

UNIVERSIDADE FEDERAL DE SÃO CARLOS  
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE  
PROGRAMA DE PÓS GRADUAÇÃO EM GENÉTICA EVOLUTIVA  
E BIOLOGIA MOLECULAR

TALITA YURI TAKAHASHI

TRIAGEM DE ISOLADOS CLÍNICOS DE *Leishmania* sp.  
PARA SEQUENCIAMENTO GENÔMICO

SÃO CARLOS -SP  
2020

TALITA YURI TAKAHASHI

TRIAGEM DE ISOLADOS CLÍNICOS DE *Leishmania* sp. PARA SEQUENCIAMENTO  
GENÔMICO

Dissertação apresentada ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular do Centro de Ciências Biológicas e da Saúde da Universidade Federal de São Carlos, como parte dos requisitos para obtenção do Título de Mestre em Genética Evolutiva e Biologia Molecular.

Orientadora: Dra. Sandra Regina Costa Maruyama

São Carlos-SP  
2020



# UNIVERSIDADE FEDERAL DE SÃO CARLOS

Centro de Ciências Biológicas e da Saúde  
Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular

---

## Folha de Aprovação

---

Defesa de Dissertação de Mestrado da candidata Talita Yuri Takahashi, realizada em 22/06/2020.

### Comissão Julgadora:

Profa. Dra. Sandra Regina Costa Maruyama (UFSCar)

Prof. Dr. Otávio Henrique Thiemann (IFSC/USP)

Profa. Dra. Laila Alves Nahum (FIOCRUZ)

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

O Relatório de Defesa assinado pelos membros da Comissão Julgadora encontra-se arquivado junto ao Programa de Pós-Graduação em Genética Evolutiva e Biologia Molecular.

Dedico este trabalho

Aos meus pais, os dois maiores  
incentivadores das realizações  
dos meus sonhos. Sem eles  
nada seria possível.

## AGRADECIMENTOS

Em primeiro lugar, gostaria de agradecer minha mãe, pai, irmã, madrinha, namorado e avós queridas pelo apoio incondicional, paciência em dias difíceis e compreensão nos momentos de ausência. Agradeço por segurarem a minha mão e caminharem comigo nessa trajetória chamada vida.

À minha orientadora Dra. Sandra Regina Costa Maruyama pela oportunidade de integrar esse trabalho e confiança depositada em mim. Obrigada pela orientação nesses últimos anos, discussões construtivas, contribuição para meu crescimento profissional, por todos os conselhos e incentivos que me permitiram ver o mundo científico com outros olhos.

Ao meu co-orientador Dr. Daniel Jeffares da University of York, Reino Unido, pela recepção amigável durante o período de estágio sanduíche, pelas discussões construtivas sobre o projeto e pela prontidão em me ajudar sem medir esforços.

À minha tutora Dra. Sarah Forrester da University of York, Reino Unido, pelos ensinamentos em análises bioinformática e pelos momentos agradáveis que fizeram me sentir em casa.

À Dra. Sally James pela contribuição no sequenciamento genômico por Oxford Nanopore Technology.

Ao Dr. John Davey pela assistência na manipulação do sequenciamento *long reads*, pelas reuniões e discussões de resultados e por seu profissionalismo.

Ao Prof. Dr. Felipe Roberti Teixeira por ser sempre solícito, presente e por compartilhar seu laboratório para o desenvolvimento desse trabalho.

Aos meus colegas do laboratório de Bioquímica Celular e Genômica Funcional e Comparativa pela contribuição direta e indireta da minha formação profissional, em especial ao Caio Oliveira, Camila Rolemberg, Juliana Garcia, Luana Rogerio, Nayore Takamiya, Patrícia

Passos e Thaiane Robeldo, agregada do laboratório, pela amizade e descontração no ambiente de trabalho.

Aos professores Dr. Iran Malavazi, Dr. Anderson F. Cunha, Dra. Lisandra G. Borges, Dr. Ricardo Borra, Dr. Pedro Galleti, Dr. Marcelo de Bello Cioffi e Dr. Marcos Roberto Chiaratti pela disponibilização de equipamentos em seus laboratórios.

Aos colaboradores desse projeto Dr. João Santana da Silva, Dr. José Marcos Ribeiro e Dr. Roque Pacheco por tornar este trabalho possível.

Aos membros da banca examinadora pela disponibilidade em participar da minha defesa e pela contribuição importante para a melhoria deste trabalho.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES pela bolsa de estudos durante o primeiro ano do mestrado e pelo fomento ao Programa de Pós-Graduação.

À Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP pelo financiamento de recursos destinados ao desenvolvimento deste projeto (processo 2016/20258-0) e pelas concessões de bolsa de estudos durante o último ano do mestrado (processo 2019/03095-9), incluindo a Bolsa de Estágio de Pesquisa no Exterior (processo 2019/12142-0) na University of York, Reino Unido.

A todos os docentes e funcionários do Programa de Pós-Graduação em Genética Evolutiva pelo suporte, empenho e coordenação que possibilitaram a realização deste trabalho.

*Biologists must constantly keep in mind that what they see was not designed, but rather evolved.*

Francis Crick

What Mad Pursuit (1990)

## RESUMO

As leishmanioses são um grupo de doenças infecciosas tropicais causadas por protozoários do gênero *Leishmania* e transmitidas ao ser humano pela picada do mosquito flebotomíneo infectado. A leishmaniose visceral (LV) é a forma clínica mais grave da doença e pode ser letal quando não tratada ou o tratamento falha. O Brasil constitui um dos 13 países com maior número de novos casos de LV no mundo causada por *Leishmania infantum*, com grande foco no Nordeste. Os tratamentos são pouco desenvolvidos e o número de fatalidades expande todo ano, acometendo principalmente crianças de até 10 anos de idade. Estudo recente apontou a ocorrência de parasito semelhante a *Crithidia fasciculata* em isolados clínicos (designados LVH60 e LVH60a) obtidos de um caso fatal de LV. O desenvolvimento da doença está relacionado com fatores da interface parasito/hospedeiro e a identificação correta do parasito, bem como análises de conteúdo genômico de isolados clínicos possibilitam a compreensão dos fatores da doença e resposta ao tratamento relacionadas à diversidade genética do parasito. Desse modo, um dos objetivos desse trabalho foi realizar a triagem de isolados clínicos obtidos de pacientes diagnosticados com LV na região de Aracaju - Sergipe, a fim de elencar quais amostras serão submetidas para sequenciamento completo do genoma (WGS) na plataforma Illumina para obter sequências genômicas completas dos parasitos relacionados à LV nesta área endêmica. A triagem englobou o estudo da morfologia dos parasitos, caracterização de regiões genômicas conservadas por meio de análise de sequência de DNA e filogenética molecular. Os dados gerados foram úteis para a construção de um painel informativo sobre as amostras analisadas, as quais orientaram na seleção de isolados clínicos para sequenciamento genômico. Paralelamente, este trabalho também teve como objetivo sistematizar a análise de dados de sequenciamento genômico obtidos de clones e culturas policlonais de isolados clínicos para a análise de Genômica Comparativa entre *Leishmania* e *Crithidia-like* em estudos futuros do grupo. Por fim, um genoma refinado de *Crithidia-like* foi obtido para um clone de LVH60a por sequenciamento *Oxford Nanopore Technology* (ONT), que resultou na montagem genômica final de 38 contigs perfazendo o tamanho predito de genoma em 34,4 Mb. Portanto, esse trabalho foi essencial para compreender a ocorrência de parasitos não-*Leishmania* em região endêmica de LV em Aracaju e estudar o genoma e variação genética para basear futuras estratégias de vigilância e intervenção na saúde pública decorrente desse novo parasito.

Palavra-chave: Leishmaniose Visceral Humana, *Crithidia-like*, Análise Genômica

## ABSTRACT

Leishmaniasis is a group of infectious tropical disease caused by protozoan from genus *Leishmania* and transmitted to humans through bites of infected sand flies. Visceral Leishmaniasis (VL) is the most severe clinical form of the disease and can be lethal if untreated or treatments fail. Brazil composes one of the 13 countries with the highest number of new VL cases in the world caused by *Leishmania infantum*, with a great focus on Northeastern. Treatments are poorly developed and the fatality rate expands every year affecting mainly children under 10 years-old. Recent study has shown the occurrence of a similar *Crithidia fasciculata* parasite in clinical isolates (LVH60 and LVH60a) in a fatal VL case. The disease onset is related to parasite/host interface factors. Parasite identification and genomic analyses of the pathogen enable to understand the intrinsic factors and genetic diversity of parasite related to the disease outcome and response to treatment. Thus, one of the aims in this study was to perform the screening of clinical isolates from patients diagnosed with VL in Aracaju - Sergipe region, to select samples for Whole-Genome Sequencing (WGS) in Illumina platform for obtaining complete genomic sequences of parasites capable of causing VL in this endemic area. The screening encompassed the parasite morphological analysis, molecular characterization of conserved genomic regions through DNA sequencing analysis and molecular phylogenetics. Engendered data were useful for constructing an informative panel about the analyzed samples, which guided to the selection of clinical isolates for genome sequencing. In parallel, this study also had as aim to systematize the genome analyzes of WGS data obtained from previous work, which will assist the Comparative Genomic analyses between *Leishmania* and *Crithidia*-like to our group's further studies. Finally, a polished genome of *Crithidia*-like parasite was obtained from a clonal sample of LVH60a with Oxford Nanopore Technology (ONT), which resulted in a final genome assembly of 38 contigs and a predicted genome size of 34.4 Mb in length. This study was essential to understand the occurrence of non-*Leishmania* parasites in the endemic region of LV in Aracaju and for studying the genome and genetic variation to lead future strategy analyses and intervention in public health caused by this new parasite.

Keyword: Human Visceral Leishmaniasis, *Crithidia*-like, Genomic Analysis

## LISTA DE FIGURAS

FIGURA 1 - O ciclo de vida do parasito <i>Leishmania</i> .....	20
FIGURA 2 - Classificação taxonômica de protozoários parasitas da família Trypanosomatidae..	21
FIGURA 3 - Fluxograma do delineamento da triagem para seleção de isolados clínicos <i>Crithidia</i> -like para sequenciamento genômico .....	29
FIGURA 4 - Ilustração do processamento e mapeamento dos arquivos de sequenciamento Illumina de culturas policlonais dos isolados clínicos LVH60 e LVH60a .....	39
FIGURA 5 - Fluxograma do preparo da biblioteca para sequenciamento por <i>Oxford Nanopore Technology</i> .....	47
FIGURA 6 - Fluxograma do processamento de dados de sequenciamento <i>Oxford Nanopore Technology</i> de LVH60a-c1 para obtenção de genoma final montado e melhorado.....	48
FIGURA 7 - Curva de crescimento in vitro de promastigotas em temperaturas simulando o hospedeiro invertebrado (25 °C) e hospedeiro vertebrado (35 °C) .....	57
FIGURA 8 - Formas promastigotas extracelulares em temperatura de 35 °C .....	58
FIGURA 9 - Formas extracelulares de diferentes culturas coradas em lâmina .....	61
FIGURA 10 - Árvore filogenética de região da subunidade 18S de rRNA (SSU rRNA) de amostras policlonais de isolados clínicos .....	68
FIGURA 11 - Alinhamento representativo das sequências 18S do RNA ribossomal SSU de culturas clonais e policlinais de isolados clínicos de leishmaniose visceral humana .....	69
FIGURA 12 - Alinhamento representativo das sequências LinJ31seq de culturas policlinais de isolados clínicos de leishmaniose visceral humana .....	70
FIGURA 13 - Alinhamento representativo das sequências Crid2.1seq de culturas clonais e policlinais de isolados clínicos de leishmaniose visceral humana .....	71
FIGURA 14 - Profundidade de cobertura de <i>reads</i> do isolado LVH60 policlonal nos genomas de referência de <i>L. infantum</i> e <i>C. fasciculata</i> concatenados .....	77
FIGURA 15 - Profundidade de cobertura de <i>reads</i> do isolado LVH60a policlonal nos genomas de referência de <i>L. infantum</i> e <i>C. fasciculata</i> concatenados .....	78
FIGURA 16 - Porcentagens dos cromossomos (cálculo pela extensão em pb) das referências <i>L. infantum</i> e <i>C. fasciculata</i> mapeadas por <i>reads</i> do isolado LVH60 policlonal .....	79

FIGURA 17 - Porcentagens dos cromossomos (cálculo pela extensão em pb) das referências <i>L. infantum</i> e <i>C. fasciculata</i> mapeadas por <i>reads</i> do isolado LVH60a policlonal .....	79
FIGURA 18 - Perfil de qualidade da <i>reads</i> de sequenciamento Oxford Nanopore Technology do clone <i>Crithidia</i> -like LVH60a-C1 .....	81
FIGURA 19 - Relatório Tapestry indicando os <i>contigs</i> únicos gerados na montagem bruta .....	83
FIGURA 20 - Alinhamento genômico dos <i>contigs</i> filtrados do clone <i>Crithidia</i> -like LVH60a-C1 com a referência <i>C. fasciculata</i> .....	84
FIGURA 21 - Comparação dos genomas <i>Crithidia fasciculata</i> e <i>Crithidia</i> -like em representação por circosplot .....	85
FIGURA 22 - Alinhamento completo do <i>tig.</i> 13 do clone <i>Crithidia</i> -like LVH60a-C1 com o cromossomo 1 da referência <i>C. fasciculata</i> . A referência foi obtida no banco de dados TrytripDB e o alinhamento foi produzido por NUCmer Dotplot.....	88
FIGURA 23 - Alinhamento dos <i>contigs</i> <i>tig.</i> 7 e <i>tig.</i> 28 do clone <i>Crithidia</i> -like LVH60a-C1 com o cromossomo 22 da referência <i>C. fasciculata</i> . A referência foi obtida no banco de dados TrytripDB e o alinhamento foi produzido por NUCmer Dotplot.....	88
FIGURA 24 - Alinhamento do <i>contig</i> <i>tig.</i> 37 do clone <i>Crithidia</i> -like LVH60a-C1 com os cromossomos 19 e 29 da referência <i>C. fasciculata</i> . A referência foi obtida no banco de dados TrytripDB e o alinhamento foi produzido por NUCmer Dotplot.....	89
FIGURA 25 - Profundidade de cobertura de <i>reads</i> Illumina de amostras clonais de LVH60 e LVH60a mapeadas em genoma referência <i>Crithidia</i> -like .....	92
FIGURA 26 - Porcentagem dos <i>contigs</i> (cálculo pela extensão em pb) da referência <i>Crithidia</i> -like mapeadas por <i>reads</i> de amostras clonais de LVH60 e LVH60a .....	94
FIGURA 27 - Alinhamento do <i>contig</i> <i>tig.</i> 1 de comprimento 12 Kb de <i>Crithidia</i> -like (LVH60a-C1) com <i>reads</i> do isolado LVH60a-C1 .....	95
FIGURA 28 - Alinhamento do <i>contig</i> <i>tig.</i> 2 de comprimento 19 Kb de <i>Crithidia</i> -like (LVH60a-C1) com <i>reads</i> do isolado LVH60a-C1 .....	95
FIGURA 29 - Ploidia estimada de amostras clonais de LVH60 e LVH60a em relação aos 38 <i>contigs</i> da referência <i>Crithidia</i> -like .....	97

## LISTA DE TABELAS

TABELA 1 - Sequências iniciadoras e reversas para amplificação de marcadores genéticos .....	34
TABELA 2 - Determinação do tempo e temperatura para amplificação por PCR .....	35
TABELA 3 - Parasitos clonais de LVH60 e LVH60a para sequenciamento genômico .....	45
TABELA 4 - Painel informativo da caracterização molecular e análise morfológica para triagem de isolados clínicos .....	62
TABELA 5 - Modelos de substituição de nucleotídeos pelo método de Máxima Verossimilhança... ..	67
TABELA 6 - Genoma predito dos parasitos <i>Leishmania infantum</i> , <i>Crithidia fasciculata</i> e <i>Crithidia-like</i> .....	72
TABELA 7 - Alinhamento genômico entre <i>Crithidia-like</i> e referências <i>Leishmania infantum</i> e <i>Crithidia fasciculata</i> .....	75
TABELA 8 - Matriz de identidade entre os genomas sequenciados e as referências <i>Crithidia fasciculata</i> e <i>Leishmania infantum</i> .....	76
TABELA 9 - Características gerais dos dados de sequenciamento <i>Oxford Nanopore Technology</i> do clone <i>Crithidia-like</i> LVH60a-C1 .....	80
TABELA 10 - Características gerais da montagem bruta do genoma do clone <i>Crithidia-like</i> LVH60a-C1 .....	82
TABELA 11 - Características do genoma <i>Crithidia-like</i> em comparação a referência <i>C. fasciculata</i> .....	86
TABELA 12 - Evolução dos processos de montagem e melhoramento do genoma <i>Crithidia-like</i> sequenciado por <i>Oxford Nanopore Technology</i> .....	90

## LISTA DE SIGLAS E ABREVIATURAS

LV – Leishmaniose Visceral

LVH – Leishmaniose Visceral Humana

HUUFS – Hospital Universitário da Universidade Federal de Sergipe

DGE – Departamento de Genética e Evolução

sp. – espécie

PCR – reação de polimerase em cadeia

WGS – sequenciamento genômico total

RNAse – ribonuclease

DNA – ácido desoxirribonucleico

rk39 – Recombinant K39

Mb – megabase

Kb – kilobase

pb – pares de base

M – Molar

mM – milimolar

uM – micromolar

pmol – picomolar

mL – mililitro

uL – microlitro

µg – micrograma

ng – nanograma

g – unidade de medida de força centrífuga relativa

s – segundo

°C – graus Celsius

pH – potencial hidrogeniônico

PBS – tampão fosfato salino

DMSO – dimetilsulfóxido

SBF – soro bovino fetal

rRNA - ácido ribonucleico ribossomal

SSU – menor subunidade ribossomal

ITS – espaçador do transcrito interno

GAPDH - gliceraldeído 3-fosfato-desidrogenase

Tris – hidroximetil aminometano

HCl –ácido clorídrico

KCl –cloreto de potássio

MgCl<sub>2</sub> – Cloreto de magnésio

dNTP – desoxinucleosídeos trifosfatos

EB – tampão de eluição

EDTA – ácido etilenodiamino tetracético

MSA – alinhamento múltiplo de sequência

TriTrypDB – banco de dados de tripanossomatídeos

k-mer – subsequências de comprimento contida em uma sequência biológica

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b> .....	<b>17</b>
1.1	LEISHMANIOSES .....	17
1.2	O PARASITO.....	19
1.3	CONTEXTO E RELEVÂNCIA DO ESTUDO.....	24
<b>2</b>	<b>OBJETIVOS</b> .....	<b>26</b>
<b>3</b>	<b>MATERIAL E MÉTODOS</b> .....	<b>28</b>
3.1	TRIAGEM DE ISOLADOS CLÍNICOS.....	28
3.1.1	Amostras .....	28
3.1.2	Cultura de Promastigotas.....	30
3.1.3	Obtenção de clones a partir das culturas promastigotas.....	31
3.1.4	Crescimento Celular .....	31
3.1.5	Caracterização Morfológica .....	32
3.1.6	Extração de DNA.....	32
3.1.7	Reação em Cadeia da Polimerase (PCR) .....	33
3.1.7.1	Marcadores genéticos .....	33
3.1.7.2	Reações .....	34
3.1.8	Sequenciamento dos Produtos de PCR .....	36
3.1.9	Análise das sequências moleculares.....	37
3.2	ANÁLISE GENÔMICA .....	38
3.2.1	Isolados Clínicos Previamente Sequenciados .....	38
3.2.1.1	Mapeamento de Reads .....	39
3.2.1.2	Alinhamento de Contigs em Genomas Referência .....	43
3.2.2	Sequenciamento de Clones de Isolados Clínicos .....	44
3.2.2.1	Sequenciamento Genômico por Oxford Nanopore Technology (ONT) .....	45
3.2.2.1.1	Recursos Computacionais .....	47
3.2.2.1.2	Análise de Long Reads.....	48
3.2.2.2	Sequenciamento Genômico por Illumina.....	52
3.2.2.2.1	Estimativa de Ploidia.....	53
<b>4</b>	<b>RESULTADOS</b> .....	<b>56</b>
4.1	CRESCIMENTO CELULAR .....	56
4.2	CARACTERIZAÇÃO MORFOLÓGICA DE PROMASTIGOTAS.....	59
4.3	PCR DE PARASITO ISOLADO EM CULTURA .....	62
4.4	ANÁLISES DE SEQUÊNCIA E FILOGENÉTICA MOLECULAR .....	65
4.5	ANÁLISE GENÔMICA .....	72
4.5.1	Análise de cultura Policlonal de LVH60 e LVH60a.....	72
4.5.1.1	Mapeamento de Amostras Policlonais em Genoma de Referência Concatenado.....	76
4.5.2	Sequenciamento Oxford Nanopore (Long Reads) .....	80
4.5.2.1	Montagem do genoma do clone de LVH60a .....	82

4.5.2.2	Validação e edição da primeira montagem genômica.....	82
<b>4.5.3</b>	<b>Análise de clones de LVH60 e LVH60a sequenciados por Illumina .....</b>	<b>91</b>
4.5.3.1	Estimativa de Ploidia.....	96
<b>5</b>	<b>DISCUSSÃO .....</b>	<b>99</b>
<b>6</b>	<b>CONCLUSÃO .....</b>	<b>107</b>
<b>7</b>	<b>REFERÊNCIAS BIBLIOGRÁFICAS .....</b>	<b>109</b>

## ***INTRODUÇÃO***

# 1 INTRODUÇÃO

## 1.1 LEISHMANIOSES

As leishmanioses compreendem um conjunto de doenças infecciosas negligenciadas distribuídas em regiões tropicais e subtropicais, ocorrentes em 98 países da África, América, Ásia e Europa (CHAPPUIS et al., 2007). Apesar da vasta distribuição, 90% de novos casos predominam em apenas 13 países (Afeganistão, Argélia, Bangladesh, Bolívia, Brasil, Colômbia, Etiópia, Índia, Iran, Peru, Sudão do Sul, Sudão e Síria) (STEVERDING, 2017). A Organização Mundial da Saúde (2020) incluiu as leishmanioses entre as seis endemias mais graves do mundo, estimando a ocorrência anual de 700.000 a 1 milhão de novos casos, sendo que desse número, apenas uma pequena parcela de pessoas desenvolve o quadro clínico da doença, e o restante permanecem assintomáticas. A doença é causada por mais de 20 espécies de protozoários *Leishmania* (ROSS, 1903) e transmitidas pela picada de flebotomíneos fêmeas infectados (CANTACESSI et al., 2015). Além disso, as leishmanioses têm relação direta com a precariedade de saneamento básico, moradia e escassez de recursos financeiros, afetando principalmente regiões e populações mais pobres do mundo, além de estarem associadas a transformações no ambiente, como desmatamento, construção de barragens, esquemas de irrigação e urbanização (WHO, 2020).

De acordo com as diferentes manifestações clínicas, as leishmanioses são identificadas como cutânea, muco-cutânea ou visceral (CHAPPUIS et al., 2007) e o tropismo do parasito causando os diferentes quadros clínicos é associado com as espécies infectantes de *Leishmania*. A leishmaniose cutânea (LC) tem como característica o aparecimento de úlceras e nódulos na superfície da pele, sendo classificada como leishmaniose muco-cutânea (LMC) quando há prevalência de ulceração progressiva das mucosas nasais e bucais até atingir a laringe e faringe (CHAPPUIS et al., 2007). As leishmanioses cutânea e muco-cutânea são causadas por *L. major* e *L. aethiopica* no Velho Mundo, e por *L. amazonensis* e *L. braziliensis* no Novo Mundo (CANTACESSI et al., 2015).

Em contrapartida, a leishmaniose visceral (LV) ou kala-azar é a forma mais grave da doença levando o indivíduo a óbito quando não tratada. É responsável por causar infecções sistêmicas no baço, fígado e medula óssea além de prejudicar as células do sistema fagocítico mononuclear (CHAPPUIS et al., 2007; STEVERDING, 2017). Os agentes causadores da leishmaniose visceral são os parasitos *L. donovani* no Velho Mundo, e *L. chagasi* e *L. infantum* no Novo Mundo

(CANTACESSI et al., 2015) tendo a predominância de novos casos concentrada no Brasil, Índia, Bangladesh, Sudão, Sudão do Sul e Etiópia.

Dentre os países da América Latina, o Brasil é o que apresenta maior incidência de LV com 90% de casos registrados e distribuídos em 23 estados. Em 2017, foi registrado 4.103 casos e o coeficiente de detecção foi de 1,98 casos por 100.000 habitantes. Do total de infectados, quase 45% ocorreu na Região Nordeste e 64,8% dos casos corresponderam a indivíduos do sexo masculino. Além disso, a doença acometeu 41% de crianças dentre 0 e 9 anos de idade (Ministério da Saúde, 2017). Até 1980, a LV no Brasil era considerada uma enfermidade restrita a regiões rurais, mas desde a crescente urbanização, esse cenário tem se modificado e intensificado a ocorrência da doença em grandes centros urbanos (WERNECK, 2014). Em 2017, foi registrado 2.744 internações de pacientes diagnosticados com LV e os casos fatais expandiram de 3,1% para 8,8% entre 2000 e o final da década (Ministério da Saúde, 2017).

Além dos casos clínicos ligados as leishmanioses, hospedeiros silvestres como raposas (*Lycalopex vetulus* e *Cerdocyon thous*) e marsupiais (*Didelphis albiventris*) são infectados naturalmente por *Leishmania* sp. em regiões do Ceará (DEANE; DEANE, 1962) e da Bahia (SHERLOCK et al., 1984), respectivamente. Esses animais sinantrópicos podem ser os responsáveis por estabelecerem a ligação entre o ambiente silvestre e urbano. Cães e humanos atuam como principais reservatórios do parasito em áreas urbanizadas, sendo a principal forma de infecção para os vetores (MAIA-ELKHOURY et al., 2008). Apesar da disseminação da doença e número de casos atingir todas as faixas etárias e diversas regiões do Brasil, essa tendência varia geograficamente e altas taxas de infecção são registradas em crianças de até 10 anos e na região Nordeste (SILVA et al., 2001).

Além da doença ocorrer majoritariamente em regiões pobres, e por conseguinte, dentre populações socialmente desprivilegiadas, não há programas de intervenção eficiente para melhorar o diagnóstico e tratamento da leishmaniose visceral. Do mais, a LV é difícil de ser diagnosticada clinicamente por apresentar sintomas iniciais muito inespecíficos, como febre e anemia; e um quadro clínico com severidade muito ampla, desde pacientes assintomáticos até aqueles que vão à óbito. Sendo assim, atualmente o diagnóstico da leishmaniose visceral é feito através de exame direto ou por cultivo de tecidos infectados (WHO, 2011). Exames relacionados ao sistema imune também podem identificar a doença, como o de imunofluorescência ou imunoenzimático, para constatar anticorpos reativos ao antígeno de *Leishmania*. O tratamento é bastante invasivo, feito através de um quimioterápico com antimoniais pentavalentes (antimoniato de N-metil glucamina

e estibogluconato de sódio), não havendo ainda vacinas licenciadas para as leishmanioses (exceto para leishmaniose visceral canina, que dispõe de dois produtos veterinários, Leish-Tec e Leishmune), sendo então a principal forma de controle da doença, o controle do vetor e dos reservatórios (GONTIJO; MELO, 2004).

## 1.2 O PARASITO

Os parasitos *Leishmania* possuem características biológicas capazes de interagir com seus hospedeiros. Esses organismos são digenéticos e alternam entre um hospedeiro invertebrado e vertebrado (Figura 1). O dimorfismo durante seu ciclo de vida está associado a duas formas principais: promastigota extracelular e amastigota intracelular, encontradas no trato digestivo do hospedeiro inseto e nas células do sistema imunológico fagocitário do hospedeiro vertebrado, nesta ordem (GHOSHAL; MANDAL, 2011). A transmissão do parasito para o hospedeiro humano ocorre pela picada do inseto fêmea infectada. O vetor é conhecido popularmente por mosquito-palha, pertence a Ordem Díptera, subfamília Phlebotominae e podem ser dos gêneros *Phlebotomus* no Velho Mundo e *Lutzomyia* no Novo Mundo (CANTACESSI et al., 2015). As formas amastigostas são ingeridas pelo flebotomíneo hematófago fêmea durante o repasto sanguíneo de um vertebrado infectado. No intestino do flebotomíneo, ocorre a diferenciação para promastigota e multiplicação desses protozoários. Ao atingir a forma infectante, promastigota metacíclica, os parasitos estão prontos para serem transferidos para hospedeiros vertebrados que ocorre através da picada do vetor infectado em um vertebrado saudável. Os protozoários do gênero *Leishmania* são caracterizados por invadir e multiplicar dentro das células do sistema fagocítico, provocando a diferenciação em amastigotas e causando as enfermidades conhecidas por leishmanioses (GHOSHAL; MANDAL, 2011)

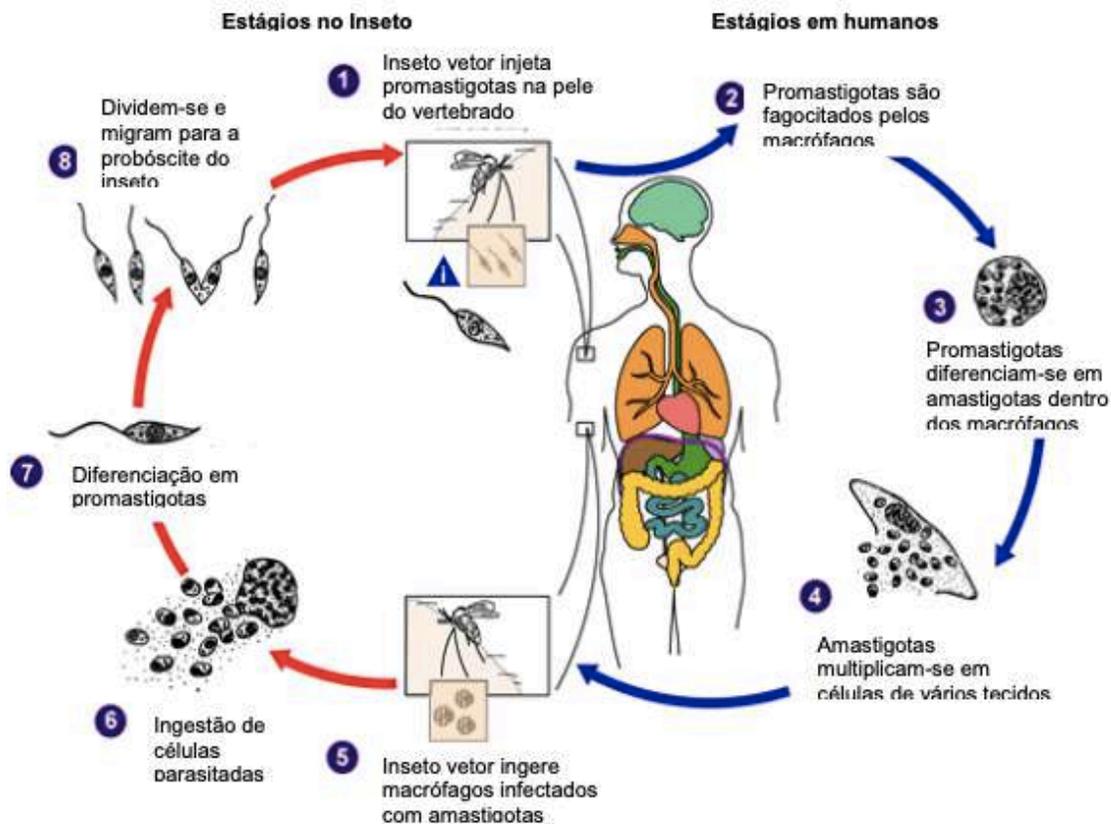


FIGURA 1 - O ciclo de vida do parasito *Leishmania*. (1,5,6,7,8) Fases do ciclo no inseto vetor; (2,3,4) Fases do ciclo nos macrófagos do hospedeiro vertebrado. Adaptado de Centro de Controle e Prevenção de Doenças (CDC) dos EUA. <<http://www.dpd.cdc.gov/DPDx/HTML/Leishmaniasis.htm>>

Os protozoários do gênero *Leishmania* compõem a ordem Kinetoplastida que possuem como característica chave a presença da organela cinetoplasto (MASLOV; PODLIPAEV; LUKES, 2001; VICKERMAN; PRESTON, 1976), a qual é representada pelo agrupamento único de DNA mitocondrial próximo a região flagelar. Os kinetoplastidas, juntamente com organismos da ordem Euglenida e Diplonemida, constituem o grupo monofilético Euglenozoa (CAVALIER-SMITH, 1993; CORLISS, 1994) que foram associados através de análises filogenéticas da subunidade menor (SSU) 18S de RNA ribossomal (PHILIPPE; ADOUTTE, 1998). A atual taxonomia de Kinetoplastida é baseada tanto em aspectos morfológicos quanto sob perspectiva do ciclo de vida dos organismos, os quais apresentam uma vasta variedade biológica desde seres de vida livre até parasitas. Os protozoários que dependem de um ambiente hospedeiro fazem parte da família Trypanosomatidae que agrupa organismos uniflagelados com cinetoplasto pequeno e obrigatoriamente parasitas (MASLOV; PODLIPAEV; LUKES, 2001). Existem diferentes

estratégias de vida dentro do parasitismo dos tripanosomatídeos (Figura 2): os organismos monoxênicos possuem ao longo de seu ciclo de vida um único hospedeiro invertebrado, como por exemplo os gêneros *Herpetomonas*, *Leptomonas* e *Crithidia* (LUKEŠ et al., 2014); já os parasitos que necessitam de mais de um tipo de hospedeiro para completar seu ciclo de vida são classificados como dixênicos, tendo como hospedeiros um invertebrado e outro vertebrado mamífero.

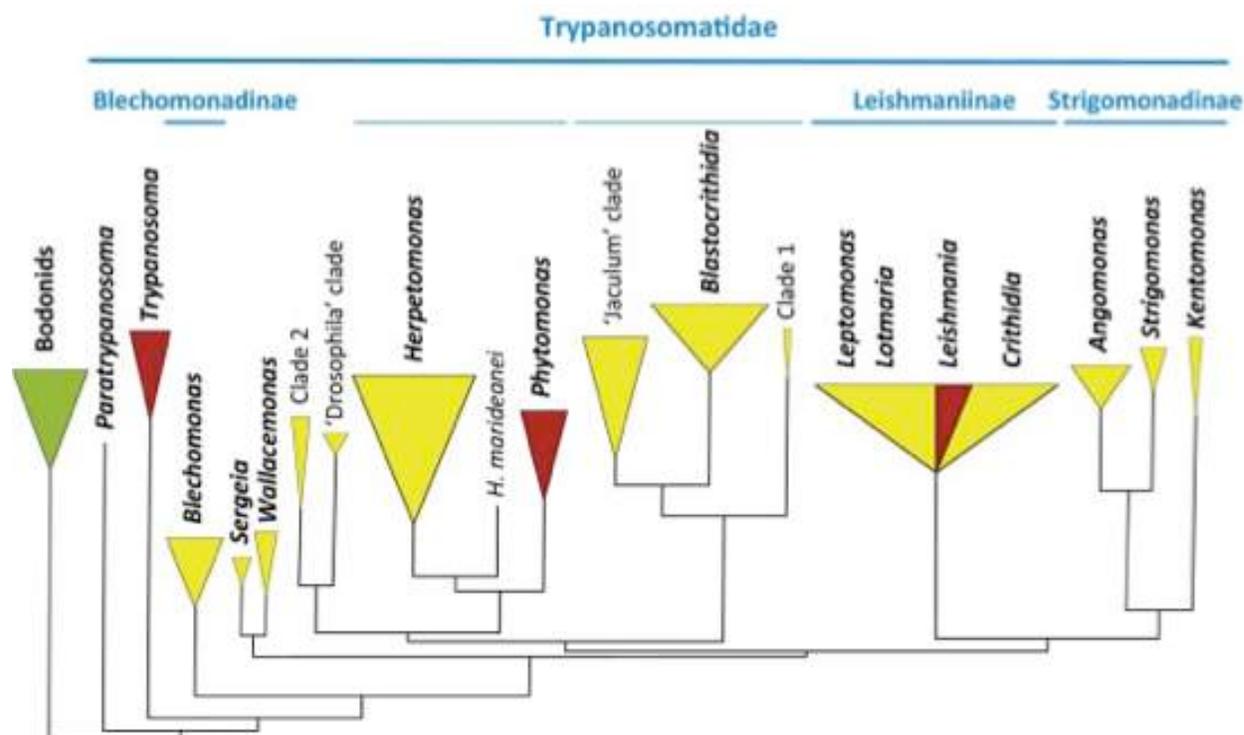


FIGURA 2 - Classificação taxonômica de protozoários parasitas da família Trypanosomatidae. Os gêneros monoxênicos (um hospedeiro) estão representados em amarelo, enquanto os gêneros dixênicos (dois hospedeiros) representados em vermelho e o grupo irmão de vida livre (Bodonids) está representado em verde. Fonte: VOTÝPKA et al., **Protist**, 2015.

