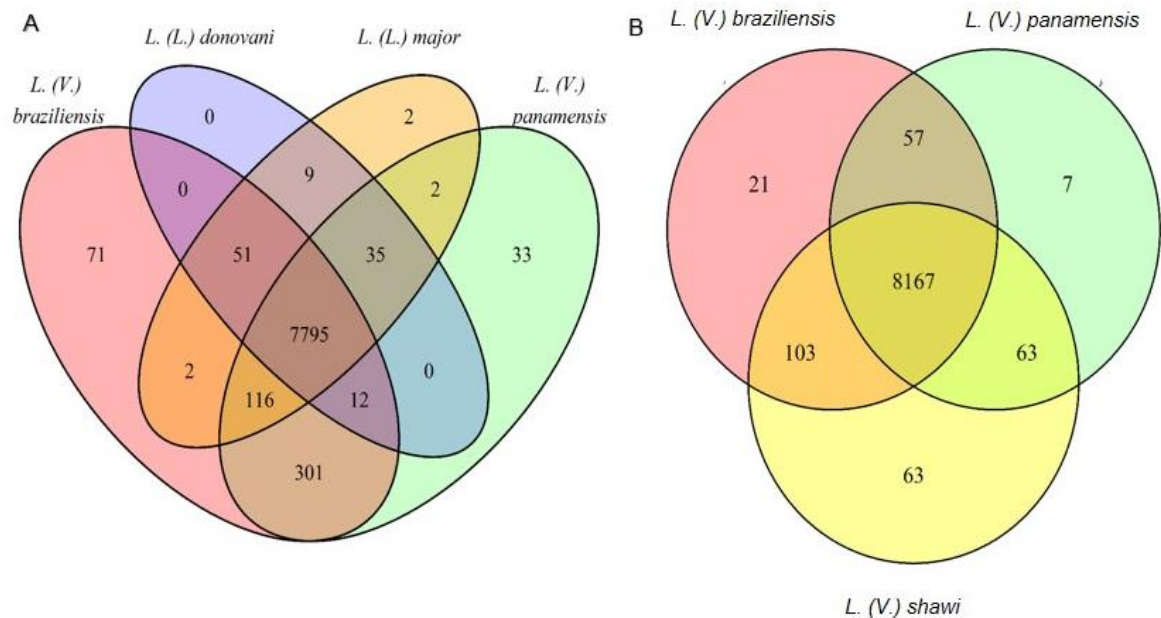


**FIGURA 4** - Identidade entre as proteínas de *L. (V.) guyanensis* e *L. (V.) shawi* em comparação com *L. (L.) donovani* (A e B), *L. (V.) guyanensis* com *L. (L.) major* (C) e *L. (V.) shawi* com *L. (L.) major* (D).

