

FUNDAÇÃO OSWALDO CRUZ
INSTITUTO AGGEU MAGALHÃES
MESTRADO ACADÊMICO EM BIOCÊNCIAS E BIOTECNOLOGIA EM SAÚDE

ADRIANA NEUMAN ALBUQUERQUE LINS MOURA DE BRITO

AVALIAÇÃO COMPARATIVA DO PERFIL DE EXPRESSÃO DE GENES
PARÁLOGOS CODIFICANTES DA PROTEÍNA DE VIRULÊNCIA GP63 DE
Leishmania braziliensis

RECIFE
2019

ADRIANA NEUMAN ALBUQUERQUE LINS MOURA DE BRITO

**AVALIAÇÃO COMPARATIVA DO PERFIL DE EXPRESSÃO DE GENES
PARÁLOGOS CODIFICANTES DA PROTEÍNA DE VIRULÊNCIA GP63 DE
*Leishmania braziliensis***

Dissertação apresentada ao Curso de Mestrado em Biociências e Biotecnologia em Saúde do Instituto de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, como requisito para obtenção do título de Mestre em ciências.

Orientador: Dr. Osvaldo Pompílio de Melo Neto

Coorientador: Dr. Artur Leonel de Castro Neto

RECIFE

2019

Catálogo na fonte
Elaine Barroso
CRB 1728

Brito, Adriana Neuman Albuquerque Lins Moura de
Avaliação da variabilidade gênica de múltiplos parálogos da proteína de virulência GP63 de *Leishmania braziliensis*/ Adriana Neuman Albuquerque Lins Moura de Brito- Recife: O Autor, 2016.

59 folhas : il., fig., tab.

Orientador: Osvaldo Pompílio de Melo Neto

Coorientador: Artur Leonel de Castro

Monografia (graduação) – Universidade Federal de Pernambuco. Centro de Biociências. Biomedicina, 2016.

Inclui referências

- 1. Leishmaniose 2. Análise cladística 3. Genes I. Melo Neto, Osvaldo Pompílio de (orientador) II. Castro Neto, Artur Leonel de (coorientador) III. Título**

616.9364

CDD (22.ed.)

UFPE/CCB-2017-032

ADRIANA NEUMAN ALBUQUERQUE LINS MOURA DE BRITO

**AVALIAÇÃO COMPARATIVA DO PERFIL DE EXPRESSÃO DE GENES
PARÁLOGOS CODIFICANTES DA PROTEÍNA DE VIRULÊNCIA GP63 DE
*Leishmania braziliensis***

Dissertação apresentada ao Curso de Mestrado em Biociências e Biotecnologia em Saúde do Instituto de Pesquisas Aggeu Magalhães, Fundação Oswaldo Cruz, como requisito para obtenção do título de Mestre em ciências.

Aprovada em: ___/___/_____

BANCA EXAMINADORA

Dr. Osvaldo Pompílio de Melo Neto
Instituto Aggeu Magalhães (IAM/FIOCRUZ-PE)

Dra. Tatiany Patrícia Romão Pompílio de Melo
Instituto Aggeu Magalhães (IAM/FIOCRUZ-PE)

Dr. Will de Barros Pita
Universidade Federal de Pernambuco (UFPE)

Ao meu padrasto (*in memoriam*)

Dedico

AGRADECIMENTOS

Ao meu orientador, Osvaldo Pompílio de Neto, por toda paciência e ensinamentos passados ao longo desses anos na Microbiologia.

Ao meu coorientador, Artur Leonel de Castro Neto, por estar presente em cada etapa da construção desse projeto, por acreditar em mim, e por dividir tanta coisa comigo diariamente (mesmo distante). Obrigada também pela parceria (não só acadêmica) e amizade ao longo desses anos.

Ao grupo GP63, Artur Leonel (novamente) e Allana Dutra, por trabalharem lado a lado comigo! Se tem uma coisa que eu sinto orgulho na vida, é do nosso grupo de pesquisa GP63. Obrigada pela nossa união e jamais vou esquecer o quanto avançamos e crescemos no estudo dessa molécula. E ainda vamos continuar crescendo.

Aos pesquisadores, Antônio Rezende, pelo suporte na parte do RNASeq e Danielle Moura, por responder a todas as minhas dúvidas praticamente diárias, com muita paciência.

A minha mãe, Maria das Graças, por cuidar de mim e me apoiar em seja qual for a decisão que eu tomar na vida. Eu te amo muito!

Ao meu pai, Adriano Brito, por todo o suporte e por sentir tanto orgulho de mim. E a minha madrastra, que me deu meu segundo maior presente, Giovanninha!

Ao meu padrasto, que cuidou de mim por muitos anos, por ser meu segundo e eterno pai, e por saber que onde você estiver estará olhando e cuidando de nós, suas mulheres.

À minha irmã, Ana Júlia, por todo companheirismo, por ser meu lado bom, e principalmente por me aguentar diariamente. Uma tarefa árdua.

Ao meu tio, Everton Lins, sua esposa Renata, e Bianquinha, por terem sido meu chão em diversos momentos, e por serem a minha parte preferida da família.

As minhas biomédicas (que são minha segunda família): minha irmã de alma (Natália Kryzia), Larissa Krovovsky, Thaysa Aragão e Jade Ribeiro. À Leandro Andrade, à Camilla Padilha, à Virgínia Patricia, à Fábio Durval.

As amizades que a Microbiologia e o mestrado me deram. Sou eternamente grata a vocês. Vocês me entendem e me apoiam, me dão conselhos, acompanham os meus dramas, sofrem e se alegram comigo a cada conquista! Cada um de vocês tem a sua importância na minha vida. Thaíse Yasmine, Mayara Barbalho, Adriana Roberto, Wagner Tenório, Camila Cavalcanti, Ísis Cirino, Maria Ribeiro, Gustavo Lima, Irassandra, Ada, Moezio, Sthephanny, Yallen, Ludmila, Renata, Tallyta, Beatriz, Izabel, Romário, Lílian, Maressa, Kleisson. Além dos vários não citados, por me proporcionarem muitos momentos de descontração e alegria no ambiente de trabalho.

À todos os funcionários do Departamento de Microbiologia e da FIOCRUZ-PE, em especial a Maria Carolina, por sempre estar disposta a nos ajudar.

Ao Biotério da FIOCRUZ-PE, e aos seus funcionários, por serem fundamentais para a realização desse projeto.

Aos órgãos de fomento CNPq e FACEPE.

“Eu estou entre aqueles que acreditam que a ciência tem uma grande beleza. Um cientista em seu laboratório não é apenas um técnico: é também uma criança colocada frente a um fenômeno natural que a impressiona tanto quanto um conto de fadas”.

(Marie Curie)

RESUMO

Ao longo da sua evolução as espécies de *Leishmania* desenvolveram métodos de escape do sistema imune dos hospedeiros. Um dos principais mecanismos de escape é mediado pela proteína de virulência GP63, que possui uma diferença significativa na quantidade de genes que a codificam, em diferentes espécies de *Leishmania*. Dentre as principais espécies patogênicas, a *L. braziliensis* possui o maior número de genes codificantes (39), quando comparada a *L. infantum* (7) e *L. major* (6). Análises genômicas indicam que esses genes podem possuir diferentes momentos de expressão ao longo do ciclo celular do parasita. Este estudo objetivou avaliar a expressão de grupos de proteínas de GP63, nas diferentes fases do ciclo de vida de *Leishmania* sp. Para tal, anticorpos policlonais anti-GP63 produzidos frente a peptídeos específicos dessas proteínas foram obtidos, purificados e utilizados em ensaios de *western blotting* (WB) frente a proteínas recombinantes GP63, juntamente com ensaios de curvas de crescimento de *L. braziliensis* e *L. infantum*. A *L. braziliensis* possui uma grande quantidade e variabilidade de genes de GP63, que quando testadas frente a anticorpos feitos contra peptídeos específicos da mesma demonstraram, tanto reações cruzadas contra outras GP63 como especificidade para determinados genes. Os ensaios utilizando curvas de crescimento de *L. braziliensis* sugeriram a hipótese de que existe uma maior quantidade de genes sendo expressos em determinados momentos do ciclo do parasita. Ensaios de RNASeq foram realizados e não foi possível identificar um grupo de genes GP63 sendo claramente expressos, apesar de existir uma maior quantidade de transcritos na fase estacionária do parasita. Os resultados observados fornecem uma visão geral no padrão do reconhecimento de anticorpos anti-GP63 específicos e no perfil de expressão dos genes de GP63 na fase promastigota de *L. braziliensis*, evidenciando a plasticidade desse parasita, responsável pela maioria dos casos de LC no Brasil.

Palavras-chave: Metalloendopeptidases, glicoproteína gp63-Leishmania, *Leishmania braziliensis*

ABSTRACT

Throughout its evolution the species of *Leishmania* developed methods of escape of the immune system of the hosts. One of the main escape mechanisms is mediated by the GP63 virulence protein, which has a significant difference in the number of genes encoding it in different *Leishmania* species. Among the major pathogenic species, *L. braziliensis* has the highest number of coding genes (39) when compared to *L. infantum* (7) and *L. major* (6). Genomic analyzes indicate that these genes may have different expression moments throughout the cell cycle of the parasite. This study aimed to evaluate the expression of groups of GP63 proteins in the different phases of the life cycle of *Leishmania* sp. To that end, anti-GP63 polyclonal antibodies produced against specific peptides of these proteins were obtained, purified and used in western blotting (WB) assays against recombinant GP63 proteins, together with growth curve assays of *L. braziliensis* and *L. infantum*. *L. braziliensis* has a large amount and variability of GP63 genes, which when tested against antibodies against specific peptides of the same, demonstrated both cross-reactions against other GP63 and specificity for certain genes. The assays using *L. braziliensis* growth curves suggested the hypothesis that there is a greater amount of genes being expressed at certain times of the parasite cycle. RNASeq assays were performed and it was not possible to identify a group of GP63 genes being clearly expressed, although there is a greater amount of transcripts in the stationary phase of the parasite. The observed results provide an overview of the pattern of recognition of specific anti-GP63 antibodies and the expression profile of GP63 genes in the promastigote phase of *L. braziliensis*, evidencing the plasticity of this parasite, responsible for the majority of cases of CL in Brazil.

Keywords: Metalloendopeptidases, glycoprotein gp63-*Leishmania*, *Leishmania braziliensis*

LISTA DE ILUSTRAÇÕES

Figura 1- Formas evolutivas de <i>Leishmania sp.</i>	18
Figura 2- Ciclo de vida dos parasitas causadores das Leishmanioses.....	19
Figura 3- Hepatoesplenomegalia causada por LV.....	23
Figura 4- Leishmaniose Cutânea.....	24
Figura 5- Leishmaniose Mucocutânea.....	25
Quadro 1- Fatores de virulência de <i>Leishmania sp.</i> e suas respectivas funções.....	27
Figura 6- Vias de sinalização dos macrófagos sendo influenciadas por GP63.....	31
Figura 7- Estrutura tridimensional da proteína GP63.....	35
Figura 8- Expressão de genes GP63 em <i>L. major</i>	37
Figura 9- Esquema metodológico.....	42
Figura 10- Filogenia dos genes de GP63.....	52
Figura 11- Verificação de subclonagens dos genes de GP63.....	54
Figura 12- Expressão em <i>E. coli</i> das proteínas GP63 recombinantes fusionadas a sequência de poli-histidinas.....	55
Figura 13- Análise da expressão da proteína GP63 recombinante LbrM.10.0540 fusionada a uma sequência de poli-histidinas.....	55
Figura 14- Características proteicas da proteína GP63.....	57
Quadro 2- Sequências de aminoácidos escolhidas para obtenção de peptídeos sintéticos.....	57
Figura 15- Localização dos peptídeos no alinhamento de sequências.....	58
Figura 16- Soro anti-peptídeo conservado frente ao peptídeo conjugado ao BSA e frente a proteínas recombinantes.....	59
Figura 17- Extratos de <i>E. coli</i> contendo as proteínas recombinantes GP63.....	60
Figura 18- Reconhecimento de proteínas GP63 recombinantes pelos anticorpos anti-peptídeo.....	62
Figura 19- Anticorpos anti-peptídeos frente a curvas de <i>L. braziliensis</i>	64
Figura 20- Análise das bibliotecas de cDNA sequenciadas.....	65
Figura 21- TPM – Fases Estacionária e Logarítmicas de <i>L. braziliensis</i>	67
Figura 22- <i>Heatmap</i> da expressão dos genes de GP63.....	68
Figura 23- Reads localizadas na C-terminal do gene 10.0590.....	69

LISTA DE ABREVIATURAS E SIGLAS

BSA	- <i>Bovine serum albumin</i>
EROs	- Espécies reativas de oxigênio
G	- Força gravitacional
GIPLs	- Glicoinositolfosfolipideos
GP63	- Glicoproteína 63
GPI	- Glicosilfosfatidilinositol
GST	- Glutathione S-transferase
His	- Histidina
HT	- Hipersensibilidade tardia
IFN γ	- Interferon gama
IgG	- Imunoglobulinas
IL-12	- Interleucina 12
IPTG	- Isopropil-tio- β -D-galactopiranosídeo
IPTG	- Isopropil-tio- β -D-galactopiranosídeo
kDa	- KiloDalton
KLH	- Keyhole limpet hemocyanin
LB	- Luria Bertani
LC	- Leishmaniose Cutânea
LM	- Leishmaniose Mucocutânea
LMW	- <i>Low molecular weight marker</i>
LPG	- Lipofosfoglicano
LV	- Leishmaniose Visceral
MARCKs	- Substrato de quinase C rico em alanina miristoilada
mL	- mililitro
mM	- milimolar
NO	- Óxido Nítrico
Pb	- Pares de base
PBS	- <i>Phosphate buffered saline</i>
PCK	- Proteína quinase C
PCR	- Reação em cadeia de polimerase

PPGs	- Proteofosfoglicanos
PTPs	- Proteínas tirosina fosfatase
Rpm	- Rotações por minuto
SDS-PAGE	- Sodium <i>Dodecyl Sulfate</i> – <i>Polyacrylamide Gel Electrophoresis</i>
SFB	- Soro fetal bovino
TNF α	- Fator de necrose tumoral alfa
X – gal	- 5-bromo-4-chloro-3-indolyl- β -D-galactopyranoside
μ g	- micrograma
μ L	- microlitro
VSG	- <i>Variant surface glycoprotein</i>
WB	- <i>Western blot</i>

SUMÁRIO

1 INTRODUÇÃO	15
2 FUNDAMENTAÇÃO TEÓRICA	17
2.1 Agente etiológico das Leishmanioses	17
2.2 Ciclo de vida de <i>Leishmania</i> sp.	19
2.3 Vetor de Transmissão das Leishmanioses	19
2.4 Epidemiologia das leishmanioses	20
2.4.1 Epidemiologia da Leishmaniose Visceral	21
2.4.2 Epidemiologia da Leishmaniose Cutânea	21
2.5 Manifestações clínicas da Leishmaniose Visceral	22
2.6 Manifestações clínicas da Leishmaniose Cutânea	23
2.7 Manifestações clínicas da Leishmaniose Mucocutânea	24
2.8 Resposta imune do hospedeiro e patogenia nas leishmanioses	25
2.9 Fatores de virulência de <i>Leishmania</i> sp.	27
2.10 Metaloproteases	29
2.11 GP63	29
2.11.1 Principais funções	30
2.11.2 GP63 no inseto vetor	32
2.11.3 Organização gênica	33
2.11.4 Estrutura	34
2.11.5 Variações na região C-terminal	35
2.11.6 Expressão	36
2.11.7 Classificação dos genes de GP63 em <i>L. braziliensis</i>	37
3 PERGUNTA CONDUTORA	39
4 HIPÓTESE	40
5 OBJETIVO GERAL	41
5.1 Objetivos Específicos	41
6 METODOLOGIA	42
6.1 Alinhamento e análise filogenética	42
6.2 Subclonagem de proteínas GP63 recombinantes	43
6.3 Expressão de proteínas GP63 recombinantes	44
6.4 Purificação das proteínas GP63 recombinantes	44
6.5 Imunização em coelhos	45

6.5.1	Imunização com a proteína 10.0540 fusionada a poli-histidina	45
6.5.2	Imunização do peptídeo conservado conjugado ao KLH	46
6.6	Obtenção comercial de anticorpos.....	46
6.7	Purificação dos anticorpos por imunoadsorção	46
6.8	Cultivo celular de formas promastigotas e curvas de crescimento	47
6.9	Obtenção de extrato celular	48
6.10	Ensaio de <i>Western blot</i>.....	48
6.11	Extração de RNA total para análise de RNASeq.....	49
6.12	Avaliação quantitativa do RNA e preparação da biblioteca de cDNA.....	49
6.13	Verificação da qualidade das sequências obtidas na análise de RNASeq	49
6.14	Trimagem e mapeamento das leituras obtidas na análise de RNASeq	50
7	RESULTADOS	51
7.1	Avaliação da diversidade das sequências dos genes de GP63 no cromossomo 10 de <i>Leishmania sp.</i>.....	51
7.2	Expressão de proteínas recombinantes para a produção de anticorpos	53
7.3	Estrutura das proteínas GP63 e escolha de peptídeos sintéticos para a produção de anticorpos.....	56
7.4	Reconhecimento dos anticorpos obtidos a partir dos peptídeos sintético..	58
7.5	Reconhecimento dos anticorpos anti-peptídeos frente a curvas de <i>L. braziliensis</i>.....	63
7.6	Quantidade de mRNA de diferentes genes de GP63 em diferentes fases de crescimento de <i>L. braziliensis</i>	65
8	DISCUSSÃO	70
9	CONCLUSÕES.....	74
	REFERÊNCIAS.....	75
	ANEXO A – ALINHAMENTO DE SEQUÊNCIAS GP63	82
	ANEXO B – PARECER DO COMITÊ DE ÉTICA.....	85
	ANEXO C – ARTIGO PUBLICADO	86

1 INTRODUÇÃO

As leishmanioses são infecções parasitárias causadas por protozoários flagelados da família *Trypanosomatidae* e do gênero *Leishmania*. A infecção provocada por estes patógenos pode levar ao desenvolvimento de diferentes manifestações clínicas, dependendo de complexas interações entre o parasita e o sistema imune do hospedeiro mamífero. As formas clínicas clássicas são divididas em três categorias principais: cutânea, mucocutânea e visceral. Dentre essas formas, a cutânea é a de maior incidência no Brasil e no mundo, causadas em boa parte por espécies do subgênero *Viannia* e geralmente caracterizada por úlceras na pele ou granulomas cutâneo.

Ao longo da sua evolução, as espécies de *Leishmania* desenvolveram métodos de escape ao sistema imune dos hospedeiros. Esse escape é mediado principalmente por proteínas e macromoléculas de virulência, que tem como principais representantes: lipofosfoglicanos (LPG), cisteína proteinases e a glicoproteína 63 (GP63). Entre suas variadas funções, durante o curso da infecção no hospedeiro vertebrado, elas agirem primariamente na instalação e manutenção do parasita, através de ações que contribuirão para a evasão do sistema imune inato.

Dos fatores de virulência citados, a proteína de superfície GP63 é considerada a principal representante. No hospedeiro mamífero, esta proteína age na inativação da cascata do sistema complemento e na clivagem de proteínas da matriz extracelular do hospedeiro, facilitando assim a sobrevivência do parasita. Uma outra função da GP63 é a capacidade de inibir a destruição da forma amastigota do protozoário pelo macrófago, através da ativação de proteínas tirosina fosfatase (PTPs), levando a inibição de cascatas bioquímicas que são responsáveis pela síntese de produtos que destroem a *Leishmania* intracelular.

Em relação a organização dos genes codificantes de GP63, foi observada a existência de múltiplos genes parálogos organizados em *tandem*, com uma diferença significativa em relação a quantidade de genes que codificam essa proteína entre as espécies de *Leishmania*. A *L. braziliensis* é a espécie que possui um maior número de cópias (39 cópias), quando comparada com outras espécies patogênicas, como *L. infantum* (15 cópias) e *L. major* (6 cópias).

Além do seu elevado número de cópias, resultados anteriores demonstraram que os genes de GP63 de espécies *Leishmania* podem ser classificados em três

grupos, baseado nas suas localizações cromossomais (cromossomos 10, 28 e 31). Os genes dos cromossomos 28 e 31 demonstraram uma maior relação com genes de GP63 encontrados em tripanossomatídeos que infectam só insetos ou plantas, sugerindo principalmente ações de adesão e nutrição. Já os genes no cromossomo 10 expandiram independentemente entre diferentes espécies e parecem estar mais associados a infecções nos hospedeiros vertebrados. Esses genes no cromossomo 10 parecem apresentar diferentes momentos de expressão durante o ciclo de vida do parasita e diferentes localizações subcelulares. Além disso apresentam uma variabilidade significativa nas suas sequências, principalmente em *L. braziliensis*, concentrada principalmente em trechos codificantes da proteína que estão expostos ao sistema imune, sugerindo um mecanismo de proteção do parasita contra o hospedeiro. Portanto, este projeto buscou avaliar a influência dessa variação de sequência na expressão diferenciada de parálogos codificantes de GP63, além de estabelecer uma relação entre essa expressão e a fase do ciclo de vida da *L. braziliensis*.

2 FUNDAMENTAÇÃO TEÓRICA

As leishmanioses são doenças causadas por protozoários intracelulares e caracterizadas por sua diversidade e complexidade (ORYAN; AKBARI, 2016). Estas doenças são conhecidas há centenas de anos, com uma das suas primeiras descrições clínicas feita em 1756 por Alexander Russell, sendo chamadas de “Febre Aleppo”. Vários nomes correspondem a esse grupo de enfermidades, tais como: *kala-azar*, *febre Dum-dum* e *lepra branca* (HIDE et al., 2007). De acordo com a Organização Mundial da Saúde (OMS), as leishmanioses estão entre as sete doenças tropicais mais importantes, representando um grave problema de saúde pública em áreas endêmicas, possuindo um amplo espectro de manifestações clínicas e potencialmente fatais (TORRES-GUERRERO et al., 2017).

Atualmente, existem 18 espécies do gênero *Leishmania* responsáveis pelo desenvolvimento das leishmanioses em seres humanos (STEVERDING, 2017). Esses protozoários são difundidos pela picada do “mosquito-palha”, onde fêmeas de flebotomíneos infectadas são as responsáveis pela transmissão, porém, já foram relatados casos de infecções congênitas e transmissão sanguínea (CLEM, 2010). Do ponto de vista dos seus hospedeiros, existem dois ciclos principais de leishmanioses: o ciclo zoonótico, que ocorre quando animais domésticos ou selvagens atuam como reservatórios e o homem acaba sendo um hospedeiro acidental; e o ciclo antroponótico, quando o homem é o único ou principal reservatório e atua como fonte de infecção para o vetor (GRAMICCIA, 2011; ROTUREAU, 2006).

As formas clínicas das leishmanioses dependem da interação complexa entre a resposta imune do hospedeiro e as espécies específicas de protozoários e vetores (KEVRIC; CAPPEL; KEELING, 2015). Desta forma são classificadas em três principais categorias: Leishmaniose Cutânea (LC), Leishmaniose Mucocutânea (LM) e Leishmanioses Visceral (LV) (LM) (CECÍLIO et al., 2014).

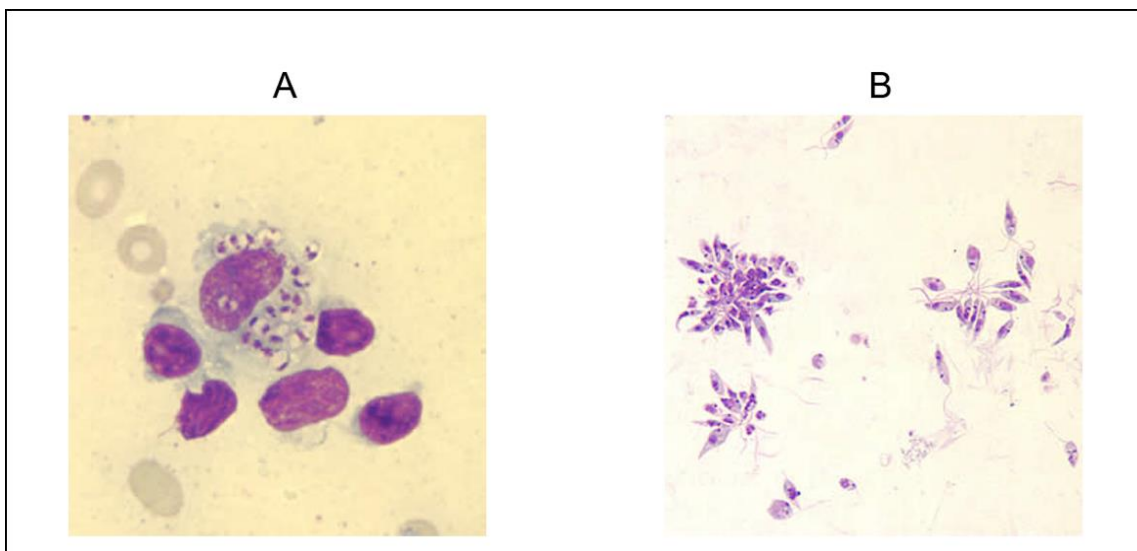
2.1 Agente etiológico das Leishmanioses

Até o momento, aproximadamente 50 espécies de *Leishmania* foram descritas em todo o mundo causadoras de leishmanioses em humanos e em animais (ALEMAYEHU; ALEMAYEHU, 2017). Estas espécies pertencem a família *Trypanosomatidae* (ordem Kinetoplastida), e as espécies patogênicas para o homem

são classificadas em dois subgêneros, *Leishmania* e *Viannia*. Estes se diferenciam pela posição na qual os parasitas das espécies associadas se desenvolvem no intestino do inseto vetor e pela localização geográfica em que evoluíram. O subgênero *Leishmania* evoluiu tanto no Novo Mundo como no Velho Mundo, abrangendo as espécies *L. major*, *L. tropica*, *L. aethiopica*, *L. donovani*, *L. infantum*, *L. mexicana*, *L. amazonensis*, entre outras. Já o subgênero *Viannia* se desenvolveu exclusivamente no Novo Mundo, e suas principais espécies de interesse médico são: *L. braziliensis*, *L. peruviana*, *L. panamensis* e *L. guyanensis* (BATES, 2007).

Os protozoários do gênero *Leishmania* possuem um único flagelo e são caracterizados por possuir um cinetoplasto, uma estrutura que consiste em DNA mitocondrial. Estes protozoários tem um ciclo de vida heteróxico e exibem duas formas morfológicas em seu ciclo de vida, que podem ser observadas na Figura 1: a forma amastigota, com formas redondas ou ovais que medem cerca de 3 a 5 μm e são encontrados em macrófagos do hospedeiro mamífero; e a forma promastigota, organismos móveis, alongados, medindo aproximadamente 10 a 15 μm de comprimento, com um único flagelo anterior e encontrados no intestino dos vetores (NEVES, 2002; ALEMAYEHU; ALEMAYEHU, 2017; DAWIT, 2013).

Figura 1 - Formas evolutivas de *Leishmania* sp.

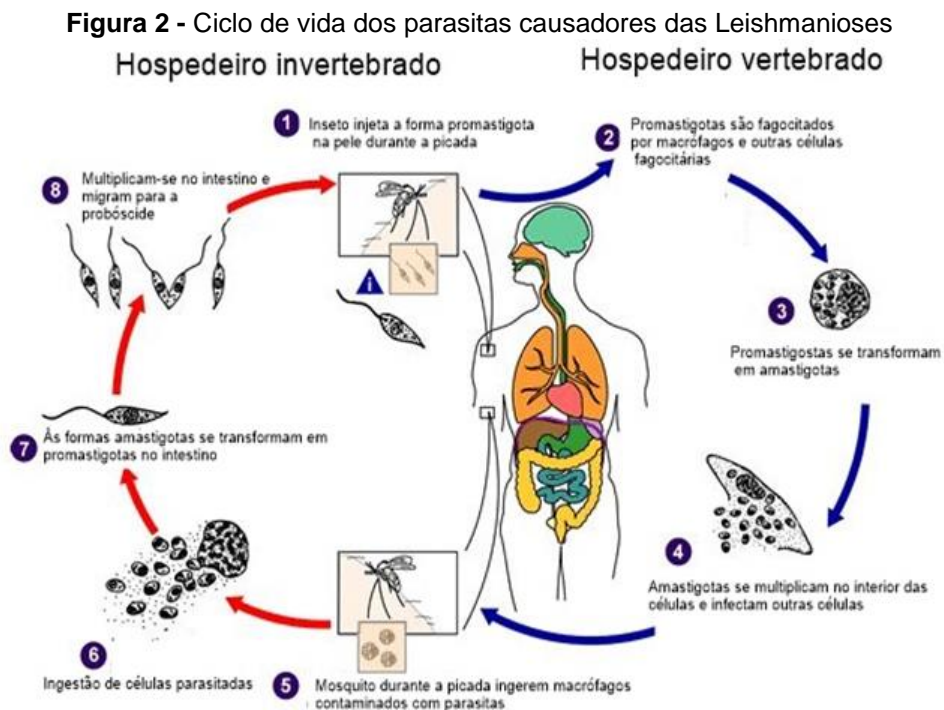


Fonte: A autora a partir de Centers for Disease Control and Prevention (2018)

Legenda: Amastigotas intracelulares de *Leishmania* sp. (A). Promastigostas de *Leishmania* sp. obtidos por cultura (B).

2.2 Ciclo de vida de *Leishmania* sp.

No ciclo de vida desses protozoários, a fêmea de *Phlebotomus* ao picar o animal infectado, absorve as formas amastigotas presentes na corrente sanguínea do reservatório. No intestino do inseto, o parasita inicia um processo de transformação e as amastigotas mudam para o estágio de promastigotas procíclicas, que se multiplicam e penetram no intestino médio para migrar para a glândula salivar e probóscide, se transformando em promastigotas metacíclicas, as quais tem capacidade infecciosa. Ao picar um animal saudável, o vetor inocula o parasita, regurgitando tais promastigotas. No local da picada os promastigotas são fagocitados por macrófagos, onde eles retornam a forma amastigota (Figura 2). Os macrófagos infectados se espalham no local da picada (no caso de LC) ou migram para diferentes tecidos e órgãos internos (no caso da LV), como o baço, fígado e medula óssea (BESSAT; SHANAT, 2015; TORRES-GUERRERO et al., 2017).



Fonte: Center for Disease Control and Prevention (2018, tradução nossa)

2.3 Vetor de Transmissão das Leishmanioses

Aproximadamente 93 espécies de fêmeas do inseto díptero popularmente conhecido como “mosquito-palha”, pertencentes ao gênero *Phlebotomus* no Velho

Mundo e *Lutzomya* nos trópicos do Novo Mundo, são os vetores responsáveis por transmitir as leishmanioses. O gênero *Lutzomya* inclui 90% das espécies patogênicas, a maioria das quais infecta humanos. Espécies desse gênero absorvem açúcares que podem desenvolver um papel importante no desenvolvimento das espécies de *Leishmania* (PACE, 2014; TORRES-GUERRERO et al., 2017). Os adultos são pequenos e raramente ultrapassam 3,5 mm de comprimento corporal, variando de quase branco a quase preto e sendo cobertos por pelos densos, mantendo suas asas sobre suas costas em formato característico de “V” quando em repouso. Tanto os machos quanto as fêmeas alimentam-se de secreções açucaradas de plantas, porém as fêmeas precisam se alimentar de sangue para completar o desenvolvimento dos ovos. São mais ativos durante a noite e, ao contrário dos mosquitos, seu ataque ao hospedeiro é silencioso (MAROLI et al., 2012).

Os vetores das leishmanioses são amplamente espalhados pelas regiões tropicais, subtropicais e temperadas, incluindo desertos, florestas tropicais, savanas e planaltos. Eles estão presentes principalmente nas zonas quentes da Ásia, África, Austrália, sul da Europa e Américas, estando ausentes na Nova Zelândia e ilhas do Pacífico (MAROLI et al., 2012; NADLER; ENK; LEON, 2014). A distribuição geográfica das leishmanioses depende diretamente da ação dos seus vetores, sua ecologia e das condições de desenvolvimento interno do parasita (DAWIT, 2013).

2.4 Epidemiologia das leishmanioses

Em escala global, aproximadamente 350 milhões de pessoas vivem em áreas endêmicas das leishmanioses, com 14 milhões de pessoas sendo diretamente afetadas. A incidência global destas é de aproximadamente dois milhões de novos casos por ano, dentre esses 1,5 milhões é de LC e 500.000 de LV (COSTA-SILVA et al., 2014; SAVOIA, 2015). No Hemisfério Oriental as leishmanioses são endêmicas na Ásia, África e sul da Europa (chamadas “leishmanioses de Velho Mundo”). Também são endêmicas no Hemisfério Ocidental, onde sua distribuição se estende desde o centro-sul do Texas até a América Central e do Sul, com exceção do Chile e Uruguai (“leishmanioses de Novo Mundo”). A doença não é encontrada na Antártida ou nas Ilhas do Pacífico e a Austrália aparentemente estava livre de *Leishmania* spp., porém infecções em cangurus e outros marsupiais foram relatadas em 2004 (DUTHIE et al., 2012; KEVRIC; CAPPEL; KEELING, 2015).

2.4.1 Epidemiologia da Leishmaniose Visceral

A LV é endêmica em 98 países, sendo que entre 20.000 a 400.000 mortes humanas são atribuídas a essa doença anualmente em todo o mundo e o risco de ressurgimento ou disseminação em novas áreas é provável de aumentar devido às mudanças climáticas. Mais de 67% dos casos da doença são encontrados no subcontinente indiano, afetando em grande parte os mais pobres da população (ABDULLAH et al., 2017).

Epidemias recorrentes de LV na África Oriental (Etiópia, Quênia, Sudão do Sul e Sudão) causam morbidade e mortalidade nas áreas afetadas (ORGANIZAÇÃO MUNDIAL DA SAÚDE, 2018). A distribuição geográfica de cada espécie de *Leishmania* influencia no tipo e na gravidade da doença que ocorre em cada região. A LV é causada por *L. donovani* no Sul da África e Ásia, enquanto *L. infantum* causa LV no Mediterrâneo, Oriente Médio, América Latina e partes da Ásia (DUTHIE et al., 2012).

Ambientes de baixo nível socioeconômico, predominantemente rurais e na periferia de grandes centros urbanos, são propícios a ocorrência de LV no território brasileiro, porém essa doença está se urbanizando, com a ocorrência de casos nos estados das regiões Sudeste e Centro Oeste. A predominância de LV ainda está na região Nordeste, com 77% dos casos ocorrendo nesta região e a doença é mais frequente em crianças menores de 10 anos (54,4%), sendo o sexo masculino o mais afetado (60%) (BRASIL, 2014).

2.4.2 Epidemiologia da Leishmaniose Cutânea

A LC é uma doença endêmica em mais de 70 países, porém 90% dos casos ocorrem no: Afeganistão, Argélia, Brasil, Paquistão, Peru, Arábia Saudita, América do Sul, Árabia e Síria (REITHINGER et al., 2007). A LC é endêmica na Síria há mais de dois séculos e a recente crise dos refugiados sírios ocasionou um surto catastrófico no “Velho Mundo”, afetando agora as centenas de milhares de pessoas que vivem em campos refugiados ou que estão retidas em zonas de conflito (DU et al., 2016). Nas Américas, uma média de 60.000 casos são diagnosticados por ano, com a Nicarágua e a Costa Rica tendo a maior incidência de LC em 2014 (RODRIGUES et al., 2018).

A LC é causada por *L. major* na África, Oriente Médio e partes da Ásia, por *L. tropica* no Oriente Médio, no Mediterrâneo e partes da Ásia, e por *L. aethiopica* em partes da África. Muitas espécies diferentes podem estar envolvidas nas Américas, onde a LC ocorre em toda a América do Sul e no extremo norte do México (DUTHIE et al., 2012).

No Brasil, a LC apresenta uma ampla distribuição, com registro de casos em todas as regiões brasileiras, afetando ambos os sexos, embora predomine em maiores de 10 anos (MINISTÉRIO DA SAÚDE, 2007). Em 2014, o Brasil registrou 19.402 casos de LC, representando 20% dos casos notificados nas Américas. Já em 2016, houve uma redução de 34% (6.705) de casos ocorridos no país (RODRIGUES et al., 2018). Existem seis espécies responsáveis por causar a LC no Brasil: *L. amazonensis*; *L. braziliensis*; *L. guyanensis*; *L. lainsoni*; *L. naiffi*; *L. shawi*. A *L. braziliensis* apresenta ampla distribuição no território nacional e é a principal espécie causadora de LC no país (BASANO; CAMARGO, 2004; COSTA-SILVA et al., 2014).