Apesar das ultraestruturas de tripanosomatídeos serem altamente conservadas, há uma ampla variação morfológica que se difere tanto de uma espécie para outra, quanto dentre os estágios do ciclo de vida de uma mesma espécie (WHEELER; GLUENZ; GULL, 2013). As seis formas mais comuns estão associadas a características morfológicas como tamanho e formato do corpo, comprimento do flagelo e, posição em que o flagelo está inserido no corpo celular. Mesmo que essa variação morfológica não seja bem conhecida ainda, há estudos que revelam que a morfogênese está estreitamente relacionada a patogenicidade do organismo (JESUS et al., 1993; WHEELER; GLUENZ; GULL, 2013). Essas particularidades tornam os tripanosomatídeos

organismos modelos de estudo de parasitas, principalmente grupos importantes a saúde pública e causadores de enfermidades, como *Trypanosoma e Leishmania*.

Por retratarem um problema de saúde pública mundial, os parasitos causadores de leishmaniose visceral se tornaram foco de estudos biológicos e moleculares por pesquisadores de diversas áreas. Muitas são as abordagens moleculares empregadas no estudo das leishmanioses (SCHÖNIAN; KUHLS; MAURICIO, 2011) e utilizam diversas técnicas capazes de identificar agentes patológico. A PCR convencional é atualmente o método mais utilizado para a detecção molecular da leishmaniose visceral, sendo que o produto amplificado pode ser sequenciado e empregado em filogenia molecular (SCHÖNIAN; KUHLS; MAURICIO, 2011). Para isso, marcadores genéticos são utilizados na identificação molecular de patógenos com base nas sequências conservadas de Trypanosomatidae (CÁSSIA-PIRES et al., 2017; MARCILI et al., 2014; RANGEL et al., 2019), como a região da subunidade 18S do RNA ribossomal (SSU rRNA) (BORGHESAN et al., 2013; NOYES et al., 1999; VOTÝPKA et al., 2013; WILFERT et al., 2011; YURCHENKO et al., 2006), a região do RNA ribossomal do espaçador transcrito interno (ITS1) (SCHÖNIAN et al., 2003) e o gene gliceraldeído 3-fosfato desidrogenase (GAPDH) (BORGHESAN et al., 2013). A partir dessas informações moleculares, a filogenética possibilita relacionar organismos por sistema de árvores (SCHÖNIAN; KUHLS; MAURICIO, 2011) e mostrar que o parasitismo em *Leishmania* e outros tripanossomatídeos é um evento independente e representa uma parcela muito pequena dentre toda a diversidade existente no grupo (JACKSON; QUAIL; BERRIMAN, 2008).

O advento das tecnologias de sequenciamento genômico permitiu a investigação de características envolvidas com a patogenicidade da doença e a caracterização do patógeno a nível molecular (CANTACESSI et al., 2015), com o objetivo de identificar os mecanismos envolvidos na interação parasito-hospedeiro e relacionar alvos genéticos com diversas manifestações clínicas. A Genômica é uma abordagem que estuda funções e estruturas de diversas regiões do genoma de um organismo, sendo muito eficiente para promover o conhecimento dos parasitas e analisar resistência a drogas (DOWNING et al., 2011). O gênero *Leishmania* apresenta genomas com variação tanto em tamanho total quanto em números cromossômicos, variando de 29 a 33 Mb de comprimento e 34 a 36 cromossomos. Os genes codificadores de proteína são agrupados em *clusters* (IVENS et al., 2005) e possuem em média 8.000 genes com sistema organizacional policistrônico que é fundamentada pela transcrição do RNA polimerase que transcreve todos os genes estruturais em molécula única, e regula a expressão de genes por mecanismos pós-

transcricionais (MASLOV et al., 2019). O primeiro genoma de *Leishmania* sequenciado foi de *Leishmania major* e proporcionou um melhor entendimento e conhecimento da relação e função de seus 36 cromossomos com quase 33 megabases de comprimento (IVENS et al., 2005). Em seguida, os genomas de *Leishmania braziliensis* e *Leishmania infantum* foram descritas por genômica comparativa (PEACOCK et al., 2007), mostrando que, ao contrário do que se pensava, poucos genes espécie-específico do parasita estavam associadas as manifestações clínicas.

O rastreamento molecular de agentes etiológicos da leishmaniose visceral é executado frequentemente no Velho Mundo. As análises de genômica de população no subcontinente Asiático é bem fundamentado com o estudo da espécie *Leishmania donovani* (DOWNING et al., 2011; FRANSSEN et al., 2020; IMAMURA et al., 2016), em que a ampla variação no número de cópias cromossomais foram relacionadas a patogenicidade e adaptação de *Leishmania* a ambiente hostil como o hospedeiro humano (MANNAERT et al., 2012). Downing e colaboradores (2011) mostraram que a diversidade genômica intra-específica está relacionada ao contexto epidemiológico, demonstrando mecanismos de resistência do patógeno aos alvos terapêuticos utilizados em tratamentos. Além do mais, pesquisadores se preocuparam em desenvolver técnicas utilizadas para sequenciamento genômico de *L. donovani* diretamente da amostra do hospedeiro como forma de minimizar o viés causado pela cultura laboratorial desses patógenos (DOMAGALSKA et al., 2019).

Estudos parecidos, ainda que poucos, foram conduzidos para leishmaniose visceral causada por *Leishmania infantum* no Brasil com base na fonte da variação genética (TEIXEIRA et al., 2017) e resistência em tratamentos convencionais (CARNIELLI et al., 2018). O aumento de casos atípicos de leishmaniose visceral resultou da emergência de novos patógenos nunca associados anteriormente a doença, implicando no controle e tratamento adequado (THAKUR et al., 2018). Além disso, no Brasil a leishmaniose é considerada uma antropozoonose, infectando cães e humanos (WERNECK, 2014) e análises genômicas têm investigado esses parasitos em isolados clínicos de cães (BATISTA et al., 2018; TEIXEIRA et al., 2017; VALDIVIA et al., 2017). Portanto, para esse importante patógeno humano é essencial compreender o seu genoma e sua variação genética para aprimorar as estratégias de vigilância e intervenção em saúde pública de áreas endêmicas.

### 1.3 CONTEXTO E RELEVÂNCIA DO ESTUDO

A caracterização molecular seguida de sequenciamento genômico são análises importantes para triagem dos isolados clínicos e certificação da identidade de parasitos patogênicos como os causadores de leishmaniose visceral. A ocorrência de espécies não *Leishmania* são relatadas em casos de leishmaniose como eventos esporádicos desde a década de 80, geralmente co-infectadas por espécies monoxênicas (i.g. *Leptomonas* e *Herpetomonas*) em pacientes imunossuprimidos (revisado por KAUFER et al., 2017). Entretanto, é possível que esta observação esteja subestimada e que parasitos pouco estudados ou ainda desconhecidos possam estar envolvidos no desenvolvimento das leishmanioses.

Análises prévias realizadas com isolados clínicos obtidos do Hospital Universitário da Universidade Federal de Sergipe em Aracaju, apresentaram tipagem molecular por PCR e conteúdo genômico diferente das espécies de *Leishmania* (MARUYAMA et al., 2019). Os resultados indicaram que os isolados clínicos se aproximaram mais da espécie monoxênica *Crithidia fasciculata* que parasitam apenas insetos e não são consideradas patogênicas ao homem ou qualquer outro mamífero.

Por conta disso, estudos genômicos foram conduzidos a partir do caso atípico de um paciente do sexo masculino de 63 anos, HIV-negativo diagnosticado com LV que morreu após complicações da doença. O paciente apresentou perda de peso, febre, anemia, baixa contagem de leucócitos e plaquetas além do aumento do fígado e baço. A verificação da LV ocorreu por exame microscópico de luz de amastigotas em aspirados de medula óssea e promastigotas em meio de cultura após o isolamento do parasita. Também foi obtido um resultado positivo para o teste sorológico de *Leishmania* rK39. O tratamento com três doses de anfotericina B lipossômica não teve resultado e no terceiro tratamento refratário, foram encontradas lesões cutâneas papulares disseminadas. Uma biópsia de pele detectou amastigotas nos macrófagos, bem como na biópsia do fígado. Cepas de parasitas foram isoladas de aspirado de medula óssea e lesão cutânea. Como o tratamento não teve resposta positiva, foi realizada esplenectomia e a complicação cirúrgica levou o paciente à óbito. As amostras vieram do Hospital Universitário da Universidade Federal de Sergipe, em Aracaju - Sergipe, região endêmica da LV no Brasil (MARUYAMA et al., 2019). Resultados preliminares destes mesmos autores evidenciaram que parasitos não *Leishmania* tem sido isolado de pacientes diagnosticados com LV mostrando a necessidade de investigar a identidade das espécies causadoras da doença na região endêmica em estudo.

## ***OBJETIVOS***

## 2 OBJETIVOS

Este trabalho teve como objetivo geral realizar a triagem de isolados clínicos obtidos de pacientes diagnosticados com leishmaniose visceral e sistematizar a análise de dados de sequenciamento genômico obtidos de clones e culturas policlonais de isolados clínicos sequenciados por Illumina para a análise de Genômica Comparativa entre *Leishmania* e parasitos semelhantes a *Crithidia fasciculata* (*Crithidia-like*).

Os objetivos específicos foram:

- a. Identificar por PCR amostras por meio de marcadores genéticos que detectam tripanossomatídeos a partir das sequências de subunidade menor do RNA ribossomal (SSU rRNA), região do RNA ribossomal do espaçador transcrito interno (ITS1) e o gene gliceraldeído 3-fosfato desidrogenase (GAPDH);
- b. Identificar por PCR amostras por meio de marcadores genéticos espécie específico de *Leishmania infantum* e *Crithidia-like*;
- c. Analisar padrões morfológicos dos isolados clínicos por microscopia de luz;
- d. Construir um painel informativo para selecionar os isolados clínicos para sequenciamento genômico;
- e. Realizar a montagem do genoma *Crithidia-like* a partir do sequenciamento por Oxford Nanopore Technology (ONT);
- f. Executar análise comparativa entre sequenciamento genômico de clone e culturas policlonais com genomas de referências das espécies *Leishmania infantum* e *Crithidia fasciculata*.

## ***MATERIAL E MÉTODOS***

### **3 MATERIAL E MÉTODOS**

#### **3.1 TRIAGEM DE ISOLADOS CLÍNICOS**

##### **3.1.1 Amostras**

O Comitê de Ética do Hospital Universitário da Universidade Federal de Sergipe (HUUFS, Aracaju-SE) aprovou a coleta de material biológico para fins de pesquisa e um termo de consentimento foi obtido de cada participante maior de 18 anos ou de seus pais/guardiões legais (protocolo CAAE 04587312.2.0000.0058 – Plataforma Brasil). A obtenção das amostras se deu sob responsabilidade do Prof. Dr. Roque Pacheco de Almeida da Universidade Federal de Sergipe, colaborador desse projeto.

Em torno de 50 isolados clínicos de pacientes do Hospital Universitário de Aracaju – Sergipe, diagnosticados com leishmaniose visceral, foram analisados durante a triagem (Figura 3). O diagnóstico da LV envolveu resultados positivos no teste de Montenegro, na sorologia por ELISA e teste rápido (rK39), e apresentação de sinais clínicos da doença (hepatoesplenomegalia, exames bioquímicos e hematológicos alterados). Os critérios utilizados para a triagem, de acordo com o genótipo desses isolados clínicos, foram:

- a. Apresentação da manifestação clínica, em relação a severidade da doença e resposta ao tratamento.
- b. Resultado da tipagem molecular que discrimina as diferentes espécies de parasitos.

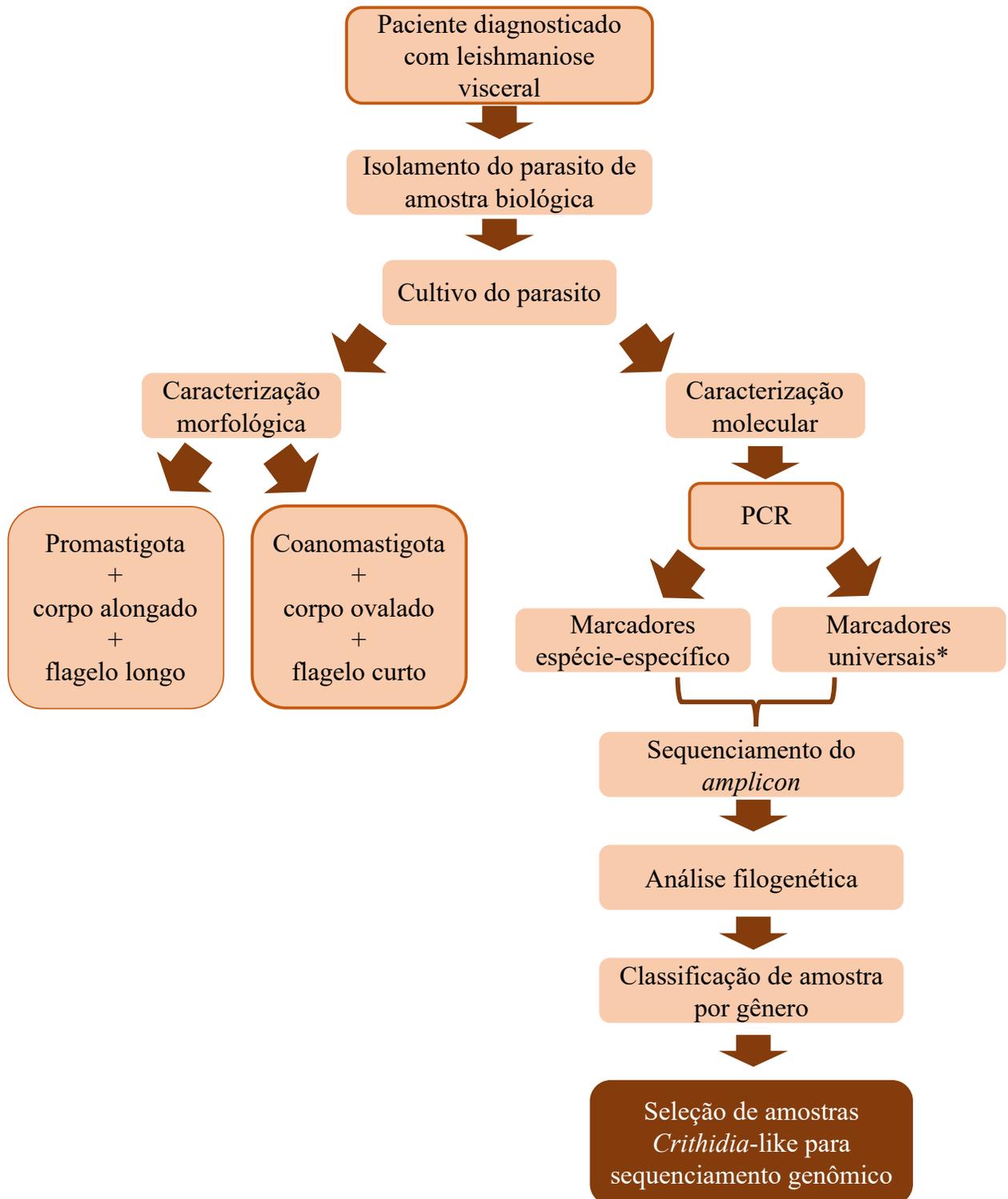


FIGURA 3 - Fluxograma do delineamento da triagem para seleção de isolados clínicos *Crithidia*-like para sequenciamento genômico. \*Iniciadores universais que amplificam sequências de tripanossomatídeos em geral.

### 3.1.2 Cultura de Promastigotas

Criotubos contendo os parasitos foram retirados do container criogênico de nitrogênio líquido e descongelados em gelo, passados para tubo estéril de 50 mL (Corning). Após o descongelamento, os protozoários foram transferidos para um tubo estéril de 15 mL contendo 10 mL de meio Schneider incompleto (Gibco), o qual foi centrifugado a 3000 rpm por 10 minutos a 4 °C. Em seguida, o sobrenadante foi descartado e o procedimento de lavagem com meio Schneider incompleto foi realizado por repetidas vezes. Posteriormente, o precipitado foi ressuscitado em 5 mL de meio Schneider contendo 10% de SBF (soro bovino fetal), 1% de antibiótico (Penicilina-Streptomomicina e L-glutamina) e 2% de urina masculina. O material foi transferido para uma garrafa de cultivo e o crescimento foi realizado a 25 °C por quatro a sete dias em estufa microbiológica.

Os frascos de cultivo celular foram mantidos por repiques *in vitro* de formas promastigotas em desenvolvimento, realizando a troca do meio Schneider e passagem da cultura de uma garrafa para outra nova após três a quatro dias do cultivo. A contagem das células foi feita na Câmara de Neubauer para ajustar a concentração celular em  $10^5$  promastigotas/mL por garrafa. O conteúdo da garrafa (cultura) foi transferido para um tubo de 50 mL contendo 30 mL de PBS 1x (Hyclone) estéril, realizando em seguida uma centrifugação de 800 rpm por três minutos a temperatura ambiente. Esse procedimento sedimentou as células mortas no fundo do tubo formando um *pellet* e as células vivas ficaram no sobrenadante. O sobrenadante foi coletado cuidadosamente e colocado em um tubo novo, acrescentando 10 mL de PBS 1x seguido de centrifugação a 3000 rpm por 10 minutos. Foi adicionado 1 mL de meio Schneider no *pellet*, ressuscitando as células e transferindo para garrafas novas com 5 mL de Schneider completo para crescimento a 25 °C. Após a cultura atingir o crescimento ideal, os parasitos foram coletados para extração de DNA e confecção de lâminas para microscopia. O procedimento completo foi realizado dentro de uma cabine de segurança biológica nível 2 em fluxo laminar, utilizando materiais estéreis e descartáveis.

Com a finalidade de manter a viabilidade e o estoque de células a baixas temperaturas, foi realizado o congelamento das culturas que estavam entre o 3° ou 4° dia de cultivo após o repique (promastigotas em fase log de crescimento). A cultura foi transferida para um tubo de 15 mL e centrifugado a 750 rpm por três minutos a 4 °C. Em seguida, transferiu-se o sobrenadante para um tubo novo de 15 mL, acrescentando sobre ele PBS 1x até completar 10 mL e centrifugar a 3000 rpm por 10 minutos a 4 °C repetindo novamente esta lavagem com PBS 1x. O “*pellet*” formado foi ressuscitado em 1 mL de meio Schneider Completo, para realizar a contagem. As células foram

congeladas em criotubos de 1 mL (Corning) contendo meio de congelamento contendo dimetilsulfoxido (DMSO) associado ao SBF inativado filtrado, sendo composto por 90% de SBF e 10% DMSO. Antes de preencher os criotubos com o composto de congelamento as células foram contabilizadas em câmara de Neubauer para ajustar a concentração em 1 mL contendo de  $5,0 \times 10^6$  a  $5,0 \times 10^7$  células/mL. As células foram mantidas overnight (aproximadamente 14 horas) no ultra freezer  $-80^\circ\text{C}$  e transferidas no dia seguinte para o container de nitrogênio líquido.

### **3.1.3 Obtenção de clones a partir das culturas promastigotas**

Para obter colônias clonais, técnicas bacteriológicas foram modificadas e enriquecidas com a finalidade de selecionar colônias a partir de uma única célula. Placas de Petri foram preparadas vertendo cerca de 15 mL de meio ágar bacteriológico autoclavado de concentração de 1,25% em meio Schneider completo. Após solidificação, as placas foram vedadas com filme de parafina plástica (Parafilm®) e mantidas em estufa a  $37^\circ\text{C}$  overnight para garantir a esterilidade do meio. Os isolados LVH60 e LVH60a foram plaqueados utilizando 40  $\mu\text{L}$  da cultura espalhada por toda a placa com alça de inoculação microbiológica estéril descartável. A placa foi vedada e mantida em estufa microbiológica a  $25^\circ\text{C}$ . A formação de colônias foi observada diariamente e o isolamento foi realizado com palitos de dente estéreis e homogeneizados diretamente em garrafa de cultivo com 5 mL de Schneider completo. Foram selecionados dois clones de cada isolado para análises molecular e genômico.

### **3.1.4 Crescimento Celular**

Os frascos de cultivo celular contendo culturas de promastigotas foram plaqueados em triplicata em placas de 12 poços com a contagem inicial de  $10^5$  células para cada 1 mL de meio de cultura. Cada poço foi preparado com 2 mL de meio, as placas foram mantidas em duas temperaturas distintas ( $25^\circ\text{C}$  para simular a temperatura no inseto vetor e  $35^\circ\text{C}$  para simular a temperatura no hospedeiro mamífero) e a contagem celular foi realizada por nove dias consecutivos. A amostra foi diluída 100 vezes em PBS 1X + formaldeído 2%, e 10  $\mu\text{L}$  dessa mistura foi colocada na câmara de Neubauer para realizar a contagem. A contagem foi realizada como segue a fórmula [Contagem =  $(C1+C2+C3+C4)/4 \times 10^2 \times 10^4$ ], em que C refere-se ao

quadrante da Câmara de Neubauer. Os valores foram anotados em tabela Excel, e os gráficos e teste estatístico *two-way* ANOVA foram gerados pelo programa GraphPrism5.

### 3.1.5 Caracterização Morfológica

Esfregaços de culturas promastigotas foram preparadas em lâminas de microscópio e coradas com um kit de coloração Panoptic Rapid (Laborclin). As lâminas a serem coradas foram colocadas em sequência por 15 segundos na solução A; 20 segundos na solução B e 30 segundos na solução C e, por último, lavadas em água destilada. Em seguida, as lâminas foram alocadas em suporte para secagem a temperatura ambiente.

O kit de coloração tem por objetivo evidenciar o flagelo, cinetoplasto, núcleo e o formato celular. As lâminas secas foram analisadas no laboratório de Citogenética de Peixes (LCP) do Departamento de Genética e Evolução (DGE) na UFSCar através de imagens de parasitos capturadas com câmera DP72 Olympus acoplada ao microscópio de luz Olympus BX50F em objetiva de imersão (100X). As formas promastigotas das cepas de referência MHOM/BR/74/PP75 (*L. infantum*) e MHOM/BR/75/M2903 (*L. braziliensis*) foram utilizadas para realizar a comparação.

### 3.1.6 Extração de DNA

Para realizar o isolamento do DNA, as formas promastigotas foram cultivadas de quatro a sete dias até atingirem cerca de  $2 \times 10^8$  parasitas/mL. As culturas foram lavadas com tampão fosfato-salino (PBS), centrifugadas e colhidas como *pellet* (extrato celular depositado no fundo do tubo) que foram usadas para extração de DNA genômico diretamente com o Kit Wizard SV Genomic® - DNA Purification System, da marca Promega. De acordo com o protocolo para extração de DNA de culturas de células, foi adicionado 150 µL de tampão de lise no *pellet* possuindo entre  $1 \times 10^4$  células e  $5 \times 10^7$  células, ressuspensando a amostra lisada. A amostra lisada foi colocada em um tubo contendo uma coluna identificada e centrifugada a 13000 g por três minutos, seguido de mais uma centrifugação a 14000 g por um minuto para descartar o que não foi passado por ela anteriormente. Acrescentou-se 650 µL de solução de lavagem da coluna (CWS) na coluna, seguido de centrifugação a 13000 g por um minuto, repetindo esta lavagem com CWS por quatro vezes. A

última lavagem foi realizada com centrifugação a 13000 g por dois minutos e 30 segundos, com as colunas já dispostas nos tubos definitivos devidamente identificados. Para o procedimento de eluição, foi colocada 2 µL de RNase ajustada em 250 µL de água sem nuclease e aplicada na coluna, deixando em repouso a temperatura ambiente por 10 minutos, seguido de centrifugação a 13000g por um minuto e armazenando a -20 °C, de acordo com as instruções do fabricante. As amostras de DNA foram quantificadas em espectrofotômetro NanoVue (GE) e armazenadas a -20 °C até serem utilizadas como moldes para PCR.

### 3.1.7 Reação em Cadeia da Polimerase (PCR)

#### 3.1.7.1 Marcadores genéticos

Os marcadores genéticos universais foram usados para identificar tripanossomatídeos de qualquer espécie. Os marcadores TRY e SSU são correntemente usados em estudos de taxonomia molecular de tripanosomatídeos (BORGHESAN et al., 2013; VOTÝPKA et al., 2013; WILFERT et al., 2011; YURCHENKO et al., 2006) pela técnica de *Nested* PCR (NOYES et al., 1998). Esses pares amplificam uma região da subunidade 18S do RNA ribossomal (SSU), correspondendo de iniciadores externos TRY927 e iniciadores internos SSU561. Iniciadores que amplificam uma outra região do rRNA ITS1 (internal transcribed spacer 1) (SCHÖNIAN et al., 2003) e que amplifica o gene gliceraldeído 3-fosfato desidrogenase (GAPDH) (BORGHESAN et al., 2013) também foram usados para identificar tripanossomatídeos de forma geral. Além dos iniciadores universais, também utilizamos dois iniciadores espécie-específicos, LinJ31 e Crid2.1, desenhados especialmente para detecção de genótipos de *L. infantum* e *Crithidia*-like, respectivamente (Tabela 1).

TABELA 1 - Sequências iniciadoras e reversas para amplificação de marcadores genéticos

Primer	Sequência (5' - 3')	Produto (nt)	Referência
TRY927	F GAAACAAGAAACACGGGAG R CTACTGGGCAGCTTGA	927	NOYES et al., 2000
SSU561	F TGGGATAACAAAGGAGCA R CTGAGACTGTAACCTCAAAGC	555	NOYES et al., 2000
gGAPDH.Tryp	F GGBCGCATGGTSTTCCAG R CCCCACTCGTTRTCTACC	~ 1000	BORGHESAN et al., 2012
ITS1	F CTGGATCATTTTCCGATG R TGATACTTATCGCACTT	300 - 350	SCHONIAN et al., 2003
LinJ31Seq*	F 20 bases R 20 bases	444	não publicado
Crid2.1Seq*	F1 20 bases R1 20 bases	500	não publicado

\* Sequências de iniciadores de amplificação desenhadas pelo grupo, dados não publicados.

### 3.1.7.2 Reações

A amplificação do DNA por reação em cadeia da polimerase foi realizada em um volume final de 25 uL, contendo cerca de 100 ng de DNA genômico de promastigotas como *template*, tampão 1X (100 mM Tris HCl (pH 8.8), 500 mM KCl, 15 mM MgCl<sub>2</sub>), 0,2 mM de cada desoxinucleosídeos trifosfatos (dNTP), 0,4 uM de iniciador, 0,4 uM de iniciador reverso, 2 mM de MgCl<sub>2</sub>, 0,5 unidades (0,25 uL) de Taq DNA Polimerase (Thermo Scientific) e água ultra pura em quantidade suficiente para completar o volume final de reação. As reações de PCR foram cicladas em termociclador (Analytik Jena, Biometra Tone), conforme especificado na tabela abaixo:

TABELA 2 - Determinação do tempo e temperatura para amplificação por PCR:

Etapas	TRY/SSU		GAPDH		ITS1*		LinJ31/Crid2.1	
	Temp. (°C)	Tempo (s)	Temp. (°C)	Tempo (s)	Temp. (°C)	Tempo (s)	Temp. (°C)	Tempo (s)
1.Desnaturação inicial	95	180	94	180	94	180	95	180
2. Desnaturação	94	30	94	60	94	60	94	60
3. Hibridização	55	60	55	120	53	30	55	30
4. Polimerização	72	90	72	120	72	40	72	40
5.Polimerização final	72	600	72	600	72	300	72	40
6. Refrigeração	12	∞	12	∞	12	∞	12	∞

A abertura da dupla fita de DNA (1, 2), seguida do anelamento dos iniciadores nas fitas geradas (3), passam pela extensão da fita através da DNA polimerase (4, 5). Para que ocorra a reação em cadeia da polimerase, as etapas 2, 3 e 4 devem ser realizadas por repetidas vezes (30 ciclos para todos os primers, exceto \*ITS1 com 35 ciclos). O processo de queda de temperatura (6) faz com que a reação de PCR seja cessada.

A *Nested* PCR foi aplicada primeiramente com os iniciadores externos TRY927, os quais geralmente não geram bandas de amplificação visível na eletroforese em gel. O produto dessa PCR diluído 1:10 foi usado como molde para realizar a segunda PCR com os iniciadores internos SSU561 de bandas visíveis. A PCR convencional foi aplicada com os marcadores genéticos de GAPDH, ITS1, LinJ31 e Crid2.1. A análise e identificação do produto das reações de amplificação, foi feita por eletroforese em gel de agarose (Sigma-Aldrich) 1% corados com SYBRsafe 0,001% DNA gel stain (Invitrogen). As amostras foram corridas em cubas para eletroforese (Sub-cell GT Bio-Rad), previamente preenchidas com tampão Tris-acetato-EDTA (TAE) 1x. O padrão de corrida Gene Ruler 1Kb Plus (Thermo Scientific) foi aplicado no primeiro poço do gel, seguido da aplicação de 5 µL do produto de PCR homogeneizado com 1µl de Ladder (6x) (glicerol 60%, EDTA 60 mM, Tris-HCl 10 mM pH 7,6 e azul de bromofenol 0,03%), com aplicação sucessiva de acordo com a quantidade de amostras. A eletroforese foi submetida por 45 minutos a 80V (Bio-Rad) obtendo um gel com fragmentos do produto de PCR, que posteriormente, foram fotodocumentadas em equipamento Chemi Doc TM XRS + Systems (Bio-Rad) a partir do *software* Image Lab TM 3.0, localizado no laboratório do Prof. Dr. Marcos Chiaratti do DGE/UFSCar.

Confirmando a presença da sequência de DNA de interesse, o produto de PCR foi purificado para subsequente sequenciamento do *amplicon*.

### 3.1.8 Sequenciamento dos Produtos de PCR

A purificação direta do produto de PCR foi feita com o kit Pure Link PCR Purification Kit (Invitrogen, Thermo Scientific), que é eficiente para remoção de *primers*, dNTPs, sais e enzimas (que foram adicionadas durante a reação de amplificação) sem usar precipitação de etanol. As etapas consistiram na adição do produto de PCR em um tubo coletor juntamente com um tampão a base de isopropanol que, ao entrarem em contato com a membrana de sílica, ocorreu a ligação do DNA da amostra nela. As próximas etapas consistiram em lavagens múltiplas com um tampão de lavagem (etanol), seguidas de centrifugação. Um novo tubo foi colocado sob a coluna e, o tampão de eluição (10 mM Tris-HCl, pH 8.5) fez com que o DNA ligado a membrana se desprendesse junto ao eluente e se depositasse no tubo novo quando centrifugado.

A purificação dos produtos de PCR com banda de amplificação dupla, foi realizada através da purificação em gel de agarose com o kit Wizard SV Gel and PCR Clean-UP System (Promega). Após a eletroforese é realizada uma pequena incisão no gel retirando apenas a banda de DNA de interesse e colocando-a em um microtubo de 1.5 mL. No microtubo foi adicionado tampão de ligação no gel, agitando-o e incubando-o a 50 °C até que seja completamente dissolvido. O gel dissolvido foi adicionado em um tubo coletor com um tampão a base de isopropanol que, ligou o DNA presente na membrana de sílica do tubo. Lavagens múltiplas com um tampão de lavagem (etanol) foram feitas, seguidas de centrifugação. Um novo tubo foi colocado sob a coluna e, adicionado tampão de eluição (10 mM Tris-HCl, pH 8.5) liberando o DNA e depositando-o em um novo tubo.

A quantificação do produto purificado foi realizada por espectrofotômetro NanoVue (GE) e as amostras com concentrações positivas (em ng/uL) foram sequenciadas pela equipe do IFSC-USP São Carlos conforme o método de terminação de cadeia desenvolvido por Sanger. Os iniciadores de amplificação utilizados para o sequenciamento foram SSU561, LinJ31Seq e Crid2.1Seq, e conforme as especificações do protocolo de amostras para sequenciamento, foi recomendado utilizar de 5 a 20 ng de *template*, considerando-se que o mesmo tenha entre 500 e 1000 pb. Além do *template*, o iniciador F foi adicionado sob concentração 3,2 pmol/uL,

acrescentando água ultra pura por último até completar o volume final de 7 uL. O sequenciador utilizado foi o 3130 Genetic Analyser da Applied Biosystems, possuindo quatro capilares de 50 cm, garantindo uma resolução de cerca de 800 pb.

### 3.1.9 Análise das sequências moleculares

A qualidade dos dados brutos de sequenciamento dos *amplicons* (disponibilizados em arquivos no formato .ab1) foram analisados com o *programa Chromas*, que projetou o arquivo como uma cromatografia, mostrando sob cada pico (representado por um nucleotídeo, sendo cada tipo de nucleotídeo de cor diferente) a qualidade do sequenciamento. No *Chromas* foi mostrado a necessidade de se ter ou não uma etapa de corte de sequências que poderiam enviesar à análise e, posteriormente, exportou o arquivo em formato FASTA. As sequências FASTA foram alinhadas com o *pacote blast* usando o *programa blastn* (de nucleotídeos) usando bancos de dados de tripanossomatídeos, para obter organismos referenciados com sequências próximas ou iguais.

Por fim, foi realizada análise filogenética molecular construindo árvore filogenética com o método de Máxima Verossimilhança utilizando sequências referência como *Leishmania infantum*, *Leishmania braziliensis* e *Leishmania donovani* disponibilizadas pelo NCBI. A referência monoxênica *Crithidia fasciculata* e o grupo externo *Trypanosoma cruzi* foram adicionados aos dados de alinhamento por *programa blastn*. A reconstrução filogenética da região da subunidade 18S de rRNA (SSU rRNA) foi obtida a partir do teste de melhor modelo de substituição de nucleotídeos dentre 24 modelos testados. Os alinhamentos múltiplos de sequências (MSA) foram obtidos pelas ferramentas *Muscle* (EDGAR, 2004) ou *Clustal Omega* (SIEVERS et al., 2011). Tanto os MSAs quanto a construção das árvores filogenéticas foram gerados no programa *Mega 6* (TAMURA et al., 2013).

Após a caracterização molecular e morfológica das amostras como proposto nos itens anteriores, os resultados foram organizados de modo que possam servir como um guia para a triagem dos isolados clínicos mais informativos para obtenção do sequenciamento completo de seus genomas (WGS).

## 3.2 ANÁLISE GENÔMICA

Neste trabalho, a análise genômica foi aplicada em dois conjuntos de dados que foram sequenciados em momentos diferentes. O primeiro conjunto de dados (item 3.2.1) trata-se de sequenciamento completo de genoma (WGS) na plataforma Illumina (*short reads*) dos isolados clínicos LVH60 e LVH60a (culturas policlonais) já disponíveis e recentemente publicados (MARUYAMA et al., 2019), portanto sequenciados anteriormente. Estes dados passaram por novas análises, diferentes daquelas executadas no estudo publicado. O segundo conjunto de dados foi genuinamente gerados pelo presente trabalho (item 3.2.2) e trata-se de sequenciamento completo de genoma (WGS) na plataforma Oxford Nanopore Technology (*long reads*) de um clone de LVH60a, cuja amostra também foi re-sequenciada na plataforma Illumina. Um segundo clone de LVH60a e dois clones de LVH60 também foram sequenciados na plataforma Illumina, compondo quatro novas amostras de sequenciamento genômico do mesmo caso clínico no segundo conjunto de dados.

