

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Instituto de Ciências Biológicas

Programa de Pós-graduação em Bioinformática

Raíssa Medina Santos

**“Diversidade e evolução de Metaloproteases da peçonha da aranha
Loxosceles laeta peruana revelada por análise transcriptômica”**

Belo Horizonte

2018

Raíssa Medina Santos

**“Diversidade e evolução de Metaloproteases da peçonha da aranha
Loxosceles laeta peruana revelada por análise transcriptômica”**

Dissertação apresentada ao curso de Pós-Graduação em Bioinformática do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais como requisito para obtenção de título de Mestre.

Orientador: Carlos Delfín Chávez Olórtegui

Coorientadora: Clara Guerra Duarte

Belo Horizonte

2018

AGRADECIMENTOS

A Deus, pela vida.

Ao departamento de Bioquímica e Imunologia e ao Programa de Pós-Graduação em Bioinformática – UFMG pela oportunidade que me foi oferecida.

Ao Professor Dr. Carlos Chávez Olórtgui, pela orientação, oportunidade, confiança, experiência e aprendizado.

À Dra. Clara Guerra, pela coorientação, conselhos, paciência e grande disponibilidade.

À doutoranda Sabrina Lima por todo incentivo, disponibilidade, e auxílio nos experimentos *in vitro*.

À doutoranda Patrícia Donado por todos os ensinamentos e disposição para o esclarecimento de dúvidas.

Ao Dr. Silvio Sanchez, pela disponibilização dos soros utilizados.

Ao Dr. Ricardo Andrez, pela síntese das membranas.

A todos os colegas do Laboratório de Imunoquímica de Proteínas, pelo convívio, companheirismo e amizades.

Às agências financiadoras CAPES, CNPq e FAPEMIG, pelo fomento.

Aos amigos das disciplinas do curso Bioinformática, Daniel e Rafael, pelo apoio, conversas e risadas.

À minha amiga Maria Carolina, por toda compreensão e atenção em todos os momentos.

Aos meus amigos Christina e Silvio, pela amizade e bons momentos vividos no dia a dia do laboratório.

Aos meus pais, Cosme e Cleonícia, e meu irmão Lucas, pelo amor incondicional, compreensão e apoio em todos os momentos.

Ao Júlio Machado, pelo incentivo, paciência, apoio emocional e amor.

E a todos que de alguma forma estiveram envolvidos e colaboraram com este trabalho.

RESUMO

As aranhas do gênero *Loxosceles* expressam em suas glândulas de veneno diversas toxinas, como as Fosfolipases D, responsáveis pela maior parte dos sintomas causados pelo Loxoscelismo e as Metaloproteases, que podem estar envolvidas nas ações hemorrágicas e de disseminação do veneno. Diversos estudos demonstram que os casos cutâneos-viscerais, a forma mais grave do Loxoscelismo, são causados principalmente pela aranha *Loxosceles laeta*, sendo esta encontrada em diversos países da América de Sul, como o Brasil, Peru e Chile. Além disso, existe um número quase três vezes maior de óbitos relacionados à mordida da *L. laeta* no Peru - *L. laeta* (P) - em relação ao Brasil. Desta forma, o presente estudo objetivou a realização de uma análise dos transcritos da glândula de veneno da espécie *L. laeta* (P) por ferramentas bioinformáticas com foco nas metaloproteases, por não possuírem suas funções totalmente identificadas, a fim de descrevê-las e caracterizá-las, investigando diferenças intraespecíficas dos venenos que possam justificar a gravidade elevada dos acidentes peruanos. Foram descritas nove metaloproteases do veneno de *L. laeta* (P) e estes resultados foram validados por experimentos *in silico* e *in vitro* comparando-os com as metaloproteases do veneno de *L. laeta* brasileira - *L. laeta* (B). Os resultados demonstraram diferenças importantes entre as sequências de metaloproteases peruanas com todas as outras metaloproteases de espécies de *Loxosceles* bem como de sua atividade *in vitro*. Esses dados preliminares indicam que a investigação mais aprofundada da ação dessas enzimas no veneno Loxoscélico pode contribuir para o melhor entendimento do envenenamento causado por essas aranhas.

ABSTRACT

Spiders of the genus *Loxosceles* express in their venom glands various toxins, such as Phospholipases D, responsible for most of the symptoms caused by Loxoscelism and Metalloproteases, which may be involved in the hemorrhagic actions and the dissemination of venom. Several studies have shown that the cutaneous-visceral cases, the most severe form of Loxoscelism, are mainly caused by the *Loxosceles laeta* spider, which is found in several countries of South America, such as Brazil, Peru and Chile. In addition, there are almost three times as many bite-related deaths from *L. laeta* in Peru compared to Brazil. Thus, the present study aimed to perform an analysis of the transcripts of the venom gland of the Peruvian *L. laeta* species by bioinformatic tools with a focus on metalloproteases, because they did not have their fully identified functions, in order to describe and characterize them, investigating intraspecific differences of poisons that may justify the high severity of Peruvian accidents. Nine metalloproteases from the *L. laeta* venom from Peru were described and these results were validated by *in silico* and *in vitro* experiments comparing them with the metalloproteases from the Brazilian lax *Loxosceles* venom. The results demonstrated important differences between the Peruvian metalloprotease sequences with all other metalloproteases of *Loxosceles* species as well as their *in vitro* activity. These preliminary data indicate that further investigation of the action of these enzymes on the loxoscelic venom may contribute to a better understanding of the poisoning caused by these spiders.

SUMÁRIO

1. INTRODUÇÃO	01
1.1. ARANHAS.....	01
1.2. EPIDEMIOLOGIA DE ACIDENTES COM ARANHAS	02
1.3. ARANHAS DO GÊNERO LOXOSCELES	03
1.4. LOXOSCELISMO	05
1.5. O VENENO LOXOSCÉLICO.....	08
1.6. AS ASTACINAS	10
1.7. A TRANSCRIPTÔMICA POR RNA-Seq.....	11
1.8. PREDIÇÃO DE EPITOPOS EM TOXINAS DE VENENOS E SUAS APLICAÇÕES	15
2. JUSTIFICATIVA	17
3. OBJETIVOS	18
3.1. OBJETIVO GERAL	18
3.2. OBJETIVOS ESPECÍFICOS.....	18
4. MATERIAIS E MÉTODOS	19
4.1. ANIMAIS E VENENOS.....	19
4.2. SEQUENCIAMENTO DE NOVA GERAÇÃO – MISEQ ILLUMINA.....	19
4.3. ANÁLISE DA QUALIDADE DAS SEQUÊNCIAS.....	19
4.4. PROCESSAMENTO DAS <i>READS</i>	20
4.5. MONTAGEM DAS <i>READS</i>	20
4.6. SELEÇÃO DE SEQUÊNCIAS SIMILARES AS METALOPROTEASES	21
4.7. CARACTERIZAÇÃO DAS METALOPROTEASES	22
4.7.1. ALINHAMENTO DE SEQUÊNCIAS	22
4.7.2. ANÁLISES FÍSICO-QUÍMICAS.....	23
4.7.3. PREDIÇÃO DE EPITOPOS	23