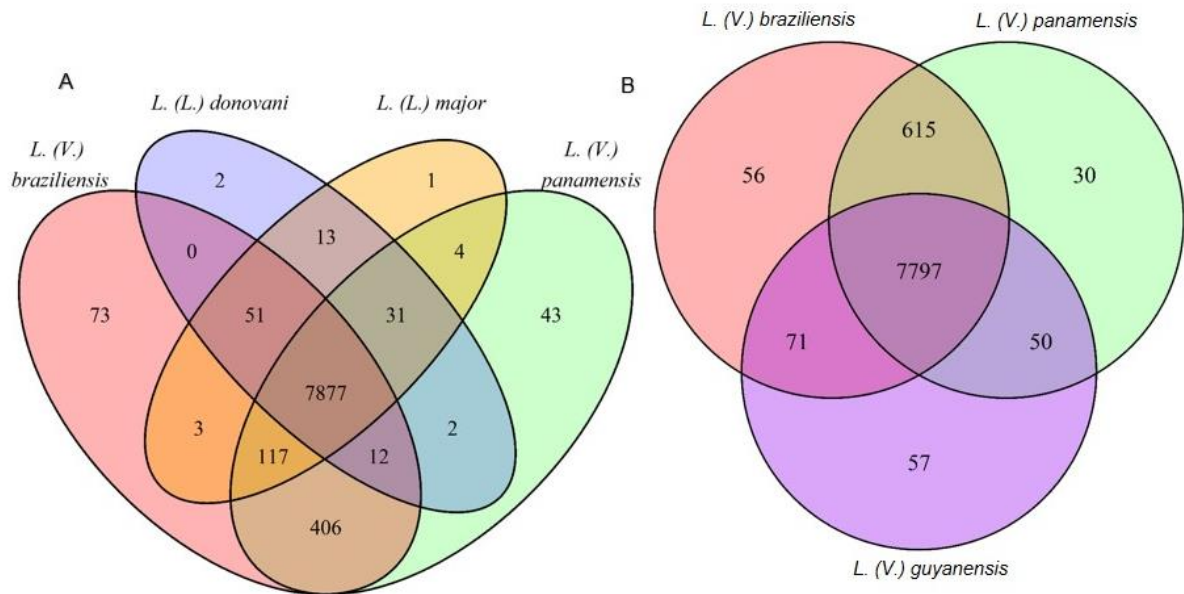
Os dados do BLASTp também foram utilizados para obtenção de diagramas de Venn, demonstrando quais proteínas de *L. (V.) guyanensis* aparecem em *L. (V.) braziliensis*, *L. (L.) donovani*, *L. (L.) major* e *L. (V.) panamensis* (Figura 5A), o total foi de 8.429 proteínas, sendo 7.795 compartilhada entre todas as espécies. O maior compartilhamento exclusivo de *L. (V.) guyanensis* foi com *L. (V.) braziliensis*, com 71 proteínas. Na comparação das espécies do subgênero *Viannia*, *L. (V.) guyanensis* com *L. (V.) shawi*, *L. (V.) braziliensis* e *L. (V.) panamensis* (Figura 5B), o total de proteínas foi de 8.481, sendo possível observar 8.167 compartilhadas entre elas, e 63 exclusivamente presentes em *L. (V.) guyanensis* e *L. (V.) shawi*.



O número de proteínas de *L. (V.) shawi* compartilhadas com *L. (V.) braziliensis*, *L. (L.) donovani*, *L. (L.) major* e *L. (V.) panamensis* (Figura 6A) chega a um total de 8.645, sendo 7.877 comum a todas; com *L. (V.) braziliensis*, observou-se exclusividade de 73 proteínas compartilhadas. Já a comparação com as espécies do subgênero *Viannia* revelou 8.676 proteínas compartilhadas, sendo 7.797 com todas as espécies avaliadas (Figura 6B). Quando se compara os dois diagramas de *Viannia* (Figura 5B e 6B), é notável a maior aproximação de *L. (V.) shawi* com *L. (V.) panamensis*, com 30 proteínas, enquanto que *L. (V.) guyanensis* com *L. (V.) panamensis* apresentam somente 7 proteínas exclusivas (Figura 5B).



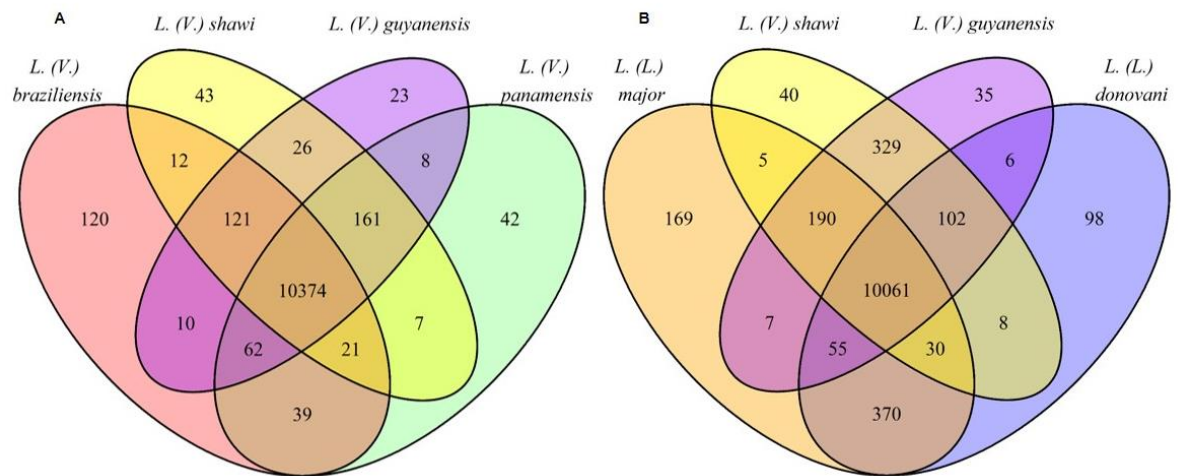
**FIGURA 5** - Compartilhamento de proteínas entre *Leishmania (V.) guyanensis* e outras espécies dos subgêneros *Leishmania* e *Viannia* (A) e entre *L. (V.) guyanensis* e outras espécies do subgênero *Viannia* (B).