### 3.2.1 Isolados Clínicos Previamente Sequenciados

A análise de sequenciamento genômico (Figura 4) foi realizada com amostras de parasito isoladas de aspirado de medula óssea (LVH60) e lesões cutâneas (LVH60a) de um paciente de caso atípico de leishmaniose visceral descrito anteriormente (MARUYAMA et al., 2019). O sequenciamento do genoma completo dessas amostras foi executado na plataforma Illumina utilizando 50 ng de DNA genômico de culturas promastigotas policlonal. O kit Nextera® XT DNA Library Preparation foi usado para fragmentar o DNA e construção da biblioteca, seguido pela adição de adaptadores nas extremidades 3' e 5' do DNA. Os adaptadores foram responsáveis por ligar o DNA no *flow-cell* (lâmina) por anelamento de bases complementares. O processo de amplificação ocorreu construindo pontes de PCR auxiliado pelo kit HiSeq® Rapid SBS Kit (500 ciclos, 2x250 bp) em equipamento HiSeq® 2500 com sistema de modo Rapid Run. O serviço de sequenciamento foi realizado pelo ESALQ Genomics Center, Universidade de São Paulo, Piracicaba, São Paulo, Brasil. As bibliotecas de DNA foram sequenciadas em duplicata gerando aproximadamente 42 milhões de *paired-reads* por amostra, perfazendo uma cobertura média de sequenciamento em torno de 600x considerando um genoma de *Leishmania* de 32 Mb.

Aproximadamente 20% dos *paired-end reads* de cada amostra sequenciada foram descartadas durante o processo de filtragem de sequências de baixa qualidade ou contaminação por iniciadores.

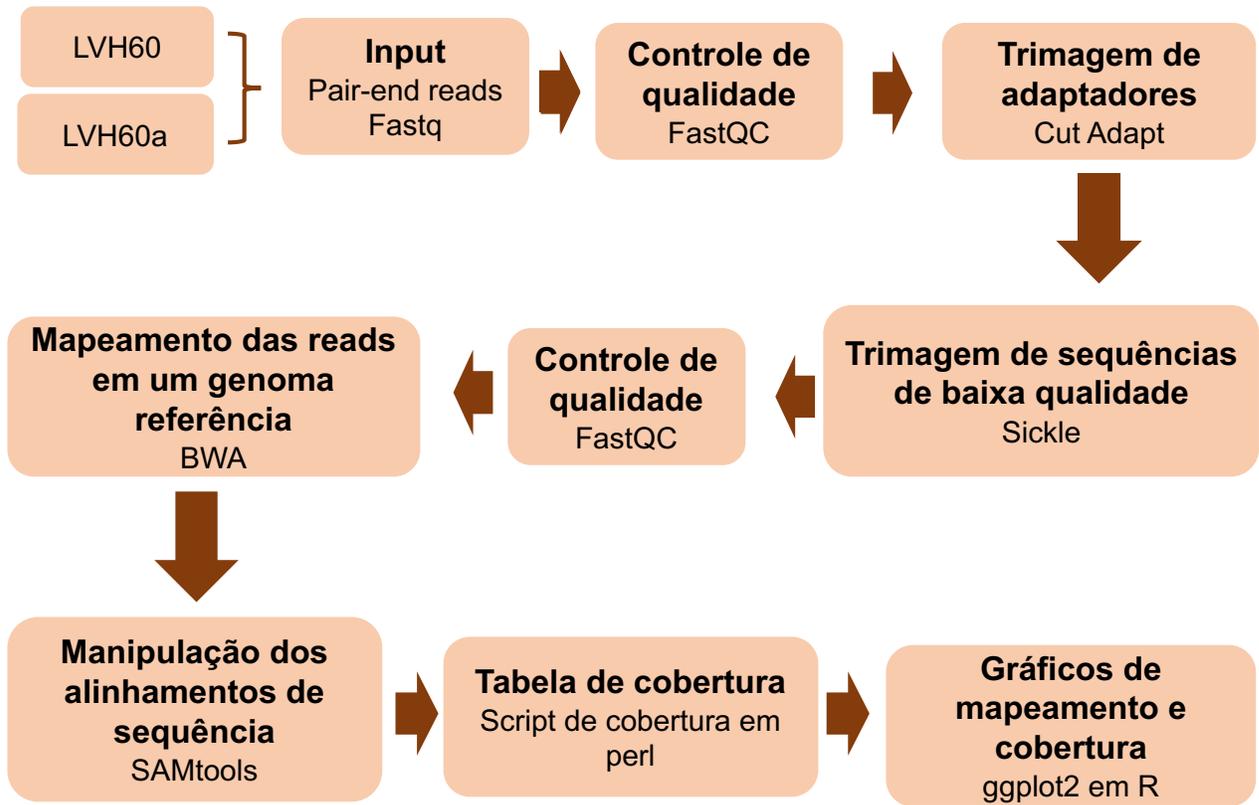


FIGURA 4 - Ilustração do processamento e mapeamento dos arquivos de sequenciamento Illumina de culturas policlonais dos isolados clínicos LVH60 e LVH60a.

### 3.2.1.1 Mapeamento de *Reads*

Foi realizado o mapeamento das *reads* de LVH60 e LVH60a usando a referência concatenada de *Leishmania infantum* (TriTrypDB) e *Crithidia fasciculata* (TriTrypDB). O controle de qualidade das *reads* brutas foi visualizado com o programa FastQC (ANDREWS, 2010) que tem por objetivo visualizar a qualidade da amostra, demonstrando ou não a necessidade de processos de trimagem e melhoramento de seqüência. O software também permitiu uma comparação da qualidade inicial e pós processamento das amostras, mostrando se elas eram boas o suficiente para análises subsequentes. Primeiramente, as *reads* em duplicata foram concatenadas antes de iniciar os passos de trimagem:

Comandos:

```
$ cat sample.L1.R1.fastq.gz > sample_R1.fastq.gz
$ cat sample_L2.R1.fastq.gz >> sample_R1.fastq.gz
```

O comando `cat` foi utilizado para concatenar dois arquivos em um único arquivo, onde `>` indica escrever um novo arquivo e `>>` indica acrescentar em um arquivo já existente. O mesmo foi realizado para as duplicatas de reverse. A remoção dos adaptadores universais foi realizado com o software `Cutadapt` (MARTIN, 2011) usando as opções abaixo:

- **a** é o adaptador *forward*

-**A** é o adaptador *reverse*

-**o** indica o arquivo de saída (*output*) das *reads forward*

-**p** indica o arquivo de saída (*output*) das *reads reverse*

`sample.1.fastq.gz` e `sample.2.fastq.gz` são os *inputs* que são aparados

`\` é para um código de formatação e significa que o comando pode ser escrito na próxima linha

Comando:

```
$ cutadapt \
  -a AGATCGGAAGAGCACACGTCTGAACTCCAGTCAC \
  -A AGATCGGAAGAGCGTCGTGTAGGGAAAGAGTGTAGATCTCGGT \
  GGTCGCCGTATCATT -o sample_trimmed.1.fastq.gz \
  -p sample_trimmed.2.fastq.gz sample_R1.fastq.gz sample_R2.fastq.gz
```

O programa `Sickle` (JOSHI; FASS, 2011) foi usado para realizar o melhorar a qualidade dos arquivos contendo *reads* que tiveram os adaptadores removidos. Este *software* remove sequências de baixa qualidade, removendo também o par da sequência que não passou nos parâmetros de qualidade. Dessa forma, a ordem das *reads* são mantidas entre os dois arquivos.

-**pe** *paired-end reads*

-**f** *forward reads file* (R1)

-**r** *reverse reads file* (R2)

-**t** método de índice de qualidade para Illumina, phred score

-**o** R1 *output*

**-p** R2 *output*

**-s** *singleton output* (sequências que não foram pareadas)

**q** limiar de pontuação de qualidade, as leituras precisam passar pela pontuação de phred Q20 (taxa de 1% de taxa de erro de sequenciamento)

**-l** é o tamanho mínimo de sequência que pode ser cortado antes que ele rejeite e coloque no arquivo R0

Comando:

```
$ sickle pe -f sample_trimmed.1.fastq.gz -r sample_trimmed.2.fastq.gz \
-t sanger -o sample_sickle_R1.fastq.gz \
-p sample_sickle_R2.fastq.gz \
-s sample_sickle__R0.fastq.gz -q 20 -l 20
```

Para o mapeamento, os genomas de referência *L. infantum* e *C. fasciculata* do banco de dados TriTryp (ASLETT et al., 2010) foram concatenados para obter uma referência combinada.

Comandos:

```
$ cat TriTrypDB-45_LinfantumJPCM5_Genome.fasta > Cfasciculata_Linfantum.fasta
$ cat TriTrypDB-45_CfasciculataCfCl_Genome >> Cfasciculata_Linfantum.fasta
```

O programa *Burrows-Wheeler Aligner* (BWA) (LI; DURBIN, 2009) é designado para mapear sequências pouco divergentes e foi usado para mapear as *reads* de alta qualidade de LVH60 e LVH60a em genoma referência concatenado. O algoritmo usado foi BWA-MEM que oferece suporte para *reads* longas e recomendado para *queries* de alta qualidade já que é mais preciso e mais rápido. A referência concatenada foi indexada e em seguida foi realizado o mapeamento.

Comandos:

```
$ bwa index Cfasciculata_Linfantum.fasta
$ bwa mem Cfasciculata_Linfantum.fasta sample_sickle_R1.fastq.gz \
sample_sickle_R2.fastq.gz > sample.sam
```

O *software* Samtools foi usado para manipular os alinhamentos no formato SAM, além de ordenar, indexar, remover duplicatas e gerar um alinhamento em um formato por posição. O comando view do Samtools foi aplicado para converter um arquivo de extensão BAM em um arquivo de extensão SAM. O comando sort foi usado para ordenar o conteúdo disposto no arquivo, enquanto o comando rmdup foi usado para remover duplicatas. Por fim, apenas as reads que alinharam com índice de mapeamento igual ou superior a 60 (isso significa 1 chance de erro a cada  $10^6$  bases) foram considerados e utilizados para gerar os dados de mapeamento e cobertura. Todos os arquivos foram indexados.

Comandos:

```
$ samtools view -bT Cfasciculata_Linfantum.fasta sample.sam > sample.bam
$ samtools sort sample.bam > sample_sort.bam
$ samtools index sample_sort.bam
$ samtools rmdup -S sample_sort.bam sample_sort_nodup.bam
$ samtools index sample_sort_nodup.bam
$ samtools view -q 60 -b -o sample_q60_sort_nodup.bam sample_sort_nodup.bam
$ samtools index sample_q60_sort_nodup.bam
```

O *output* de mapeamento e cobertura foi obtido utilizando um *script perl* desenvolvido e disponibilizado por Sarah Forrester (University of York, UK) e é demonstrada abaixo:

**Ref size:** Tamanho da referência disposto no arquivo fasta. O tamanho da referência é o comprimento em pares de base (pb).

**Mapped:** Número de bases mapeadas em uma dada referência/cromossomo. Se o tamanho da referência tem 100 pb e 100 pb são mapeadas, então todas as posições da referência é coberta pela sequência consulta de interesse (*query*) que está sendo alinhada à referência.

**%Mapped:** Número de bases mapeadas da sequência *query* dividido pelo tamanho da referência multiplicado por 100.

**Mean:** Profundidade media de cobertura vertical ao longo da referência/cromossomo

**Median:** Profundidade mediana de cobertura vertical ao longo da referência/cromossomo

**SD:** Desvio padrão de cobertura por referência

**Q1:** 25% da profundidade de cobertura

**Q3:** 75% da profundidade de cobertura

**2.5%:** 2,5% da profundidade de cobertura

**87.5%:** 87,5% da profundidade de cobertura

**Min:** Menor cobertura encontrada em qualquer posição ao longo da referência

**Max:** Maior cobertura encontrada em qualquer posição ao longo da referência

Comando:

```
$ perl coverage.pl -i sample_q60_sort_nodup.bam > sample_q60.txt
```

A linguagem de programação estatística R foi usado para gerar gráficos das reads mapeadas na referência concatenada e gráficos da profundidade media de cobertura, usando o pacote ggplot2 (<https://ggplot2.tidyverse.org>):

Comandos:

```
$ ggplot(data= sample, mapping=aes(Chromosome,Mean, col=Organism, fill=Organism)) +
geom_col(position = position_dodge2(width = 1, preserve = "single")) +
scale_color_manual(values=c("red","blue")) + ylab("Mean depth coverage") +
scale_x_continuous(breaks=seq(1,36,1)) + theme(axis.text=element_text(size=12),
axis.title=element_text(size=20))
```

```
$ ggplot(data= sample, mapping=aes(Chromosome,X._Mapped, col=Organism, fill=Organism))
+ geom_col(position = position_dodge2(width = 1, preserve = "single")) +
scale_color_manual(values=c("red","blue")) + ylab("Percentage chromosome mapped to") +
scale_x_continuous(breaks=seq(1,36,1)) + theme(axis.text=element_text(size=12),
axis.title=element_text(size=20))
```

### 3.2.1.2 Alinhamento de *Contigs* em Genomas Referência

As *reads* de alta qualidade foram usadas para montagem do genoma (gerar *contigs*) usando o ABySS (SIMPSON et al., 2009) com um tamanho k-mer de 96. As sequências de DNA codificantes foram traduzidas obtendo a sua região aberta de leitura a partir da ferramenta EMBOSS (<https://www.ebi.ac.uk/Tools/emboss/>), ajustando a metionina inicial por BLASTX contra banco de sequências de proteínas Protozoa-RefSeq do NCBI (MARUYAMA et al., 2019).

Já o alinhamento entre genomas montados foi feito com o programa MUMmer 3.0 (KURTZ et al., 2004) que alinha de forma rápida genomas completos e incompletos com um genoma

referência. O comando NUCmer tem como finalidade alinhar dois genomas através da localizando de regiões conservadas nas sequências de DNA. A partir disso é gerado alinhamentos entre sequências consulta de interesse (*query*) e a referência, produzindo uma matriz de identidade entre os organismos comparados.

Comandos:

```
$ nucmer --prefix=ref_qry ref.fasta qry.fasta  
$ show-coords -rcl ref_qry.delta > ref_qry.coords
```

Onde

**--prefix=** é o nome dado ao arquivo de saída (resultado);

**ref.fasta** é o arquivo FASTA da referência,

**qry.fasta** é o arquivo FASTA do query (no caso, os *contigs* montados para os isolados clínicos LVH60, LVH60a e HUUFS14);

**-r** ordenar a linha do output pela referência;

**-c** incluir coluna de porcentagem de cobertura no output;

**-l** inclui uma coluna de comprimento no output;

**ref\_qry.coords** refere-se ao arquivo de saída .coords que apresenta todas as coordenadas de posicionamento do alinhamento entre os genomas.

Foram utilizadas sequência genômica de referência de dois organismos para essa análise: *Leishmania infantum* (<https://www.ncbi.nlm.nih.gov/genome/249>) e *Crithidia fasciculata* (<https://www.ncbi.nlm.nih.gov/genome/14517>).

### 3.2.2 Sequenciamento de Clones de Isolados Clínicos

Novas amostras foram geradas e analisadas. Clones foram originados das culturas policlonais de LVH60 e LVH60a, sequenciados por Illumina e alinhados em genoma referência de *Crithidia*-like (LVH60a clone 1) obtido por sequenciamento na plataforma de *long reads* Oxford Nanopore Technology (ONT). Ambos sequenciamentos foram executados a partir de parasitos de um homem de 63 anos HIV negativo com LV fatal isolados de aspirado de medula óssea (LVH60)

e lesão cutânea (LVH60a) na Universidade de York em York-UK durante o período de Bolsa de Estágio de Pesquisa no Exterior (BEPE FAPESP Processo 2019/12142-0) (Tabela 3).

TABELA 3 - Parasitos clonais de LVH60 e LVH60a para sequenciamento genômico.

#Amostra	Espécie	Código da amostra	Isolado Clínico	Método de Sequenciamento
1	Crithidia-like	LVH60 C3	Medula óssea	Illumina
2	Crithidia-like	LVH60 P1F6	Medula óssea	Illumina
<b>3</b>	<b>Crithidia-like</b>	<b>LVH60a C1</b>	<b>Lesão cutânea</b>	<b>ONT e Illumina</b>
4	Crithidia-like	LVH60a C2	Lesão cutânea	Illumina

Amostra 3 foi sequenciada por ambas plataformas Oxford Nanopore Technology e Illumina.

Culturas clones de promastigotas foram isoladas de diferentes amostras e realizada igualmente ao item 3.1.3. *Pellet* contendo cerca de  $10^8$  parasitas foram mantidos em 1 mL de reagente DNA/RNA Shield (Zymo Research) diluído em água ultrapura (1:1) de acordo com a recomendação do fabricante, para serem inativados por agentes infecciosos. Este procedimento garantiu a estabilidade do DNA.

### 3.2.2.1 Sequenciamento Genômico por Oxford Nanopore Technology (ONT)

O sequenciamento de *long reads* foi realizado por Oxford Nanopore Technology com o clone LVH60a C1. O DNA das amostras em DNA/RNA Shield foi extraído usando um protocolo de extração a base de fenol-clorofórmio adaptado. O volume de 200  $\mu$ L de células armazenadas em DNA/RNA Shield foi utilizado e quando necessário o volume foi ajustado para 200  $\mu$ L total usando o tampão de eluição (EB - Tris-HCl 10 mM, pH 8,5). A solução fenol/clorofórmio/álcool isoamílico é um detergente que tem como finalidade remover proteínas e outros contaminantes presentes junto ao DNA. Foi adicionado um volume igual (200  $\mu$ L) de solução composta de fenol 100  $\mu$ L (Sigma-Aldrich), clorofórmio 96  $\mu$ L (Invitrogen) e álcool isoamílico 4  $\mu$ L (Fisher Scientific) na proporção 25: 24: 1, pH 7,8 - 8,2 (tubo 1) e a solução foi misturada vertendo o tubo cuidadosamente por período de um minuto. A solução do tubo 1 foi centrifugada em velocidade

18.000 g por cinco minutos. Em seguida, aproximadamente 180  $\mu\text{L}$  da solução aquosa que se formou na porção superior do tubo contendo o DNA foi realocada em um novo tubo, evitando extrair a fase fenol/clorofórmio/álcool isoamílico (parte inferior do tubo). Realizamos a lavagem com 200  $\mu\text{L}$  de EB no tubo 1 novamente, para retirar resquícios de DNA do tubo e misturamos a solução por inversão por um minuto. A solução do tubo 1 foi centrifugada a 18.000 g por cinco minutos e 120  $\mu\text{L}$  da fase aquosa formada na porção superior contendo o DNA foi realocada para o tubo 2, totalizando 300  $\mu\text{L}$  de solução com DNA no tubo 2. A solução clorofórmio e álcool isoamílico possibilita a extração de lipídeos, proteínas e outros resíduos, depositando-se na parte inferior do tubo enquanto o DNA permanece na porção superior. Um volume igual ao do tubo 2 (300  $\mu\text{L}$ ) foi adicionado de solução de clorofórmio 288  $\mu\text{L}$  e álcool isoamílico 12  $\mu\text{L}$  na proporção 24:1 ao tubo 2 e misturado por inversão por 1 minuto. A solução do tubo 2 foi centrifugada à 18.000 g por cinco minutos e 230  $\mu\text{L}$  da fase aquosa formada na porção superior foi realocada para um novo tubo (tubo 3).

O acetato de amônio ( $\text{NH}_4\text{CH}_3\text{CO}_2$ ) promove a precipitação do DNA e foi adicionado 17  $\mu\text{L}$  a uma concentração final de 0,75M (tubo 3). Em seguida, foi adicionado 1  $\mu\text{L}$  de glicogênio que ajuda a recuperar o DNA (20  $\mu\text{g}/\mu\text{L}$ ) vertendo o tubo para misturar a solução. Depois disso, foi acrescentado um volume de 2,5X de etanol 100% (aproximadamente 650  $\mu\text{L}$ ) usado para limpar o DNA desidratado, vertendo o tubo para misturar bem e incubando a 20 °C por 10 minutos. Após a incubação, o tubo 3 foi centrifugado a 18.000 g por 20 minutos. O sobrenadante foi removido cuidadosamente sem perturbar o *pellet* e em seguida o mesmo foi lavado com 300  $\mu\text{L}$  de etanol 80% e misturado três vezes por inversão do tubo. Em seguida, o tubo foi centrifugado por 15 minutos a 18.000 g decantando cuidadosamente o sobrenadante sem perturbar o *pellet*. Uma segunda lavagem com etanol 80% foi realizada de acordo com os passos descrito anteriormente. Foi dada uma rotação rápida para acumular resíduos de etanol no fundo do tubo. A tampa do tubo foi deixada aberta por cinco minutos para secar ao ar, e finalmente, ressuspender o *pellet* em 40  $\mu\text{L}$  de EB. A concentração de DNA foi estimada por Espectofotometria *NanoDrop* 100 (Thermo Scientific) e *Qubit* por quantificação fluorimétrica (Thermo Scientific) usando o kit Qubit ds DNA HS assay kit. O comprimento e a concentração do DNA foram examinados em aparelho 4200 *TapeStation* (Agilent Technologies) para garantir DNA de alta qualidade e peso molecular.

A biblioteca para sequenciamento *Nanopore* foi construída pela Instalação de Tecnologia do Departamento de Biologia/Universidade de York usando uma quantidade de aproximada de 1  $\mu\text{g}$  de DNA para cada amostra com os kits Nanopore ligation sequencing kit SQK-LSK109 com

SQK-NBS104 *barcode expansion* kit que é uma espécie de código de barra (sequência de DNA) acoplada em cada amostra permitindo o sequenciamento *multiplex* (amostras de diferentes procedências sequenciadas juntas em uma mesma corrida do sequenciador). Foi realizado o reparo do DNA e também o preparo das extremidades do DNA para conectar os adaptadores de sequenciamento. Em seguida, o *barcode* fornecido no kit foi ligado nas extremidades de DNA reparadas anteriormente. Por ligação química, adaptadores de sequenciamento foram conectados nas extremidades do DNA. Depois disso, a *flow-cell* foi iniciada e a biblioteca de DNA foi carregada (Figura 5).

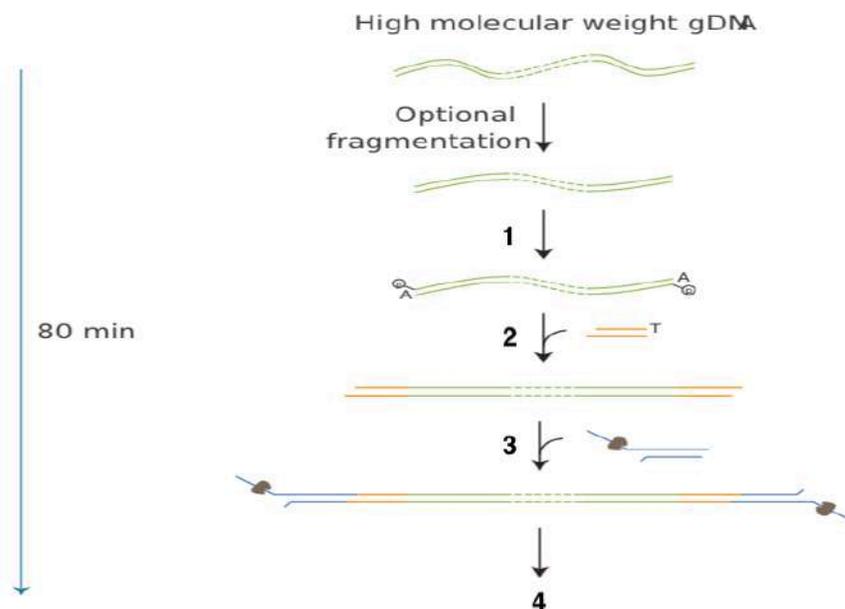


FIGURA 5 - Fluxograma do preparo da biblioteca para sequenciamento por *Oxford Nanopore Technology*. Os passos para sequenciamento a partir do DNA de alto peso molecular extraído consiste em 1- Reparo e preparo da extremidade de DNA. 2- Ligação do *barcode* fornecido no kit. 3- Ligação do adaptador. 4- Carregamento na *flow-cell* com a biblioteca pronta (adaptada de protocolo Nanopore: Native Barcoding genomic DNA (with EXP-NBD104, EXP-NBD114, and SQK-LSK109)

### 3.2.2.1.1 Recursos Computacionais

Os dados de bioinformática gerados foram processados e analisados com *Viking*, um servidor computacional *Linux* de alto desempenho, que busca o desenvolvimento e a execução de tarefas de computação. O cluster foi acessado pelo protocolo SSH permitindo executar muitos programas ao mesmo tempo e acessar mais recursos como memória e espaço em disco do que os disponíveis em um computador convencional. O *Viking* foi projetado para usar como sistema

operacional *Centos Linux 7.3* (<http://www.centos.org/>) e Grid Engine Slurm Workload Manger 18.08.4 (<https://slurm.schedmd.com/>). O node de login tem acesso a módulos com vários aplicativos e pacotes de software disponíveis.

### 3.2.2.1.2 Análise de Long Reads

O *output* contendo os *barcodes* de LVH60a-C1 foram gerados a partir do sequenciamento por ONT. Os *barcodes* foram removidos com o *software Guppy* (<https://github.com/svenil/guppy-pe/>) e o controle de qualidade da sequência bruta foi realizado e visualizado com *PycoQC* (<https://pypi.org/project/pycoQC/>). O processamento dos dados brutos de sequenciamento até a obtenção de um genoma *draft* é demonstrado na Figura 6.

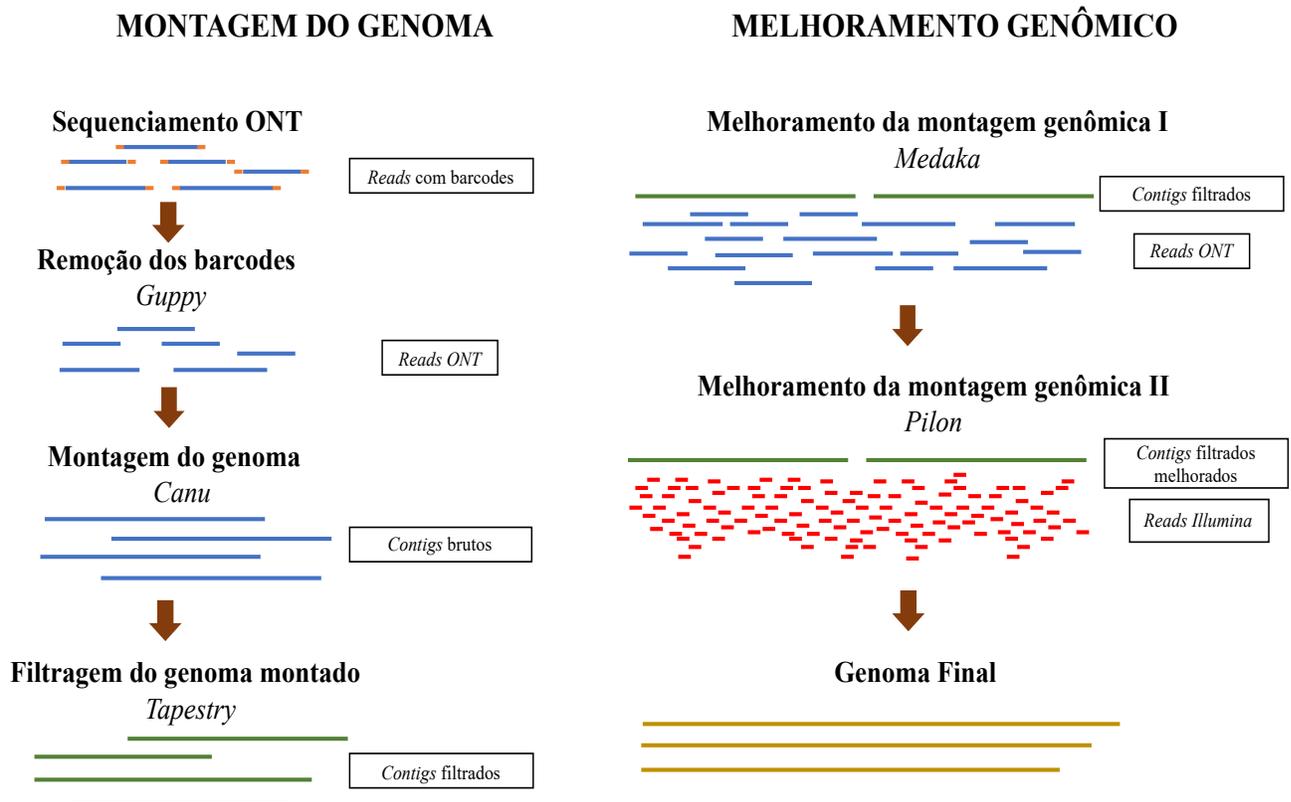


FIGURA 6 - Fluxograma do processamento de dados de sequenciamento Oxford Nanopore Technology de LVH60a-c1 para obtenção de genoma final montado e melhorado.

O programa *Canu* v.1.8 (KOREN et al., 2017) foi usado para realizar a montagem das *reads* LVH60a-c1 em *contigs* através de processos de correção, corte e montagem do genoma. As funções empregadas foram:

- p** prefixo o qual os arquivos gerados serão nomeados
- d** criação de um diretório para deposição dos arquivos gerados
- genomeSize=41.0m** tamanho aproximado do genoma
- nanopore-raw** tipo de sequenciamento

Comando:

```
$ canu -p prefix-assembly -d directory-assembly genomeSize=41.0m -nanopore-raw  
raw_reads.fastq.gz
```

Seqkit v.11.0 (WEI SHEN et al., 2016) foi usado para manipular os arquivos FASTA e gerar um resumo estatístico para o genoma montado usando o comando `stats`. Para transformar o arquivo FASTA para um formato tabular e obter informações como comprimento dos *contigs* e conteúdo CG, utilizamos o comando `fx2tab`:

Comando:

```
$ seqkit stats -a assembly_contigs.fasta  
$ seqkit fx2tab -n -l -g assembly_contigs.fasta
```

As opções foram descritas abaixo:

- n** nome dos *contigs*
- l** comprimento da sequência
- g** conteúdo CG

A ferramenta *Tapestry* (<https://github.com/johnomics/tapestry>) foi usada na validação e filtragem da montagem genômica. O comando *weave* é o responsável para produzir o *output*.

Comando:

```
$ weave -a assembly.fasta -r reads.fastq.gz -t TGATA -o assembly -c 4
$ clean -a assembly.fasta -c assembly.filtered.csv
```

As opções usadas foram:

- a *contigs* montados em formato FASTA
- r *reads* geradas com o sequenciamento ONT
- t sequência telomérica
- o diretório onde o output será depositado
- c número de *cores* usado

O relatório resumido referente a montagem do genoma em formato HTML foi usado para ordenar, filtrar e descrever os *contigs* baseados nas informações do relatório. O comando *clean* gerou um novo conjunto de *contigs* que foram exportados do relatório e salvos em arquivo CSV. Assim, foi criado um arquivo FASTA com os *contigs* da montagem filtrados.

O *Nucmer* (KURTZ et al., 2004) é um algoritmo usado para comparar os nucleotídeos do genoma montado filtrado de LVH60a-c1 com o genoma referência *C. fasciculata* (TritrypDB) e também para mapear esse genoma LVH60a-C1 filtrado no seu próprio genoma. O *output* Nucmer foi um arquivo delta, o qual foi usado como input para rodar um script *Python* (<https://dnanexus.github.io/dot/DotPrep.py>) para gerar arquivos compatíveis para serem visualizados no Dot (<https://github.com/dnanexus/dot>).

Comando:

```
$ nucmer -p LVH60a-c1-clean_Cfasc reference_genome.fasta clean_assembly.fasta
$ nucmer -p LVH60a-clean-c1_LVH60a-c1-clean clean_assembly.fasta clean_assembly.fasta
```

O *script* DotPrep.py foi aplicado para obter um algoritmo único de filtragem para marcar os alinhamentos como únicos ou de caráter repetitivo. O *output* DotPrep.py inclui os arquivos

coords e coords.idx que foram carregados na plataforma Dot em um navegador de internet (<https://dnanexus.github.io/dot/>).

Comando:

```
$ python DotPrep.py --delta LVH60a-c1-clean_Cfasc.delta --overview 10000 --out LVH60a-c1-clean.nucmer.delta
```

O melhoramento do genoma filtrado foi realizado em duas etapas. Primeiramente, o genoma foi processado para melhoramento usando as *reads* do sequenciamento ONT com o programa *Medaka*. Posteriormente, esse genoma previamente processado por *Medaka* foi usado para uma segunda etapa de melhoramento usando reads geradas de seuqenciamento Illumina com o programa *Pilon*. A ferramenta *Medaka* (OXFORD NANOPORE TECHNOLOGY, 2018) foi usada para obter uma sequência consenso dos dados de sequenciamento ONT (<https://nanoporetech.github.io/medaka/>). O *input* (montagem do genoma filtrado e *reads* de sequenciamento ONT) em formato FASTA gerou um output consensus.fasta.

Comando:

```
$ medaka_consensus -i nanopore_reads.fasta -d clean_assembly.fasta -o medaka-output_directory -m r941_min_high -t 40
```

O programa *Pilon* (WALKER et al., 2014) tem como característica identificar regiões de alinhamento ruins e melhorar a montagem do genoma, corrigindo erros das *reads Illumina*. A ferramenta foi projetada para encontrar diferenças de base única, inserções, deleções e preenchimento de lacunas. O arquivo consensus.fasta gerado por *Medaka* e as *pair-end reads* Illumina foram usadas como *input*. O *output* gerado foi resultado da correção do *Pilon* visando erros de base e como produto obtivemos um genoma final montado e melhorado de LVH6a-c1.

Comando:

```
$ sbatch slurm_pilon.sh final-assembly.medaka.pilon ilumina-reads.R1.trimmed.fastq.gz ilumina-reads.R2.trimmed.fastq.gz medaka-output_directory/consensus.fasta 40 3
```

Para comparar as características do genoma bruto e processado, usamos a ferramenta Seqkit. Comparamos os tamanhos da montagem bruta de Canu (LVH60a-c1-contigs.fasta), montagem filtrada por Tapestry (LVH60a-c1\_clean\_assembly), genoma melhorado por Medaka (consensus.fasta) e genoma melhorado por Pilon (pilon1.fasta, pilon2.fasta e pilon3.fasta).

Comando:

```
$ seqkit stat -a LVH60a-c1-contigs.fasta LVH60a-c1_clean_assembly consensus.fasta pilon1.fasta  
pilon2.fasta pilon3.fasta
```

Sequências codificadoras de proteínas (CDS) foram extraídas de *contigs* consenso por os 6 frames de leitura (ORF) usando a ferramenta EMBOSS (RICE et al., 2000).

### 3.2.2.2 Sequenciamento Genômico por Illumina

O sequenciamento *pair-end reads* de quatro amostras provenientes de diferentes isolados clínicos foi realizado com a plataforma Illumina.

A extração do DNA foi realizada usando o kit DNeasy Blood and Tissue (Quiagen). Foram colocadas 500 µL de amostra em DNA/RNA Shield (Zymo Research) em um novo tubo e adicionadas 50 µL de proteinase K. As amostras foram incubadas a 56 °C por 10 minutos. Foi adicionado 500 µL de etanol (90-100%) e agitado por vortex. O produto foi adicionado sobre a coluna de rotação e colocada em um tubo de coleta para centrifugar a 8000 rpm por 1 minuto. A coluna de rotação foi colocada em um novo tubo de capacidade 2 mL e foram adicionados 500 µL de tampão AW1, seguidos de centrifugação a 8000 rpm por 1 minuto. O tubo de coleta foi descartado. A coluna de rotação foi colocada em um novo tubo de 2 mL e 500 µL de tampão AW2 foram adicionados, seguidos de centrifugação a 18.000 g por 3 minutos. O tubo de passagem de coleta foi descartado e a coluna de rotação foi transferida para um novo tubo de 1,5 mL. Foi realizada a eluição do DNA adicionando-se 50 µL de tampão AE no centro da membrana de sílica presente na coluna. A solução resultante foi incubada por 10 minutos à temperatura ambiente e centrifugada posteriormente a 8000 rpm por 1 minuto. A concentração de DNA foi verificada pelos equipamentos *NanoDrop* e *Qubit* antes das amostras serem enviadas ao serviço de sequenciamento.

A biblioteca de DNA para sequenciamento *Illumina* foi realizado a partir de 100 ng de DNA, com quatro ciclos de PCR e preparados com o kit NEBNext Ultra II FS DNA Library Prep para *Illumina*. As amostras foram sequenciadas em um instrumento *Illumina* HiSeq 3000 visando 30-40x de cobertura por amostra.