4.7.4.	FILOGENIA.....	24
4.7.5.	MODELAGEM.....	25
4.8.	VALIDAÇÃO DA PRESENÇA DE METALLOPROTEASES <i>IN VITRO</i>	25
4.8.1.	QUANTIFICAÇÃO DAS PROTEÍNAS DOS VENENOS	25
4.8.2.	SDS-PAGE E <i>WESTERN BLOTTING</i>	25
4.8.3.	ENSAIOS ENZIMÁTICOS.....	26
4.8.3.1.	ATIVIDADES FIBRINOGENOLÍTICAS	26
4.8.3.2.	ENSAIO DE GELATINASE/COLAGENASE	27
4.8.4.	MAPEAMENTO DE EPITOPOS PELO MÉTODO DE SPOT	27
5.	RESULTADOS E DISCUSSÃO	30
5.1.	ANÁLISE DA QUALIDADE DAS SEQUÊNCIAS.....	30
5.2.	PROCESSAMENTO DAS <i>READS</i>	31
5.3.	MONTAGEM DAS <i>READS</i>	32
5.4.	SELEÇÃO DE SEQUÊNCIAS SIMILARES AS METALLOPROTEASES	33
5.5.	CARACTERIZAÇÃO DAS METALLOPROTEASES	35
5.5.1.	ALINHAMENTO DE SEQUÊNCIAS	35
5.5.2.	ANÁLISES FÍSICO-QUÍMICAS.....	37
5.5.3.	PREDIÇÃO DE EPITOPOS	45
5.5.4.	FILOGENIA.....	48
5.5.5.	MODELAGEM.....	49
5.6.	VALIDAÇÃO DA PRESENÇA DE METALLOPROTEASES <i>IN VITRO</i>	51
5.6.1.	SDS-PAGE E <i>WESTERN BLOTTING</i>	51
5.6.2.	ENSAIOS ENZIMÁTICOS.....	52
5.6.2.1.	ATIVIDADES FIBRINOGENOLÍTICAS	52
5.6.2.2.	ENSAIO DE GELATINASE/COLAGENASE	54
5.6.3	MAPEAMENTO DE EPITOPOS PELO MÉTODO DE SPOT	55
6.	CONCLUSÕES	68

7. **PERSPECTIVAS**69

8. **BIBLIOGRAFIA**70

ÍNDICE DE FIGURAS

Figura 1 - Morfologia externa das aranhas	02
Figura 2 - Aspecto da aranha marrom	03
Figura 3 - Distribuição mundial de espécies <i>Loxosceles</i> de importância médica.	04
Figura 4 - Espécimes adultos de <i>L. laeta</i> coletados no Município do Rio de Janeiro, RJ, Brasil.....	05
Figura 5 - Loxoscelismo cutâneo.....	06
Figura 6 - Loxoscelismo cutâneo-visceral.....	06
Figura 7 - Gráfico de identificação da abundância de transcritos presentes na glândula de veneno de <i>L. laeta</i>	09
Figura 8 - Predição da estrutura tridimensional da LALP3	11
Figura 9 - Um experimento típico de RNA-Seq.....	12
Figura 10 - Visão geral da montagem e análise em Trinity	21
Figura 11 - Valores de qualidade Q de Phred por base de reads brutas em (A) R1 e (B) R2	30
Figura 12 - Valores de qualidade Q de Phred por base de reads após trimagem em (A) R1 e (B) R2	32
Figura 13 – Similaridade entre nove LALP-LPs pelo programa Emboss Stretcher.....	34
Figura 14 - Alinhamento de astacina de <i>Astacus astacus</i> , LALPs 1 a 3 (metaloproteases de <i>L. intermedia</i>) LALP4 (metaloprotease de <i>L. laeta</i> (B)), LALP5 (metaloprotease de <i>L. gaúcho</i>) e metaloproteases de <i>L. laeta</i> (P) (LALP-LP) pelo Clustal Ômega e Jalview ..	36
Figura 15 - Análise das sequências de metaloproteases de <i>L. laeta</i> (P) LALP-LPs, pelo Expassy - SignalP.....	40
Figura 16 - Predição dos sítios de N-glicosilação de metaloproteases de <i>L. laeta</i> (P) (LALP-LPs) pelo NetNGlyc.....	42
Figura 17 - Predição de epitopos de LALP-LP3 utilizando ABCPred.....	45
Figura 18 - Predição de epitopos de LALP-LP4 utilizando ABCPred.....	46
Figura 19 - Predição de epitopos de LALP-LP8 utilizando ABCPred.....	47
Figura 20 - Predição de epitopos de LALP-LP3, LALP-LP4 e LALP-LP8 utilizando BepiPred-2.0.....	47
Figura 21 - Árvore filogenética de máxima verossimelhança de astacina de <i>Astacus astacus</i> , LALPs 1 a 3 (metaloproteases de <i>L. intermedia</i>) LALP4 (metaloprotease de <i>L. laeta</i> (B)) e metaloproteases de <i>L. laeta</i> (P) (LALP-LP) pelo Mega 7.0	48

Figura 22 - Modelagem das metaloproteases identificadas por transcriptômica utilizando o programa Phyre2 e análise dos resultados feita pelo programa PyMOL	49
Figura 23 - Modelagem das nove metaloproteases identificadas por transcriptômica utilizando o programa Phyre2 e sobrepostas pelo programa PyMOL.....	50
Figura 24 – Modelagem de LALP1 e LALP-LP3 utilizando o programa Phyre2 sobrepostas pelo programa PyMOL.....	50
Figura 25 - SDS-PAGE de venenos de <i>L. intermedia</i> , <i>L. laeta</i> brasileira e peruana.....	51
Figura 26 - Western blotting de venenos de <i>L. intermedia</i> , <i>L. laeta</i> brasileira e peruana.....	52
Figura 27 - Atividade fibrinogenolítica de <i>L. laeta</i> brasileira e peruana.....	53
Figura 28 - Atividade fibrinogenolítica tempo-dependente de <i>L. laeta</i> brasileira (A) e peruana (B)	54
Figura 29 - Gráfico de ensaio de gelatinase/colagenase de venenos de <i>L. laeta</i> brasileira e peruana.....	55
Figura 30 - Mapeamento de epitopos de LALP-LP3 com soro anti-metaloprotease de <i>L. intermedia</i>	56
Figura 31 - Mapeamento de epitopos de LALP-LP4 com soro anti-metaloprotease de <i>L. intermedia</i>	57
Figura 32 - Mapeamento de epitopos de LALP-LP8 com soro anti-metaloprotease de <i>L. intermedia</i>	58
Figura 33 - Gráfico com a quantificação de coloração dos spots medido de LALP-LP3, LALP-LP4 e LALP-LP8 com soro anti-metaloprotease de <i>L. intermedia</i> medidos por ImageJ.....	59
Figura 34 - Alinhamento dos epitopos de metaloproteases de rMEPlox com sequências completas de LALP-LP3, LALP-LP4 e LALP-LP8	60
Figura 35 – Mapeamento de epitopos de LALP-LP3 com soro anti rMEPlox (Quimera) de <i>L. intermedia</i>	61
Figura 36 – Mapeamento de epitopos de LALP-LP4 com soro anti rMEPlox (Quimera) de <i>L. intermedia</i>	62
Figura 37 – Mapeamento de epitopos de LALP-LP8 com soro anti rMEPlox (Quimera) de <i>L. intermedia</i>	62
Figura 38 - Gráfico com a quantificação de coloração dos spots medido de LALP-LP3, LALP-LP4 e LALP-LP8 com soro anti rMEPlox de <i>L. intermedia</i> medidos por ImageJ.....	64