2.5 Manifestações clínicas da Leishmaniose Visceral

Também chamada de “calazar” ou “febre negra” no subcontinente indiano, a LV é uma síndrome associada a progressão e a infecção sistêmica de protozoários de *Leishmania*. Sua apresentação pode incluir ou não pápulas cutâneas no local da inoculação (GUERIN et al., 2002; GRADONI, 2018). Após um período de incubação que geralmente dura entre dois e seis meses, os pacientes com LV apresentam sintomas de infecção sistêmica persistente (incluindo febre, fadiga, fraqueza, perda de apetite e perda de peso) e sinais de invasão parasitária no sangue e no sistema retículo endotelial, com gânglios linfáticos, baço e fígado alargados. A febre é geralmente associada com rigor e calafrios, podendo ser intermitente. A fadiga e a fraqueza são agravadas pela anemia, que é causada pelo estado inflamatório persistente, hiperesplenismo (aumento da destruição periférica de eritrócitos no baço) e, às vezes, por sangramentos (AGUIAR; RODRIGUES, 2017; CHAPPUIS et al., 2007).

A LV pode ser dividida em diferentes períodos: Fase Aguda, sendo o período inicial da doença e caracterizando-se por febre, com duração inferior a 4 semanas, palidez cutâneo-mucosa e hepatoesplenomegalia (Figura 3), e esporadicamente, por tosse e diarreia; Período de Estado, onde o indivíduo apresenta febre irregular,

geralmente associada a emagrecimento progressivo, palidez e um maior aumento da hepatoesplenomegalia; Período Final, caso a doença não seja diagnosticada e tratada, quando evolui progressivamente com comprometimento mais intenso do estado geral do paciente, que pode ir a óbito (BRASIL, 2014).

Figura 3 - Hepatoesplenomegalia causada por LV



Fonte: Pace (2014)

Legenda: Hepatoesplenomegalia em criança de 20 meses com LV causada por *Leishmania infantum*.

2.6 Manifestações clínicas da Leishmaniose Cutânea

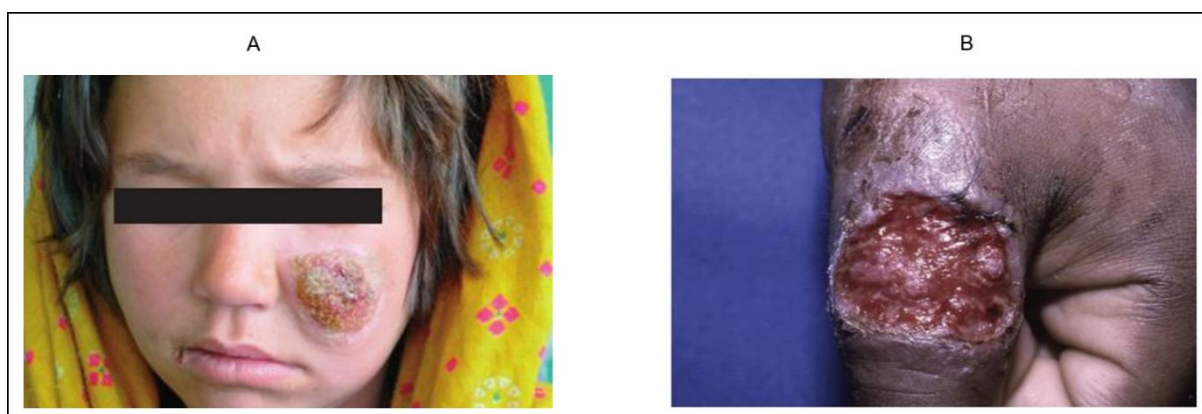
A LC, forma mais comum de leishmaniose, não apresenta risco de vida, mas pode levar à morbidade estética substancial, estigmatização social e efeitos psicológicos. É caracterizada por uma única pápula na pele na região em que o vetor infectado picou o hospedeiro e que se desenvolve após um período de latência variável, progredindo ao longo de semanas a meses (BURZA; CROFT; BOELAERT, 2018; REITHINGER et al., 2007). Durante a evolução da doença, o parasita pode atingir a corrente linfática e se difundir para outras partes do corpo, causando várias lesões cutâneas ulceradas, forma denominada de leishmaniose cutânea difusa. As lesões geralmente se apresentam de forma indolor e costumam-se localizar em áreas expostas da pele, com formato arredondado, base eritematosa, infiltrada e de

consistência firme, com bordas bem-delimitadas e elevadas, como pode ser observado na Figura 4. Caso não tratadas, as lesões tendem à cura espontânea em períodos de meses ou anos (BRASIL, 2007). A LC em humanos pode ser causada por várias espécies de *Leishmania* spp., mas a maioria das infecções permanece assintomática (GONTIJO, 2003; REITHINGER et al., 2007; GRADONI, 2018).

Na LC do Velho Mundo, as lesões podem evoluir para placas hiperkeratóticas ou semelhantes a verrugas. As lesões causadas por *L. tropica* e *L. major* são auto-cicatrizadas em um ano, mas tendem a deixar cicatrizes permanentes. Lesões por *L. aethiopica* levam anos para cicatrizar e podem evoluir para leishmaniose mucocutânea oronasal grave e formas difusas de leishmaniose cutânea (BURZA; CROFT; BOELAERT, 2018).

Aproximadamente 3% dos pacientes com LC desenvolvem lesões desfigurantes que geralmente envolvem a mucosa nasal ou orofaríngea (GIUDICE et al., 2012). Um estudo recente relatou o tropismo na mucosa nasal da infecção por *L. (Viannia) braziliensis* e um fenótipo de casos clínicos associados ao DNA do parasita na mucosa nasal, os quais eram pacientes que precisavam de mais ciclos de tratamento para alcançar a cura clínica (CANÁRIO et al., 2019)

Figura 4 - Leishmaniose Cutânea



Fonte: Reithinger et al. (2007); Kevric, Cappel e Keeling (2015).

Legenda: Em (A) observamos uma lesão causada por *L. tropica*, em Kabul, Afeganistão, em (B) apresenta-se uma lesão primária típica de LC no dedo esquerdo

2.7 Manifestações clínicas da Leishmaniose Mucocutânea

A LM é uma forma grave de leishmaniose, causada pela disseminação na mucosa facial de espécies de *Leishmania* do Novo Mundo, pertencentes ao

subgênero *Viannia*. A LM causa lesões desfigurantes e destrutivas, observadas na Figura 5, e é potencialmente fatal (HIDE et al., 2007; REITHINGER et al., 2007). As lesões por LM localizam-se na mucosa das vias aéreas superiores. Acredita-se que a lesão mucosa metastática aconteça por disseminação hematogênica ou linfática. Na maioria dos casos a LM resulta de LC não tratada ou com tratamento inadequado (MINISTÉRIO DA SAÚDE, 2007). Esta infecção é mais comumente associada a *L. braziliensis* embora vários outros agentes etiológicos tenham sido evidenciados, tais como as *L. panamensis*, *L. guyanensis*, *L. amazonensis*, *L. major*, *L. tropica* e *L. infantum*. Nesta forma da doença a identificação da espécie causadora é de grande importância para avaliação do risco de progressão de uma LC para uma LM, principalmente porque a LM geralmente se instala com maior facilidade quando o indivíduo já possui lesões ativas de LC (REITHINGER et al., 2007; GRADONI, 2018).

Figura 5 - Leismaniose Mucocutânea



Fonte: Reithinger et al. (2007)

Legenda: Destruição da mucosa causada por *L. braziliensis* em Cochabamba, Bolívia.

2.8 Resposta imune do hospedeiro e patogenia nas leishmanioses

Na infecção, os promastigotas metacíclicos de *Leishmania* obtêm acesso ao hospedeiro, onde eles encontram neutrófilos, macrófagos e monócitos (GOLLOB; VIANA; DUTRA, 2014). Uma vez que essas células são especializadas em destruir patógenos invasores e iniciar a resposta imune, as espécies de *Leishmania* tiveram

que evoluir uma gama de mecanismos para subverter a função normal dos macrófagos, garantindo sua sobrevivência e manutenção da infecção (OLIVIER; GREGORY, 2005).

O estado de saúde da pessoa infectada, incluindo fatores nutricionais e idade, o vetor de transmissão; a espécie de *Leishmania* e a constituição genética do parasita, são fatores determinantes na apresentação clínica resultante da infecção (AL-KAMEL, 2017; BURZA; CROFT; BOELAERT, 2018; KEVRIC; CAPPEL; KEELING, 2015). A interação entre esses elementos modela a resposta imunológica individual e gera o quadro clínico. Durante a infecção as formas promastigotas metacíclicas são processadas por células dendríticas e apresentadas a células T-*naïve*, que por sua vez produzem um padrão de citocinas resultando na produção de diferentes populações de células T (KEVRIC; CAPPEL; KEELING, 2015). Os dois maiores mecanismos responsáveis pelo controle dos parasitas são a produção de EROs (espécies reativas de oxigênio), gerados pela explosão respiratória que ocorre na fagocitose, e o NO (óxido nítrico), gerado pela ativação de células por INF γ (interferon gama). Embora os parasitas de *Leishmania sp.* sejam sensíveis a EROs, a explosão respiratória em macrófagos não-ativados é insuficiente para matar o parasita no fagolisossomo (SCOTT; NOVAIS, 2016). Este fato pode ser devido a atuação de algumas proteínas de virulência do parasita, quando no interior do fagolisossomo, que inibem a geração de compostos que matariam o parasita (MATHEOUD et al., 2013).

O espectro imunológico observado em pacientes com leishmaniose varia de indivíduos com HT (hipersensibilidade tardia) mediada por células T, com altos níveis de INF γ , a indivíduos que não tem uma HT, mas um alto nível de anticorpos. As células de *Leishmania* são mortas por macrófagos ativados por INF γ e não são neutralizadas por anticorpos. Indivíduos com uma forte HT, tem poucos parasitas em suas lesões, enquanto aqueles com apenas resposta humoral, são incapazes de controlar sua carga parasitária (SCOTT; NOVAIS, 2016). Foi visto que pacientes com LV produzem uma grande quantidade de anticorpos específicos, porém não protetivos contra certos antígenos do parasita. As espécies de *Leishmania* possuem moléculas que permitem estabelecer um parasitismo intracelular bem-sucedido em fagolisossomos ou vacúolos parasitóforos de macrófagos e outras células. Essas moléculas interagem com o sistema imunitário do hospedeiro favorecendo o parasita, sendo diretamente responsáveis pelos sintomas clínicos vistos nas leishmanioses (CHANG; MCGWIRE, 2002).

Em estudos comparando a infecção de *L. braziliensis* com a de *L. amazonensis*, a primeira demonstrou um eficiente desencadeamento da resposta imune em relação a *L. amazonensis*. A infecção por *L. braziliensis* leva a ativação de células dendríticas, acompanhadas de uma regulação positiva de uma das vias que controlam a atividade gênica ligada a resposta imune (JAK-STAT), o que não foi observado em *L. amazonensis*. Além disso, células dendríticas infectadas por *L. braziliensis* são mais eficientes na iniciação de células-T *naive*, que irão proliferar e produzir INF γ , e interleucina 17 (IL-17) (SOONG; MELBY, 2012).

2.9 Fatores de virulência de *Leishmania* sp.

Os parasitas causadores das leishmanioses desenvolveram diferentes estratégias para sobreviver às ações da resposta imune inata durante os primeiros momentos de infecção, bloqueando rapidamente a indução e a regulação das principais funções da célula hospedeira, tais como: produção de NO, TNF α (fator de necrose tumoral alfa), IL-12 (interleucina 12) e EROs (GÓMEZ; OLIVIER, 2016). Também atuam influenciando vias de sinalização dos macrófagos, interferindo na fosforilação e no estado de ativação de várias quinases das famílias MAPK e JAK. Essas quinases atuam sobre alguns fatores de transcrição que são críticos para a ativação de vários mediadores anti-microbianos e inflamatórios necessários para controlar os agentes infecciosos (OLIVIER et al., 2012). Diferentes fatores de virulência estão ligados às estratégias de defesa das espécies de *Leishmania* contra a resposta imune dos hospedeiros. Alguns deles estão descritos no Quadro 1.

Quadro 1 -Fatores de virulência de *Leishmania* sp. e suas respectivas funções

(continua)

Fator de Virulência	Funções
Glicoinositolfosfolídeos (GIPLs)	Ajudam espécies de <i>Leishmania</i> , incluindo <i>L. major</i> e <i>L. braziliensis</i> , a sobreviverem dentro do ambiente hostil dos macrófagos, inibindo a óxido nítrico sintetase e a proteína quinase C (PROUDFOOT, et al. 1995)

Quadro 1 -Fatores de virulência de *Leishmania sp.* e suas respectivas funções

(conclusão)

Fator de Virulência	Funções
Proteofosfoglicanos (PPGs)	São polipeptídeos altamente glicosilados. Suas funções não são totalmente elucidadas, porém como os PPGs de membrana revestem a superfície do parasita, acredita-se que eles desempenham algum papel na ligação aos receptores de macrófagos. Também podem contribuir para a manutenção do vacúolo parasitóforo. (PETER et al 1997; ILG, 2000)
Lipofosfoglicano (LPG)	É um ligante de macrófagos que está diretamente envolvido nos estágios iniciais da infecção. Ensaio realizado em <i>L. major</i> demonstraram que na ausência de genes codificantes da biossíntese de LPG, os parasitas ficam mais susceptíveis a ação do sistema complemento e a oxidantes produzidos pelas células infectadas, ficando assim sua virulência altamente prejudicada. O LPG de diferentes espécies de <i>Leishmania</i> é altamente polimórfico, onde partes da molécula podem ser substituídas em algumas espécies (SPATH et al 2000; GAUR et. al 2009; SVÁROVSKÁ et al 2010)
Proteína de membrana de cinetoplastídeo 11	É uma proteína hidrofóbica que é associada ao LPG que apresenta propriedades imunorreguladoras e é capaz de induzir a expressão de IL-10 em células de pacientes com LC e LM (JARDIM, 1995)
Proteases	São enzimas que hidrolisam as ligações peptídicas, podendo ser classificadas com base em seus domínios catalíticos em: serina-, treonina-, aspartil-, metalo- e cisteína-proteases (SILVA-ALMEIDA et al., 2012). Tais enzimas têm sido implicadas em diversas atividades parasitárias, como invasão tecidual, sobrevivência em macrófagos e modulação da resposta imune do hospedeiro(SILVA-ALMEIDA et al., 2014).

Fonte: A autora

2.10 Metaloproteases

As metaloproteases compreendem um grupo de enzimas necessárias para a proliferação celular, diferenciação, remodelação extracelular, vascularização e migração celular em processos fisiológicos e patológicos. São divididas em dois grupos: metaloproteases de matriz (MMP), tais como MMP-9 e MMP-2 produzidas principalmente por macrófagos ativados, e as desintegrinas metaloproteases, como a GP63, presente na membrana de espécies do gênero *Leishmania* (MURASE et al., 2018).

O zinco é o metal essencial para a atividade catalítica das metaloproteases. Estas são produzidas por vários microrganismos patogênicos e têm a sequência de consenso HEXXH, na qual dois resíduos de histidina funcionam como ligantes do íon zinco. As proteases produzidas por patógenos oportunistas mostram uma ampla variedade de ações patológicas, porque elas têm atividade proteolítica em diversos tipos de proteínas hospedeiras, tais como componentes da matriz extracelular e proteínas plasmáticas (HOOPER, 1994; MIYOSHI; SHINODA, 2000).

No genoma de *L. (Viannia) braziliensis*, 52% dos genes de protease correspondem a metaloproteases, porém, ainda não está completamente claro como essa diversidade favorece a infecção e a sobrevivência desses parasitas (SUTTER et al., 2017). Dentre as proteases, a que mais merece destaque é a GP63, devido ao seu crítico papel na relação parasito-hospedeiro (ISNARD et al., 2012).

2.11 GP63

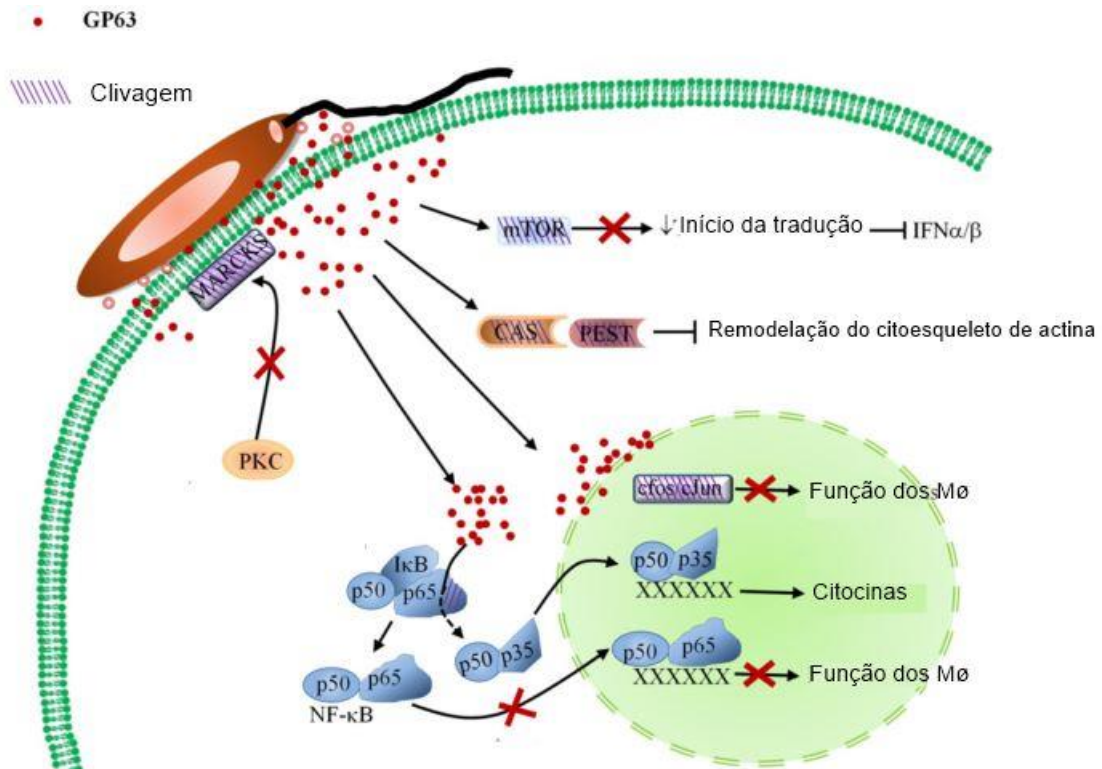
Entre os conhecidos fatores de virulência envolvidos na interação parasita-hospedeiro, as proteases GP63 são compartilhadas pelos três tipos de tripanossomatídeos mais conhecidos (*Leishmania sp.*; *T. brucei*; *T. cruzi*), embora o *T. brucei* empregue uma glicoproteína variante de superfície (VSG) como seu principal fator de virulência e o *T. cruzi* utilize as mucinas. As proteases GP63 de *T. brucei* são predominantemente expressas na forma sanguínea do parasita, e as evidências experimentais implicam que elas podem proteger o parasita da corrente sanguínea contra a lise mediada pelo complemento por prejudicar uma função de processamento de proteínas. Já em *Leishmania sp.*, a GP63 (também conhecida como

leishmanolisina) é considerada um dos principais fatores de virulência (GÓMEZ; OLIVIER, 2016; MA et al., 2011).

2.11.1 Principais funções

Desde 1980, vários papéis relevantes na patogênese da *Leishmania* são atribuídos a GP63 e é sabido que esta proteína desempenha um papel importante durante a ligação e invasão ao macrófago. A opsonização *in vivo* de promastigotas metacíclicos de *Leishmania* por C3b e C3bi permite a interação com o receptor de complemento de macrófagos 1 (CR1) e receptor do complemento tipo 3 (CR3) respectivamente. No entanto, como o C3b é rapidamente convertido em C3bi por GP63, aparentemente o CR3 é o receptor mais importante e a interação com CR1 é apenas transitória. A ligação via CR3 em vez de CR1 é vantajosa para o parasita, uma vez que não irá desencadear a explosão oxidativa durante a fagocitose (OLIVIER; GREGORY, 2005). A GP63 também se liga a outros receptores, tais como o receptor de fibronectina (JOSHI et al., 2002; MCKERROW et al., 2005; YAO; DONELSON; WILSON, 2003). O equilíbrio entre fosforilação e desfosforilação de resíduos de serina, treonina e tirosina na estrutura e regulação de proteínas quinases e fosfatases, é fundamental para o controle de mecanismos intracelulares de células eucarióticas. Formas amastigotas de *Leishmania*, com o auxílio da proteína GP63, são capazes de controlar as vias de sinalização e as funções antimicrobianas dos macrófagos hospedeiros, explorando o papel das fosfatases como reguladores negativos dessas vias. A GP63 cliva então vários alvos dos macrófagos infectados, tais como: o substrato de quinase C rico em alanina miristoilada (MARCKS), um substrato crítico para a sinalização dependente de proteína quinase C (PKC); moléculas envolvidas na ativação do citoesqueleto de actina; a quinase mTOR, que afeta o início da tradução e conseqüentemente a produção de IFN tipo 1; e fatores de transcrição, como os NF- κ B e AP – 1. A GP63 parece então atuar de diferentes formas anulando a capacidade de transcrição dos macrófagos (OLIVIER et al., 2012a). Algumas dessas vias podem ser observadas na Figura 6.

Figura 6 - Vias de sinalização dos macrófagos sendo influenciadas por GP63.



Fonte: Olivier et. al (2012, modificado)

Legenda: As setas indicam alvos GP63 envolvidos em vias de sinalização; cruzes vermelhas indicam alteração na cascata de sinalização e inibição funcional.

Outro papel fundamental da GP63 é a ativação das proteínas tirosina fostatases (PTPs) de macrófagos. Essa ativação vai resultar na inibição de diversas vias que culminariam na formação de produtos antimicrobianos, tais como NO, promovendo a sobrevivência do parasita e manutenção da doença (GÓMEZ; OLIVIER, 2016; OLIVIER et al., 2012b). Outras funções da GP63 incluem: (I) facilitação da fagocitose da forma promastigota pelos macrófagos, (II) inibição da função das células *natural-killer*, (III) resistência a peptídeos antimicrobianos e (IV) promoção da sobrevivência de amastigotas intracelulares (AVILA-LEVY et al., 2014). Em *L. tarentolae*, foi observado ainda que proteínas GP63 são responsáveis por clivar proteínas vesiculares (VAMP8) e proteínas de superfície de vesículas (SNARE), presentes em fagossomos fagocíticos. O fenômeno de clivagem impede a ligação do complexo NADPH oxidase aos fagossomos alterando as propriedades de pH e degradação dos fagossomos. Então, a apresentação dos antígenos exógenos de *Leishmania* por moléculas de MHC classe I não ocorre, diminuindo a ativação dos

linfócitos T, persistindo a infecção e reduzindo significativamente a morte parasitária (MATHEOUD et al., 2013; MURASE et al., 2018)..

2.11.2 GP63 no inseto vetor

Para realizar seu ciclo de vida, as diferentes espécies de *Leishmania* devem enfrentar várias condições adversas no hospedeiro vertebrado e mesmo no vetor invertebrado. No vetor, essas incluem a ação de enzimas digestivas, o espaço endoperitrófico e a ligação a um receptor de microvilosidades. Para contornar essas barreiras, os parasitas desenvolveram uma série complexa de glicoconjugados de superfície, conhecidos por estarem envolvidos em muitas etapas da interação parasita-hospedeiro (SOARES et al., 2017).

O desenvolvimento destes parasitas nos vetores flebotomíneos começa então com a liberação de amastigotas de macrófagos infectados ingeridos durante o repasto sanguíneo, que então se transformam em promastigotas procíclicas móveis com flagelos curtos. Estas formas crescem dentro dos insetos rodeadas por uma matriz peritrófica semi-permeável. Depois que a matriz peritrófica começa a se degradar, promastigotas procíclicas se transformam em formas nectomonadas alongadas, altamente móveis, cuja função é escapar do espaço endoperitrófico e ancorar-se ao epitélio do intestino médio. Essa ligação de promastigotas ao epitélio do intestino médio é considerada uma parte essencial do ciclo de vida do vetor de flebotomíneos, permitindo que os parasitas persistam além da fase inicial do repasto sanguíneo. Entre as moléculas de superfície de *Leishmania* que fazem parte desta e de outras etapas fundamentais para a interação parasito-hospedeiro invertebrado, duas moléculas ancoradas a GPI são as mais importantes: as LPG e as GP63 (JECNA et al., 2013; SOARES et al., 2017).

As moléculas de GP63 também desempenham funções de adesão e atividade proteolítica no hospedeiro invertebrado, atuando na aquisição de nutrientes, na migração através da matriz extracelular do inseto, e na degradação de moléculas de defesa do mesmo (AVILA-LEVY et al., 2014).

2.11.3 Organização gênica

Os genes de GP63 são genes organizados em *tandem* e não estão restritos ao gênero *Leishmania*, uma vez que estudos demonstram a presença de sequências codificantes de GP63 conservadas e funcionais em espécies de tripanossomatídeos que parasitam insetos, plantas e também em protozoários mais distantes evolutivamente como *Trichomonas vaginalis* (SMITH; PEACOCK; CRUZ, 2007).

Entre as espécies de *Leishmania* foi encontrada uma grande variação na quantidade de cópias de genes que codificam a GP63, sendo que estes foram encontrados em três cromossomos (10, 28 e 31). Um estudo realizado em *L. major* demonstrou que o conjunto de genes que codificam a GP63 possui quatro cópias no cromossomo 10, com um único gene no cromossomo 28 e um gene relacionado no cromossomo 31. Em *L. infantum* e *L. braziliensis* existem cinco e 33 genes, respectivamente, no cromossomo 10 (ISNARD et al., 2012). Em relação a *L. braziliensis* (M2904), os dados mais recentes sugerem a presença de 33 genes de GP63 no cromossomo 10, dois no cromossomo 31 e nenhum no cromossomo 28 (FUENTE et al., 2019).

Estudos com múltiplos isolados de *L. braziliensis* e *L. peruviana* (duas espécies do subgênero *Viannia*) mostraram uma variação extensiva no tamanho dos genes e evidência de rearranjos gênicos nesse locus, sugerindo que o locus GP63 é suscetível a mudanças dinâmicas neste subgênero (PEACOCK et al., 2007; SMITH; PEACOCK; CRUZ, 2007).

Do ponto de vista evolutivo, são observadas diferenças entre os genes de GP63 do subgênero *Viannia* e do subgênero *Leishmania*, em número e em estrutura (VICTOIR; DUJARDIN, 2002). Existe uma grande variedade de genes de GP63 em *L. braziliensis* localizados em sua maioria no cromossomo 10, porém o motivo da expansão desse número de genes especificamente nesse cromossomo ainda não está totalmente elucidado, mas pode ser atribuído a falta de dados estruturais para esse parasita (SUTTER et al., 2017). As espécies do gênero *Leishmania* que possuem a mais complexa organização nos genes codificantes de GP63 são *L. braziliensis* e *L. guyanensis*, ambas do subgênero *Viannia* (HAMERS et al., 1995; VICTOIR; DUJARDIN, 2002).

Buscando entender essa divergência entre a expansão desses genes no cromossomo 10, estudos de filogenia utilizando sequências proteicas de

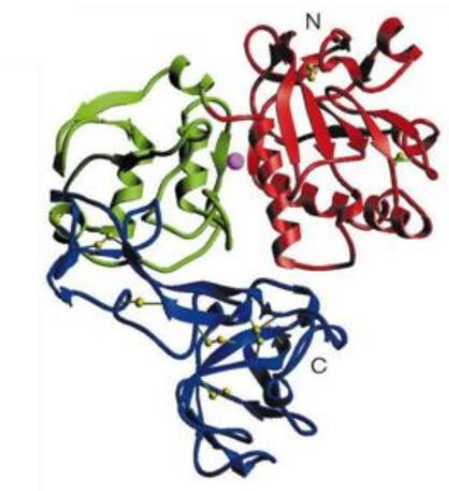
tripanossomatídeos mais distantes evolutivamente demonstraram a clara divergência de genes de GP63 mapeados no cromossomo 10 de *Leishmania*, dos genes mapeados nos cromossomos 28 e 31. Também observou-se que os genes nos cromossomos 28 e 31 de *Leishmania* provavelmente surgiram antes das espécies ancestrais colonizarem vertebrados e podem não estar envolvidos em processos patogênicos no hospedeiro mamífero (CASTRO NETO et al, 2019).

2.11.4 Estrutura

A GP63 é uma metaloprotease de zinco ancorada a membrana plasmática por meio de uma molécula de glicofosfatidilinositol (GPI). Esta enzima pertence à família de peptidases M8, e estima-se que cada célula promastigota tem 5×10^5 moléculas desta proteína. A estrutura tridimensional da GP3 foi resolvida (PDB ID: 1L ML) (SCHLAGENHAUF; ETGES; METCALF, 1998) e revelou uma estrutura de três domínios (N-terminal, região central e C-terminal), que podem ser observados na Figura 7. O domínio N-terminal contém uma região ligante de zinco correspondente aos sítios catalíticos das metaloproteases de zinco, com o motivo de sequência HEXXHXXGXXH. Diferente de outras proteases desta classe, entretanto, existe um inserto de 62 aminoácidos entre a glicina e o terceiro resíduo de histidina (SUTTER et al., 2017; YAO; DONELSON; WILSON, 2003).

As GP63 partilham de alta identidade de sequência entre si, embora já foram encontradas tanto formas anfífilas, como formas hidrofílicas. Elas diferem principalmente em suas regiões codificantes C-terminais e nas regiões 3' não traduzida (3'-UTRs) dos seus mRNAs (YAO; DONELSON; WILSON, 2003).

Figura 7 - Estrutura tridimensional da proteína GP63



Fonte: Schlagenhauf et. al (1998, adaptado)

Legenda: Representação da estrutura da molécula de GP63, com o domínio C-terminal mostrado em azul, o domínio N-terminal mostrado em vermelho e o domínio central mostrado em verde. Ligações dissulfeto estão em amarelo. O átomo de zinco no sítio ativo está representado como uma esfera em rosa.

2.11.5 Variações na região C-terminal

A maioria das proteínas GP63 possuem diferenças entre suas regiões C-terminais. Essas diferenças, tais como deleções e extensões podem alterar a ligação proteica, se esta região for um domínio trans-membranar (MAURICIO et al., 2006; RAMAMOORTHY et al., 1992; VOTH et al., 1998). A região C-terminal, caso deletada, diminui em aproximadamente 50% a atividade catalítica da enzima (MAZUMDER; GANGULY; ALI, 2010).

Utilizando-se de alinhamentos de sequências proteicas codificadas por genes do cromossomo 10, três grupos de proteínas foram identificados no total, para *L. major* e *L. infantum*, com possíveis diferenças de expressão (KELLY; NELSON; MCMASTER, 2001), relacionadas a diferenças na sua região C-terminal. Em *L. braziliensis* esta análise revelou seis grupos de genes de GP63 no cromossomo 10 compatíveis com alterações funcionais que podem estar ligadas à sua localização na membrana ou liberação em meio extracelular (CASTRO NETO et al., 2019).

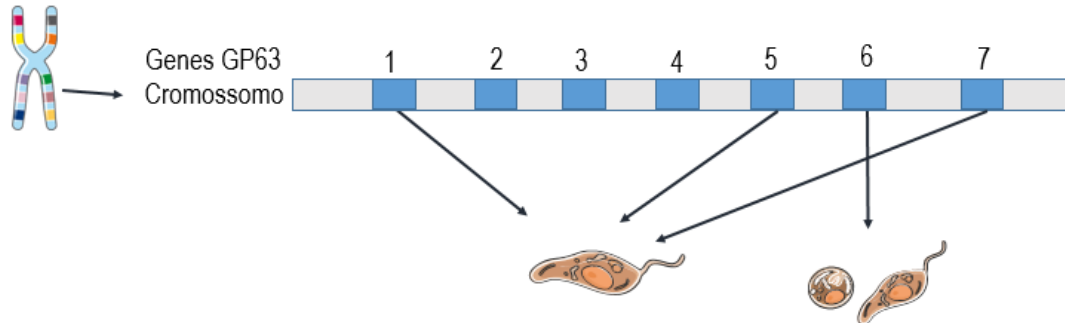
2.11.6 Expressão

A síntese de GP63 é realizada no retículo endoplasmático, onde sua sequência-sinal N-terminal é clivada e os resíduos de aminoácidos C-terminais são substituídos por uma estrutura de membrana de ligação a âncora GPI. O processamento proteolítico subsequente envolve a clivagem de um pró-peptídeo para produzir a cadeia polipeptídica madura (SCHLAGENHAUF; ETGES; METCALF, 1998). Em média, 500.000 cópias de GP63 são expressas por célula, o que representa cerca de 1% do total de proteínas celulares (NOVOZHILOVA; BOVIN, 2010). Estudos demonstraram que a expressão de produtos dos genes de GP63 variam de acordo com a fase de crescimento e da forma do ciclo de vida de *Leishmania*. A grande quantidade de genes de GP63 geram proteínas abundantes que variam entre espécies de *Leishmania* levando a diferentes efeitos biológicos (ISNARD et al., 2012).

Uma cepa de *L. braziliensis* virulenta demonstrou um perfil proteolítico complexo e distinto de uma cepa de *L. braziliensis* não-virulenta. Diferenças encontradas na produção de proteases estão possivelmente associadas ao mecanismo de infecção e virulência de *L. braziliensis*. A atividade proteolítica em cepas de diferentes padrões de virulência pode estar diretamente relacionada com a participação dessas enzimas durante a infecção no hospedeiro (LIMA et al., 2009).

Estudos feitos em *L. major*, demonstraram que os genes denominados GP63 1-5 e 7, são estágio-específicos, sendo altamente expressos somente na fase promastigota, já o gene 6 é constitutivamente expresso em ambas as fases (Figura 8), porém em uma quantidade menor, sendo a GP63 uma proteína preferencialmente expressa no estágio promastigota (SCHNEIDER et al., 1992; VOTH et al., 1998). Foi visto que moléculas de GP63 são também expressas pela forma amastigota de *L. major* (FROMMEL et al., 1990).

Figura 8 - Expressão de genes GP63 em *L. major*



Fonte: À autora

Legenda: Na imagem podemos observar um segmento de cromossomo, com os genes de GP63 em organizados em *tandem*, onde os genes 1, 5 e 6 são expressos somente na fase promastigota, enquanto o gene 6 está sendo expresso constitutivamente (fase promastigota e fase amastigota).

As regiões não traduzidas dos mRNAs e as regiões intergênicas dos genes organizados em *tandem*, na maioria das vezes, regulam a expressão gênica em tripanossomatídeos. Visto isso, em *L. major* foi demonstrado que a expressão diferencial de GP63 em amastigotas e promastigotas é regulada por sequências únicas preditas nas regiões 3' não traduzidas dos mRNAs de GP63 (YAO; DONELSON; WILSON, 2003).

Foi demonstrado em *L. donovani chagasi* que existem três grupos distintos de genes de GP63, classificados de acordo com o padrão de expressão dos seus mRNAs (mRNAs de fase log, fase estacionária e constitutivos) e contendo regiões 3' não traduzidas únicas específicas em cada fase. Os RNAs de fase log e estacionária possuem regiões codificantes idênticas. Já os RNAs constitutivos, possuem um C-terminal com uma adição de 41 aminoácidos, que substitui um sítio de ligação a âncora GPI por uma sequência que pode ser um domínio transmembrana (RAMAMOORTHY et al., 1992).

2.11.7 Classificação dos genes de GP63 em *L. braziliensis*

Para se entender melhor a real diversidade e função dos inúmeros genes de GP63 localizados no cromossomo 10 de *L. braziliensis*, em um estudo prévio foi realizada uma análise comparativa entre estes genes e aqueles no mesmo

cromossomo de *L. major* e de *L. infantum*. Nesta análise foram considerados apenas três dos grupos de identificados de GP63 de *L. braziliensis* uma vez que os demais apresentavam truncagens na sua extremidade C-terminal. Destes três grupos, com extremidades C-terminais nas respectivas proteínas claramente distintas, um destes se mostrou relacionado aos genes expressos na fase estacionária em *L. infantum* enquanto um segundo, com grande expansão em números de genes, se mostrou relacionado ao gene de *L. infantum* expresso constitutivamente. O terceiro grupo de *L. braziliensis*, também expandido em número de cópias, não parece ter equivalente em *L. infantum* no que concerne ao sua extremidade C-terminal (CASTRO NETO et al., 2019). Esses resultados reforçam a existência de características únicas nos genes de *L. braziliensis* com expansão em número de cópias e levantam alguns pontos relevantes para investigação. O primeiro deles seria qual a diferença entre esses três grupos de genes no que concerne a sua expressão e localização subcelular, tendo visto as diferenças observadas nas extremidades 3' não traduzidas nos mRNAs e nos sinais de localização/secreção nas extremidades C-terminais das proteínas. O segundo diz respeito a existência ou não de diferentes padrões de expressão para os múltiplos genes classificados dentro de um mesmo grupo no cromossomo 10. Este trabalho buscou avançar nessa investigação.

3 PERGUNTA CONDUTORA

Com base na existência de múltiplos genes parálogos de GP63 localizados no mesmo cromossomo e de estrutura similar, existe variação no perfil de expressão das diferentes proteínas GP63 parálogas ao longo do ciclo de vida da *L. braziliensis*?