**FIGURA 6** - Compartilhamento de proteínas entre *Leishmania (V.) shawi* e outras espécies dos subgêneros *Leishmania* e *Viannia* (A) e entre *L. (V.) shawi* e outras espécies do subgênero *Viannia* (B).

Os dados do RPS-BLAST também foram utilizados para construção de diagramas de Venn, para demonstrar o compartilhamento de domínios entre as espécies do subgênero *Viannia*: *L. (V.) guyanensis*, *L. (V.) shawi*, *L. (V.) panamensis* e *L. (V.) braziliensis* (Figura 7A). De um total de 11.069 domínios de proteínas conservados, 10.374 são compartilhados entre as espécies do subgênero *Vianna*, sendo *L. (V.) braziliensis* a que possui mais singularidade com seus 120 domínios, enquanto *L. (V.) guyanensis* e *L. (V.) shawi* apresentem somente 23 e 43 domínios, respectivamente.

No diagrama das espécies sequenciadas neste estudo (*L. (V.) guyanensis* e *L. (V.) shawi*), com as espécies do subgênero *Leishmania* *L. (L.) major* e *L. (L.) donovani*, o total de domínios foi de 11.505, porém o número compartilhado entre elas foi 10.061, sendo *L. (L.) major* com 169 domínios singulares enquanto que *L. (V.) guyanensis* e *L. (V.) shawi* apresentam somente 35 e 40, respectivamente.



**FIGURA 7** – Compartilhamento de domínios entre espécies de *Leishmania*. Subgênero *Viannia*: *L. (V.) braziliensis*, *L. (V.) shawi*, *L. (V.) guyanensis* e *L. (V.) panamensis* (A). *L. (L.) major*, *L. (V.) shawi*, *L. (V.) guyanensis* e *L. (L.) donovani* (B).

Quando observamos *L. (V.) guyanensis* e *L. (V.) shawi* comparadas separadamente à *L. (V.) braziliensis*, *L. (V.) panamensis*, *L. (L.) major* e *L. (L.) donovani*, mesmo em digramas distintos, nota-se números similares, pois ambas ultrapassam os 14.000 domínios compartilhados (Figura Suplementar 1). Dentre estes domínios de proteínas estudados, quatro aparecem mais de 100 vezes nas seis espécies de *Leishmania* comparadas: são quinases, com funções básicas de manutenção da célula como regulação do ciclo celular, e fosfotransferases e alguns dos domínios de proteínas presentes só em *L. (V.) shawi* e *L. (V.) guyanensis* são domínios que ocorrem também em bactérias. Foi possível observar domínios presentes somente nas espécies do subgênero *Viannia* (Tabela Suplementar 1).

## DISCUSSÃO

A realização, pela primeira vez, do sequenciamento genômico de *Leishmania (Viannia) guyanensis* e *Leishmania (Viannia) shawi*, associada à anotação de proteínas permitiu ampliar o conhecimento das espécies deste subgênero, importantes agentes da forma cutânea da leishmaniose.

A comparação da montagem dos genomas realizados neste estudo com os genomas de *Leishmania (Viannia) braziliensis* (PEACOCK et al., 2007), *Leishmania (Leishmania) amazonensis* (REAL et al., 2013) e *Leishmania (Viannia) panamensis* (LLANES et al., 2015) permite notar que há consistência no tamanho dos genomas, (em torno de 30 Mb), na porcentagem de CG (aproximadamente 57%), bem como no número de genes preditos (próximos de 8.000), confirmando a qualidade dos procedimentos adotados.