Figura 39 - Controle negativo das membranas LALP-LP3, LALP-LP4 e LALP-LP8 no ensaio de SPOT	65
Figura 40 – Alinhamento das três sequências de LALP-LPS com as marcações de prováveis epitopos identificados através de técnicas bioinformáticas e de membrana de spot	66
Figura 41 – Modelagem das LALP-LPs identificadas por transcriptômica com marcação dos epitopos identificados na membrana de SPOT com soro anti metaloprotease de <i>L. intermedia</i> utilizando o programa PyMOL	66
Figura 42 – Modelagem das LALP-LPs identificadas por transcriptômica com marcação dos epitopos identificados na membrana de SPOT com soro anti quimera de <i>L. intermedia</i> utilizando o programa PyMOL.....	67

ÍNDICE DE TABELAS

Tabela 1 - Distribuição dos tipos de acidentes com tipo de aranha causadora de acordo com a faixa etária de 2007 a 2016	03
Tabela 2 - Casos reportados de Loxoscelismo no Peru (1943-1997)	07
Tabela 3 - Índice de score Phred	20
Tabela 4 - Sequências de possíveis metaloproteases de <i>L. laeta</i> (P) selecionadas a partir de BlastP e alinhadas por Jalview em relação à LALP1	34
Tabela 5 - Análise das sequências de metaloproteases de <i>L. laeta</i> (P), <i>L. intermedia</i> (LALP1) e <i>L. laeta</i> (B) (LALP4) pelo Expasy - Protparam	38
Tabela 6 - Predição dos sítios de N-glicosilação de metaloproteases de <i>L. laeta</i> (P) (LALP-LPs) e score pelo NetNGlyc	43
Tabela 7 - Predição de domínios de metaloproteases de <i>L. laeta</i> (P), <i>L. intermedia</i> (LALP1) e <i>L. laeta</i> (B) (LALP4) pelo BlastP	44
Tabela 8 - Predição de propeptídeos de metaloproteases de <i>L. laeta</i> (P) pelo ProP	44

LISTA DE ABREVIACÕES

Å	Ångström
BLAST	<i>Basic Local Alignment Search Tool</i>
BSA	Soro albumina bovina
cDNA	DNA complementar
DNA	Ácido desoxirribonucleico
E	<i>E-value</i> (valor esperado)
EDTA	Ácido etilenodiamino tetra-acético
ELISA	<i>Enzyme Linked Immuno Sorbent Assay</i>
EST	<i>Expressed Sequence Tags</i>
Gb	Gigabases
ICB	Instituto de Ciências Biológicas
kDa	Quilodaltons
Kg	Quilograma
LALP	<i>Loxosceles astacin-like proteinases</i>
LALP-LP	<i>Loxosceles astacin-like proteases peruvian</i>
<i>L. laeta</i>	<i>Loxosceles laeta</i>
<i>L. laeta</i> (B)	<i>Loxosceles laeta</i> brasileira
<i>L. laeta</i> (P)	<i>Loxosceles laeta</i> peruana
Mg	Miligrama
µg	Micrograma
µL	Microlitro
NCBI	Centro Nacional de Informação Biotecnológica
NGS	Sequenciamento de nova geração
P	Probabilidade de erro
Pb	Pares de base
PBS	Tampão fosfato salino
PBS-T	Tampão fosfato salino com Tween-20
PCR	Reação em cadeia de polimerase
Q	Valor de qualidade
rMEPLox	Proteína multiepitópica recombinante derivada de toxinas loxoscélicas
RNA	Ácido ribonucleico

RNASeq	Sequenciamento de RNA
SDS-Page	Eletroforese em gel de poliacrilamida com SDS
SMase D	Esfingomielinase D
UFMG	Universidade Federal de Minas Gerais
UFPR	Universidade Federal do Paraná
V	Volts
Vol	Volume

1. INTRODUÇÃO

1.1. ARANHAS

Os artrópodes são um filo conhecido por possuir grande capacidade de adaptação a diversos ambientes, sendo os mais populares as aranhas, escorpiões, insetos, centopeias, caranguejos e camarões (OLIVEIRA; SILVA, 2009). Dentro deste filo, as aranhas são agrupadas na classe Arachnida, na ordem Araneae. Elas estão divididas em 113 famílias, com 4.075 gêneros e aproximadamente 47.116 espécies (World Spider Catalog, 2018).

Morfologicamente, a maioria das aranhas possui exoesqueleto fino, sem segmentação externa visível e pedipalpos menos evidentes. Seu corpo é organizado em dois tágmas: cefalotórax anterior e abdômen posterior, ligados pelo pedicelo (CARVALHO; BONALDO; BRESCOVIT, 2007). O cefalotórax é provido de quatro pares de apêndices e as quelíceras são os primeiros pares de apêndices da cabeça. Os olhos são localizados na região cefálica, podendo ter oito, seis ou três olhos, dependendo da família e espécie. Na região ventral posterior são localizadas as fiandeiras, que possuem a função de externalização da seda produzida por glândulas localizadas no interior do abdômen (JUNIOR; POLEGATTO, 2008).

A alimentação das aranhas é carnívora, baseando-se principalmente em insetos, logo poucas espécies representam algum perigo aos seres humanos. Além disso, a maioria das espécies é sensível a diversos fatores físicos, como umidade, temperatura e luminosidade, e a fatores biológicos, como disponibilidade de alimento e estrutura da vegetação (Figura 1) (FOELIX, 1996).

