



**INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO PIAUÍ
CAMPUS PEDRO II
CURSO CIÊNCIAS BIOLÓGICAS**

JOSÉ ÍTALO SILVA NASCIMENTO

**IDENTIFICAÇÃO E VALIDAÇÃO IN SILICO DE POTENCIAIS MARCADORES
MOLECULARES PARA O GÊNERO LEISHMANIA**

PEDRO II

2021

JOSÉ ÍTALO SILVA NASCIMENTO

IDENTIFICAÇÃO E VALIDAÇÃO IN SILICO DE POTENCIAIS MARCADORES
MOLECULARES PARA O GÊNERO LEISHMANIA

Trabalho de Conclusão de Curso
apresentado como exigência parcial para
obtenção do diploma do Curso de Ciências
biológicas do Instituto Federal de Educação,
Ciência e Tecnologia do Piauí. Campus
Pedro II.

Orientador: Prof. Dr. Willame Rodrigues do
Nascimento Sousa.

PEDRO II

2021

Dados Internacionais de Catalogação na Publicação (CIP) de acordo com ISBD

Nascimento, José Ítalo Silva

N244i Identificação e validação in silico de potenciais marcadores moleculares para o gênero Leishmania / José Ítalo Silva Nascimento. - 2021.

59 f.: il. color.

Trabalho de Conclusão de Curso (Licenciatura em Ciências Biológicas) Instituto Federal do Piauí, Campus Pedro II, 2021.

Orientador : Prof. Dr. Willame Rodrigues do N. Sousa.

1. Leishmaniose - diagnóstico molecular . 2. Genética - marcadores moleculares . 3. Bioinformática . I.Título.

CDD - 570

Elaborado por Wirllanna Naira da Silva Torres CRB 3/1088

FOLHA DE APROVAÇÃO

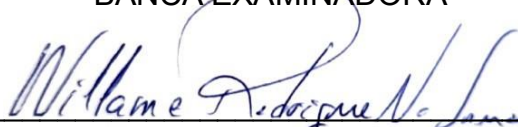
JOSÉ ÍTALO SILVA NASCIMENTO

IDENTIFICAÇÃO E VALIDAÇÃO IN SILICO DE POTENCIAIS MARCADORES MOLECULARES PARA O GÊNERO LEISHMANIA

Trabalho de Conclusão de Curso
(monografia) apresentado como exigência
parcial para obtenção do diploma do Curso
de Ciências Biológicas do Instituto Federal
de Educação, Ciência e Tecnologia do Piauí.
Campus Pedro II -PI.

Aprovada em: 26 / 04 / 2021.

BANCA EXAMINADORA



Prof. Dr. Willame Rodrigues do Nascimento Sousa (Orientador)
Instituto Federal do Piauí (IFPI)



Dr. Geice Ribeiro da Silva
Universidade Federal do Piauí - UFPI



Me. Raimundo Leoberto Torres de Sousa
Fio Cruz-PI

Dedico este trabalho e
minha vida a minha família, por sua capacidade de acreditar em me e investir
em mim. Mãe, seu cuidado e dedicação foi que deu, em alguns momentos, a
esperança para seguir. Pai, sua presença significou segurança e certeza de
que não estou sozinho nessa caminhada.

AGRADECIMENTOS

Começo por agradecer a Deus por, ao longo deste processo complicado e desgastante, me ter feito ver o caminho, nos momentos em que pensei em desistir.

Quero agradecer a minha mãe em que todos os momentos difíceis sempre estiveram ao meu lado ao meu pai que me apoio em tudo na minha vida, ao meu irmão por estar presente em minha vida.

Ao meu orientador o professor Dr. Willame Rodrigues por me auxiliar nesta jornada.

Aos meus familiares que de alguma forma me ajudou com palavras de incentivo e motivação.

Aos meus amigos e colegas de curso que sempre estiveram ao meu lado na correria das aulas na formação e apresentação de seminários e entrega de trabalho.

E por fim, agradecer a Instituição que me recebeu de braços abertos, desde da minha entrada na Instituição todos servidores sempre a disposição a ajudar.

“O homem não teria alcançado o possível se, repetidas vezes, não tivesse tentado o impossível.”
Max Weber

RESUMO

As leishmanioses englobam um largo espectro de doenças causadas pelo parasita do gênero *Leishmania*. Métodos moleculares estão sendo cada vez mais empregados para diagnóstico, estudos taxonômicos, filogenéticos e epidemiológicos envolvendo este parasita. O desenvolvimento de marcadores moleculares vem se tornando fundamental para diagnósticos mais precisos na identificação de patógenos de interesse ao homem. O objetivo deste trabalho foi desenvolver marcadores moleculares *in silico* para *Leishmania* e avaliar sua aplicabilidade o que foi desenvolvido utilizando programas de bioinformática gratuitos disponíveis na internet. Foram selecionadas 11 sequências do NCBI para *Leishmania infantum* e o complexo *Leishmania donovani*, no qual foi realizado um alinhamento múltiplo com a obtenção de uma sequência consenso. Identificada as regiões conservadas, foi definido a região do primer a ser desenvolvido. Foi desenvolvido três primers, levando em conta o número de pares de bases, a porcentagem de guanina e citosina, que variou de 37,5 a 47,62 e temperatura de Melting que apresenta um intervalo de 56,48°C a 59,08°C. Os primers desenvolvidos apresentam uma taxa de validação de 85% a 100%. A PCR-RFLP *in silico*, mostrou-se adequada para a identificação de *Leishmania* a nível de espécie. A variação genética nas sequências em mini círculos de kDNA entre os isolados *L. donovani* e *L. infantum*, pode ser detectado através dos SNP presentes nos respectivos genomas. A enzima de restrição *RsaI* foi considerada ideal para determinar a variação de bandas entre as duas espécies.

Palavras-chave: Bioinformática, leishmanioses, primers, enzima de restrição.

ABSTRACT

Leishmaniasis encompasses a wide spectrum of diseases caused by the parasite of the genus *Leishmania*. Molecular methods are being increasingly used for diagnosis, taxonomic, phylogenetic and epidemiological studies involving this parasite. The development of molecular markers has become essential for more accurate diagnoses in the identification of pathogens of interest to man. The objective of this work was to develop molecular markers in silico for *Leishmania* and evaluate its applicability which was developed using free bioinformatics programs available on the internet. Eleven NCBI sequences were selected for *Leishmania infantum* and *Leishmania donovani*, in which a multiple alignment was performed with the achievement of a consensus sequence. Once the conserved regions were identified, the region of the primer to be developed was defined. Three primers were developed, taking into account the number of base pairs, the percentage of guanine and cytosine, which ranged from 37,5% to 47,62% and Melting temperature that has a range of 56,48°C to 59,08°C. The developed primers have a validation rate of 85% to 100%. The PCR-RFLP in silico, proved to be adequate for the identification of *Leishmania* at the species level. The genetic variation in sequences in mini kDNA circles between *L. donovani* and *L. infantum* isolates, can be detected through the SNPs present in the respective genomes. The restriction enzyme *RsaI* was considered ideal to determine the band variation between the two species.

Keywords: Bioinformatics, leishmaniasis, primers, restriction enzyme

LISTA DE FIGURAS

Figura 1 – Situação da endemicidade da Leishmaniose Cutânea (LC) no mundo, 2018.....	19
Figura 2 – Situação da endemicidade da Leishmaniose Visceral (LV) no mundo, 2018.....	20
Figura 3 – <i>Lutzomyia longipalpis</i> fazendo o repasto sanguíneo em ser humano.....	21
Figura 4 – <i>Lutzomyia cruzi</i>	21
Figura 5 – Casos de LV por UF de infecção. Brasil, 2019.....	22
Figura 6 – Classificação atualizada de espécies de <i>Leishmania</i> patogênicas para humanos. As espécies conhecidas por causar a forma visceral da doença estão destacadas em negrito.....	24
Figura 7 – Morfologia da <i>Leishmania</i> nas formas promastigota e amastigota.....	25
Figura 8 – Forma promastigota de <i>Leishmania</i> sp. com a localização do kDNA, na porção anterior.....	27
Figura 9 – Minicírculo de kDNA de <i>Leishmania</i> sp. com seus componentes. BSC: Bloco de Sequência Conservada.....	28
Figura 10 – Organização das fibras de DNA do cinetoplasto em maxicírculos e minicírculos em <i>L. tarentolae</i>	30
Figura 11 – Sequencia realizada pelo NCBI, selecionando as 11 primeiras sequencia.....	43
Figura 12 – Validação do primeiro par de primers observando os parâmetros de T _m , GC%, formação de harpin, heterodímeros, self dímeros e ΔG	44
Figura 13 – Validação do segundo par de primers observando os parâmetros de T _m , GC%, formação de harpin, heterodímeros, self dímeros e ΔG	45
Figura 14 – Análise dos valores de ΔG na formação de self dímeros.....	46

Figura 15 – Análise de Valor de ΔG na formação de harpin.....	46
Figura 16 – Análise do Valor de ΔG na formação de heterodímeros.....	46
Figura 17 – Múltiplas sequências alinhadas pelo BioEdit e posicionamento do primer ITA1-F.....	47
Figura 18 – Múltiplas sequências alinhadas pelo BioEdit e posicionamento do primer ITA2-R1.....	47
Figura 19 - Múltiplas sequências alinhadas pelo BioEdit e posicionamento do primer ITA2-R2.....	47
Figura 20 - Sequência consenso obtido a partir do alinhamento de 11 acessos pelo BioEdit.....	48
Figura 21 - Sequência consenso de <i>L. infantum</i>	48
Figura 22 - Sequência consenso de <i>L. donovani</i>	49

LISTA DE TABELAS

Tabela 1 - Distribuição dos casos de Leishmaniose Tegumentar e Visceral nos municípios piauienses com maiores registros de casos, entre 2015 e 2019....22

Tabela 2 - Sequências utilizadas para o desenho e validação de primers (NCBI).....43

Tabela 3 - Sequência dos primers desenvolvidos para a detecção das espécies em estudo.....44

Tabela 4 - Descrição dos primers, sequências, tamanho do Amplicon e enzima para o genoma de *L. infantum*.....49

Tabela 5 - Descrição dos primers, sequências, tamanho do Amplicon e enzima para o genoma de *L. donovani*.....50

LISTA DE ABREVIATURAS E SIGLAS

BIOINFORM - Serviço global de notícias sobre bioinformática

BLAST - Basic Local Alignment Search Tool

DNA - Ácido Desoxirribonucleico

EBI - Instituto Europeu de Bioinformática

ENSEMBL - Banco de dados de anotação automática sobre genomas

GENBANK - Banco de dados de sequências genéticas, uma coleção anotada de todas as sequências de DNA disponíveis ao público

LC - Leishmania cutânea

LTA - Leishmania Tegumentar Americana

LV - Leishmania Visceral

NCBI - Centro Nacional de Informações sobre Biotecnologia

PCR - Reação em cadeia da polimerase

RFLP - Polimorfismo de Comprimento de Fragmentos de Restrição

RNA - Ácido ribonucleico

SNP - Polimorfismo de nucleotídeo simples

SUMÁRIO

1. INTRODUÇÃO	15
2. OBJETIVOS.....	18
2.1 OBJETIVO GERAL	18
2.2 OBJETIVO ESPECÍFICO.....	18
3. REFERENCIAL TEORICO.....	19
3.1 LEISHMANIA NO BRASIL E NO MUNDO.....	19
3.2 LEISHMANIOSE E SEUS TIPOS.....	23
3.3 KDNA.-DNA de cinetoplasto (DNA mitocondrial).....	26
3.4 MARCADORES MOLECULARES.....	30
3.5 RFLP (RESTRICTION FRAGMENT LENGHT POLIMORPHISM) ..	33
3.6 BIOINFORMATICA.....	34
3.7 PCR <i>IN SILICO</i>	36
4. MATERIAIS E MÉTODOS.....	39
5. RESULTADOS E DISCURSSÕES.....	42
6. CONCLUSÃO.....	52
7. REFERENCIAS.....	54

1. INTRODUÇÃO

As leishmanioses são doenças que acometem o homem e várias espécies de animais, podendo ser domésticos ou silvestres. São causadas por protozoários pertencentes à família Trypanosomatidae, gênero *Leishmania*. Esses protozoários apresentam-se na forma promastigota ou amastigota, sendo que os promastigotas se replicam no inseto vetor e, as amastigotas são as formas infectivas, presentes em macrófagos do hospedeiro vertebrado (SOARES-BEZERRA et al., 2004). A leishmaniose vem sendo considerada como uma das doenças em que há bastante ocorrência no Brasil. Com relação à Leishmaniose Visceral, estima-se que 90% dos casos registrados na América Latina, ocorram no Brasil, sendo que a maioria deles em crianças (DNDI, 2013).

A principal forma de transmissão do parasita ao hospedeiro vertebrado ocorre pela picada do inseto fêmea. Ao picar o vertebrado infectado, ingere com o sangue, com as formas amastigotas, presente nas células. No trato digestivo dos insetos as promastigostas ainda não são infectivas, elas passam pela fase de transição de paramastigotas para promastigotas procíclicas então após chegarem na região anterior do inseto é que atingem a metaciclogênese tornando-se infectante. Estas formas infectadas se multiplicam e, são inoculadas por meio do repasto sanguíneo, sobre o novo hospedeiro vertebrado não infectado. (BRASIL, 2006; CAMARGO-NEVES, 2006).

Uma das razões que provavelmente dificultam avanços nestas áreas, encontra-se na diversidade do próprio parasitas, reflexo de sua plasticidade genética. (DUJARDIN et al., 2007; ROGERS et al., 2011). Trata-se de um táxon com pelo menos 21 espécies de *Leishmania*, das mais de 30 já descritas, que foram apontadas como infectantes para humanos (WHO, 2010). Os parasitas podem ser transmitidos por diferentes espécies (COURA, 2005) e populações (FERREIRA et al., 2012) de vetores, insetos flebotomíneos, apresentando ciclos de transmissão complexos, envolvendo hospedeiros/reservatórios de diferentes ordens de mamíferos (COURA, 2005).

Sendo assim, o estudo da variabilidade genética do parasita, aparece como um elemento importante, para compreender os mecanismos que levam à complexidade clínica e epidemiológica.