O parâmetro mais divergente foi o tamanho dos *scaffolds*, provavelmente por ter sido utilizada estratégia de biblioteca única (*paired-end*) sequenciadas em plataforma *Illumina*<sup>®</sup>, diferentemente da utilizada na montagem da *L. (L.) amazonensis* (REAL et al., 2013), onde foram hibridizados dados de duas plataformas de sequenciamento, a *Illumina*<sup>®</sup> e a *Titanium*<sup>®</sup>. Já a estratégia para montagem de *L. (V.) panamensis* (LLANES et al., 2015), além de usar estas mesmas plataformas também inclui duas bibliotecas, a *paired-end* e a *mate-pair*. Por outro lado, a estratégia de adicionarmos os dados disponíveis no reservatório público do NCBI resultou na obtenção de 3.999 *scaffolds* para *L. (V.) guyanensis* contra 3.199 no caso de *L. (L.) amazonensis*.

Além da adição dos dados externos, a escolha dos programas de bioinformática foi importante para o êxito do trabalho, a começar pelo montador SPAdes, o qual utilizou um conjunto de K-mers para diminuir as falhas decorrentes da cobertura não uniforme dos *reads* e, na intenção de melhorar ainda mais a montagem, foi utilizado o SSPACE, um extensor de *scaffolds*. Além disto, todos os programas usados na montagem são de uso livre, diminuindo assim o custo da pesquisa.

As espécies sequenciadas no presente estudo são do subgênero *Viannia* que apresenta somente os genomas de *L. (V.) braziliensis* (PEACOCK et al., 2007) e *L. (V.) panamensis* (LLANES et al., 2015) montados. A análise de similaridade das proteínas anotadas permitiu observar que o subgênero *Viannia* apresenta semelhanças entre os seus representantes e se diferencia das espécies que compõe o subgênero *Leishmania*. Assim, as sequências proteicas de *L. (V.) guyanensis* e *L. (V.) shawi*, quando comparadas entre si e com *L. (V.) braziliensis* e *L. (V.) panamensis*, mantiveram similaridade superior a 95%, enquanto que, quando comparadas com *(L.) major* e *L. (L.) infantum*, apresentaram similaridade menor (75-77%). Dados semelhantes foram observados entre *L. (V.) braziliensis* e *L. (L.) major* ou entre *L. (V.) braziliensis* e *L. (L.) infantum*, revelando 77% de similaridade, enquanto que entre *L. (L.) major* e *L. (L.) infantum*, estas do subgênero *Leishmania*, a similaridade foi de 92% (PEACOCK et al., 2007).

A análise de compartilhamento das proteínas entre estas espécies sustenta a classificação destas em dois subgêneros distintos dado que, mesmo compartilhando muitas proteínas entre si, o número de proteínas partilhadas entre as espécies do subgênero *Viannia* é maior que entre as representantes do subgênero *Leishmania*. A análise de distribuição dos grupos de genes ortólogos de *L. (V.) panamensis* realizado por Llanes et al. (2015) também revelou maior proximidade desta com *L. (V.) braziliensis* do que com as demais espécies pertencentes ao subgênero *Leishmania*, corroborando com a distinção dos subgêneros.

Além disto, quando observamos os domínios das proteínas anotadas podemos notar conservação com outras análises assim, mesmo as espécies de *Leishmania* sendo bem próximas e partilhando muito em comum, há diferenças que permitem separá-las em subgêneros distintos, por pequenos detalhes, os quais podem ser importantes para explicar as diferentes formas clínicas. As espécies representantes do subgênero *Viannia*: *L. (V.) guyanensis*, *L. (V.) shawi*, *L. (V.) braziliensis* e *L. (V.) panamensis* se mostraram mais próximas entre si com uma conservação de domínios mais singular entre elas.

## CONCLUSÃO

A comparação de duas espécies do subgênero *Viannia*, que causam leishmaniose tegumentar, com as espécies de genomas de *Leishmania* já montados e anotados permitiu elucidar as similaridades e o alto grau de conservação entre as espécies deste parasito, além de evidenciar aspectos singulares dos subgêneros que apresentam distribuição geográfica e aspectos clínicos diferentes. O sequenciamento de *L. (V.) guyanensis* e *L. (V.) shawi*, seguido da montagem e anotação de proteínas ampliaram a informação genômica destas espécies de *Leishmania*.