4 HIPÓTESE

Diferentes proteínas GP63 são expressas diferencialmente em fases específicas do ciclo de vida de *L. braziliensis* e essa expressão pode variar de acordo com aspectos diferenciados da estrutura dessas proteínas e como reflexo das necessidades do parasita para sua sobrevivência.

5 OBJETIVO GERAL

Avaliar de forma comparativa o reconhecimento por soro policlonal específico e a expressão de diferentes proteínas parálogas GP63 de *L. braziliensis*, codificadas por genes localizados no seu cromossomo 10.

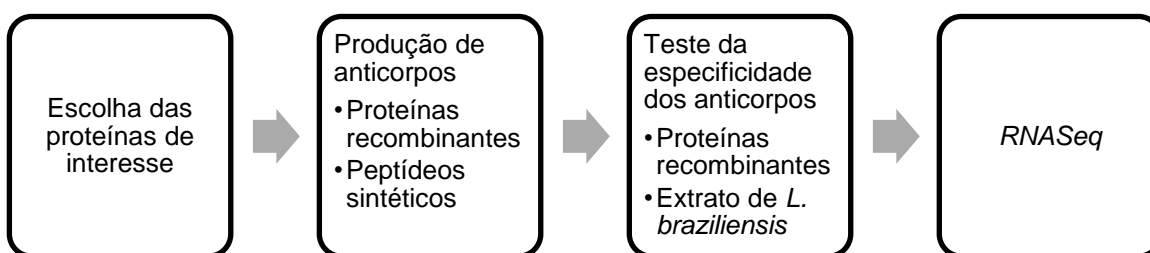
5.1 Objetivos Específicos

- a) Avaliar o reconhecimento e especificidade de múltiplos lotes de anticorpos gerados contra diferentes parálogos de GP63 de *L. braziliensis*;
- b) Analisar, de forma comparativa, a expressão de parálogos selecionados ao longo de curvas de crescimento da forma promastigota do parasita;
- c) Comparar a expressão de transcritos de GP63 em fase logarítmica e estacionária de crescimento de promastigotas de *L. braziliensis*.

6 METODOLOGIA

Para a produção de anticorpos foram selecionadas através de alinhamentos e análises filogenéticas, proteínas e peptídeos GP63 com sequências de aminoácidos divergentes. Foram utilizadas duas metodologias para a produção de anticorpos para as proteínas GP63 em questão. A primeira utilizando proteínas recombinantes e a segunda utilizando-se de peptídeos sintéticos conjugados a molécula KLH. Os anticorpos foram testados frente a proteínas recombinantes e frente a extratos de *L. braziliensis* em diferentes fases de crescimento da forma promastigota do parasita. Para analisar qual das proteínas GP63 estava sendo mais expressa na fase promastigota, foram feitas análises de RNASeq (Figura 9).

Figura 9 - Esquema metodológico



Fonte: A autora

6.1 Alinhamento e análise filogenética

As sequências selecionadas disponíveis em bancos de dados, juntamente com sequências obtidas por PCR de estudos anteriores, foram alinhadas no programa *MAFFT version 7* (KATO et al., 2002). Em seguida, esse alinhamento foi editado pelo Trimal (CAPELLA-GUTIERREZ, SILLA-MARTINEZ, GABALDON, 2009), para manter somente os sítios filogenéticos informativos. Foi-se utilizado para construir a árvore filogenética o programa PhyML 3.0 (GUINDON et al., 2010) e usando o método de máxima verossimilhança e o MrBayes (HUELSENBECK; RONQUIST, 2001), aplicando o método bayesiano. O suporte dos ramos para as árvores foi dado por análises de *bootstrap* usando 1000 replicatas.

6.2 Subclonagem de proteínas GP63 recombinantes

Foram feitas subclonagens utilizando construções plasmidiais no vetor pGEMT-easy contendo os genes de GP63 de interesse já obtidas previamente pelo grupo. As construções e o plasmídeo de expressão pET21a(+) (Novagen) foram digeridas com enzimas de restrição *Bam*HI/ *Xho*I, para obtenção dos fragmentos de DNA correspondentes. Os fragmentos obtidos foram avaliados por meio de eletroforese em gel de agarose 0,8%, corado com *Syber*® *Safe* (Invitrogen®) e visualizados em aparelho emissor de ultravioleta de ondas longas. As bandas de interesse foram excisadas e purificadas utilizando-se o sistema de purificação GFX (GE Helthcare®) conforme recomendação do fabricante.

A partir do produto da digestão purificado e quantificado no *NanoDrop 2000, UV-Vis Spectrophotometer* (Thermo Scientific), foram realizadas ligações com a enzima T4 DNA ligase (*NEW ENGLAND Biolabsinc.*), a 16°C por 18 horas. Reações contendo apenas o vetor linearizado, ou seja, sem os genes, foram utilizadas como controles negativos da ligação. As construções foram transformadas em *Escherichia coli* cepa TOP10 quimiocompetentes, através de choque térmico (30 minutos imersas em gelo e 5 minutos em banho-maria à 37°C) em solução tamponada de cloreto de magnésio.

Estas células, foram plaqueadas em meio LB (Luria-Bertani) sólido contendo ampicilina na concentração de 100 µg/mL, e foram colocadas em estufa bacteriológica por 16 horas a 37 °C. Após análise das placas, colônias de cada gene foram escolhidas aleatoriamente, isoladas e inoculadas em tubos contendo 3 mL de meio LB líquido com ampicilina (100 µg/mL). Os inóculos foram então colocados sob agitação constante (180 rpm) à 37°C por aproximadamente 16 horas. Após o crescimento, foram retiradas alíquotas de 1,5 mL dos inóculos, para obtenção de um sedimento através de centrifugação à 12.000 × g por 3 minutos. A partir do sedimento foi realizado o a extração de DNA plasmidial por lise alcalina, e a análise dos plasmídeos obtidos realizada em gel de agarose a 1%.

A confirmação das subclonagens foi feita através da análise pela digestão (usando as enzimas *Bam*HI e *Xho*I) e sequenciamento das minipreparações, realizado por método de sanger automático, no equipamento ABI Prism 3100 *Genetic Analyzer* (*Applied Biosystems*) do Núcleo de Plataformas Tecnológicas (NPT) do IAM-FIOCRUZ.

6.3 Expressão de proteínas GP63 recombinantes

Para a expressão das proteínas recombinantes o vetor de expressão pET21a(+) (Novagen), contendo os fragmentos gênicos de interesse e o vetor vazio (para o controle de expressão) foram transformados em cepas de células BL21 star (Invitrogen®) e Rosetta (Novagen) de *E. coli* através de transformação por choque térmico. Para a transformação, as bactérias foram incubadas por 30 minutos a 4°C, seguido de choque térmico a 37°C por 5 minutos, com subsequente semeio das células transformadas em placa de Petri contendo meio Agar LB sólido na presença de ampicilina (100 µg/µL) e incubadas à 37°C por 18 horas. As colônias obtidas foram inoculadas em 2 mL de meio LB líquido contendo antibiótico, e foram incubadas sob agitação constante (180 rpm) à 37°C por aproximadamente 16 horas. Para indução da expressão das proteínas recombinantes, alíquotas de 2 mL das culturas crescidas foram inoculadas em 20 mL de meio LB com antibiótico e incubadas à 37°C sob agitação constante (180 rpm) e o crescimento bacteriano foi acompanhado pela aferição em espectrofotômetro (comprimento de onda de 600 nm) até atingir uma densidade ótica (D.O.) entre 0,5 nm e 0,8 nm. Em seguida foi adicionado o indutor de síntese de proteínas, IPTG (isopropil-β-D-tio-galactosídeo - Sigma-Aldrich®), na concentração final de 0,1 mM e a cultura permaneceu sob agitação por 4 horas à 30°C.

A visualização das expressões foram realizadas através de eletroforese em gel de *SDS-PAGE* 15% (15x8cm) (VESTERBERG, 1993). Após a expressão em pequena escala, foi feita uma expressão em grande escala, com um inóculo de 20 mL crescendo em 500 mL de meio nas mesmas condições descritas anteriormente.

6.4 Purificação das proteínas GP63 recombinantes

As culturas bacterianas foram centrifugadas a ~10000 × g durante 10min e o sedimento bacteriano foi ressuspenso em 20 mL de tampão de lise e equilíbrio (Fosfato de sódio 100 mM, Tris 10 mM, Uréia 8 M, Imidazol 20 mM – pH 8,0). Em seguida a lise das bactérias foi feita por ultra-sonicação (6 pulsos de 30 segundos com intervalos de 1 minuto). Para a purificação foi utilizada a resina Ni-NTA-Agarose (Qiagen) através de ligação de coordenação do cátion Ni²⁺ imobilizado das histidinas presentes na proteína recombinante. O produto sonicado foi incubado com 400 µL da

resina por 1 hora à 4°C. Após centrifugação à 5000 rpm por 10min o sobrenadante foi guardado, e foi adicionado 1 mL de tampão de lise e equilíbrio a resina. A solução foi centrifugada à velocidade máxima por 30 segundos. O sobrenadante foi descartado e o último processo foi repetido. Em seguida foi colocado 1 mL do tampão de lavagem (Fosfato de sódio 100 mM, Tris 10 mM, Uréia 8 M, Imidazol 30 mM – pH 6,0) no tubo contendo a resina, o tubo foi agitado no vórtex e incubado sob agitação lenta à 4°C por 5 min, após isso a solução foi centrifugada em velocidade máxima por 30 segundos. O processo de lavagem foi repetido, e em ambas as vezes o sobrenadante foi guardado à -80°C. Foram adicionados 500 µL do tampão de eluição (100 mM fosfato de sódio, 10 mM Tris, 8 M de Uréia, 1 M de imidazol – pH 4,5) nas resinas, o tubo foi agitado no vórtex e incubado sob agitação lenta à 4°C por 30 minutos. O processo de eluição foi repetido, e em ambas as vezes o sobrenadante foi guardado à -80°C.

6.5 Imunização em coelhos

Antes da imunização, foram coletados aproximadamente 1,5 mL de sangue da orelha de cada coelho New Zealand, para o controle pré-imune dos experimentos. Após a coagulação, este sangue foi centrifugado a 5000 rpm por 10 min para obtenção do soro que foi armazenado à -80 °C. Todas as imunizações foram feitas em duplicata. O presente estudo possui aprovação pela Comissão de Ética no Uso de Animais (CEUA) do Instituto Aggeu Magalhães, sob o protocolo nº 93/2015 (ANEXO B).

6.5.1 Imunização com a proteína 10.0540 fusionada a poli-histidina

Após quantificação da proteína presente no produto sonificado por curva de BSA, cerca de 150 µg de proteínas fusionadas a sequência de poli-histidinas foram submetidas a uma corrida eletroforética em gel *SDS-PAGE* 10% e após a visualização com *Comassie Blue*, o fragmento de interesse foi excisado. Estes fragmentos foram então macerados e foram adicionados 200µL de adjuvante completo de Freund e 300 µL de PBS (NaCl 137 mM, KCl 2.7 mM, Na₂HPO₄ 10 mM, KH₂PO₄ 1.8 Mm). Esta mistura foi inoculada por via subcutânea no dorso de coelhos em múltiplos pontos de aplicação. Nas imunizações seguintes, foi utilizado adjuvante incompleto de Freund.

O processo foi composto de quatro imunizações, com intervalo de quinze dias entre cada uma delas.

6.5.2 Imunização do peptídeo conservado conjugado ao KLH

Peptídeos sintéticos foram produzidos pela empresa “GenOne – Soluções em Biotecnologia”. O peptídeo conservado foi encaminhado a empresa “Célula B – Serviço de produção de anticorpos do Centro de Biotecnologia do Estado do Rio Grande do Sul” para a conjugação a molécula KLH. Para cada imunização foram utilizados 250 µg de peptídeo, 200 µL de adjuvante completo de Freund (para a primeira imunização) ou adjuvante incompleto de Freund (para as demais imunizações) e 3500 µL de PBS. Esta mistura foi inoculada por via subcutânea no dorso de coelhos em múltiplos pontos de aplicação. O processo foi composto de quatro imunizações, com intervalo de quinze dias entre cada uma delas.

6.6 Obtenção comercial de anticorpos

Três peptídeos (10.0520, 10.0590, 10.1610) foram encaminhados a empresa “Célula B – Serviço de produção de anticorpos do Centro de Biotecnologia do Estado do Rio Grande do Sul” para além da conjugação a molécula KLH, a produção de anticorpos em coelhos. Os soros recebidos e a amostra pré-imune foram aliquotados e armazenado à -80°C.

6.7 Purificação dos anticorpos por imunoadsorção

Foram utilizadas duas metodologias para a purificação dos anticorpos. Uma utilizando os respectivos peptídeos conjugados a molécula BSA e outra utilizando proteínas recombinantes. Para a purificação dos anticorpos policlonais anti-peptídeo conservado, anti-peptídeo 10.0590 e 10.1610, cerca de 250 µg dos respectivos peptídeos conjugados ao BSA, foram fracionados em gel *SDS-PAGE* 20% e transferidos para membrana de transferência de PVDF (Fluoreto de Polivinilideno, Immobilon-P, Millipore Corporation) e cada membrana, após transferência, corada com *Ponceau S* 0,2% (Sigma-Aldrich) em solução de ácido tricloroacético (1%). As bandas correspondentes aos peptídeos de interesse foram excisadas da membrana,

cortadas em pequenos pedaços e transferidas para tubos eppendorf. Os fragmentos de membrana foram lavados três vezes com 1 mL de PBS, por 10 minutos cada, e bloqueados com solução de leite 5% em PBS-Tween 20 durante 30 minutos a 4°C. A solução bloqueadora foi removida e a membrana lavada rapidamente uma vez com PBS. Foi então adicionados 750 µL do soro obtido da imunização e 500 µL de PBS e os fragmentos de membrana permaneceram incubando por 24 horas, sob agitação constante, a 4°C. Ao término da incubação, os fragmentos de membrana foram então lavados três vezes com 1 mL de PBS-Tween 20 por 10 minutos cada e os anticorpos adsorvidos às membranas foram eluídos com 200 µL de solução de glicina ácida (glicina-HCl 0,1 M, pH 2,5) por agitação vigorosa durante cinco minutos à temperatura ambiente. Posteriormente, a solução de glicina com os anticorpos eluídos foi recuperada para um novo tubo eppendorf e o pH da solução equilibrado pela adição de 20 µL de Tris-HCl (1M, pH 8,0). Os anticorpos purificados foram estocados a -80°C. Já para a purificação do anticorpo policlonal anti-peptídeo 10.0520 obtido a partir dos soros obtidos por imunização em coelhos, cerca de 100 µg de proteína recombinante purificada 10.0520 fusionada a GST, foram fracionadas em gel *SDS-PAGE* 20% e transferidas para membrana de transferência de PVDF (Fluoreto de Polivinilideno, Immobilon-P, Millipore Corporation) e a membrana, após transferência, corada com *Ponceau S* 0,2% (Sigma-Aldrich) em solução de ácido tricloroacético (1%). As bandas correspondentes as proteínas de interesse foram excisadas da membrana, cortadas em pequenos pedaços e transferidas para tubos eppendorf. A partir dessa etapa, a purificação seguiu com os mesmos passos já citados anteriormente para o outro anticorpo.

6.8 Cultivo celular de formas promastigotas e curvas de crescimento

A manutenção das formas promastigotas de *Leishmania braziliensis* foi feita em meio Schneider pH 7,2 suplementado com 20% de soro fetal bovino (SFB). Foram realizados inóculos de 10⁶ células em meio Schneider pH 7,2 e o crescimento celular foi acompanhado diariamente até a fase de morte celular, todos os dias eram retiradas alíquotas para obtenção de extratos celulares para os ensaios de WB.

6.9 Obtenção de extrato celular

Os extratos totais do parasita, utilizados para os ensaios de *Western blot*, foram obtidos a partir de culturas quantificadas, em fase exponencial de crescimento (10^7), das quais as células foram coletadas por centrifugação ($1000 \times g$ / 5 minutos / 4°C) e o sedimento foi lavado com PBS (NaCl 137 mM, KCl 2.7 mM, Na_2HPO_4 10 mM, KH_2PO_4 1.8 mM) e em seguida ressuspendido em 100 μL de tampão de proteína SDS-desnaturante (Laemmli 2X).

6.10 Ensaios de *Western blot*

Para os ensaios de *Western blot*, extratos protéicos de *E. coli* provenientes da indução das proteínas recombinantes de GP63 e extratos celulares de *L. braziliensis* foram fracionados em gel *SDS PAGE* 15% e transferidos para membranas de PVDF (Fluoreto de Polivinilideno, Immobilon-P, Millipore Corporation) em sistema de transferência semi-seco.

Na primeira etapa do *Western blot*, as membranas foram bloqueadas em solução de leite 5% em TBS/Tween (20mM Tris, 500mM NaCl, Tween 20 1% pH 7,5) por 1 hora. Posteriormente foram incubadas *over-night* com os soros contra as diferentes proteínas em estudo, e com os anticorpos purificados a diluições variadas. As membranas, após a incubação, foram lavadas, por 10 minutos, três vezes com TBS/Tween 20 1%. Em seguida, uma nova incubação de 1 hora foi realizada com o segundo anticorpo (anti-IgG de coelho, Jackson Immunoresearch Laboratories), marcado com peroxidase, numa diluição de 1:10.000. Decorrido esse tempo, as membranas foram novamente lavadas três vezes com TBS/Tween 20 1% por 10 minutos. Após as lavagens, as membranas foram banhadas em luminol 1,2 nM (Sigma-Aldrich), iodofenol 0,4 mM (Sigma-Aldrich) e peróxido de hidrogênio 0,03% (Vetec) por 2 minutos para realização da reação de quimioluminescência. Em seguida, foram expostas a um filme de quimioluminescência durante 5 minutos. O filme foi revelado com solução Dektol (Kodak) por 3 minutos e fixado em solução de ácido acético por 1 minuto.

6.11 Extração de RNA total para análise de RNASeq

O RNA de parasitas na fase Logarítmica e Estacionária foi extraído utilizando o RNeasy Mini Kit (Qiagen) de acordo com as instruções do fabricante e as amostras resultantes foram armazenadas a -80 °C. O processo foi executado em triplicata para todas as amostras.

6.12 Avaliação quantitativa do RNA e preparação da biblioteca de cDNA

Para avaliação quantitativa do RNA extraído foi realizada a quantificação de todas as amostras em aparelho fluorímetro Qubit[®] 2.0 (Thermo Fisher) com o reagente RNA HS Assay Kit (*Thermo Fisher*) e foram utilizadas para biblioteca de DNA complementar (cDNA) as amostras que tinham um total de mais de 100 ng de RNA. A biblioteca de cDNA foi obtida através do TruSeq Stranded mRNA Library Prep Kit (illumina Inc.) utilizando rigorosamente as instruções do fabricante. Resumidamente esse processo consiste na purificação do mRNA através cauda poliA, utilizando-se beads magnéticas ligadas a oligos poli-T, seguida da fragmentação dos mRNAs purificados. Posteriormente foi feita a síntese do cDNA utilizando hexâmetros randômicos e a enzima Superscript II (Invitrogen). O material gerado a partir dessa reação foi purificado utilizando-se esferas magnéticas AMPure XP beads (Beckman Coulter). Por fim para obtenção da biblioteca foram então ligados os adaptadores em ambas extremidades do cDNA e executada uma PCR para um enriquecimento seletivo dos fragmentos obtidos. O sequenciamento das amostras foi realizado com o cartucho MiSeq[®] Reagent Kit v3 (illumina Inc.) de 150 ciclos em aparelho MiSeq (illumina Inc.).

6.13 Checagem da qualidade das sequências obtidas na análise de RNASeq

As leituras obtidas tiveram sua qualidade checada pela ferramenta FastQC versão 0.11.5 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc>), esse programa oferece um conjunto de análises para cada biblioteca onde avalia sua qualidade, sendo capaz de identificar problemas na etapa de construção ou sequenciamento, trazendo diagnóstico sobre a qualidade das bases, quantidade de

bases não identificadas, comprimento médio das sequências, identificação de sequências duplicadas ou superexpressas, além de outros parâmetros.

6.14 Trimagem e mapeamento das leituras obtidas na análise de RNASeq

Após a verificação da qualidade das leituras, foi realizada a trimagem para remoção das sequências de adaptadores também eliminação das sequências ruins identificadas. Para isso foi utilizado o programa Trimmomatic versão 0.36 (<http://www.usadellab.org/cms/?page=trimmomatic>) onde foram retiradas as bases de baixa qualidade (Phred < 20) e com tamanho menor que 40 pb.

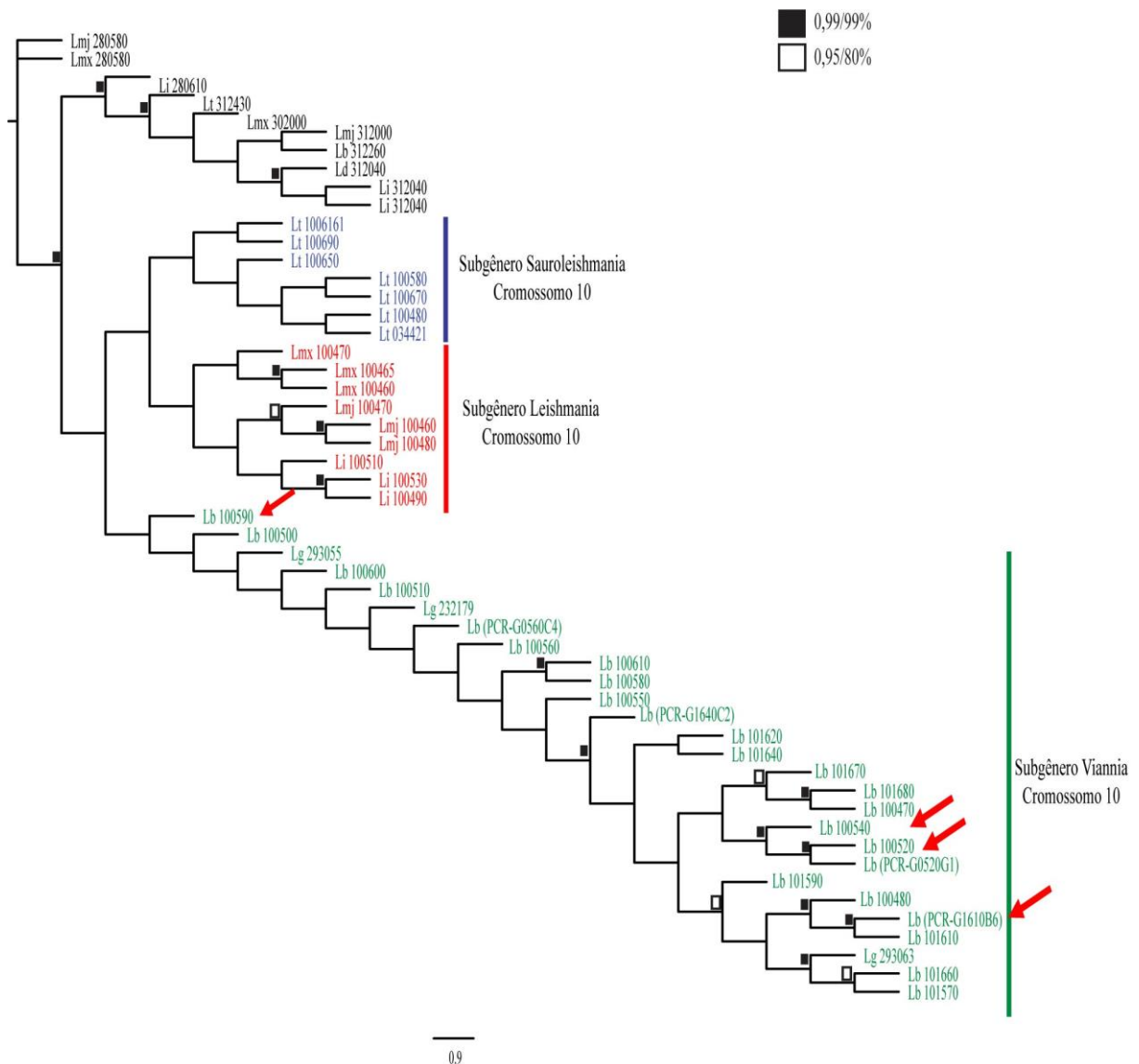
Para o mapeamento das sequências foi utilizado o programa Bowtie 2 versão 1.2.1 (LANGMEAD; SALZBERG, 2012), utilizando como referência o genoma disponível na plataforma TriTrypDB *Leishmania braziliensis* M2904 (MHOM/BR/75M2904) e também utilizando sequências de PCR obtidas por estudos anteriores do grupo. Para essa etapa foram utilizados os parâmetros padrões do *software*. Por fim a contagem das leituras mapeadas foram obtidas utilizando o pacote HTSeq Python versão 0.6.1p2 (<http://www-huber.embl.de/users/anders/HTSeq/doc/index.html>) utilizando a opção `stranded = yes` (indicando que as leituras foram consideradas dependentes do fato de serem mapeadas pela fita oposta) os outros parâmetros utilizados seguiram o padrão estabelecido pelo *software*. Os dados foram analisados na plataforma RStudio, onde foram construídos heatmaps com dados normalizados. O mapeamento foi visualizado utilizando o programa IGV versão 2.4

7 RESULTADOS

7.1 Avaliação da diversidade das sequências dos genes de GP63 no cromossomo 10 de *Leishmania sp.*

Este trabalho teve como foco principal o estudo de parálogos de GP63 do cromossomo 10 de *Leishmania braziliensis*, buscando entender o por que do número elevado de parálogos dessa proteína apenas neste cromossomo, contribuindo para o conhecimento sobre a patogenicidade do parasita. Como etapa preliminar, foi inicialmente considerada a expansão desses genes em linhagens de *Leishmania sp.*, com ênfase em *L. braziliensis* e outras espécies de *Viannia*. Análises filogenéticas prévias puderam separar os genes de GP63 mapeados no cromossomo 10, dos genes mapeados no cromossomo 28 e 31, indicando que esses últimos genes estejam relacionados ao estágio do inseto no ciclo de vida do parasita, enquanto os do cromossomo 10 estariam mais relacionados a infecção no hospedeiro vertebrado (CASTRO NETO et al., 2019) (ANEXO C). A partir disso foi feita uma análise com sequências do cromossomo 10 de espécies relevantes de *Leishmania* usando como grupos externos os genes selecionados dos cromossomos 28 e 31. Para esta análise, também incluímos sequências de *L. tarentolae*, onde ocorre uma expansão semelhante no cromossomo 10. A relevância em incluir as sequências de *L. tarentolae* deve-se ao fato de que essa espécie não parasita mamíferos, apenas lagartos, o que significa que qualquer papel potencial na patogênese associada aos genes GP63 do cromossomo 10 não é específica para seus hospedeiros mamíferos. Nesta análise também foi incluída novas sequências de GP63 de *L. braziliensis* geradas previamente pela autora desta dissertação através da técnica de PCR. A árvore filogenética mostrada na Figura 10 resume os resultados destas análises com base em alinhamentos usando as sequências de comprimento total para todas as proteínas (ou os fragmentos de PCR de comprimento total). Para maior clareza, apenas as sequências representativas mais divergentes foram usadas para construir essa árvore.

Figura 10 - Filogenia dos genes de GP63



Fonte: Castro Neto et al. (2019).

Nota: Árvore filogenética comparando múltiplos genes GP63 do cromossomo 10 de espécies selecionadas de *Leishmania*. A árvore destaca a separação dos genes GP63 de acordo com os três principais subgêneros de *Leishmania* (*Sauroleishmania*, *Leishmania* e *Viannia*). Os valores para nós altamente suportados foram substituídos por quadrados pretos e brancos, o que representa as probabilidades Bayesianas e o suporte de *bootstrap* para PhyML, respectivamente.

Legenda: As setas em vermelho indicam os genes que foram utilizados para a produção de soros. Lb: *Leishmania braziliensis* M2904, Lmj: *Leishmania major*, Lmx: *Leishmania mexicana*, Li: *Leishmania infantum*, Lt: *Leishmania tarentolae*, Lg: *Leishmania guyanensis*.

Dentro de cada um dos três subgêneros de *Leishmania* analisados, todas as sequências de GP63 do cromossomo 10 estão mais intimamente relacionadas com sequências da mesma espécie ou espécies relacionadas em relação as sequências encontradas em espécies pertencentes aos outros subgêneros. Mesmo dentro dos

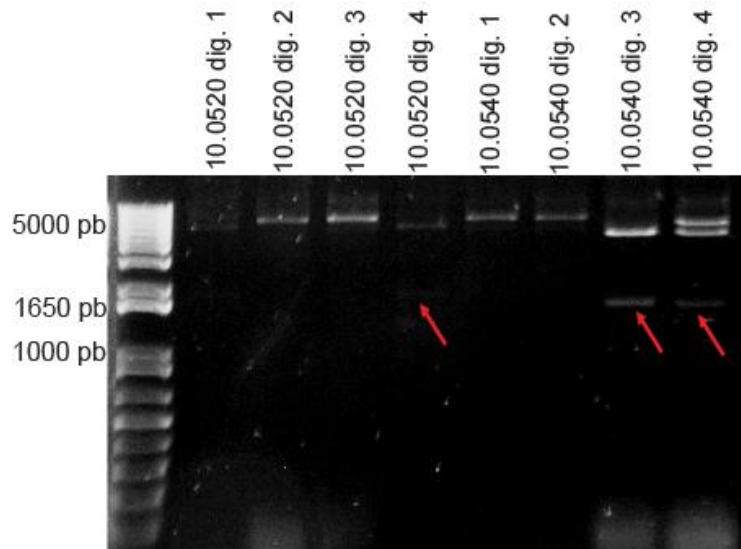
clados do subgênero *Leishmania*, os genes de *L. infantum* (em vermelho) pareciam estar mais intimamente relacionados entre si, em relação aos seus homólogos em *L. major*. Para as duas espécies de *Viannia* analisadas (*L. braziliensis* e *L. guyanensis*) foi visto que os genes estão relacionados entre si. Estes resultados estão de acordo com expansões independentes no número de sequências de GP63 do cromossomo 10 em cada clado, com grandes expansões ocorrendo para os subgêneros *Sauroleishmania* (em azul) e *Viannia*. Para as espécies do subgênero *Viannia*, o início dessa expansão pode ter precedido a divisão entre *L. braziliensis* e *L. guyanensis*, mas posteriormente continuou e pode ser um processo contínuo.

Em sequência, a avaliação de diferenças funcionais associadas aos múltiplos parálogos de GP63 identificados no cromossomo 10 de *L. braziliensis* depende da capacidade de se detectar a expressão de genes individuais ao nível das respectivas proteínas. Para isso se é necessário a disponibilidade de soros/anticorpos específicos para parálogos selecionados. Este trabalho teve então como foco importante a produção destes soros/anticorpos e sua análise quanto a capacidade de reconhecer parálogos específicos recombinantes expressos em *Escherichia coli* ou nativos em extratos totais de *L. braziliensis*.

7.2 Expressão de proteínas recombinantes para a produção de anticorpos

Após a análise filogenética, utilizamos genes de clados independentes para guiar a escolha de proteínas para a produção de soros anti-GP63 específicos para diferentes parálogos de GP63, entre aqueles que sofreram expansão no cromossomo 10. Inicialmente optou-se pela utilização de proteínas recombinantes para a produção destes soros. Para isso, partiu-se para a expressão de proteínas recombinantes correspondentes a dois genes de GP63, LbrM.10.0520 e LbrM.10.0540. Os respectivos fragmentos gênicos, previamente clonados, foram subclonados no vetor de expressão pET21a(+) (Novagen). Na Figura 11 podemos observar a verificação da subclonagem no vetor pET21a(+) (Novagen), dos genes LbrM.10.0520 e LbrM.10.0540, após a digestão com as enzimas de restrição *Bam*HI e *Xho*I. Apenas as amostras que liberaram plasmídeos e insertos no tamanho predito (5441 pb e 1650 pb, respectivamente) foram utilizadas para a produção de proteínas recombinantes fusionadas a histidina por expressão heteróloga em *E. coli*.

Figura 11 - Verificação de subclonagens dos genes de GP63

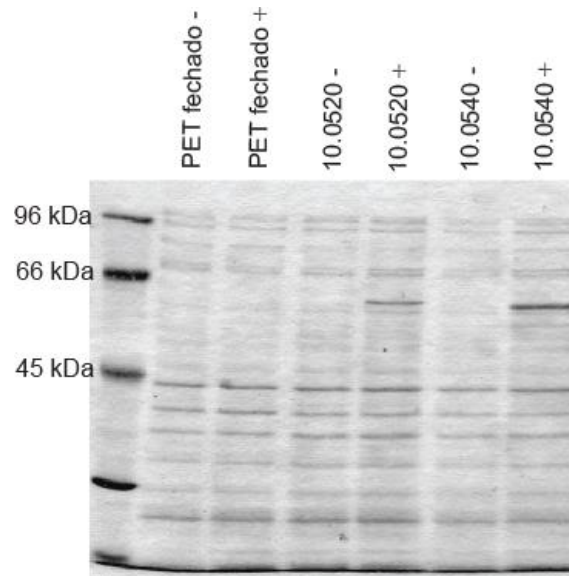


Fonte: A autora

Legenda: Figura de verificação de subclonagem após digestão da construção 10.0520-pET21a(+) e 10.0540-pET21a(+). As setas em vermelho indicam as digestões que liberaram um inserto no tamanho predito (aproximadamente 1650 pb para o inserto e aproximadamente 5000 pb para o plasmídeo). Marcador (M): 1 Kb plus (Invitrogen).

A expressão das proteínas recombinantes correspondentes aos genes selecionados pode ser observada através da Figura 12, demonstrando bandas nos tamanhos preditos de 60.41 kDa, para LbrM.10.0520, e 61.59 kDa para LbrM.10.0540. Foi feita a tentativa de expressão em grande escala dessas proteínas, porém somente a LbrM.10.0540, foi expressa (Figura 13 (A)). A confirmação da expressão da 10.0540 foi detectada através de *Western blot* frente a anticorpos específicos anti-histidina (Figura 13 (B)). No ensaio, foi observada uma banda no tamanho correspondente a proteína de interesse (61.59 kDa para a LbrM.10.0540).

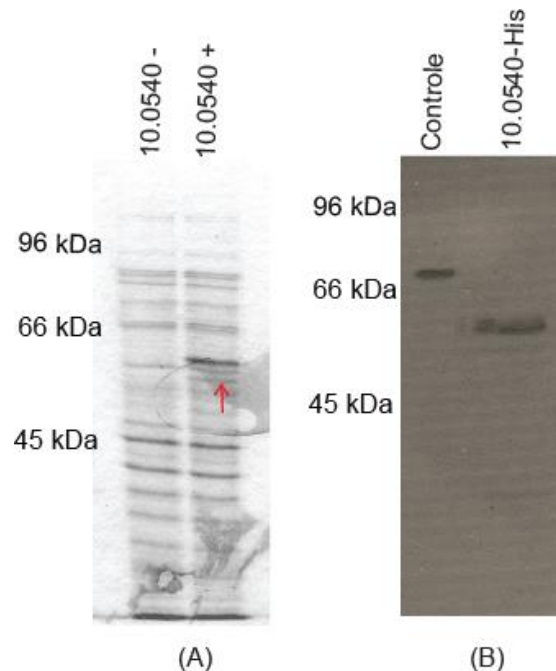
Figura 12 - Expressão em *Escherichia coli* das proteínas GP63 recombinantes fusionadas a sequência de poli-histidinas.



Fonte: A autora

Legenda: Gel de *SDS-PAGE* 15% corado com Comassie Blue. Os tamanhos dos marcadores de peso molecular estão expressos em kilodaltons (kDa) a esquerda. (-) = pré-IPTG; (+) = pós IPTG.

Figura 13 - Análise da expressão da proteína GP63 recombinante LbrM.10.0540 fusionada a uma sequência de poli-histidinas



Fonte: A autora.

Legenda: (A) Gel de *SDS-PAGE* 20% corado com Comassie Blue. A proteína recombinante 10.0540 obtida pode ser observada na seta. (-) = pré-IPTG; (+) = pós IPTG. (B) Análise por *WB* com o anticorpo monoclonal Anti-His (diluição 1:10000). Como controle positivo foi utilizada a proteína recombinante Lci13, também fusionada a sequência de poli-histidinas.

Para a produção de anticorpos anti-LbrM.10.0540, foi feita a tentativa de purificação da respectiva proteína, porém, visto que proteínas transmembrana são de difícil purificação, apenas aproximadamente 0,12 µg de proteína recombinante foi obtido no ensaio. Essa quantidade não é o suficiente para a imunização em coelhos (150 µg por pulso). Mesmo assim, seguimos com a tentativa de produzir soros utilizando a banda de GP63 do extrato de *E. coli* sonicado. Os soros produzidos não reconheceram nenhuma proteína recombinante (resultados não mostrados).