O nível de discriminação da tipagem, baseada em reação em cadeia da polimerase (PCR) depende do marcador molecular utilizado. Dessa forma, identificar regiões alvo adequadas para o anelamento do primer, é um passo crucial, pois essas regiões devem ser conservadas dentro do táxon alvo e, diferir em táxons relacionados (LUCCHI et al., 2013; WONG et al., 2014). Os recentes avanços na tecnologia de sequenciamento, permitem a realização de estudos genéticos a um custo significativamente mais baixo, mesmo para organismos não modelo. A exploração dos dados podem melhorar os métodos de detecção de infecções assintomáticas com baixas cargas parasitárias, uma vez que facilita a busca por genes multi-cópias, e a diferenciação de espécies, ao identificar sequencias específicas (WONG et al., 2014).

Para uma classificação apropriada dos isolados deste gênero, necessitam-se de ferramentas moleculares e, a identificação da espécie causadora da doença, é muito importante. Com a popularização da PCR, muitas abordagens moleculares vêm sendo utilizadas, com o objetivo de identificar espécies e cepas de *Leishmania* (DA SILVA et al., 2010; GRACA et al., 2012). O diagnóstico preciso da leishmaniose é essencial para o controle da doença, pois representa um ganho em sensibilidade e especificidade de diagnóstico, e é muito útil, sobretudo nos casos não resolvidos por métodos clássicos (GOMES et al., 2007). O conhecimento da estrutura genética é de grande importância para o desenvolvimento de *primers* exclusivos nestes organismos, pois permite a correlação entre a distribuição geográfica e padrões filogenéticos com a dinâmica de transmissão e epidemiologia dessa doença.

É provável que a variabilidade existente nestes parasitas, favoreçam a sobrevivência em sistemas ecológicos diversos e, determine a distribuição das formas clínicas observadas da doença (CALVOPINA et al., 2004; LUCAS et al., 1998). Distinguir as diferentes espécies de *Leishmania* e compreender sua heterogeneidade, podem contribuir para o diagnóstico e prognóstico da doença, assim como para a tomada de decisões referentes ao tratamento e combate à doença.

Diante do exposto, acredita-se que o desenvolvimento de marcadores moleculares, que permitam tanto a identificação de isolados de *Leishmania*, quanto o estudo da diversidade do gênero, e que funcionem também como ferramenta epidemiológica, representam uma contribuição valiosa para o estudo, controle e diagnóstico das leishmanioses.

2. OBJETIVOS

2.1 OBJETIVO GERAL

Identificar marcadores moleculares em potencial *in silico*, através das sequências armazenadas no NCBI (Centro Nacional de Informação e Biotecnologia), para identificação e diferenciação da *Leishmania infantum* de *Leishmania donovani*.

2.2 OBJETIVOS ESPECIFICOS

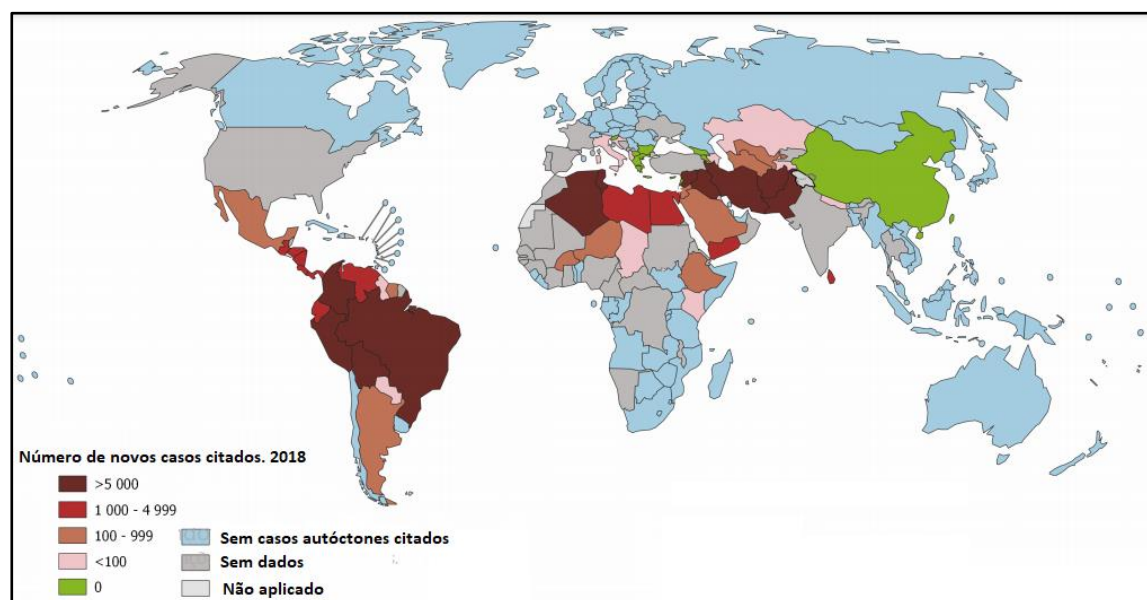
- Localizar sequências de DNA (ácido desoxirribonucleico) do genoma de cinetoplastos (kDNA) de *Leishmania infantum* e *Leishmania donovani*, depositadas no NCBI (National Center of Biotechnology Information);
- Identificar regiões conservadas e variáveis no minicírculo do genoma kDNA de *Leishmania sp*;
- Determinar nas sequências em estudo, as regiões polimórficas que possam diferenciar *Leishmania infantum* de *Leishmania donovani*;
- Desenvolver e validar os novos marcadores moleculares *in silico*;
- Comparar os diferentes fragmentos gerados (polimorfismo), a partir do padrão de clivagem das enzimas de restrição, devido às variações nucleotídicas por SNP (Polimorfismo de nucleotídeo simples) entre as espécies.

3. REFERENCIAL TEÓRICO

3.1 Leishmaniose no Brasil e no mundo

Sua presença afeta principalmente países de baixa renda em regiões tropicais e subtropicais está diretamente ligada à pobreza, mas fatores sociais, ambientais e climatológicos, também influenciam diretamente na epidemiologia da doença (DE RUITER et al., 2014).

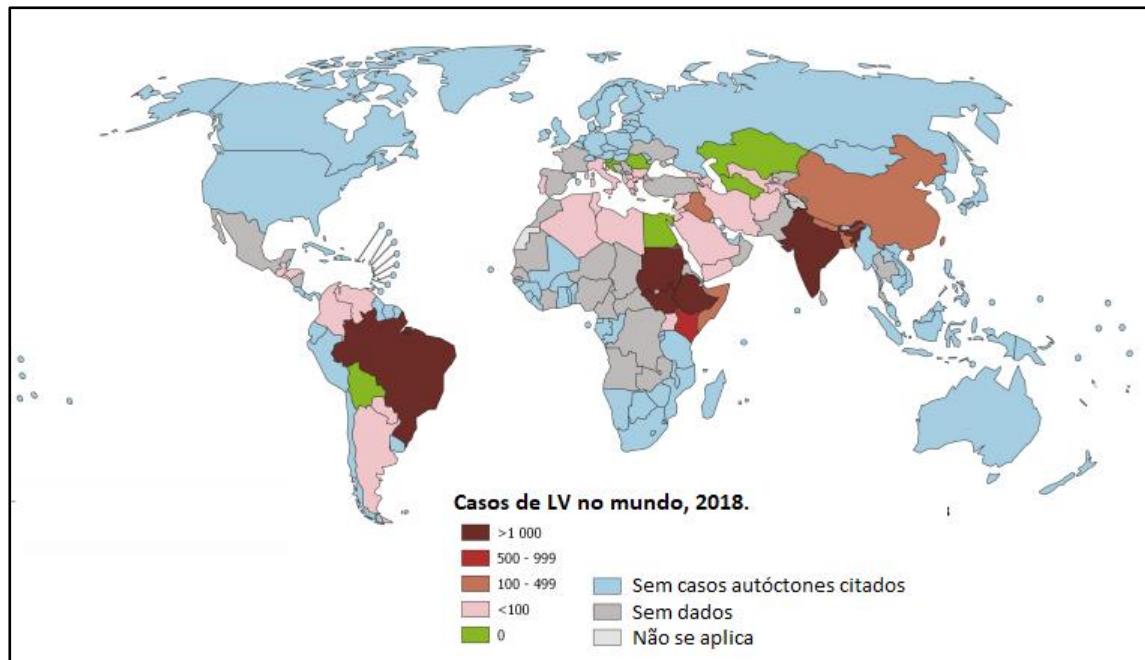
Existem aproximadamente 350 milhões de pessoas em risco e, mais de 12 milhões de casos, com uma incidência anual estimada em todo o mundo de 700.000 a 1.200.000 de casos de leishmaniose cutânea (LC) (Figura 1) e, 200 a 400 mil casos de leishmaniose visceral (LV) (Figura 2) (ALVAR et al., 2012).



FONTE: <https://www.who.int/leishmaniasis/burden/en/>

Figura 1 – Situação da endemidade da Leishmaniose Cutânea (LC) no mundo, 2018.

Portanto, as leishmanioses constituem um importante problema de saúde pública, com crescente aumento na última década (BERN et al., 2008). Na América Latina, com relação à Leishmaniose Visceral, estima-se que 90% dos casos registrados ocorram no Brasil, sendo a maioria em crianças (DNDI, 2013).



FONTE: <https://www.who.int/leishmaniasis/burden/en/>

Figura 2 – Situação da endemicidade da Leishmaniose Visceral (LV) no mundo, 2018.

As leishmanioses representam um grupo de doenças com distribuição mundial e, um amplo espectro de manifestações clínicas, constituindo um importante problema de saúde (OMS, 2018). Com isso, faz-se necessário um diagnóstico amplo, para identificar o parasito, devido à variabilidade do táxon.

No Brasil, duas espécies de insetos, até o momento, estão relacionadas com a transmissão da doença: *Lutzomyia longipalpis* e *Lutzomyia cruzi* (Figura 3). A primeira é considerada a principal espécie transmissora da *L. infantum chagasi* no Brasil e, recentemente o flebotomíneo *Lutzomyia cruzi* foi definido como vetor e, é mais restrito a região Centro-Oeste e em alguns municípios de Mato Grosso do Sul em Corumbá e Camapuã. (FALCÃO et al., 20217).

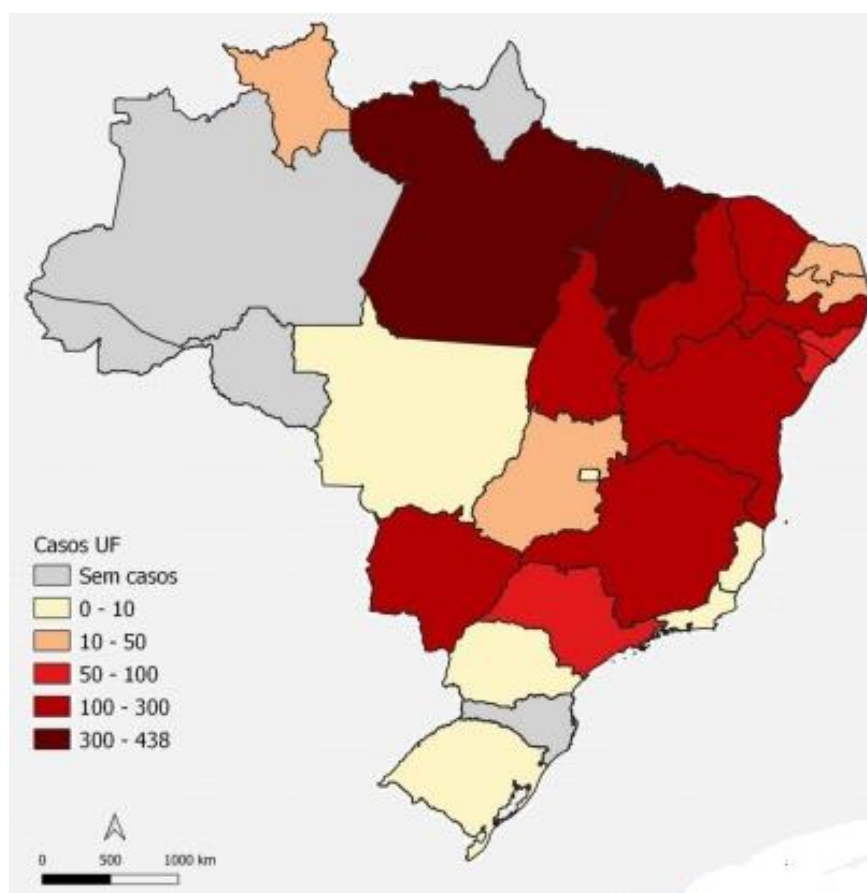
Os canídeos selvagens e cães domésticos, são considerados os principais reservatórios que atuam no ciclo zoonótico na região nordeste, sendo a região mais afetada, com 50% dos 3 a 4 mil casos anuais do país. A doença, que era tipicamente rural, a partir de 1980 passou a ser responsável por epidemias urbanas (COSTA et al., 1990). O homem se torna hospedeiro quando entra acidentalmente no ciclo silvestre do protozoário, em regiões próximas de mata fechada ou florestas (ALMEIDA et al., 2018).



FONTE: <https://www.bio.fiocruz.br>

Figura 3 – *Lutzomyia longipalpis* fazendo o repasto sanguíneo em ser humano.

O Brasil é dos quatro países com o maior número de casos de LV, representando 14% de casos no mundo e 90% no continente americano, de acordo com dados da Secretaria de Vigilância e Saúde (SVS), do Ministério da Saúde (Figura 5). Entre 2010 a 2019, o país apresentou de 3 mil a 3,5 mil casos por ano (quadro abaixo), com letalidade (mortalidade) em torno de 7%. A região Nordeste, associado aos fatores sociais e ambientais, concentra as maiores áreas de incidência da LV do Brasil (CRUZ et al., 2016).



FONTE: SVS/MS

Figura 5 – Casos de LV por UF de infecção. Brasil, 2019

No estado do Piauí, no período de 2015 a 2019, foram notificados 958 casos de Leishmaniose Visceral e 245 casos de Leishmaniose Tegumentar Americana (Quadro 1). Dos casos notificados de LV, os municípios que registraram os maiores números foram Teresina, Floriano e Parnaíba. Os casos de LTA, os municípios de Teresina, Pedro II e Luzilândia registraram as maiores incidências no estado.

Tabela 1. Distribuição dos casos de Leishmaniose Tegumentar e Visceral nos municípios piauienses com maiores registros de casos, entre 2015 e 2019.