## AGRADECIMENTOS

À equipe do Laboratório de Bioquímica e Biologia Molecular, FMVA – UNESP Araçatuba, à equipe de Bioinformática do Laboratório Central de Tecnologias de Alto Desempenho em Ciências da Vida (LaCTAD), parte da Universidade Estadual de Campinas (UNICAMP) e a Università Cattolica Del Sacro Cuore, em Piacenza na Itália. À Capes pelo consentimento da bolsa de Mestrado.

## REFERÊNCIAS

ALTSCHUL, S. F.; MADDEN, T. L.; SCHÄFFER, A. A.; ZHANG, J.; ZHANG, Z.; MILLER, W.; LIPMAN, D. J. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. **Nucleic Acids Research**, v. 25, n. 17, p. 3389–3402, 1997.

BANKEVICH, A.; NURK, S.; ANTIPOV, D.; GUREVICH, A. A.; DVORKIN, M.; KULIKOV, A. S.; LESIN, V. M.; NIKOLENKO, S. I.; PHAM, S.; PRJIBELSKI, A. D.; PYSHKIN, A. V.; SIROTKIN, A. V.; VYAHHI, N.; TESLER, G.; ALEKSEYEV, M. a.; PEVZNER, P. a. SPAdes: A new genome assembly algorithm and its applications to single-cell sequencing. **Journal of Computational Biology**, v. 19, n. 5, p. 455–477, 2012.

BATES, P. A. Transmission of *Leishmania* metacyclic promastigotes by phlebotomine sand flies. **International Journal for Parasitology**, v. 37, n. 10, p. 1097–1106, 2007.

BOETZER, M.; HENKEI, C. V.; JANSEN, H. J.; PIROVANO, W. Scaffolding pre-assembled contigs using SSPACE. **Bioinformatics**, v. 27, n. 4, p.578-579, 2010.

BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: a flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 2014.

DESJEUX, P. Leishmaniasis: current situation and new perspectives. **Comparative Immunology, Microbiology and Infectious Diseases**, v. 27, n. 5, p. 305–318, 2004.

GOTO H.; LINDOSO, J. A. L. Current diagnosis and treatment of cutaneous and mucocutaneous leishmaniasis. **Expert Review Anti-Infective Therapy**, v. 8, n. 4, p. 419–433, 2010.

GRAMICCIA, M.; GRADONI, L. The current status of zoonotic leishmaniases and approaches to disease control. **International Journal for Parasitology**, v. 35, n. 11–12, p. 1169–1180, 2005.

GUREVICH, A.; SAVELIEV, V.; VYAHHI, N.; TESLER, G. QUASt: Quality assessment tool for genome assemblies. **Bioinformatics**, v. 29, n. 8, p. 1072–1075, 2013.

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, p. 1-20, 1987.

LANGMEAD, B.; TRAPNELL, C.; POP, M.; SALZBERG, S. Ultrafast and memory-

efficient alignment of short DNA sequences to the human genome. **Genome Biol**, [s.l.], v. 10, n. 3, p.25-34, 2009. Springer Nature.

LLANES, A.; RESTREPO, C. M.; DEL VECCHIO, G.; ANGUIZOLA, F. J.; LLEONART, R. The genome of *Leishmania panamensis*: insights into genomics of the L. (Viannia) subgenus. **Scientific reports**, v. 5, p. 8550, 2015.

MARCHLER-BAUER, A.; ZHENG, C.; CHITSAZ, F.; DERBYSHIRE, M. K.; GEER, L. Y.; GEER, R. C.; GONZALES, N. R.; GWADZ, M.; HURWITZ, D. I.; LANCZYCKI, C. J.; LU, F.; LU, S.; MARCHLER, G. H.; SONG, J. S.; THANKI, N.; YAMASHITA, R. A.; ZHANG, D.; BRYANT, S. H. CDD: Conserved domains and protein three-dimensional structure. **Nucleic Acids Research**, v. 41, n. D1, p. 348–352, 2013.

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

NOYES, H.; PRATLONG, F.; CHANCE, M.; ELLIS, J.; LANOTTE, G.; DEDET, J. P. A previously unclassified trypanosomatid responsible for human cutaneous lesions in Martinique ( French West Indies ) is the most divergent member of the genus *Leishmania* ss. **Parasitology**, p. 17–24, 2002.