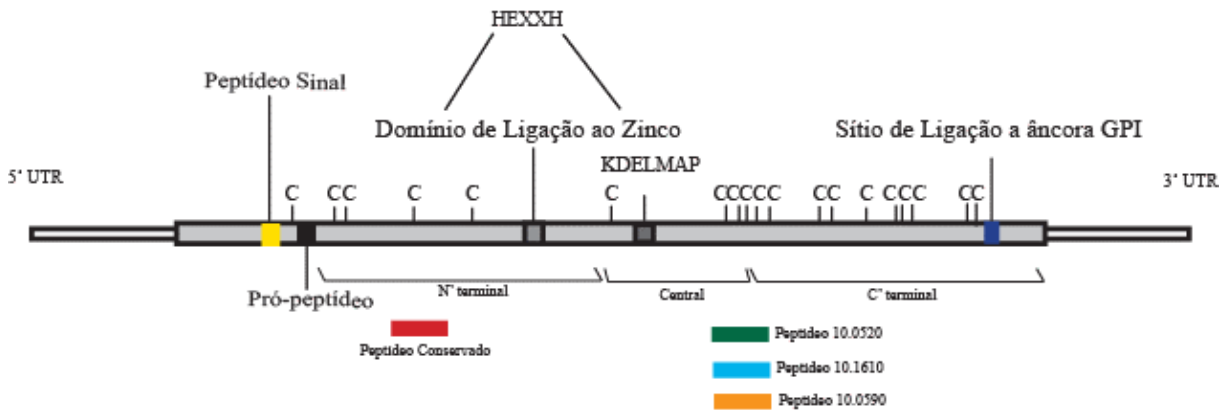
7.3 Estrutura das proteínas GP63 e escolha de peptídeos sintéticos para a produção de anticorpos

Visto que a tentativa da produção de anticorpos utilizando a metodologia de proteínas recombinantes não foi bem-sucedida, partimos então para avaliar a produção de soros policlonais utilizando peptídeos sintéticos. Inicialmente investigamos a estrutura das proteínas GP63, para, com isto, escolher regiões vantajosas para o desenho destes peptídeos e específicas para parálogos selecionados. Em alinhamento das sequências proteicas utilizadas nesse estudo (ANEXO A), e com base em dados descritos na literatura, foi possível observar os três domínios descritos das proteínas GP63 (N-terminal, central e C-terminal), sítios de remoção do peptídeo sinal e pró-peptídeo, regiões de ligação a âncora GPI, região ligante de zinco com o motivo de sequência HEXXHXXGXXH e os resíduos conservados de cisteína. No alinhamento também foi observado que as sequências 10.0610, 10.0520, 10.1630, 10.1610 de *L. braziliensis*, 10.0470 de *L. major* e 10.0510 de *L. infantum* não possuem o peptídeo sinal para adição da âncora GPI (presente na posição 577 em *L. major* e 563 em *L. braziliensis*). Ainda no alinhamento, foi visto que a sequência proteica 10.0540 perde uma parte do peptídeo sinal, presente na região N-terminal da proteína.

No esquema da Figura 14 estão representadas, de maneira ilustrativa, as diferentes regiões comuns às proteínas GP63, bem como a posição de peptídeos selecionados específicos dos parálogos LbrM.10.0520, LbrM.10.0590 e LbrM.10.1610. Estes peptídeos foram selecionados por apresentarem pouca similaridade entre si (genes apontados na árvore filogenética da Figura 10). Também foi identificado um peptídeo correspondente a uma sequência conservada na maioria

dos genes de GP63, que foi utilizada buscando-se um reconhecimento geral dessas proteínas (Quadro 2 e Figura 15).

Figura 14 - Características proteicas da proteína GP63



Fonte: A autora

Legenda: O quadrado em amarelo na sequência indica o sítio de remoção do peptídeo sinal; O quadrado preto na sequência indica o sítio de remoção do pró-peptídeo. Estão delimitados os três domínios da proteína (N-terminal, central, C-terminal). O quadrado em cinza claro na sequência indica o motivo de sequência motivo de sequência HEXXH. O quadrado em cinza escuro indica a sequência conservada KDEL MAP. O quadrado em azul-escuro indica o sítio de ligação a âncora GPI. Os retângulos coloridos abaixo da sequência demonstram a região em que os peptídeos sintéticos estão na sequência.

Quadro 2 - Sequências de aminoácidos escolhidas para obtenção de peptídeos sintéticos

Genes codificantes	Região no alinhamento	Sequência de aminoácidos
10.0590	~413	IVLTHTSVPQYFQYFTDPT
10.1610	~430	KFTQRKRPLPTYFQYFTDSFI
10.0520	~415	IIVTHDDPLPTYFQYFTEPS
	Região no alinhamento	Sequência de aminoácidos
Sequência de aminoácidos conservada	~180	VRQVQGSWKVTGMTG

Fonte: A autora

Figura 15 - Localização dos peptídeos no alinhamento de seqüências

LbrM.10.0590	146	YLIPQALQLH ^{ERL} LRVQVQGS ^W KVVTGM ^S SDVCG ^S FKV ^{FA} HV ^{AG} GN ^{SD} DFVLYVASVPS ^{ES} GVLA ^W AITCO ^Q FS ^D GH ^F	225
LbrM.10.0610	156	YLIPQALQLH ^{ERL} LRVQVQGS ^W KVVTGM ^S SDVCG ^S FKV ^{FA} HV ^{AG} GN ^{SD} DFVLYVASVPS ^{ES} GVLA ^W AITCO ^Q FS ^D GH ^F	235
100520PCR	148	SLIPQALQLH ^{ERL} LRVQVQGS ^W KVVTGM ^S SDVCG ^S FKV ^{FA} HV ^{AG} GN ^{SD} DFVLYVASVPS ^{ES} GVLA ^W AITCO ^Q FS ^D GH ^F	227
LbrM.10.0540	130	SLIPQALQLH ^{ERL} LRVQVQGS ^W KVVTGM ^S SDVCG ^S FKV ^{FA} HV ^{AG} GN ^{SD} DFVLYVASVPS ^{ES} GVLA ^W AITCO ^Q FS ^D GH ^F	209
LbrM.10.1630	157	YLIPQALQLH ^{ERL} LRVQVQGS ^W KVVTGM ^S SDVCG ^S FKV ^{FA} HV ^{AG} GN ^{SD} DFVLYVASVPS ^{ES} GVLA ^W AITCO ^Q FS ^D GH ^F	236
LbrM.10.1610	157	YLIPQALQLH ^{ERL} LRVQVQGS ^W KVVTGM ^S SDVCG ^S FKV ^{FA} HV ^{AG} GN ^{SD} DFVLYVASVPS ^{ES} GVLA ^W AITCO ^Q FS ^D GH ^F	236
LmjF.10.0470	159	YLIPQALQLH ^{ERL} LRVQVQGS ^W KVVTGM ^S SDVCG ^S FKV ^{FA} HV ^{AG} GN ^{SD} DFVLYVASVPS ^{ES} GVLA ^W AITCO ^Q FS ^D GH ^F	238
LmjF.10.0460	159	YLIPQALQLH ^{ERL} LRVQVQGS ^W KVVTGM ^S SDVCG ^S FKV ^{FA} HV ^{AG} GN ^{SD} DFVLYVASVPS ^{ES} GVLA ^W AITCO ^Q FS ^D GH ^F	238
LinJ.10.0500	156	YLIPQALQLH ^{ERL} LRVQVQGS ^W KVVTGM ^S SDVCG ^S FKV ^{FA} HV ^{AG} GN ^{SD} DFVLYVASVPS ^{ES} GVLA ^W AITCO ^Q FS ^D GH ^F	235
LinJ.10.0510	156	YLIPQALQLH ^{ERL} LRVQVQGS ^W KVVTGM ^S SDVCG ^S FKV ^{FA} HV ^{AG} GN ^{SD} DFVLYVASVPS ^{ES} GVLA ^W AITCO ^Q FS ^D GH ^F	235
LbrM.10.0590	386	TQW ^E MP ^{EN} TD ^{EN} AI ^{RC} PT ^{ER} LR ^{LG} CA ^{IV} W ^{HT} -S ^{WE} YF ^Q Y ^{FT} D ^{PI} L ^{IG} S ^D F ^{MD} Y ^{CP} W ^{VE} Y ^D G ^{SC} H ^{QR} S ^{ET} SS ^{IM}	464
LbrM.10.0610	396	TQW ^E MP ^{EN} TD ^{EN} RR ^{RC} PT ^{ER} LR ^{LG} CA ^{IV} W ^{HT} -S ^{WE} YF ^Q Y ^{FT} D ^{PI} L ^{IG} S ^D F ^{MD} Y ^{CP} W ^{VE} Y ^D G ^{SC} H ^{QR} S ^{ET} SS ^{IM}	475
100520PCR	388	TQW ^E MP ^{EN} TD ^{EN} RR ^{RC} PT ^{ER} LR ^{LG} CA ^{IV} W ^{HT} -S ^{WE} YF ^Q Y ^{FT} D ^{PI} L ^{IG} S ^D F ^{MD} Y ^{CP} W ^{VE} Y ^D G ^{SC} H ^{QR} S ^{ET} SS ^{IM}	467
LbrM.10.0540	370	TQW ^E MP ^{EN} TD ^{EN} RR ^{RC} PT ^{ER} LR ^{LG} CA ^{IV} W ^{HT} -S ^{WE} YF ^Q Y ^{FT} D ^{PI} L ^{IG} S ^D F ^{MD} Y ^{CP} W ^{VE} Y ^D G ^{SC} H ^{QR} S ^{ET} SS ^{IM}	449
LbrM.10.1630	397	TQW ^E MP ^{EN} TD ^{EN} RR ^{RC} PT ^{ER} LR ^{LG} CA ^{IV} W ^{HT} -S ^{WE} YF ^Q Y ^{FT} D ^{PI} L ^{IG} S ^D F ^{MD} Y ^{CP} W ^{VE} Y ^D G ^{SC} H ^{QR} S ^{ET} SS ^{IM}	476
LbrM.10.1610	397	TQW ^E MP ^{EN} TD ^{EN} RR ^{RC} PT ^{ER} LR ^{LG} CA ^{IV} W ^{HT} -S ^{WE} YF ^Q Y ^{FT} D ^{PI} L ^{IG} S ^D F ^{MD} Y ^{CP} W ^{VE} Y ^D G ^{SC} H ^{QR} S ^{ET} SS ^{IM}	476
LmjF.10.0470	399	TQW ^E MP ^{EN} TD ^{EN} RR ^{RC} PT ^{ER} LR ^{LG} CA ^{IV} W ^{HT} -S ^{WE} YF ^Q Y ^{FT} D ^{PI} L ^{IG} S ^D F ^{MD} Y ^{CP} W ^{VE} Y ^D G ^{SC} H ^{QR} S ^{ET} SS ^{IM}	478
LmjF.10.0460	399	TQW ^E MP ^{EN} TD ^{EN} RR ^{RC} PT ^{ER} LR ^{LG} CA ^{IV} W ^{HT} -S ^{WE} YF ^Q Y ^{FT} D ^{PI} L ^{IG} S ^D F ^{MD} Y ^{CP} W ^{VE} Y ^D G ^{SC} H ^{QR} S ^{ET} SS ^{IM}	477
LinJ.10.0500	396	TQW ^E MP ^{EN} TD ^{EN} RR ^{RC} PT ^{ER} LR ^{LG} CA ^{IV} W ^{HT} -S ^{WE} YF ^Q Y ^{FT} D ^{PI} L ^{IG} S ^D F ^{MD} Y ^{CP} W ^{VE} Y ^D G ^{SC} H ^{QR} S ^{ET} SS ^{IM}	474
LinJ.10.0510	396	TQW ^E MP ^{EN} TD ^{EN} RR ^{RC} PT ^{ER} LR ^{LG} CA ^{IV} W ^{HT} -S ^{WE} YF ^Q Y ^{FT} D ^{PI} L ^{IG} S ^D F ^{MD} Y ^{CP} W ^{VE} Y ^D G ^{SC} H ^{QR} S ^{ET} SS ^{IM}	475

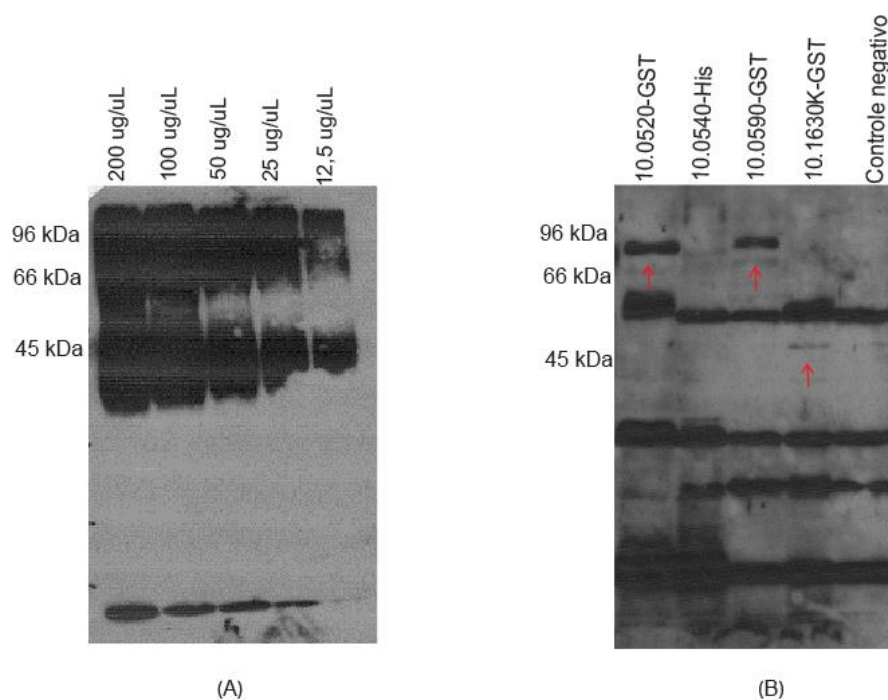
Fonte: A autora.

Legenda: No quadro vermelho podemos observar as seqüências conservadas entre as proteínas de *L. braziliensis*. No quadro verde temos a seqüência peptídica 10.0590, no quadro azul seqüência peptídica 10.0520 e no quadro roxo a seqüência peptídica 10.1610.

7.4 Reconhecimento dos anticorpos obtidos a partir dos peptídeos sintéticos

Quatro soros foram gerados, um para cada peptídeo, a partir de sua imunização conjugada a proteína carreadora em coelhos. Como primeira etapa para se avaliar a eficiência do procedimento, soros selecionados foram previamente testados frente ao próprio peptídeo conjugado ao BSA nas concentrações de 200 µg/µL; 100 µg/µL; 50 µg/µL; 25 µg/µL e 12,5 µg/µL (Figura 16A). Os soros foram então testados frente a proteínas GP63 recombinantes conforme aqui demonstrado para o soro anti-peptídeo conservado (Figura 16B), confirmando então sua capacidade de reconhecer esta proteína. Entretanto pode-se notar um alto reconhecimento inespecífico de proteínas de *Escherichia coli* o que levou a necessidade de se purificar, a partir dos soros policlonais, anticorpos específicos anti-GP63.

Figura 16 - Soro anti-peptídeo conservado frente ao peptídeo conjugado ao BSA e frente a proteínas recombinantes.



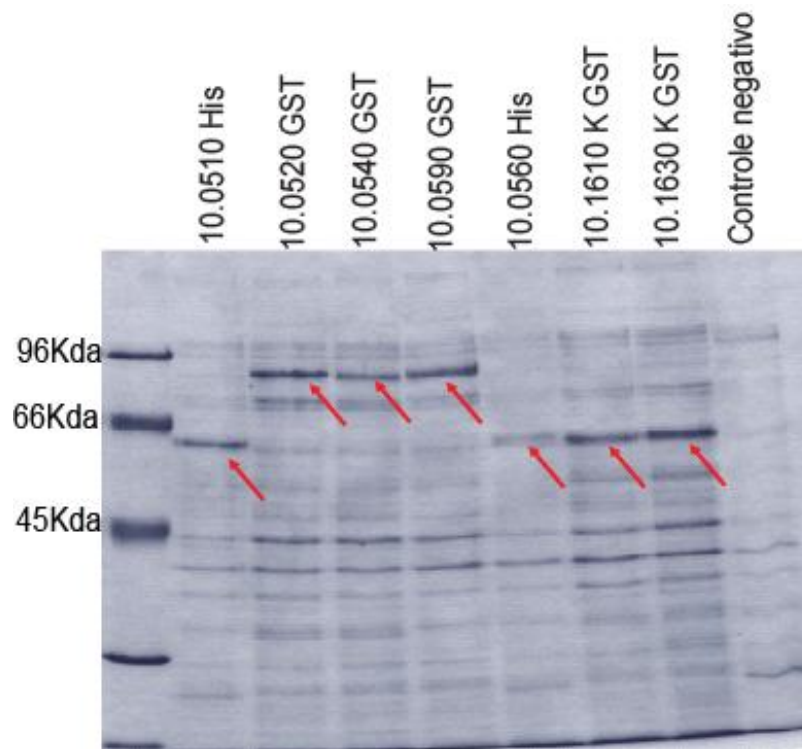
Fonte: A autora

Legenda: (A) Os peptídeos conjugados ao BSA nas concentrações de 200 $\mu\text{g}/\mu\text{L}$; 100 $\mu\text{g}/\mu\text{L}$; 50 $\mu\text{g}/\mu\text{L}$; 25 $\mu\text{g}/\mu\text{L}$ e 12,5 $\mu\text{g}/\mu\text{L}$; foram fracionados em gel *SDS-PAGE* 20% e analisados por *Western blot* com os anticorpos anti-peptídeo sintético 10.0520 (diluição 1:1000). (B) Extratos celulares de *E. coli* contendo as proteínas recombinantes (10.0520-GST; 10.0540 – H; 10.0590 – GST; e 10.1630- GST) foram fracionados em gel *SDS-PAGE* 15% e analisadas por *Western blot* contra o soro pós-imune anti-peptídeo conservado (1:1000). As setas indicam as bandas correspondentes as proteínas GP63.

Após os ensaios de imunoabsorção para cada anticorpo, os quatro conjuntos de anticorpos obtidos foram testados contra extratos de *E. coli* expressando proteínas recombinantes previamente produzidas correspondentes a diferentes parálogos de GP63 de *L. braziliensis*. De acordo com a disponibilidade limitada de proteínas disponíveis, optou-se por um painel mais diverso possível e incluindo proteínas contendo os diferentes peptídeos utilizados na produção de soro, visando confirmar a especificidade dos anticorpos contra parálogos específicos de GP63. Quatro proteínas recombinantes, fusionadas a Glutathione-S transferase (GST) ou a poli-histidinas, foram derivadas de fragmentos correspondentes a maior parte da região codificante dos respectivos genes, terminando na âncora GPI (10.0510-His, 10.0520-GST; 10.0540-GST; 10.0560-His; 10.0590-GST). Outras duas proteínas, fusionadas a GST (10.1610-GST e 10.1630-GST), são derivadas de fragmentos gênicos que terminam

no peptídeo KDELMAP, conservado em todas as sequências de GP63 (ver Figura 13). Para um melhor entendimento das figuras de *Western blot*, observamos na Figura 17 um gel de *SDS-PAGE* 20% corado com *Comassie Blue*, que obtém as mesmas proteínas utilizadas nos ensaios, além de um controle negativo que consiste em um extrato celular bacteriano.

Figura 17 - Extratos de *E. coli* contendo as proteínas recombinantes GP63



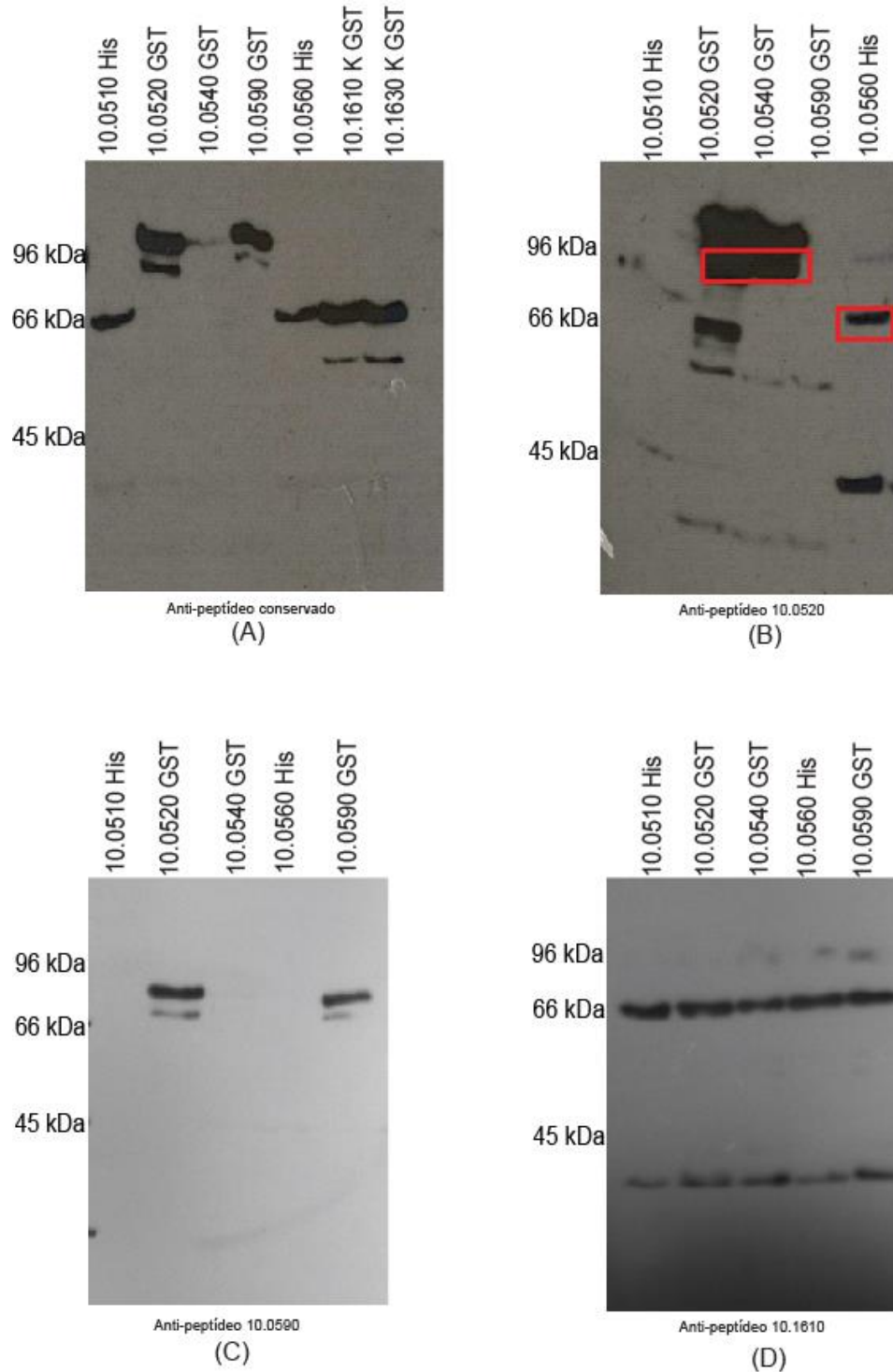
Fonte: A autora.

Legenda: Gel de *SDS-PAGE* 15% corado com *Comassie Blue*. As setas indicam as proteínas expressas. Os tamanhos dos marcadores de peso molecular estão expressos em kilodaltons (kDA) a esquerda.

A Figura 18 resume os resultados alcançados testando os quatro lotes de anticorpos obtidos contra o painel de diferentes proteínas recombinantes. Inicialmente observamos o reconhecimento dos anticorpos anti-peptídeo conservado (Figura 18A). Conforme o esperado esses anticorpos reconheceram todas as proteínas recombinantes, visto que este peptídeo está presente em todas as GP63 na região N-terminal (anterior a região KDELMAP). Entretanto, este anticorpo reconhece fracamente a proteína 10.0540-GST, talvez devido a diferenças conformacionais, onde uma região pode estar mais exposta em uma proteína em relação as outras.

Em seguida foi avaliado o reconhecimento das proteínas recombinantes pelos anticorpos anti-peptídeo 10.0520 (Figura 18B). Esse anticorpo reconheceu as proteínas 10.0520-GST e 10.0540-GST; que são duas proteínas que agrupam bem próximas na árvore filogenética devido a sua similaridade (Figura 10) e também a proteína 10.0560-His. O anticorpo não ficou totalmente específico, reconhecendo algumas bandas do extrato bacteriano, porém em relação ao soro pós-imune, obtivemos um reconhecimento mais limpo (resultados não mostrados). Já na Figura 18C, temos o reconhecimento dos anticorpos anti-peptídeo 10.0590, que reconhece a proteína 10.0520-GST e 10.0590-GST. É possível que os anticorpos anti-peptídeo 10.0590 (sequencia: IVLTHTSVPQYFQYFTDPT) também possam reconhecer a proteína recombinante 10.0520-GST por partilhar alguns aminoácidos nessa sequência específica "IIVTH" e "YFQYFT". No ensaio de *Western blot* para os anticorpos anti-peptídeo 10.1610, não foi possível observar o reconhecimento desses anticorpos, visto que, não temos a proteína LbrM.10.1610 recombinante no ensaio, observando que esses anticorpos não possuem reação cruzada com outras proteínas GP63 (Figura 18D).

Figura 18 - Reconhecimento de proteínas GP63 recombinantes pelos anticorpos anti-peptídeo gerados neste estudo.



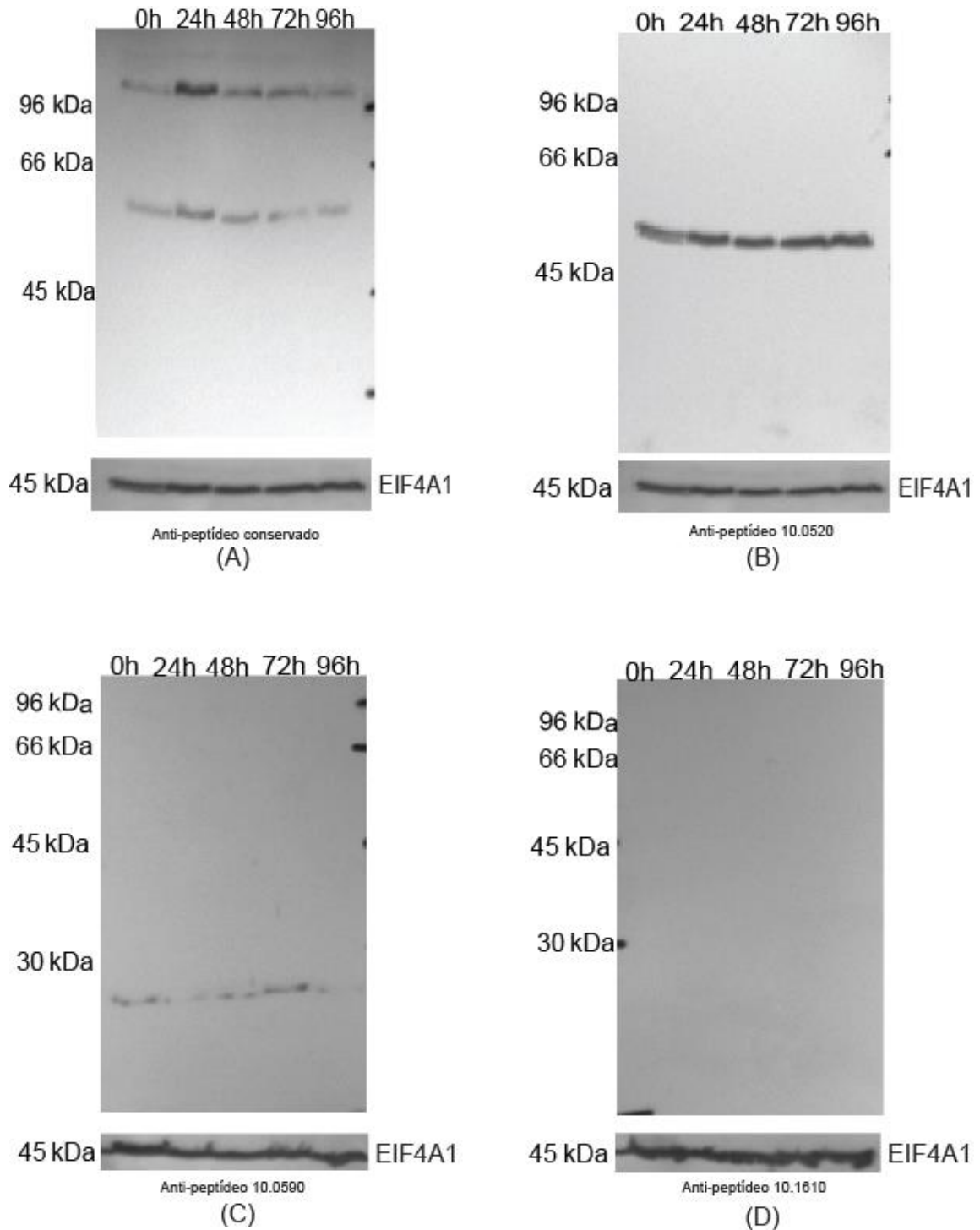
Fonte: A autora

Legenda: Extratos celulares de *E. coli* expressando diferentes proteínas recombinantes fracionadas em gel *SDS-PAGE* 20% e analisadas por *WB* contra os anticorpos anti-peptídeo conservado (A), anti-peptídeo 10.0520 (B), anti-peptídeo 10.0590 (C) e anti-peptídeo 10.1610 (D). Todos os anticorpos estão na concentração de 1:1000.

7.5 Reconhecimento dos anticorpos anti-peptídeos frente a curvas de *L. braziliensis*

Para avaliar o reconhecimento em extratos de *L. braziliensis* dos quatro lotes de anticorpos obtidos neste estudo, foram feitos ensaios de *Western blot* com cada lote de anticorpos frente a curvas de crescimento de *L. braziliensis* contendo células em fase Logarítmica (24h a 48h) e células em fase Estacionária (72h a 96h). Com os anticorpos anti-peptídeo conservado foi observado o reconhecimento de duas bandas, uma de aproximadamente 55 kDa e outra de 96 kDa, havendo uma diminuição da expressão das duas bandas na fase estacionária (Figura 19A). Os anticorpos anti-peptídeo 10.0520 reconheceram apenas uma banda específica, ou duas de tamanhos bem próximos, de aproximadamente 50 kDa (Figura 19B). Nesse caso não foi possível observar alterações significativas no padrão de expressão dessas proteínas ao longo da curva de crescimento. Com os anticorpos anti-peptídeo 10.0590 pode-se observar uma banda de aproximadamente 30 kDa, o que pode ser um reconhecimento inespecífico uma vez que este tamanho seria incompatível com a GP63 (Figura 19C). Já os anticorpos anti-peptídeo 10.1610 não reconheceram nenhuma banda no extrato de *L. braziliensis* (Figura 19D).

Figura 19 - Anticorpos anti-peptídeos frente a curvas de *L. braziliensis*



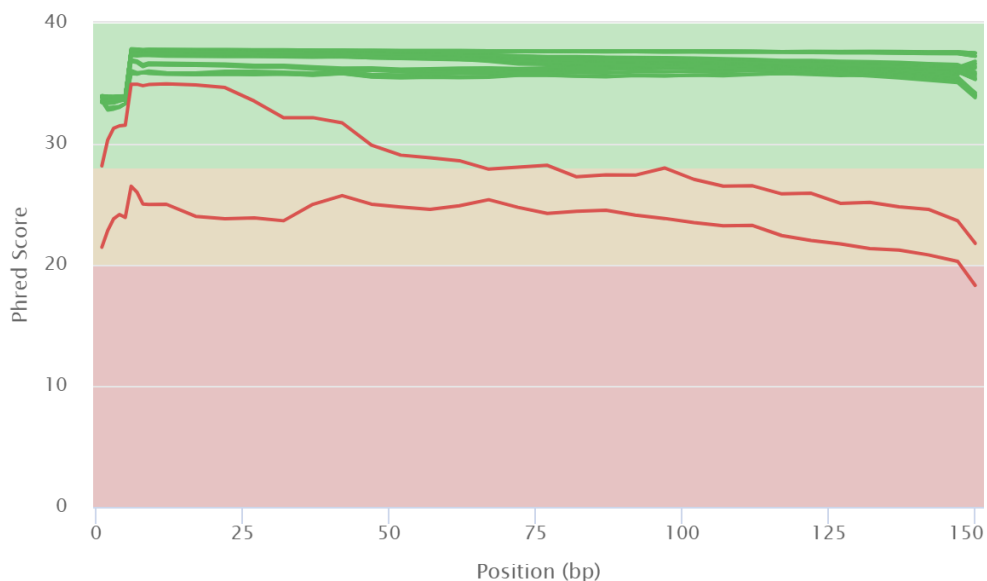
Fonte: A autora

Legenda: Extratos celulares de curvas de crescimento de *L. braziliensis* (0h, 24h, 48h, 72h e 96h) fracionados em gel *SDS-PAGE* 20% e analisados por *Western blot* contra os anticorpos purificados anti-peptídeo conservado (A), anti-peptídeo 10.0520 (B), anti-peptídeo 10.0590 (C) e anti-peptídeo 10.1610 (D). Todos os anticorpos estão na concentração de 1:500. Abaixo de cada figura podemos observar a expressão constitutiva da proteína EIF4A1.

7.6 Quantidade de mRNA de diferentes genes de GP63 em diferentes fases de crescimento de *L. braziliensis*

Buscando uma análise independente da quantidade de mRNA de cada GP63, foram feitas análises de RNASeq de *L. braziliensis* realizado em trabalho anterior, focando em uma maior avaliação da presença dos mRNAs de GP63. Antes de realizá-las, contudo, foi avaliada a qualidade das reads obtidas. O FastQC fornece um conjunto modular de análises para termos uma noção sobre a qualidade e sobre possíveis problemas que possam estar ocorrendo com dados brutos de sequenciamento. Analisar os dados no FastQC é uma etapa fundamental para evitar possíveis problemas nas análises futuras. A avaliação da qualidade do sequenciamento é baseada em scores. Valores de qualidade 30 para uma base significam, por exemplo, que a chance dessa base estar incorreta é de 1 em 1000, possuindo acurácia igual a 99,9%. Na análise de RNASeq, só foram utilizadas as bibliotecas com sequências que continham um score acima de 30. As amostras (linhas em vermelho) com um valor abaixo de 30 não foram utilizadas nas análises (Figura 20).

Figura 20 - Análise das bibliotecas de cDNA sequenciadas.

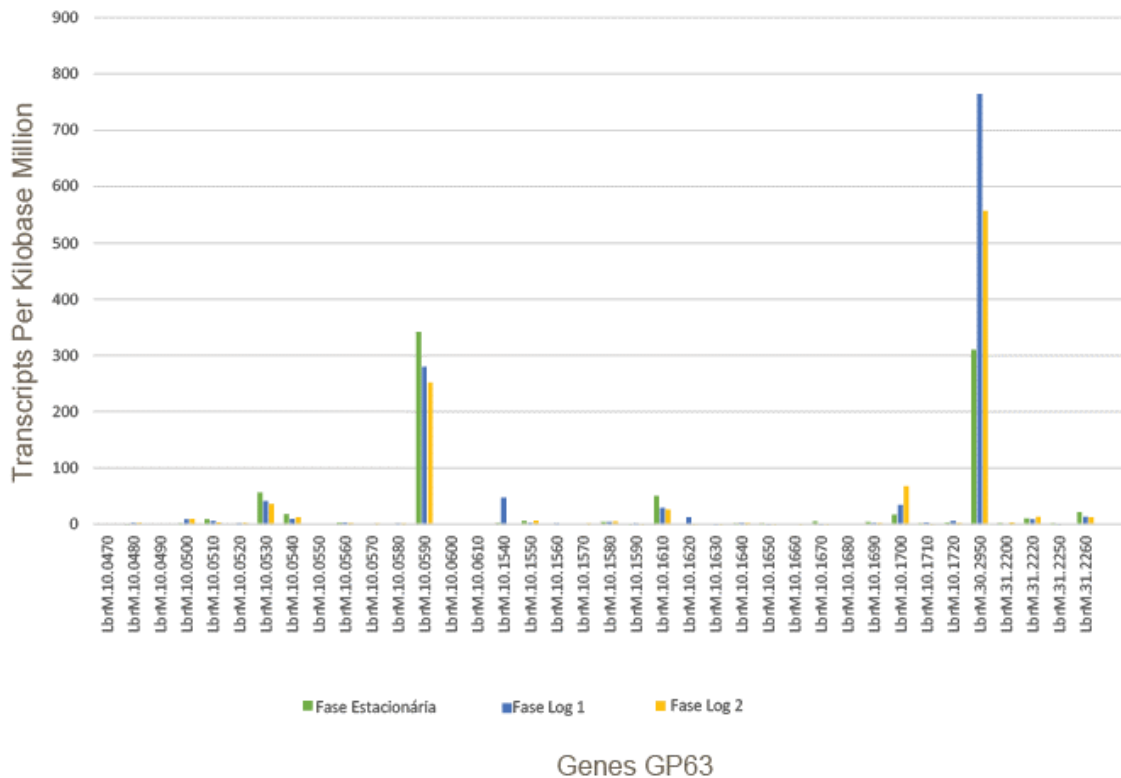


Fonte: A autora

Legenda: Phred Score: indica a probabilidade de uma base nucleotídica indicada pelo sequenciador estar errada. Position (bp): Posição de cada nucleotídeo nas sequências das bibliotecas avaliadas. Cada linha verde representa uma amostra analisada. As cores verde, bege e vermelho no gráfico, representam scores de amostras de alta, média e baixa qualidade.