	LTA	TOTAL	LV	TOTAL
	Municípios	245	Município	958
1	Teresina	133	Teresina	778
2	Pedro II	26	Floriano	41
3	Luzilândia	10	Parnaíba	24

4	Altos	8	Picos	15
5	Barras	8	SRN	12
6	Uruçuí	6	Piripiri	7
7	Inhuma	4	Bom Jesus	7
8	Floriano	4	Pedro II	5

Fonte: SINAN/DATASUS

3.2 - Leishmaniose e seus tipos

A estrutura taxonômica do gênero *Leishmania* é convencionalmente organizada nos níveis de subgênero, complexo de espécies e subespécies (HARKINS et al. , 2016). A parte mais bem caracterizada do gênero *Leishmania* é a seção *Euleishmania*, composta por pelo menos quatro subgêneros: *Leishmania*, *Viannia* , *Sauroleishmania* e *Mundinia* (Figura 6).

O subgênero *Leishmania* possui quatro complexos de espécies principais: *L. donovani*, *L. major*, *L. mexicana* e *L. tropica* (AKHOUNDI et al., 2016).

A diversidade genética observada neste grupo de parasitas é relatada em diversos estudos (BOITE et al., 2012; CUPOLILLO et al., 1997; CUPOLILLO et al., 1998) As formas clínicas da leishmaniose são particularmente diversas, com diferentes fenótipos relacionados: leishmaniose visceral (LV), que geralmente é fatal se não for tratada; leishmaniose mucocutânea (LMC), é caracterizada por lesões mucosas agressivas que afetam as regiões nasofaríngeas; Leishmaniose cutânea difusa (LCD), uma doença caracterizada pelo aparecimento de lesões nodulares não-ulceradas; leishmaniose cutânea (LC), caracterizadas por lesões ulcerosas, indolores, únicas ou múltiplas. (RANGEL&LAISON 2003; WHO 2018).

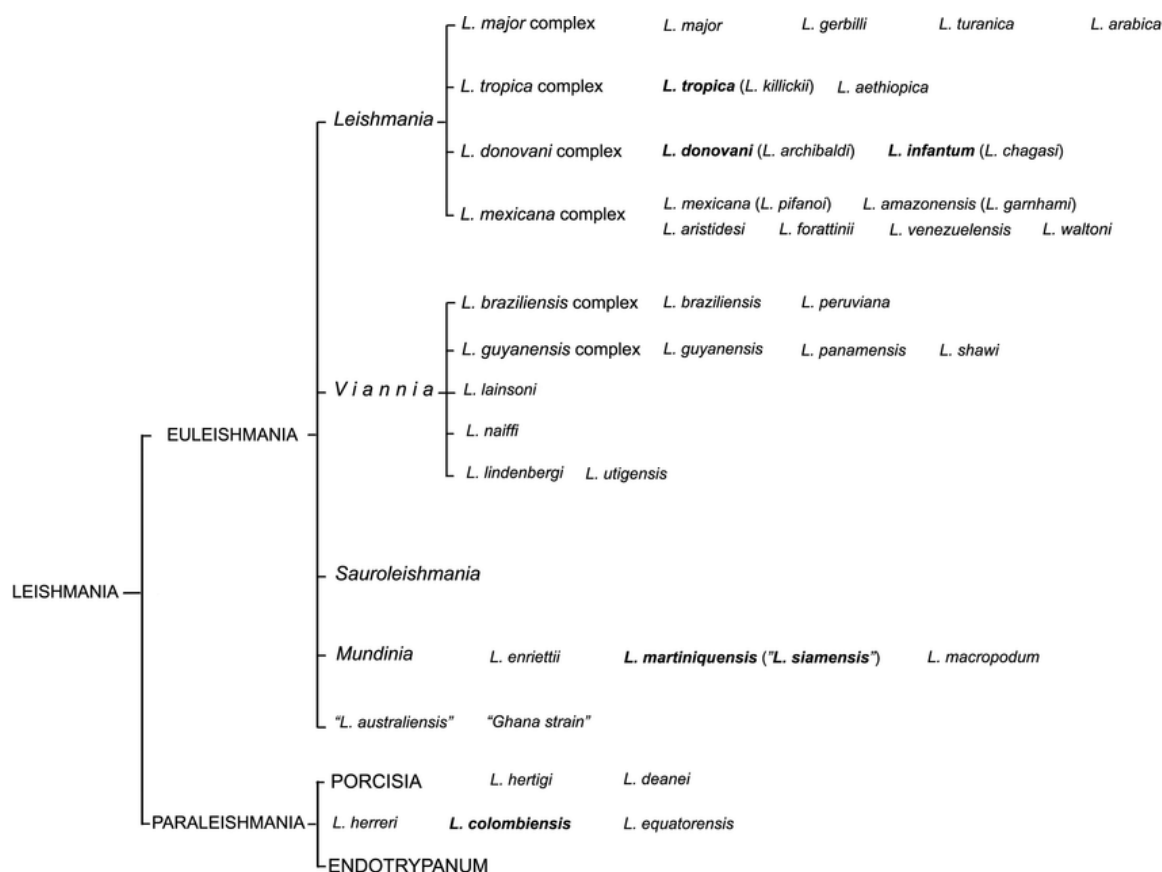


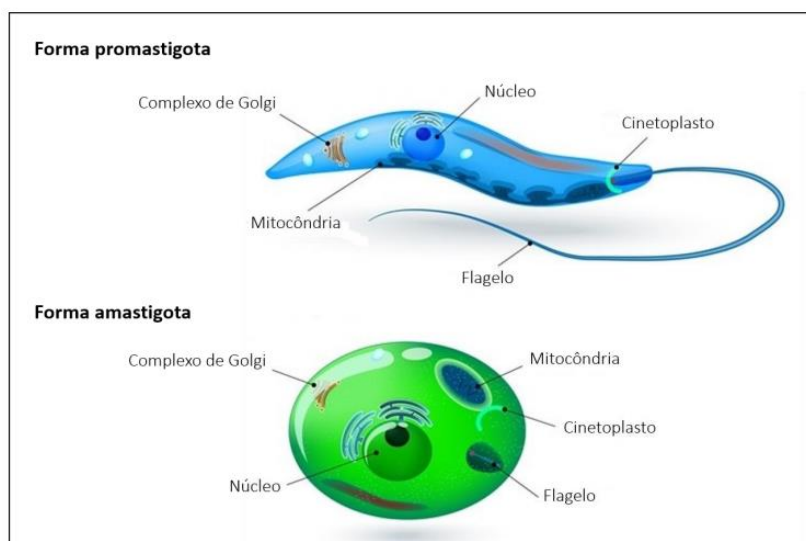
Figura 6 - Classificação atualizada de espécies de *Leishmania* patogênicas para humanos. As espécies conhecidas por causar a forma visceral da doença estão destacadas em negrito. FONTE: SERENO *et al.*, 2017

A presença de poucos genes específicos, sugere que as diferentes manifestações clínicas observadas, tenham como explicação a expressão gênica diferencial, a diferença no número de cópias dos genes e a resposta imune diferencial do hospedeiro (PEACOCK *et al.*, 2007; ROGERS *et al.*, 2011; TEIXEIRA *et al.*, 2012). A epidemiologia é extremamente diversa: 20 espécies de *Leishmania* são patogênicas para o homem, e 30 espécies de flebotomíneos são vetoras. Existem duas classificações epidemiológicas para a doença: zoonose, que inclui animais hospedeiros como reservatórios no ciclo de transmissão, e antroponose, na qual o homem é a única fonte de infecção para o mosquito vetor (KAYE; SCOTT, 2011)

Embora a virulência possa ser ajustada por fatores ambientais e genéticos, relacionados ao hospedeiro mamífero e ao vetor, as características genéticas do parasita são um fator chave no estabelecimento de uma infecção (PEACOCK *et al.*, 2007).

De todas as formas clínicas, a mais grave é a Leishmaniose Visceral, também chamada de calazar ou febre muda. É caracterizada por febre intermitente, hepatoesplenomegalia, anemia e fraqueza progressiva, que geralmente leva à morte. As principais espécies responsáveis pela LV são *L. donovani* e *L. infantum* e *L. (L) chagasi* (MCCALL *et al.* , 2013 ; BURZA *et al.* , 2018), que juntas formam o complexo de espécies de *L. donovani*.

L. donovani é considerada amplamente antroponótica, em que o ser humano é o único reservatório, suscetível e hospedeiro, mesmo que os parasitas possam ser encontrados em animais (BHATTARAI *et al.*, 2010). Em contraste, *L. infantum* - como a maioria das espécies de *Leishmania* - causa uma doença zoonótica, em que os cães são o principal reservatório doméstico, mas uma variedade de mamíferos selvagens também estão envolvidos na transmissão (DÍAZ-SÁEZ *et al.*, 2014; QUINNELL & COURTENAY, 2009). Ambas as espécies estão distribuídas em todo o mundo.



Fonte: Adaptado de News Medical Life Sciences, 2018.

Figura 7 - Morfologia da *Leishmania* nas formas promastigota e amastigota.

A infecção do vetor ocorre quando uma fêmea suga o sangue de um mamífero infectado e ingere os macrófagos parasitados com a forma não-flagelada. No trato digestivo anterior, se rompem, liberando essas formas. Eles se reproduzem por divisão binária e rapidamente se diferenciam na forma de promastigotas procíclica. Também se reproduzem por meio de um processo de divisão binária contínua. O ciclo do parasita em insetos é concluído em

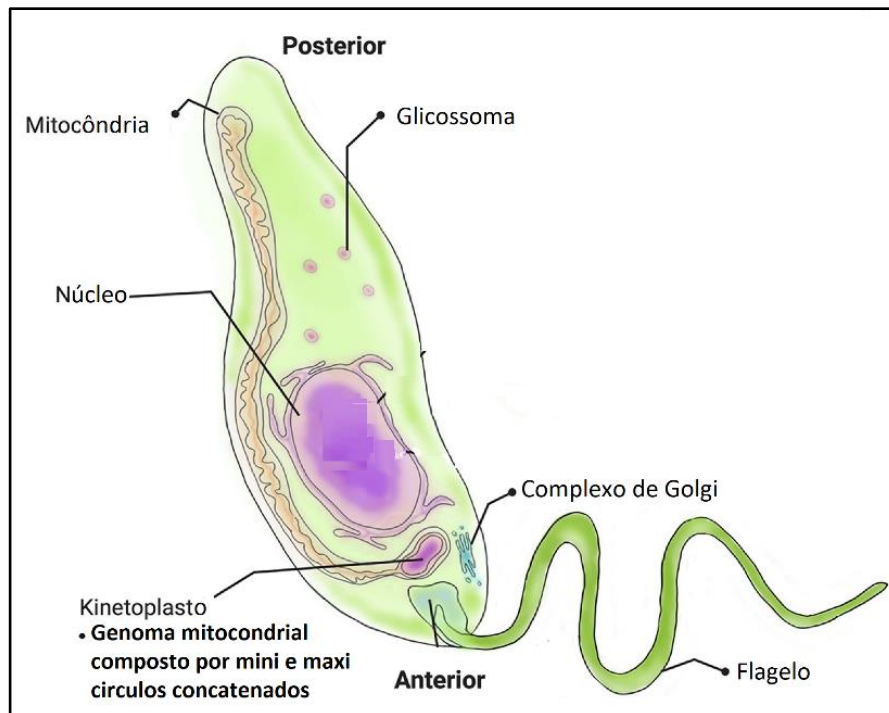
aproximadamente 7 dias. Após esse período, quando a fêmea infectada é alimentada com uma nova refeição de sangue no hospedeiro vertebrado, ela irá liberar a forma promastigota metacíclica entre as partículas na saliva do inseto.

No interior dos macrófagos diferenciam-se em amastigotas e multiplicam-se até o rompimento dos mesmos, ocorrendo à liberação destas formas que serão fagocitadas por novos macrófagos num processo contínuo, ocorrendo então a disseminação hematogênica para outros tecidos ricos em células do sistema mononuclear fagocitário, como linfonodos, fígado, baço e medula óssea (Ministério da Saúde, 2006).

A identificação da *Leishmania* causadora da doença é muito importante. E com a popularização da PCR muitas abordagens moleculares vêm sendo utilizadas com o objetivo de identificar espécies e cepas de *Leishmania* (DA SILVA et al., 2010; GRACA et al., 2012).

3.3 - kDNA - DNA de cinetoplasto (DNA mitocondrial)

Leishmania spp. e outros tripanossomatídeos, têm características únicas de organização genômica em comparação com os eucariotos, como genes sem íntrons, policistrons e pequenos cromossomos com alta densidade genética. Além disso, esses flagelados possuem uma única mitocôndria chamada cinetoplasto (Figura 8), que contém uma grande rede de DNA do cinetoplasto (kDNA) (LUKEŠ et al., 2002).

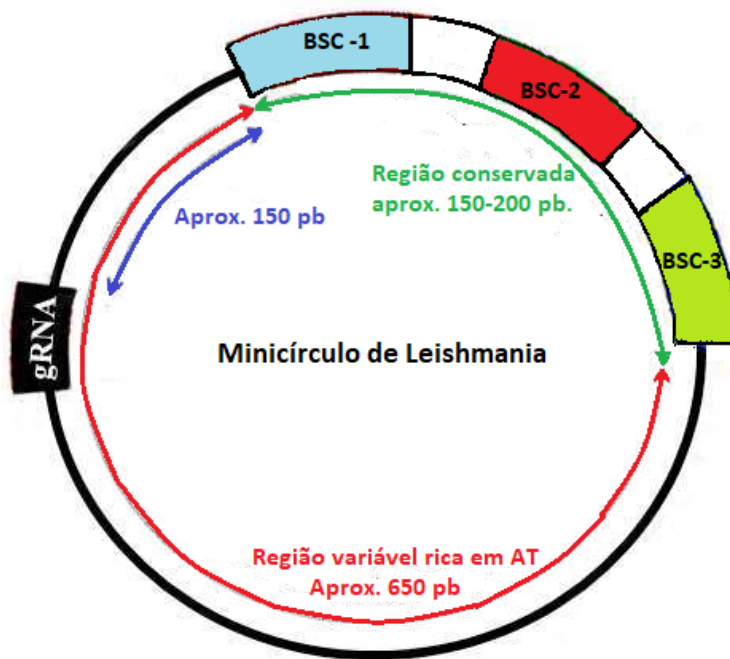


FONTE: <http://www.dbbm.fiocruz.br/tropical/leishman/leishext/html/morfologia.htm>

Figura 8 - Forma promastigota de *Leishmania sp.* com a localização do kDNA, na porção anterior.