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. O.; 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, n. 7, p. 839–47, 2007.

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.

SHAW, J. J.; ISHIKAWA, E. A.; LAINSON, R.; BRAGA, R. R.; SILVEIRA, F.T. Cutaneous leishmaniasis of man due to *Leishmania (Viannia) shawi* Lainson, de Souza, Povoá, Ishikawa & Silveira, in Para State, Brazil. **Ann Parasitol Hum Comp**. 66, 243–246, 1991.

SLATER, G. S. C.; BIRNEY, E. Automated generation of heuristics for biological



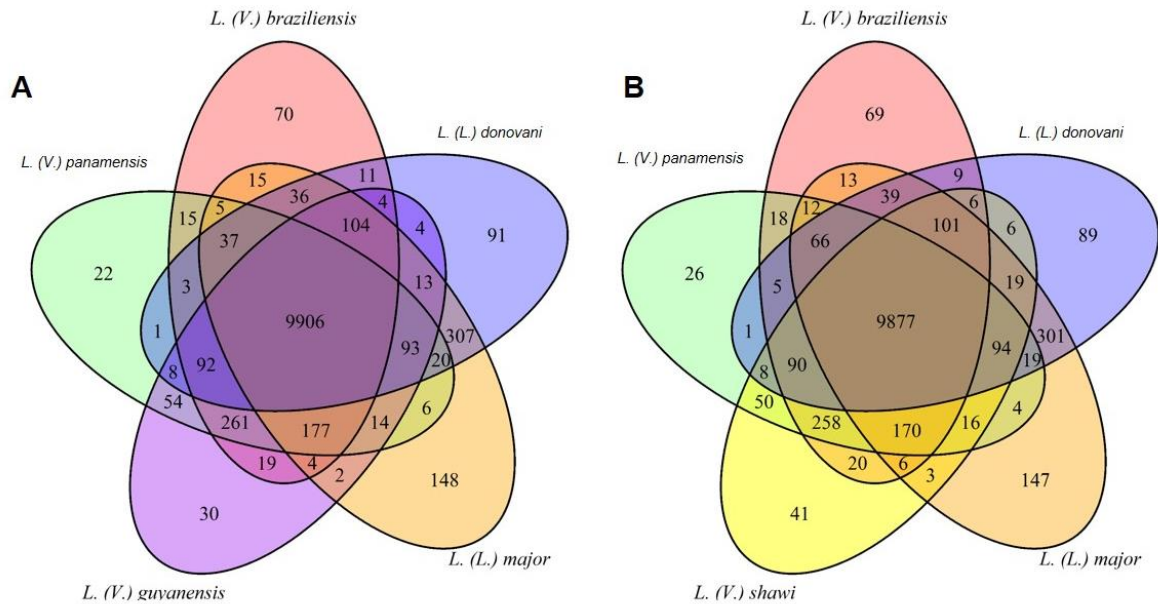
sequence comparison. **BMC bioinformatics**, v. 6, p. 31, 2005.

STANKE, M.; MORGENSTERN, B. AUGUSTUS: A web server for gene prediction in eukaryotes that allows user-defined constraints. **Nucleic Acids Research**, v. 33, n. SUPPL. 2, p. 465–467, 2005.

WARREN, R. L.; SUTTON, G. G.; JONES, S. J. M.; HOLT, R. A. Assembling millions of short DNA sequences using SSAKE. **Bioinformatics**, v. 23, n. 4, p. 500–501, 2007.

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.

## APÊNDICES



**FIGURA SUPLEMENTAR 1** - Compartilhamento de proteínas entre espécies de *Leishmania*. *L. (V.) guyanensis* e outras espécies: *L. (V.) braziliensis*, *L. (V.) panamensis*, *L. (L.) major*, e *L. (L.) donovani*. (A). *L. (V.) shawi* e outras espécies *L. (V.) guyanensis*, *L. (V.) braziliensis*, *L. (V.) panamensis*, *L. (L.) major* e *L. (L.) donovani* (B).