Para realizar o mapeamento e determinar a profundidade média de cobertura, foi executado as mesmas análises descritas em 3.2.1.1, usando amostras clonais e policlonais como query e o genoma *Crithidia*-like como referência.

#### 3.2.2.2.1 Estimativa de Ploidia

A estimativa da ploidia foi empregada pela média de cobertura obtida para cada *contig* da referência *Crithidia*-like. Para isso, usamos o alinhamento das *reads* *Illumina* com a referência *Crithidia*-like (genoma *draft*) pelo programa BWA. Arquivos em formato BED (Browser Extendable Data) foram construídos como no exemplo abaixo:

Exemplo:

```
tig.1 1 comprimento-do-tig.1
tig.2 1 comprimento-do-tig.2
tig.3 1 comprimento-do-tig.3
.
.
.
tig.38 1 comprimento-do-tig.38
```

Para calcular a contagem total de bases das *reads* por *contig*, utilizamos o comando do *SAMtools* descrito abaixo:

Comando:

```
$ samtools bedcov yourbed.bed yourbam.bam > output.txt
```

Os gráficos de ploidia provenientes do arquivo output.txt foram gerados pelo pacote ggplot em linguagem de programação estatística em R.

Comando:

```
#Ler os dados
d <- read.table("arquivo.bed.txt",header = FALSE, sep="\t",stringsAsFactors=FALSE, quote="")

#Renomear o nome das colunas
names(d) <- c("chr", "chr_start", "chr_end", "read_count")

#barplot de readcount por contig dividido pelo tamanho do contig
barplot(d$read_count/d$chr_end)

#Criar uma nova variável de reads por base dividido pelo tamanho do contig
d$coverage <- d$read_count/d$chr_end

#Fazer um subset excluindo contigs que não são diplóides
d2 <- subset(d, chr != "chr8")
d3 <- subset(d2, chr != "chr9")
d4 <- subset(d3, chr != "chr16")
d5 <- subset(d4, chr != "chr22")
d6 <- subset(d5, chr != "chr27")

#Calcular a média para obter a profundidade de cobertura de um diplóide
diploid = median(d6$coverage)
haploid = diploid/2
triploid = diploid*1.5
tetraploid = diploid*2

pdf("LvH60a-c2_somy.pdf", width=18, height=8)
barplot(d6$coverage,
main = "LVH60a-clone2 Somy Estimation",
xlab = "Contig",
ylab = "Coverage depth", names.arg = 1:38)
axis(side = 4, at = c(0,haploid, diploid, triploid, tetraploid), labels = c(0,1,2,3,4))
abline(h = c(0,haploid,diploid,triploid,tetraploid), lty = 2)
dev.off()
```

## ***RESULTADOS***

## 4 RESULTADOS

### 4.1 CRESCIMENTO CELULAR

O ciclo de vida dixênico do gênero *Leishmania* permite que estes organismos tenham dois hospedeiros para completar seu ciclo, ocorrendo na forma promastigota no seu hospedeiro invertebrado (vetor) e amastigota intracelular em seu hospedeiro vertebrado (humanos ou mamíferos reservatórios). Para verificar a viabilidade e crescimento celular desses parasitos, realizamos o experimento em duplicata, através de contagem celular diária por período de nove dias em temperaturas que simulavam o ambiente dos hospedeiros invertebrados (25 °C) e vertebrados (35 °C) (Figura 7). Os seguintes resultados foram obtidos a partir da contagem diária do número de promastigotas de diversas linhagens de *Leishmania* e *Crithidia*-like em diferentes condições de temperatura. O padrão de crescimento *in vitro* de formas promastigotas foi analisado diante da curva de crescimento de isolados clínicos policlonais e clonais de LVH60, LVH60a, LVH60 C3, LVH60 P1E6, HUUFS14 e PP75 (OMS/MHOM/BR/74/PP75). Todas as amostras iniciaram o experimento com  $2 \times 10^5$  células em 2 mL de meio de cultura enriquecido e os dados foram analisados com teste estatístico Two-Way ANOVA.

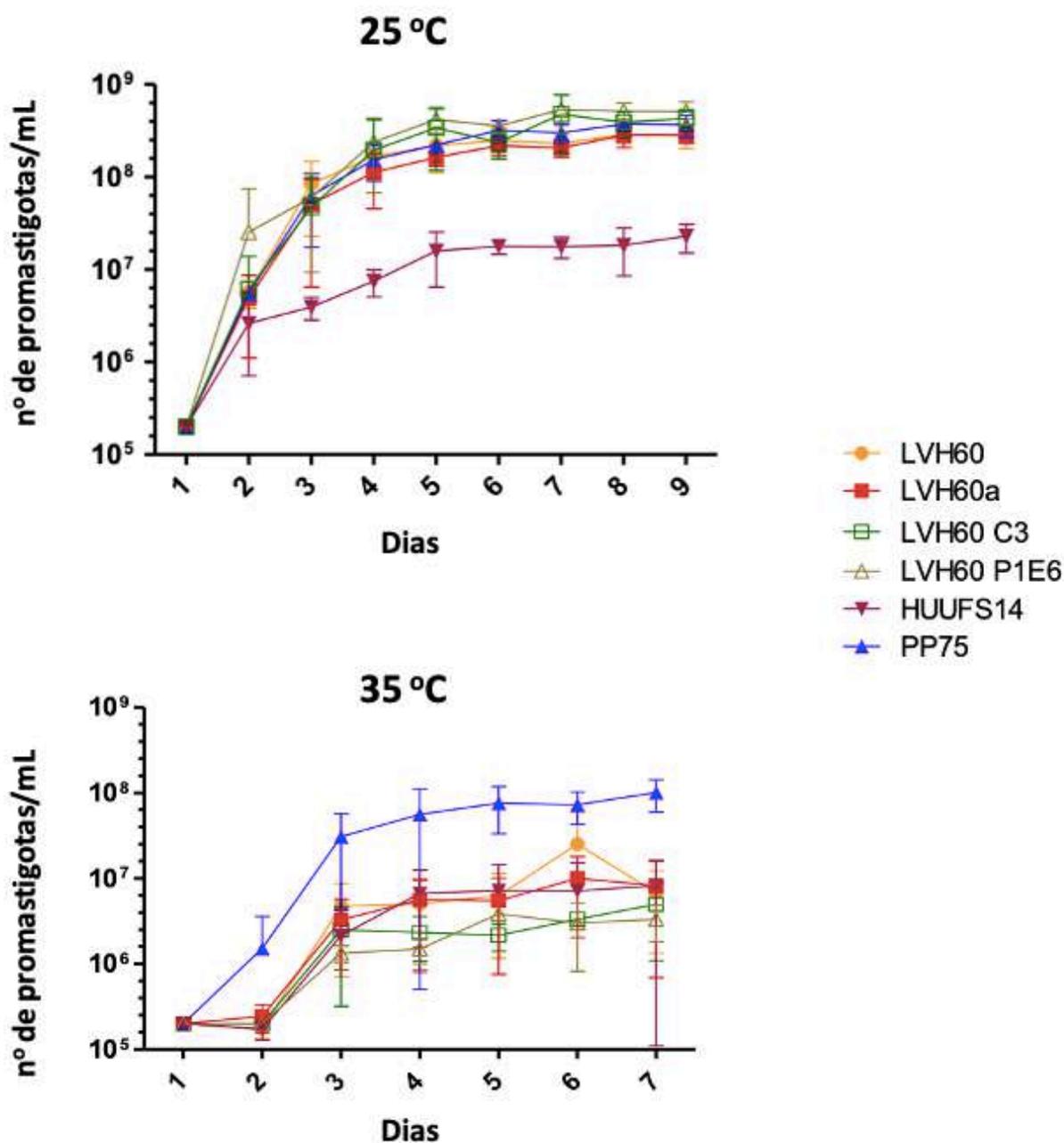


FIGURA 7 - Curva de crescimento *in vitro* de promastigotas em temperaturas simulando o hospedeiro invertebrado (25 °C) e hospedeiro vertebrado (35 °C). LVH60: isolado clínico obtido de amostra de medula óssea; LVH60a: isolado clínico obtido de amostra cutânea (sendo LVH60 e LVH60a colhidos do mesmo paciente); HUUFS14: isolado clínico de *L. infantum* referência da região de Sergipe; PP75: isolado clínico de *L. infantum* (OMS/MHOM/BR/74/PP75) As culturas de promastigotas possuíam concentração inicial de 10<sup>5</sup> células/mL e foram contadas em câmara de Neubauer e microscópio de luz diariamente por nove dias. Cada amostra foi cultivada em triplicata e dois experimentos independentes foram realizados. Os dados foram analisados com teste estatístico two-way ANOVA usando o programa GraphPrism 5.0. Não houve diferença estatística entre as amostras, apenas entre os dias de crescimento dentro de cada amostra exceto por HUUFS14 a 25 °C (p<0,0001).

O gráfico de crescimento celular em temperatura a 25 °C simulou temperatura similar ao hospedeiro invertebrado, indicou crescimento das amostras LVH60, LVH60a, LVH60 C3, LVH60 P1E6 e MHOM/BR/74/PP75 igualmente desde o dia da inoculação das promastigotas em concentração inicial de  $2 \times 10^5$  células até atingir aproximadamente  $5 \times 10^8$  células no 5º dia. A partir do 6º dia todas as amostras entraram em fase estacionária mantendo fixa a quantidade de células sem crescimento visível até o último dia de contagem. Em contrapartida, a linhagem de *Leishmania infantum* HUUFS14 apresentou variação quantitativa quando comparada as outras cepas, chegando ao platô com  $1 \times 10^7$  células no 5º dia de contagem. As promastigotas em temperatura de 25 °C apresentaram corpo celular alongado, com flagelo longo inserido na região anterior da base flagelar.

Já o gráfico de proliferação celular a 35 °C, que representa temperatura do hospedeiro vertebrado de sangue quente, mostrou crescimento de LVH60, LVH60a, LVH60 C3, LVH60 P1E6 e HUUFS14 igualmente desde o dia inicial contendo  $2 \times 10^5$  células até atingir  $\sim 5 \times 10^7$  no 3º dia, onde não houve mais multiplicação das células. No entanto as células de MHOM/BR/74/PP75 (*L. infantum*) atingiram a fase estacionária apenas no 4º dia de contagem com  $6 \times 10^7$  células. As células cultivadas a 35 °C apresentaram corpo celular arredondado e pequeno, com um flagelo curto (e pouco nítido em imagens de microscopia de luz) de pouca movimentação (Figura 8) e tornaram-se mais semelhantes a forma amastigota imóvel de *C. fasciculata* encontrada no intestino de dípteros.

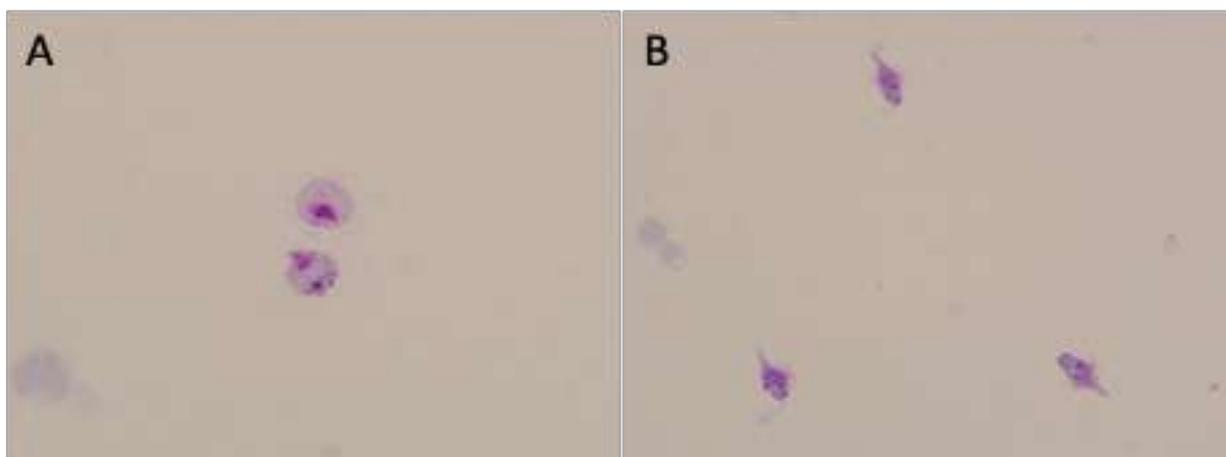
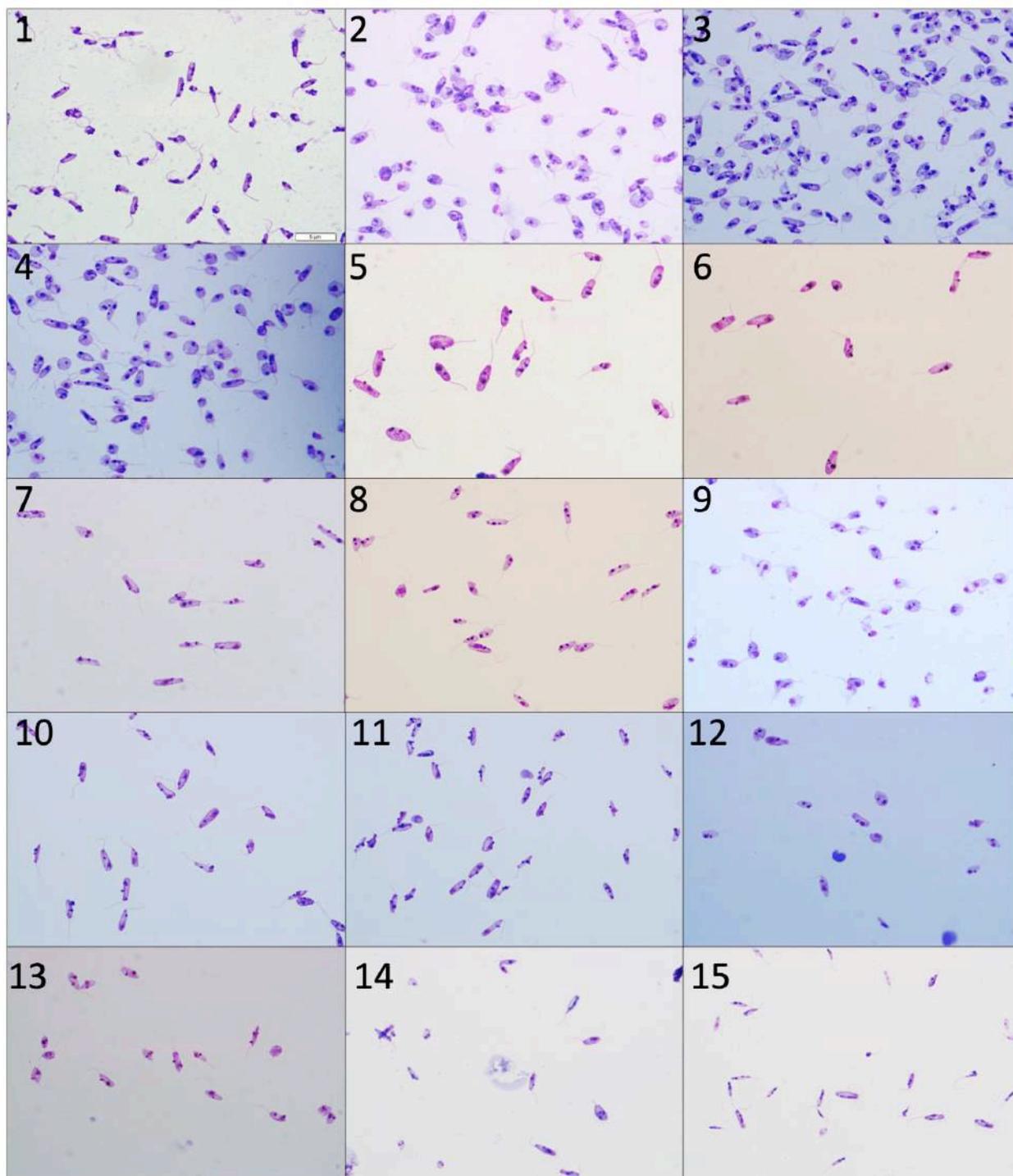
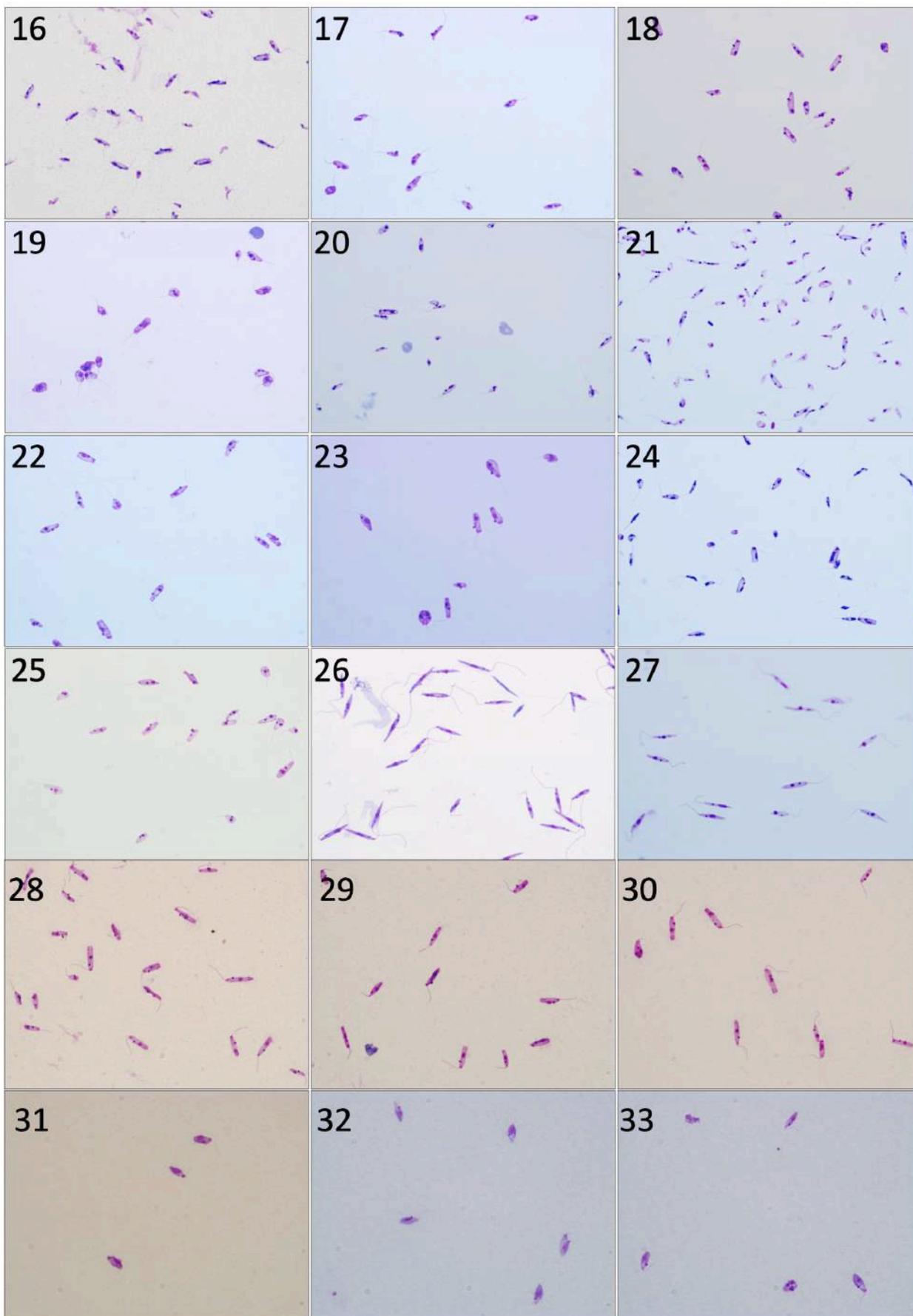


FIGURA 8 - Formas promastigotas extracelulares em temperatura de 35 °C. **A.** Cultura de LVH60: morfotipo apresentou-se arredondado, sem flagelo aparente e imóvel. **B.** Cultura de PP75 (*L. infantum*): morfotipo apresentou-se levemente arredondado, com flagelos curtos e imóvel.

#### 4.2 CARACTERIZAÇÃO MORFOLÓGICA DE PROMASTIGOTAS

O total de 27 lâminas de cultura de promastigotas retiradas de pacientes diagnosticados com leishmaniose visceral humana (LVH) foram coradas. As lâminas foram fotografadas evidenciando estruturas celulares como cinetoplasto, núcleo e flagelo.





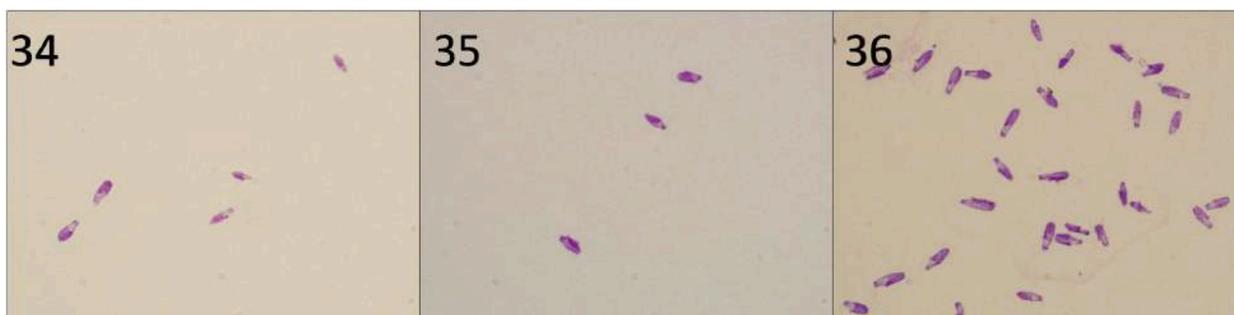


FIGURA 9 - Formas extracelulares de diferentes culturas coradas em lâmina. (1) LVH09; (2) LVH27; (3) LVH29; (4) LVH32; (5) LVH33; (6) LVH42; (7) LVH43; (8) LVH44; (9) LVH49; (10) LVH55; (11) LVH56; (12) LVH57; (13) LVH59; (14) LVH60 clone 1; (15) LVH60 clone 2; (16) LVH60 clone 3; (17) LVH61; (18) LVH62; (19) LVH66; (20) LVH91; (21) LVH92; (22) LVH94; (23) LVH94a; (24) LVH96; (25) LVH102; (26) PP75 clone 1; (27) *Zelonia*. (28) LVH23 clone1; (29) LVH23 clone 2; (30) LVH23 clone 3; (31) LVH09 clone 1; (32) LVH09 clone 2; (33) LVH09 clone 3; (34) LTCP393 clone 1; (35) LTCP393 clone 2; (36) LTCP393 clone 3. Imagens obtidas por microscópio de luz Olympus BX50F, com câmera acoplada Olympus DP72.

A forma promastigota em *Leishmania* sp. é encontrada no hospedeiro vetor e quando no estágio metacíclico representa a forma infectante do parasita. Apesar dos tripanossomatídeos apresentarem uma organização celular básica e morfologia muito semelhante entre eles, a relação existente entre tamanho do flagelo, inserção do flagelo, posição do cinetoplasto e posição do núcleo determinam e diferenciam uma espécie da outra (WHEELER; GLUENZ; GULL, 2013). O exemplar de *Leishmania infantum* PP75 e HUUFS14 na (Figura 9.26 e 9.27) apresentam células alongadas, com flagelos longos inseridos na região anterior do corpo celular com cinetoplasto visível anteriormente ao núcleo, típico da forma promastigota. Já nas fotografias de LVH27 (Figura 9.2), LVH29 (Figura 9.3) e LVH32 (Figura 9.4) e LVH49 (Figura 9.9) por exemplo, mostram células mais arredondadas, com flagelos mais curtos e cinetoplasto ao lado do núcleo parecendo uma coanomastigota (KAUFER et al., 2017). Ainda que seja necessário realizar análise morfológica mais detalhada, já foi possível notar diferenças dentre os diferentes isolados clínicos. Quando comparado com o exemplar *Leishmania infantum chagasi* (MHOM/BR/74/PP75, cepa referência da WHO), observamos que a maioria dos isolados clínicos apresentam o corpo celular mais arredondado e flagelo mais curto.

## 4.3 PCR DE PARASITO ISOLADO EM CULTURA

As PCRs utilizaram como DNA molde amostras de cultura (LVH) tendo como amostras referência HUUFS14 (*L. infantum*), PP75 (*L. infantum*) e LbrM2903 (*L. braziliensis*). As regiões da subunidade menor do RNA ribossômico SSU (NOYES et al., 1999), o gene GAPDH (BORGHESAN et al., 2013) e a região ribossomal do espaçador transcrito interno ITS1 (SCHÖNIAN et al., 2003) foram aplicadas para detectar qualquer espécie dentro de tripanossomatídeos. Além disso, regiões gênicas designadas como LinJ31 e Crid2.1 foram amplificadas especificamente para detectar *Leishmania infantum* e *Crithidia*-like, respectivamente. Após confirmada amplificação dos fragmentos no tamanho esperado por meio de gel de eletroforese, os dados foram anotados no painel informativo (Tabela 4).

TABELA 4 - Painel informativo da caracterização molecular e análise morfológica para triagem de isolados clínicos.

	Parasito							
	Tipagem Molecular					Morfologia		
	Universal			Espécie-específico				
	GAPDH	SSU rRNA	ITS1	LinJ31	Crid2.1	Promastigota	Coanomastigota	Outros
<sup>3</sup> LVH09	P	P	P	N	P		P	
<sup>3</sup> LVH21	Pf	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH23	Pf	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH27	P	P	P	N	P		P	
<sup>3</sup> LVH29	P	P	P	N	P		P	
<sup>3</sup> LVH32	Pf	P	P	N	P		P	
<sup>3</sup> LVH 33	Pf	P	P	N	P		P	
<sup>3</sup> LVH41	P	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH42	P	P	P	N	N		P	
<sup>3</sup> LVH43	P	P	P	N	P		P	

<sup>3</sup> LVH44	Pff	P	P	N	P		P	
<sup>3</sup> LVH49	P	P	P	N	P		P	
<sup>3</sup> LVH52	Pf	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH54	P	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH55	P	P	P	N	P		P	
<sup>3</sup> LVH56	P	P	P	N	P		P	
<sup>2</sup> LVH57	P	P	P	P	P		P	
<sup>3</sup> LVH59	Pf	P	P	P	N		P	
<sup>3</sup> LVH60	P	P	P	P	P		P	
<sup>3</sup> LVH60a	P	P	NA	P	P		P	
<sup>3</sup> LVH61	P	P	P	N	P		P	
<sup>3</sup> LVH62	P	P	NA	N	P		P	
<sup>3</sup> LVH65	P	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH66	P	P	P	P	P		P	
<sup>3</sup> LVH72	P	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH79	Pf	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH80	N	P	P	N	P	NA	NA	NA
LVH81	P	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH82	Pf	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH84	Pf	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH85	P	P	P	N	P	NA	NA	NA
<sup>3</sup> LVH87	P	NA						
<sup>3</sup> LVH90	P	NA						
<sup>3</sup> LVH91	P	NA	P	NA	NA		P	
<sup>3</sup> LVH92	P	NA	NA	NA	NA		P	
<sup>3</sup> LVH93	P	NA						
<sup>3</sup> LVH94	Pf	P	P	N	P		P	
<sup>3</sup> LVH94a	Pf	P	P	N	P		P	
<sup>3</sup> LVH95	N	NA						
<sup>3</sup> LVH96	P	NA	NA	NA	NA		P	
<sup>3</sup> LVH101	Pf	P	NA	N	P	NA	NA	NA

<sup>2</sup> LVH102	N	P	P	N	N	P		
<sup>2</sup> LVH103	NA	P	NA	N	N	NA	NA	NA
<sup>3</sup> LVH105	NA	P	NA	N	NA	NA	NA	NA
<sup>2</sup> LVH106	NA	N	NA	N	N	NA	NA	NA
<sup>3</sup> LVH107	NA	N	NA	N	P	NA	NA	NA
<sup>2</sup> LVH109	NA	N	NA	N	N	NA	NA	NA
<sup>2</sup> LVH110	NA	P	NA	N	N	NA	NA	NA
<sup>2</sup> LVH114	NA	N	NA	N	N	NA	NA	NA
<sup>2</sup> LVH115	NA	N	NA	P	N	NA	NA	NA
<sup>3</sup> LVH117	P	P	NA	N	Pf	NA	NA	NA
<sup>2</sup> LVH119	NA	P	NA	P	N	NA	NA	NA
<sup>3</sup> LVH120	P	P	NA	N	Pf	NA	NA	NA
<sup>3</sup> LVH122	NA	P	NA	N	N	NA	NA	NA
<sup>3</sup> LVH123	N	NA						
<sup>2</sup> LVH132	NA	N	NA	P	N	NA	NA	NA
<sup>2</sup> LVH133	NA	P	NA	P	N	NA	NA	NA
<sup>3</sup> LVH134	NA	N	NA	N	Pf	NA	NA	NA
<sup>3</sup> LVH135	NA	Pf	NA	P	P	NA	NA	NA
<sup>3</sup> LVH137	NA	N	NA	N	P	NA	NA	NA
<sup>3</sup> LVH138	NA	P	NA	N	Pf	NA	NA	NA
<sup>2</sup> LVH139	NA	Pf	NA	N	N	NA	NA	NA
<sup>3</sup> LVH140	NA	P	NA	N	P	NA	NA	NA
<sup>3</sup> LVH142	NA	N	NA	N	P	NA	NA	NA
<sup>2</sup> HUUF514	P	P	P	P	N	P		
<sup>2</sup> PP75	P	P	P	P	N	P		
<sup>1</sup> LbrM2903	P	P	P	N	N	NA	NA	NA

Parasito: Amostras de cultura de pacientes diagnosticados com leishmaniose humana. 1. Cepas de *Leishmania braziliensis*, causadora de leishmaniose cutânea. 2. Cepa de *Leishmania infantum*, causadora de leishmaniose visceral. 3. Possível cepa de *Crithidia*-like, a serem investigadas. Amostras de PCR positivas (P), positivas fracas (Pf) negativas (N) e não avaliadas (NA) para os primers GAPDH, SSU, ITS, HPS70, LinJ31 e Crid2.1 representadas para todas as amostras.

A triagem de isolados clínicos foi realizada com 64 amostras provenientes de pacientes diagnosticados com leishmaniose visceral humana (LVH) e utilizamos como controle positivo para

*L. infantum* as cepas HUUFS14 e PP75 e, para *L. braziliensis*, a cepa LbrM2903. A tipagem molecular realizada com o gene GAPDH apresentou 41 amostras positivas, 4 amostras negativas e 19 amostras não analisadas. Com o gene SSU rRNA obtivemos 47 amostras positivas, 9 amostras negativas e 8 amostras não analisadas. O marcador da região ITS1 foi recentemente utilizada e apresentou 33 amostras positivas e 31 amostras não analisadas. Esse conjunto de marcadores genéticos devem ser positivos para qualquer tripanossomatídeo. Amostras negativas podem indicar parasitos não tripanossomatídeos ou i) baixa concentração de DNA com resíduos que dificultaram a amplificação dessas regiões e/ou ii) integridade comprometida do material.

A tipagem molecular realizada com o marcador específico para *L. infantum* identificou 9 amostras positivas, 47 amostras negativas e 8 amostras não analisadas, sendo que 4% das amostras positivas não foram exclusivas para o marcador LinJ31, sendo positiva para Crid2.1 também. Já a análise com o marcador específico para *Crithidia*-like, 41 amostras foram positivas, 14 amostras deram negativas e 9 amostras não foram analisadas até o presente momento. Dentre as amostras positivas, 6% amplificaram para LinJ31, marcador de *L. infantum*. Esse resultado foi apontado nas amostras LVH57, LVH60, LVH60a e LVH66, indicando possível co-infecção dos parasitos *L. infantum* e *Crithidia*-like nesses isolados clínicos. Análise genômica da amostra LVH60 policlonal sugeriu possível co-infecção com *L. infantum*, corroborando os resultados aqui apresentados.

#### 4.4 ANÁLISES DE SEQUÊNCIA E FILOGENÉTICA MOLECULAR

Os produtos de PCR positivos foram purificados e enviados para serviço de sequenciamento. Todos os dados brutos de sequenciamento foram analisados quanto à sua qualidade e uma etapa de corte de sequências não informativas foi realizada. Ambas as etapas de controle de qualidade e cortes foram realizadas pelo programa *Chromas*, com o intuito de melhorar as sequências e manter dados mais robustos para fazer o alinhamento dessas sequências. Após realizado o controle de qualidade, os alinhamentos foram gerados pelo programa *Muscle* (EDGAR, 2004) por meio do programa *Mega 6* (TAMURA et al., 2013).

Marcadores genéticos são rotineiramente aplicados para reconstrução filogenética molecular de protozoários Trypanosomatidae (VOTÝPKA et al., 2013). A técnica de nested PCR tem grande sucesso com amostras sanguíneas e foi realizada com o marcador do RNA ribossomal obtido de *Trypanosoma brucei* (TRY927) com uma sequência de 927 pb (NOYES et al., 1999).

Após a amplificação de TRY927, uma segunda PCR foi aplicada na porção interior de TRY927, região da subunidade 18S do RNA ribossomal (SSU rRNA) com 555 pb (BORGHESAN et al., 2013; NOYES et al., 1999; VOTÝPKA et al., 2013; WILFERT et al., 2011; YURCHENKO et al., 2006). Os iniciadores foram adequados para detectar sequências homólogas e conservadas de todas as espécies de tripanossomatídeos. A região do RNA ribossomal do espaçador transcrito interno (ITS1) com 300-350 pb (SCHÖNIAN et al., 2003) e o gene gliceraldeído 3-fosfato desidrogenase (GAPDH) de produto aproximado de 1000pb (BORGHESAN et al., 2013) são altamente conservados dentro da família Trypanosomatidae e usados para identificar organismos pertencentes a esse grupo e realizar análises de filogenia. Em taxonomia molecular, essas sequências são vastamente aplicadas em tripanossomatídeos (CÁSSIA-PIRES et al., 2017; MARCILI et al., 2014; RANGEL et al., 2019). Além disso, dois marcadores espécie-específicos foram aplicados nos isolados clínicos para identificação qualitativa de parasitos *Leishmania infantum* (LinJ31Seq) e *Crithidia-like* (Crid2.1Seq). O marcador LinJ31Seq pertence a uma região do cromossomo 31 de *Leishmania infantum* e possui 444 pb e o marcador Crid2.1Seq possui 500 pb. A análise produzida da região da subunidade 18S de rRNA (SSU rRNA) foi comparada com outros tripanossomatídeos através do *programa blast*, que usa um algoritmo capaz de comparar sequências biológicas (nucleotídeos e proteínas). Esta sequência é frequentemente usada em estudos de taxonomia molecular por se tratar de região conservada e presente em todos os tripanossomatídeos (CÁSSIA-PIRES et al., 2017).

A reconstrução filogenética da região da subunidade 18S de rRNA (SSU rRNA) foi realizada levando-se em conta o melhor modelo de substituição de nucleotídeos dentre 24 modelos testados (Tabela 5). Os modelos com as pontuações BIC (*Bayesian Information Criterion*) mais baixas foram considerados para descrever o padrão de substituição da melhor *fit model*. Para cada modelo, o valor AICc (*Akaike Information Criterion corrected*), o valor de Máxima Verossimilhança (lnL) e o número de parâmetros (incluindo comprimentos dos ramos) também foram apresentados (NEI; KUMAR, 2000). Outros dados também foram reportados durante a análise do melhor modelo de substituição - porém omitidos na tabela 5, como valores presumidos ou estimados de viés de transição/transversão (R) seguidos por frequências de nucleotídeos (f) e taxas de substituições de bases (r) para cada par de nucleotídeos. Os valores relativos de r instantâneo devem ser considerados ao avaliá-los. Para simplificar, a soma dos valores de r é igual a 1 para cada modelo. Para estimar os valores de máxima verossimilhança, uma topologia em árvore foi calculada automaticamente. A análise envolveu 18 sequências de nucleotídeos e 587 pb

no conjunto de dados final. As análises de filogenética molecular foram realizadas no Mega 6 (TAMURA et al., 2013) e o critério usados para escolha de quais sequências seriam inclusas como representativas de todos os isolados clínicos foi o comprimento das sequências mais próximo ao tamanho de SSU rRNA (555 pb).