Todos os flagelados do cinetoplastídeo, possuem um único genoma mitocondrial conhecido como DNA do cinetoplasto (kDNA), que consiste em milhares de moléculas circulares de DNA, ligadas entre si em uma rede concatenada (LUKEŠ et al., 2002).

O cinetoplasto constituída por uma massa de DNA circular com milhares de minicírculos (~1 Kb cada) (Figura 9) e várias dezenas de maxicírculos (~23 Kb cada). Os maxicírculos codificam genes homólogos presentes no DNA mitocondrial de outros eucariotos. Os minicírculos constituem aproximadamente 95% do kDNA e codificam pequenas moléculas de RNA denominadas RNAs guia, que fornecem informações para a edição de RNA dos transcritos codificados pelo maxicírculo (STUART et al., 2005).



FONTE: https://repositorio.unb.br/bitstream/10482/15362/1/2013_ManuelaMacielBrittoAragao.pdf

Figura 9 - Minicírculo de kDNA de *Leishmania* sp. com seus componentes.
BSC: Bloco de Sequência Conservada

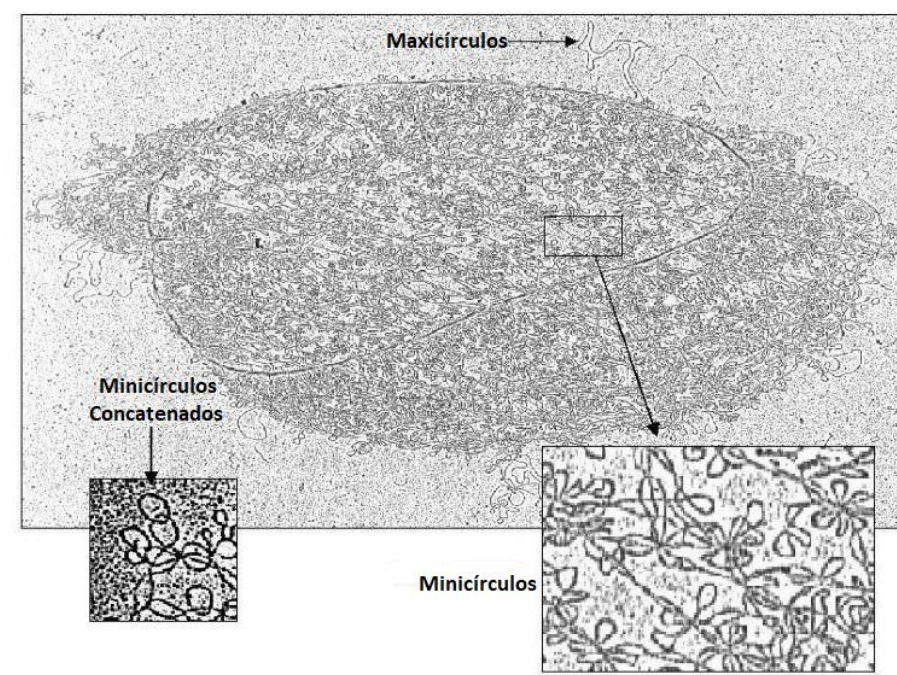
O kDNA é tradicionalmente o alvo mais usado para a detecção de *Leishmania* devido à sua natureza de multicópia, embora seja de alta sensibilidade (JARA et al., 2013). É composta por aproximadamente 10.000 minicírculos e 20–50 maxicírculos que formam a intrincada rede do DNA do cinetoplasto (Figura 10) (TELLERIA et al., 2006). Os minicírculos são pequenas moléculas circulares com aproximadamente 800 bp, contendo uma região conservada de 120 nucleotídeos e, uma região altamente variável, que codifica RNAs-guia envolvidos na edição de genes do maxicírculo (APHASIZHEV & APHASIZHEVA, 2011).

A diversidade de sequências entre o kDNA do minicírculo e do maxicírculo é conhecida por ser de natureza variável (REITHINGER, R.; DUJARDIN, 2007).

Em Trypanosomatidae, todos os minicírculos de kDNA, são organizados por regiões altamente conservadas e regiões variáveis (Figura 9). Nos minicírculos, há três blocos altamente conservados: BSC1 (GGGCGT), BCS2

(CCCCGTTC) e BSC3 (GGGGTTGGTGTA), com homologia interespecífica de 90% a 100%). BSC1 e BSC3 são quase idênticos em todas as espécies investigadas até o momento, enquanto BSC2 é menos universal (RAY, 1989; SHELINE & RAY, 1989).

Em laboratórios de diagnóstico molecular, o DNA do cinetoplasto (kDNA) é comumente utilizado, sendo que há duas abordagens (GHASEMLOO et al., 2016). Esse kDNA corresponde a 30% do DNA celular, sendo então um bom alvo para ser utilizado para o diagnóstico molecular desses parasitos (DEGRAVE et al., 1994).



FONTE: <https://kdna.net/> (ADAPTADO)

Figura 10 - Organização das fibras de DNA do cinetoplasto em maxicírculos e minicírculos em *L. tarentolae*.

O teste de PCR para detectar cinetoplasto ou DNA nuclear é, de longe a técnica de diagnóstico mais sensível e, também permite a identificação em nível de espécie. A visualização de formas amastigotas de *Leishmania* usando microscopia direta com coloração de Giemsa é o principal método usado fora dos centros de referência e em ambientes com poucos recursos (AL-JAWABREH et al., 2006). No entanto, devido à baixa sensibilidade do microscópio direto, exame de muitos pacientes não recebem a confirmação do

diagnóstico definitivo (AL-JAWABREH et al., 2006). Os testes sorológicos incluem o ensaio de imunofluorescência indireta (IFA) e ensaio imunoenzimático (ELISA). Contudo, essas técnicas podem produzir resultados falsos negativos e reações cruzadas com outras doenças sendo necessárias técnicas mais eficazes de avaliação (MORENO & NIETO, 2004).

3.4- Marcadores moleculares

Os marcadores moleculares são comumente empregados na biologia molecular de plantas, animais, fungos e microrganismos, por apresentar múltiplas finalidades, como mapeamento de genoma, marcação de genes, análise filogenética, taxonomia, seleção assistida e estudos de diversidade genética. (RIBEIRO et al, 2016). Nas últimas quatro décadas, o número de estudos e pesquisas utilizando marcadores de DNA, aumentou exponencialmente em todo o mundo (GROVER & SHARMA, 2016). Os marcadores moleculares surgiram a partir de uma necessidade para detectar o polimorfismo genético presente no DNA dos organismos. Define-se um marcador molecular como qualquer fenótipo molecular procedente de um gene expresso ou de um segmento específico de DNA (FERREIRA & GRATTAPAGLIA, 1998). Um marcador genético é um gene ou sequência de DNA com um local conhecido em um cromossomo e, associado a um gene ou característica particular. Pode ser descrita como uma variação, que pode surgir a partir de uma mutação ou alteração nos locais genômicos, que podem ser observados. Um marcador genético pode ser uma curta sequência de DNA, ou uma sequência envolvendo uma única alteração no par de bases (Polimorfismo de nucleotídeo único, SNP) ou longo, como mini e microsatélites (WANG, 2004). Nos últimos anos, as pessoas têm demonstrado grande interesse no uso de marcadores moleculares, que revelam polimorfismos no nível do DNA por desempenharem um papel importante no estudo da diversidade genética biológica.

Um marcador molecular é definido como “qualquer fenótipo molecular oriundo de um gene expresso ou de um segmento específico de DNA” (FERREIRA e GRATTAPAGLIA, 1998) e que assim permite a distinção de indivíduos ou grupo de indivíduos geneticamente diferentes. Há uma vasta quantidade de marcadores moleculares com varias finalidades contendo muitas vantagens sobre os marcadores morfológicos tradicionais, como neutralidade fenotípica, epistasia rara ou interações pleiotrópicas, e podem ser detectados em qualquer estágio do ciclo de vida. Estes marcadores possuem um amplo potencial para responder muitas questões em biologia que vão desde relações de parentesco, conectividade entre populações, diferenciação taxonômica até respostas sobre relações filogenéticas (AVISE, 1994).

O nível de discriminação da tipagem baseada em PCR depende do marcador molecular utilizado. Dessa forma, identificar regiões alvo adequadas para o anelamento do primer é um passo crucial, pois essas regiões devem ser conservadas dentro da taxa alvo e diferir em taxa relacionados (LUCCHI et al., 2013; WONG et al., 2014). Os recentes avanços na tecnologia de sequenciamento permitem a realização de estudos filogenéticos bem como a identificação de parasitas a um custo significativamente mais baixo, mesmo para organismos não modelo. A exploração dos dados genômicos pode melhorar os métodos de detecção de infecções assintomáticas com baixas cargas parasitárias, uma vez que facilita a busca por genes multi-cópias, e a diferenciação de espécies ao identificar sequencias específicas (WONG et al., 2014).

O diagnóstico molecular pode, também, permitir a identificação de características específicas do parasita, como a virulência ou resistência às drogas (LUCCHI et al., 2013). Sendo assim, a quantidade de dados genômicos disponíveis combinada com ferramentas de bioinformática abrem a possibilidade de rastrear marcadores altamente informativos, como microsatélites e genes ortólogos, a fim de discriminar as diferentes espécies do gênero e características específicas (LI et al., 2009b; REITHINGER; DUJARDIN, 2007).

Os marcadores moleculares podem ser considerados como uma revolução dentro de genética de populações, uma vez que, não se faz mais necessário a existência de organismos mutantes que manifestem diferenças visíveis, para se estudar a sua genética (PAVINATO, 2010). Por meio de uma imensidade de dados gerados pelos métodos moleculares em estudos de polimorfismos genéticos, populações podem ser estudadas independentemente da espécie ou habitat sem que haja a necessidade de cruzamentos controlado (HARTL, 2008).

O surgimento da biologia molecular tem promovido o desenvolvimento dessa tecnologia, que se utiliza das diferenças genéticas, e para a criação de um marcador espécie-específicos, que tem como finalidade a diferenciação entre espécies. Do ponto de vista da diversidade populacional, estrutura, ecologia e relações com outros grupos, esta ferramenta melhorou muito a pesquisa populacional. Mesmo com a ajuda de dados fósseis, a história evolutiva da população pode ser reconstruída.

Um grande passo metodológico acerca do uso de marcadores moleculares foi o desenvolvimento da PCR (Reação em Cadeia da Polimerase). Esta técnica descrita por Kary Mullis no final dos anos 80 revolucionou a Biologia Molecular, pois possibilitou uma nova estratégia na análise de genes por meio de um método simples e rápido de amplificação de sequências, além de requerer quantidades muito pequenas de DNA (FARAH, 2000). A técnica da Reação em Cadeia da Polimerase (PCR) é essencial para a biologia molecular e, é a prática laboratorial mais utilizada na amplificação de ácidos nucleicos. A PCR permite multiplicar sequências específicas de nucleotídeos (DNA) de forma exponencial, a partir de um complexo genômico (GARIBYAN & AVASHIA, 2013).

Para que o marcador tenha sucesso, no diagnóstico molecular, deve ser selecionada a região com maior diferença entre os grupos, ou seja, a região com alta taxa de mutação flanqueados por setores conservados que serão complementares aos primers.

Em resumo, o minicírculo de kDNA de *Leishmania*, é o alvo ideal para a genotipagem e desenvolvimento de novos primers, pois as diferenças

genéticas estão permitindo uma correta discriminação entre as espécies do táxon *Leishmania*. (RANASINGHE et al., 2015).

3.5 RFLP (Restriction fragment lenght polymorphism) como alternativa ao sequenciamento

O princípio básico dos marcadores RFLP é que qualquer DNA genômico pode ser diferenciado de acordo com a presença ou ausência de locais de enzimas de restrição. O polimorfismo para comprimento de fragmentos de restrição (RFLP), proposto por BOTSTEIN et al. (1980), é um tipo de marcador que resulta da variação na sequência de nucleotídeos do DNA, utilizando endonucleases que cortam a molécula de DNA em locais específicos, chamados sítios de restrição.

O polimorfismo genético esta presente, e devido ao acúmulo de mutações de nucleotídeo único no DNA genômico, nos locais das enzimas de restrição no DNA mudam, resultando em uma diferença nos padrões de restrição de dois genomas intimamente relacionados (SARASWATHY & RAMALINGAM, 2011).

O resultando de RFLP é determinado pela formação de diferentes fragmentos de comprimentos variáveis, capaz de distinguir indivíduos em função de alterações nos nucleotídeos, devido a mutação, deleção, inserção e inversão (DIBA et al., 2014). Essas diferenças na sequência de ácido nucleico nos locais de ligação da enzima de restrição, apenas confirma que a sequência entre dois ou mais indivíduos mudou, constituindo assim polimorfismo genético. Esse marcador molecular possui grandes vantagens, pois estão disponíveis em grande número, exibindo uma ampla variabilidade natural, não influenciados pelo ambiente. (DIBA et al., 2014).

O método molecular PCR-RFLP combina a sensibilidade da PCR com os meios de diferenciação das espécies para o diagnóstico e identificação das espécies de *Leishmania*. Além disso, quando o RFLP está associado à PCR,

este requer pequenos volumes de DNA do parasita, que são necessários para uma análise comparativa com os métodos convencionais de RFLP (MCMANUS, D.P.; BOWLES.; 1996).

Apesar do desenvolvimento de métodos mais atualizados em um ambiente de laboratório de diagnóstico e, na ausência de uma alternativa econômica que possa oferecer um tempo de resposta eficiente, desde a coleta até o diagnóstico, a PCR-RFLP continua sendo um método útil para a detecção e identificação de espécies de *Leishmania* (MCMANUS, D.P.; BOWLES.; 1996).

3.6- Bioinformática

Bioinformática é um campo interdisciplinar da ciência, que desenvolve métodos computacionais e pacotes de software para análise de dados. A bioinformática combina tecnologias das ciências da computação, estatística, biologia molecular e bioquímica, para otimizar e processar os dados biológicos obtidos (HE, 2015).