**Tabela Suplementar 1** – Número de alguns domínios de proteínas selecionadas por singularidades únicas ou abundância em todas as espécies

Subgêneros	<i>L. Viannia</i>				<i>L. Leishamnia</i>		ID CDD
CDD_nome	<i>guyanensis</i>	<i>shawi</i>	<i>panamensis</i>	<i>braziliensis</i>	<i>major</i>	<i>donovani</i>	CDDnum
2-desacetyl-2-hydroxyethyl_bacteriochlorophyllide_like	0	1	0	0	0	0	176217
APC10-ZZEF1	1	1	1	1	0	0	176488
Borrelia_P83	0	1	0	0	0	1	114011
C2_Dock-C	1	1	1	1	0	0	176078
C2_Dock-D	1	0	1	1	0	0	176079
C2B_RasGAP	1	0	1	1	0	0	176057
Cyt_b561	1	1	0	1	0	0	176489
DinB_2	0	1	0	0	0	0	289624
dTDP_HR_like_SDR_e	1	0	0	0	0	0	187564
EZ_HEAT	1	1	1	1	0	0	128837
Gag_spuma	0	1	0	0	0	0	281293
GDPD_GDE_2_3_6	1	0	0	0	1	1	176516
GH18_CFLE_spore_hydrolase	1	1	0	0	0	0	119353
Glutenin_hmw	0	1	0	0	0	0	281191
GST_C_2	0	1	0	0	0	0	290146
KaiC	0	1	0	0	0	0	284219
LbH_eIF2B_epsilon	0	1	1	1	1	1	100061
MDR_like	0	1	0	2	0	0	176204
miaA	1	1	1	1	0	0	173191
minC	1	1	1	0	0	0	178806
nitrilase_4	1	1	1	1	0	0	143606

BBP1_ABC_transporter_LIVBP_like	1	1	1	1	0	0	107263
Peptidases_S8_Subtilisin_like_2	1	0	1	0	0	0	173793
PHA02736	1	1	1	1	0	0	165103
PHA02859	1	1	1	1	0	0	165195
PHA02884	1	1	1	1	0	0	165212
PI-PLCc_beta1	1	1	1	0	0	0	176560
Pkinase	172	178	168	167	174	166	278497
PLCYc	1	2	1	1	0	0	128454
PLN02152	1	1	1	1	0	0	177813
PLN02165	1	1	1	1	0	0	177823
PLN02207	1	1	1	1	0	0	177857
PLN02555	1	1	1	1	0	0	178170
PLN02795	1	1	0	1	0	0	178392
PLN02992	1	1	1	1	0	0	178572
PLN03015	1	1	1	1	0	0	178589
PLN03212	1	1	1	1	0	0	178751
PRK03515	1	1	0	0	0	0	179587
PRK05870	1	1	1	1	0	0	180298
PRK06143	1	0	0	0	0	0	180423
PRK06202	1	1	1	1	0	0	180466
PRK06475	1	1	1	1	0	0	180582
PRK07246	0	1	1	1	1	1	180905
PRK07511	1	0	1	1	0	0	181009
PRK07594	1	1	1	1	0	0	136438
PRK07786	1	0	0	0	0	0	169098
PRK08974	0	1	0	0	0	0	236359

PRK12492	0	1	0	0	0	0	171539
PRK12705	0	1	0	0	0	0	237178
PRK13391	0	1	0	0	0	0	184022
PRK14664	1	1	1	1	0	0	173127
PTZ00449	0	1	0	0	0	1	185628
Rph	1	1	1	1	0	0	178914
rps2	1	1	2	1	0	0	177007
S_TKc	189	195	185	185	189	185	214567
S1_Tex	0	1	0	0	0	0	240190
sdhA	0	1	0	0	0	0	236319
secoisolariciresinol-DH_like_SDR_c	1	1	0	0	0	0	187587
SIN1_PH	0	1	0	0	0	0	293584
Smc	147	152	144	151	160	145	224117
START_STARD12-like	1	1	1	1	0	0	176917
START_STARD5-like	1	1	1	1	0	0	176912
START_STARD6-like	1	1	1	1	0	0	176913
STKc_CAMK	109	111	104	104	107	101	270687
ttLC_FACS_AIKK_like	1	0	0	0	0	0	213327

---