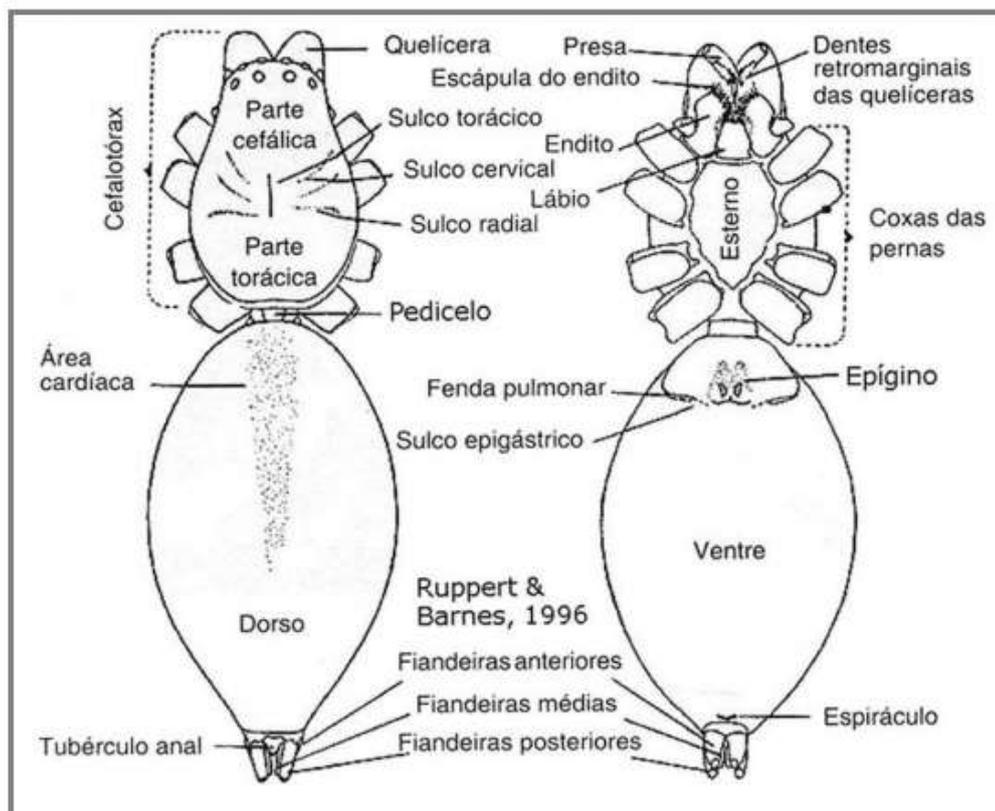


Figura 1: Morfologia externa das aranhas. Representação esquemática geral de Araneae com visão dorsal (esquerda) e ventral (direita). (Fonte: RUPPERT; BARNES, 1996)

1.2. EPIDEMIOLOGIA DE ACIDENTES COM ARANHAS

Acidentes com animais peçonhentos são considerados emergências clínicas em vários países tropicais, principalmente em áreas rurais. Apesar disso, o crescimento urbano se associando à destruição do habitat natural das aranhas as levam a buscar abrigo e alimento nas cidades e, por possuírem hábitos alimentares generalistas, a dispersão das mesmas ocorre com muita facilidade. Além disso, a urbanização ocasiona uma diminuição da quantidade dos predadores naturais das aranhas, como aves e répteis, o que contribui para um crescente índice de acidentes causados pelas aranhas (ABREU; BERNARDE; MOTA DA SILVA, 2015).

No Brasil, três gêneros de aranhas possuem maior importância na Saúde Pública: *Phoneutria* (armadeira), *Loxosceles* (aranha-marrom) e *Latrodectus* (viúva-negra), sendo *Loxosceles* responsável pela maioria dos acidentes (Loxoscelismo) e apresenta maior letalidade (Tabela 1) (ABREU; BERNARDE; MOTA DA SILVA, 2015).

Faixa etária	Phoneutria (armadeira) N° casos	Loxosceles (aranha-marrom) N° casos	Latrodectus (viúva negra) N° casos
< 1	326	954	24
1 a 4	1.395	4.005	46
5 a 9	1.609	3.601	55
10 a 14	1.700	3.771	71
15 a 19	2.323	6.394	93
20 a 39	10.857	30.379	379
40 a 59	12.463	22.845	286
60 a 64	2.584	3.120	42
65 a 69	2.007	2.208	50
70 a 79	2.418	2.206	47
> 80	663	581	18
TOTAL	38.334	80.076	1.111

Tabela 1: Distribuição dos tipos de acidentes com tipo de aranha causadora de acordo com a faixa etária de 2007 a 2016: Fonte SINAN/SVS/MS3.

1.3. ARANHAS DO GÊNERO LOXOSCELES

As aranhas do gênero *Loxosceles* (Figura 2) Heineken & Lowe, 1832, foco do estudo deste trabalho, são mundialmente distribuídas, com ocorrências nas Américas, África, Europa e Oceania. Aproximadamente 100 espécies diferentes desse gênero já foram descritas (Figura 3).

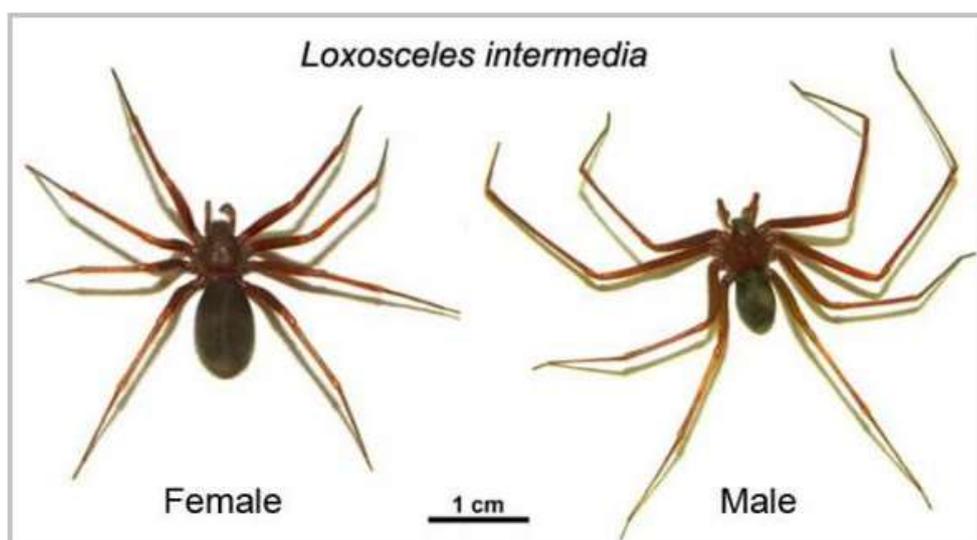


Figura 2: Aspecto da aranha marrom. Espécimes adultos de *L. intermédia* – fêmea (esquerda) e macho (direita). (Fonte: CHAIM et al., 2011)

Na América do Sul, o Brasil registra a maior parte dos acidentes Loxoscélicos que se concentram nas regiões Sul e Sudeste (Ministério da Saúde, 2011), sendo *Loxosceles intermedia*, *Loxosceles gaucho* e *Loxosceles laeta* as principais espécies relevantes na saúde pública brasileira (HOGAN; BARBARO; WINKEL, 2004). Entretanto, pode-se encontrar *Loxosceles* em outras regiões do continente, como Argentina, Chile e principalmente no Peru, onde a *Loxosceles laeta* (*L. laeta*) (Figura 4) é conhecida como “aranha assassina”, devido ao grande número de relatos de acidentes loxoscélicos fatais (MAGUINÃ-VARGAS et al., 2004).