A partir dos dados resultantes, foram então feitos gráficos utilizando a média dos valores de TPM (*Transcripts Per Kilobase Million*), onde normalizamos as leituras por comprimento do gene e por profundidade de sequenciamento, entre as triplicadas para cada situação: fase estacionária; fase logarítmica de uma célula recém descongelada; fase logarítmica de uma célula após repiques sucessivos (Figura 21). Foi possível observar que o mRNA de 10.0590 é o mais abundante, nas duas fases de crescimento, seguido dos genes 10.0530, 10.1610 (porém em menor quantidade). Desses três genes, foi observado uma maior quantidade de transcritos na fase estacionária do parasita. Nota-se a reduzida ou quase nula expressão dos genes de GP63 codificados no cromossomo 31. Nas análises foi adicionado como controle o gene LbrM 30.2950, que codifica a enzima glicerol 3-fosfato, muito expressa na fase promastigota dos parasitas. Chama a atenção a redução brusca nos níveis desse transcrito controle na fase estacionária do parasita, ao contrário do que se observa nos transcritos mencionados de GP63 do cromossomo 10, onde na verdade os maiores níveis são observados justamente na fase estacionária de crescimento.

Figura 21 - TPM – Fases Estacionária e Logarítmicas de *L. braziliensis*

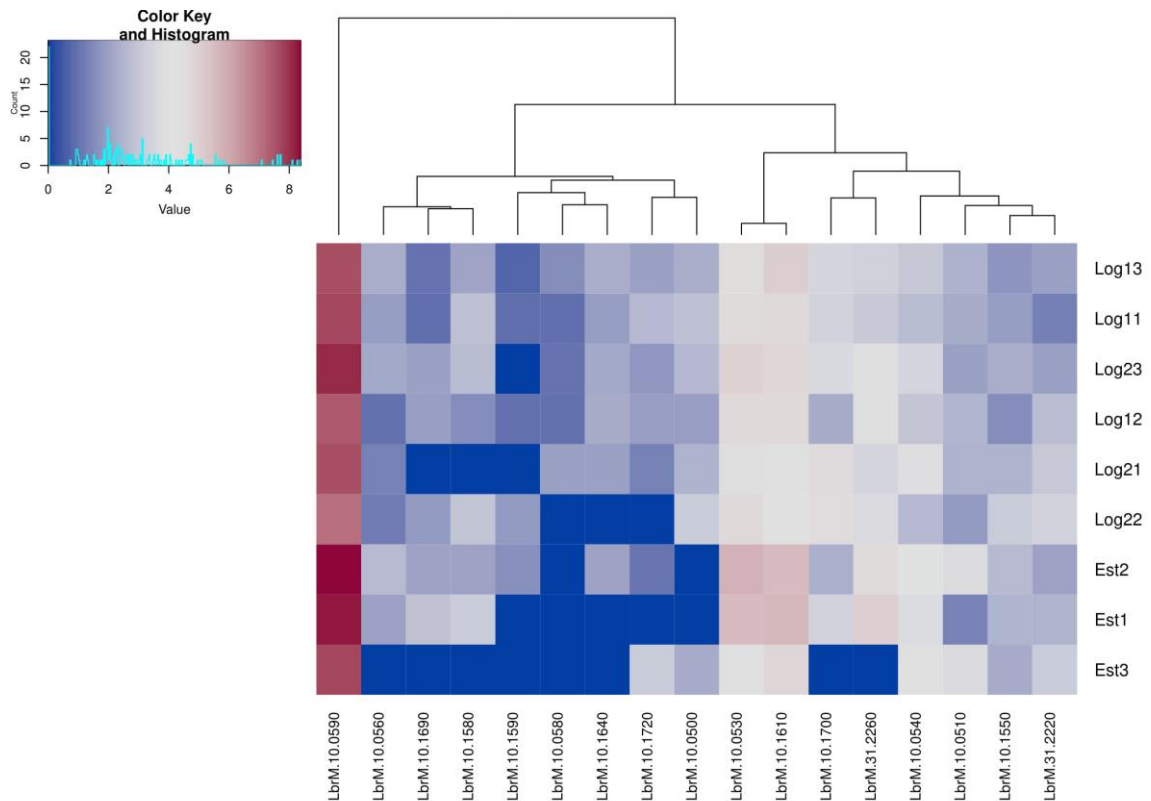


Fonte: A autora

Legenda: No eixo vertical podemos observar o número de *Transcripts Per Kilobase Million* – leituras normalizadas por comprimento do gene e por profundidade de sequenciamento de cada gene de GP63 analisado (eixo horizontal). As barras em verde representam os genes na fase Estacionária; As barras em azul representam os genes na Fase Logarítmica de uma célula recém descongelada (Fase Log 1); As barras em amarelo representam os genes na Fase Logarítmica de uma célula após repiques sucessivos (Fase Log 2).

Para avaliar graficamente a expressão de transcritos por outra forma, foram construídos gráficos *heatmaps*. Conforme observado no gráfico de barras, após a normalização dos dados, foi observado que o gene 10.0590 é mais expresso em relação aos outros genes de GP63, seguido dos genes 10.1610, 10.0530 (observados em tons de vermelho) (Figura 22).

Figura 22 - *Heatmap* da expressão dos genes de GP63



Fonte: A autora

Legenda: O *heatmap* representa os valores normalizados dos transcritos dos genes de GP63. Eixo vertical: Log 11, Log 12, Log 13: Replicatas das análises com as células da Fase Logarítmica (célula recém descongelada). Log 21, Log 22, Log 23: Replicatas das análises com as células da Fase Logarítmica (célula após repiques sucessivos). Est1, Est2, Est3: Replicatas das análises com as células da Fase Estacionária. No quadro acima da imagem, observamos os valores dos padrões de cores observados na análise.

Para observar a localização das reads no cromossomo 10 de *L. braziliensis*, foi feito o mapeamento das reads obtidas na análise de RNASeq, no genoma de *L. braziliensis*. Foi observado existe uma grande quantidade de reads alinhadas na região C-terminal do gene LbrM.10.0590 (LbrM.10:220,702-222,871), conforme pode ser observado na Figura 23. A região C-terminal é muito variável e específica em relação a grupos de genes de GP63. O gene 10.0590 é o representante do grupo 3 de *L. braziliensis*, contendo apenas 2 parálogos aos quais sugere-se serem genes expressos na fase promastigota do parasita, não tendo relação ou funções relevantes no hospedeiro mamífero.

Figura 23 - Reads localizadas na C-terminal do gene 10.0590



Fonte: A autora

Legenda: Os números acima representam em pares de base, a localização cromossômica das reads em todo o genoma de *L. braziliensis*. Podemos observar uma maior cobertura de reads (círculo em vermelho) na parte C-terminal do gene 10.0590 (LbrM.10:220,702-222,871).

8 DISCUSSÃO

Estudos de caracterização demonstraram sítios e motivos de sequências característicos de GP63 de *Leishmania sp.*, como o sítio de remoção do pró-peptídeo, resíduos conservados de arginina e histidina, sítio de ligação a âncora GPI, o domínio HEXXHXXGXXH, e sequências variáveis nas regiões C-terminal e N-terminal (YAO; DONELSON; WILSON, 2003). Essas características se mostraram conservadas e também foram identificadas nas sequências de GP63 de *L. braziliensis* presentes neste estudo. Além disso, essas sequências apresentaram variações significativas em suas regiões C-terminais, com algumas delas contendo o sítio de ligação a âncora GPI e outras podendo apresentar um domínio transmembrana, já identificado em outras espécies de *Leishmania* (RAMAMOORTHY et al., 1992; VOTH et al., 1998).

A variabilidade gênica existente entre os genes de GP63 foi associada a eventos de mosaicismos (MAURICIO et al., 2006) e recombinação intragênica (VICTOIR; DUJARDIN, 2002) que resultou na grande diversidade de genes de GP63. Expansões gênicas no cromossomo onde as sequências codificantes dessas proteínas estão localizadas foram relatadas no nosso estudo e podem estar relacionadas a pressão seletiva causada pela resposta imune dos seus hospedeiros durante a infecção por *Leishmania sp* (MEDINA et al., 2016). Além disso, a presença desses arranjos de múltiplas cópias pode levar à especiação (LYNCH; CONERY, 2003) ou indicar a possível necessidade de genes específicos de cada estágio do ciclo de vida do parasita (ROGERS et al., 2011). Até agora, é provável que essa expansão possa estar relacionada a novos aspectos da patogênese da *Leishmania* aos hospedeiros vertebrados, mas isso ainda precisa ser melhor definido.

Neste estudo, os resultados variáveis (reconhecimentos específicos e inespecíficos) das proteínas do cromossomo 10 de *L. braziliensis* frente a anticorpos específicos podem estar relacionados a variabilidade gênica já mencionada. Isto também pode estar relacionado a presença dessa variação de sequência localizada principalmente na parte externa da proteína, que interage com o sistema imune hospedeiro (CASTRO NETO et al., 2019).

Em estudo realizado com as principais proteínas de virulência de *T. brucei*, as VSG's, foi observado que a antigenicidade das mesmas está associada com a região C-terminal, principalmente com o domínio de ligação da âncora GPI. Quando retirado este domínio, várias alterações conformacionais na proteína foram observadas

(BUTIKOFER et al., 2001). As proteínas recombinantes utilizadas neste ensaio de reconhecimento pelos anticorpos não possuem a região de ligação a âncora GPI. Isso poderia ter gerado alguma conformação diferencial que afetaria o reconhecimento das mesmas, mas, apesar disso, os anticorpos anti-peptídeos deste estudo foram capazes de reconhecer as proteínas recombinantes de interesse. O fato do anti-10.0540, feito com a proteína recombinante não ter demonstrado eficiência no reconhecimento das GP63 pode estar associado a uma baixa antigenicidade desta proteína e conseqüentemente a ausência de uma reação imunológica satisfatória durante a produção do anticorpo.

Apesar do esperado reconhecimento gene-específico pelos anticorpos anti-peptídeo, alguns apresentaram reação com outras proteínas. Esse reconhecimento pode estar associado a similaridade de sequência e relação evolutiva entre essas proteínas. O anticorpo anti-peptídeo 10.0520, reconheceu as proteínas (10.0520, 10.0540 e 10.0560) que estavam localizadas em clados próximos, visualizadas na árvore filogenética demonstrada neste estudo, que é corroborada por árvores de outros estudos filogenéticos da expansão dos genes de GP63 no cromossomo 10 das espécies do subgênero *Viannia* (CASTRO NETO et al., 2019; VALDIVIA et al., 2015). O mesmo não pode ser aplicado para a reação cruzada do anticorpo anti-peptídeo 10.0590 e a proteína 10.0520, que são divergentes em estudos filogenéticos (CASTRO NETO et al., 2019; VALDIVIA et al., 2015). Contudo essas proteínas compartilham 11 aminoácidos na região do peptídeo, que podem estar envolvidos na reação cruzada dos anticorpos.

Esses diferentes anticorpos também foram utilizados em ensaios de WB com extrato de *L. braziliensis* e demonstraram diferentes padrões de bandas entre si. Isso indica a variabilidade de isoformas dessas proteínas, que é corroborado por um estudo realizados em *Trypanosoma cruzi*. No referido trabalho, foram utilizados soros pós-ímmunes para estudar a expressão e localização dessas proteínas em diferentes fases do parasita (epimastigota, amastigota e tripomastigotas), que também apresentaram um padrão de reconhecimento diferenciado (KULKARNI et al., 2009).

Foi observado em um estudo que proteínas GP63 são mais expressas em fase estacionária (RAMAMOORTHY et al., 1992). Esse padrão de expressão, de maneira não foi observado com nossos anticorpos o que sugere que podem existir genes que são mais expressos em diferentes fases do ciclo de vida do parasita, dependendo da sua necessidade e função. Esse dado está de acordo com o RNASeq realizado neste

estudo, que também encontrou uma maior quantidade de mRNA de GP63 em parasitas na fase estacionária, porém sem uma diferença significativa em relação aos transcritos de fase Log. Vale ressaltar que a *L. braziliensis* cepa MHOM/BR/75/M2904 utilizada em nossas análises, já passou por diversos repiques, fato que pode alterar a infectividade do parasita e seu padrão de expressão gênica.

No estudo de Marshall et al., 2018, utilizando-se da hipótese de que existem grupos de genes de GP63 sendo expressos numa fase específica do ciclo de vida do parasita (RAMAMOORTHY et al., 1992; WILSON et al., 1993), as GP63 de *L. infantum* foram separadas em classes, baseados em suas características proteicas, e foram analisados dados de espectrometria de massas dos exossomos desses parasitas. Eles relataram que nenhum grupo de genes era expresso especificamente somente em uma fase de crescimento de promastigotas, apenas alguns genes pontuais se destacavam, com expressão em uma única fase. Estes dados corroboram com as nossas análises de RNASeq, onde também não foi observado um grupo específico de genes sendo expresso em uma fase determinada de crescimento de promastigotas. Além disso, Marshall et al., 2018, também relatou uma maior quantidade de genes de GP63 sendo expressos em exossomos de *L. infantum* em fase estacionária, o que também foi observado nas análises de RNASeq.

O mesmo não acontece em *Trypanosoma cruzi*, onde os genes de GP63 estão divididos em grupos. O grupo I e II estão presentes com um alto número de cópias. Dados de RNASeq e *Northern blot* demonstraram que o grupo I é mais expresso que o grupo II, que quase não apareceu nas análises de *Northern blot* (CUEVAS; CAZZULO; SÁNCHEZ, 2003). O grupo II possui alto grau de homologia com *Leishmania guyanensis* (espécie do subgênero *Viannia*), enquanto o grupo I (mais expresso), foi confirmado após digestão por fosfolipase C específica de fosfatidilinositol, que são proteínas de superfície ligadas à membrana por uma âncora GPI, e, é sugerido a esse grupo a participação na infecção (BERNÁ et al., 2017; CUEVAS; CAZZULO; SÁNCHEZ, 2003).

Baseando-se nas diferenças das regiões C-terminais de GP63, sugere-se que o grupo 3 de *L. braziliensis* possui sua expressão estágio-específica, estando presente na forma promastigota do parasita e não possui papéis relevantes no hospedeiro mamífero (CASTRO NETO et al, 2019). No grupo 3, que tem apenas dois parálogos, está a proteína 10.0590 que foi a proteína com a maior quantidade de mRNA observada nesse estudo. Porém, os genes do grupo 1 e 2 de *L. braziliensis* foram os

genes que sofreram uma grande expansão no cromossomo 10 dessa espécie, então, não observar um grande número de transcritos nesses genes pode significar que estes genes estão relacionados com a infecção e podem ser encontrados sendo expressos na fase amastigota e não na fase promastigota do parasita.

Análises de RNASeq de diferentes fases do ciclo de vida de *Leishmania major*, demonstraram que genes de GP63 no cromossomo 10 aumentam da fase amastigota para fase promastigota (procíclicos e metacíclicos). Porém um gene do cromossomo 28 tem apenas um pico na fase de promastigotas metacíclicos (INDAR et al., 2017). Mais estudos utilizando as diferentes fases do ciclo de vida de *L. braziliensis* precisam ser realizados para esclarecer o mecanismo de expressão dos genes de GP63, considerando que nossa análise somente na fase promastigota não foi suficiente para identificar possíveis genes estágio específicos. Outros estudos também são necessários na fase amastigota ou parasitas isolados diretamente de infecções, visto que foi relatada uma maior expressão de genes de GP63 em infecções em andamento (SANYAL; GHOSH; SARKAR, 1994).

9 CONCLUSÕES

- a) A falta de reconhecimento uniforme das proteínas recombinantes por um único anticorpo, pode indicar que apenas um anticorpo não é suficiente para reconhecer todas as proteínas de GP63, durante o processo de infecção;
- b) Existem variações no padrão de expressão dos genes de GP63 de *L. braziliensis*, porém mais estudos serão necessários para detalhar quais são esses genes, bem como a sua localização subcelular e suas possíveis implicações na interação com o sistema imune;
- c) Genes de GP63 dos grupos 1 e 2 foram pouco expressos na análise de RNASeq em relação ao gene do grupo 3, sugerindo que esses genes podem ter funções no hospedeiro mamífero. Sendo assim, é preciso a realização de análises utilizando-se das outras formas do ciclo de vida do parasita, tais como a forma amastigota.

REFERÊNCIAS

- ABBAS, A. K.; LICHTMAN, A. H. H.; PILLAI, S. **Cellular and molecular immunology E-book**. [S. l.]: Elsevier Health Sciences, 2014.
- ABASCAL, F.; ZARDOYA, R.; POSADA, D. ProtTest: selection of best-fit models of protein evolution. **Bioinformatics**, Oxford, v. 21, n. 9, p. 2104-2105, 2005.
- ABDULLAH, A. Y. et al. Environmental factors associated with the distribution of visceral leishmaniasis in endemic areas of Bangladesh: modeling the ecological niche. **Tropical Medicine and Health**, Oxford, v. 45, n. 13, p. 1–15, 2017.
- AGUIAR, P. F.; RODRIGUES, R. K. Leishmaniose Visceral no Brasil: Artigo de Revisão. **Unimontes Científica**, Montes Claros, v. 19, n. 1, p. 191–203, 2017.
- ALEMAYEHU, B.; ALEMAYEHU, M. Leishmaniasis: A Review on Parasite, Vector and Reservoir Host. **Health Science Journal**, Londres, v. 11, p. 1–6, 2017.
- AL-KAMEL, M. A. N. Leishmaniasis and Malignancy: A Review and Perspective. **Clinical Skin Cancer**, Sana'a, v. 2, n. 1/2, p. 54–58, 2017.
- AVILA-LEVY, C. M. et al. GP63 Function in the Interaction of Trypanosomatids with the Invertebrate Host : Facts and Prospects. In: SANTOS, A. L. S. et al. (Ed.). **Proteins and Proteomics of Leishmania and Trypanosoma, Subcellular Biochemistry**. Rio de Janeiro: Springer Science, 2014. p. 253–270.
- BASANO, S. de A.; CAMARGO, L. M. A. Leishmaniose tegumentar americana: histórico, epidemiologia e perspectivas de controle. **Revista Brasileira de Epidemiologia**, São Paulo, v. 7, p. 328–337, 2004.
- BATES, P. A. Transmission of *Leishmania* metacyclic promastigotes by phlebotomine sand flies. **International Journal for Parasitology**, Oxford, v. 37, n. 10, p. 1097–1106, 2007.
- BERNÁ, L. et al. Transcriptomic analysis reveals metabolic switches and surface remodeling as key processes for stage transition in *Trypanosoma cruzi*. **PeerJ: Life and Environmental**, [S. l.], v. 5, p. e3017, 2017.
- BESSAT, M.; SHANAT, S. EL. Leishmaniasis : Epidemiology , Control and Future Perspectives with Special Emphasis on Egypt. **Journal of Tropical Diseases**, [S. l.] v. 3, n. 1, p. 1–10, 2015.
- BURZA, S.; CROFT, S. L.; BOELAERT, M. Leishmaniasis. **The Lancet**, London, v. 392, n. 10151, p. 951–970, 2018.
- BUTIKOFER, P. et al. GPI-anchored proteins: now you see “em, now you don”t. **The FASEB Journal**, Bethesda, v. 15, n. 2, p. 545–548, 2001.
- CANÁRIO, A. et al. Presence of parasite DNA in clinically unaffected nasal mucosa during Cutaneous Leishmaniasis caused by *Leishmania (V.) braziliensis*. **Clinical**

Microbiology and Infection, 2019, no prelo

CASTRO NETO, A. L. et al. In silico characterization of multiple genes encoding the GP63 virulence protein from *Leishmania braziliensis*: identification of sources of variation and putative roles in immune evasion. **BMC Genomics**, Londres, v. 20, p. 1–17, 2019.

CECÍLIO, P. et al. Deception and manipulation: The arms of *Leishmania*, a successful parasite. **Frontiers in Immunology**, Bethesda, v. 5, p. 1–16, 2014.

CENTERS FOR DISEASE CONTROL AND PREVENTION (Estados Unidos). **Biology**. Atlanta, 27 Jul. 2018. Disponível em: <<https://www.cdc.gov/parasites/leishmaniasis/biology.html>>. Acesso em: 15 dez. 2018.

CHANG, K.; MCGWIRE, B. S. Molecular determinants and regulation of *Leishmania* virulence. **Kinetoplastid Biology and Disease**, Bulgaria, v. 7, p. 1–7, 2002.

CHAPPUIS, F. et al. Visceral leishmaniasis: What are the needs for diagnosis, treatment and control? **Nature Reviews Microbiology**, Londres, v. 5, n. 11, p. 873–882, 2007.

CLEM, A. A Current Perspective on Leishmaniasis. **Journal of Infectious Diseases**, Oxford, v. 2, n. 2, p. 124–126, 2010.

COSTA-SILVA, M. F. et al. Gene expression profile of cytokines and chemokines in skin lesions from Brazilian Indians with localized cutaneous leishmaniasis. **Molecular Immunology**, Elmsford, v. 57, p. 74–85, 2014.

CUEVAS, I. C.; CAZZULO, J. J.; SÁNCHEZ, D. O. gp63 homologues in *Trypanosoma cruzi*: Surface antigens with metalloprotease activity and a possible role in host cell infection. **Infection and Immunity**, Washington, v. 71, n. 10, p. 5739–5749, 2003.

DAWIT, G. A Review on Biology, Epidemiology and Public Health Significance of Leishmaniasis. **Journal of Bacteriology & Parasitology**, Barcelona, v. 4, p. 2–7, 2013.

DU, R. et al. Old World Cutaneous Leishmaniasis and Refugee Crises in the Middle East and North Africa. **PLoS Neglected Tropical Diseases**, São Francisco, v. 10, n. 5, p. 1–11, 2016.

DUTHIE, M. S. et al. The development and clinical evaluation of second-generation leishmaniasis vaccines. **Vaccine**, Kidlington, v. 30, n. 2, p. 134–141, 2012.

FROMMEL, T. O. et al. The major surface glycoprotein (GP63) is present in both life stages of *Leishmania*. **Molecular and Biochemical Parasitology**, Amsterdam, v. 38, n. 1, p. 25–32, 1990.

- FUENTE, S. G. et al. Complete and de novo assembly of the *Leishmania braziliensis* (M2904) genome. **Memorias do Instituto Oswaldo Cruz**, Rio de Janeiro, v. 114, p. 1–6, 2019.
- GIUDICE, A. et al. Macrophages participate in host protection and the disease pathology associated with *Leishmania braziliensis* infection. **BMC Infectious Diseases**, Londres, v. 12, p. 1–9, 2012.
- GOLLOB, K. J.; VIANA, A. G.; DUTRA, W. O. Immunoregulation in Human American Leishmaniasis: Balancing Pathology and Protection. **Parasite Immunology**, Oxford, v. 36, n. 8, p. 367–376, 2014.
- GÓMEZ, M. A.; OLIVIER, M. Proteases and phosphatases during leishmania-macrophage interaction: Paving the road for pathogenesis. **Virulence**, Austin, v. 1, n. 4, p. 314–318, 2016.
- GONTIJO, B. Leishmaniose tegumentar americana. **Medicina Tropical**, Uberaba, v. 36, n. 13, p. 71–80, 2003.
- GRAMICCIA, M. Veterinary Parasitology Recent advances in leishmaniosis in pet animals : Epidemiology , diagnostics and anti-vectorial prophylaxis. **Veterinary Parasitology**, Amsterdam, v. 181, n. 1, p. 23–30, 2011.
- GUERIN, P. J. et al. Personal view Visceral leishmaniasis Visceral leishmaniasis : current status of control , diagnosis , and treatment , and a proposed research and development agenda Personal view. **The Lancet Infectious Diseases**, Londres, v. 2, p. 494–501, 2002.
- HAMERS, R. et al. Plasticity of gp63 gene organization in *Leishmania* (*Viannia*) *braziliensis* and *Leishmania* (*Viannia*) *peruviana*. **Parasitology**, Cambridge, v. 111, n. 3, p. 265–273, 1995.
- HIDE, M. et al. Understanding Human Leishmaniasis: The Need for an Integrated Approach. In: TIBAYRENC, M. (Ed.). **Encyclopedia of Infectious Diseases: Modern Methodologies**. New Jersey: Wiley-Liss, 2007. p. 87–123.
- HOOPER, N. M. Families of zinc metalloproteases. **FEBS letters**, Amsterdam, v. 354, n. 1, p. 1–6, 1994.
- HUELSENBECK, J. P.; RONQUIST, F. MRBAYES : Bayesian inference of phylogenetic trees. **Bioinformatics**, Oxford, v. 17, n. 8, p. 754–755, 2001.
- INDAR, E. et al. The Transcriptome of *Leishmania major* Developmental Stages in Their Natural Sand Fly Vector. **American Society For Microbiology**, Washington, v. 8, n. 2, p. 1–18, 2017.
- ISNARD, A. et al. Impact of *Leishmania* metalloprotease GP63 on macrophage signaling. **Frontiers in Cellular and Infection Microbiology**, Louisville, v. 2, n. 72, p. 1–9, 2012.

JECNA, L. et al. The role of surface glycoconjugates in *Leishmania* midgut attachment examined by competitive binding assays and experimental development in sand flies. **Parasitology**, Cambridge, v. 140, n. 8, p. 1026–1032, 2013.

JOSHI, P. B. et al. Targeted gene deletion in *Leishmania major* identifies leishmanolysin (GP63) as a virulence factor. **Molecular & Biochemical Parasitology**, Amsterdam, v. 120, p. 33–40, 2002.

KELLY, B. L.; NELSON, T. N.; MCMASTER, W. R. Stage-specific expression in *Leishmania* conferred by 3' untranslated regions of *L. major* leishmanolysin genes (GP63). **Molecular and Biochemical Parasitology**, Amsterdam, v. 116, n. 1, p. 101–104, 2001.

KEVRIC, I.; CAPPEL, M. A.; KEELING, J. H. New World and Old World *Leishmania* Infections: A Practical Review. **Dermatologic Clinics**, New York, v. 33, n. 3, p. 579–593, 2015.

KULKARNI, M. M. et al. *Trypanosoma cruzi* GP63 proteins undergo stage-specific differential posttranslational modification and are important for host cell infection. **Infection and Immunity**, Washington, v. 77, n. 5, p. 2193–2200, 2009.

LANGMEAD, B.; SALZBERG, S. Fast gapped-read alignment with Bowtie 2. **Nature Methods**, New York, v. 9, n. 4, p. 357–359, 2012.

LIMA, A. K. C. et al. Dissimilar peptidase production by avirulent and virulent promastigotes of *Leishmania braziliensis*: inference on the parasite proliferation and interaction with macrophages. **Parasitology**, Amsterdam, v. 136, p. 1179–1191, 2009.

LYNCH, M.; CONERY, J. S. The Origins of Genome Complexity. **Science**, Washington, v. 302, 2003.

MA, L. et al. An evolutionary analysis of trypanosomatid GP63 proteases. **Parasitology research**, Berlin, v. 109, n. 4, p. 1075–84, 19 out. 2011.

MAROLI, M. et al. Phlebotomine sandflies and the spreading of leishmaniases and other diseases of public health. **Medical and Veterinary Entomology**, [S. l.], v. 27, n. 2, p. 123–147, 2012.

MARSHALL, S. et al. Extracellular release of virulence factor major surface protease via exosomes in *Leishmania infantum* promastigotes. **BMC Genomics and Vectors**, Londres, v. 11, n. 1, p. 1–10, 2018.

MATHEOUD, D. et al. *Leishmania* Evades Host Immunity by Inhibiting Antigen Cross-Presentation through Direct Cleavage of the SNARE VAMP8. **Cell Host & Microbe**, New York, v. 14, n. 1, p. 15–25, 2013.

MAURICIO, I. L. et al. Glycoprotein 63 (gp63) genes show gene conversion and reveal the evolution of Old World *Leishmania*. **International Journal For Parasitology**, Oxford, v. 37, n. 5, p. 565–576, 2006.

MAZUMDER, S.; GANGULY, A.; ALI, N. The effect of C-terminal domain deletion on the catalytic activity of *Leishmania donovani* surface proteinase GP63: Role of Ser446 in proteolysis. **Biochimie**, Paris, v. 92, n. 12, p. 1876–1885, 2010.

MCKERROW, J. H. et al. Proteases in Parasitic Diseases. **The Annual Review of Pathology: Mechanisms of Disease**, Palo Alto, v. 1, p. 497–536, 2005.

MEDINA, L. S. et al. The gp63 gene cluster is highly polymorphic in natural leishmania (*viannia*) *braziliensis* populations, but functional sites are conserved. **PLoS ONE**, [S. l.], v. 11, n. 9, p. 1–13, 2016.

MINISTÉRIO DA SAÚDE. Manual de vigilância da leishmaniose tegumentar americana. **Brasília-DF: Ministério da Saúde, Brasil**, p. 71, 2007.

MIYOSHI, S. ICHI; SHINODA, S. Microbial metalloproteases and pathogenesis. **Microbes and Infection**, Paris, v. 2, n. 1, p. 91–98, 2000.

MURASE, L. S. et al. The role of metalloproteases in *Leishmania* species infection in the New World: a systematic review. **Parasitology**, Amsterdam, p. 1–11, 2018.

NADLER, C.; ENK, C. D.; LEON, G. T. Diagnosis and Management of Oral Leishmaniasis — Case Series and Literature Review. **Journal of Oral Maxillofacial Surgery**, Filadélfia, v. 72, n. 5, p. 927–934, 2014.

NOVOZHILOVA, N.; BOVIN, N. Structure, Functions, and Biosynthesis of Glycoconjugates of *Leishmania* spp. Cell surface. **Biochemistry**, Easton, v. 75, n. 6, p. 686–694, 2010.

OLIVIER, M. et al. *Leishmania* virulence factors: Focus on the metalloprotease GP63. **Microbes and Infection**, Paris, v. 14, p. 1377–1389, 2012.

OLIVIER, M.; GREGORY, D. J. Subversion Mechanisms by Which *Leishmania* Parasites Can Escape the Host Immune Response : a Signaling Point of View. **Clinical Microbiology Reviews**, Washington, v. 18, n. 2, p. 293–305, 2005.

ORYAN, A.; AKBARI, M. Worldwide risk factors in leishmaniasis. **Asian Pacific Journal of Tropical Medicine**, Haikou, v. 9, n. 10, p. 925–932, 2016.

PACE, D. Leishmaniasis. **Journal of Infection**, Londres, v. 69, p. 10–18, 2014.

PEACOCK, C. S. et al. Comparative genomic analysis of three *Leishmania* species that cause diverse human disease. **Nature Genetics**, Londres, v. 39, n. 7, p. 839–847, 2007.

RAMAMOORTHY, R. et al. Three distinct RNAs for the surface protease gp63 are differentially expressed during development of *Leishmania donovani* chagasi promastigotes to an infectious form. **Journal of Biological Chemistry**, Bethesda, v. 267, n. 3, p. 1888–1895, 1992.

REITHINGER, R. et al. Cutaneous leishmaniasis. **Lancet infectious disease**,

Londres, v. 7, p. 581–596, 2007.

RODRIGUES, M. G. DE A. et al. The role of deforestation on American Cutaneous Leishmaniasis incidence: spatial-temporal distribution, environmental and socioeconomic factors associated in the Brazilian Amazon. **Tropical Medicine & International Health**, Oxford, v. 24, p. 348-355, 2019.

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

ROTUREAU, B. Are New World leishmaniases becoming anthroponoses? **Medical Hypotheses**, Penrith, v. 67, p. 1235–1241, 2006.

SANYAL, T.; GHOSH, D. K.; SARKAR, D. Immunoblotting identifies and antigen recognized by anti gp63 in the immune complexes of Indian kala-azar patient sera. **Molecular and Cellular Biochemistry**, Dordrecht, v. 130, n. 1, p. 11–17, 1994.

SAVOIA, D. Recent updates and perspectives on leishmaniasis. **The Journal Of Infection in Developing Countries**, [S. l.], v. 9, n. 6, p. 588–596, 2015.

SCHLAGENHAUF, E.; ETGES, R.; METCALF, P. The crystal structure of the Leishmania major surface proteinase leishmanolysin (gp63). **Structure**, [S. l.], v. 6, n. 8, p. 1035–1046, 1998.

SCHNEIDER, P. et al. Leishmania major : Differential Regulation of the Surface Metalloprotease in Amastigote and Promastigote Stages. **Experimental Parasitology**, San Diego, v. 206, p. 196–206, 1992.

SCOTT, P.; NOVAIS, F. O. Cutaneous leishmaniasis : immune responses in protection and pathogenesis. **Nature reviews. Immunology**, Londres, v. 16, n. 9, p. 581–592, 2016.

SILVA-ALMEIDA, M. et al. Proteinases as virulence factors in Leishmania spp . infection in mammals. **Parasites & vectors**, Londres, v. 5, n. 1, p. 1–10, 2012.

SILVA-ALMEIDA, M. et al. Overview of the organization of protease genes in the genome of Leishmania spp. **Parasites and Vectors**, Londres, v. 7, n. 1, p. 1–7, 2014.

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

SOARES, R. P. et al. In Vitro Inhibition of Leishmania Attachment to Sandfly Midguts and LL-5 Cells by Divalent Metal Chelators, Anti-gp63 and Phosphoglycans. **Protist**, Jena, v. 168, n. 3, p. 326–334, 2017.

SOONG, L.; MELBY, P. C. Immunopathogenesis of non-healing American cutaneous leishmaniasis and progressive visceral leishmaniasis. **Seminars in**

Immunopathology, Berlin, v. 34, p. 735–751, 2012.

STEVERDING, D. The history of leishmaniasis. **Parasites & vectors**, Londres, v.10, n. 1, p. 1–10, 2017.

SUTTER, A. et al. Structural insights into leishmanolysins encoded on chromosome 10 of *Leishmania (Viannia) braziliensis*. **Memorias do Instituto Oswaldo Cruz**, Rio de Janeiro, v. 112, n. 9, p. 617–625, 2017.

TORRES-GUERRERO, E. et al. Leishmaniasis : a review. **F1000Research**, Londres, v. 6, p. 1–15, 2017.

VALDIVIA, H. O. et al. Comparative genomic analysis of *Leishmania (Viannia) peruviana* and *Leishmania (Viannia) braziliensis*. **BMC Genomics**, Londres, p. 1–10, 2015.

VESTERBERG, O. A short history of electrophoretic methods. **Electrophoresis**, Weinheim, v. 14, n. 1, p. 1243–1249, 1993.

VICTOIR, K.; DUJARDIN, J. How to succeed in parasitic life without sex ? Asking *Leishmania*. **Trends in parasitology**, Oxford, v. 18, n. 2, p. 81–85, 2002.

VOTH, B. R. et al. Differentially expressed *Leishmania major* gp genes encode cell surface leishmanolysin with distinct signals for 63 glycosylphosphatidylinositol attachment. **Molecular & Biochemical Parasitology**, Amsterdam, v. 93, n. 1, p. 31–41, 1998.

WILSON, M. E. et al. The Effect of Ongoing Protein Synthesis on the Steady State Levels of Gp63 RNAs in *Leishmania chagasi*. **The Journal of Biological Chemistry**, Bethesda, v. 268, n. 21, p. 15731–15736, 1993.

YAO, C.; DONELSON, J. E.; WILSON, M. E. The major surface protease (MSP or GP63) of *Leishmania* sp . Biosynthesis , regulation of expression , and function. **Molecular & Biochemical Parasitology**, Amsterdam, v. 132, n. 1, p. 1–16, 2003.