Ao longo dos últimos 20 anos, a quantidade de dados genômicos e proteômicos produzidos, tem gerado uma grande abundância de informações biológicas que necessitam ser interpretados (RAZA, 2010). Dentre as áreas da bioinformática, destinadas ao processamento e análises em bancos genômicos ou de proteínas, destaca-se a mineração de dados, que visa desenvolver algoritmos com potencial para analisar e interpretar uma grande quantidade de informações (ALPAYDIN, 2004; CARVALHO, 2005).

Com o rápido desenvolvimento da bioinformática global, vários projetos internacionais de sequenciamento de genoma foram desenvolvidos e redes de computadores que facilitam o acesso a dados biológicos foram criadas. Bancos de dados biológicos (genes e proteínas) são projetados para fornecer sequências de nucleotídeos ou aminoácidos gratuitamente a toda a comunidade científica, de forma que diagnósticos precisos possam ser feitos a partir do desenvolvimento de programas ou softwares para análises mais precisas.

Vários programas de bioinformática já foram descritos na literatura para seleção de pares de *primers* a partir de uma sequência alvo, como um marcador molecular. A maioria destes estão disponível gratuitamente, entretanto, a qualidade e manutenção das ferramentas são variáveis. Isso muitas vezes resulta na perda de *links* e incompatibilidades, portanto, o que anteriormente se mostrava útil pode não ser mais funcional (ABD-ELSALAM, 2003).

Há uma vasta quantidade de sites disponíveis em que pode se utilizar na bioinformática dentre eles pode-se acessar (disponíveis gratuitamente na internet)

- Centro Nacional de Informações sobre Biotecnologia (www.ncbi.nlm.nih.gov). Contém ferramentas de bioinformática e bancos de dados de DNA, RNA e Proteínas;
- National Center for Genome Resources (www.ncgr.org/) - vincula cientistas a soluções de bioinformática por meio de colaborações, dados e desenvolvimento de software;
- Genbank (www.ncbi.nlm.nih.gov/Genbank) - armazena e arquiva sequências de DNA de grande projetos genomas e de laboratórios individuais;
- Instituto Europeu de Bioinformática (www.ebi.ac.uk) - centro de pesquisa e serviços em bioinformática; gerencia bancos de dados biológicos;
- Ensembl (www.ensembl.org) - banco de dados de anotação automática sobre genomas;
- BioInform (www.bioinform.com) - serviço global de notícias sobre bioinformática;
- SWISS-PROT (www.expasy.org) - Importante banco de dados de proteínas com dados de sequência de todos os organismos, que possui um alto nível de anotação (inclui função, estrutura e variações) e é minimamente redundante (poucas cópias duplicadas);

A solução desses problemas envolve o desenvolvimento de ferramentas computacionais que organizam, armazenam e recuperam

informações de banco de dados (ABDELSALAM, 2003; ATTWOOD et al., 2011; FULLER et al., 2013). A tecnologia de serviços web foi especialmente concebida para proporcionar interoperabilidade entre diferentes plataformas. A ideia de construir um serviço web ao invés de encapsular scripts na forma de um programa torna-se, portanto, uma opção mais vantajosa (CAVALCANTI et al., 2005).

Na bioinformática, esses serviços têm sido apontados como uma tecnologia capaz de contribuir significativamente na exploração plena dos dados biológicos e muitos projetos da área têm concentrado esforços nessa direção, viabilizando a execução, visualização e gestão dos dados através da internet (CAVALCANTI et al., 2005).

Uma das tendências mais conhecidas na bioinformática é o crescimento exponencial da quantidade de dados genômicos disponíveis, devido principalmente aos crescentes avanços na tecnologia de sequenciamento (PAGANI et al., 2012). Para fazer pleno uso desses dados, a sequência do genoma deve ser convertida em conhecimento biológico. Por esse motivo, vários tipos de estudos pós-genômicos têm sido conduzidos para identificar as características dessas sequências de nucleotídeos. Dessa forma, ferramentas de bioinformática têm sido desenvolvidas para realizar uma infinidade de funções, como por exemplo: predição de genes, domínios e homologia, e rastreamento de sequências repetitivas e polimorfismos (DURAN et al., 2009).

A bioinformática vem sendo cada vez mais utilizando por pesquisadores nos avanços das pesquisas e contribuindo no processo do desenvolvimento da ciência, pois a mesma pode ser utilizada no dia-a-dia como ferramenta para a detecção de doenças.

3.7 - PCR *in silico*

Com o progresso da bioinformática, devido a crescente ampliação de sequências depositadas no GenBank do NCBI (banco central de dados genômico) desde 1982 (NCBI, 2015), pôde-se dispor de ferramentas e

softwares de bioinformática de qualidade, para aplicação no desenho e validação de *primers* (ARBEX; SILVA; COSTA, 2006).

O desenho de primer é uma fase crítica em todos os procedimentos de PCR, para garantir a amplificação específica e eficiente de uma sequência alvo e, portanto, muitos pesquisadores se empenharam a desenvolver conjuntos de software eficientes para o *design*, seleção e verificação de *primers* (FREDSLUND & LANGE, 2007; BEKAERT & TEELING, 2008; KALENDAR et al., 2011; KALENDAR et al., 2014).

PCR de computador eletrônico refere-se a uma PCR virtual realizada por um software ao inserir um par ou um conjunto de iniciadores do genoma desejado e / ou uma biblioteca de sequências disponível na Internet. Com o desenvolvimento da tecnologia de sequenciamento e a rápida redução de custos, muitos genomas foram sequenciados e anotados no banco de dados.

Devido ao conhecimento prévio de pseudogenes ou polimorfismos de nucleotídeo único homólogos (SNPs) e / ou amplicons indesejados, a PCR simulada por computador visa avaliar a especificidade da PCR para identificar os locais de ligação do primer.

A reação em cadeia da polimerase (PCR) é uma técnica *in vitro*, que é a base da prática da biologia molecular. No entanto, a utilidade e eficiência dos métodos de PCR dependem da identificação e desenvolvimento de primers eficazes. Estes *primers* são fitas curtas de DNA de fita simples (18-23 pb) correspondendo à sequência no final do fragmento ou à sequência alvo. Os primers são essenciais para iniciar a síntese de DNA, e o mesmo é combinado com uma sequência modelo mais longa para iniciar a síntese de DNA. Assim, a PCR *in silico* pode auxiliar na seleção de primers recém-projetados, simulando em computador, a identificação de possíveis regiões de incompatibilidade de ligação específica, evitando dificuldades na amplificação de produtos de PCR indesejados, antes de qualquer experimento real na prática (BARTLETT & STIRLING, 2003).

A detecção e diferenciação de espécies de *Leishmania* contribuem para a determinação do prognóstico do paciente e tratamento e pode ser uma

ferramenta inestimável para a vigilância epidemiológica (SCHRIEFER et al., 2008; MEDLEY et al., 2015). A PCR *in silico* como ferramenta nos laboratórios no desenvolvimento e avanços da ciência se mostra essencial na biologia molecular para aprimoramento e desenvolvimento de marcadores para detecção das leishmania.

4 MATERIAIS E MÉTODOS

A análise por PCR *in silico*, pode ser útil e eficiente na identificação de possíveis problemas, antes de qualquer experimento real ou, na validação de *primers*. Para análise utiliza-se um computador com acesso à Internet é fundamental. Acessar os sites com pacotes de ferramentas para análises da PCR *in silico* e, fazer o *download* do programa *BioEdit* (<https://www.mybiosoftware.com/bioedit-7-0-9-biological-sequence-alignment-editor.html>), disponível gratuitamente na Web. Acessar os sites da *CLUSTAL Omega* (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) *Neb Cutter2* (<http://nc2.neb.com/NEBcutter2/help/cite.html>), *Net Primer* (<http://www.premierbiosoft.com/netprimer>), *Sequence Manipulation Suite* https://www.bioinformatics.org/sms2/pcr_products.html ou uma ferramenta disponível na Internet utilizada e discutida neste estudo são principalmente as “PCR *in silico*”.

Para o desenvolvimento dos *primers*, foi selecionada uma região complementar conservada e, sem variação nas espécies-alvo, para garantir repetibilidade das reações de amplificação. Foram considerados como homologias, os alinhamentos com similaridades baseada na estatística E-value, e que apresentaram valores inferiores a 1e-10. O valor do E-value estima a probabilidade de que o alinhamento entre sequências tenha ocorrido ao acaso.

A primeira etapa no desenvolvimento dos *primers* foi à busca por sequencias de nucleotídeos, através da base de dados do NCBI (<https://www.ncbi.nlm.nih.gov/>). Encontrar a sequência codificadora de um determinado gene ou região de interesse, através do BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), que permite localizar regiões de similaridade entre sequências biológicas. O programa compara sequências de nucleotídeos ou proteínas no bancos de dados de sequências.

Foram selecionados 11 sequencias de DNA do genoma KDNA mitocondrial completo, sendo 2 de *L. donovani* e 9 de *L. infantum*, exibidas no formato *Genbank* (Tabela 1). Estas sequências foram convertidas para o formato *FASTA Text*, no próprio sistema do banco de dados e, em seguida, realizado o *download*. As sequencias de cada espécie, foram agrupadas e

exportadas para um *software* de edição de textos simples (bloco de notas do sistema operacional *Windows*), dando origem a dois bancos de sequencias.

As diferentes sequências foram alinhadas e comparadas manualmente para que seja possível reconhecer a região de maior similaridade entre as espécies pesquisadas. Uma vez identificada região, que guarda uma maior semelhança entre as espécies, pode-se, então, utilizá-la para o desenho dos *primers*, mantendo-se as regiões idênticas e, considerando-se todas as opções de nucleotídeos nas regiões que diferem entre si. Os resultados apresentados foram selecionados para um melhor desenvolvimento e, confirmação dos *primers*. Estes foram convertidos para o formato FASTA, para um download contendo as sequencias, sendo posteriormente exportada para um software de edição de texto e utilizado no bloco de notas, para a realização desta etapa.

A segunda etapa se dá por meio do alinhamento. As sequências foram alinhadas utilizando o algoritmo Clustal Omega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>), onde se originou um novo resultado contendo as sequencias resultantes. Com a realização do mesmo, foram analisadas as regiões do genoma quanto a sua conservação ou variabilidade. As mesmas sequências alinhadas no Clustal Omega, foram complementados no software BioEdit Sequence Alignment Editor versão 7.2.6 (HALL, 1999).

Os *primers* foram desenvolvidos *online* no Primer3Plus (<https://www.bioinformatics.nl/cgi-bin/primer3plus/primer3plus.cgi>), determinando alguns parâmetros como critérios de inclusão ou exclusão. A região escolhida deveria apresentar nucleotídeos entre 20-23 pb e, com distribuição de GC máxima em 60%, para evitar a elevação da temperatura de anelamento e dissociação. As regiões com muitas AT também foram evitadas, para não haver a redução da Tmelting, que possibilitasse o auto anelamento, sendo recomendado valores inferiores a 60°C. A Tmelting foi calculada, estimando-se valores entre cada par de *primers* mais próximos, visando estabelecer uma média na padronização da PCR.

A validação final foi feita com uso da ferramenta online NetPrimer (<http://www.premierbiosoft.com/NetPrimer/AnalyzePrimer.jsp>). Foi verificada a possível formação de self dímeros, heterodímeros e hairpin. Na análise de dímeros, os *primers* não poderiam apresentar energia livre de Gibbs abaixo de

-9 kcal/mole-1, ou muitas bases complementares. Na análise de Hairpin, além dos valores de Delta G, foi verificada a temperatura de melting, que possibilitasse o auto anelamento, sendo recomendado valores inferiores a 60° C e, taxa de validação mínima de 80%.

As sequências dos potenciais marcadores foram testados para gerar fragmentos previstos, que resultariam numa digestão por PCR-RFLP. As sequências obtidas no GenBank NCBI, foram alinhadas de acordo com a espécie e, geradas sequências consenso em separado. Depois incluídas para gerar um modelo contendo todas as sequências de *Leishmania* disponíveis. Os procedimentos de digestão virtual foram realizadas, usando o programa on line NEBcutter V2.0 (<https://nc2.neb.com/NEBcutter2/>) realizadas com todas as enzimas disponíveis comercialmente, para avaliar a mais adequada, na promoção de diferenciar as espécies em estudo.

5 RESULTADOS E DISCUSSÕES

Várias regiões do genoma do KDNA foram selecionadas e, calculadas o possível tamanho dos fragmentos de amplificação na PCR *in silico*, até que se pudesse chegar a um conjunto de *primers* que se encaixassem nos parâmetros descritos na metodologia. Os *primers* obtidos estão representados na Tabela 3. As sequências foram escolhidas de acordo com diversos critérios, como região conservada, área interprimer com polimorfismo de nucleotídeo único (SNP). Além destes requisitos, outros foram adicionados para melhorar a eficiência dos iniciadores. Caixeta *et al.* (2016), resume em 8 itens:

- I - comprimento dos oligonucleotídeos entre 18 a 30 pb;
- II - composição de GC entre 40 e 60%;
- III - número de repetições de purinas/ pirimidinas menores do que 4;
- IV- Ausência de sequências palindrômicas e sem estruturas secundárias;
- V - sem complementaridade na ponta de um *primer* e qualquer parte do outro (evitar dímeros);
- VI - Extremidade 3' terminando em G ou C, mas sem região formadora de grampo;
- VII - temperatura de *melting* muito próxima entre os oligonucleotídeos;
- VIII - temperatura de anelamento abaixo de 60° C.

A obtenção das sequencias de alinhamento significativos, no qual foram selecionadas 11 sequencias com maiores níveis de porcentagem de identificação (Tabela 2).

Após a seleção, fez-se download das sequencias no formato TXT utilizando o Clustal Omega para realizar múltiplas sequencias de alinhamento e observar a similaridade entre as regiões conservadas e variáveis por SNP, para serem avaliadas no desenvolvimento de primer.

Depois foi gerado uma sequência consenso (figura 20) , utilizando o software Bioedit (www.mybiosoftware.com/bioedit-7-0-9-biological-sequence-alignment-editor.html), onde foi possível observar e validar os primer desenvolvidos.