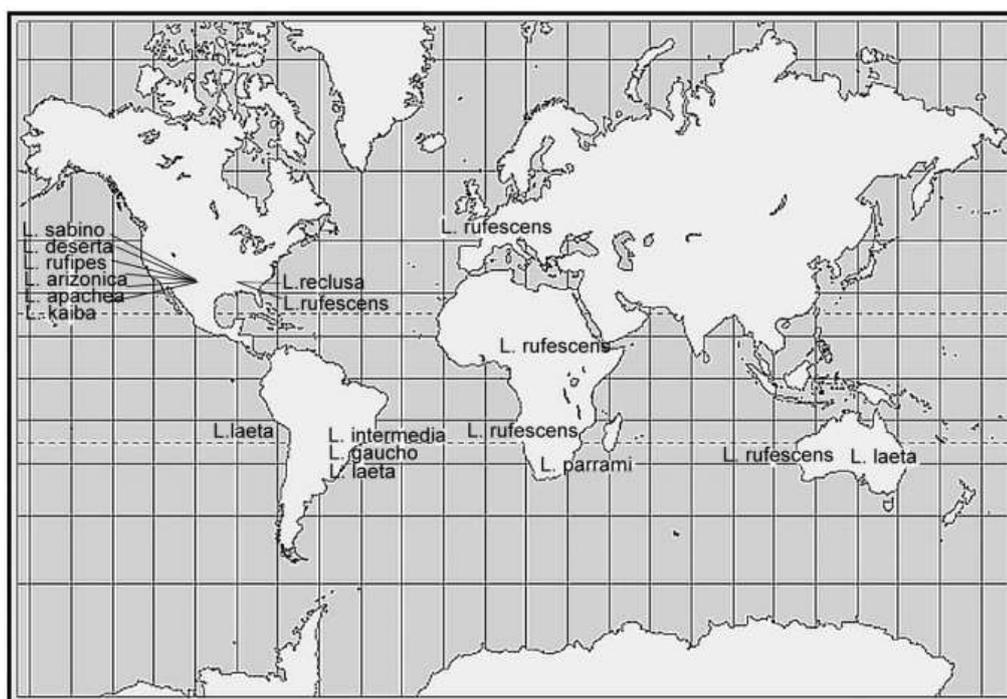


Figura 3: Distribuição mundial de espécies *Loxosceles* de importância médica. (Fonte: HOGAN; BARBARO; WINKEL, 2004)

Popularmente conhecidas como “aranha marrom” ou “aranha violino” (VETTER; VISSCHER; D, 1998), as aranhas *Loxosceles* apresentam coloração marrom, com variações mais escuras ou mais claras, de acordo com a espécie, e uma mancha escura no cefalotórax em forma de violino. Seu tamanho é único, medindo cerca de 2 a 3 cm, sendo as aranhas machos menores. Possuem 6 olhos distribuídos em 3 pares (CHATZAKI et al., 2012; DA SILVA et al., 2004; SWANSON; VETTER, 2006).

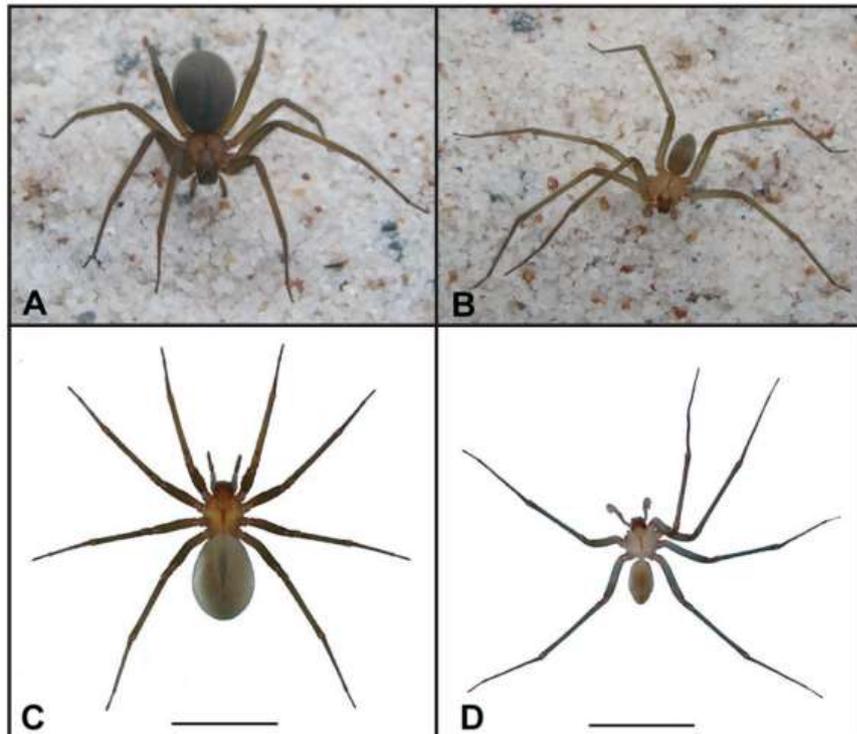


Figura 4: Espécimes adultos de *L. laeta* coletados no Município do Rio de Janeiro, RJ, Brasil. A: fêmea em vida, B: macho em vida, C: fêmea preservada, D: macho preservado. Barra de escala: 1cm (Fonte: SILVEIRA, 2009).

1.4. LOXOSCELISMO

O termo utilizado para representar os acidentes envolvendo aranhas do gênero *Loxosceles* é o *Loxoscelismo* e o quadro clínico pode se apresentar de duas formas: cutâneo e cutâneo-visceral (hemolítica). A maioria dos casos, cerca de 87% a 98%, desenvolve a forma cutânea, que se caracteriza após 2 a 8 horas por dor, edema endurecido e eritema no local da mordida. Após 24 a 72 horas, os sintomas da lesão podem variar de incharacterístico, onde se apresenta uma bolha de conteúdo seroso, calor e rubor, com ou sem dor em queimação; sugestiva, que apresenta endureção, bolha, equimoses e dor em queimação; ou característica, com dor em queimação, lesões hemorrágicas focais, mescladas com áreas pálidas de isquemia e necrose. Em cerca de 7 a 12 dias a lesão pode evoluir para uma necrose seca (escara) que com algumas semanas, ao se destacar, deixa uma úlcera de difícil cicatrização (Figura 5) (DA SILVA et al., 2004; HOGAN; BARBARO; WINKEL, 2004; MINISTERIO DE SAUDE, 2001; SILVEIRA, 2009).