ANEXO A – ALINHAMENTO DE SEQUÊNCIAS GP63

LbrM.10.0590	1	MPLDSSSTHRRRSVAARLVRLAAAGVAALAVGTAAAWA	x	AAATP	HRCIHDKLQARV	RDSAAHRRMPPSAVSAVGLPYIA	80
LbrM.10.0610	1	MPLDSSSTHRRRSDAARLMRLAAAG--LVMAVGAAAWAQAG--		HC	CIHDKLQARV	QSVAQQRSPAVS	VSAVGLPYVS 75
100520PCR	1	MSRDRSSTPLRRSDAARLMRLAAAG--LVMAVGAAAVWQAAG--		HC	CIHDKLQTRV	LQSVQQHRPPG	SVSALGLPYVS 76
LbrM.10.0540	1	-----MRLAAAG--LVMAVGAAAVWQAAG--		HC	CIHDKLQTRV	LQSVQQHRPPG	SVSALGLPYVS 58
LbrM.10.1630	1	MSRDRSSTHRRRSVAARLMRLAAAG--LVMAVGAAAVWQAAG--		HC	CIHDKLQARV	QSVAQQHRPPG	SVSALGLPYVS 76
LbrM.10.1610	1	MSRDRSSTHRRRSVAARLMRLAAAG--LVMAVGAAAVWQAAG--		HC	CIHDKLQARV	QSVAQQHRPPG	SVSALGLPYVS 76
LmjF.10.0470	1	MSVDSSSTHRRRCVAARLVRLAAAGAAVTVAVGTAAWA		HAGALQ	HRCVHDAMQARV	RQSVADH	HKAPGAVSAVGLPYVT 80
LmjF.10.0460	1	MSVDSSSTHRRRCVAARLVRLAAAGAAVTVAVGTAAWA		HAGALQ	HRCVHDAMQARV	RQSVADH	HKAPGAVSAVGLPYVT 80
LinJ.10.0500	1	MSVDSSSTHRRRSVAARLVRLAAAGAAVIAVGTAAAWA		HAGAVQ	HRCIHDAMQARV	RQSVARH	HHTAPGAVSAVGLPYVT 80
LinJ.10.0510	1	MSVDSSSTHRRRSVAARLVRLAAAGAAVIAVGTAAAWA		HAGAVQ	HRCIHDAMQARV	RQSVARH	HHTAPGAVSAVGLPYVT 80

LbrM.10.0590	81	ID-----AADTVARAADWGL	O	LRIAVSTADLTD	PDYHCTRV	QVRVSNHAGAI	IVTCTAEDV	LTEEKRD	IILVS 145
LbrM.10.0610	76	AGTISSAHTVDWALADSTSPSVVRAADWGL		LRIAVSTADLTD	PDYHCTRV	QVRVSNHAGAF	VFTCTAEDIL	TEEKRD	IILVT 155
100520PCR	77	A-----DWALADSTSPSVVRAADWGL		LRIAVSTADLTD	PDYHCTRV	QVRVSNHAGAL	DMCKAEDILT	DAKRNT	ILVG 147
LbrM.10.0540	59	A-----DWALADSTSPSVVRAADWGL		LRIAVSTADLTD	PDYHCTRV	QVRVSNHAGAL	DMCKAEDILT	DAKRNT	ILVG 129
LbrM.10.1630	77	ADPISSAHTVDWAQADSTSPSVVRAADWGL		LRIAVSTADLTD	PDYHCTRV	QVRVSNHAGAL	DICKEDILT	DAKRNT	ILIS 156
LbrM.10.1610	77	ADPISSAHTVDWAQADSTSPSVVRAADWGL		LRIAVSTADLTD	PDYHCTRV	QVRVSNHAGAL	DICKEDILT	DAKRNT	ILIN 156
LmjF.10.0470	81	LDAAHTAAAAD--PRPGSARVSVVVDVNWGAL		LRIAVSTEDLTD	PAYHCA	RVQVRVSNHAGAI	IVTCTAEDILT	NEKRN	IILVK 158
LmjF.10.0460	81	LDAAHTAAAAD--PRPGSARVSVVVDVNWGAL		LRIAVSTEDLTD	PAYHCA	RVQVRVSNHAGAI	IVTCTAEDILT	NEKRN	IILVK 158
LinJ.10.0500	81	ID---TAAAAD--RRPGSAPTVVRAANWGAL		LRIAVSTEDLTD	PAYHCA	RVQVRVSNHAGAI	IVTCTAEDILT	DEKRD	IILVK 155
LinJ.10.0510	81	ID---TAAAAD--RRPGSAPTVVRAANWGAL		LRIAVSTEDLTD	PAYHCA	RVQVRVSNHAGAI	IVTCTAEDILT	DEKRD	IILVK 155

LbrM.10.0590	146	YLIPQALQLHAEERL	RVRQVQGS	SWKVTGMT	SDVCGE	EFKVP	EAHVAKGV	SNAD	DFVLYVASVPSE	EGVLAWATTCQV	FSDDHP	225
LbrM.10.0610	156	YLIPQALQLHAEERL	RVRQVQGS	SWKVTGMT	SDVCGE	EFKVP	EAHVAKGV	SNAD	DFVLYVASVPSE	EGVLAWATTCQV	FSDDHP	235
100520PCR	148	SIIPQALQLHATERL	KVRQVQGS	SWKVTGMT	SDVCGE	EFKVP	EAHVAKGV	SNAD	DFVLYVASVPSE	EGVLAWATTCQV	FSDDHP	227
LbrM.10.0540	130	SIIPQALQLHATERL	KVRQVQGS	SWKVTGMT	SDVCGE	EFKVP	EAHVAKGV	SNAD	DFVLYVASVPSE	EGVLAWAVMCQAF	FD	DRP 209
LbrM.10.1630	157	DLPLALQLHATERL	EVRQVQGT	TKWVTGME	SDVCGE	EFKVP	EAHVAKGV	SNAD	DFVLYVASVPSE	EGVLAWAVMCQAF	FD	DRP 236
LbrM.10.1610	157	ILPLALQLHATERL	KVRQVQGT	TKWVTGME	SDVCGE	EFKVP	EAHVAKGV	SNAD	DFVLYVASVPSE	EGVLAWAVMCQAF	FD	DRP 236
LmjF.10.0470	159	YLIPQAVQLHAEERL	KVRQVQGS	SWKVTGMT	SDVCGE	EFKVP	EAHVAKGV	SNAD	DFVLYVASVPSE	EGVLAWATTCQV	FSDGHP	238
LmjF.10.0460	159	YLIPQAVQLHATERL	KVRQVQGS	SWKVTGMT	SDVCGE	EFKVP	EAHVAKGV	SNAD	DFVLYVASVPSE	EGVLAWATTCQV	FSDGHP	238
LinJ.10.0500	156	HLIPQALQLHATERL	KVRQVQGS	SWKVTGMT	SDVCGE	EFKVP	EAHVAKGV	SNAD	DFVLYVASVPSE	EGVLAWATTCQV	FSDGHP	235
LinJ.10.0510	156	YLIPQALQLHATERL	KVRQVQGS	SWKVTGMT	SDVCGE	EFKVP	EAHVAKGV	SNAD	DFVLYVASVPSE	EGVLAWATTCQV	FSDGHP	235

LbrM.10.0590	226	AVGVMNIPAAANIVSRYDQ	STTRIVT	THEVAHALGFSSV	FFENAG	IIVKSV	NTNLRGKPF	AA	VPVINSST	VAKAREQYGC	P	TLE 305			
LbrM.10.0610	236	AVGVMNIPAAANIVSRYDQ	STTRIVT	THEVAHALGFSSV	FFENAG	IIVKSV	NTNLRGKPF	AA	VPVINSST	VAKAREQYGC	P	TLE 315			
100520PCR	228	AVGVMNIPAAANIVSRYDQ	IMLRTVA	HAHEVAHALGF	DLTV	FE	DVGIV	GDV	NTNLRGRG	YF	VPVLS	SPTVA	VAKAREQYGC	P	TLE 307

LbrM.10.0540 210 AVGVINIPAAV IQSAYDQFMLRSVTHEMAHALGF DLLIFE ELDLLDVVSNLRGKDYEPVVLNSPTVAVAKAREQYGCA TLT 289
 LbrM.10.1630 237 AVGVVNIPTAYIQSAYDQIMVRTVAHEVAHALGFDLTA FE DALELIHDVKDLERNYEPVVLSSPTVAVAKAREQYGCTLT 316
 LbrM.10.1610 237 AVGVINIPAVNIRSPYDQLMIRTVVAHEVAHALGFNRIF D SFGMVTAAGIRGKDYYPV LNTPTVAVAKAREQYGCTL S 316
 LmjF.10.0470 239 AVGVINIPAAV IQSAYDQLVTRVVTHEMAHALGFGPFFEDARIVANVSNVRGKNFDVPVINSSTAVAKAREQYGCDTLE 318
 LmjF.10.0460 239 AVGVINIPAAV IQSAYDQLVTRVVTHEMAHALGFGPFFEDARIVANVSNVRGKNFDVPVINSSTAVAKAREQYGCDTLE 318

LinJ.10.0500 236 AVGVINIPAAV IQSAYDQLVTRVVTHEMAHALGFGVFFEGARILESISNVRHKDFDVPVINSSTAVAKAREQYGCDTLE 315
 LinJ.10.0510 236 AVGVINIPAAV IQSAYDQLVTRVVTHEMAHALGFGTFFTEILLVTQMMNIRGKDFNVSINSSTAVAKAREQYGCDTLE 315

LbrM.10.0590 306 YLEVEDQGGSGSAGSHLKGRNAKDELMAPASAAGYYTALTMAVFE DLGFKADFTKAEVMPWGRNASCDFLTKKCMERNI 385
 LbrM.10.0610 316 YLEVEDQGGSGSAGSHLKGRNAKDELMAPASAAGYYTALTMAVFE DLGFKADFSMAEVMPWGRNASCDFITNKCMEDNI 395
 100520PCR 308 FLELEDTGGSTAGSHLKGRNAKDELMAPAVTAGYYTALTMAVFQDLGFYQVDFSMAEVMPWAEHASCDFITNKCMVNI 387
 LbrM.10.0540 290 FLEVEDTGGSTAGSHLKGRNAKDELMAPVMGAGYYTALTMATFE DLGFRVDFTKAEVMPWAEHASCDFITNKCMEDNI 369
 LbrM.10.1630 317 FLELEDTGGSGFAGSHLKGRNAKDELMAPASDAGYYTALTMAVFQDLGFYKADFTKAEVMPWANLASCDFITKCMENNI 396
 LbrM.10.1610 317 FLELEDKGGFASLGSHLKGRNAKDELMSSVVKGGYYTALTMAVFQDLGFYQADFSMAEVMPWAYLATCDFLTNKCMERNI 396
 LmjF.10.0470 319 YLEVEDQGGAGSAGSHIKMRNAQDELMAAAAAGYYTALTMATFQDLGFYQADFSKAEVMPWQNAACAFLTNKCMEQNI 398
 LmjF.10.0460 319 YLEVEDQGGAGSAGSHIKMRNAQDELMAAAAAGYYTALTMATFQDLGFYQADFSKAEVMPWQNAACAFLTNKCMEQSV 398
 LinJ.10.0500 316 YLEIEDQGGAGSAGSHIKMRNAQDELMAAAAAGYYTALTMATFQDLGFYQADFSKAEVMPWGRNACCAFLSEKCMERNI 395
 LinJ.10.0510 316 YLEIEDQGGAGSAGSHIKMRNAKDELMAPAAAAGYYTALTMATFQDLGFYQADFSKAEVMPWGRNACCAFLSEKCMERNI 395

LbrM.10.0590 386 TQWPEMFCNTE DENALRCPTDRLGLGGCIVLTHT-SVPEYFOYFTDPTITGLSDFMDYCPVVPYDDGSCAQORASETSSDM 464
 LbrM.10.0610 396 TQWPEMFCNTEERRYRCPTDRLKLGTCIRTYSTPMPYFOYFTNAI LGGFSEFLDYCPFIVDYSNACNODPMSMASPSL 475
 100520PCR 388 TQWPGMFCNSTESLYRCFSDRLKIGRCITVTHDDPLPTVYFOYFTEPSVGGRLTFMDFCPVIGYSNTACDQDPSTAPSF 467
 LbrM.10.0540 370 TQWPEMFCNSTESSYRCPTDRLKIGRCISLATYDDPLPTVYFOYFTKTS LGGSEFMDFCPIVIGYSNACNODPSTASPAL 449
 LbrM.10.1630 397 TQWPEMFCNTEERRYRCFSDRLKIGTCISVTKHDMPLPTVYRYFTKTSVGGSEFMDFCPIVVPYGTAAACNODPSTASPSL 476
 LbrM.10.1610 397 TQWPGMFCNTE DVLLRCSTDRLLTGTCIFQQRKRPLPTVYFOYFTDSEF LGGSPFMDYCPVVDITYPDGACNODPMSMASPSL 476
 LmjF.10.0470 399 TQWPA MFCNESEDAIRCPTSRLLLGTCIFREYELPLPRVYQYFTNASLGGYSPFLDYCPFVIDYADGSCNODASSAEFF 478
 LmjF.10.0460 399 TQWPA MFCNESEDAIRCPTSRLLSLGCGVTRHP-GLPPYQYFTDPSL AGLS AFMDYCPVVPYSDGSCTQRASEA HASL 477
 LinJ.10.0500 396 TKWPA MFCNESEVTMRCPTSRLLSLGCGVTRHP-DLPPYQYFTDPSL AGLS AFMDYCPVVPYSDGSCTQRASEA GAPF 474
 LinJ.10.0510 396 TKWPA MFCNESEVTMRCPTSRLLMVGTCIFRGYSTPFSLYQYFTNASLGGYSPFLDYCPFVIGYSNACNODASL AAGFF 475

LbrM.10.0590 465 QAFNVFSDAARCLDGAFRPT-TTREDVTYAGMCANVKCDTAAARTYSVQVRGSSGYVACTPGESVELATLSAAAFVNGSYIT 543
 LbrM.10.0610 476 QAFNVFSDAARCFDGVFQPRNSNARSEPNNALCANVMCDTAAARTYSVQVRGSSGYVACTPGESIDLATLSAAAFVNGSYIT 555
 100520PCR 468 KEFNVFSDSSRCFDGTFERPKHDETPPGPYNGMCANVKCDRDHHTYSVQVRGSTGYVACTPGERLDLATLSTTFVEGSYIM 547
 LbrM.10.0540 450 MEFNVFSDSSRCFDGILLTPKHNTGPEQYNALCANVMCDRAHHTYSVQVRGSSGYVACTPGERVELTTSTAFVEGSYIT 529
 LbrM.10.1630 477 KEFNVFSDAARCLDGAFRPKHSTGPPGPNGLCANVKCDRAHHTYSVQVRGSSGYVACTPGERLELATLSTAFVEGSYIM 556
 LbrM.10.1610 477 QAFNVFSDAARCLDGVFRPKHGVHGNBYNGLCANVKCDRVHHRYSVQVRGSSGYVACTPQSIELATLSTDAFVEGSYIM 556
 LmjF.10.0470 479 TAFNVFSDAARCLDGAFRPKATNGIVKSYAAGLCANVCDTATRTYSVQVRGSSNDYTNCTPGLRVELSTVSKTFEEGYSYIT 558

```

LmjF.10.0460 478 LFNVFSDAARCIIDGAFRPKATDGVKSYAGLCANVQCDTATRTYSVQVHGSGNDYTNCTPGLRVELSTVSNAFEEGGYIT 557
LinJ.10.0500 475 KGFNVFSDAARCIIDGAFRPKTSHGIIKSYAGLCANVRCDTATRTYSVQVHGSGYANCTPGLRVELSTVSSAFEEGGYIT 554
LinJ.10.0510 476 SAFNVFSDAARCIIDGAFRPKNRTAANGYAGLCANVRCDTATRTYSVQVHGSMDYVNCCTPGLRVELSTVSSAFEEGGYIT 55

```

```

* * + + C' f
LbrM.10.0590 544 CAPYVEVCQANVQGATS-----SGNAAAGRRGPRAAATALLVAALLAIACA----- 589
LbrM.10.0610 556 CAPYVEVCQANIKGLIDFE-----RDAADTAAMRQWSERMYVIAFTVTAVLLGIVLAAMAGLVVGLLVISLS---- 621
100520PCR 548 CPPYVEVCQA----- 557
LbrM.10.0540 530 CAPYVEVCQANIKGVIDFE-----RDAADTAAV----- 557
LbrM.10.1630 557 CPLLYVEVCQANIKGVIDFE-----CDAADTAAV----- 584
LbrM.10.1610 557 CPLLYVEVCQANIKGVIDFE-----CDAADTAAV----- 584
LmjF.10.0470 559 CPPYVEVCQGNVQAAKDDGDSDSSSSSDAADKAAIERWNERMAGLATAAATVLLGVVLSLMALVVVWLLLLVSCPRWCCK 638
LmjF.10.0460 558 CPPYVEVCQGNVQAAKD-----GGNTAAGRRGPRAAATALLVAALLAVAL----- 602
LinJ.10.0500 555 CPPYVEVCQGNVQAAKD-----GGNAAAGRRGPRAAATALLVAALLAVAL----- 599
LinJ.10.0510 556 CPPYVEVCQANVKGAKDFAGDSDSSSSAGDAADRAAMQRWNDRMAGLATAAMVLLGMVLSLMALVVVWLLLLLTCPPWWCCK 635

```

```

LbrM.10.0590 589 ----- 589
LbrM.10.0610 621 ----- 621
100520PCR 557 ----- 557
LbrM.10.0540 557 ----- 557
LbrM.10.1630 584 ----- 584
LbrM.10.1610 584 ----- 584
LmjF.10.0470 639 VGGLPT 644
LmjF.10.0460 602 ----- 602
LinJ.10.0500 599 ----- 599
LinJ.10.0510 636 FGGLPT 641

```

+: Sítio de ligação GPI
HEXXHXXGXXH - sunlinhado

X = Sítio de remoção - peptídeo sinal
O = Sítio de remoção - pró-peptídeo
* = cisteínas

ANEXO B – PARECER DO COMITE DE ÉTICA



Ministério da Saúde
FIOCRUZ
 Fundação Oswaldo Cruz
 Centro de Pesquisas Aggeu Magalhães

COMISSÃO DE ÉTICA NO USO DE ANIMAIS

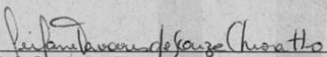
Certificado de Aprovação

Certificamos que o projeto intitulado: “ESTUDO DOS MÚLTIPLOS PARÁLOGOS DE GP63 DE *leishmania braziliensis* E SEU PAPEL NA RELAÇÃO PARASITO-HOSPEDEIRO” protocolado sob nº 93/2015 pelo (a) pesquisador (a) Dr **Oswaldo Pompílio de Melo Neto** está de acordo com a Lei 11.794/2008 e foi aprovado pela COMISSÃO DE ÉTICA NO USO DE ANIMAIS do Centro de Pesquisas Aggeu Magalhães / Fundação Oswaldo Cruz (CEUA/CPqAM) em 15/06/2016. Na presente versão, este projeto está licenciado e tem validade até junho de 2020,

Quantitativo de Animais Aprovados	
Espécie	Nº de Animais
Camundongo heterogênico Swiss fêmea	60
Camundongo isogênico BALB/c fêmea	60
Coelho Nova Zelândia Macho e/ou Fêmea	20
TOTAL	140

We certify that project “ESTUDO DOS MÚLTIPLOS PARÁLOGOS DE GP63 DE LEISHMANIA BRAZILIENSIS E SEU PAPEL NA RELAÇÃO PARASITO-HOSPEDEIRO” (CEUA Protocol nº 93/2015) coordinated by **OSVALDO POMPÍLIO DE MELO NETO** is according to the ethical principles in animal research adopted by the Brazilian law 11.794/2008 and so was approved by the Ethical Committee for Animal Research of the Centro de Pesquisas Aggeu Magalhães / Fundação Oswaldo Cruz on June, 15, 2016. In present version this project is licensed and valid until June 2020.

Recife (PE, BRAZIL) May, 20, 2016


 Dr Gerlane Tavares de Souza Chioratto
 Vice-Coordenadora CEUA/CPqAM

Dr Gerlane Tavares de Souza Chioratto
 Vice-coordenadora da Comissão de Ética
 no Uso de Animais – CEUA
 Mat. SIAPE 1631158 / CRMV: PE 2073
 e-mail: gerlane@cpqam.fiocruz.br
 CPqAM/Fiocruz

Av. Professor Moraes Rego, s/n - Cidade Universitária – Campus da UFPE
 Recife - PE - CEP: 50.670-420
 Telefone: (81) 2101-2500/2101-2600 Fax: (81) 3453-1911
 www.cpqam.fiocruz.br

ANEXO C – ARTIGO PUBLICADO

Castro Neto et al. BMC Genomics _____
<https://doi.org/10.1186/s12864-019-5465-z>

BMC Genomics

RESEARCH ARTICLE

Open Access

1
 2 **In silico characterization of multiple genes**
 3 **encoding the GP63 virulence protein from**
 4 ***Leishmania braziliensis*: identification of**
 5 **sources of variation and putative roles**
 6 **in immune evasion**

7 Artur L. Castro Neto^{1,2}, Adriana N. A. L. M. Brito², Antonio M. Rezende², Franklin B. Magalhães³ and
 8 Osvaldo P. de Melo Neto^{2*}

Abstract

13 **Background:** The leishmaniasis are parasitic diseases caused by protozoans of the genus *Leishmania*, highly divergent
 14 eukaryotes, characterized by unique biological features. To survive in both the mammalian hosts and insect vectors,
 15 these pathogens make use of a number of mechanisms, many of which are associated with parasite specific proteases.
 16 The metalloprotease GP63, the major *Leishmania* surface antigen, has been found to have multiple functions required
 17 for the parasite's survival. GP63 is encoded by multiple genes and their copy numbers vary considerably between
 18 different species and are increased in those from the subgenus *Viannia*, including *L. braziliensis*.

19 **Results:** By comparing multiple sequences from *Leishmania* and related organisms this study sought to characterize
 20 paralogs in silico, evaluating their differences and similarities and the implications for the GP63 function. The *Leishmania*
 21 GP63 genes are encoded on chromosomes 10, 28 and 31, with the genes from the latter two chromosomes
 22 more related to genes found in insect or plant parasites. Those from chromosome 10 have experienced independent
 23 expansions in numbers in *Leishmania*, especially in *L. braziliensis*. These could be clustered in three groups associated
 24 with different mRNA 3' untranslated regions as well as distinct C-terminal ends for the encoded proteins, with presumably
 25 distinct expression patterns and subcellular localizations. Sequence variations between the chromosome 10 genes were
 26 linked to intragenic recombination events, mapped to the external surface of the proteins and predicted to
 27 be immunogenic, implying a role against the host immune response.

28 **Conclusions:** Our results suggest a greater role for the sequence variation found among the chromosome 10
 29 GP63 genes, possibly related to the pathogenesis of *L. braziliensis* and closely related species within the mammalian host.
 30 They also indicate different functions associated to genes mapped to different chromosomes. For the chromosome 10
 31 genes, variable subcellular localizations were found to be most likely associated with multiple functions and
 32 target substrates for this versatile protease.

33 **Keywords:** *Leishmania braziliensis*, GP63, Virulence proteins

* Correspondence: opmn@cpqam.fiocruz.br

²Instituto Aggeu Magalhães, Fundação Oswaldo Cruz (Fiocruz-Pernambuco),
 Recife, Pernambuco, Brazil

Full list of author information is available at the end of the article



© The Author(s). 2019 **Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated.

34 Background

35 The leishmaniasis are parasitic infectious diseases caused
36 by flagellated protozoa belonging to the genus *Leishmania*,
37 family Trypanosomatidae, and which are transmitted by
38 sandflies of the genera *Phlebotomus* or *Lutzomyia*. These
39 diseases are found as two major clinical forms, named as
40 cutaneous leishmaniasis (CL) and visceral leishmaniasis
41 (VL), with a global incidence for each in the range of
42 hundreds of thousands of cases per year [1]. Multiple
43 *Leishmania* species are associated with the leishman-
44 iasis and distinct species, closely related or not, are
45 responsible for the disease in different parts of the
46 world. Those belonging to the subgenus *Viannia* are
47 restricted to the New World (including *L. braziliensis* and
48 *L. guyanensis*), have evolved separately from better known
49 species belonging to the subgenus *Leishmania* (*L. major*,
50 *L. infantum*, *L. mexicana* and others) and are associated
51 with the mucocutaneous leishmaniasis (MCL), a more
52 aggressive variation of CL [2].

53 As successful pathogens, the various *Leishmania* species
54 have developed effective mechanisms to escape the mam-
55 malian host immune response and proliferate [3, 4]. Some
56 of these evasion mechanisms are dependent on proteases,
57 which help ensure that the parasites can invade the
58 mammalian tissue, survive, differentiate and multiply [5].
59 The GP63 protease, also known as leishmanolysin or
60 major surface protease (MSP), was first discovered in 1980
61 as the major surface antigen of the promastigote form of
62 many species of *Leishmania* [6]. It was later found to be
63 bound to the cell membrane through a GlycosylPhosphatidyl-
64 inositol (GPI) anchor and was also identified as an
65 important virulence factor. This is a zinc-dependent met-
66 alloproteinase, which belongs to the peptidase family M8
67 and the metzincin class and includes conserved features
68 such as the motif HEXXHXXGXXH and a pro-peptide lo-
69 cated in the protein's N-terminal region that renders the
70 proenzyme inactive during translation and is removed
71 during its maturation and activation. The GP63 proteins
72 also include an N-terminal signal sequence which directs
73 them to the endoplasmic reticulum and to the *Leishmania*
74 secretory pathway [7, 8].

75 GP63 has been found to play multiple roles during
76 *Leishmania* infection in mammals, starting in the extra-
77 cellular environment where it acts inactivating the
78 complement cascade, by cleaving C3b into iC3b. This
79 inactivation prevents the formation of the membrane
80 attack complex (MAC), despite allowing the opsonisa-
81 tion of the *Leishmania*, mediated by iC3b, and facilitat-
82 ing its phagocytosis. GP63 can also facilitate the
83 binding of the parasite to the macrophage through fi-
84 bronectin receptors, cleaving proteins from the host's
85 extracellular matrix. Within the macrophages, it also
86 acts to reduce the production of TNF, IL-12 and nitric
87 oxide, which contributes to the protection and survival

88 of the parasite, and provides the *Leishmania* with a faster
89 entry into the macrophage, through the activation of a host
90 tyrosine phosphatase [7, 9, 10]. GP63 has also been shown
91 to be released through exosomes into the extracellular
92 medium and this may facilitate its uptake by the macro-
93 phage even before the internalization of the *Leishmania*
94 parasite [11]. Lack of GP63 drastically reduces the
95 *Leishmania's* ability to establish and maintain an
96 infection, since the hosts are more likely to induce an
97 innate immunity inflammatory responses [12]. Within the
98 host cell cytoplasm GP63 has been shown to cleave the
99 transcriptional factor AP-1, which regulates the production
100 of pro-inflammatory cytokines and nitric oxide by the
101 macrophage [11, 13]. GP63 was also shown to be associated
102 with the inactivation of the mTOR kinase, leading to the
103 inhibition of protein synthesis in the macrophage and pro-
104 viding an ideal environment for the proliferation of the
105 pathogen [14].

106 Early studies have shown that GP63 is more abundantly
107 expressed in the promastigote stage of the *Leishmania* life
108 cycle, the proliferative stage within the insect vector. This
109 expression may peak during metacyclogenesis, when the
110 parasite prepares to infect the mammalian hosts, and is
111 subsequently reduced again upon differentiation into amas-
112 tigotes, the intracellular stage that multiplies within the
113 mammalian macrophages [7, 15, 16]. The abundant GP63
114 expression in promastigotes indicates relevant functions
115 also in the insect vector, presumably needed for survival
116 and proliferation. Indeed, a potential involvement in the
117 degradation of protein components that would lead to the
118 adhesion of the parasite in the insect gut epithelium has
119 been shown [17, 18]. Due to its wide substrate specificity,
120 GP63 may also perform a nutritional role for the parasite,
121 acting as an endopeptidase [19, 20], or even protect the
122 *Leishmania* against the insect defences [19].

123 Concerning the GP63 gene organization, there is a no-
124 ticeable variation in the number of gene copies encoding
125 these proteins among different *Leishmania* species. In *L.*
126 *major* these genes are present in more than one chromo-
127 some and multiple copies have been detected arranged in
128 tandem [21], with the same multi-copy arrangement also
129 found in *L. infantum* and *L. braziliensis* [22]. Noteworthy,
130 however, is the substantial increase in the number of gp63
131 genes reported for *L. braziliensis* and other species
132 belonging to the subgenus *Viannia*, when compared with
133 the subgenus *Leishmania*. This was reported in early stud-
134 ies [23–25] and has been confirmed more recently by
135 results derived from a screening for cosmids harboring
136 multiple GP63 genes from *L. braziliensis* [26], as well as
137 by genome sequencing data for different *Leishmania* spe-
138 cies [27–29]. No clear biological reasons are known, how-
139 ever, to explain this expansion in the GP63 gene copy
140 number. Here, aiming to contribute further to the under-
141 standing of the role of GP63 in *Leishmania* pathogenesis

142 in general but with a focus on *Viannia* species, we sought
 143 to investigate the GP63 gene expansion further, using a
 144 range of in silico tools. We started by better defining the
 145 extent of GP63 gene diversity in *L. braziliensis*, followed by
 146 an in-depth analysis of the similarities and differences
 147 between different genes from this and related *Leishmania*
 148 species. The GP63 genes were first grouped according to
 149 their chromosomal localization followed by phylogenetic
 150 comparisons between different trypanosomatid species.
 151 Further grouping according to sequence similarities or dif-
 152 ferences within non-coding and coding elements was also
 153 carried out, in order to define putative functional distinc-
 154 tions. Possible mechanisms associated with the gene expan-
 155 sion due to DNA recombination were then investigated
 156 and variations in sequence mapped on the GP63 structure
 157 and linked with predictions of immunogenic potential. Our
 158 results are consistent with a selective expansion of a subset
 159 of GP63 genes in *L. braziliensis* that might be linked to
 160 mammalian pathogenesis and might be required for a bet-
 161 ter protection against the host immune system.

162 Results

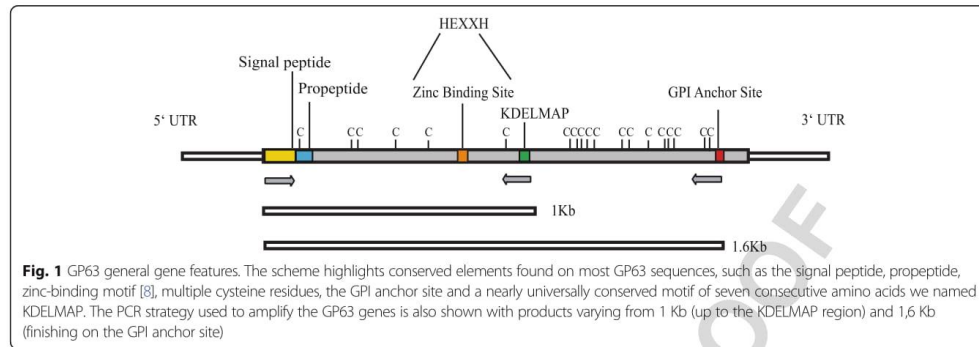
163 Search for new *L. braziliensis* GP63 paralogs

164 The early studies based on hybridization assays [23, 24]
 165 had suggested that the total number of GP63 genes found
 166 within *Leishmania* species belonging to the *Viannia* sub-
 167 genus is greater than the number of genes available at the
 168 TriTrypDB database and identified after the *L. braziliensis*
 169 genome sequencing and annotation. Recent data based on
 170 next generation sequencing have also suggested major vari-
 171 ations in copy number of GP63 genes between species
 172 within the same subgenus, *Leishmania* or *Viannia*, that
 173 have not yet been included on the annotated genomes
 174 [28–30]. Here, to begin to understand the true diversity of
 175 the *L. braziliensis* GP63 genes, we first sought to reevaluate
 176 the available *L. braziliensis* GP63 gene sequences consid-
 177 ering that the automatic annotation methods might have
 178 missed further genes. We therefore performed a reanalysis
 179 of the *L. braziliensis* genome sequences and searched for
 180 possible new GP63 paralogs that might not have been
 181 annotated. To do this we performed a search in the *L.*
 182 *braziliensis* genome using the Hidden Markov Models
 183 (HMMs) methodology [31], carried out after a grouping of
 184 the entire proteome set from different *Leishmania* species
 185 (described in methods). Nine subsets of GP63 sequences
 186 were created using the OrthoMCL tool in order to group
 187 these sequences and allow the search to be performed, as
 188 shown in Additional file 1: Table S1, with the number of
 189 genes in each subset varying in size from 56 to only two.
 190 All nine subsets were used to build HMMs and these were
 191 then applied for the search of new paralogs in the pre-
 192 dicted proteome from *L. braziliensis* 2904. In general, all
 193 HMMs were able to find the GP63 sequences assigned to
 194 each subset, however no new paralogs were found during

195 the search. The genome of *L. braziliensis* strain 2904 depos-
 196 ited on TriTrypDB lists 39 GP63 genes and, in total, the
 197 HMMs identified the presence of 38 related sequences. A
 198 single gene (LbrM.10.1720) was not recovered using these
 199 models and indeed its coding sequence did not provide an
 200 alignment with a score high enough to be considered as a
 201 GP63. The results of the search for each HMM are
 202 summarized in Additional file 2: Table S2 and confirm
 203 the gene count number for GP63 genes derived from
 204 the *L. braziliensis* genome sequencing, 38 genes, lower than
 205 earlier estimates based on the hybridization studies [23].

206 Next, we considered that the shot gun nature of the se-
 207 quencing strategy used for the assembly of the best genome
 208 available from a *Viannia* species, from the *L. braziliensis*
 209 2904, might have led to the grouping of similar GP63 genes
 210 together, causing in turn a reduction in the number of
 211 genes found. In order to obtain as many natural GP63
 212 sequences as possible, and therefore have a clearer idea of
 213 the true number of genes present in the *L. braziliensis* ge-
 214 nome, we opted to amplify these genes using primers
 215 directed to conserved regions of representative genes iden-
 216 tified in the genome analysis. The PCR strategy used to
 217 amplify the GP63 genes can be seen in Fig. 1, superimposed
 218 on a schematic representation of a typical GP63. The
 219 scheme highlights conserved elements found on all GP63
 220 sequences, such as the zinc-binding motif, multiple cysteine
 221 residues, the GPI anchor site and a nearly universally con-
 222 served motif of seven consecutive amino acids we named
 223 KDELMAP. Six oligonucleotides annealing to sequences
 224 encoding the N-terminal ends of the GP63 sequences were
 225 used as 5' primers, considering the variation previously
 226 observed within the N-terminus of the various GP63 genes
 227 and in order to maximize the number of genes amplified.
 228 As 3' primers, two sets of two oligonucleotides annealing
 229 to the more conserved KDELMAP or the GPI anchor site
 230 motifs were alternatively used. Individual PCR reactions
 231 were set up with different pairs of oligonucleotides, always
 232 with a single 5' and a single 3' primer. After amplification,
 233 cloning and sequencing a total of 40 different GP63 gene
 234 fragments were obtained, with thirty-four of those having
 235 sequences different from the ones described in the data-
 236 bases. The new GP63 DNA fragments obtained by PCR are
 237 listed in Additional file 3: Table S3, which also includes the
 238 set of oligonucleotide primers used to amplify each se-
 239 quence. In all 31 new gene fragments were found, since
 240 some were duplicates (also indicated in the table). The new
 241 sequences have all been submitted to GenBank and
 242 were compared with the already known GP63 genes
 243 from *L. braziliensis* and these analyses will be discussed
 244 further below. They are consistent with a higher copy
 245 number for the *L. braziliensis* GP63 genes than pre-
 246 dicted based on the genome sequencing alone and
 247 more in agreement with the original estimates based
 248 on Southern-blot, although no precise quantification

F1



f1.1 **Fig. 1** GP63 general gene features. The scheme highlights conserved elements found on most GP63 sequences, such as the signal peptide, propeptide,
 f1.2 zinc-binding motif [8], multiple cysteine residues, the GPI anchor site and a nearly universally conserved motif of seven consecutive amino acids we named
 f1.3 KDELMAP. The PCR strategy used to amplify the GP63 genes is also shown with products varying from 1 Kb (up to the KDELMAP region) and 1,6 Kb
 f1.4 (finishing on the GPI anchor site)
 f1.5

249 is possible either way. Despite the fact that both the gen-
 250 omic and PCR data used for our analyses are derived from
 251 the same *L. braziliensis* 2904 strain, the question of cultur-
 252 ing in different laboratories being responsible for the
 253 differences observed regarding gene copy number and
 254 identification between the two sets of results can be raised.
 255 Nevertheless, considering the limited time frame for the
 256 culturing procedures, these could only have any impact
 257 on gene or chromosomal duplication events and would
 258 not lead to the different gene sequences that were found
 259 through PCR and/or DNA sequencing.