Sequences producing significant alignments					Download	New Select columns	Show	100	
☐ select all 11 sequences selected					GenBank	Graphics	Distance tree of results	New MSA Viewer	
	Description	Scientific Name	Max Score	Total Score	Query Cover	E value	Per Ident	Acc. Len	Accession
☒	Leishmania infantum genome assembly chromosome 1	Leishmania inf...	2287	2287	100%	0.0	100.00%	17897	gi 1756572067 LR697137.1
☒	TPA: Leishmania infantum strain JPCM5 kinetoplast complete sequence	Leishmania inf...	2287	2287	100%	0.0	100.00%	18277	gi 1751371458 BK010877.1
☒	Leishmania infantum isolate ZK47 kinetoplast complete genome	Leishmania inf...	2287	2287	100%	0.0	100.00%	17462	gi 1897352072 MT762287.1
☒	Leishmania infantum isolate ZK5 kinetoplast complete genome	Leishmania inf...	2287	2287	100%	0.0	100.00%	17557	gi 1897352059 MT762286.1
☒	Leishmania infantum isolate ZK43 kinetoplast complete genome	Leishmania inf...	2287	2287	100%	0.0	100.00%	17631	gi 1897352046 MT762285.1
☒	Leishmania infantum isolate ZK28 kinetoplast complete genome	Leishmania inf...	2287	2287	100%	0.0	100.00%	17831	gi 1897352033 MT762284.1
☒	Leishmania infantum isolate ZK25 kinetoplast complete genome	Leishmania inf...	2287	2287	100%	0.0	100.00%	17631	gi 1897352020 MT762283.1
☒	Leishmania infantum isolate LIPA83 kinetoplast complete genome	Leishmania inf...	2287	2287	100%	0.0	100.00%	17825	gi 1897352007 MT762282.1
☒	Leishmania infantum isolate LIPA60 kinetoplast complete genome	Leishmania inf...	2287	2287	100%	0.0	100.00%	17523	gi 1897351994 MT762281.1
☒	Leishmania donovani strain pasteur kinetoplast complete genome	Leishmania do...	2237	2237	100%	0.0	99.27%	40931	gi 1229082218 CP022652.1
☒	Leishmania donovani isolate MHOM/SD/62/1S-C12D maxicircle partial sequence kinetoplast	Leishmania do...	2237	2237	100%	0.0	99.27%	17028	gi 213492206 FJ416603.1
☐	Leishmania major strain Friedlin genome assembly chromosome 1	Leishmania m...	1701	1701	100%	0.0	91.54%	18998	gi 1756572068 LR697138.1
☐	Leishmania major isolate MHOM/SU/73/5ASKH maxicircle partial sequence mitochondrial	Leishmania m...	1679	1679	100%	0.0	91.22%	8738	gi 195970463 EU140338.1
☐	TPA: Leishmania aethiopica strain MHOMET kinetoplast complete sequence	Leishmania ae...	1653	1653	99%	0.0	90.87%	18065	gi 1751371523 BK010882.1
☐	TPA: Leishmania pifanoi strain LPROD kinetoplast complete sequence	Leishmania pif...	1635	1635	99%	0.0	90.55%	19040	gi 1751371549 BK010884.1
☐	Leishmania amazonensis isolate MPRP/BR/72/M1845 kinetoplast partial genome	Leishmania a...	1622	1622	99%	0.0	90.41%	9374	gi 298289304 HM439238.1

Figura 11 - Sequencia realizada pelo NCBI, selecionando as 11 primeiras sequencias.

Tabela 2 - Sequências utilizadas para o desenho e validação de primers (NCBI).

Espécie	Acesso GenBank
<i>L. donovani</i>	FJ416603.1
	CP022652.1
<i>L. infantum</i>	LR697137.1
	BK010877.1
	MT762287.1
	MT762286.1
	MT762285.1
	MT762284.1
	MT762283.1
	MT762282.1
	MT762281.1

Os três *primers* desenvolvidos, testados e validados nos programas utilizados, todos os produtos obtidos (*amplicons*) pelos 11 acessos (Tabela 2) apresentaram homologias entre elas. O desenho do *primer* e a análise de PCR-RFLP *in silico* foram baseados nos alinhamentos das sequências nucleotídicas consenso de regiões que apresentam polimorfismo entre as várias sequências nas bases de dados do *GenBank* (Figura 21).

Tabela 3 - Sequência dos primers desenvolvidos para a detecção das espécies em estudo.

	Nome	Primer	pb	% GC	Tm °C	Amplicon (pb)	TAXA VALIDAÇÃO (%)
1	ITA1-F	CTTGTAGTTGTGGCGTGTTTG	21	47,62	56,79	306	100
2	ITA2-R1	TTTCTTTCTCTCCGCTTATCTT	23	39,13	59,08		100
3	ITA2-R2	ACGACATCATCGTACATGGTAAT	23	39,13	56,48	672	85

Os *primers* escolhidos apresentaram variação de tamanho entre 21-23 pb, sendo que as sequências respectivas, bem como a posição no gene de interesse estão descritas na Tabela 3. As temperaturas de fusão (*Tmelting*) apresentaram um intervalo entre 56,48 °C e 59,08 °C, dentro da variação esperada entre os pares de no máximo 5°C e, a concentração de GC%, de 39,13% a 47,62% (Tabela 3).

Analysis Results #1: CTTGTAGTTGTGGCGTGTTTG					
Rating	:	100.0		3' end stability	: -7.18 kcal/mol
Molecular Wt	:	6505.29		ΔH	: -153.5 kcal/mol
Tm	:	56.79 °C		ΔS	: -0.4 kcal/°K/mol
GC%	:	47.62		5' end ΔG	: -6.84 kcal/mol
GC Clamp	:	1		Self Dimer (ΔG)	: kcal/mol
nmol/A ₂₆₀	:	5.19		Hairpin (ΔG)	: kcal/mol
ug/A ₂₆₀	:	33.76		Repeats (# of pairs)	: kcal/mol
ΔG	:	-33.73	kcal/mol	Run (# of bases)	: 3 kcal/mol
Analysis Results #2: TTTCTTTCTCTCCGCTTATCTT					
Rating	:	100.0		3' end stability	: -6.59 kcal/mol
Molecular Wt	:	6848.56		ΔH	: -181.2 kcal/mol
Tm	:	59.08 °C		ΔS	: -0.48 kcal/°K/mol
GC%	:	39.13		5' end ΔG	: -7.06 kcal/mol
GC Clamp	:	1		Self Dimer (ΔG)	: kcal/mol
nmol/A ₂₆₀	:	5.35		Hairpin (ΔG)	: kcal/mol
ug/A ₂₆₀	:	36.62		Repeats (# of pairs)	: kcal/mol
ΔG	:	-38.36	kcal/mol	Run (# of bases)	: 3 kcal/mol
Cross Dimer (ΔG)	:		kcal/mol		

Figura 12: Validação do primeiro par de primers observando os parâmetros de Tm, GC%, formação de harpin, heterodímeros, self dímeros e ΔG .

Analysis Results #1: CTTGTAGTTGTGGCGTGTTC					
Rating	:	100.0		3' end stability	: -7.18 kcal/mol
Molecular Wt	:	6505.29		ΔH	: -153.5 kcal/mol
Tm	:	56.79	°C	ΔS	: -0.4 kcal/°K/mol
GC%	:	47.62		5' end ΔG	: -6.84 kcal/mol
GC Clamp	:	1		Self Dimer (ΔG)	: kcal/mol
nmol/A ₂₆₀	:	5.19		Hairpin (ΔG)	: kcal/mol
ug/A ₂₆₀	:	33.76		Repeats (# of pairs)	: kcal/mol
ΔG	:	-33.73	kcal/mol	Run (# of bases)	: 3 kcal/mol

Analysis Results #2: ACGACATCATCGTACATGGTAAT					
Rating	:	85.0		3' end stability	: -5.72 kcal/mol
Molecular Wt	:	7031.69		ΔH	: -162.8 kcal/mol
Tm	:	56.48	°C	ΔS	: -0.43 kcal/°K/mol
GC%	:	39.13		5' end ΔG	: -7.87 kcal/mol
GC Clamp	:	1		Self Dimer (ΔG)	: -6.53 kcal/mol
nmol/A ₂₆₀	:	4.3		Hairpin (ΔG)	: -2.13 kcal/mol
ug/A ₂₆₀	:	30.24		Repeats (# of pairs)	: kcal/mol
ΔG	:	-34.8	kcal/mol	Run (# of bases)	: kcal/mol

Cross Dimer (ΔG)	:	-4.26	kcal/mol
----------------------------	---	-------	----------

Figura 13: Validação do segundo par de primers observando os parâmetros de Tm, GC%, formação de harpin, heterodímeros, self dímeros e ΔG .

Estes marcadores moleculares foram desenhados propondo a sua utilização em duas PCRs. Na 1ª PCR utiliza-se os *primers* ITA1-F e ITA2-R1 que amplificam um fragmento de 306 pb (Figura 13) e, na 2ª PCR, os *primers* ITA1-F e ITA2-R2, geram um produto de 672 pb (Figura 14). Os valores de validação variaram entre 85% e 100%, representando uma alta especificidade. Os primers ITA1-F e ITA2-R1 obtiveram um valor com 100% de identidade, o que pode ser explicado pela ausência em ambos de heterodímeros, self dímeros e formação de harpin (Figura 13). Na combinação entre ITA1-F e ITA2-R2, já apresentam formação de heterodímeros, self dímeros e harpin (Figura 14). O heterodímero se caracteriza pela hibridização de primers diferentes (Figura 16), já os self dímeros (Figura 14) é definido pelo auto anelamento do primer (ITA2-R2), produzindo numa PCR *in vitro* e visualizado numa corrida em gel de agarose. A formação de harpin (grampo) em um primer é devido a ligação entre nucleotídeos da sequência, originando uma dobra e consequentemente inabilitando a função de se anelar na sequência pretendida (Figura 15).

Em todas as possíveis combinações de hibridização por heterodímeros, selfdímeros e formação de harpin, os valores obtidos no ΔG

não ultrapassaram o limiar de -9Kcal.mole⁻¹, conforme esperado. Quanto aos resultados referentes a avaliação de self dímeros e heterodímeros (Figuras 14 e 16) se mantiveram baixo e não comprometeram a qualidade dos primers.

```

Total Found = 4
1.  ΔG = -6.53 kcal/mol
    5' ACGACATCATCGTACATGGTAAT 3'
      ||||  ||||
    3' TAATGGTACATGCTACTACAGCA 5'

2.  ΔG = -5.38 kcal/mol
    5' ACGACATCATCGTACATGGTAAT 3'
      ||||
    3' TAATGGTACATGCTACTACAGCA 5'

3.  ΔG = -3.65 kcal/mol
    5' ACGACATCATCGTACATGGTAAT 3'
      |  |  ||||  |  |
    3' TAATGGTACATGCTACTACAGCA 5'

4.  ΔG = -3.3 kcal/mol (3' Dimer)
    5' ACGACATCATCGTACATGGTAAT 3'
      |||  |||
    3' TAATGGTACATGCTACTACAGCA 5'

```

Figura 14 - Análise dos valores de ΔG na formação de self dímeros.

```

Total Found = 1
1.  ΔG = -2.13 kcal/mol
    TACAGCA 5'
    T ||||
    LCATCGTACATGGTAAT 3'

```

Figura 15 - Análise de Valor de ΔG na formação de harpin.

```

Total Found = 1
1.  ΔG = -4.26 kcal/mol
    5' CTTGTAGTTGTGGCGTGTGTTG 3'
      |  ||||  ||
    3' TAATGGTACATGCTACTACAGCA 5'

```

Figura 16 - Análise do Valor de ΔG na formação de heterodímeros.

Estes parâmetros são importantes no desenho de primers, pois evitam desperdício no momento da execução real numa PCR *in vitro*.

Nas figuras 17, 18 e 19 foram identificadas regiões conservadas e ideais para o desenvolvimento dos *primers* ITA1-F, ITA2-R1 e ITA2-R2, que são específicos para a região definida.



Figura 17: Múltiplas seqüências alinhadas pelo BioEdit e posicionamento do primer ITA1-F.

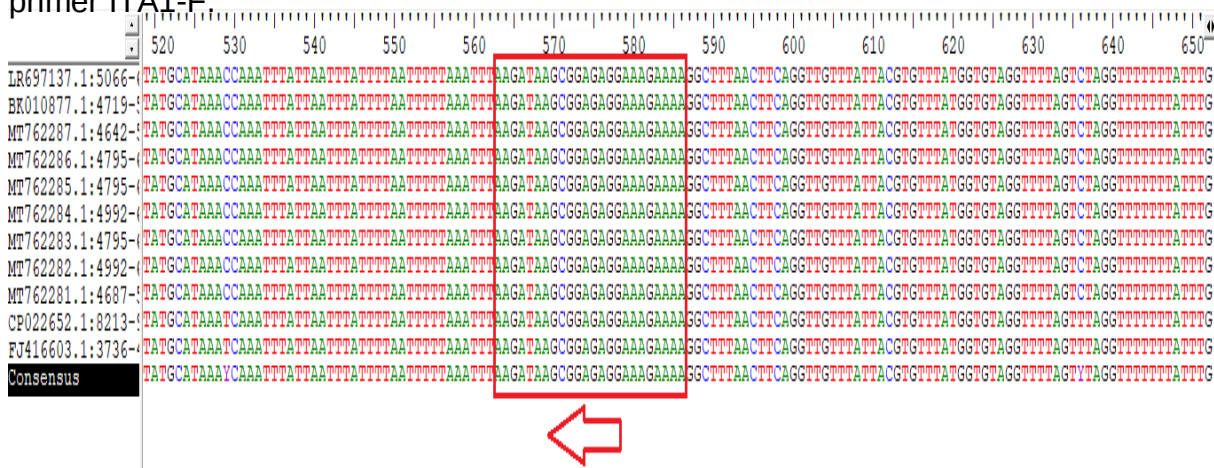


Figura 18: Múltiplas seqüências alinhadas pelo BioEdit e posicionamento do primer ITA2-R1.



Figura 19: Múltiplas seqüências alinhadas pelo BioEdit e posicionamento do primer ITA2-R2.