Figura 5: Loxoscelismo cutâneo. Evolução de um paciente com quadro de Loxoscelismo cutâneo, mostrando os dias 5 (A), 27 (B), 41 (C) e 76 (D) após a picada. (Fonte: WEN, 2009)

O Loxoscelismo cutâneo-visceral está presente em 1% a 13% dos casos. É comumente causado pela espécie *L. laeta*, apresentando manifestações clínicas em virtude de hemólise intravascular. Nas primeiras 24 horas, é possível observar icterícia, anemia e hemoglobinúria, podendo ser acompanhados por petéquias e equimoses. Tornando-se mais grave, os casos podem evoluir para insuficiência renal aguda, sendo a principal causa de óbito pelo Loxoscelismo (MINISTERIO DE SAUDE, 2001) (Figura 6).



Figura 6: Loxoscelismo cutâneo-visceral. Evolução da urina de um paciente com Loxoscelismo cutâneo-visceral. (Fonte: MAGUIÑA-VARGAS C., 2006)

Os casos de acidentes loxoscélicos são classificados de acordo com as alterações clínicas-laboratoriais e a identificação da aranha. Os casos ditos como leves apresentam lesão incharacterística, sem alterações clínicas e laboratoriais. Estes pacientes devem

realizar um acompanhamento de 72 horas caso haja necessidade de mudança da sua classificação por alterações nas características da lesão. Em casos moderados, a lesão é sugestiva ou característica e em casos graves se observa a presença da lesão característica acompanhada de alterações clínico-laboratoriais de hemólise intravascular (MINISTERIO DE SAUDE, 2001).

Apesar do Loxoscelismo cutâneo-visceral (casos graves) estar presente na menor porcentagem de casos, é observado que o veneno de *L. laeta* é responsável por uma maior quantidade de casos com efeitos sistêmicos do que apenas locais, sendo responsável por um maior número de óbitos e conseqüentemente por uma necessidade de estudos mais aprofundados sobre a mesma (GONÇALVES-DE-ANDRADE; VILARINHO TAMBOURGI, 2003).

No Brasil, os casos de Loxoscelismo acarretados pela mordida de *L. laeta* se concentram na região Sul, principalmente no estado de Santa Catarina, mas no Peru é considerada uma espécie endêmica, onde os casos de Loxoscelismo cutâneo-visceral são responsáveis por quase o dobro em relação aos apenas cutâneos (HERNAN; ZAVALETA, 1997) (Tabela 2).

AUTORES	LUGAR	PERIODO	CASOS		
			Cutâneo	Sistêmico	Total
Yzu	Lima	1943-52	18	13	31
Pesce y Lumberas	Lima	1954	00	01	01
Vellard	Lima	1954	00	01	01
Vellard	Lima	1955	02	01	03
Cuadra Calle	Lima-Ica	1956	00	03	03
Montoya	Camana	1954-59	10	11	21
Burga	Lima	1956-60	21	00	21
Nâquira y Col	Arequipa	1959-62	06	00	06
Maya	Lima	1961-62	00	03	03
Chávez	Trujillo	1960-63	03	03	06
Lândazuri y Col	Lima	1957-64	05	00	05
Delgado	Lima	1962-69	49	03	52
Mendoza	Trujillo	1963-70	00	01	01
Barrantes	Trujillo	1964-70	07	07	14
Gonzales	Lima	1973	01	01	01
Salazar y Col	Lima	1975-76	00	04	04
Zavaleta A. y Col	Lima	1971-80	28	05	33
Gonzales y Col	Lima	1969-81	11	03	14
Zavaleta A	Lima	1981	02	00	02
Zavaleta R.**	Trujillo	1983	11	04	15
Kemper	Lima	1985	75	25	100
Maguiña y Col	Lima	1984-86	16	00	16
Maguiña y Col	Lima	1985-86	13	00	13
Maguiña y Col	Lima	1987-91	39	00	39
Mangiante**	Lima	1980-90	39	50	89

* Modificado del reporte de Sanabria y Zavaleta ¹¹
 ** Estudios en que sólo se incorporaron niños.

Tabela 2: Casos reportados de Loxoscelismo no Peru (1943-1997). (Fonte: HERNAN; ZAVALETA, 1997).

1.5. O VENENO LOXOSCÉLICO

O veneno das aranhas do gênero *Loxosceles* é um líquido cristalino formado por uma mistura complexa de proteínas e toxinas peptídicas produzidos por um par de glândulas situadas no cefalotórax ligadas ao par de quelíceras, com um perfil com predominância de proteínas de baixa massa molecular que variam de 1 a 40 kDa (CHAIM et al., 2011). Este veneno vem sendo estudado durante muitos anos por diversos grupos de pesquisas e muitas famílias de toxinas foram identificadas, como as hialuronidasas, fosfolipases-D (FUTRELL, 1992), metaloproteases, serinoproteases e peptídeos com atividade inseticida, que chega a representar mais de 50% da expressão no veneno de *Loxosceles intermedia*, identificadas por atividades fibrinogenolíticas e gelatinolíticas (FEITOSA et al., 1998). Apesar disso, a composição do veneno loxoscélico não está totalmente esclarecida. Em 2008, Fernandes-Pedrosa e colaboradores observaram que em um transcriptoma das glândulas de veneno de *L. laeta* brasileira, através da técnica de Sanger de sequenciamento para identificação de ESTs, possuía uma porcentagem total de 30,9% de transcritos que indicavam ser toxinas e, além disso, que 25% do total de transcritos não possuía nenhuma correspondência em bancos de dados (FERNANDES-PEDROSA et al., 2008) (Figura 7).

O principal grupo de proteínas com grande expressão no veneno de *Loxosceles laeta* é o da Fosfolipase-D, ou Proteínas Loxtox. Elas são as principais causadoras de dermonecrose *in vivo*, além de serem importantes na manifestação dos sintomas clínicos de Loxoscelismo cutâneo e sistêmico (DA SILVA et al., 2004; SWANSON; VETTER, 2006). Experimentalmente já foi observada sua capacidade de produzir uma resposta inflamatória, agregação plaquetária, hemólise, aumento da permeabilidade vascular, nefrotoxicidade, letalidade em camundongos e efeitos citotóxicos em linhagens celulares estabelecidas (CHAIM et al., 2011; DA SILVA et al., 2004; FUTRELL, 1992; RIBEIRO et al., 2015).

Outro grupo de proteínas que também já foi identificado no veneno de *Loxosceles* é denominado hialuronidase. Ela possui uma massa molecular de 33 a 63kDa e foi reconhecida por aumentar o efeito das fosfolipases recombinantes em pele de coelhos, confirmando sua ação de propagação do veneno (RIBEIRO et al., 2015). Já as serinoproteases foram primeiramente identificadas em *L. laeta* por Fernandes-Pedrosa e colaboradores em 2008. Elas são proteases que clivam ligações peptídicas e se caracterizam por possuírem um resíduo de aminoácido serina no sítio catalítico, mas

apesar disso não se sabe qual sua função na ação do veneno (FERNANDES-PEDROSA et al., 2008).