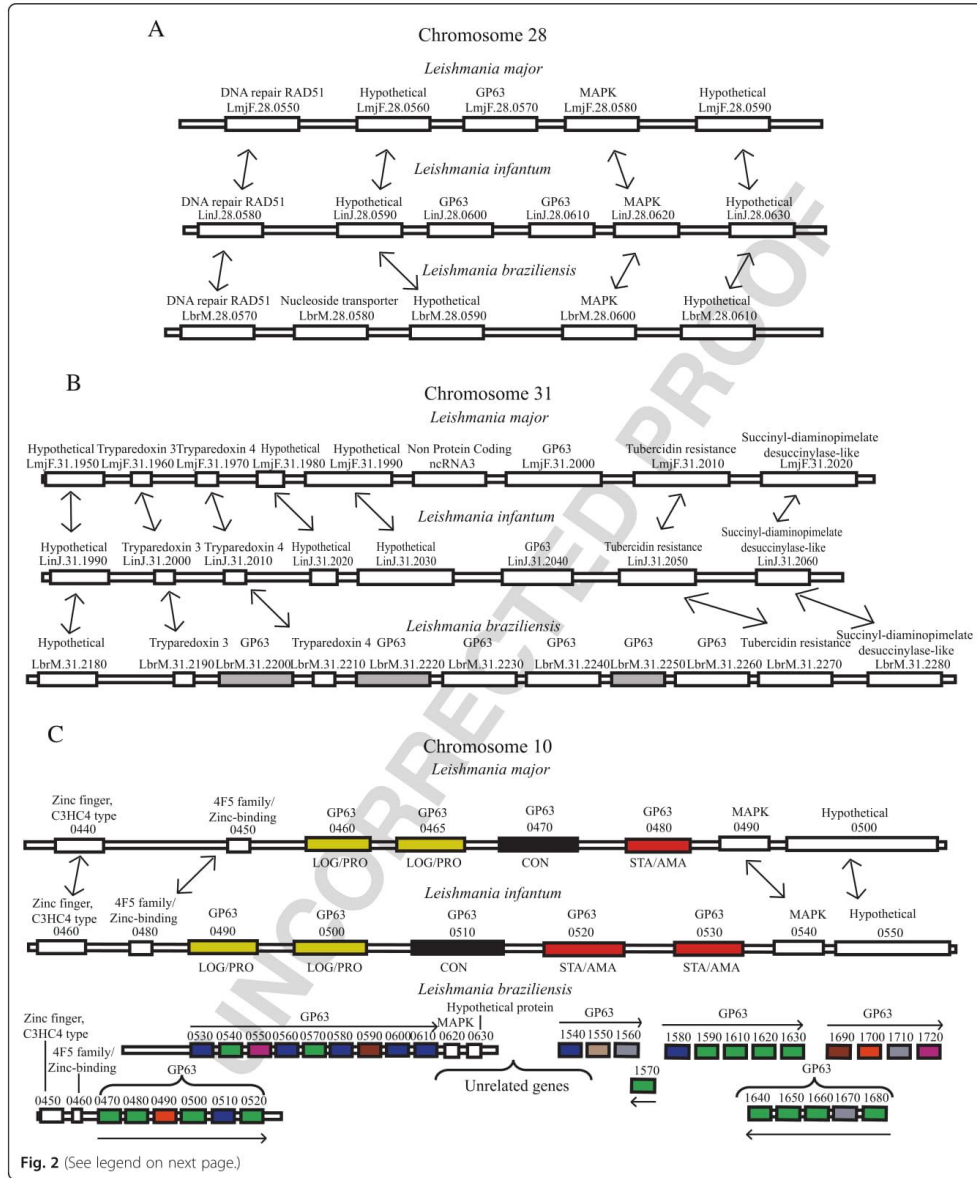
260 Genomic analysis of known *Leishmania* GP63 genes

261 To clarify the relationship between the multiple GP63
 262 genes in *Leishmania*, we opted to review their chromo-
 263 somal organization within the major lineages of pathogenic
 264 *Leishmania*. For *L. major*, the best studied of the available
 265 *Leishmania* genome sequences, three sets of GP63 genes
 266 were found distributed in chromosomes 28 (one gene), 31
 267 (one gene) and 10 (four genes) [21]. Based on the
 268 sequences available at the Trityp database, a similar
 269 organization is also observed for *L. infantum* (summarized
 F2 270 in Fig. 2) and *L. mexicana* (not shown), represented by two
 271 and one genes for chromosome 28 in *L. infantum* and *L.*
 272 *mexicana*, respectively, and one gene for chromosomes 31
 273 in both species, considering that the *L. mexicana* chromo-
 274 some 30 is equivalent to the *L. major* chromosome 31. Five
 275 GP63 genes are also found in chromosome 10 for both
 276 species and in agreement with previously reported data for
 277 *L. infantum* [22].

278 In *L. braziliensis*, based on the available genomic data for
 279 the 2904 strain, major differences in the organization of
 280 the GP63 genes are observed when they are compared with
 281 those found in species from the *Leishmania* subgenus.
 282 First, no GP63 gene is found on chromosome 28, as
 283 highlighted before for other *Viannia* species [32], despite
 284 the presence of orthologues to the same genes flanking the
 285 single GP63 sequence from *L. major* and *L. infantum*. In

286 contrast, six GP63 genes or gene fragments are found on
 287 chromosome 31, again generally flanked by orthologues to
 288 the same genes found flanking the GP63 gene found in the
 289 *L. major* and *L. infantum* chromosome 31. Even more
 290 noteworthy, however, are the 33 GP63 genes found clus-
 291 tered on chromosome 10. Again, these are localized to the
 292 same region seen harboring the other *Leishmania* chromo-
 293 some 10 genes, as confirmed by the presence of neighbor-
 294 ing sequences encoding orthologues to those found
 295 flanking the *L. major* and *L. infantum* GP63 genes from
 296 chromosome 10. However, the precise gene organization
 297 cannot be properly defined and many of the genes
 298 sequenced are assembled in relatively short contigs, as indi-
 299 cated in the scheme from Fig. 2c. Again, this might be due
 300 to the high similarity between the gene sequences and the
 301 nature of the sequencing strategy which might have pre-
 302 vented a proper assembly of repeated sequences.

303 The significantly greater number of *L. braziliensis* GP63
 304 genes from chromosome 10 is supported by our PCR data
 305 where primers sets directed to the chromosome 10 genes
 306 were able to amplify more genes than the ones originally
 307 used for their synthesis. For example, a primer pair
 308 designed to amplify the gene LbrM.10.0470 allowed the
 309 amplification of eight different gene fragments (G0510B2;
 310 G0560B1; G0560B2; G1610B3; G1610B4; G1610B5;
 311 G1610B6; G1620B1) and similarly, the primer pair directed
 312 to gene LbrM.10.0540 amplified fragments from six differ-
 313 ent genes (G0510C1; G0510C2; G0540C1; G0560C4;
 314 G1640C1; G1640C2). In contrast, two sets of PCR reac-
 315 tions directed to a single GP63 gene from chromosome 31,
 316 using the same 5' primer and two distinct 3' primers, only
 317 led to the amplification of the same gene, LbrM.31.2260
 318 (Additional file 3: Table S3). Indeed, we believe that most
 319 of the six GP63 genes annotated from the *L. braziliensis*
 320 chromosome 31 might not exist and in fact are either
 321 pseudogenes or derived from genome assembly errors.
 322 Only one of those genes (LbrM.31.2260) has GP63-related
 323 protein features, such as the propeptide domain (HEXXH),



f2.6
f2.7
f2.8
f2.9
f2.10
f2.11
f2.12
f2.13
f2.14
f2.17
f2.18

(See figure on previous page.)

Fig. 2 Genomic organization of *Leishmania* GP63 genes. Genomic organization of GP63 genes from *Leishmania* sp., showing the general distribution, location and synteny of these genes along chromosomes 10, 28 and 31. **a** Synteny analysis on chromosome 28 showing the localization of the GP63 gene in *L. major* and *L. infantum* and its absence in the equivalent position from the *L. braziliensis* chromosome 28. **b** Distribution of GP63 genes in *Leishmania* sp., chromosome 31. Pseudogenes from the *L. braziliensis* genome are highlighted in gray. **c** Expansion of GP63 genes in the *L. braziliensis* chromosome 10, when compared to *L. major* and *L. infantum*, within the same chromosome position, as indicated by the flanking genes on the left (some unrelated genes are also included). The horizontal arrows indicate the transcription sense of the genes and the yellow, black and red colors define the 3' UTR groups identified for the *L. major* and *L. infantum* genes. The diagram for *L. infantum* does not include the preliminary annotation derived from the recent resequencing of its genome [30]. For the *L. braziliensis* genes, identical colors group those with similar 3'UTRs, as classified in Table 1

and shares a high similarity (85%) with the *L. major* and *L. infantum* chromosome 31 genes. The LbrM.31.2200, LbrM.31.2220, LbrM.31.2230, LbrM.31.2240 and LbrM.31.2250 genes have stop codons in the middle of their sequences and/or in alignments showed identical N-terminal or C-terminal regions to LbrM.31.2260 (data not shown). It is possible that these genes may represent parts of LbrM.31.2260 not properly assembled and this in agreement with our PCR data finding only LbrM.31.2260. Overall these results are consistent with the expansion in the number of GP63 genes in *L. braziliensis*, and other species belonging to the *Viannia* subgenus, being mainly directed to the chromosome 10 genes.

GP63 evolutionary analyses

Sequences encoding GP63 related genes are also found in other trypanosomatids and more distantly related kinetoplastids and these include multiple genes from *T. brucei*, *T. cruzi*, and others. The number of genes in these parasites is quite variable. *T. cruzi* has over 150 GP63 genes annotated in the TriTrypDB database, but with many pseudogenes among them. *T. brucei* and *C. fasciculata* have a smaller amount with 10 and 18 genes respectively. This multiplicity of GP63 genes along the various kinetoplastids lineages reinforce the multiple roles this protein has, independent of the life cycle of the organism involved. Here, we next sought to assess how the *Leishmania* GP63 genes are related to those found in more distantly related kinetoplastids and whether some function can be inferred based on which genes are found in each organism. To do this we built a phylogenetic tree comparing the most divergent and representative sequences from the three major sets of *Leishmania* GP63 genes (from chromosomes 10, 28 and 31) with genes from different *Trypanosoma* species (*T. brucei*, *T. cruzi* and *T. theileri*) and more distantly related organisms. These included species that parasitize reptiles (*L. tarentolae*) and plants (*Phytomonas* sp.), have monoxenous life-cycles in insects (*Crithidia fasciculata* and *Leptomonas pyrrocoris*) and are free living (*Bodo saltans*). As shown in Fig. 3, the phylogenetic analysis could separate the GP63 genes mapped to the *Leishmania* chromosome 10 from those genes mapped to chromosomes 28 and 31. We also could observe a clear separation between the *Leishmania* subgenus based on

the genes located on chromosome 10. Noteworthy are the *T. cruzi* and *T. brucei* GP63 sequences more closely associated with the GP63 genes from *B. saltans* and *T. theileri*. Also, when we observe the clustering of the genes present in chromosomes 28 and 31 from *Leishmania*, they generally show more proximity to the genes from *L. pyrrocoris*, *C. fasciculata* and *Phytomonas* sp. Nevertheless, one *L. pyrrocoris* and two *C. fasciculata* genes are more closely related to those from the *L. braziliensis* chromosome 10.

Overall, the gene clusters shown in the tree highlight the higher similarity between the *Leishmania* sp. genes from chromosomes 28 and 31 with the GP63 genes found in organisms that live in insects only or parasitize plants. For instance, the 38 annotated GP63 genes from *Phytomonas* are more closely related to the *Leishmania* chromosome 28 GP63. It is then possible to hypothesize that these genes might be more involved in the insect stage of the parasite life cycle. Genes more closely related to the chromosome 10 GP63 genes can be found in the insect parasites *L. pyrrocoris* and *C. fasciculata*, but in general these genes seem to have suffered a substantial expansion within *Leishmania* species.

Evaluation of the sequence diversity of the *Leishmania* GP63 genes from chromosome 10

Considering the expansion of the chromosome 10 GP63 genes in the *Leishmania* lineage in general, and even more so in *L. braziliensis* and other *Viannia* species, we then opted to investigate the origins of their diversity further. To do this we compared the full extent of the chromosome 10 GP63 sequences from relevant *Leishmania* species using as outgroups selected genes from chromosomes 28 and 31. For this analysis, we also included sequences from *L. tarentolae* (based on the published genome sequence [33]), where a similar expansion in the chromosome 10 GP63 genes was noticed, with 49 genes found in this chromosome while only one gene was found in chromosome 31 and another in chromosome 33. *L. tarentolae* is currently classified within the *Sauroleishmania* subgenus, but it is likely to be more closely related to the *Leishmania* subgenus than to *Viannia* [2, 34]. The relevance in including the *L. tarentolae* sequences is due to the fact that it does not parasitize mammals, only lizards, meaning that any potential role in pathogenesis associated with the

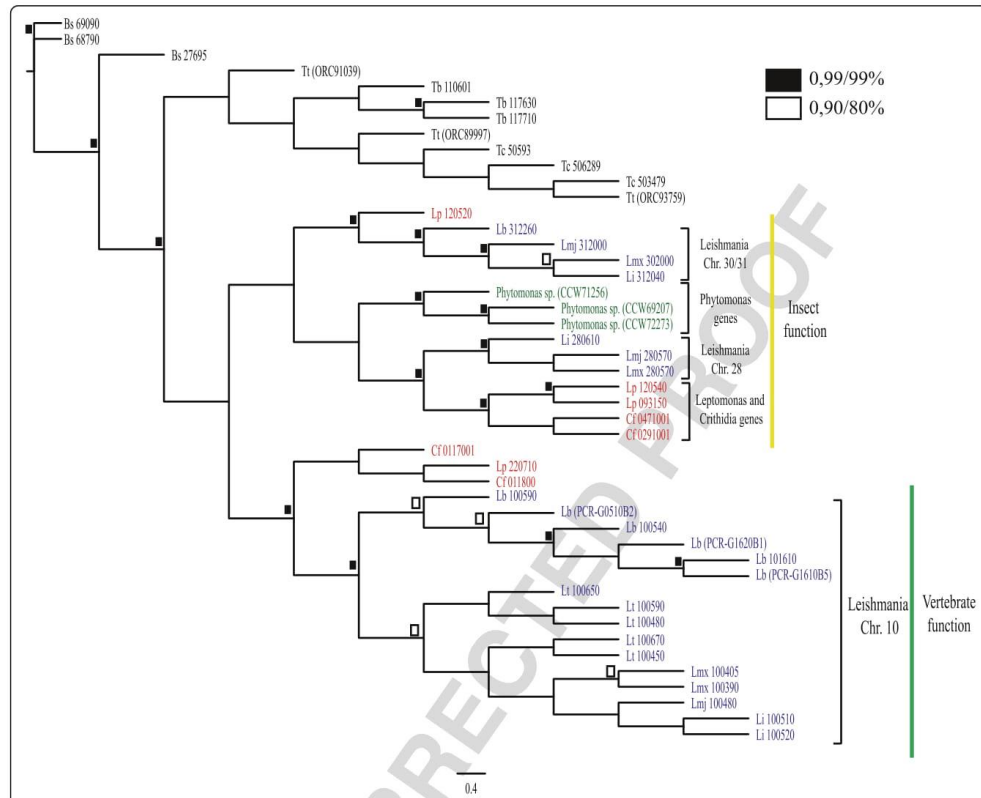
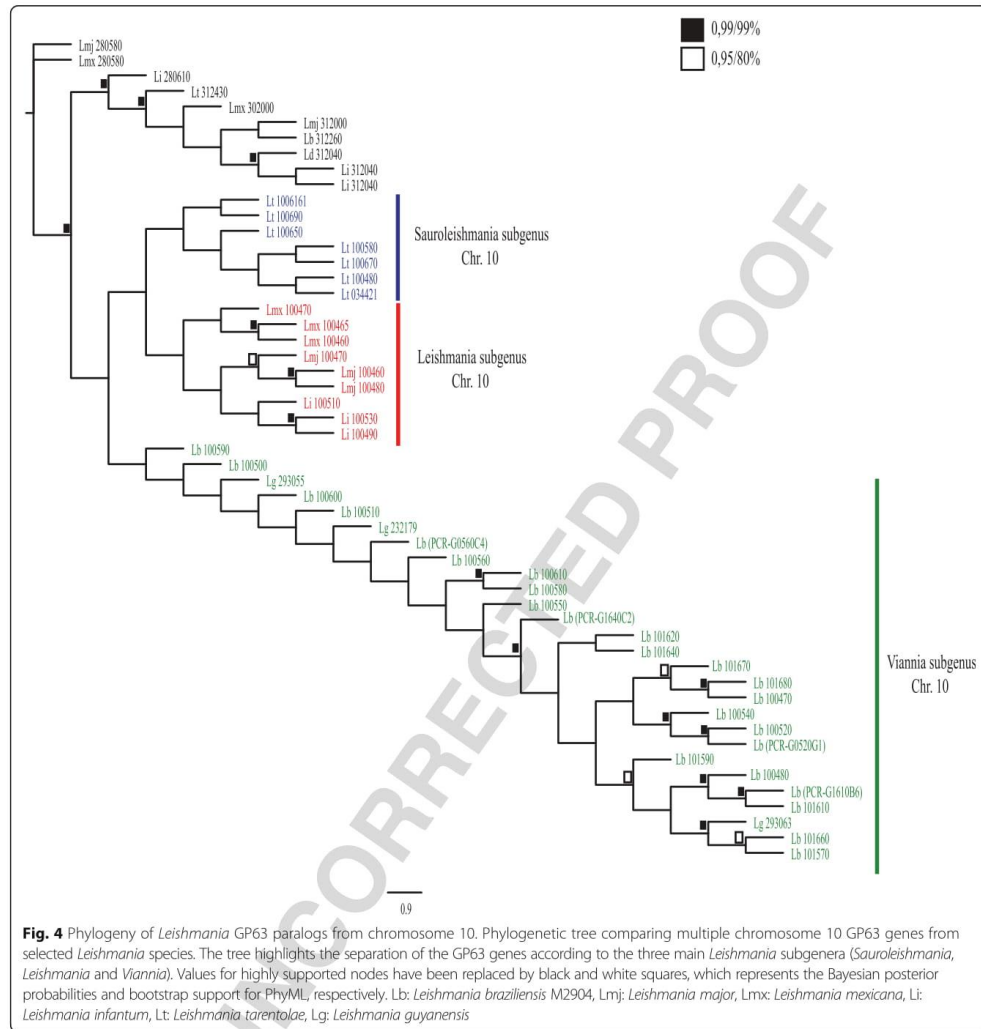


Fig. 3 Phylogenetic tree comparing selected *Leishmania* GP63 paralogs with sequences from more distantly related organisms. The tree demonstrates the proximity between the GP63 genes from the *Leishmania* chromosomes 28 and 31, with GP63 genes from monoxenes trypanosomatids genes, such as *Leptomonas pyrrocoris* and *Crithidia fasciculata*, as well as, genes from a plant tripanosomatid, like *Phytomonas sp.* Values for highly supported nodes have been replaced by black and white squares, which represents the Bayesian posterior probabilities and bootstrap support for PhyML, respectively. *Bs*: *Bodo saltans*, *Tt*: *Trypanosoma theileri*, *Tb*: *Trypanosoma brucei*, *Tc*: *Trypanosoma cruzi*, *Lp*: *Leptomonas pyrrocoris*, *Cf*: *Crithidia fasciculata*, *Lb*: *Leishmania braziliensis* M2904, *Lmj*: *Leishmania major*, *Lmx*: *Leishmania mexicana*, *Li*: *Leishmania infantum*, *Lt*: *Leishmania tarentolae*

410 chromosome 10 GP63 genes is not dependent on their
 411 mammalian hosts. We also included in this analysis the
 412 new *L. braziliensis* sequences generated by us through the
 F4 413 PCR approach. The phylogenetic tree shown in Fig. 4 sum-
 414 marizes the results from these analyses based on align-
 415 ments using the full-length sequences for all proteins (or
 416 the full-length PCR fragments). For clarity, only the most
 417 divergent representative sequences were used to build this
 418 tree, with those very similar or nearly identical to the ones
 419 shown purportedly removed from the final figure. In the
 420 original analysis all chromosome 10 sequences from the
 421 selected species were used but with similar results (not
 422 shown). Within each of the three *Leishmania* subgenera
 423 analysed, all GP63 sequences from chromosome 10 are
 424 more closely related to sequences from the same or related

425 species than to sequences found in species belonging to
 426 the other subgenera. Even within the *Leishmania* clades,
 427 the *L. infantum* genes (in red) seemed to be more closely
 428 related to each other than to their *L. major* counterparts,
 429 although for the two *Viannia* species (*L. braziliensis* and
 430 *L. guyanensis*) analysed genes (in green) more closely
 431 related between the two species were found. These results
 432 are in agreement with independent expansions on the
 433 number of the chromosome 10 GP63 sequences in each
 434 clade, with major expansions occurring for both
 435 *Sauvoleishmania* (in blue) and *Viannia* species. For the
 436 latter species, at least, the start of this expansion may have
 437 preceded the split between *L. braziliensis* and *L. guyanensis*
 438 but has subsequently continued and may be an on-
 439 going process.



f4.1 **Fig. 4** Phylogeny of *Leishmania* GP63 paralogs from chromosome 10. Phylogenetic tree comparing multiple chromosome 10 GP63 genes from
 f4.2 selected *Leishmania* species. The tree highlights the separation of the GP63 genes according to the three main *Leishmania* subgenera (*Sauroleishmania*,
 f4.3 *Leishmania* and *Viannia*). Values for highly supported nodes have been replaced by black and white squares, which represents the Bayesian posterior
 f4.4 probabilities and bootstrap support for PhyML, respectively. Lb: *Leishmania braziliensis* M2904, Lmj: *Leishmania major*, Lmx: *Leishmania mexicana*, Li:
 f4.5 *Leishmania infantum*, Lt: *Leishmania tarentolae*, Lg: *Leishmania guyanensis*
 f4.6

440 **Identification of functional differences between the**
 441 **various *Leishmania* GP63 genes from chromosome 10**

442 So far not much is known regarding possible functional
 443 differences between the various GP63 genes found within
 444 any specific trypanosomatid. In *Leishmania*, with the goal
 445 of defining functional differences between multiple GP63,
 446 and even prior to the completion of the first *Leishmania*
 447 genomes, early studies investigated the expression pattern
 448 of selected genes attempting to identify differences in ex-
 449 pression during the parasite life cell cycle [8, 15, 35]. In *L.*

450 *infantum* and *L. major*, three growth stage-specific pat-
 451 terns of expression were observed for the then known
 452 GP63 genes, with one gene constitutively expressed, a sec-
 453 ond gene (or genes) expressed during the log phase of
 454 promastigote growth and the third gene expressed only in
 455 stationary phase cells and/or amastigotes [15, 25]. Here,
 456 by comparing their coding sequences and 3' UTRs with
 457 those from the available genomes, we were able to map
 458 those genes within the annotated genome sequences (indi-
 459 cated in Fig. 2): LinJ.10.0510 and LmjF.10.0470 are the

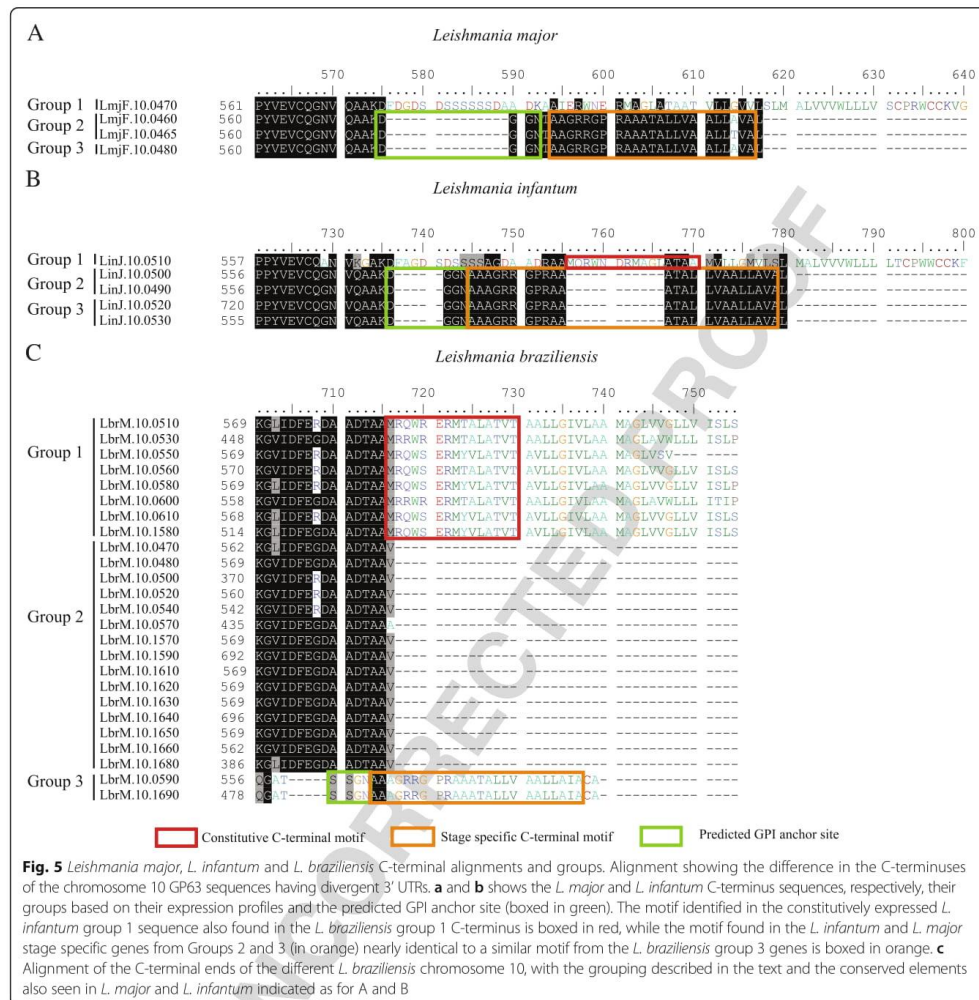
constitutively expressed genes (here named Group 1 – colored in black in the figure); LinJ.<https://doi.org/10.0490/LinJ.10.0500> and LmjF.<https://doi.org/10.0460/LmjF.10.0465> are equivalent to the previously described log phase promastigote genes (Group 2 – highlighted in yellow); LinJ.<https://doi.org/10.0520/LinJ.10.0530> from *L. infantum* and the *L. major* LmjF.10.0480 gene correspond to the stationary phase/amastigote specific GP63 (Group 3 – colored in red). The Group 1 genes are characterized by unique 3' UTRs and 3' intergenic regions absent from the remaining chromosome 10 genes, while the Group 2 and 3 genes share very similar sequences within the first ~ 400 nucleotides of their 3'UTRs, although these subsequently diverge into two distinct patterns that correlate with the two groups (these genes and their groups are highlighted by different colors in the scheme from Fig. 2). We also looked at differences within the coding sequences that could be typical of GP63 genes belonging to any particular group. As previously reported for *L. major* [25], a clear distinction is observed between the C-terminus of the Group 1 proteins and those from Group 2 and 3. Both *L. major* and *L. infantum* Group 1 proteins are characterized by a longer C-terminus enriched in hydrophobic and positively charged residues and lack the typical asparagine required for the GPI anchor. In contrast, the shorter C-terminus from the Group 2 and 3 proteins include the GPI anchor site and end in a stretch of mostly hydrophobic amino acids (Fig. 5).

F5 Next, we attempted to group the *L. braziliensis* chromosome 10 genes based first on similarities and differences within the putative 3'UTRs. Thirty genes were analyzed based on the sequences available from the reference genome sequence and these were classified into six groups according to their 3'UTR, with the first two groups represented by 8 and 15 genes, respectively, while the remaining

T1 groups included only one or two genes (Table 1 – also colored differently in Fig. 2). When these 3'UTRs were compared with the three *L. infantum* and *L. major* groups no clear similarities were found with any of the *L. braziliensis* groups, likely due to the large sequence variation observed between species from the two distinct subgenera. We also looked at the amino acid sequences, looking for features in common for *L. braziliensis* genes sharing similar UTRs. No association with features such as signal peptide, transmembrane domains and isoelectric points was found, however distinct C-terminus were observed which also separated the *L. braziliensis* into six groups, with a clear association observed between each 3'UTR group and nearly all of the proteins' C-terminal ends (Table 1). Since Groups 4, 5 and 6 consists of truncated proteins they will not be considered further here, but for the remaining three groups their C-terminal ends were also compared with those seen to be associated with the *L. infantum* and *L. major* groups. A clear association between the two proteins from the *L. braziliensis* Group 3 and the *L. major/L. infantum* Groups 2 and 3 can be seen, since they share a nearly identical C-terminus that includes the GPI anchor attachment motif (DGGN [36]). The Group 1 genes from *L. braziliensis* also share conserved elements with the *L. major/L. infantum* Group 1, such as the lack of a typical GPI anchor site and the presence of a hydrophilic set of amino acids followed by a hydrophobic region resembling a transmembrane segment. The motif "MRQWRERMTALATVT" found in the *L. braziliensis* sequences is also very similar to the "MQRWNRMAGLATAA" motif found in *L. infantum* LinJ.10.0510 gene. Only the *L. braziliensis* Group 2 genes then, characterized by a shorter C-terminus missing entirely the GPI anchor site or related hydrophobic sequences, do not seem to have counterparts in *L. infantum* nor in *L. major*. Nevertheless, it seems likely that, as observed in *L. infantum* and *L. major*, the different groups of *L. braziliensis* chromosome 10 GP63 genes are also differentially regulated during the parasite growth in culture and this is in agreement with the different 3'UTRs seen associated with each group.

Gene recombination in GP63 sequences from chromosome 10

Through analyzes of the alignments generated in this study, we identified that specific regions of certain GP63 gene sequences were very similar to equivalent regions from other GP63 genes which otherwise were more divergent. For example, certain small motifs generally seen only on the *L. braziliensis* Group 1 genes were also found in one or more of the group 2 genes and vice-versa, an indication of gene recombination. Indeed, the locus for these genes is reported as having high plasticity [23, 37] and the data from the literature shows that this gene family can be influenced by mosaic or fragmental gene conversion [26, 38]. Here, in order to understand why the expansion of the GP63 genes occurs mainly on chromosome 10, we performed an *in-silico* search for recombination events targeting these genes so as to better evaluate whether their variability was related to intragenic recombination. The software chosen to find the recombination events (RDP4) uses several tools to determine events such as the likely position of recombination breakpoints and the identity of sequences most closely related to the gene being evaluated. In this study we only considered recombination events that were detected by at least two of the tools tested. Therefore, we decided to perform a gene recombination analysis with all the GP63 genes of *L. braziliensis* present in databases and the ones generated by us through PCR. We first targeted the chromosome 31 GP63 genes, but no recombination events were detected by the software. In contrast, when the 38 PCR sequences from chromosome 10 were analyzed 30 (or 79%) were reported as recombinant genes. Regarding the database genes, 32 recombination events were found,



Q75.1
 f5.2
 f5.3
 f5.4
 f5.5
 f5.6
 f5.7
 f5.8

567 sometimes with more than one event for the same gene
 F6 568 (Fig. 6 and Additional file 4: Table S4). Gene duplication
 569 and recombination events then are possibly the major
 570 source of the novel GP63 sequences seen in the chromo-
 571 some 10 from *L. braziliensis* and closely related sequences.

572 Protein structural modeling, mapping of variable regions 573 and B-cell epitope prediction

574 Based on the crystallized structure of a membrane GP63
 575 from *L. major* promastigotes, GP63 was identified as a
 576 compact protein consisting predominantly of β sheet
 577 secondary structure elements divided into three distinct

578 domains (N-terminus, Central domain and C-terminus) 578
 579 and with features typical of the catalytic modules of zinc 579
 580 proteases [39]. After observing the recombination events 580
 581 and sequence variations between the multitudes of *L. brazi-* 581
 582 *liensis* GP63 genes, we decided to investigate where these 582
 583 variations are found along the 3D structure of the protein. 583
 584 Through three-dimensional protein structure predictions, 584
 585 we were able to model the structure of eight divergent 585
 586 GP63 sequences with high modeling scores, as can be seen 586
 587 in Fig. 7. We then mapped on the models the most variable 587
 588 motifs identified by the previous multiple alignments 588
 589 (highlighted in blue in the structures shown in the figure). 589

F7

Table 1 *Leishmania braziliensis* GP63 3' UTR and C-terminal end gene groups. Table showing the Gp63 gene groups from chromosome 10 defined according to the 3' UTR and C-terminus sequence similarity

Group	<i>L. braziliensis</i> 3'UTRs	<i>L. braziliensis</i> chromosome 10 GP63 paralogs
1	LbrM.10.0510, LbrM.10.0530, LbrM.10.0560, LbrM.10.0580, LbrM.10.0600, LbrM.10.0610, LbrM.10.1540, LbrM.10.1580	LbrM.10.0510, LbrM.10.0530, LbrM.10.0550, LbrM.10.0560, LbrM.10.0600, LbrM.10.0610, LbrM.10.1540, LbrM.10.1580
2	LbrM.10.0470, LbrM.10.0480, LbrM.10.0500, LbrM.10.0520, LbrM.10.0540, LbrM.10.0570, LbrM.10.1570, LbrM.10.1590, LbrM.10.1610, LbrM.10.1620, LbrM.10.1630, LbrM.10.1640, LbrM.10.1650, LbrM.10.0600, LbrM.10.1680	LbrM.10.0470, LbrM.10.0500, LbrM.10.0520, LbrM.10.0540, LbrM.10.0570, LbrM.10.1570, LbrM.10.1590, LbrM.10.1610, LbrM.10.1620, LbrM.10.1630, LbrM.10.1640, LbrM.10.1650, LbrM.10.1660, LbrM.10.1680
3	LbrM.10.0590, LbrM.10.1690	LbrM.10.0590, LbrM.10.1690
4	LbrM.10.0550, LbrM.10.1720	LbrM.10.1720
5	LbrM.10.0490, LbrM.10.1700	LbrM.10.0490, LbrM.10.1700
6	LbrM.10.1550	LbrM.10.1550

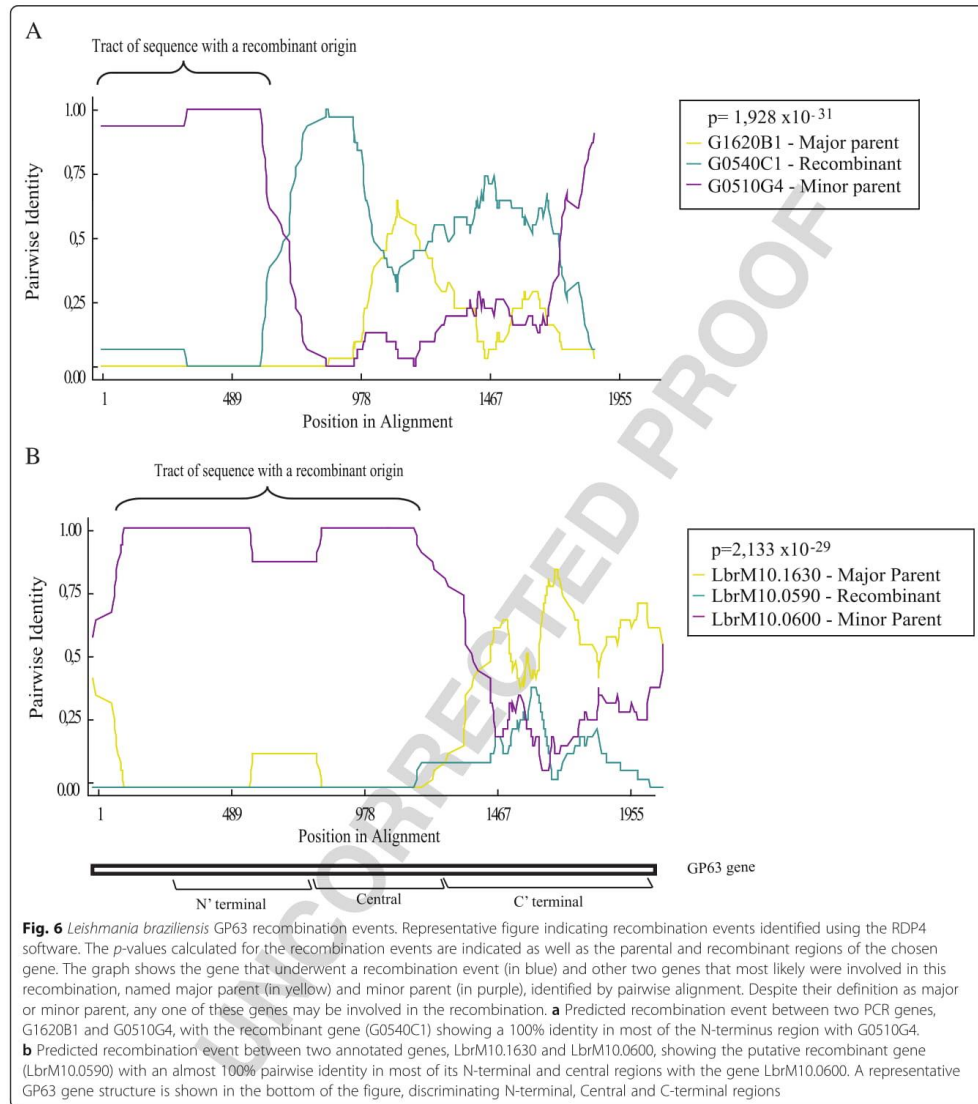
As can be observed, most of the variable regions were positioned externally on the structures.