TTGTTAGTGTAGAGTTTGCGTTTG TATTTGTTATTCCTATTATGTTTAGTTGTTAAATTTGTGATTTTGGTTTTGTATTTATTGGTA
TTTTATTGATATTTTAAATTTATTAATTAATACATTTTATTATTTGTAAGTGTTTATTTGTCAATTTTGTATTTTATTGGTT
TCGTTTTTTTTATGTGTTTTATTATGTTGTGAGTAGGTATATTATTTGGATTTTATTTTATGAAATCAAGTTTGAGAGTTTTCA
CTATTATTTGTTACTTGTAGTTGTGGCGTGTGGTTCKATATTATTTTAATAGATTATTGCATTTTAGTCATGTGTTTTAGGTA
TATTTTATTATTTTATGTTTAGYAGATGTTTAACTTTTATGTATGGAYACAGTTTGTATTTTATATGTTGTATGTTTATAT
TGACATTTTGTGATTGTGTATGATTTTTTTTATTACGATTTGTATTTTGATGTTTTAAGTGTAGTGTAYTTGTATGCATAAAYC
AAATTTATTAATTTATTTTAAATTTTAAATTTAAGATAAGCGGAGAGGAAAGAAAAGGCTTTAACTTCAGGTTGTTTATTACGTG
TTATGGTGTAGGTTTTAGTYTAGGTTTTTATTGTATGCAGATAATATGTGGYGTGTGTTAGCGTGGTTATTTTATTAGTTGTT
TATATGTACAAATTGATATTTTGTATTTTATGAGATTTGACTTAGGTTTTGAATACGTAGTACACATATTTGTTTACATC
ATTATATTTTCTCTTTATGTCCATATATTTAAGTGTATAGTATTAATAATATTGTTTGATACACATATTTAGTATGAGTGGTA
GGTTTTGTTATATATATTTATAGTAATAATAGGTTTTATTGGTTATGTATTACCATGTACGATGATGTCGT

Figura 20: Sequência consenso obtido a partir do alinhamento de 11 acessos pelo BioEdit.

>LR697137.1:5353-6303 *Leishmania infantum* genome assembly, chromosome: 1

CTTGAGTTGTGGCGTGTGGTTGTTCTATATTATTTTAATAGATTATTGCATTTTAGTCATGTGTTTTAGGTATATTTTATTATT
TTTATGTTTACAGATGTTTAACTTTTATGTATGGACACACGTTTGTATTTTATATGTTGTATGTTTATATTGACATTTGTT
GATTGTGTATGATTTTTTTTATTACGATTTGTATTTTGATGTTTAAAGTGTAGTGTACGTTGTATGCATAAACCAAAATTTATTAAT
TTATTTTAAATTTTAAATTTAAGATAAGCGGAGAGGAAAGAAAAGGCTTTAACTTCAGGTTGTTTATTACGTGTTTATGGTGTAG
GTTTATGCTAGGTTTTTTTATTGTATGCAGATAATATGTGGTGTGTGTTAGCGTGGTTATTTTATTAGTTGTTTATATGTACAA
ATTGATATTTTGTATTTTATTTTATGAGATTTGACTTAGGTTTTGAATACGTAGTACACATATTTGTTTACATCATTATTATTTT
CCTTCTTTATGTCCATATATTTAAGTGTATAGTATTAATAATATTGTTTGATACACATATTTAGTATGAGTGGTAGGTTTTGTTAT
ATATATATTTATAGTAATAATAGGTTTTATTGGTTATGTATTACCATGTACGATGATGTCGT

Figura 21- Sequência consenso de *L. infantum*

Observando os resultados nas Tabelas 4 e 5, pode ser verificado que os valores obtidos na avaliação dos parâmetros, estão em concordância com os estudos já realizados. Na Tabela 4, os primers geram um produto de 306 pb, que sob a ação da endonuclease de restrição RsaI, origina 2 fragmentos (72 e 234 pb) no genoma de *Leishmania infantum* (Figura 21). Numa segunda combinação de primers (ITA1-F - ITA2-R2), do mesmo genoma, o amplicon obtido é mais amplo (672 pb), gerando 5 fragmentos de comprimentos variados, sob ação da mesma enzima de restrição (Tabela 4).

Tabela 4 – Descrição dos primers, sequências, tamanho do Amplicon e enzima para o genoma de *L. infantum*.

Par	Primer	Sequência	Amplicon	Enzima	Coordenadas	Comprimento (pb)
1	ITA1-F	CTTGTAGTTGTGGCGTGTTTG	306 pb	Rsal	1 - 234	234
	ITA2-R1	TTTCTTTCTCTCCGCTTATCTT			235 - 306	72
2	ITA1-F	CTTGTAGTTGTGGCGTGTTTG	672 pb		1 - 234	234
	ITA2-R2	ACGACATCATCGTACATGGTAAT			235 - 431	197
					432 - 491	60
					492 - 659	168
					660 - 672	13

>CP022652.1:8500-9450 *Leishmania donovani* strain pasteur kinetoplast, complete genome

CTTGTAGTTGTGGCGTGTTTG GTTCGATATTATTTTAATAGATTATTGCATTTTAGTCATGTGTTTTAGGTATATTTTATTATT
TTTATGTTTTAGTAGATGTTTTAACTTTTATGTATGGATACACGTTTTGTTTTTATATGTTGTATGTTTATATTGACATTTTGT
GATTGTGTATGATTTTTTTTATTACGATTGTATATTTGATGTTTTAAGTGTAGTGTATTTGTATGCATAAATCAAATTTATTAATT
TATTTAAATTTTAAATTTAAGATAAGCGGAGAGGAAAGAAAAGGCTTAACTTCAGGTTGTTTATTACGTGTTTATGGTGTAGG
TTTTAGTTTAGGTTTTTTTATTGTATGCAGATAATATGTGGCGTGTTTAGCGTGTTATTTTTAGTTGTTTTATATGTACAAA
TTGATATTTTGTTTATTTTATGAGATTTCGACTTAGGTTTTGTAATACGTAGTACACATATTTGTTTACATCATTATTATTTTC
CTTCTTTATGTCCATATATTTAAGTGTATAGTATTAATAATATTGTTTGATACACATATTTAGTATGAGTGGTAGGTTTTGTTATA
TATATATTTATAGTAATAAATAGGTTTTATTGGTTATGTATTACCATGTACGATGATGTCGT

Figura 22 – Sequência consenso de *L. donovani*.

Na Tabela 4, a sequência de 306 pb se mantém íntegra, mesmo com a presença da *RsaI*. A região específica de atuação da enzima *RsaI* (GTAC), não está disponível neste produto, caracterizando um polimorfismo de nucleotídeo único (SNP). No segundo par de iniciadores (Tabela 5), sob a ação da endonuclease de restrição *RsaI*, houve uma redução no número de fragmentos no genoma de *L. donovani* (Figura 12). A partir destes resultados, pode prever com uma PCR-RFLP *in silico*, a identificação das espécies de *L. donovani* e *L. infantum* através do número de fragmentos gerados pela ação da enzima *RsaI*.

Tabela 5 - Descrição dos primers, sequências, tamanho do Amplicon e enzima para o genoma de *L. donovani*.

Par	Primer	Sequência	Amplicon	Enzima	Coordenadas	Comprimento (pb)
1	ITA1-F	CTTGTAGTTGTGGCGTGTTTG	306 pb	RsaI	Sem cortes	306
	ITA2-R1	TTTCTTTCTCTCCGCTTATCTT				
2	ITA1-F	CTTGTAGTTGTGGCGTGTTTG	672 pb		1 - 431	431
	ITA2-R2	ACGACATCATCGTACATGGTAAT			432 - 491	60
					492 - 659	168
					660 - 672	13

Cortes et al (2006), monitorando 136 isolados de *L. infantum*, e investigando a necessidade do uso de marcadores moleculares na diversidade genética por PCR-RFLP. Os primers de cinetoplastídeo MC1 e MC2, específicos para *Leishmania Complexo donovani*, amplificou uma sequência parcial de kDNA em 447 bp. O marcador molecular em análise amplificou uma sequência parcial de DNA cinetoplastídeo (kDNA) em 447 bp. O ensaio concluiu que a população portuguesa de *L. infantum* apresentou um elevado grau de polimorfismo com 13 perfis identificados. No ensaio cinetoplastídeo DNA PCR-RFLP, foram utilizados 9 endonucleases, com sítios de restrição variando de 4 a 6 bp (*BglII*, *Bme1390I*, *DdeI*, *HpaII*, *PstI*, *RsaI*, *Sfcl*, *Vspl* e *XapI*). Os padrões de restrição obtidos com *PstI*, *Sfcl* e *XapI* eram monomórficos, enquanto os padrões obtidos com *BglII*, *Bme1390I*, *DdeI*, *HpaII*, *RsaI* e *Vspl*, eram polimórficos. Estas últimas seis enzimas foram consideradas ideais para genotipagem.

O uso da técnica de PCR no diagnóstico de leishmaniose tegumentar americana, foi utilizada por MARTINS et al (2009), onde o DNA de 119 amostras de sangue foi extraído e submetido a reação em cadeia da polimerase (PCR). Utilizaram-se primers do minicírculo do DNA do cinetoplasto (kDNA) de *Leishmania braziliensis*, cuja sequência-alvo gera um fragmento de 750 pares de bases. No total 58 (48,7%) indivíduos apresentaram amplificação positiva e 61 (51,3%) negativa.

A análise do polimorfismo do comprimento do fragmento de restrição (RFLP) pode prejudicar e tornar difícil uma interpretação, pois pequenas

diferenças no tamanho dos fragmentos e digestão incompleta, podem interferir na leitura correta dos padrões resultantes..

Leishmania (Viannia) guyanensis foi por muitos anos a única espécie identificada na Guiana Francesa, mas a identificação precisa da espécie não era confiável. SIMON *et al.* (2010) utilizou sete cepas de referência (*Leishmania (Leishmania) amazonensis*, *Leishmania (Viannia) lainsoni*, *Leishmania (Viannia) braziliensis*, *L. (V.) guyanensis*, *Leishmania (Viannia) naiffi*, *Leishmania (Leishmania) major*, *Leishmania (Leishmania) infantum*) e 112 amostras clínicas de lesões positivas. Desenvolveu a técnica de reação em cadeia da polimerase, usando o polimorfismo de comprimento de fragmento de restrição – com um fragmento de 615 pb do gene da RNA polimerase II e duas enzimas de restrição, *TspRI* e *HgaI*. A identificação do parasita *Leishmania (Viannia) guyanensis* foi confirmado em 84,4%, das 199 amostras de diferentes pacientes com várias *Leishmania spp.*

6 CONCLUSÃO

- As ferramentas de bioinformática adotadas para alinhamento, desenho e validação de primers *in silico*, foram eficazes para definir a região ideal para o anelamento dos iniciadores específicos;
- A técnica da PCR-RFLP *in silico*, foi eficaz para detecção de polimorfismo de única base (SNP) e sugestão com identificação das enzimas de restrição atuando no amplicon;
- A enzima de restrição RsaI, foi escolhida e utilizado para promover a identificação das espécies de Leishmanias, através da divergência genética existente entre *L. donovani* e *L. infantum*;
- Foram desenvolvidos e validados os seguintes marcadores moleculares:
F1 - 5'- CTTGTAGTTGTGGCGTGTTTG - 3'
R1 - 5'- TTTCTTTCCTCTCCGCTTATCTT - 3'
R2 - 5'- ACGACATCATCGTACATGGTAAT - 3'
- O produto amplificado da PCR *in silico* com os marcadores moleculares F1 e R1 (306 pb) do *L. infantum* é clivado, gerando dois fragmentos (72 pb e 234 pb);
- O produto amplificado da PCR *in silico* com os marcadores moleculares F1 e R2 (672 pb) do *L. infantum* é clivado, gerando cinco fragmentos (234,197, 168, 60 e 13 pb);
- O produto amplificado da PCR *in silico* com os marcadores moleculares F1 e R1 (306 pb) do *L. donovani*, não foi clivado;
- O produto amplificado da PCR *in silico* com os marcadores moleculares F1 e R2 (672 pb) do *L. donovani* é clivado, gerando quatro fragmentos (431, 168, 60 e 13 pb);
- Foi possível observar uma maior área do genoma mitocondrial e identificar através do alinhamento das sequências consenso, muitas regiões conservadas e outras com marcas de SNP;

- A PCR-RFLP *in silico*, mostrou-se adequada para a identificação de *Leishmania* a nível de espécie. Os marcadores moleculares desenvolvidos, foram validados com uma taxa de 85% a 100%, permitindo um excelente parâmetro na qualidade na confirmação dos *primers*. A variação genética nas sequências de mini círculos de kDNA entre isolados *L. donovani* e *L. infantum*, pode ser detectado através dos SNP presentes nos respectivos genomas, utilizando a endonuclease de restrição RsaI.

7 REFERENCIAS

- ABD-ELSALAM, K. A. **Bioinformatic tools and guideline for PCR primer design**. African Journal of Biotechnology, 2003. v. 2, n. 5, p. 91–95.
- ALPAYDIN, E. Introduction to machine learning. **Cambrige**: MIT Press, 2004.
- ALMEIDA, S.C.B.; LEITE, I.S.; CARDOSO, C.O. **Leishmaniose tegumentar americana: perfil epidemiológico no município de rio branco - acre (2007-2015)**. South American J of Basic Education, Technical and Technological. 2018. p. 20-23.
- AL-JAWABREH, A.; SCHOENIAN, G.; HAMARSHEH, O.; PRESBER, W. **Clinical diagnosis of cutaneous leishmaniasis: a comparison study between standardized graded direct microscopy and ITS1-PCR of Giemsa-stained smears**. Acta Trop. Alvar J, Cañavate C, Molina R, 2006. p.55-61.
- ARBEX, W.; COSTA, V.S.; SILVA, M.V.G. Bioinformática como ferramenta nas pesquisas atuais. **Anais do III Encontro de Genética e Melhoramento, Viçosa, UFV, 2006**.
- ATTWOOD, T. K. et al. **Concepts , Historical Milestones and the Central Place of Bioinformatics in Modern Biology : A European Perspective**. In: Bioinformatics - Trends and Methodologies. [s.n.]. p. 3–39.
- AVISE, J.C. Molecular markers, Natural history and Evolution. **Chapman & Hall**. 1994.
- BASIC local **alignment search tool**. Disponível em: <<http://www.blast.ncbi.nlm.nih.gov/>> Acesso em: 29 dez. 2019.
- BIFELD, E.; CLOS, J. The genetics of Leishmania virulence. **Medical Microbiology and Immunology**, 2015.
- BOTSTEIN, D.; WHITE, R.L.; SKOLNICK, M.; DAVIS, R.W. Construction of a Genetic Linkage Map in Man Using Restriction Fragment Length Polymorphisms. **Am JHum Genet**, 1980. p. 314-331,
- BOITE MC, MAURICIO IL, MILES MA, CUPOLILLO E. **New insights on taxonomy, phylogeny and population genetics of leishmania (viannia) parasites based on multilocus sequence analysis**. PLoS Negl Trop Dis. 2012.
- BRASIL. **Manual de Recomendação para Diagnóstico, Tratamento e Acompanhamento da Co-infecção Leishmania/HIV**. Ministério da Saúde (MS)/ Secretária de Vigilância em Saúde (SVS). Programa Nacional de DST e AIDS. Brasília. DF, 2004.
- BRASIL. **Ministerio da saude. Manual da vigilância e controle da leishmania visceral**. Brasilia, 2006. p. 120.