Os peptídeos inseticidas possuem baixa massa molecular, variando de 2 a 12 kDa e são classificadas de acordo com seu modo de ação: Afetando a transmissão glutamatérgica de cálcio, sódio, potássio e cloreto. São amplamente encontradas no veneno de *L. intermedia* e possuem de 53 a 69 resíduos de aminoácidos, mas apesar disso, ainda não foram descritos seus mecanismos de ação.

Já as atividades das metaloproteases (*astacin-like*), escolhidas como principal alvo deste trabalho, foram descritas nas famílias de *L. intermedia* e *L. gaucho* como ações caseinolíticas, gelatinolíticas e fibrinogenolíticas, sugerindo participação no processo de dermonecrose (FEITOSA et al., 1998).

Identificação da abundância de transcritos presentes na glândula de veneno de *Loxosceles laeta* (em %)

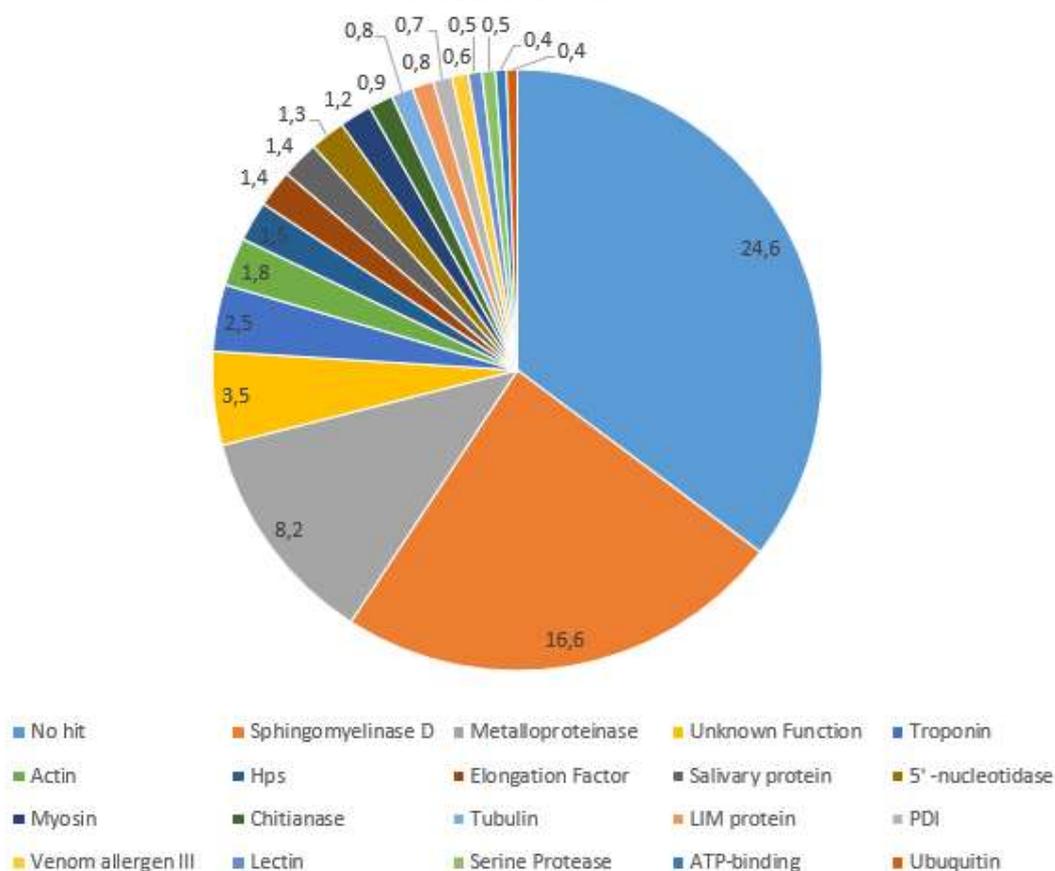


Figura 7: Gráfico de identificação da abundância de transcritos presentes na glândula de veneno de *L. laeta*. (Fonte: Adaptado de FERNANDES-PEDROSA et al., 2008)

1.6. AS ASTACINAS

Membros da família astacina de metaloproteases são formados por endopeptidases de zinco e são encontrados em todo o reino animal, como nos seres humanos, camundongos, anfíbios, peixes, insetos e moluscos, tendo papéis desde a digestão de alimentos até o processamento dos componentes da matriz extracelular. Já nas aranhas, foram identificadas proteases semelhantes as astacinas (*Loxosceles astacin-like proteases* - LALPs), sendo considerada a primeira astacina descrita em venenos (TREVISAN-SILVA et al., 2010).

Esta família foi nomeada de acordo com uma enzima digestiva astacina e primeiramente identificada em lagostim, *Astacus astacus*, sendo composta por proteases extracelulares ou ligadas à superfície celular envolvidas em funções digestivas, desempenhando papel na ativação de fatores de crescimento, degradação de polipeptídeos e processamento de moléculas extracelulares (MÖHRLÉN et al., 2006; TREVISAN-SILVA et al., 2010).

Os membros da família de metaloproteases apresentam dois elementos essenciais para essas moléculas: Zinc-binding Signature (HEXXHXXGXXHE), que é o domínio de ligação ao zinco, e Met-turn (MXY) (BODE; GOMIS-RÜTH; STÖCKLER, 1993) (Figura 7).

Além disso, no veneno das aranhas do gênero *Loxosceles*, as metaloproteases possuem papel importante na hemorragia local e na cicatrização, aumentando a permeabilidade de tecidos e permitindo facilidade de disseminação de outras toxinas. Elas possuem massa molecular de cerca de 29kDa, são dependentes de zinco e possuem peptídeo sinal. Essas enzimas foram descritas como componentes do veneno de *L. intermedia* e denominadas como LALP1, e outras duas proteoformas (forma molecular específica de um produto proteico proveniente de um gene específico) de *L. intermedia* foram denominadas LALP2 e LALP3 (Figura 8). O mesmo trabalho também denominou de LALP4 a proveniente de *L. laeta* e LALP5 de *L. gaucho* (GUIMARÃES et al., 2013).