We next sought to evaluate how the various GP63 sequence would be recognized by the B cells from the mammalian immune system and also to predict their ability to induce the production of specific antibodies. Linear and conformational B-cell epitope predictions were carried out using the various chromosome 10 GP63 sequences from *L. braziliensis*. The linear epitope predictions returned 56 epitopes from the sequences of the modeled proteins (Table 2). Regarding their localization within the various structures, from the total of 56 epitopes, 41 were mapped to the proteins' external regions, whereas 6 were localized internally and nine could not be evaluated due to the comparative nature of the tridimensional structure modeling. The modeling is based on a mature GP63 crystallized structure, which lost part of its N-terminal region, during the protein posttranslational modification, which prevented the assessment of epitopes localized to this region. Out of the 47 linear epitopes localized externally, 37 coincide with motifs that display sequence variation, while 10 are found in regions conserved between the different GP63 sequences. Considering only the epitopes that were predicted to localize internally, four coincide with variable sequences while two were associated with conserved regions. As for the prediction of conformational B-cell epitopes, the analysis returned 40 epitopes, with 32 mapped to the proteins' external region. Twenty-three of those were in motifs with sequence variation, while nine were in conserved regions. Regarding the eight epitopes localized internally, five coincided with variable sequence motifs and three were in conserved motifs. We also identified 14 motifs that were present in both linear and conformational epitope predictions (not shown). As above, most of the epitopes coincided with variable motifs localized externally, as shown in Table 2. Noteworthy also is the fact that a peptide from *L. infantum* GP63 that has been previously shown to react strongly with sera from dogs infected with visceral leishmaniasis [40] is nearly identical to one of the *L. braziliensis* B-cell epitopes predicted by our analysis

within the Lbr10.0590 polypeptide (not shown). Overall, these results are consistent with variable regions localizing externally and being more capable of inducing an immune response.

Discussion

The in-silico analysis carried out here highlights the strong selective pressure for the expansion in copy number of the chromosome 10 GP63 genes within *Leishmania* species, and in particular in the *Viannia* and *Sauroleishmania* subgenera. The increased number of the chromosome 10 GP63 genes in different *Leishmania* species evolved independently generating a wide range of paralogs, which display sequence variations and may be generated by recombination. This expansion seems to be an ongoing process that might be related to pathogenesis or defense mechanisms directed to the vertebrate host and the parallel expansion in both *Viannia* and *Sauroleishmania* species is something that must be taken into account. Such expansion of multiple genes arranged *in tandem*, originating from duplication and recombination events, demonstrates the adaptability of *Leishmania* species to the environment, associated with the evolutionary pressure suffered by the GP63 genes [28]. As a result, the presence of these multi-copy arrays may lead to speciation [41] or indicate the possible need for stage-specific genes [28]. As previously highlighted [42], an expansion in GP63 sequences also occurred independently in other trypanosomatids, such as *T. cruzi* and *T. brucei*, and this led to novel GP63 domains which might be associated with species-specific or group-specific functions. It is likely then that this expansion in *Leishmania* GP63 genes might be related to novel aspects of the pathogenesis of these parasites to the vertebrate hosts, but this still needs to be better defined. When multiple strains from a single species is considered the overall GP63 diversity might be even greater, as recently evaluated [43], and this might be associated with possibly different virulence phenotypes and clinical outcomes for the disease. The recent release of a new *L. infantum* genome based on data using two distinct methodologies of

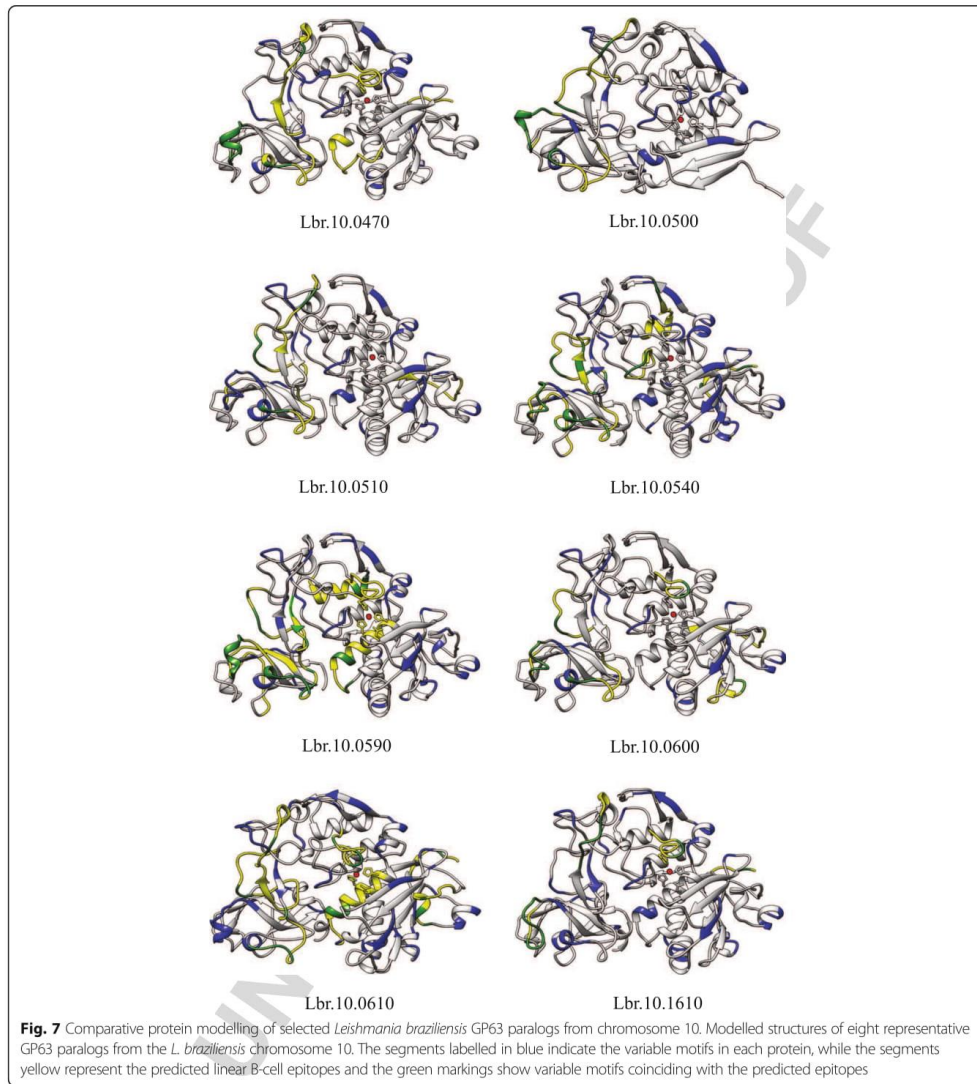


f6.1
f6.2
f6.3
f6.4
f6.5
f6.6
f6.7
f6.8
f6.9
f6.10

669 next generation sequencing, and showing a higher copy
670 number of GP63 genes for this species [30], also high-
671 lights the need for better quality genomes in order to
672 properly define the true diversity of these genes for
673 multiple *Leishmania* and trypanosomatid species.

674 The expansion observed in the chromosome 10 genes
675 are concentrated in the Group 1 and Group 2 genes and,

676 if we extrapolate the expression for Group 1 based on the
677 data with the *L. major* and *L. infantum* genes [15, 25], they
678 are likely to be expressed constitutively with likely func-
679 tions during the mammalian infection. In contrast, for the
680 Group 3 genes, with only two paralogs, their expression
681 might be restricted to the promastigote stage of the para-
682 site life cycle, therefore with minor or no relevant function



f7.1
f7.2
f7.3
f7.4

683 in the mammalian host. For all three groups their expression
684 will have to be confirmed but the large sequence
685 variations observed for both Groups 1 and 2 genes, con-
686 centrated on potentially immunogenic regions localized
687 on the surface of the GP63 molecules, also imply related
688 expression patterns during the mammalian stage of the
689 *Leishmania* life cycle. As previously shown in *L. major*
690 and *L. mexicana* [44], these expression patterns should be

691 linked to the mRNA 3'UTRs and sorting out the molecular
692 mechanisms associated will be a major endeavor. A major
693 question that emerges regarding the expression of the
694 genes from Groups 1 and 2 is related to the multiple para-
695 logs. Are multiple genes belonging to the same group
696 expressed simultaneously or some are expressed more effi-
697 ciently than others or alternatively? This will also need to
698 be investigated further.

Table 2 *Leishmania braziliensis* GP63 B-cell epitope prediction. Table showing the distribution of the predicted epitopes along the studied GP63 paralogs

Total Number of Epitopes	Conserved Region	Sequence Variation Region	Type of Epitopes	Protein Region
41	8	33	Linear	External
6	2	4	Linear	Internal
32	9	23	Conformational	External
8	3	5	Conformational	Internal

Another relevant question remaining deals with the functional roles for the distinct groups and how these might be associated with sequence differences between the paralogs. A possible link with the proteins' subcellular localization is presumed based on the differences in the C-terminus of the subsets identified and the presence or absence of a typical GPI anchor. These differences regarding the presence or absence of GPI anchor sites have been suggested before based on comparisons between the *L. major* and *L. infantum* GP63 sequences [8]. Here, the C-terminal Group 2 of *L. braziliensis* GP63 sequences lacks the GPI anchor signal entirely, which is consistent with proteins that are directly secreted into the extracellular medium, as previously reported for *L. mexicana* GP63 [45]. This release into the extracellular environment might contribute at the early stages of infection, due to the ability of GP63 to digest the extracellular matrix proteins, facilitating parasite mobility and invasion [46]. Alternatively, these proteins might be selectively transferred to exosomes and later to the macrophages in order to influence its metabolism and promote *Leishmania* growth [11]. For the *L. braziliensis* Group 1 proteins, they all share a C-terminus having a likely transmembrane domain with no clear GPI anchor site. Lack of a typical GPI anchor site however, with a more likely transmembrane domain identified, was also seen in the *L. major* Group 1 gene, which was nevertheless seen to have a GPI anchor [25]. The distinct C-terminal ends nevertheless clearly suggest critical differences in subcellular localization for the distinct GP63 groups, but these need further experimental confirmation.

In early studies performed with *L. guyanensis*, it was suggested that new GP63 genes may be generated by events of mosaicism through recombination between 5' and 3' UTRs and protein coding regions [24], and mosaicism in GP63 sequences was subsequently also found in *L. braziliensis* genes [26]. The data obtained by us corroborate with other studies investigating GP63 recombination that found it to target mainly the N-terminal and C-terminal regions of the gene [37, 38]. The impact of the GP63 sequence variability in its structure has been investigated in a wider scale, comparing *Trypanosoma* and *Leishmania* sequences, and found to be associated with variability in its zinc binding site and presumably

activity [42]. In a recent study targeting *L. braziliensis* GP63 sequences, structural differences have also been found to be associated with sequence variability, implying functional differences, such as during substrate binding, which may affect the interaction with the host [47]. Alternatively, the variability in protein structure could mainly affect recognition by the host immune system and promote infection mainly because the host would need to produce different antibodies to neutralize a single group of proteins. Our results, showing variability concentrated on antigenic regions on the protein's surface, is in agreement with previously reported data based on *L. major* and *L. infantum* sequence analysis where regions of GP63 sequence variation were mapped to the surface of the protein and were associated with immunodominant epitopes [37]. However, more studies are needed to better understand the recognition of different GP63 paralogs by the host immune system.

Overall the data presented here highlights novel and relevant aspects related to the expansion of GP63 genes in *L. braziliensis* and related *Viannia* species and raises specific issues regarding the role of GP63 in the parasite pathogenesis during the infection in mammals. It is possible that species belonging to the subgenus *Viannia* may have added a new level of complexity to GP63 function and this may somehow be related to the capacity of some species to cause the more aggressive mucocutaneous form of the disease. The new questions raised here then, when solved, shall provide novel and relevant knowledge regarding the very unique mechanisms of pathogenesis associated with these parasites.

Conclusions

Our results suggest a greater role for the sequence variation found among the chromosome 10 GP63 genes for the pathogenesis of *L. braziliensis* and closely related species within the mammalian host. The variation in sequence and the expansion in number of these GP63 genes have occurred independently in different *Leishmania* lineages, is associated with intragenic recombination events and has a likely role against the host immune response. They also indicate different functions associated to genes mapped to different chromosomes and, for the chromosome 10 genes at least, variable subcellular localizations likely associated with multiple functions and target substrates for this versatile protease.

Methods

Parasites and culture conditions

In this study, we used *Leishmania (Viannia) braziliensis* (MHOM/BR/75/M2904) in its promastigote form. This is a reference strain from the Evandro Chagas Institute, Belém, Brazil. The cells were cultured at 26 °C in Schneider (Sigma) pH 7.2 supplemented with 20% fetal

- 794 bovine serum (FBS), antibiotics (Streptomycin / Penicillin
795 0.1%) and 0.1% Hemin.
- 796 **PCR, cloning and sequencing**
- 797 Approximately, 10⁸ *L. braziliensis* promastigotes were
798 used for total genomic DNA extraction using DNAzol
799 (Invitrogen) and standard procedures. PCR reactions for
800 the amplification of the GP63 sequences were performed
801 using Phusion[®] High-Fidelity DNA Polymerase (New Eng-
802 land Biolabs), following the manufacturer's protocol and
803 with the oligonucleotides used as primers listed in the
804 Additional file 5: Table S5. After amplification, cloning and
805 sequencing of the PCR products, a nomenclature was cre-
806 ated for the newly generated sequences in order to identify
807 from which set of primers they were derived, whether from
808 those encoding the KDELMAP or GPI regions, and defin-
809 ing which annotated GP63 gene it most closely resembles.
810 The newly generated sequences derived from the PCR am-
811 plifications were deposited on the GenBank and all acces-
812 sion numbers are listed in Additional file 3: Table S3.
- 813 **Search for new GP63 paralogs through hidden Markov**
814 **models**
- 815 First, the predicted proteomes of the following organisms
816 were downloaded from TritypDB in August 25, 2014: *L.*
817 *braziliensis* strain 2903 [taxid: 1295825], *L. braziliensis*
818 strain 2904 [taxid: 420245], *L. infantum* [taxid: 435258], *L.*
819 *major* [taxid: 347515], *L. donovani* [taxid: 981087], *L.*
820 *mexicana* [taxid: 929439] and *L. tarentolae* [taxid: 5689].
821 GP63 genes were then identified within the downloaded
822 proteomes, considering only genes annotated as GP63,
823 encoding proteins longer than 30 amino acids and with no
824 more than one stop codon per sequence. All of the protein
825 sequences derived from genes that met these inclusion cri-
826 teria were submitted to the analysis of the OrthoMCL
827 program [48], and grouped according to homology using
828 the Markov Cluster algorithm [49]. Protein sequences
829 from each group were aligned using the MAFFT software
830 (default settings) [50] and the multiple alignments used as
831 input for the hmmbuild, version 3.0, a tool from the
832 HMMER package [51] to build Hidden Markov Models
833 (HMMs). The models were then used with the hmmsearch
834 tool to search for new paralogs within the *L. braziliensis*
835 strain 2904 proteome. A cutoff of 0.001 for hit significance
836 (e-value <= 0.001) was applied.
- 837 **GP63 phylogenetic analysis and detection of**
838 **recombination events**
- 839 A phylogenetic tree was built with GP63 protein sequences
840 from genes encoded within chromosomes 10, 28 and 31
841 from diverse *Leishmania* species and more distantly related
842 organisms. These include the *Phytomonas* sp. isolate
843 Hart11 [taxid: 134014], *Crithidia fasciculata* [taxid: 5656],
844 *Leptomonas pyrrhocoris* [taxid: 157538], *Trypanosoma*
845 *cruzi* [taxid: 353153], *Trypanosoma brucei* [taxid: 185431],
846 *Trypanosoma theileri* [taxid: 67003] and *Bodo saltans*
847 [taxid: 75058]. Another tree was made with selected GP63
848 sequences used in the previous analyses plus the ones
849 obtained by PCR from *L. braziliensis* as well as GP63
850 sequences from *Leishmania guyanensis* [taxid: 5670]. For
851 all trees, the selected sequences were aligned by MAFFT
852 (default settings) and the alignments automatically edited
853 by Trimal [52] to keep just phylogenetically informative
854 sites. ProtTest [53] was then used to predict the best evolu-
855 tionary model which was subsequently used as a setting to
856 build phylogenetic trees with PhyML, applying the Max-
857 imum Likelihood (ML) method [54], and MrBayes, apply-
858 ing the Bayesian method [55, 56]. The branch support for
859 the ML tree was given by non-parametric bootstrap ana-
860 lysis using 1000 replicates. The Bayesian inferred trees were
861 determined by 5,000,000 chains to check for convergence
862 and a 100% burn-in was discarded. The aligned nucleotide
863 sequences from *L. braziliensis*, obtained from the Tri-
864 trypDB database and through PCR, were analyzed for
865 recombination using the RDP4 program [57].
- 866 **Modelling of GP63 homologs and searches for non-**
867 **conserved regions**
- 868 Eight of the most variable paralogs from different *L.*
869 *braziliensis* C-terminal groups were chosen for the
870 three-dimensional modelling. The modelling was per-
871 formed for the amino acid sequences previously obtained
872 from TriTrypDB and applying the SWISS MODEL plat-
873 form [58]. When the models were completed, their qual-
874 ities were assessed through Procheck [59]. Specific regions
875 of the protein models were then evaluated using the initial
876 alignment information, highlighting the non-conserved re-
877 gions which were characterized by amino acid exchanges.
- 878 **B-cell epitope prediction**
- 879 Linear B-cell epitope predictions were performed for the
880 protein sequences used in the 3D modeling step. The pre-
881 dictions were carried out using the following programs:
882 AAP12 [60], BCPred12 [61] and BepiPred [62]. Only epi-
883 topes predicted by at least two programs, with lengths
884 equal to or greater than 10 amino acids and with scores
885 greater than 0.8 were considered as positive predictions
886 on AAP12 and BCPred12. Epitopes with scores over 0.5
887 obtained by BepiPred were also included in the analysis.
888 In addition to the linear prediction, a conformational
889 prediction of epitopes was also performed to evaluate if
890 the protein structures were also able to generate inter-
891 action with the immune system. The conformational epi-
892 topes were predicted by the CBTOPE web server [63],
893 where only epitopes that with more than 10 amino acids
894 and a score above 4 were considered for this study. After
895 the prediction, an assessment was performed to map the
896 localization of all the epitopes on the modeled proteins.

Q5 897 **Additional files**
 898
 900 **Additional file 1: Table S1.** Subsets of GP63 sequences used to build the
 901 Hidden Markov Models (HMMs) (DOCX 13 kb)
 902 **Additional file 2: Table S2.** GP63 genes identified by HMM from the *L.*
 903 *braziliensis* M2904 proteome. Table showing the number of GP63 genes
 904 identified by each HMM after the search for new paralogs within the *L.*
 905 *braziliensis* M2904 genome sequences. (DOCX 11 kb)
 906 **Additional file 3: Table S3.** Set of oligonucleotides, GP63 sequences
 907 obtained by PCR and their respective GenBank accession number.
 908 (DOCX 15 kb)
 909 **Additional file 4: Table S4.** Table showing the recombination events
 910 found for the *L. braziliensis* GP63 genes, followed by the programs that
 911 detected it and their *p*-value. Detailed description of the recombination
 912 events throughout the *L. braziliensis* GP63 genes. (XLSX 16 kb)
 913 **Additional file 5: Table S5.** Oligonucleotides used for the PCR reactions.
 914 A list of the oligonucleotides used for the PCR reactions in this
 915 study. (DOCX 13 kb)

917 Abbreviations

918 CL: Cutaneous Leishmaniasis; FBS: Fetal Bovine Serum; GP63: Glycoprotein 63;
 919 GPI: GlycosylPhosphatidylinositol; HMMs: Hidden Markov Models;
 920 MAC: Membrane Attack Complex; MCL: Mucocutaneous Leishmaniasis;
 921 MSP: Major Surface Protease; NPT: Nucleo de Plataformas Tecnológicas;
 922 UTR: Untranslated Region; VL: Visceral Leishmaniasis

923 Acknowledgements

924 We thank members of Dr. O. P. de Melo Neto's laboratory for helpful discussions
 925 and support. The authors thank the Nucleo de Plataformas Tecnológicas (NPT)
 926 at the Institute Aggeu Magalhães (FIOCRUZ-PE) for the use of its automatic
 927 sequencing facility.

928 Funding

929 The *Leishmania* work in Dr. de Melo Neto's lab was more recently funded with
 930 grants provided by the Brazilian funding agencies FACEPE (APQ-0239-2.02/12
 931 and APQ-1662-2.02/15), CNPq (480899/2013-4, 313934/2013-4 and 401282/
 932 2014-7) and CAPES (23038.007656/2011-92). Studentships for the graduate
 933 students (ALCN, and ANALMB) were provided by CAPES and FACEPE.

934 Availability of data and materials

935 The GP63 sequence datasets analysed in this study are available at TriTrypDB
 936 (<http://tritrypdb.org/tritrypdb/>) and GenBank (<https://www.ncbi.nlm.nih.gov/genbank/>) under the accession number provided in the Additional file 3: Table S3.

Q4 Q3

938 Authors' contributions

939 ALCN – experimental and bioinformatic analysis, writing original draft.
 940 ANALMB - experimental and bioinformatic analysis. AMR – bioinformatics
 941 experimental design, FBM – results discussions. OPMN – conceptualization,
 942 supervision, funding acquisition, writing, review and editing. All authors read
 943 and approved the final manuscript.

944 Ethics approval and consent to participate

945 Not applicable.

946 Consent for publication

947 Not applicable.

948 Competing interests

949 The authors declare that they have no competing interests.

950 Publisher's Note

951 Springer Nature remains neutral with regard to jurisdictional claims in published
 952 maps and institutional affiliations.

953 Author details

954 ¹Universidade Federal de Pernambuco, Recife, Pernambuco, Brazil. ²Instituto
 955 Aggeu Magalhães, Fundação Oswaldo Cruz (FioCruz-Pernambuco), Recife,

Pernambuco, Brazil. ³Centro Universitário Tabosa de Almeida – ASCES/UNITA, 956
 Caruaru, Pernambuco, Brazil. 957

Received: 6 June 2018 Accepted: 21 January 2019 958

959

References

1. Pace D. Leishmaniasis. *J Inf Secur.* 2014;69(Suppl 1):10–8. 961
2. Akhouni M, Kuhls K, Cannet A, Votyčka J, Marty P, Delaunay P, et al. A 962
historical overview of the classification, evolution, and dispersion of 963
Leishmania parasites and sandflies. *PLoS Negl Trop Dis.* 2016;10:1–40. 964
3. Kaye P, Scott P. Leishmaniasis: complexity at the host–pathogen interface. 965
Nat Rev Microbiol. 2011;9:604–15. 966
4. Arango Duque G, Descoteaux A. *Leishmania* survival in the macrophage: 967
where the ends justify the means. *Curr Opin Microbiol.* 2015;26:32–40. 968
5. Silva-Almeida M, Pereira BA, Ribeiro-Guimarães M, Alves C. Proteinases as 969
virulence factors in *Leishmania* spp. infection in mammals. *Parasit Vectors.* 970
2012;5:160. 971
6. Bordier C. The promastigote surface protease of leishmania. *Parasitol Today.* 972
1987;3:151–3. 973
7. Olivier M, Atayde VD, Isnard A, Hassani K, Shio MT. *Leishmania* virulence factors 974
: focus on the metalloprotease GP63. *Microbes Infect.* 2012;14:1377–89. 975
8. Yao C, Donelson JE, Wilson ME. The major surface protease (MSP or GP63) 976
of *Leishmania* sp. biosynthesis, regulation of expression, and function. *Mol* 977
Biochem Parasitol. 2003;132:1–16. 978
9. Gupta G, Oghum S, Satoskar AR. Mechanisms of immune evasion in 979
Leishmaniasis. *Adv Appl Microbiol.* 2014;82:155–84. 980
10. Pódnovskaia M, Descoteaux A. *Leishmania* and the macrophage: a 981
multifaceted interaction. *Future Microbiol.* 2015;10:111–29. 982
11. Atayde VD, Hassani K, da Silva Lira Filho A, Borges AR, Adhikari MC, et al. 983
Leishmania exosomes and other virulence factors: impact on innate 984
immune response and macrophage functions. *Cell Immunol.* 2016;309:7–18. 985
12. Olivier M, Hassani K. Protease inhibitors as prophylaxis against leishmaniasis: 986
new hope from the major surface protease gp63. *Future Med Chem.* 2010;2: 987
539–42. 988
13. Contreras I, Gómez MA, Nguyen O, Shio MT, McMaster RW, Olivier M. 989
Leishmania-induced inactivation of the macrophage transcription factor AP- 990
1 is mediated by the parasite metalloprotease GP63. *PLoS Pathog.* 2010;6: 991
e1001148. 992
14. Jaramillo M, Gomez MA, Larsson O, Shio MT, Topisirovic I, Contreras I, 993
et al. *Leishmania* repression of host translation through mTOR cleavage 994
is required for parasite survival and infection. *Cell Host Microbe.* 2011;9: 995
331–41. 996
15. Ramamoorthy R, Donelson E, Paetz E, Maybodi M, Roberts SC, Wilson ME. 997
Three distinct RNAs for the surface protease gp63 are differentially 998
expressed during development of *Leishmania donovani chagasi* 999
promastigotes to an infectious form. *J Biol Chem.* 1992;267:1888–95. 1000
16. Schneider P, Rosat JP, Bouvier J, Louis J, Bordier C. *Leishmania major*: 1001
differential regulation of the surface metalloprotease in amastigote and 1002
promastigote stages. *Exp Parasitol.* 1992;75:196–206. 1003
17. D'Avila-Levy CM, De Almeida Dias F, ACN DM, Martins JL, De Carvalho 1004
Santos Lopes AH, ALS DS, et al. Insights into the role of gp63-like proteins in 1005
lower trypanosomatids. *FEMS Microbiol Lett.* 2006;254:149–56. 1006
18. Pereira FM, Bernardo PS, Dias Junior PFF, Silva BA, Romanos MTV, D'Avila- 1007
Levy CM, et al. Differential influence of gp63-like molecules in three distinct 1008
Leishmania species on the adhesion to insect cells. *Parasitol Res.* 2009;104: 1009
347–53. 1010
19. D'Avila-Levy CM, Altoé ECF, Uehara LA, Branquinho MH, Santos ALS. 1011
GP63 function in the interaction of Trypanosomatids with the 1012
invertebrate host: facts and prospects. In: *Proteins and Proteomics of* 1013
Leishmania and *Trypanosoma*, vol. 74; 2014. p. 253–70. 1014
20. Santos ALS, Branquinho MH, D'Avila-Levy CM. The ubiquitous gp63-like 1015
metalloprotease from lower trypanosomatids: in the search for a function. 1016
An Acad Bras Cienc. 2006;78:687–714. 1017
21. Ivens AC, Peacock CS, Worthey EA, Murphy L, Berriman M, Sisk E, et al. The 1018
genome of the kinetoplastid parasite, *Leishmania major*. *Science.* 2005;309: 1019
436–42. 1020
22. Peacock CS, Seeger K, Harris D, Murphy L, Ruiz JC, Quail MA, et al. 1021
Comparative genomic analysis of three *Leishmania* species that cause 1022
diverse human disease. *Nat Genet.* 2007;39:839–47. 1023

- 1024 23. Victoir K, Dujardin JC, de Doncker S, Barker DC, Arevalo J, Hamers R, et al. 1025
1026 Plasticity of gp63 gene organization in *Leishmania (Viannia) braziliensis* and
1027 *Leishmania (Viannia) peruviana*. Parasitology. 1995;111:265–73.
- 1028 24. Steinkraus H, Greer J. DC. S, PJ L. sequence heterogeneity and polymorphic 1029
1030 gene arrangements of the *Leishmania guyanensis* gp63 genes. Mol Biochem
1031 Parasitol. 1993;62:173–86.
- 1032 25. Voth BR, Kelly BL, Joshi PB, Ivens AC, McMaster WR. Differentially expressed 1033
1034 *Leishmania major* gp63 genes encode cell surface leishmanolysin with
1035 distinct signals for glycosylphosphatidylinositol attachment. Mol Biochem
1036 Parasitol. 1998;93:31–41.
- 1037 26. Victoir K, Arevalo J, De Doncker S, Barker DC, Laurent T, Godfroid E, et al. 1038
1039 Complexity of the major surface protease (msp) gene organization in
1040 *Leishmania (Viannia) braziliensis*: evolutionary and functional implications. 1041
1042 Parasitology. 2005;131:207–14.
- 1043 27. Llanes A, Restrepo CM, Del Vecchio G, Anquiza FJ, Lleone R. The 1044
1045 genome of *Leishmania panamensis*: insights into genomics of the *L.*
1046 (*Viannia*) subgenus. Sci Rep. 2015;5:8550.
- 1047 28. Rogers MB, Hilley JD, Dickens NJ, Wilkes J, Bates PA, Depledge DP, et al. 1048
1049 Chromosome and gene copy number variation allow major structural
1050 change between species and strains of *Leishmania*. Genome Res. 2011;21:
1051 2129–42.
- 1052 29. Coughlan S, Taylor AS, Feane E, Sanders M, Schonian G, Cotton JA, et al. 1053
1054 *Leishmania naiffi* and *Leishmania guyanensis* reference genomes highlight
1055 genome structure and gene evolution in the *Viannia* subgenus. R Soc Open
1056 Sci. 2018;5:172212.
- 1057 30. González-De La Fuente S, Peiró-Pastor R, Rastrojo A, Moreno J, Carrasco- 1058
1059 Ramiro F, Requena JM, et al. Resequencing of the *Leishmania infantum*
1060 (strain JPCMS) genome and de novo assembly into 36 contigs. Sci Rep. 1061
1062 2017;7:18050.
- 1063 31. Potter SC, Luciani A, Eddy SR, Park Y, Lopez R, Finn RD. HMMER web server: 1064
1065 2018 update. Nucleic Acids Res. 2018;46:200–4.
- 1066 32. Valdivia HO, Scholte LLS, Oliveira G, Gabaldón T, Bartholomeu DC. The 1067
1068 *Leishmania* metapathome: a comprehensive survey of *Leishmania* protein
1069 phylogenetic relationships. BMC Genomics. 2015;16:887.
- 1070 33. Raymond F, Boisvert S, Roy G, Ritt J-F, Légaré D, Isnard A, et al. Genome 1071
1072 sequencing of the lizard parasite *Leishmania tarentolae* reveals loss of genes
1073 associated to the intracellular stage of human pathogenic species. Nucleic 1074
1075 Acids Res. 2012;40:1131–47.
- 1076 34. Mauricio IL, Taxonomy L. In: Bruschi F, Gradoni L, editors. The Leishmaniases - old 1077
1078 neglected tropical diseases: Springer International Publishing; 2018. p. 15–30.
- 1079 35. McCoy JJ, Beetham JK, Ochs DE, Donelson JE, Wilson ME. Regulatory 1080
1081 sequences and a novel gene in the msp (GP63) gene cluster of
1082 *Leishmania chagasi*. Mol Biochem Parasitol. 1998;95:251–65.
- 1083 36. Schneider P, Ferguson MAJ, McConville MJ, Mehlert A, Homans SW, Bordier 1084
1085 C. Structure of the glycosyl-phosphatidylinositol membrane anchor of the
1086 *Leishmania major* promastigote surface protease. J Biol Chem. 1990;265:
1087 16955–64.
- 1088 37. Victoir K, Dujardin JC. How to succeed in parasitic life without sex? Asking 1089
1090 *Leishmania*. Trends Parasitol. 2002;18:81–5.
- 1091 38. Mauricio IL, Gaunt MW, Stothard JR, Miles MA. Glycoprotein 63 (gp63) genes 1092
1093 show gene conversion and reveal the evolution of Old World *Leishmania*.
1094 Int J Parasitol. 2007;37:565–76.
- 1095 39. Schlagenhauf E, Etges R, Metcalf P. The crystal structure of the *Leishmania* 1096
1097 *major* surface proteinase leishmanolysin (gp63). Structure. 1998;6:1035–46.
- 1098 40. Morales G, Carrillo G, Requena JM, Guzman F, Gomez LC, Patarroyo ME, 1099
1100 et al. Mapping of the antigenic determinants of the *Leishmania infantum*
1101 gp63 protein recognized by antibodies elicited during canine visceral
1102 leishmaniasis. Parasitology. 1997;114:507–16.
- 1103 41. Lynch M, Conery JS. The origins of genome complexity. Science. 2003;302:
1104 1401–4.
- 1105 42. Ma L, Chen K, Meng Q, Liu Q. An evolutionary analysis of trypanosomatid 1106
1107 GP63 proteases. Parasitol Res. 2011;109:1075–84.
- 1108 43. Medina LS, Souza BA, Queiroz A, Guimarães LH, Machado PRL, Carvalho EM, 1109
1110 et al. The gp63 gene cluster is highly polymorphic in natural *Leishmania*
1111 (*Viannia*) *braziliensis* populations, but functional sites are conserved. PLoS
1112 One. 2016;11:1–13.
- 1113 44. Kelly BL, Nelson TN, McMaster WR. Stage-specific expression in *Leishmania* 1114
1115 conferred by 3' untranslated regions of *L. major* leishmanolysin genes
1116 (GP63). Mol Biochem Parasitol. 2001;116:101–4.
- 1117 45. Ellis M, Sharma DK, Hilley JD, Coombs GH, Mottram JC. Processing and 1118
1119 trafficking of *Leishmania mexicana* GP63. Analysis using GPI8 mutants
1120 deficient in glycosylphosphatidylinositol protein anchoring. J Biol Chem. 1121
1122 2002;277:27968–74.
- 1123 46. Mcgwire BS, Chang K, Engman DM. Migration through the extracellular 1124
1125 matrix by the parasitic protozoan *Leishmania* is enhanced by surface
1126 metalloprotease GP63. Infect Immun. 2003;71:1008–10.
- 1127 47. Sutter A, Antunes D, Silva-Almeida M, Costa MG de S, Caffarena ER. 1128
1129 Structural insights into leishmanolysins encoded on chromosome 10 of
1130 *Leishmania (Viannia) braziliensis*. Mem Inst Oswaldo Cruz. 2017;112:617–25.
- 1131 48. Li L, Stoeckert C, OrthoMCL RDS. Identification of ortholog groups for 1132
1133 eukaryotic genomes. Genome Res. 2003;13:2178–89.
- 1134 49. Enright AJ, Van DS, Ouzounis CA. An efficient algorithm for large-scale 1135
1136 detection of protein families. Nucleic Acids Res. 2002;30:1575–84.
- 1137 50. Katoh K, Misawa K, Kuma K, Miyata TMAFFT. a novel method for rapid 1138
1139 multiple sequence alignment based on fast Fourier transform. Nucleic Acids
1140 Res. 2002;30:3059–66.
- 1141 51. Eddy S. Profile hidden Markov models. Bioinformatics. 1998;14:755–63.
- 1142 52. Capella-Gutiérrez S, Silla-Martinez JM, Gabaldón T. trimAl: a tool for 1143
1144 automated alignment trimming in large-scale phylogenetic analyses.
1145 Bioinformatics. 2009;25:1972–3.
- 1146 53. Abascal F, Zardoya R, Posada D. ProtTest: selection of best-fit models of 1147
1148 protein evolution. Bioinformatics. 2005;21:2104–5.
- 1149 54. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New 1150
1151 algorithms and methods to estimate maximum-likelihood phylogenies:
1152 assessing the performance of PhyML 3.0. SystBiol. 2010;59:307–21.
- 1153 55. Huelsenbeck JP, Ronquist F. MRBAYES: Bayesian inference of phylogenetic 1154
1155 trees. Bioinformatics. 2001;17:754–5.
- 1156 56. Ronquist F, Huelsenbeck JP. MrBayes 3: Bayesian phylogenetic inference 1157
1158 under mixed models. Bioinformatics. 2003;19:1572–4.
- 1159 57. Martin DP, Murrell B, Golden M, Khoosal A, Muhire B. RDP4: detection and 1160
1161 analysis of recombination patterns in virus genomes. Virus Evol. 2015;1:1–5.
- 1162 58. Biasini M, Bienert S, Waterhouse A, Arnold K, Studer G, Schmidt T, et al. 1163
1164 SWISS-MODEL: modelling protein tertiary and quaternary structure using
1165 evolutionary information. Nucleic Acids Res. 2014;42:252–8.
- 1166 59. R. a L, J. a R, MacArthur MW, Kaptein R, Thornton JM. AQUA and PROCHECK- 1167
1168 NMR: programs for checking the quality of protein structures solved by
1169 NMR. J Biomol NMR. 1996;8:477–86.
- 1170 60. Chen J, Liu H, Yang J, Chou KC. Prediction of linear B-cell epitopes using 1171
1172 amino acid pair antigenicity scale. Amino Acids. 2007;33:423–8.
- 1173 61. EL-Manzalawy Y, Dobbs D, Honavar V. Predicting linear B-cell epitopes using 1174
1175 string kernels Yasser. J Mol Recognit. 2008;21:243–55.
- 1176 62. Larsen JEP, Lund O, Nielsen M. Improved method for predicting linear B-cell 1177
1178 epitopes. Immunome Res. 2006;2:2.
- 1179 63. Ansari HR, Raghava GP. Identification of conformational B-cell epitopes in an 1180
1181 antigen from its primary sequence. Immunome Res. 2010;6:6.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