TABELA 5 - Modelos de substituição de nucleotídeos pelo método de Máxima Verossimilhança.

Modelo	Parâmetros	BIC	AICc	lnL
K2	34	3.025.307	2.784.510	-1.358.120
K2+I	35	3.034.998	2.787.126	-1.358.420
T92	35	3.035.883	2.788.011	-1.358.863
T92+I	36	3.040.496	2.785.550	-1.356.624
TN93	38	3.044.811	2.775.719	-1.349.692
JC+I	34	3.046.496	2.805.699	-1.368.715
HKY	37	3.047.897	2.785.878	-1.355.780
TN93+I	39	3.049.161	2.772.997	-1.347.322
JC	33	3.049.814	2.816.091	-1.374.919
HKY+I	38	3.050.696	2.781.604	-1.352.634
GTR+G	42	3.051.735	2.754.356	-1.334.973
GTR+I	42	3.059.104	2.761.724	-1.338.658
GTR+G+I	43	3.060.567	2.756.117	-1.334.844
GTR	41	3.064.206	2.773.898	-1.345.754
K2+G	35	71.256.067	71.008.195	-35.468.955
K2+G+I	36	71.265.157	71.010.212	-35.468.955
T92+G	36	71.267.271	71.012.326	-35.470.012
TN93+G	39	71.268.440	70.992.276	-35.456.961
T92+G+I	37	71.276.361	71.014.342	-35.470.012
TN93+G+I	40	71.277.531	70.994.294	-35.456.961
HKY+G	38	71.277.763	71.008.671	-35.466.168
JC+G	34	71.284.926	71.044.129	-35.487.930
HKY+G+I	39	71.286.853	71.010.689	-35.466.168
JC+G+I	35	71.294.017	71.046.145	-35.487.930

Abreviações: GTR: General Time Reversible; HKY: Hasegawa-Kishino-Yano; TN93: Tamura-Nei; T92: Tamura 3-parameter; K2: Kimura 2-parameter; JC: Jukes-Cantor.

A análise filogenética de SSU rRNA mostrou que isolados clínicos policlonais de leishmaniose visceral humana (LVH) agruparam-se com *C. fasciculata*, permanecendo no clado de espécies monoxênicas não patogênicas ao ser humano e distantes do clado *Leishmania sp.* Essa análise ainda mostrou que os isolados HUUFS14 e LTCP393, controles positivos para *L. infantum* e *L. braziliensis*, respectivamente, se encontram no grupo de organismos dixênicos, como previsto (Figura 10).

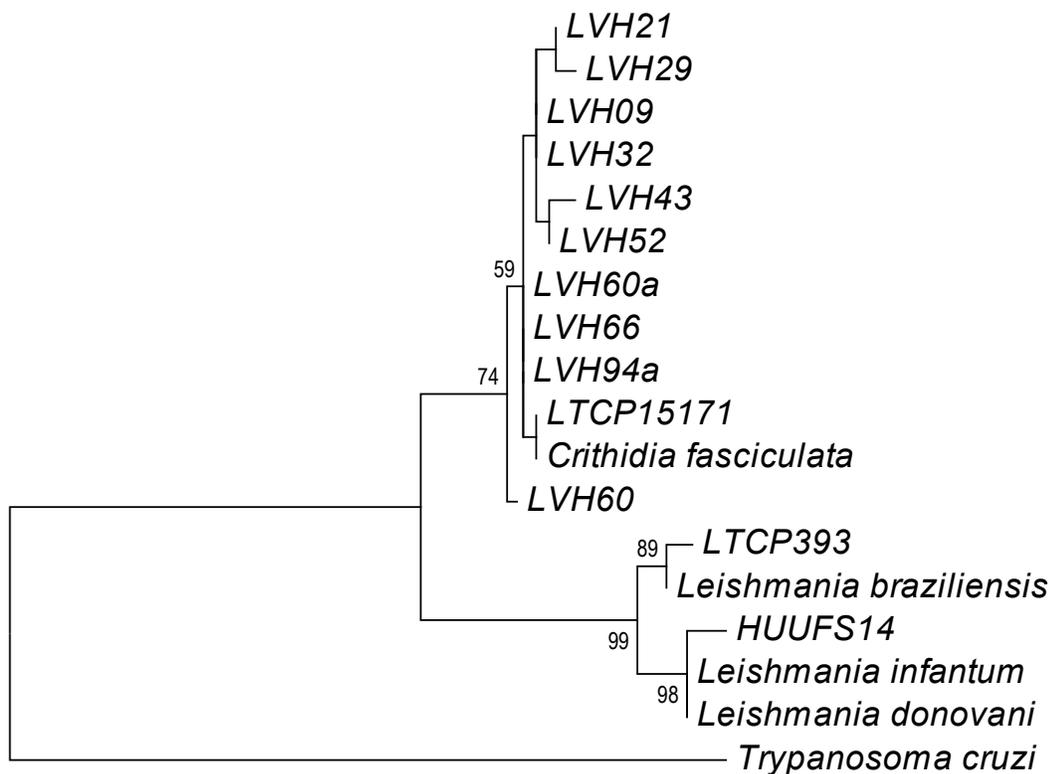


FIGURA 10 - Árvore filogenética de região da subunidade 18S de rRNA (SSU rRNA) de amostras policlonais de isolados clínicos. Sequências referência de *Leishmania*, *Leptomonas* e *Crithidia* obtidas pelo NCBI (*blastn*). Método feito por Maximum Likelihood (ML) com o melhor modelo de substituição, Kimura 2-parameter.

Além das análises filogenéticas e por matriz de identidade, os alinhamentos múltiplos de sequência foram produzidos com as sequências SSU rRNA, LinJ31 e Crid2.1 (Figuras 11, 12 e 13). O programa *Bioedit* gerou visualização gráfica definida por cores e cada cor representa uma nucleobase: adenina, timina, citocina e guanina. Dessa forma foi possível comparar a sequência da

região de interesse presente nas amostras e organismos referência obtidos por algoritmo *blastn* e verificar variação e similaridade entre as linhagens e espécies.

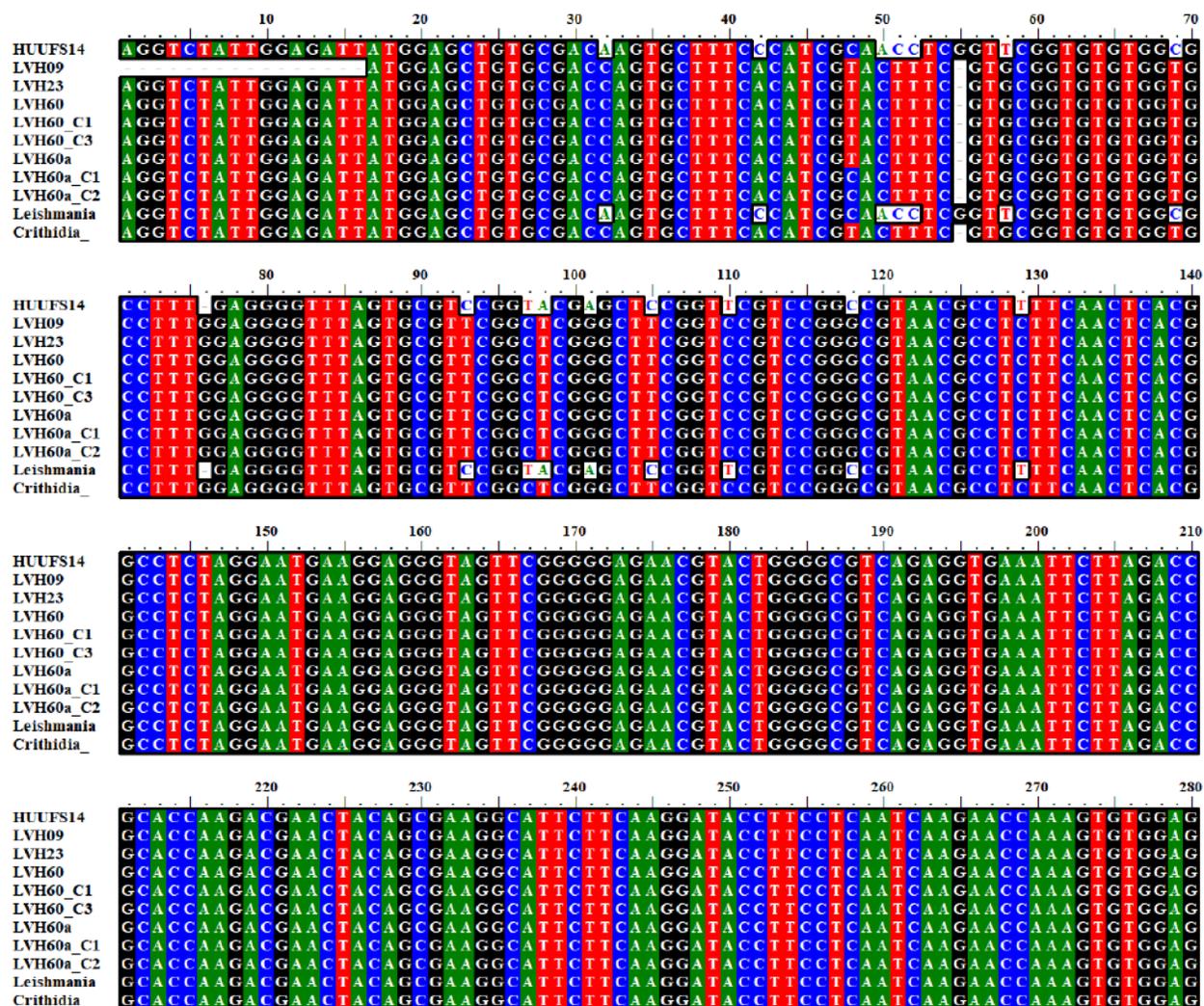


FIGURA 11 - Alinhamento representativo das sequências 18S do RNA ribossomal SSU de culturas clonais e policlinalis de isolados clínicos de leishmaniose visceral humana. As referências *Crithidia fasciculata* e *Leishmania infantum* foram obtidas a partir de sequências de nucleotídeos pelo banco de dados no NCBI com algoritmo BLASTN, alinhados pela ferramenta Muscle no programa Mega 6 e visualizados pelo programa Bioedit. O isolado HUUFS14 é o controle positivo de *L. infantum* do laboratório.

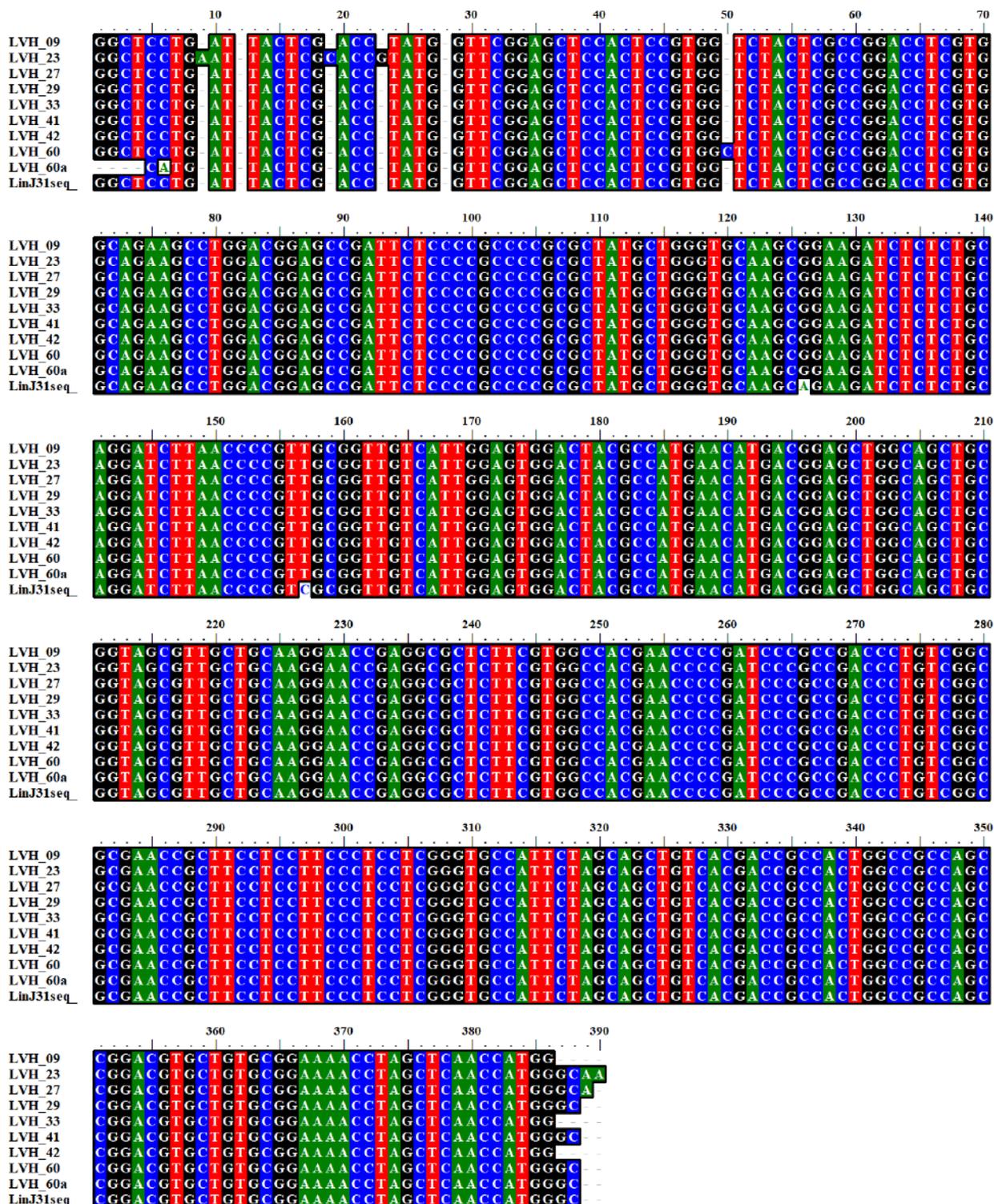


FIGURA 12 - Alinhamento representativo das seqüências LinJ31seq de culturas policlonais de isolados clínicos de leishmaniose visceral humana. A seqüência referência usada foi o próprio amplicon LinJ31seq, espécie-específico de *Leishmania infantum*. As seqüências foram alinhadas pela ferramenta Muscle no programa Mega 6 e visualizados pelo programa Bioedit.

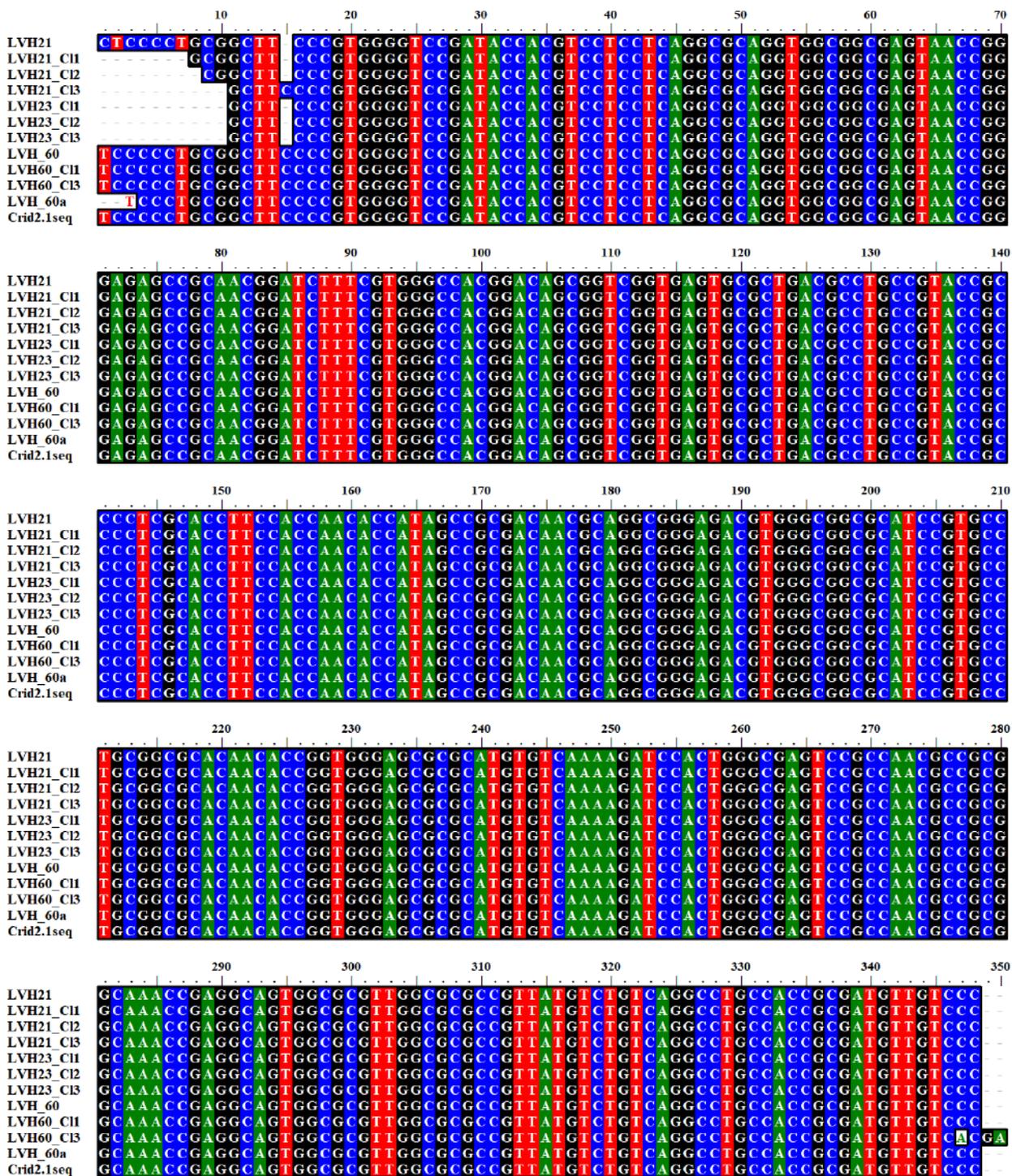


FIGURA 13 - Alinhamento representativo das sequências Crid2.1seq de culturas clonais e policlonais de isolados clínicos de leishmaniose visceral humana. A sequência referência usada foi o próprio amplicon Crid2.1seq espécie-específico de patógeno *Crithidia-like*. As sequências foram alinhadas pela ferramenta Muscle no programa Mega 6 e visualizados pelo programa Bioedit.

Coletivamente, as análises morfológicas e moleculares contribuiram para compor um painel informativo que será útil para a escolha de isolados clínicos que terão seus genomas sequenciados futuramente.

#### 4.5 ANÁLISE GENÔMICA

##### 4.5.1 Análise de cultura Policlonal de LVH60 e LVH60a

Para análise bioinformática dos genomas montados *Crithidia*-like policlonal LVH60 e LVH60a (MARUYAMA et al., 2019), duas referências foram selecionadas para conduzir os estudos comparativos com base nos resultados obtidos por análises filogenéticas no processo de triagem. Sequências genômicas de *Leishmania infantum* JPCM5 (código de acesso GeneBank:GCA\_000002875.2) e *Crithidia fasciculata* (código de acesso GCA\_000331325.2) extraídas do *National Center for Biotechnology Information* (NCBI) foram usadas na análise dos genomas "draft" LVH60 e LVH60a. O genoma de HUUFS14 foi usado como referência positiva de *L. infantum* obtida da mesma região de coleta das cepas LVH60 e LVH60a. O tamanho predito do genoma de LVH60 e LVH60a (~41 Mb) montados em 4522 e 4495 *contigs*, respectivamente, é quase 10 Mb maior que o genoma da referência *L. infantum* JPCM5 (~32 Mb) e mais próximo do tamanho do genoma de *C. fasciculata* (~41 Mb) (Tabela 6). Já o tamanho do genoma de HUUFS14 é similar a *L. infantum* JPCM5 (~32 Mb), mostrando que, as cepas LVH60 e LVH60a não apresentam o tamanho de genoma das linhagens pertencentes a espécie *L. infantum*.

TABELA 6 - Genoma predito dos parasitos *Leishmania infantum*, *Crithidia fasciculata* e *Crithidia*-like.

	Tamanho do genoma predito (pb)
<i>L. infantum</i> JPCM5	32.115.981
<i>C. fasciculata</i>	41.260.738
LVH60	41.900.000
LVH60a	41.300.000
HUUFS14	32.486.338

O genoma das referências *L. infantum* e *C. fasciculata* foram obtidos no banco de dados do NCBI e os isolados clínicos LVH60 e LVH60a policlonal foram montados com Abyss (SIMPSON et al., 2009).

O programa MUMmer 3.0 foi utilizado para analisar o genoma completo desses isolados clínicos previamente sequenciados e montados em *contigs* (MARUYAMA et al., 2019). Este sistema é capaz de realizar alinhamento rápido entre dois genomas completos ou rascunho (*draft*) de genomas (sendo uma referência e o outro *query*), identificando porcentagem de identidade entre dois genomas. O *output* NUCmer em formato coords.txt mostrou todas as características do alinhamento descritas abaixo, sendo possível sistematizar análises para obter uma matriz de identidade entre a referência e o *query*. Os genomas referência utilizados foram de *L. infantum* e *C. fasciculata*, com os respectivos códigos de acesso do GeneBank:GCA\_000002875.2 e GCA\_000331325.2; LVH60, LVH60a e HUUFS14 foram os genomas “*draft*” sequenciados (*query*) e analisados. Para a construção da matriz de identidade, os *contigs* da sequência *query* foram agrupados por cromossomo da referência. Para uma primeira análise, toda cobertura de *query* inferior a 90% foi descartada, a fim de realizar um alinhamento mais idêntico possível a referência. Em seguida, foi obtido o número total de gaps entre os *contigs* que formavam um determinado cromossomo. Feito isso, obtivemos a cobertura dos *contigs* em relação a referência pela fórmula  $COV_{total} = (\text{número total de contigs do cromossomo} - \text{gap total do cromossomo}) / \text{número total de nucleotídeos do cromossomo} * 100$ . Realizando a cobertura para todos os cromossomos, o número total de gaps de todos os cromossomos foi somado e o resultado foi subtraído pelo genoma completo da referência, dividindo por 100. Com isso, foi possível fazer a matriz de identidade dos alinhamentos comparados. Os dados gerados por NUCmer foram listados abaixo:

[S1] Início da região do alinhamento na sequência referência;

[E1] Final da região do alinhamento na sequência referência;

[S2] Início da região do alinhamento na sequência *query*;

[E2] Início da região do alinhamento na sequência *query*;

[LEN 1] Comprimento da região de alinhamento na sequência referência, medido em nucleotídeos;

[LEN 2] Comprimento da região de alinhamento na sequência *query*, medido em nucleotídeos;

[% IDY] Porcentagem de identidade do alinhamento calculado pelo (número de correspondência exata) / ([LEN 1] + inserções no *query*);

[LEN R] Comprimento da sequência referência;

[LEN Q] Comprimento da sequência *query*;

[COV R] Porcentagem de cobertura do alinhamento na sequência referência, calculado por  $[LEN 1] / [LEN R]$ ;

[COV Q] Porcentagem de cobertura do alinhamento na sequência *query*, calculado por  $[LEN 2] / [LEN Q]$ ;

[TAGS] A FastA ID da referência e do *query*.

O alinhamento genômico foi realizado em pares com uma sequência *query* e uma sequência referência pelo programa MUMmer 3.0 (KURTZ et al., 2004) com a função NUCmer. Foi utilizado três isolados clínicos que tiveram seus genomas completos sequenciados, sendo eles LVH60 e LVH60a (*Crithidia-like*) e HUUFS14 (*L. infantum*). LVH60 e LVH60a foram obtidos do mesmo paciente, que apresentou um caso atípico da doença. O paciente do sexo masculino tinha 63 anos e foi diagnosticado com LV, contudo apresentou três recidivas sem sucesso quimioterápico. Em sua última internação apresentou lesões cutâneas da qual foi isolada a cepa LVH60a e sem sucesso ao tratamento, o paciente foi submetido à retirada do baço. Por complicações cirúrgicas ele veio a óbito (MARUYAMA et al., 2019).

Os genomas referência dos organismos *L. infantum* JPCM5 e *C. fasciculata* foram alinhados com os genomas *queries* LVH60, LVH60a e HUUFS14. Além disso, os genomas de LVH60 e LVH60a foram usados como referência para realizar o alinhamento contra eles mesmos. Após a obtenção do *output* pelo NUCmer, os dados brutos foram analisados a partir do tamanho total de alinhamento (TTA) em relação a referência e a porcentagem média de identidade (idt%) do genoma referência com o *query* (Tabela 7).

TABELA 7 - Alinhamento genômico entre *Crithidia*-like e referências *Leishmania infantum* e *Crithidia fasciculata*.

Referência	Query	Genoma referência (nt)	TTA (nt)	Média idt%
Linf	HU	32.115.981	39.123.488	96,28
	Cfasc	32.115.981	1.620.975	87,68
	LVH60*	32.115.981	1.950.080	89,57
	LVH60a*	32.115.981	1.984.533	89,53
Cfasc	HU	41.260.738	1.470.826	88,87
	LVH60*	41.260.738	81.587.902	92,05
	LVH60a*	41.260.738	82.050.698	91,98
	Linf	41.260.738	1.190.966	89,27
LVH60a*	LVH60*	41.300.000	83.531.381	95,39
LVH60*	LVH60a*	41.900.000	84.483.678	95,76

(Linf) *Leishmania infantum*; (Cfasc) *Crithidia fasciculata*; (HU) HUUFS14 amostra de *L. infantum chagasi*; (LVH60) amostra isolada de medula óssea; (LVH60a) amostra isolada de lesão cutânea; \* amostras de mesmo paciente, (nt) nucleotídeos; (TTA) tamanho total alinhado redundante; (média idt%) Porcentagem da média de identidade da região alinhada.

Como observado na Tabela 6, o *output* do alinhamento genômico entre a referência *L. infantum* JPCM5 e o *query* HUUFS14 apresentou o TTA superior ao tamanho do seu próprio genoma (~39Mb), perfazendo um alinhamento completo com 100% de cobertura entre os genomas. De modo esperado, mesmo diante de sequências redundantes compondo o alinhamento, HUUFS14 é genoma de controle positivo para *L. infantum* e, desse modo, apresentaria identidade superior a 96%. Por outro lado, a referência *L. infantum* JPCM5 quando alinhada ao *query* *C. fasciculata* apresentou TTA 1,6 Mb e como esperado, esse valor correspondeu a um alinhamento de apenas 6,25% ao ser alinhado também aos *queries* LVH60 e LVH60a.

Ao contrário do alinhamento com a referência *L. infantum*, quando as amostras de LVH60 e LVH60a são alinhadas a referência *C. fasciculata*, o TTA superior foi completo (100% de cobertura) entre os genomas. A porcentagem média de identidade reflete no TTA, mostrando valores superiores a 90% para todos os alinhamentos com o tamanho total alinhado superior ao tamanho do genoma de sua referência. Os valores maiores que 90% referem-se aos alinhamentos com o TTA inferior ao tamanho do genoma de sua referência. Esses resultados sugerem que as

amostras de LVH60 e LVH60a apresentam maior similaridade com a referência *C. fasciculata*, espécie monoxênica não patogênica ao homem (ALCOLEA et al., 2014).

Para os alinhados contra os genomas referência anotado, foi criada uma matriz de identidade dos queries com suas respectivas referências (Tabela 8). Apenas cobertura superior a 50% dos *contigs* do query sobre a referência foram consideradas nessa análise. Essa análise mostra que LVH60 e LVH60a apresentam maior identidade com *C. fasciculata* do que com *L. infantum*.

Diante desses resultados concluiu-se que os isolados clínicos LVH60 e LVH60a apresentam genoma mais parecido com o genoma de *C. fasciculata* do que com o genoma de *L. infantum*, corroborando resultados de filogenômica publicados que mostram que esse isolados são *Crithidia*-like invés de *L. infantum* (MARUYAMA et al., 2019).

TABELA 8 - Matriz de identidade entre os genomas sequenciados e as referências *Crithidia fasciculata* e *Leishmania infantum*.

	<i>C. fasciculata</i>	<i>L. infantum</i>
LVH60	70,43%	0,55%
LVH60a	71,79%	0,5%
HUUFS14	0,081%	82,04%

Identidade gerada considerando do alinhamento dos contigs ao genoma referência que apresentaram acima de 50% de cobertura.

#### 4.5.1.1 Mapeamento de Amostras Policlonais em Genoma de Referência Concatenado

Para avaliar a ocorrência de sequências pertencentes a *L. infantum* nos dados de sequenciamento de LVH60 e LVH60a, as *reads* de interesse provenientes de amostras policlonais foram alinhadas às referências concatenadas de *Crithidia fasciculata* (30 cromossomos, 41 Mb, disponível publicamente no TriTrypDB) e *Leishmania infantum* JPCM5 (36 cromossomos, 32 Mb, disponível publicamente no TriTrypDB). As *reads* para essa análise foram filtradas de acordo com o índice de qualidade phred  $\geq 60$ , que garantiu *reads* com uma chance de erro a cada  $10^6$  nucleotídeos, ou seja, sequências de interesse 99,9% confiáveis. A profundidade média de cobertura expressa a quantidade de vezes que um genoma foi totalmente preenchido por um

conjunto de nucleotídeos, demonstrando a redundância de cobertura de um genoma. Os resultados mostraram que a profundidade média de cobertura por cromossomo (Figura 14 e 15) foi de 1X no genoma de *L. infantum* e 60X em *C. fasciculata*. Esses dados corroboram mais uma vez a semelhança genômica dessas cepas com organismo monoxênico *C. fasciculata*, sugerindo a ocorrência de um patógeno *Crithidia*-like associado em casos de leishmaniose visceral.

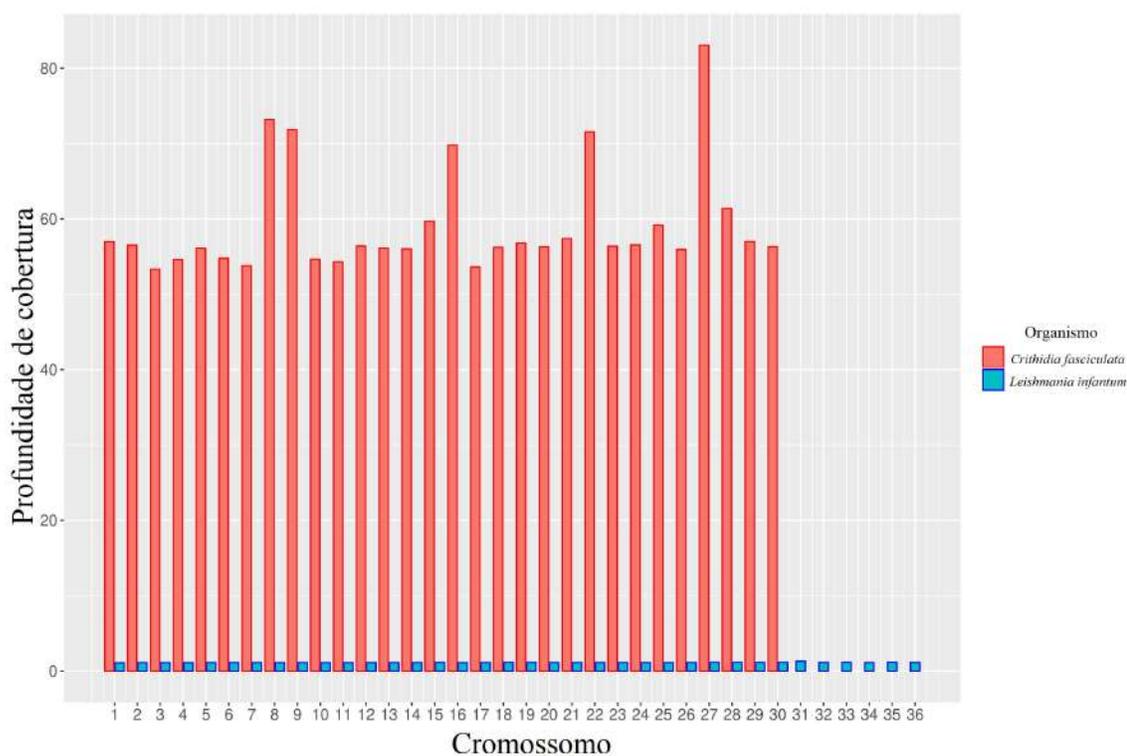


FIGURA 14 - Profundidade de cobertura de *reads* do isolado LVH60 policlonal nos genomas de referência de *L. infantum* e *C. fasciculata* concatenados.

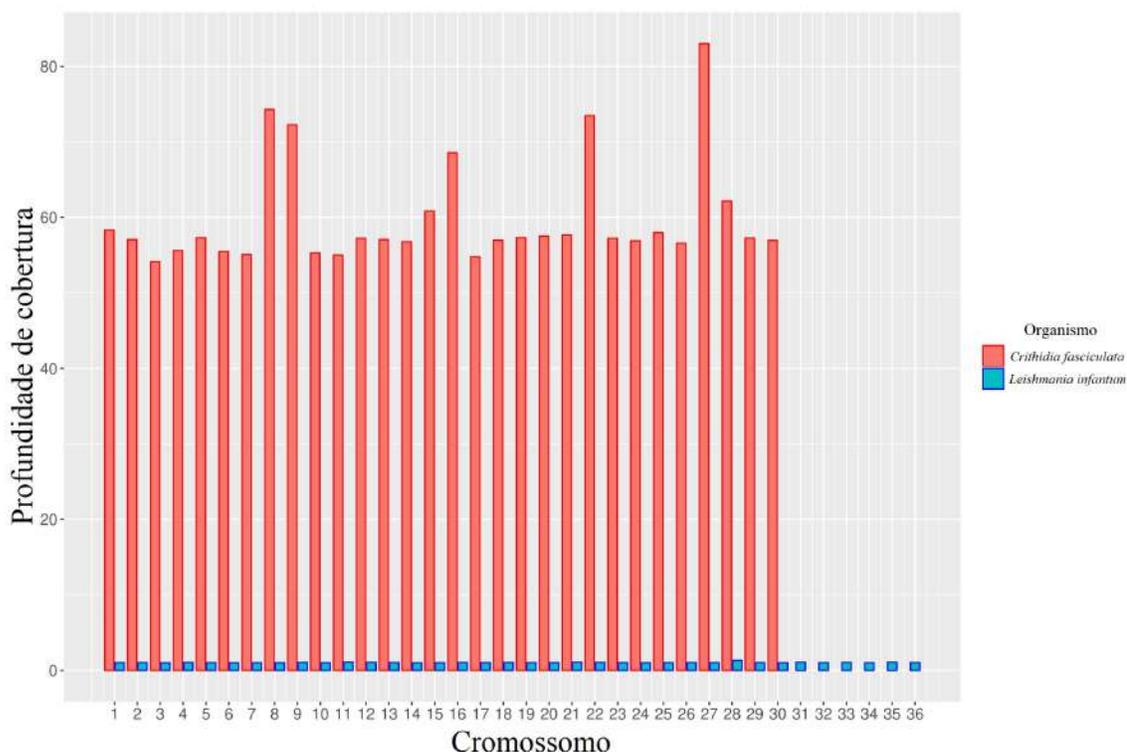


FIGURA 15 - Profundidade de cobertura de *reads* do isolado LVH60a policlonal nos genomas de referência de *L. infantum* e *C. fasciculata* concatenados.

O mapeamento foi avaliado pelo número total de nucleotídeos posicionados ao longo de cada cromossomo do genoma referência, obtendo a partir desse cálculo a porcentagem de mapeamento dos isolados clínicos de interesse com base nas referências concatenadas. Exemplificando, se um cromossomo da referência possui 100 pb e a sequência de interesse possui 90 pb mapeados a esse cromossomo, então o valor de mapeamento é 90%.

A referência de *L. infantum* JPCM5 foi mapeada pelas *reads* do isolado LVH60 preenchendo cerca de 20% do genoma referência (Figura 16) e essa porcentagem de mapeamento corresponde a apenas 1X de cobertura. O mapeamento realizado pelas *reads* do isolado LVH60a representou aproximadamente 2% do genoma da referência de *L. infantum* (Figura 17) com cobertura de 1X da referência. Essa análise mostrou um pequeno número de nucleotídeos mapeado que se estendeu com 1X de cobertura, i.g. 1 sequência, pelos cromossomos de *L. infantum*. Já o mapeamento obtido entre esses isolados clínicos e a referência genômica *C. fasciculata* foi superior a 90% e cerca de 60X de cobertura ao longo dos cromossomos.