- CALVOPINA, M.; ARMIJOS, R.X.; HASHIGUCHI, Y. **Epidemiology of leishmaniasis in Ecuador: current status of knowledge** -- a review. *Memorias do Instituto Oswaldo Cruz*. [Review]. 2004. p. 663-72
- CAVALCANTI, M. C. et al. **Managing structural genomic workflows using Web services**. *Data & Knowledge Engineering*, 2005. v. 53, n. 1, p. 45–74
- COSTA CHN, PEREIRA HF, ARAÚJO MV. Visceral leishmaniasis epidemic in Piauí State, Brazil, 1980-1986. **Revista de Saúde Pública**. outubro de 1990. 24(5):361–72.
- CRUZ, G.S.; FECHINE, M.A.B; COSTA, E.C. **Leishmaniose Tegumentar Americana: Aspectos Clínicos, Epidemiológicos e Influência de Fatores Predisponentes**. Repositório- Unilab, 2016.
- CUPOLILLO, E.; GRIMALDI, G.; MOMEN, H. **Genetic diversity among Leishmania (Viannia) parasites. Annals of tropical medicine and parasitology**. [Research Support, NonU.S. Gov't], 1997. 91(6):617-26.
- CUPOLILLO, E.; MOMEN, H.; GRIMALDI, G.J.R. **Genetic diversity in natural populations of New World Leishmania**. *Memorias do Instituto Oswaldo Cruz*, 1998. 93(5):663-8.
- DA SILVA, L.A.; DE SOUSA, C.D.; DA GRACA, 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 : journal of molecular epidemiology and evolutionary genetics in infectious diseases*. 2010. 10(1) p. 77-83
- DE ARRUDA, **Leishmanioses**. In: **manual da zoonoses**. Parana V.1, 2ed. 90p 2010. Disponível em: <<http://www.zoonoses.org.br/zoonoses>> Acesso em 14 dez 2019
- DEGRAVE, W.; FERNANDES O, CAMPBELL D, BOZZA M, LOPES UG. **Use of molecular probes and PCR for detection and typing of Leishmania—a mini-review**. *Mem. Inst. Oswaldo Cruz*. 1994. 89: p. 463–469
- DESJEUX, P., 2004. **Leishmaniasis: Current situation and new perspectives**. *Comp Immun Microbiol Infect Dis*, 27: p. 305-318.
- DIBA, K.; MIRHENDI, H.; KORDBACHEH, H.P.; REZAIE, S. Development of RFLP-PCR method for the identification of medically important Aspergillus species using single restriction enzyme MwoI. **Brazilian Journal of Microbiology**, 2014. 45, 2, p. 503-507.
- DRUGS FOR NEGLECTED DISEASES INITIATIVE. **Leishmaniose visceral**. Disponível em: <http://dndi.org.br/pt/doencasnegligenciadas/leishmanioses/estrategia-dadndi.html> Acesso em 03 dez 2019.
- DUJARDIN, J.C, DE DONCKER S, JACQUET D, BANULS AL, BALAVOINE M, VAN BOCKSTAELE D, TIBAYRENC M, AREVALO J, LE RAY D. Clonal

propagation and the fast generation of karyotype diversity: **An in vitro Leishmania model. Parasitology.** 2007. p. 3-9.

DURAN, C. et al. **Molecular Genetic Markers:** Discovery, Applications, Data Storage and Visualisation, 2009. v. 61, p. 16–27.

FARAH, S.B. **DNA segredos e mistérios.** Editora Savier. 2000. p. 276.

FALCÃO DE OLIVEIRA E, OSHIRO ET, FERNANDES WS, MURAT PG, DE MEDEIROS MJ, SOUZA AI, DE OLIVEIRA AG, GALATI EA. Experimental infection and transmission of *Leishmania* by *Lutzomyia cruzi* (Diptera: Psychodidae): Aspects of the ecology of parasite-vector interactions. *PLoS Negl Trop Dis.* 2017 Feb 24;11(2):e0005401. doi: 10.1371/journal.pntd.0005401. PMID: 28234913; PMCID: PMC5342273.

FERREIRA, M.E.; GRATTAPAGLIA, D. **Introdução ao uso de marcadores moleculares em análise genética.** Brasília: EMBRAPA-CENARGEN Documentos 20, 1998. 3.ed. p. 220.

FULLER, J. C. et al. **Biggest challenges in bioinformatics. EMBO reports,** 2013. v. 14, n. 4, p. 302–4.

GARIBYAN, L.; AVASHIA, N. Research Techniques Made Simple: Polymerase Chain Reaction (PCR). **J. Invest Dermatol.** 2013. p. 133-3.

GHASEMLOO, H.; RASTI, S.; DELAVARI, M.; DOROODGAR, A. **Molecular Diagnosis of Clinical Isolates of Cutaneous Leishmaniasis Using ITS1 and KDNA Genes and Genetic Polymorphism of Leishmania in Kashan, Iran.** *Pak J Biol Sci.* 2016;19(3):136-142. doi: 10.3923/pjbs.2016.136.142. PMID: 29023050.

GOMES, A.H.S.; FERREIRA, I.M.R.; LIMA, M.L.S.R.; et al. **PCR identification of Leishmania in diagnosis and control of canine leishmaniasis. Veterinary Parasitology,** 2007. p. 234-241.

GRACA, G.C.; VOLPINI AC, ROMERO, G.A.; OLIVEIRA, M.P.O.; HUEB, M.; PORROZZI, R.; BOITE, M.C.; CUPOLILLO, E. **Development and validation of PCR-based assays for diagnosis of American cutaneous leishmaniasis and identification of the parasite species.** *Memorias do Instituto Oswaldo Cruz.* 2012. 107(5) p. 664-74.

HARTL, D.L. **Princípios de Genética de População.** Tradução I.F. Afonso. Terceira Edição . Riberão Preto .FUNPEC, 1V. 2008.

HE, Z. Introduction to bioinformatics. **Data Mining for Bioinformatics Applications.** 2015. cap.2., p. 11-17.

IVENS, A. C. et al. **The genome of the kinetoplastid parasite, Leishmania major.** *Science (New York, N.Y.).* 2005. v. 309, n. 5733, p. 436–42.

KAYE, P.; SCOTT, P. **Leishmaniasis: complexity at the host-pathogen interface. Nature reviews. Microbiology,** 2011. v. 9, n. 8, p. 604–615.

LI, J. et al. Hundreds of microsatellites for genotyping *Plasmodium yoelii* parasites. **Molecular and Biochemical Parasitology**, 2009. v. 166, p. 153–158.

LIN, R.H.; LAI, D.H.; ZHENG, L.L.; WU, J.; LUKES, J.; HIDE, G.; LUN, Z.R. **Analysis of the mitochondrial maxicircle of *Trypanosoma lewisi*, a neglected human pathogen.** *Parasit Vectors*, 2015. 8. p. 665.

LUCAS, C.M.; FRANKE, E.D.; CACHAY, M.I.; TEJADA, A.; CRUZ, M.E.; KREUTZER, R.D.; BARKER, D.C.; MCCANN, S.H.; WATTS, D.M. Geographic distribution and clinical description of leishmaniasis cases in Peru. **The American journal of tropical medicine and hygiene**. [Research Support, U.S. Gov't, Non-P.H.S.]. 1998. 59(2). p. 312-7.

LUCCHI, N.W.; OBERSTALLER, J.; KISSINGER, J.C.; UDHAYAKUMAR, V.: Malaria diagnostics and surveillance in the post-genomic era. **Public Health Genomics** 2013, 16, p. 37–43.

Ministério da Saúde. Brasil. **Manual de Vigilância e Controle da Leishmaniose Visceral**. Brasília. 2003.

MCMANUS, D.P.; BOWLES, J. **Molecular genetic approaches to parasite identification: their value in diagnostic parasitology and systematics.** *Int J Parasitol* 1996. 26, p. 687–704.

MORENO. J.; NIETO, J. **Canine leishmaniasis.** *Adv Parasitol*. 2004. 57, p. 1-88.

OLIVEIRA, E.F.; OSHIRO, E.T.; FERNANDES, W.S.; MURAT, P.G.; DE MEDEIROS, M.J.; SOUZA, A.I.; DE OLIVEIRA, A.G.; GALATI, E.A. Experimental infection and transmission of *Leishmania* by *Lutzomyia cruzi* (Diptera: Psychodidae): Aspects of the ecology of parasite-vector interactions. **PLoS Negl Trop Dis**. 2017 Feb 24;11(2):e0005401. doi: 10.1371/journal.pntd.0005401. PMID: 28234913; PMCID: PMC5342273.

ORGANIZAÇÃO MUNDIAL DA SAÚDE/ ORGANIZAÇÃO PAN- AMERICANA DE SAÚDE. Leishmanioses: Informe Epidemiológico nas Américas: Washington: **Organização Pan-Americana da Saúde**; 2018. Disponível em: < <https://www.paho.org/leishmaniasis>>. Acesso em: Set/2019

ORGANISATION MONDIALE DE LA SANTÉ. **Lutte contre les leishmanioses. Serie de rapports techniques**. Geneve: OMS, 1990. p. 176.

PAGANI, I. et al. The Genomes OnLine Database (GOLD) v.4: status of genomic and metagenomic projects and their associated metadata. **Nucleic Acids Research**, v. 40, n. Database issue, 2012. p. 571–579.

PAVINATO, V.A.C. Tese. **Variabilidade e estrutura genética de populações de *Alabama argillacea* (Hüb.) (Lepidoptera: Noctuidae) no Brasil:** subsídios para o manejo da resistência à toxina Cry1Ac em algodão geneticamente modificado. Piracicaba, 2010. p. 121.

PEACOCK, C. S. et al. **Comparative genomic analysis of three Leishmania species 87 that cause diverse human disease.** *Nature genetics*, 2007. v. 39, n. 7, p. 839–47.

RANGEL EFR, LAINSON R. **Ecologia das Leishmanioses: Transmissores de Leishmaniose Tegumentar Americana.** In Rangel EFR, Lainson R. *Flebotomíneos do Brasil*. Rio de Janeiro: FIOCRUZ; 2003. p. 291-310.

RAY, D. S. **Conserved sequence blocks in kinetoplast minicircles from diverse species of trypanosomes.** *Mol. Cell. Biol.* 1989. 9, p. 1365–1367.

RAZA, K. Application of data mining in Bioinformatics. **Indian Journal of Computer Science and Engineering**, New Delhi, 2010. v. 1, n. 2, p. 114-118.

REITHINGER, R.; DUJARDIN, J.-C. Molecular diagnosis of leishmaniasis: current status and future applications. **Journal of clinical microbiology**, 2007. v. 45, n. 1, p. 21–5.

RIBEIRO, N.S.; JETIMANE, J.L; MILITÃO, E; MAQUIA, I; CHIRIZANE, C; DE SOUSA, C; ALVES, T; VELOSO M.M; GOULAO, L.F; RIBEIRO-BARROS A.I. Ecological characterization of an ex situ conservation plantation in southeastern Mozambique. **African Journal of Ecology**. 10p.[serial online]. 2016.

ROGERS, M.B.; HILLEY, J.D.; DICKENS NJ, WILKES J, BATES PA, DEPLEDGE DP, HARRIS D, HER Y, HERZYK P, IMAMURA H, OTTO TD, SANDERS M, SEEGER K, DUJARDIN JC, BERRIMAN M, SMITH DF, HERTZ-FOWLER C, MOTTRAM JC. **Chromosome and gene copy number variation allow major structural change between species and strains of Leishmania.** *Genome Res*, 2011. 21(12), p. 2129-42.

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

SARASWATHY, N.; RAMALINGAM, P. Concepts and techniques in genomics and proteomics. **Genome sequence assembly and annotation**, 2011. p. 109-121

SERENO, D.; A, MOHAMMAD.;D. FRANCK.; O, BRUNO.; M, HOOMAN.; P, PASCALE. **What pre-Columbian mummies could teach us about South American leishmaniasis?.** *Pathogens and Disease*. 75. 10.1093/femspd/ftx019. 2017.

SIMON, S.; VERON, V.; CARME, B. *LEISHMANIA spp.* identification by polymerase chain reaction–restriction fragment length polymorphism analysis and its applications in French Guiana, **Diagnostic Microbiology and Infectious Disease**, 2010. v. 66, Issue 2, p. 175-180.

SHELINE, C., AND D. S. RAY. Specific discontinuities in *Leishmania tarentolae* minicircles map within universally conserved sequence blocks. **Mol. Biochem. Parasitol.** 1989. 37, p.151–157.

SOARES-BEZERA, R. J.; LEON, L.; GENESTRA, M. Recent advances on leishmaniasis chemotherapy: intracellular molecules as a drug target. **Rev. Bras. Cienc. Farm.** vol.40 no.2 São Paulo 2004.

TEIXEIRA, S. M. et al. **Trypanosomatid comparative genomics:** Contributions to the study of parasite biology and different parasitic diseases. Genetics and molecular biology, 2012. v. 35, n. 1, p. 1–17.

TELLERIA, J, LAFAY B, VIRREIRA M, BARNABE C, TIBAYRENC M, SVOBODA M, 2006. ***Trypanosoma cruzi*: sequence analysis of the variable region of kinetoplast minicircles.** *Exp Parasitol*, 2006. 114, p. 279–288.

WAN, Q.H.; WU, H.; FUJIHARA, T.; FANG, S.G. Which genetic marker for which conservation genetics issue? **Electrophoresis**. 2004. 25 p. 2165–2176.

WHO. World Health **Organization Leishmaniasis Epidemiological situation.** [Online].; 2018 [cited 2018 junho 19. Available from:<http://www.who.int/leishmaniasis/burden/en/>.]

WONG, S. S. et al. Molecular diagnosis in clinical parasitology: When and why? 89 **Experimental biology and medicine (Maywood, N.J.)**, 2014.