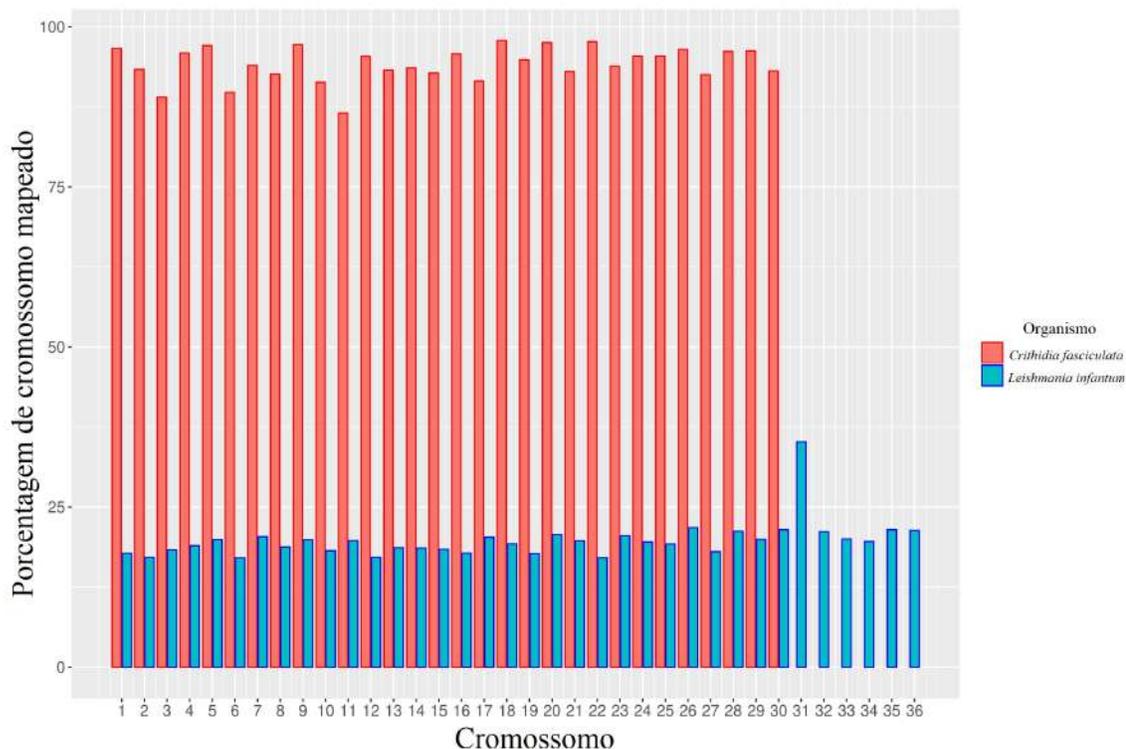


FIGURA 16 - Percentagens dos cromossomos (cálculo pela extensão em pb) das referências *L. infantum* e *C. fasciculata* mapeadas por reads do isolado LVH60 policlonal.



FIGURA 17 - Percentagens dos cromossomos (cálculo pela extensão em pb) das referências *L. infantum* e *C. fasciculata* mapeadas por reads do isolado LVH60a policlonal.

Visto que esses dados apresentaram elevado rendimento inicial de sequenciamento (aproximadamente 450X, 42 milhões de *reads*) é provável que a amostra LVH60 (oriunda da medula óssea) estivesse com DNA resquicial de *L. infantum*, indicando possível co-infecção na medula óssea. Entretanto, isso só poderia ser confirmado por meio de análise da medula óssea do paciente, invés de analisar a cultura de parasito derivada deste tecido.

#### 4.5.2 Sequenciamento Oxford Nanopore (*Long Reads*)

O genoma *Crithidia*-like sequenciado por ONT foi obtido a partir de um clone da cultura policlonal de LVH60a (LVH60a-C1), que foi escolhida por apresentar tamanho de 1Mb superior em relação a cepa de medula óssea (LVH60) predito de acordo com o sequenciamento Illumina realizado anteriormente (MARUYAMA *et al.*, 2019). O sequenciamento foi realizado em um equipamento MinION Oxford Nanopore e produziu 379,000 *reads* (Tabela 9).

TABELA 9 - Características gerais dos dados de sequenciamento Oxford Nanopore Technology do clone *Crithidia*-like LVH60a-C1.

	<b>Todas as <i>reads</i></b>	<b><i>Reads</i> filtradas</b>
<b><i>Reads</i></b>	379.039	357.617 (94 %)
<b>Bases*</b>	3,1 Gb	2,9 Gb
<b>Tamanho médio das <i>reads</i>**</b>	1.871 bp	1.956 bp
<b>N50***</b>	28.179 bp	28.430 bp

**Todas as *reads*** referem-se ao total de *reads* brutas do sequenciamento por MinION antes do processo de trimagem. ***Reads* filtradas** referem-se as *reads* que passaram por processo de qualidade pelo Guppy e pela remoção dos *barcodes*. \*Número total de bases nitrogenadas. \*\*Número que representa metade das *reads* são de comprimento inferior e a outra metade de comprimento superior. \*\*\*Comprimento mínimo de *reads* necessário para cobrir 50% do genoma, onde metade das *reads* necessariamente devem apresentar comprimento superior ou igual ao tamanho N50. Dados obtidos por com pycoQC 2.5.0.9.

A qualidade dos dados brutos obtidos a partir do sequenciamento por *Nanopore* foi verificada através do algoritmo *phred*, índice adotado pela comunidade científica que mede a qualidade da identificação dos nucleotídeos originados do sequenciamento de DNA (EWING et al., 1998). O índice *phred* foi utilizado para filtrar *reads* de baixa qualidade e selecionar *reads* com valor superior a 20 ( $\text{phred score} \geq 20$ ), ou seja, sequências que apresentam 1 chance de erro a cada 100 nucleotídeos representando 99% de confiabilidade (Figura 18).

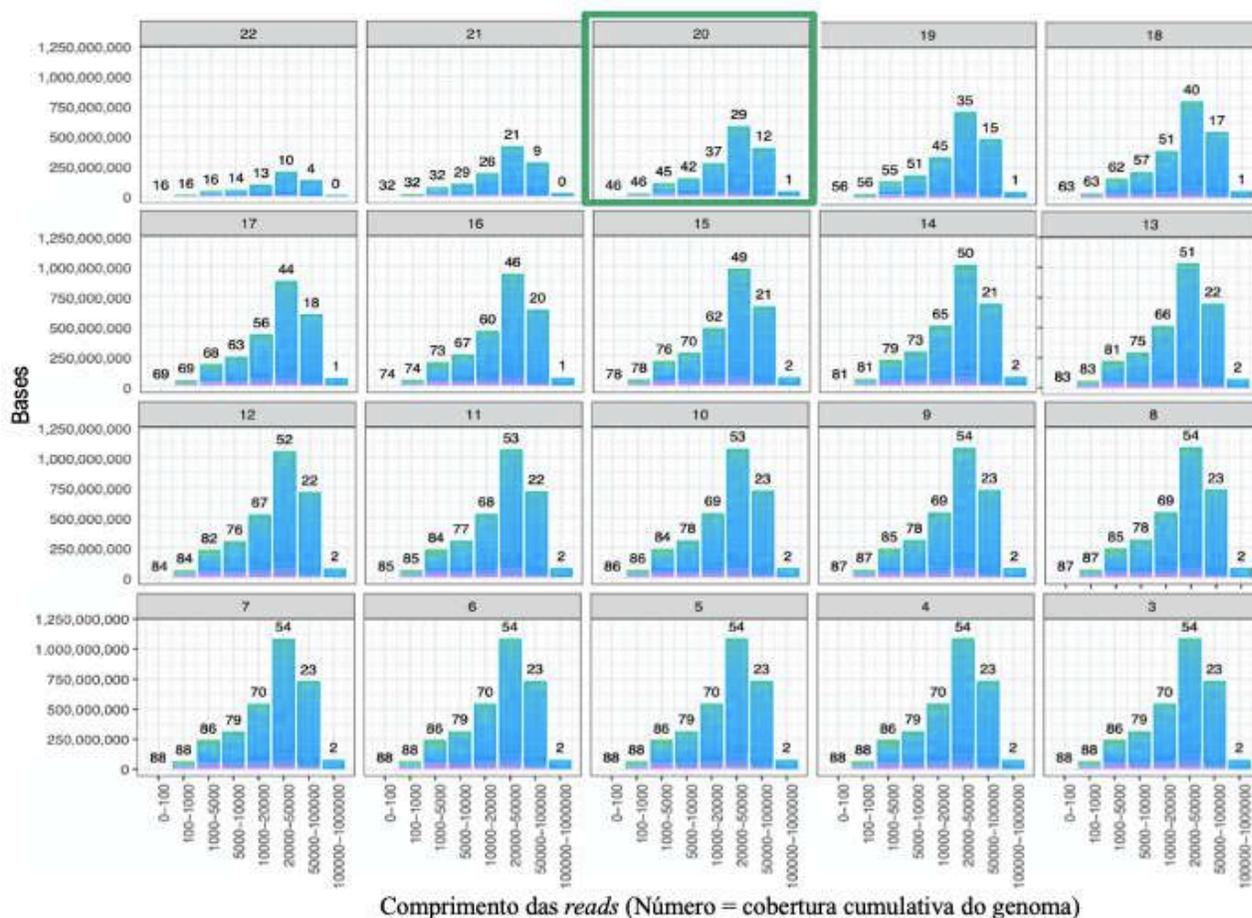


FIGURA 18 - Perfil de qualidade das *reads* de sequenciamento Oxford Nanopore Technology do clone *Crithidia*-like LVH60a-C1. Cada gráfico representa um índice médio de qualidade das *reads*, contendo nas barras verticais a cobertura de *reads* de diferente comprimento. As *reads* de comprimento mínimo de 20 Kb, cobertura de 29X com índice *phred* de valor igual e superior a 20 foi usado como corte para montagem genômica (gráfico destacado em verde).

*Reads* com índice *phred* de valor mínimo 20 apresentaram sequências de comprimento entre 20 e 50 Kb perfazendo uma cobertura de 29X do genoma e foram selecionadas para realizar a montagem bruta.

#### 4.5.2.1 Montagem do genoma do clone de LVH60a

A montagem genômica bruta produzida pelo programa Canu utilizou *reads* de sequenciamento *Nanopore* com cobertura média de 29X do tamanho do genoma a partir *reads* de 20 Kb de comprimento mínimo. A função de correção aplicada melhorou a precisão das bases das *reads* e, posteriormente, a função de corte removeu *reads* de baixa qualidade e/ou resíduos de adaptadores. A montagem bruta foi gerada através de arranjo de *reads* em sequências mais longas, os *contigs*. O comprimento médio dos *contigs* foi de 414,179 pb, apresentando como comprimento máximo *contig* com tamanho superior a 2 Mb e um genoma de 40 Mb de comprimento (Tabela 10).

TABELA 10 - Características gerais da montagem bruta do genoma do clone *Crithidia*-like LVH60a-C1

<b>Número de <i>contigs</i></b>	<b>Comprimento total*</b>	<b><i>Contig</i> Comprimento mínimo*</b>	<b>Média de comprimento dos <i>contigs</i>*</b>	<b><i>Contig</i> - Comprimento máximo*</b>	<b>N50*</b>
97	40.175.452	1.298	414.179	2.936.253	955.566

\*Comprimento em pares de bases (pb).

A partir do sequenciamento por técnicas que garantem *long reads* realizada por ONT foi produzido uma montagem bruta do parasito *Crithidia*-like proveniente da amostra clonal LVH60a-c1 com tamanho médio de 40 Mb e 97 *contigs*.

#### 4.5.2.2 Validação e edição da primeira montagem genômica

A ferramenta Tapestry foi usada para validar, editar e analisar a montagem bruta dos *contigs* LVH60a-c1 e identificar possíveis cromossomos, haplótipos, sequências teloméricas, minicírculos e outras características genômicas. Para isso, foi fornecido ao programa os *contigs* montados, as *reads* brutas do sequenciamento ONT. Após executar Tapestry no terminal, um arquivo de saída

HTML foi gerado e visualizado em um navegador. Tapestry é uma ferramenta interativa para fazer anotações e obter informações sobre o comprimento de cada *contig*, conteúdo CG e profundidade de cobertura de reads em cada *contig* montado utilizando *long reads*.

No genoma da montagem bruta de LVH60a-c1 foram identificados 33 *contigs* diplóides, 2 *contigs* triploides, totalizando 35 *contigs* únicos; 11 *contigs* representativos de minicírculo (DNA mitocondrial) e 2 *contigs* com baixa cobertura e não identificados. Também foram identificados 49 haplótipos, combinação de alelos que pertencem a um cromossomo. O total de *contigs* únicos pode inferir o número total de cromossomos, mas ainda não foi possível identificar sequência teloméricas para afirmar isso (Figura 19).

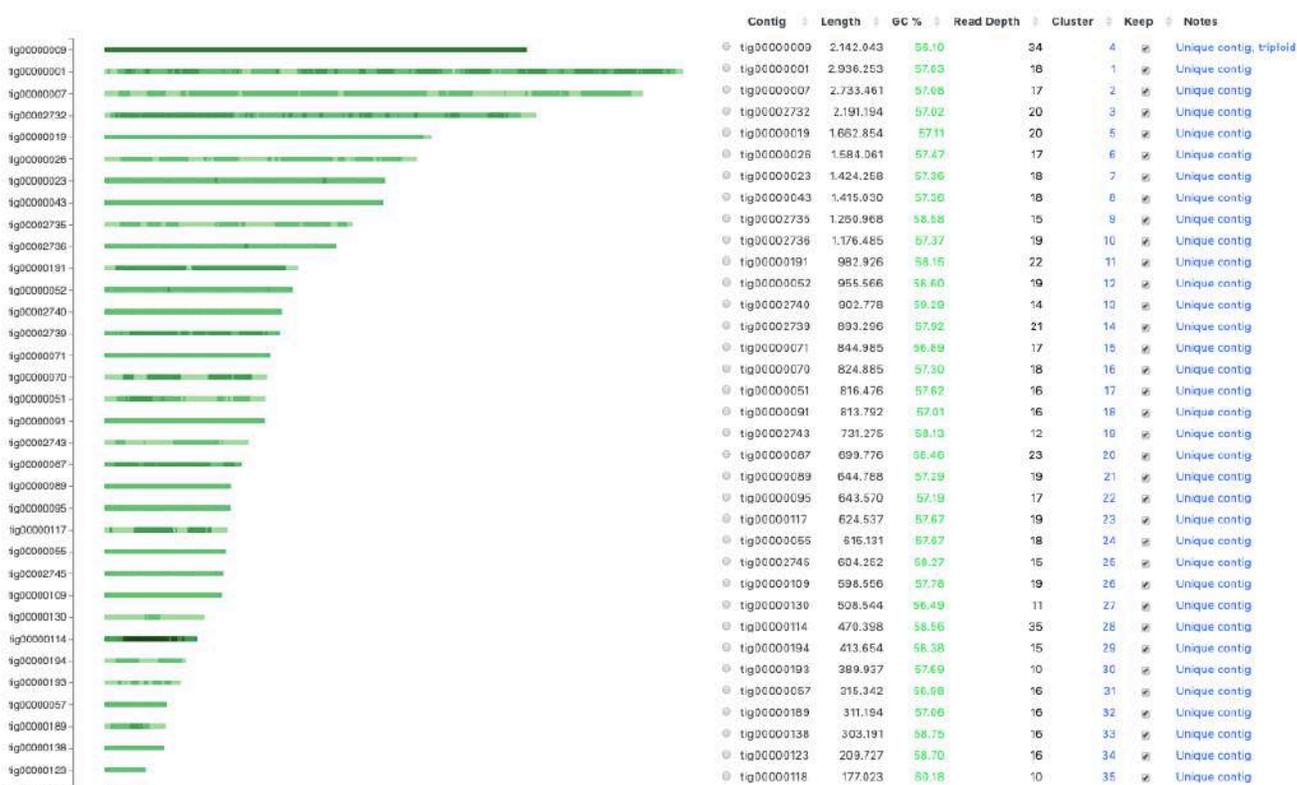


FIGURA 19 - Relatório Tapestry indicando os *contigs* únicos gerados na montagem bruta. O gráfico apresenta cada *contig* montado por linha com informações como nome do *contig*, comprimento em pares de base (pb), porcentagem de GC, cobertura das reads e anotações.

NUCmer Dotplot foi usado para alinhar os *contigs* da montagem bruta LVH60a-c1 ao genoma de *C. fasciculata*, a partir de um método clássico para visualizar alinhamentos de dois genomas, levando em consideração características importantes de ambos. Os arquivos *input* para

gerar o gráfico são arquivos coords e coords.idx. Com o Dotplot, foi possível identificar regiões sintênicas e fazer anotações genômicas estruturais. Por exemplo, o *contig* renomeado *fig.38* é sintênico ao cromossomo 30 do genoma de *C. fasciculata* (Figura 20).

A análise do NUCmer indicou que três *contigs* não se alinham ao genoma de *C. fasciculata* e existem trechos do genoma de *C. fasciculata* que não se alinham no genoma *Crithidia-like*. A porcentagem de identidade foi calculada usando o tamanho total de nucleotídeos *Crithidia-like* que se alinham em *C. fasciculata* resultou em 92% de identidade entre *Crithidia-like* e a referência (Figura 21), sugerindo que ambos genomas compartilham similaridades, mas não correspondem a mesma espécie. Para ser considerado organismos de mesma espécie, os genomas deveriam demonstrar pelo menos 95% de identidade, como observado em alinhamentos anteriores (Tabela 7). Com isso, se faz necessária a investigação de sequências únicas da amostra *Crithidia-like* para melhor compreensão de seu genoma.

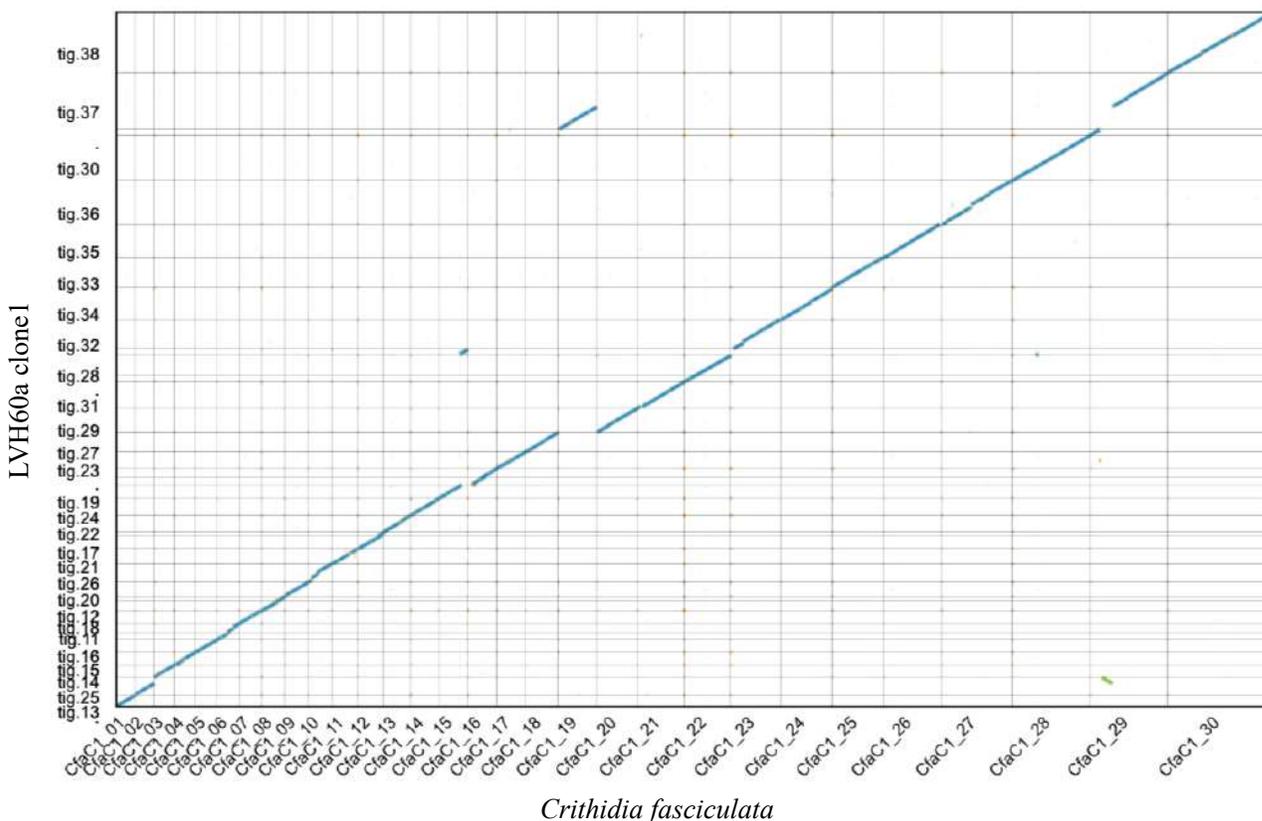


FIGURA 20 - Alinhamento genômico dos *contigs* filtrados do clone *Crithidia-like* LVH60a-C1 com a referência *C. fasciculata*. A referência foi obtida no banco de dados TrytripDB e o alinhamento foi produzido por NUCmer Dotplot.

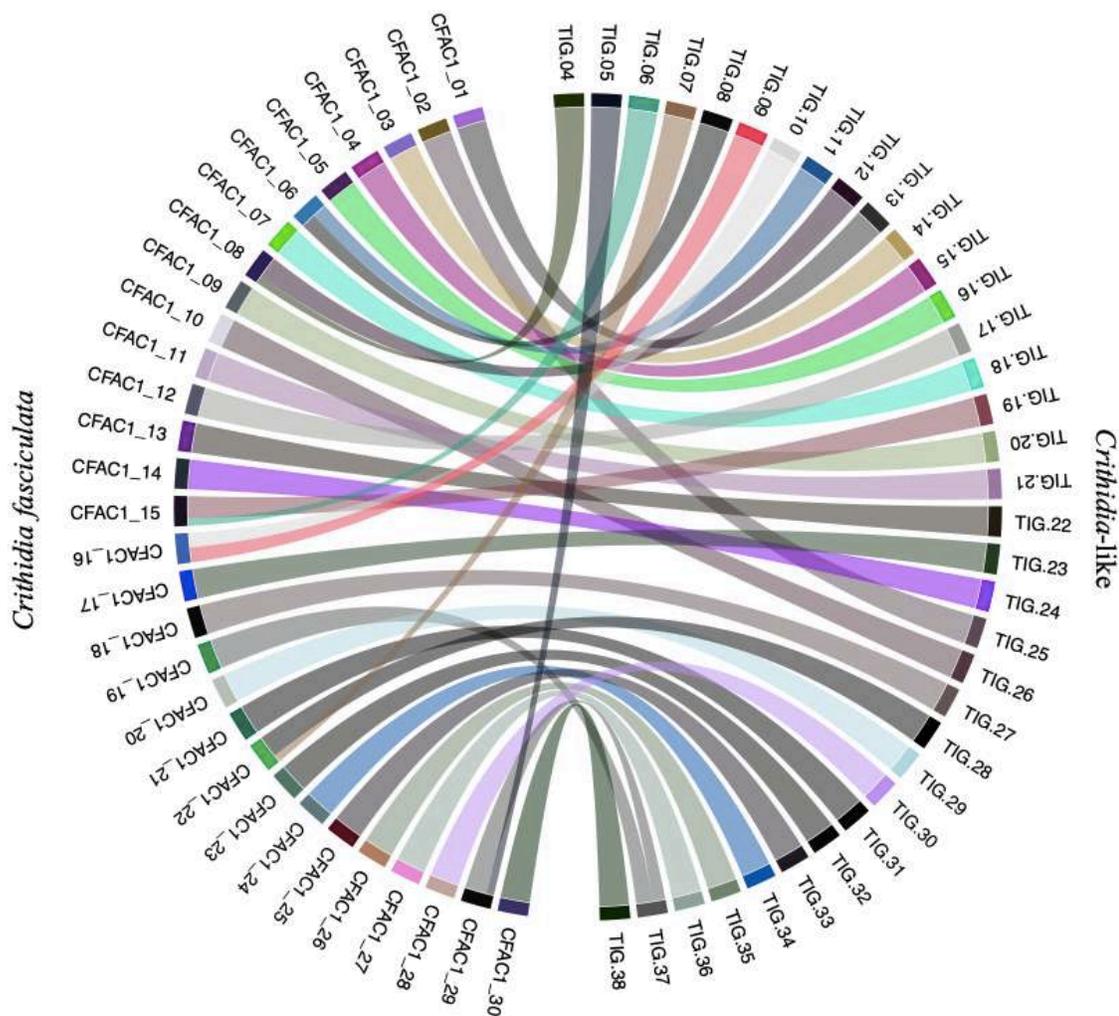


FIGURA 21 - Comparação dos genomas *Crithidia fasciculata* e *Crithidia*-like em representação por circosplot. Figura gerada por Circlize em pacote R Por Magalhães, R.D.M.

Após os processos de melhoramento da montagem genômica com os *softwares* *Medaka* (OXFORD NANOPORE TECHNOLOGY, 2018) e *Pilon* (WALKER et al., 2014) foram identificados 38 *contigs* únicos que variam de 12.018 pb a 2.988.870 pb resultando no tamanho de 34,4 Mb do genoma completo.

Os *contigs* do genoma *Crithidia*-like foram renomeados por ordem crescente em que o *contig* 1 e *contig* 38 fossem o menor e o maior, respectivamente (Tabela 11). Além disso, as primeiras anotações estruturais foram realizadas manualmente através da sintenia observada entre o genoma *Crithidia*-like e a referência *C. fasciculata*.

TABELA 11 - Características do genoma *Crithidia*-like em comparação a referência *C. fasciculata*

<b>Nome original do contig</b>	<b>Contig renomeado</b>	<b>Comprimento do contig</b>	<b>Número de CDS</b>	<b>Sintenia com <i>C. fasciculata</i></b>
tig00000019-1	tig.1	12.018	3	-
tig00000071-1	tig.2	19.095	1	-
tig00000118	tig.3	179.582	49	1/4 CFAC1_12
tig00000123	tig.4	212.911	62	1/4 CFAC1_08
tig00000138	tig.5	307.358	116	1/4 CFAC1_29
tig00000057	tig.6	318.858	97	1/4 CFAC1_15
tig00000189	tig.7	320.550	108	1/4 CFAC1_22
tig00000070-1	tig.8	333.127	104	1/2 CFAC1_6
tig00000193	tig.9	397.493	107	1/2 CFAC1_16
tig00000194	tig.10	424.022	132	1/2 CFAC1_16
tig00000070-0	tig.11	454.187	93	1/2 CFAC1_6
tig00000114	tig.12*	479.386	122	3/4 CFAC1_08
tig00000130	tig.13	522.973	174	CFAC1_01
tig00000109	tig.14	609.261	215	CFAC1_03
tig00000055	tig.15	623.946	212	CFAC1_04
tig00002745	tig.16	624.211	207	CFAC1_05
tig00000117	tig.17	642.655	192	3/4 CFAC1_12
tig00000089	tig.18	654.970	200	CFAC1_07
tig00000095	tig.19	658.402	198	3/4 CFAC1_15
tig00000087	tig.20	710.710	238	CFAC1_09
tig00002743	tig.21	751.230	227	CFAC1_11
tig00000071-0	tig.22	826.861	303	CFAC1_13
tig00000051	tig.23	829.291	294	CFAC1_17
tig00000091	tig.24	834.559	256	CFAC1_14
tig00002739	tig.25	904.112	263	CFAC1_02
tig00002740	tig.26	918.833	335	CFAC1_10
tig00000052	tig.27	971.786	294	CFAC1_18
tig00000191	tig.28	1.000.058	343	3/4 CFAC1_22

tig00002736	tig.29	1.194.408	371	CFAC1_20
tig00002732	tig.30	2.239.518	815	CFAC1_28
tig00002735	tig.31	1.292.868	399	CFAC1_21
tig00000023	tig.32	1.440.346	457	CFAC1_23
tig00000043	tig.33	1.457.499	498	CFAC1_25
tig00000026	tig.34	1.613.826	508	CFAC1_24
tig00000019-0	tig.35	1.645.447	617	CFAC1_26
tig00000009	tig.36*	2.180.050	658	CFAC1_27
tig00000007	tig.37	2.798.163	931	CFAC1_19 + 3/4 CFAC1_29
tig00000001	tig.38	2.988.870	1095	CFAC1_30

CFAC1 corresponde ao código do cromossomo de *C. fasciculata*. *Contigs* trissômicos\*

Com base na análise estrutural pelo alinhamento gerado por *NUCmer Dotplot* do genoma *Crithidia*-like, foi realizada anotação de sequências sintênicas compartilhadas com a referência *C. fasciculata*. O resultado mostrou 63 % dos *contigs* de *Crithidia*-like são completamente sintênicos a um cromossomo da referência *Crithidia fasciculata*, como por exemplo, os *contigs* tig.13, tig.14, tig.15, tig.16, tig.18, tig.20, tig.21, tig.22, tig.23, tig.24, tig.25, tig.26, tig.27, tig.29, tig.30, tig.31, tig.32, tig.33, tig.34, tig.35, tig.36 e tig.38 (Figura 22). Todavia, o *contig* tig.37 foi gerado pela combinação de mais de um cromossomo de *C. fasciculata*, os cromossomos 19 e parte do cromossomo 29 (Figura 23). Outro evento observado foi a fragmentação de um cromossomo da referência originando mais de um *contig*, como o cromossomo 22 de *C. fasciculata* alinhou-se ao tig.7 e tig.28, o cromossomo 6 da referência aos tig.8 e tig.11, o cromossomo 8 aos os tig.4 e tig.12, o cromossomo 16 aos tig.9 e tig.10, o cromossomo 12 aos tig.3 e tig.17 e o cromossomo 29 alinhou-se aos tig.5 e tig.37 (Figura 24). Considerando que os *contigs* *Crithidia*-like sejam bastante parecidos aos cromossomos da referência, possivelmente esses oito *contigs* a mais do genoma *Crithidia*-like em relação aos 30 cromossomos de *C. fasciculata* poderiam ser explicados por eventos de rearranjo cromossômico observados aqui. Para validar a ocorrência de rearranjo nos cromossomos de *Crithidia*-like e excluir a hipótese de ser artefato da montagem, será necessário realizar o alinhamento das *reads* brutas do sequenciamento ONT e comprovar a existência dessas sequências experimentalmente.

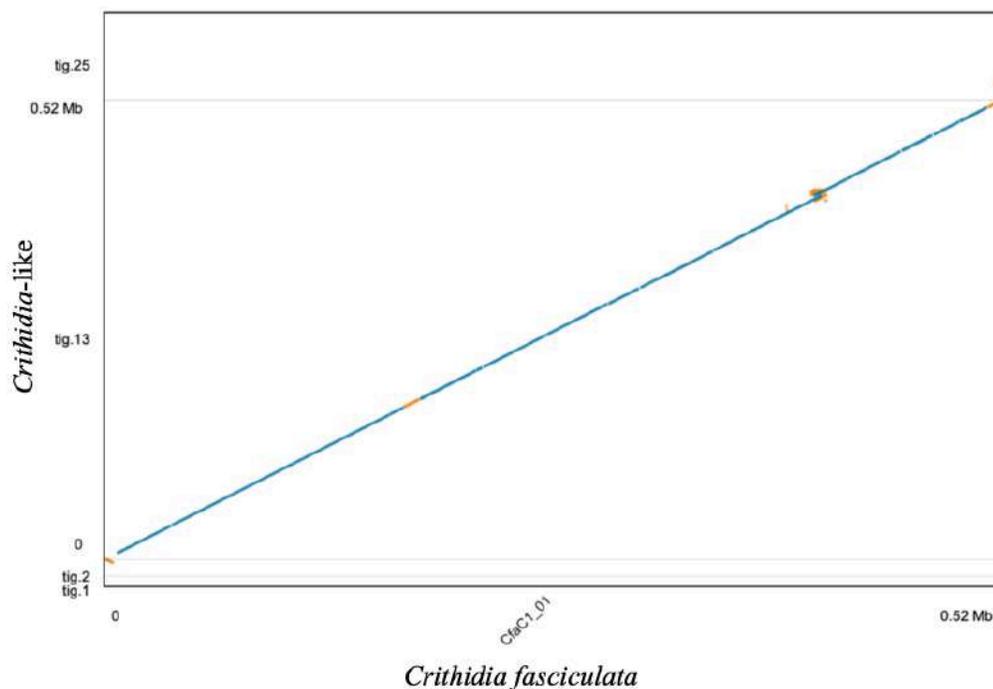


FIGURA 22 - Alinhamento completo do *fig. 13* do clone *Crithidia*-like LVH60a-C1 com o cromossomo 1 da referência *C. fasciculata*. A referência foi obtida no banco de dados TrytripDB e o alinhamento foi produzido por NUCmer Dotplot.

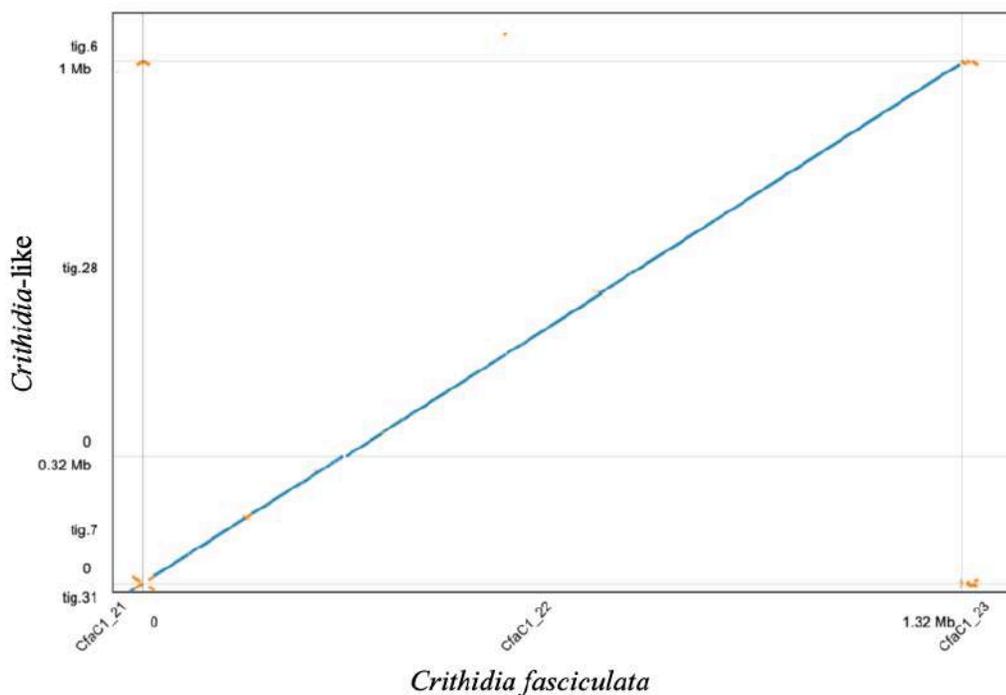


FIGURA 23 - Alinhamento dos *contigs* *fig. 7* e *fig. 28* do clone *Crithidia*-like LVH60a-C1 com o cromossomo 22 da referência *C. fasciculata*. A referência foi obtida no banco de dados TrytripDB e o alinhamento foi produzido por NUCmer Dotplot.

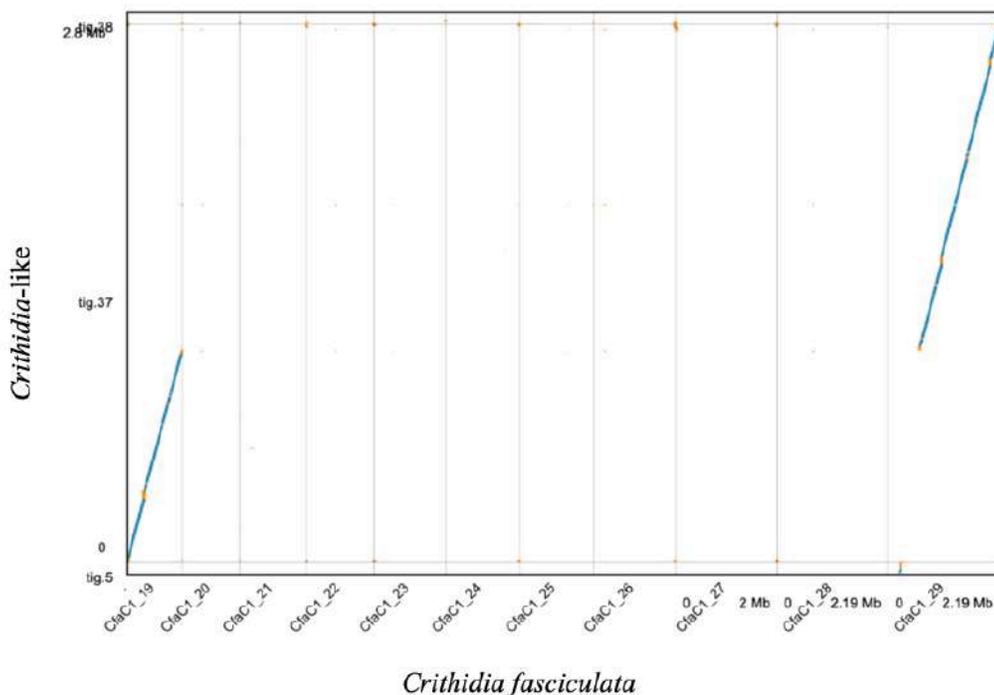


FIGURA 24 - Alinhamento do *contig* tig. 37 do clone *Crithidia*-like LVH60a-C1 com os cromossomos 19 e 29 da referência *C. fasciculata*. A referência foi obtida no banco de dados TrytripDB e o alinhamento foi produzido por NUCmer Dotplot.

Além dos *contigs* sintênicos aos cromossomos da referência, foram observadas sequências que não se alinham em nenhuma região da referência *C. fasciculata*, como ocorreu com os *contigs* tig.1 e tig.2. Analisando esses *contigs* pela ferramenta *blastn* encontramos uma região de 5000 pb dos *contigs* tig.1 e tig.2 similares a um gene de proteína hipotética de *Entamoeba dispar* de 1469 pb e aparentemente, existe cópias do mesmo gene sobrepostas em diferentes regiões dessas sequências. Interessantemente, esse gene é desconhecido em outros organismos Tripanosomatídeos, sugerindo possivelmente a troca de material genético entre *Entamoeba* e *Crithidia*-like em algum momento evolutivo. Também foi verificado nos tig.1 e tig.2 uma pequena porção similar a transposons de *L. braziliensis*.

O *contig* tig.3 possui 49 CDSs e pela ferramenta *blastx*, foi buscado por similaridades no banco de dados de proteínas do NCBI, submetendo as sequências individualmente para análise. A maioria das sequências indicou identidade com proteínas hipotéticas de tripanossomatídeos, e.g. *Leptomonas*, *Leishmania* e *Trypanosoma*. Curiosamente, quatro sequências codificadoras de proteínas não se alinham com nenhum organismo, sugerindo ser proteínas inéditas, que precisam ser estudadas para avaliar se estão associadas à vantagem.

Ao contrário de *Crithidia fasciculata* que possui 9.489 sequências codificadoras de proteína (CDS), a extração das CDSs do genoma *Crithidia*-like resultou em 11.294 sequências codificantes (Tabela 11) revelando mais de 1800 sequências proteínas extras. O número de CDS representa 17.437.822 pb, ou seja, 50,7% do genoma *Crithidia*-like.

Todas as etapas de montagem e melhoramento do genoma LVH60a-C1 foram comparadas (Tabela 12). A montagem bruta gerada com Canu apresentou 97 *contigs*, enquanto a montagem final polida apresentou 38 *contigs*. Também podemos identificar um aumento no N50 de 955 pb para 1.292 pb comparando montagem bruta e montagem polida, respectivamente. Às vezes, Medaka divide *contigs* se encontrar regiões que não são bem cobertas por leituras, explicando o número de *contigs* subindo (35 a 38 sequências de montagem limpa para montagem polida) e o comprimento mínimo diminui (177.023 Kb para 12.008 Kb). A montagem final produziu um genoma de 34,4 Mb com 38 *contigs* únicos, variando de 12.018 pb a 2.988.870 pb.

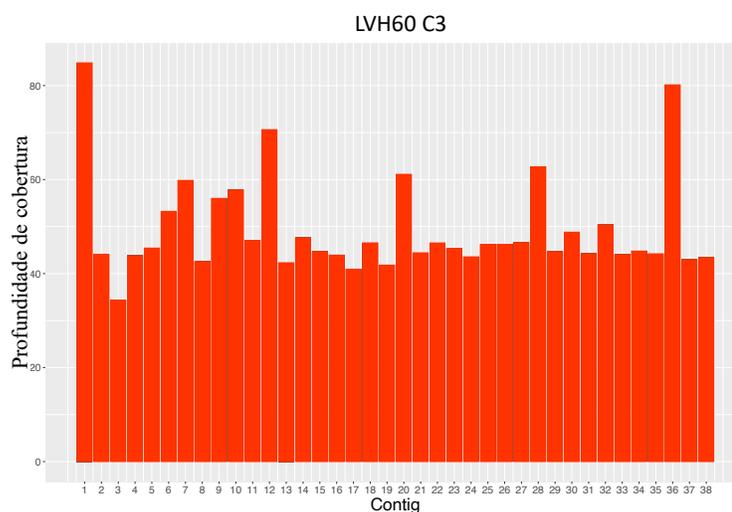
TABELA 12 - Evolução dos processos de montagem e melhoramento do genoma *Crithidia*-like sequenciado por *Oxford Nanopore Technology*

Etapas	Número de sequências	Comprimento total (pb)	Comprimento mínimo (pb)	Comprimento máximo (pb)	N50
Montagem do genoma (Canu)	97	40.175.452	1.298	2.936.253	955.566
Filtragem do genoma (Tapestry)	35	33.822.206	177.023	2.936.253	1.260.968
Melhoramento do genoma (Medaka)	38	34.473.237	12.008	2.989.743	1.293.754
Melhoramento do genoma (Pilon1)	38	34.408.674	12.015	2.989.152	1.292.924
Melhoramento do genoma (Pilon2)	38	34.397.382	12.017	2.988.899	1.292.908
<b>Melhoramento do genoma (Pilon3)</b>	<b>38</b>	<b>34.393.440</b>	<b>12,018</b>	<b>2.988.870</b>	<b>1.292.868</b>

Com isso, a montagem realizada a partir do sequenciamento *Oxford Nanopore* gerou sequências genômicas mais longas e menos fragmentadas para ser usada como genoma referência de organismos *Crithidia*-like afim de substituir a montagem Illumina do sequenciamento anterior (MARUYAMA et al., 2019).

#### 4.5.3 Análise de clones de LVH60 e LVH60a sequenciados por Illumina

O mapeamento de *reads* de amostras clonais de sequenciamento Illumina foi executado com as cepas LVH60 e LVH60a utilizando a montagem do genoma *Crithidia*-like de 38 *contigs* únicos como referência (Tabela 11). A análise de profundidade de cobertura mostrou que o resultado foi relativamente uniforme em relação ao genoma referência *Crithidia*-like. para a maioria dos *contigs*, exceto para os *contigs* 1, 12 e 36 (Figura 25). De forma geral, a cobertura média variou de 10X a 40X dentre as amostras, apesar de exibir um padrão de cobertura entre as amostras por *contig* sobre o genoma referência.



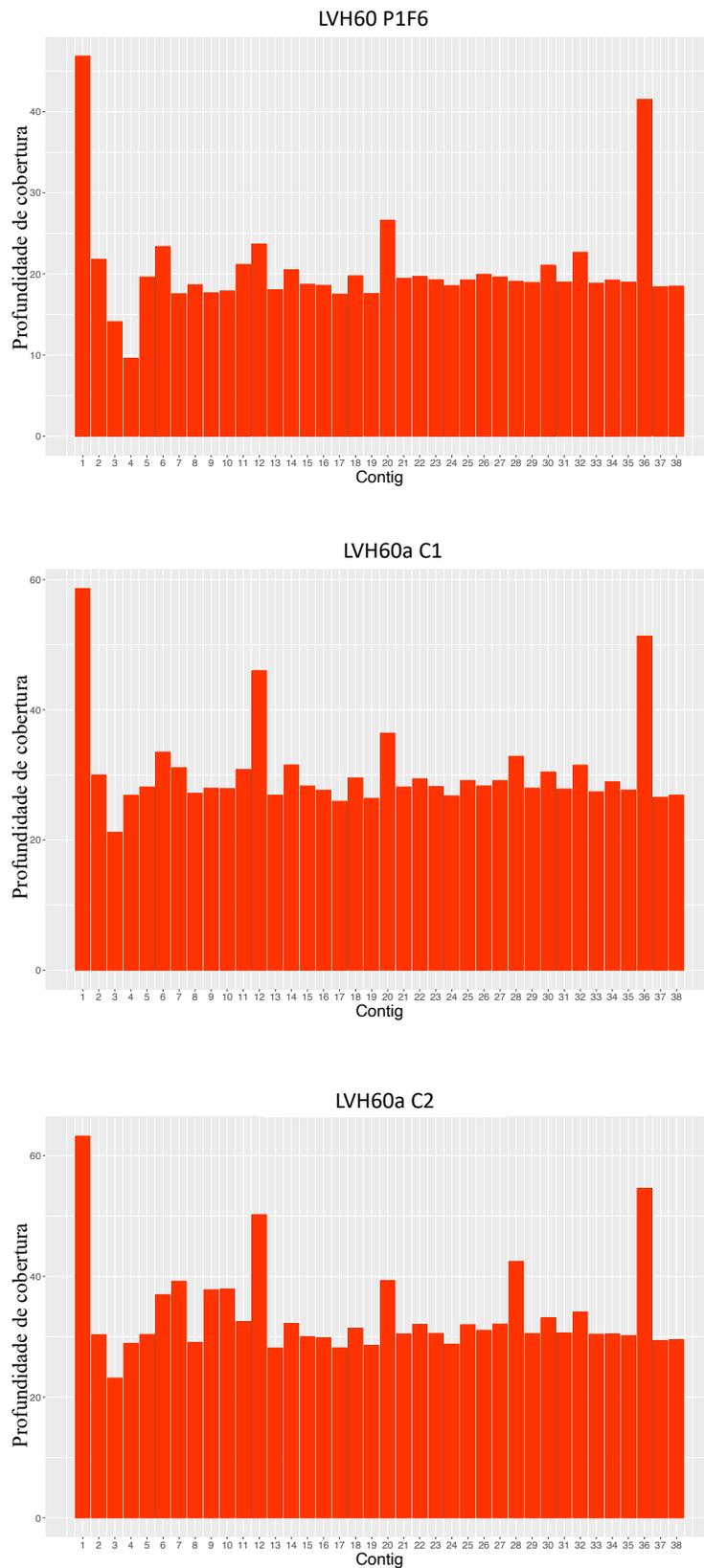
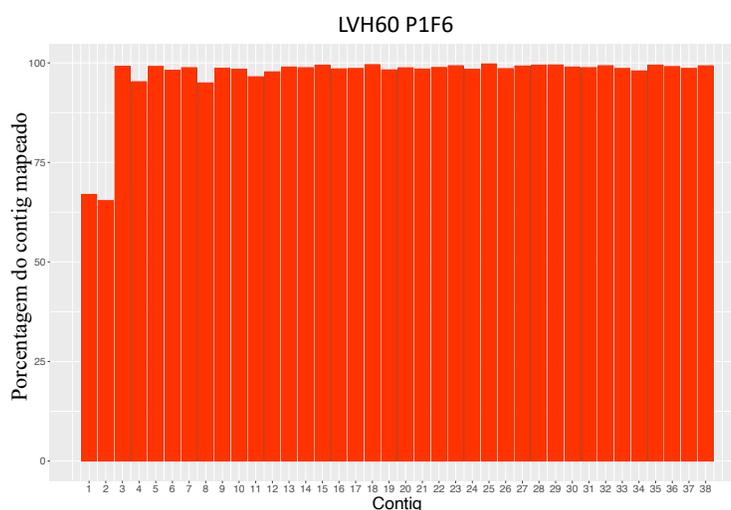
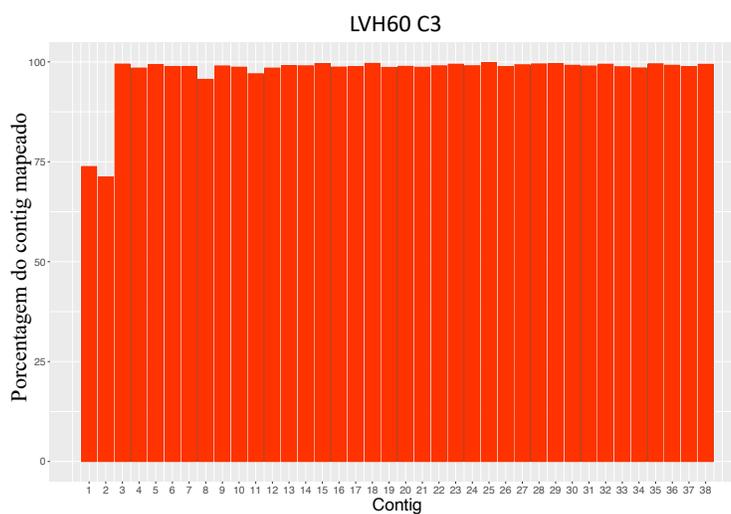


FIGURA 25 - Profundidade de cobertura de *reads* Illumina de amostras clonais de LVH60 e LVH60a mapeadas em genoma referência *Crithidia*-like.

O mapeamento referente a porcentagem total de bases de cada amostra que se alinharam no genoma referência *Crithidia*-like também apresentou uniformidade entre as amostras com média de 99% das bases mapeadas por *contig* do genoma referência, exceto pelos *contigs* 1 e 2 que apresentaram quase 75% de bases mapeadas ao longo dessas regiões da referência (Figura 26). Como observado antes, os *contigs* 1 e 2 alinham com proteína hipotética de *Entamoeba*. Esses resultados indicam que, as amostras clonais e policlonais apresentadas são pertencentes ao parasito *Crithidia*-like (MARUYAMA et al., 2019).



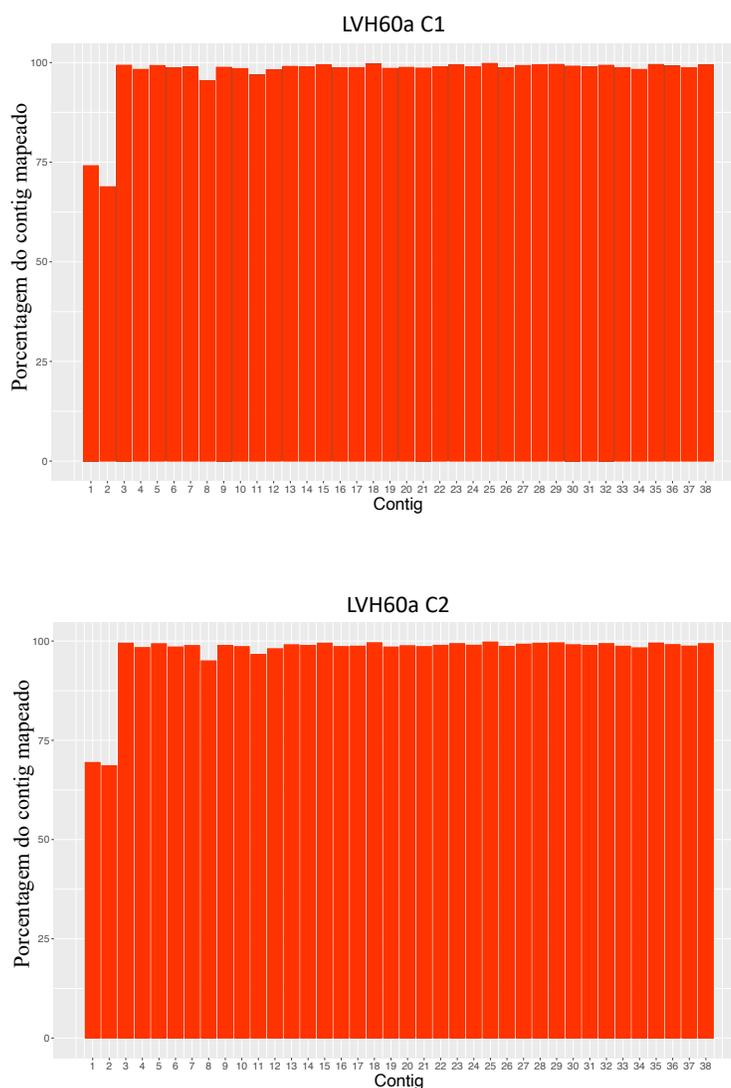


FIGURA 26 - Porcentagem dos *contigs* (cálculo pela extensão em pb) da referência *Crithidia-like* mapeadas por *reads* de amostras clonais de LVH60 e LVH60a.

Os *contigs* únicos foram investigados por BLASTN buscando por regiões similares e suportadas estatisticamente. Aproximadamente 5000 pb dos contigs *fig.1* e *fig.2* com 89.93% de identidade apresentaram similaridade com um gene de proteína hipotética de *Entamoeba dispar* de 1469 pb e aparentemente, possui cópias do mesmo gene sobrepostas; sugerindo que houve troca de material genético entre *Entamoeba* e *Crithidia-like* em algum momento evolutivo, já que esse gene é desconhecido em outros organismos tripanossomatídeos (Figuras 27 e 28). Além disso, uma outra pequena porção apresentou similaridade com um *transposon* de *L. braziliensis*. Essas regiões

podem representar artefato de sequenciamento ou regiões inéditas e precisam ser melhor compreendidas.

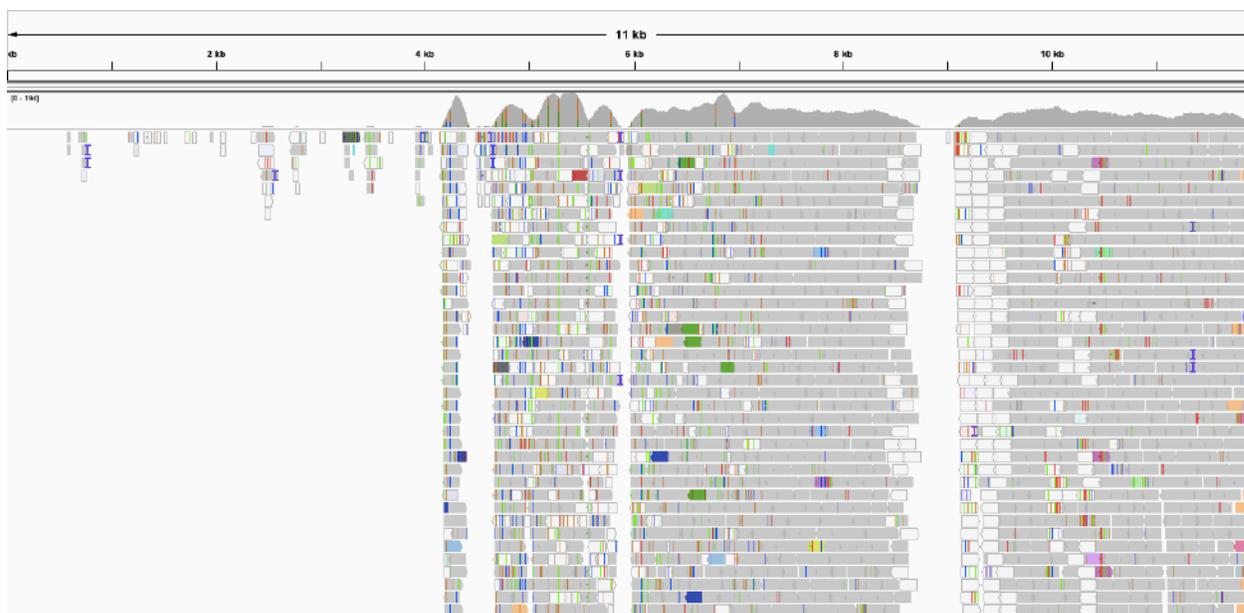


FIGURA 27 - Alinhamento do *contig* tig.1 de comprimento 12 Kb de *Crithidia*-like (LVH60a-C1) com as *reads* do isolado LVH60a-C1. O alinhamento foi visualizado pelo programa IGV. Os primeiros 4 Kb não possuem alinhamento.

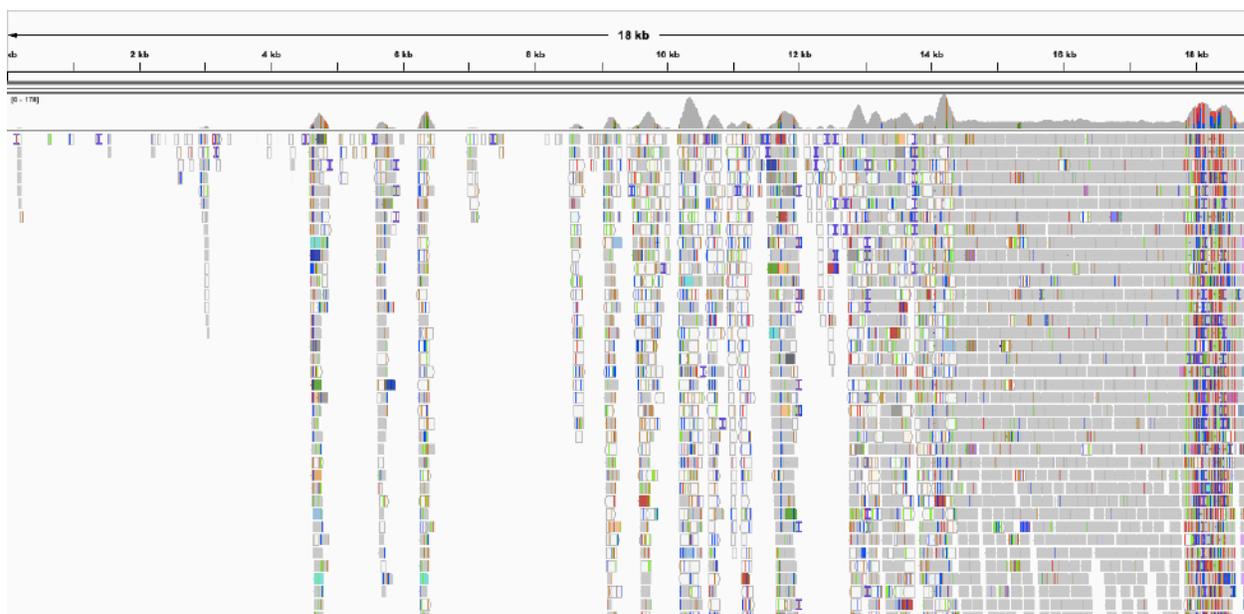
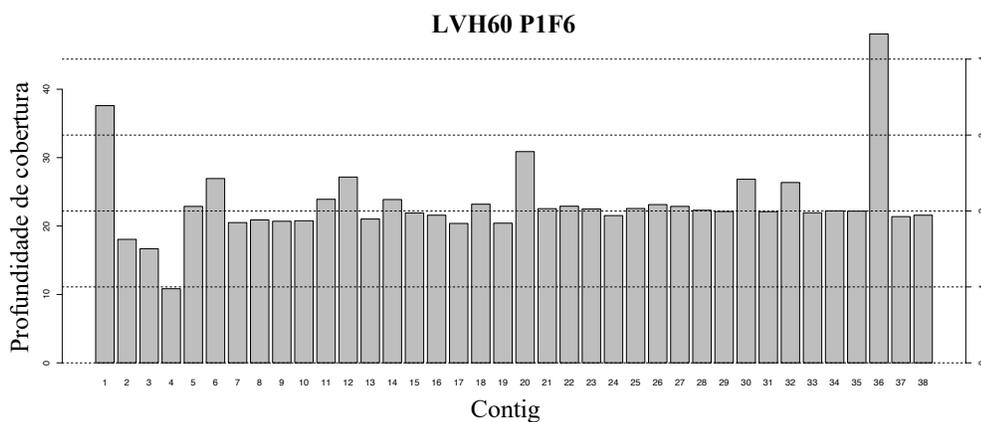
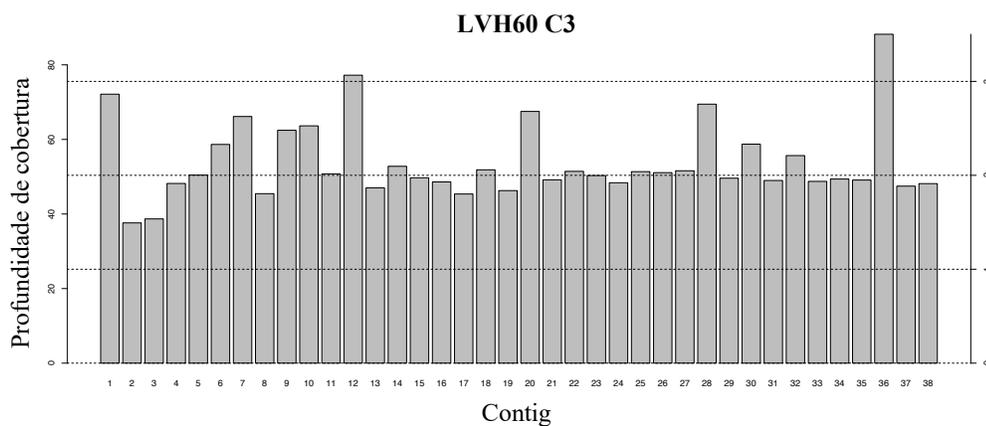


FIGURA 28 - Alinhamento do *contig* tig.2 de comprimento 19 Kb de *Crithidia*-like (LVH60a-C1) com as *reads* do isolado LVH60a-C1. O alinhamento foi visualizado pelo programa IGV. Os primeiros 4 Kb não possuem alinhamento.

#### 4.5.3.1 Estimativa de Ploidia

A ploidia foi estimada levando-se em conta os valores relativos de cobertura, ou seja, a média de cobertura obtida para cada *contig* da referência *Crithidia*-like. As *reads* de cada amostra foram alinhadas na referência *Crithidia*-like usando o *software* BWA. Em seguida, as sequências foram ordenadas com a remoção de *reads* duplicadas e indexadas para a obtenção dos dados em formato BAM, utilizando o SAMtools. Os arquivos gerados em formato BED (Browser Extendable Data) permite acessar com flexibilidade dados a serem exibidos na anotação.

O *output* é um arquivo em formato de texto indicando o total de bases das *reads* que se alinham em cada região genômica especificada e fornecida no arquivo BED. Sendo assim, o *output* representou a soma de profundidade por base de cada *read* alinhada e quanto maior o *contig*, maior será essa cobertura (extensão). Por isso, ao calcular a ploidia, assumimos que a maioria dos *contigs* eram diplóides (Figura 29).



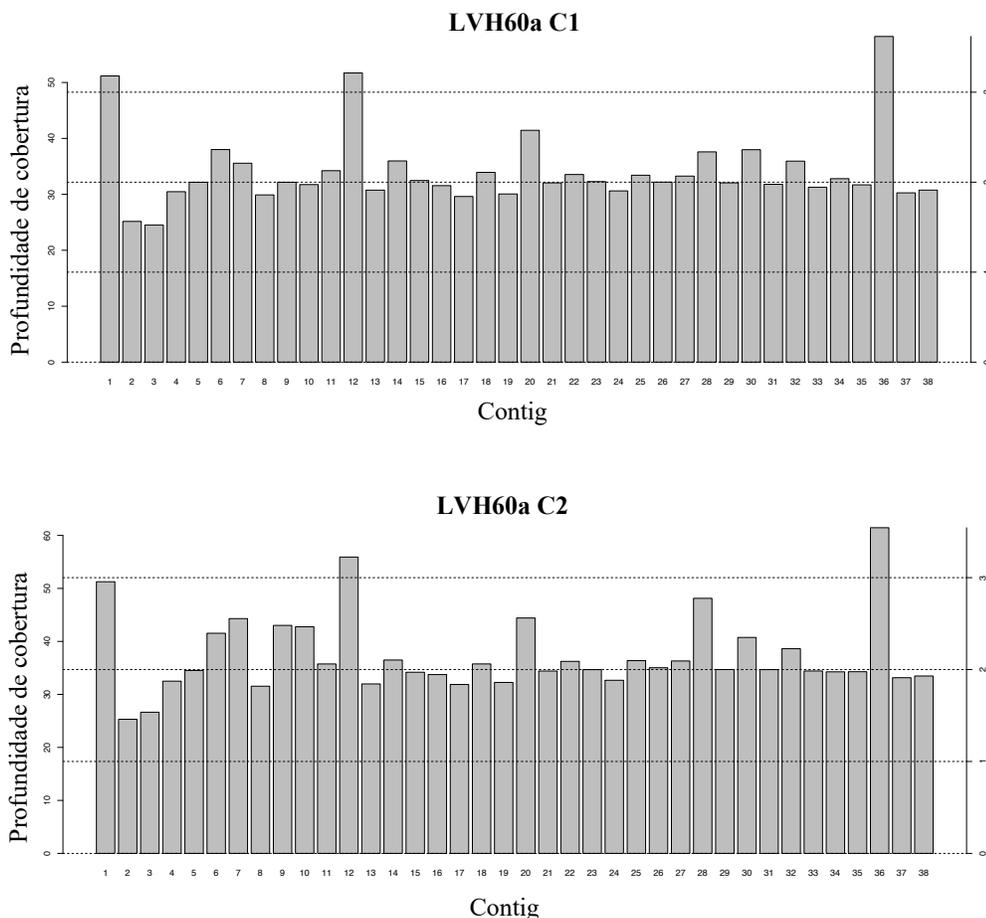


FIGURA 29 - Plodia estimada de amostras clonais de LVH60 e LVH60a em relação aos 38 *contigs* da referência *Crithidia*-like. O eixo da direita representa a ploidia, sendo 1. haplóide 2. diplóide 3. triplóide e 4. Tetraplóide.

A análise de estimativa de ploidia das amostras clonais *Crithidia*-like evidenciaram um grau de aneuploidia. Considerando que a referência *Crithidia*-like seja diplóide, os *contigs* 1, 12 e 36 de todas as amostras apresentaram aneuploidia, demonstrando a ocorrência de três cópias para esses *contigs*. Interessantemente, a amostra LVH60 P1F6 exibiu aneuploidia apenas nos *contigs* 1 e 36. Análises do genoma referência *Crithidia*-like apresentou aneuploidia apenas nos *contigs* 12 e 36 (Tabela 11). Aneuploidia é um fenômeno ainda não descrito em tripanossomatídeos monoxênicos. Esse tipo de variação no número de cópias cromossômicas pode estar relacionada com o tropismo do parasito nos diferentes tecidos do hospedeiro para garantir sua sobrevivência e capacidade de infecção (STERKERS et al., 2012).

***DISCUSSÃO***

## 5 DISCUSSÃO

Os protozoários da família Trypanosomatidae agrupam uma diversidade de organismos exclusivamente parasitários (WHEELER; GLUENZ; GULL, 2013) que possuem características como flagelo único e uma estrutura que compila o DNA mitocondrial em maxicírculos e minicírculos, o cinetoplasto (MASLOV; PODLIPAEV; LUKES, 2001; RITTIG; BOGDAN, 2000). O estudo do ciclo de vida de *Leishmania* é indispensável para o maior entendimento sobre a biologia do parasito e a capacidade de sobrevivência em diferentes hospedeiros, já que como estratégia de vida é classificada como organismo dixênico e apresenta duas morfologias principais para completar seu ciclo de vida: amastigota intracelular, no hospedeiro vertebrado e promastigota extracelular, no vetor flebotomíneo (WHEELER; GLUENZ; GULL, 2013). Além das espécies de tripanossomatídeos patogênicos ao homem, há um grupo de organismos classificados como monoxênicos que infectam exclusivamente insetos (LUKEŠ et al., 2014) e podem apresentar polimorfismo, como ocorre em *Crithidia fasciculata*, com fases coanomastigota e amastigota na inseto vetor (ALCOLEA et al., 2014; ISHEMGULOVA et al., 2017).

Considerando que os isolados clínicos deste trabalho eram provenientes de pacientes diagnosticados com leishmaniose visceral, era esperado encontrar parasitos *Leishmania*. Entretanto, a melhor caracterização biológica e as diversas análises genômica corroboraram que esse parasito é mais semelhante a espécie monoxênica *C. fasciculata*, e por conta disso, foi designado provisoriamente de parasito *Crithidia-like*, até que estudos adicionais sejam realizados para confirmar como sendo uma nova espécie. Os isolados clínicos deste trabalho tiveram origem da medula óssea (LVH60) e lesões cutâneas (LVH60a) de caso atípico e letal de leishmaniose visceral.

As linhagens clonais e policlonais utilizadas no experimentos de crescimento sob diferentes temperaturas foram exemplares de *Leishmania infantum* HUUF14 e MHOM/BR/74/PP75 e, exemplares de *Crithidia-like* LVH60, LVH60a, LVH60 C3 e LVH60 P1E6. A verificação da viabilidade e crescimento celular de parasitos foi realizada a partir da contagem celular diária por nove dias em temperaturas diferentes, já que os parasitas da leishmaniose visceral necessitam sobreviver e proliferar em dois ambientes distintos, o intestino dos flebótomos (25 °C) e macrófagos dos mamíferos (35 °C). A mudança de temperatura durante a transmissão do inseto

vetor para mamíferos é importante para a continuidade do ciclo de vida do parasito e para a síntese de proteínas *heat shock* (choque térmico), que estão relacionadas a sobrevivência em temperaturas mais elevadas (HOMBACH et al., 2014). Por conta disso, a termo tolerância é um fator fundamental para o parasitismo, principalmente em hospedeiros de sangue quente, como os seres humanos. O crescimento celular de promastigotas em meio de cultura é caracterizado pelas fases mid-logarítmica, logarítmica e estacionária além de variar entre as espécies de tripanossomatídeos (LEMESRE et al., 1988). *Leishmania infantum* tem uma taxa de multiplicação celular superior quando comparada a outras espécies do gênero, como *L. braziliensis* (LEMESRE et al., 1988).

O resultado do crescimento celular sob temperatura de 25 °C, que representa a temperatura do inseto vetor, apresentou variação estatística entre os dias de contagem, mas não houve diferença entre as linhagens, exceto por HUUF14 ( $p < 0,0001$ ). Todas as linhagens apresentaram proliferação celular, iniciando com  $2 \times 10^5$  células até atingirem a fase estacionária no 6º dia com  $5 \times 10^8$  células. Parasitos causadores da leishmaniose visceral, como *L. infantum* e *L. chagasi*, atingem o platô por volta do 5º dia de contagem, enquanto que espécies relacionadas a leishmaniose cutânea passam para a fase estacionária próximo ao 7º dia (LEMESRE et al., 1988). Os resultados do experimento de crescimento celular a 25 °C apontaram que as linhagens *Crithidia*-like, com base nas linhagens de controle positivo para *L. infantum*, possuem taxa de crescimento igual a PP75 e maior que HUUF14. Dados não publicados também apontaram que parasitos *Crithidia*-like apresentavam maior crescimento em relação a HUUF14 e isso pode ser devido a linhagem de *L. infantum* usada, já que dentro da mesma espécie é comum a ocorrência de variações na taxa de crescimento.

Em contrapartida, a proliferação celular resultante do experimento que simula o hospedeiro mamífero (35 °C) mostrou que não houve diferença estatística entre as amostras, apenas entre os dias de crescimento dentro de cada amostra, exceto pela linhagem PP75 de *L. infantum*. Todas as linhagens proliferaram apenas entre o dia da inoculação iniciado com  $2 \times 10^5$  células até o 3º dia de contagem, atingindo a fase estacionária com  $5 \times 10^6$  células. Já as células de PP75, atingiram o platô no 4º dia de contagem com  $6 \times 10^7$  células. As células cultivadas a 35 °C apresentaram corpo celular arredondado e pequeno, com um flagelo curto de pouca movimentação apresentando grande similaridade à forma amastigota axênica (ALCOLEA et al., 2014). Contudo, a forma amastigota em leishmaniose visceral corresponde a forma intracelular, sem motilidade que sobrevive e se replica obrigatoriamente dentro dos macrófagos do hospedeiro (RODRIGUES et al., 2016). Para verificar a viabilidade de infecção, faz-se necessário experimento de infecção in vitro com essas

linhagens. Além disso, estudos comprovaram que organismos dos gêneros *Crithidia* são capazes de suportar temperaturas elevadas e isso pode ser considerado uma pré-adaptação ao ciclo de vida dixênico dos parasitos que sobrevivem em ambiente hostil do hospedeiro humano (FILOSA et al., 2019). Isso explica a tolerância térmica e sobrevivência das linhagens *Crithidia*-like diante a temperaturas elevadas que simulam a temperatura do corpo do vertebrado mamífero.

No que diz respeito a caracterização morfológica, observamos diferença na forma celular e tamanho do flagelo dos isolados clínicos e isso poderia estar relacionado a diferentes espécies de parasito causando leishmaniose visceral. Apesar de compartilhar estruturas citoesqueléticas e membranas conservadas, a estrutura corporal dos protozoários tripanossomatídeos diverge tanto nos estágios do ciclo de vida do parasito, quanto dentro de cada espécie (WHEELER; GLUENZ; GULL, 2013).

A diversidade morfológica desses parasitos foi examinada com o auxílio da microscopia de luz que possibilitou a descrição e classificação desses microrganismos de acordo com seis principais classes morfológicas classificadas com base na posição e profundidade da bolsa flagelar, comprimento do flagelo e fixação lateral do flagelo ao corpo da célula (HOARE; WALLACE, 1966; revisado por KAUFER et al., 2017; WHEELER; GLUENZ; GULL, 2013). Embora a função dessas formas celulares seja obscura, há exemplos que indicam a importância da morfogênese certa para a patogenicidade (JESUS et al., 1993; WHEELER; GLUENZ; GULL, 2013). Provavelmente, dois fenômenos biológicos são responsáveis para a diversidade morfológica dos estágios de vida: as limitações intrínsecas da própria organização celular podendo reprimir as possíveis formas celulares viáveis e, pressões seletivas do ambiente hospedeiro viabilizando algumas formas e outras não. As formas promastigotas, coanomastigotas e opistomastigotas ocorrem comumente no ambiente hospedeiro invertebrado, todavia há relatos de formas amastigotas em insetos.

Os isolados clínicos de *L. infantum* MHOM/BR/74/PP75 e HUUFS14 analisados neste trabalho apresentam células alongadas, com flagelos longos inseridos na região anterior do corpo celular e cinetoplasto visível na porção anterior do núcleo, característico da forma promastigota (WHEELER; GLUENZ; GULL, 2013). Já os isolados clínicos *Crithidia*-like apresentaram células com aparência arredondada, flagelos encurtados e cinetoplasto ao lado do núcleo, aproximando-se da forma coanomastigota. Os resultados aqui apresentados foram semelhantes aos estudos obtidos por Ghobakhloo e colaboradores (2019) que mostraram divergências morfológicas entre *L. major* e *Crithidia* spp. encontradas em isolados clínicos co-infectados no Iran. Além disso, foi revelado que uma pequena parcela de isolados clínicos apresentou apenas cepas de *Crithidia* spp. em

pacientes suspeitos com leishmaniose e a partir de ensaios de infecção *in vitro*, confirmou a capacidade de organismos monoxênicos como *Crithidia* infectar humanos (GHOBAKHLOO et al., 2019). Ainda que seja necessário realizar outras análises para fundamentar os aspectos morfológicos aqui descritos, já foi possível identificar pela observação morfológica diferenças entre os isolados clínicos obtidos de pacientes diagnosticados com leishmaniose visceral, mas que se divergem morfológicamente da linhagem conhecida de *L. infantum* (MHOM/BR/74/PP75).

Com o advento das tecnologias moleculares, a caracterização morfológica e ciclo de vida deixaram de ser as principais formas de definição do sistema taxonômico que por muito tempo eram usadas para estabelecer relação entre os organismos da família Trypanosomatidae (HOARE; WALLACE, 1966). Estudos recentes apontaram que a distinção de morfologias não é mais viável devido a vasta amplitude e variação morfológica dentro de uma mesma população ou mesmo um organismo diante de hospedeiros diferentes (VOTÝPKA et al., 2013, 2015). Por outro lado, técnicas moleculares como PCR seguida por sequenciamento são amplamente usadas para caracterização e genotipagem de parasitos tripanossomatídeos (MASLOV; PODLIPAEV; LUKES, 2001; YURCHENKO et al., 2006).

A árvore filogenética obtida com base na sequência SSU rRNA das amostras LVH, mostrou uma agregação das linhagens LVH com *C. fasciculata*, permanecendo distantes do clado *Leishmania*. A linhagem HUUFS14 foi empregada como controle positivo para *L. infantum* e foi posicionada junto ao gênero *Leishmania*, como previsto. O mesmo ocorreu com a linhagem LTCP393, controle positivo para *L. braziliensis*, situada ao lado da referência *L. braziliensis*.

Esses resultados mostraram que um organismo não *Leishmania* está relacionado com a ocorrência da leishmaniose visceral na região de Aracaju - Sergipe, Brasil. De maneira interessante, esse novo parasito identificado a partir das análises filogenéticas se agrupa com a espécie monoxênica *C. fasciculata* que não causa danos a saúde humana (ALCOLEA et al., 2014). Os isolados clínicos LVH09, LVH21, LVH23, LVH52, LVH54, LVH59, LVH101 e LVH135 são fortes candidatos para o próximo sequenciamento genômico de acordo com a triagem molecular e morfológica.

O avanço dos estudos baseados em análises genômicas promoveu o conhecimento do parasito a nível molecular (DOWNING et al., 2011) contribuindo com a investigação de características envolvidas com a patogenicidade das doenças (CANTACESSI et al., 2015) e buscando identificar mecanismos associados ao parasitismo e alvos genéticos relacionados as manifestações clínicas. A genômica comparativa realizada no subcontinente Asiático é bastante

avançada no que se trata de análises com *L. donovani*, causadora da leishmaniose visceral no Velho Mundo (DOWNING et al., 2011; FRANSSEN et al., 2020; IMAMURA et al., 2016). Algumas análises genômicas também foram conduzidas com LV no Brasil causada pelo patógeno *L. infantum* (CARNIELLI et al., 2018; TEIXEIRA et al., 2017), mostrando a necessidade de estudos que identifiquem adequadamente os parasitos responsáveis pela leishmaniose visceral.

Apesar da ocorrência de parasitos de insetos em humanos ser um evento improvável, há relatos desses tripanossomatídeos monoxênicos em cenários de co-infecção com outras espécies (SRIVASTAVA et al., 2010). Estudos elencaram casos de Trypanosomatidae em pacientes imunossuprimidos, como parasitos *Leishmania* (PUIG; PRADINAUD, 2003) e também parasitos monoxênicos como *Leptomonas* e *Crithidia* (CHICHARRO; ALVAR, 2003; GHOBAKHLOO et al., 2019). Casos de co-infecção também foram reportados entre espécies monoxênicas e pacientes com leishmaniose, como identificado no subcontinente Indiano *Leptomonas* spp. em isolados de *Leishmania donovani* (GHOSH et al., 2012; KAUFER et al., 2017; SRIVASTAVA et al., 2010) e *Crithidia* spp. com *Leishmania major* no Iran (GHOBAKHLOO et al., 2019).

A identidade das cepas LVH60 e LVH60a ainda não estão bem definidas. Na avaliação dos dados de sequenciamento de LVH60 e LVH60a policlonal, analisamos os dados de mapeamento e cobertura das reads dessas cepas com as referências concatenadas de *C. fasciculata* e *L. infantum*. A profundidade de cobertura das linhagens *Crithidia*-like apresentou valor elevado com a referência *C. fasciculata* e valor baixo com *L. infantum*, corroborando mais uma vez que essas amostras são mais parecidas ao parasito monoxênico. Como o sequenciamento dos isolados clínicos policlonais apresentaram rendimento inicial de sequenciamento elevado, é provável que a amostra LVH60, oriunda da medula óssea, estivesse com DNA resquicial de *L. infantum*, indicando possível co-infecção. Um fato interessante é que o paciente dos isolados LVH60 e LVH60a era HIV negativo e a maioria dos casos que foram encontrados parasitos de insetos em humanos eram pacientes HIV positivos (CHICHARRO; ALVAR, 2003), i.e., que apresentavam sistema imunológico suprimido.

Estudos de Ghobakhloo e colaboradores (2019) revelaram que somente cepas de *Crithidia* spp. foram isoladas de pacientes diagnosticados com leishmaniose (sem a presença de *Leishmania*), sugerindo que organismos monoxênicos como *Crithidia* são capazes de infectar humanos. Para validarmos o parasito *Crithidia*-like foi necessário obter um genoma mais refinado e representativo para realizar análises de genômica estrutural e compreender o conteúdo do genoma desse potencial nova parasito. A partir do sequenciamento realizado pela tecnologia ONT obtivemos um genoma

final de 34,4 Mb com 38 *contigs* (pseudocromossomos) e mais de 11 mil sequências codificadoras de proteínas. O parasito *Crithidia*-like apresentou 92% de identidade com *C. fasciculata* e provou que não pertencem a mesma espécie. Análises de mapeamento genômico e profundidade de cobertura de isolados clonais originários de LVH60 e LVH60a mostraram que são pertencentes ao parasito *Crithidia*-like. Além do mais, o resultado de genômica comparativa mostrou que o parasito *Crithidia*-like apresentou regiões únicas de seu genoma que não se alinharam com nenhum cromossomo de *C. fasciculata*. As sequências inéditas necessitam ser experimentalmente corroboradas por amplificação de PCR e testadas em outros isolados clínicos *Crithidia*-like.

Coerente com esses resultados, também a estimativa de ploidia foi conferida nos isolados clonais de *Crithidia*-like. A aneuploidia é um fenômeno habitualmente expressado como anomalias severas ou diminuição do fitness, todavia alguns organismos se beneficiam desse evento para adaptar-se rapidamente a ambientes em mudança, como os fungos patogênicos e células cancerígenas (MANNAERT et al., 2012). Estudos de genômica recentes relataram que parasitos *Leishmania* possuem variação no número de cópias cromossomais aparecendo como uma tendência para esse gênero (CARNIELLI et al., 2018; IMAMURA; DUJARDIN, 2019; TEIXEIRA et al., 2017; VALDIVIA et al., 2015). Contudo, esse fenômeno permanece desconhecido em outros gêneros da família Trypanosomatidae. A análise de estimativa de ploidia das amostras clonais *Crithidia*-like exibiram um grau de aneuploidia em três *contigs*, levando em conta que a referência *Crithidia*-like é diplóide. Essa tipo de variação no número de cópias cromossomais pode estar relacionada com a adaptação entre o ambiente dos diferentes hospedeiros e tropismo em diferentes tecidos do hospedeiro humano (tegumento e vísceras) para garantir sua sobrevivência e capacidade de infecção (STERKERS et al., 2012).

Portanto, nossos resultados e o conjunto de dados já publicados (MARUYAMA et al., 2019), sugerem que parasitos *Crithidia*-like estão relacionados à doença infecciosa grave parecida com leishmaniose visceral na região de Aracaju - Sergipe. Estudos ainda são necessários para verificar a taxa de co-infecção entre *Crithidia*-like e *L. infantum* afim de estabelecer um perfil de ocorrência. Também é necessário validar sequências codificadoras de proteínas inéditas de *Crithidia*-like para serem testadas e utilizadas como alvo de detecção do parasito e busca por suas funções, que eventualmente seriam relacionadas com a patogenicidade e sobrevivência no hospedeiro humano. Por fim, é necessário realizar sequenciamento genômico de parasitos obtido diretamente de tecido infectados, como efetuado com *L. donovani* na Índia (DOMAGALSKA et

al., 2019) como forma de minimizar qualquer viés relacionado ao cultivo laboratorial desses parasitos.

***CONCLUSÃO***

## 6 CONCLUSÃO

- Este trabalho forneceu um painel informativo para a seleção de amostras para sequenciamento genômico;
- A análise de crescimento celular verificou a capacidade de termo tolerância e proliferação das cepas *Crithidia*-like em temperatura que simula a do hospedeiro vertebrado, característica indispensável para a adaptação do parasitismo em seres humanos;
- A morfologia dos isolados clínicos *Crithidia*-like é divergente dos isolados clínicos referência de *Leishmania infantum*;
- A filogenia molecular da região SSU rRNA formou o clado do parasito *Crithidia*-like com a espécie monoxênica *Crithidia fasciculata*, que é diferente do clado *Leishmania*;
- Foi possível sistematizar a análise de dados genômicos para futuros estudos de Genômica Comparativa de isolados clínicos;
- A montagem genômica do parasito *Crithidia*-like obtido para um clone de LVH60a por Oxford Nanopore Technology resultou em 38 *contigs* (genoma de 34 Mb) que possivelmente representam o cariótipo deste novo parasito;
- Na montagem refinada, foi observado dois *contigs* inéditos que precisam ser investigados quanto à sua relação a emergência desse novo parasito.

## ***REFERÊNCIAS BIBLIOGRÁFICAS***

## 7 REFERÊNCIAS BIBLIOGRÁFICAS

- ALCOLEA, Pedro J. et al. An Insight into the Proteome of *Crithidia fasciculata* Choanomastigotes as a Comparative Approach to Axenic Growth, Peanut Lectin Agglutination and Differentiation of *Leishmania* spp. Promastigotes. **PLOS ONE**, [S. l.], v. 9, n. 12, p. e113837, 2014. Disponível em: <https://doi.org/10.1371/journal.pone.0113837>
- ANDREWS, Simon. **FastQC A Quality Control tool for High Throughput Sequence Data**. [s. l.], 2010. Disponível em: <https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>. Acesso em: 22 nov. 2018.
- ASLETT, Martin et al. TriTrypDB: a functional genomic resource for the Trypanosomatidae. **Nucleic Acids Research**, [S. l.], v. 38, n. Database issue, p. D457–D462, 2010. Disponível em: <https://doi.org/10.1093/nar/gkp851>
- BATISTA, Luís F. S. et al. Canine leishmaniasis: Genome-wide analysis and antibody response to *Lutzomyia longipalpis* saliva. **PLoS ONE**, [S. l.], v. 13, n. 5, 2018. Disponível em: <https://doi.org/10.1371/journal.pone.0197215>. Acesso em: 24 maio. 2020.
- BORGHESAN, Tarcilla C. et al. Molecular Phylogenetic Redefinition of *Herpetomonas* (Kinetoplastea, Trypanosomatidae), a Genus of Insect Parasites Associated with Flies. **Protist**, [S. l.], v. 164, n. 1, p. 129–152, 2013. Disponível em: <https://doi.org/10.1016/j.protis.2012.06.001>
- CANTACESSI, Cinzia et al. The past, present, and future of *Leishmania* genomics and transcriptomics. **Trends in Parasitology**, [S. l.], v. 31, n. 3, p. 100–108, 2015. Disponível em: <https://doi.org/10.1016/j.pt.2014.12.012>
- CARNIELLI, Juliana B. T. et al. A *Leishmania infantum* genetic marker associated with miltefosine treatment failure for visceral leishmaniasis. **EBioMedicine**, [S. l.], v. 36, p. 83–91, 2018. Disponível em: <https://doi.org/10.1016/j.ebiom.2018.09.029>
- CÁSSIA-PIRES, Renata de et al. Multiplex PCR as a tool for the diagnosis of *Leishmania* spp. kDNA and the *gapdh* housekeeping gene of mammal hosts. **PLOS ONE**, [S. l.], v. 12, n. 3, p. e0173922, 2017. Disponível em: <https://doi.org/10.1371/journal.pone.0173922>
- CAVALIER-SMITH, T. Kingdom protozoa and its 18 phyla. **Microbiology and Molecular Biology Reviews**, [S. l.], v. 57, n. 4, p. 953–994, 1993.
- CHAPPUIS, François et al. Visceral leishmaniasis: what are the needs for diagnosis, treatment and control? **Nature Reviews. Microbiology**, [S. l.], v. 5, n. 11, p. 873–882, 2007. Disponível em: <https://doi.org/10.1038/nrmicro1748>
- CHICHARRO, C.; ALVAR, J. Lower trypanosomatids in HIV/AIDS patients. **Annals of Tropical Medicine and Parasitology**, [S. l.], v. 97 Suppl 1, p. 75–78, 2003. Disponível em: <https://doi.org/10.1179/000349803225002552>
- CORLISS, J. O. An interim utilitarian ('User-friendly') hierarchical classification and characterization of the protists. **Acta Protozoologica**, [S. l.], 1994. Disponível em: <http://www.vliz.be/en/imis?module=ref&refid=67543&printversion=1&dropIMISitle=1>. Acesso em: 22 nov. 2018.
- DEANE, L. M.; DEANE, M. P. Visceral leishmaniasis in Brazil: geographical distribution and transmission. **Revista do Instituto de Medicina Tropical de Sao Paulo**, [S. l.], v. 4, n. 3, 1962. Disponível em: <https://www.cabdirect.org/cabdirect/abstract/19641000202>. Acesso em: 22 nov. 2018.
- DOMAGALSKA, Malgorzata A. et al. **Genomes of *Leishmania* parasites directly sequenced**

- from patients with visceral leishmaniasis in the Indian subcontinent.** [S. l.]: Microbiology, 2019. preprint. Disponível em: <https://doi.org/10.1101/676163>. Acesso em: 16 jan. 2020.
- DOWNING, Tim et al. Whole genome sequencing of multiple *Leishmania donovani* clinical isolates provides insights into population structure and mechanisms of drug resistance. **Genome Research**, [S. l.], v. 21, n. 12, p. 2143–2156, 2011. Disponível em: <https://doi.org/10.1101/gr.123430.111>
- EDGAR, Robert C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. **Nucleic Acids Research**, [S. l.], v. 32, n. 5, p. 1792–1797, 2004. Disponível em: <https://doi.org/10.1093/nar/gkh340>
- EWING, Brent et al. Base-Calling of Automated Sequencer Traces Using Phred. I. Accuracy Assessment. **Genome Research**, [S. l.], v. 8, n. 3, p. 175–185, 1998. Disponível em: <https://doi.org/10.1101/gr.8.3.175>
- FILOSA, John N. et al. Dramatic changes in gene expression in different forms of *Crithidia fasciculata* reveal potential mechanisms for insect-specific adhesion in kinetoplastid parasites. **PLOS Neglected Tropical Diseases**, [S. l.], v. 13, n. 7, p. e0007570, 2019. Disponível em: <https://doi.org/10.1371/journal.pntd.0007570>
- FRANSSSEN, Susanne U. et al. Global genome diversity of the *Leishmania donovani* complex. **eLife**, [S. l.], v. 9, 2020. Disponível em: <https://doi.org/10.7554/eLife.51243>. Acesso em: 5 abr. 2020.
- GHOBAKHLOO, Nafiseh et al. Isolation of *Crithidia* spp. from lesions of immunocompetent patients with suspected cutaneous leishmaniasis in Iran. **Tropical Medicine & International Health**, [S. l.], v. 24, n. 1, p. 116–126, 2019. Disponível em: <https://doi.org/10.1111/tmi.13042>
- GHOSH, Susmita et al. Coinfection of *Leptomonas seymouri* and *Leishmania donovani* in Indian Leishmaniasis. **Journal of Clinical Microbiology**, [S. l.], v. 50, n. 8, p. 2774–2778, 2012. Disponível em: <https://doi.org/10.1128/JCM.00966-12>
- GHOSHAL, Angana; MANDAL, Chitra. **A Perspective on the Emergence of Sialic Acids as Potent Determinants Affecting Leishmania Biology.** [s. l.], 2011. Research article. Disponível em: <https://doi.org/10.4061/2011/532106>. Acesso em: 22 nov. 2018.
- GONTIJO, Célia Maria Ferreira; MELO, Maria Norma. Leishmaniose visceral no Brasil: quadro atual, desafios e perspectivas. **Revista Brasileira de Epidemiologia**, [S. l.], v. 7, n. 3, p. 338–349, 2004. Disponível em: <https://doi.org/10.1590/S1415-790X2004000300011>
- HOARE, Cecil A.; WALLACE, Franklin G. Developmental Stages of Trypanosomatid Flagellates: a New Terminology. **Nature**, [S. l.], v. 212, n. 5068, p. 1385–1386, 1966. Disponível em: <https://doi.org/10.1038/2121385a0>
- HOMBACH, Antje et al. A small heat shock protein is essential for thermotolerance and intracellular survival of *Leishmania donovani*. **Journal of Cell Science**, [S. l.], v. 127, n. 21, p. 4762–4773, 2014. Disponível em: <https://doi.org/10.1242/jcs.157297>
- IMAMURA, Hideo et al. Evolutionary genomics of epidemic visceral leishmaniasis in the Indian subcontinent. **eLife**, [S. l.], v. 5, 2016. Disponível em: <https://doi.org/10.7554/eLife.12613>. Acesso em: 24 maio. 2020.
- IMAMURA, Hideo; DUJARDIN, Jean-Claude. A Guide to Next Generation Sequence Analysis of *Leishmania* Genomes. **Methods in Molecular Biology (Clifton, N.J.)**, [S. l.], v. 1971, p. 69–94, 2019. Disponível em: [https://doi.org/10.1007/978-1-4939-9210-2\\_3](https://doi.org/10.1007/978-1-4939-9210-2_3)
- ISHEMGULOVA, Aygul et al. Molecular mechanisms of thermal resistance of the insect trypanosomatid *Crithidia thermophila*. **PLOS ONE**, [S. l.], v. 12, n. 3, p. e0174165, 2017.

- Disponível em: <https://doi.org/10.1371/journal.pone.0174165>
- IVENS, Alasdair C. et al. The Genome of the Kinetoplastid Parasite, *Leishmania major*. **Science (New York, N.Y.)**, [S. l.], v. 309, n. 5733, p. 436–442, 2005. Disponível em: <https://doi.org/10.1126/science.1112680>
- JACKSON, Andrew P.; QUAIL, Michael A.; BERRIMAN, Matthew. Insights into the genome sequence of a free-living Kinetoplastid: *Bodo saltans* (Kinetoplastida: Euglenozoa). **BMC Genomics**, [S. l.], v. 9, n. 1, p. 594, 2008. Disponível em: <https://doi.org/10.1186/1471-2164-9-594>
- JESUS, A. R. de et al. Gene deletion suggests a role for *Trypanosoma cruzi* surface glycoprotein GP72 in the insect and mammalian stages of the life cycle. **Journal of Cell Science**, [S. l.], v. 106, n. 4, p. 1023–1033, 1993.
- JOSHI, NA; FASS, JN. **Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]**. [S. l.: s. n.] Disponível em: Available at <https://github.com/najoshi/sickle>
- KAUFER, Alexa et al. The evolution of trypanosomatid taxonomy. **Parasites & Vectors**, [S. l.], v. 10, 2017. Disponível em: <https://doi.org/10.1186/s13071-017-2204-7>. Acesso em: 22 nov. 2018.
- KOREN, Sergey et al. Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. **Genome Research**, [S. l.], 2017. Disponível em: <https://doi.org/10.1101/gr.215087.116>. Acesso em: 31 maio. 2019.
- KURTZ, Stefan et al. Versatile and open software for comparing large genomes. **Genome Biology**, [S. l.], v. 5, n. 2, p. R12, 2004. Disponível em: <https://doi.org/10.1186/gb-2004-5-2-r12>
- LEMESRE, J. L. et al. Requirements of defined cultivation conditions for standard growth of *Leishmania promastigotes* in vitro. **Acta Tropica**, [S. l.], v. 45, n. 2, p. 99–108, 1988.
- LI, Heng; DURBIN, Richard. Fast and accurate short read alignment with Burrows-Wheeler transform. **Bioinformatics (Oxford, England)**, [S. l.], v. 25, n. 14, p. 1754–1760, 2009. Disponível em: <https://doi.org/10.1093/bioinformatics/btp324>
- LUKEŠ, Julius et al. Evolution of parasitism in kinetoplastid flagellates. **Molecular and Biochemical Parasitology**, [S. l.], v. 195, n. 2, Special Issue on the 35th Anniversary of Molecular and Biochemical Parasitology, p. 115–122, 2014. Disponível em: <https://doi.org/10.1016/j.molbiopara.2014.05.007>
- MAIA-ELKHOURY, Ana Nilce Silveira et al. Visceral leishmaniasis in Brazil: trends and challenges. **Cadernos de Saúde Pública**, [S. l.], v. 24, n. 12, p. 2941–2947, 2008. Disponível em: <https://doi.org/10.1590/S0102-311X2008001200024>
- MANNAERT, An et al. Adaptive mechanisms in pathogens: universal aneuploidy in *Leishmania*. **Trends in Parasitology**, [S. l.], v. 28, n. 9, p. 370–376, 2012. Disponível em: <https://doi.org/10.1016/j.pt.2012.06.003>
- MARCILI, Arlei et al. Phylogenetic relationships of *Leishmania* species based on trypanosomatid barcode (SSU rDNA) and gGAPDH genes: Taxonomic revision of *Leishmania (L.) infantum chagasi* in South America. **Infection, Genetics and Evolution: Journal of Molecular Epidemiology and Evolutionary Genetics in Infectious Diseases**, [S. l.], v. 25, p. 44–51, 2014. Disponível em: <https://doi.org/10.1016/j.meegid.2014.04.001>
- MARTIN, Marcel. Cutadapt removes adapter sequences from high-throughput sequencing reads. **EMBnet.journal**, [S. l.], v. 17, n. 1, p. 10–12, 2011. Disponível em: <https://doi.org/10.14806/ej.17.1.200>
- MARUYAMA, Sandra R. et al. Non-*Leishmania* Parasite in Fatal Visceral Leishmaniasis-Like Disease, Brazil. [S. l.], v. 25, n. 11, p. 2088–2092, 2019. Disponível em:

- <https://dx.doi.org/10.3201/eid2511.181548>
- MASLOV, Dmitri A. et al. Recent advances in trypanosomatid research: genome organization, expression, metabolism, taxonomy and evolution. **Parasitology**, [S. l.], v. 146, n. 1, p. 1–27, 2019. Disponível em: <https://doi.org/10.1017/s0031182018000951>
- MASLOV, Dmitri A.; PODLIPAEV, Sergei A.; LUKES, Julius. Phylogeny of the kinetoplastida: taxonomic problems and insights into the evolution of parasitism. **Memórias do Instituto Oswaldo Cruz**, [S. l.], v. 96, n. 3, p. 397–402, 2001. Disponível em: <https://doi.org/10.1590/S0074-02762001000300021>
- NEI, Masatoshi; KUMAR, Sudhir. **Molecular Evolution and Phylogenetics**. Oxford, New York: Oxford University Press, 2000. E-book.
- NOYES, H. a et al. A nested PCR for the *ssrRNA* gene detects *Trypanosoma binneyi* in the platypus and *Trypanosoma* sp. in wombats and kangaroos in Australia1. **International Journal for Parasitology**, [S. l.], v. 29, n. 2, p. 331–339, 1999. Disponível em: [https://doi.org/10.1016/S0020-7519\(98\)00167-2](https://doi.org/10.1016/S0020-7519(98)00167-2)
- NOYES, Harry A. et al. A Nested-PCR-Based Schizodeme Method for Identifying *Leishmania* Kinetoplast Minicircle Classes Directly from Clinical Samples and Its Application to the Study of the Epidemiology of *Leishmania tropica* in Pakistan. **Journal of Clinical Microbiology**, [S. l.], v. 36, n. 10, p. 2877–2881, 1998.
- OXFORD NANOPORE TECHNOLOGY. **Medaka — Medaka 0.11.1 documentation**. [s. l.], 2018. Disponível em: <https://nanoporetech.github.io/medaka/>. Acesso em: 18 dez. 2019.
- PEACOCK, Christopher S. et al. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. **Nature Genetics**, [S. l.], v. 39, n. 7, p. 839–847, 2007. Disponível em: <https://doi.org/10.1038/ng2053>
- PHILIPPE, H.; ADOUTTE, A. **The molecular phylogeny of Eukaryota: solid facts and uncertainties**. [s. l.], 1998. Disponível em: <https://ci.nii.ac.jp/naid/10015054422/#cit>. Acesso em: 22 nov. 2018.
- PUIG, L.; PRADINAUD, R. *Leishmania* and HIV co-infection: dermatological manifestations. **Annals of Tropical Medicine and Parasitology**, [S. l.], v. 97 Suppl 1, p. 107–114, 2003. Disponível em: <https://doi.org/10.1179/000349803225002589>
- RANGEL, Diana Azeredo et al. Isolation and characterization of trypanosomatids, including *Crithidia mellificae*, in bats from the Atlantic Forest of Rio de Janeiro, Brazil. **PLOS Neglected Tropical Diseases**, [S. l.], v. 13, n. 7, p. e0007527, 2019. Disponível em: <https://doi.org/10.1371/journal.pntd.0007527>
- RITTIG, M. G.; BOGDAN, C. *Leishmania*–Host-cell Interaction: Complexities and Alternative Views. **Parasitology Today**, [S. l.], v. 16, n. 7, p. 292–297, 2000. Disponível em: [https://doi.org/10.1016/S0169-4758\(00\)01692-6](https://doi.org/10.1016/S0169-4758(00)01692-6)
- RODRIGUES, Vasco et al. Regulation of immunity during visceral *Leishmania* infection. **Parasites & Vectors**, [S. l.], v. 9, 2016. Disponível em: <https://doi.org/10.1186/s13071-016-1412-x>. Acesso em: 18 maio. 2020.
- SCHÖNIAN, G.; KUHLS, K.; MAURICIO, I. L. Molecular approaches for a better understanding of the epidemiology and population genetics of *Leishmania*. **Parasitology**, [S. l.], v. 138, n. 4, p. 405–425, 2011. Disponível em: <https://doi.org/10.1017/S0031182010001538>
- SCHÖNIAN, Gabriele et al. PCR diagnosis and characterization of *Leishmania* in local and imported clinical samples11Part of this work has been presented at the Second World Congress on Leishmaniasis in Crete, Greece, May 2001. **Diagnostic Microbiology and Infectious Disease**, [S. l.], v. 47, n. 1, p. 349–358, 2003. Disponível em: [https://doi.org/10.1016/S0732-8893\(03\)00093-2](https://doi.org/10.1016/S0732-8893(03)00093-2)

- SHERLOCK, Ítalo A. et al. **Natural infection of the opossum *Didelphis albiventris* (Marsupialia, Didelphidae) with *Leishmania donovani*, in Brazil.** [s. l.], 1984. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0074-02761984000400020&lng=en&nrm=iso&tlng=en](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0074-02761984000400020&lng=en&nrm=iso&tlng=en). Acesso em: 22 nov. 2018.
- SIEVERS, Fabian et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. **Molecular Systems Biology**, [S. l.], v. 7, p. 539, 2011. Disponível em: <https://doi.org/10.1038/msb.2011.75>
- SILVA, Eduardo S. et al. Visceral leishmaniasis in the Metropolitan Region of Belo Horizonte, State of Minas Gerais, Brazil. **Memórias do Instituto Oswaldo Cruz**, [S. l.], v. 96, n. 3, p. 285–291, 2001. Disponível em: <https://doi.org/10.1590/S0074-02762001000300002>
- SIMPSON, Jared T. et al. ABySS: A parallel assembler for short read sequence data. **Genome Research**, [S. l.], v. 19, n. 6, p. 1117–1123, 2009. Disponível em: <https://doi.org/10.1101/gr.089532.108>
- SRIVASTAVA, Pankaj et al. Detection of *Leptomonas* sp. parasites in clinical isolates of Kala-azar patients from India. **Infection, genetics and evolution: journal of molecular epidemiology and evolutionary genetics in infectious diseases**, [S. l.], v. 10, n. 7, p. 1145–1150, 2010. Disponível em: <https://doi.org/10.1016/j.meegid.2010.07.009>
- STERKERS, Yvon et al. Novel insights into genome plasticity in Eukaryotes: mosaic aneuploidy in *Leishmania*. **Molecular Microbiology**, [S. l.], v. 86, n. 1, p. 15–23, 2012. Disponível em: <https://doi.org/10.1111/j.1365-2958.2012.08185.x>
- STEVERDING, Dietmar. The history of leishmaniasis. **Parasites & Vectors**, [S. l.], v. 10, 2017. Disponível em: <https://doi.org/10.1186/s13071-017-2028-5>. Acesso em: 16 maio. 2020.
- TAMURA, Koichiro et al. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. **Molecular Biology and Evolution**, [S. l.], v. 30, n. 12, p. 2725–2729, 2013. Disponível em: <https://doi.org/10.1093/molbev/mst197>
- TEIXEIRA, D. G. et al. Comparative analyses of whole genome sequences of *Leishmania infantum* isolates from humans and dogs in northeastern Brazil. **International Journal for Parasitology**, [S. l.], v. 47, n. 10–11, p. 655–665, 2017. Disponível em: <https://doi.org/10.1016/j.ijpara.2017.04.004>
- THAKUR, Lovlesh et al. Atypical leishmaniasis: A global perspective with emphasis on the Indian subcontinent. **PLoS neglected tropical diseases**, [S. l.], v. 12, n. 9, p. e0006659, 2018. Disponível em: <https://doi.org/10.1371/journal.pntd.0006659>
- VALDIVIA, Hugo O. et al. Comparative genomic analysis of *Leishmania* (*Viannia*) *peruviana* and *Leishmania* (*Viannia*) *braziliensis*. **BMC Genomics**, [S. l.], v. 16, n. 1, 2015. Disponível em: <https://doi.org/10.1186/s12864-015-1928-z>. Acesso em: 23 maio. 2020.
- VALDIVIA, Hugo O. et al. Comparative genomics of canine-isolated *Leishmania* (*Leishmania*) *amazonensis* from an endemic focus of visceral leishmaniasis in Governador Valadares, southeastern Brazil. **Scientific Reports**, [S. l.], v. 7, n. 1, p. 40804, 2017. Disponível em: <https://doi.org/10.1038/srep40804>
- VICKERMAN, K.; PRESTON, T. M. Comparative cell biology of the kinetoplastid flagellates. **Biology of the Kinetoplastida**, [S. l.], 1976. Disponível em: <http://agris.fao.org/agris-search/search.do?recordID=US201302971463>. Acesso em: 22 nov. 2018.
- VOTÝPKA, Jan et al. Diversity of Trypanosomatids (Kinetoplastea: Trypanosomatidae) Parasitizing Fleas (Insecta: Siphonaptera) and Description of a New Genus *Blechomonas* gen. n. **Protist**, [S. l.], v. 164, n. 6, p. 763–781, 2013. Disponível em: <https://doi.org/10.1016/j.protis.2013.08.002>
- VOTÝPKA, Jan et al. New Approaches to Systematics of Trypanosomatidae: Criteria for

- Taxonomic (Re)description. **Trends in Parasitology**, [S. l.], v. 31, n. 10, p. 460–469, 2015. Disponível em: <https://doi.org/10.1016/j.pt.2015.06.015>
- WALKER, Bruce J. et al. Pilon: An Integrated Tool for Comprehensive Microbial Variant Detection and Genome Assembly Improvement. **PLOS ONE**, [S. l.], v. 9, n. 11, p. e112963, 2014. Disponível em: <https://doi.org/10.1371/journal.pone.0112963>
- WEI SHEN et al. **SeqKit: A Cross-Platform and Ultrafast Toolkit for FASTA/Q File Manipulation**. [s. l.], 2016. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0163962>. Acesso em: 17 dez. 2019.
- WERNECK, Guilherme Loureiro. Visceral leishmaniasis in Brazil: rationale and concerns related to reservoir control. **Revista de Saúde Pública**, [S. l.], v. 48, n. 5, p. 851–856, 2014. Disponível em: <https://doi.org/10.1590/S0034-8910.2014048005615>
- WHEELER, Richard John; GLUENZ, Eva; GULL, Keith. **The Limits on Trypanosomatid Morphological Diversity**. [s. l.], 2013. Disponível em: <https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0079581>. Acesso em: 22 nov. 2018.
- WILFERT, L. et al. Trypanosomatids are common and diverse parasites of *Drosophila*. **Parasitology**, [S. l.], v. 138, n. 7, p. 858–865, 2011. Disponível em: <https://doi.org/10.1017/S0031182011000485>
- YURCHENKO, Vyacheslav et al. An Integrated Morphological and Molecular Approach to a New Species Description in the Trypanosomatidae: the Case of *Leptomonas podlipaevi* n. sp., a Parasite of *Boisea rubrolineata* (Hemiptera: Rhopalidae). **Journal of Eukaryotic Microbiology**, [S. l.], v. 53, n. 2, p. 103–111, 2006. Disponível em: <https://doi.org/10.1111/j.1550-7408.2005.00078.x>