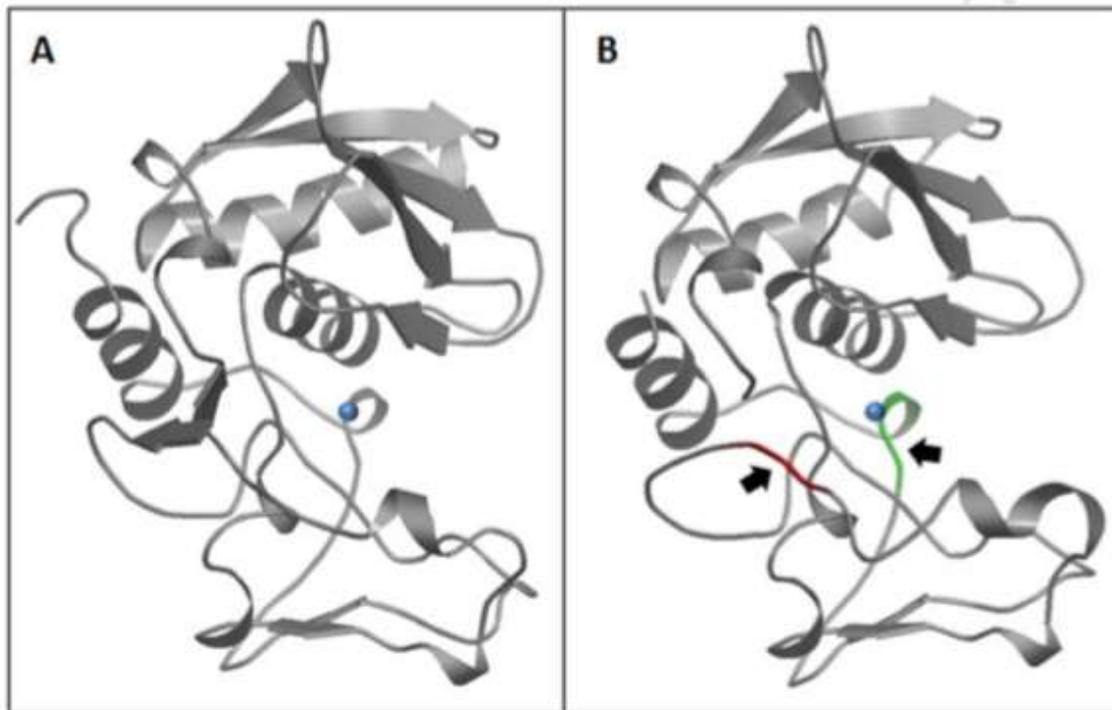


Figura 8: Predição da estrutura tridimensional da LALP3. (A) Estrutura de resolução de 1,8Å da astacina de *Astacus. astacus* utilizada como modelo para a geração do modelo LALP3. (B) Predição da estrutura de LALP3. A esfera azul representa o ion zinco interagindo com o sítio catalítico das enzimas. As alças vermelha e amarela representam os dois potenciais locais de N-glicosilação da LALP3 (setas). (Fonte: MORGON et al., 2016)

Para identificação de outras LALPs uma das possíveis estratégias é a realização do sequenciamento dos RNAs mensageiros presentes da glândula de veneno destas aranhas, e por análise transcriptômica, analisar o conjunto completo dos transcritos (RNAs), que é o reflexo direto da expressão gênica.

1.7. A TRANSCRIPTÔMICA POR RNA-Seq

O transcriptoma é um conjunto completo de transcritos em forma de moléculas de RNA de parte de um genoma (RNAs mensageiros, RNAs ribossômicos, RNAs transportadores e os microRNAs) (WANG; GERSTEIN; SNYDER, 2009), que a partir da utilização do RNA-Seq, uma metodologia NGS (Next Generation Sequencing), o revela pelo sequenciamento do DNA complementar (cDNA), gerando milhões de fragmentos de sequências curtas (*reads*) (Figura 9).

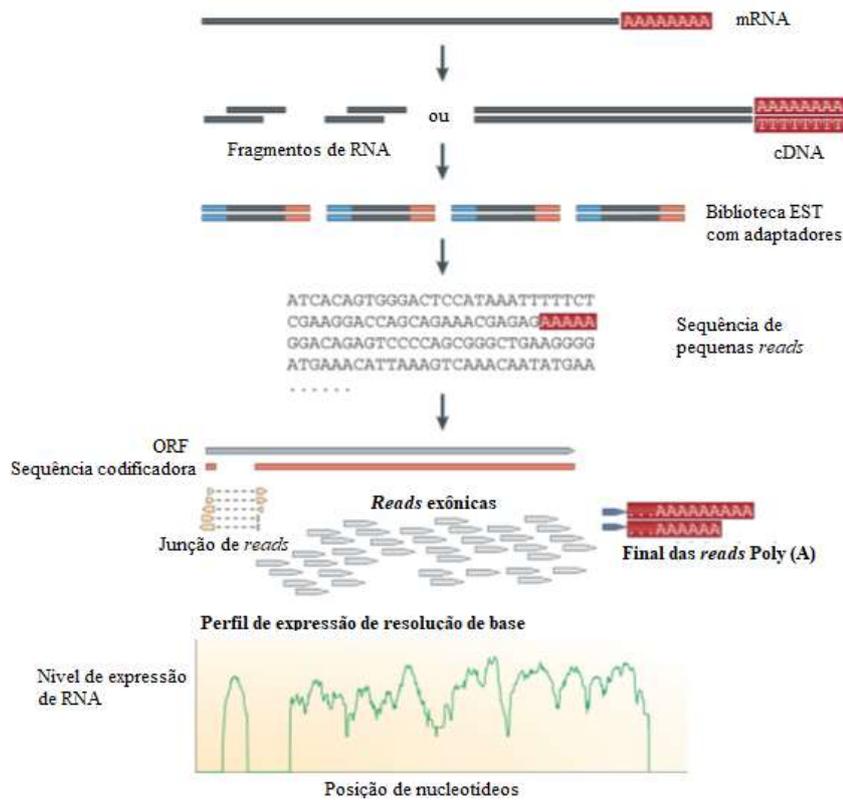


Figura 9: Um experimento típico de RNA-Seq. Resumidamente, os RNAs longos são primeiramente convertidos em uma biblioteca de fragmentos de cDNA através da fragmentação do RNA ou de DNA. Os adaptadores do sequenciamento (azul) são posteriormente adicionados a cada fragmento de cDNA e uma pequena sequência é obtida de cada cDNA utilizando a tecnologia de sequenciamento de alto rendimento. As sequências resultantes são alinhadas com um genoma ou transcriptoma de referência e classificadas em três tipos: *Reads* exônicas, junção de *reads* e final das *reads* Poly (A). Estes três tipos são utilizados para gerar um perfil de expressão de resolução de base para cada gene. (Fonte: Adaptado de WANG; GERSTEIN; SNYDER, 2009)

O primeiro estudo de transcriptoma ocorreu em 1991 na descoberta de genes humanos (ADAMS et al., 1991) e a partir deste momento, houve um crescimento exponencial na geração de dados em bancos de dados públicos de diversos organismos. Estes resultados são formados por inúmeras ESTs (*Expressed sequence tag*), que são sequências curtas versáteis com várias aplicações geradas a partir do sequenciamento pelo método de Sanger, que em 1970 era definido como padrão ouro. Este mesmo método foi utilizado para o sequenciamento da glândula de veneno de *L. laeta* brasileira por Fernandes-Pedrosa e colaboradores em 2008 para a identificação da presença de transcritos de alta abundância (FERNANDES-PEDROSA et al., 2008).

Apesar disso, muitas ESTs disponibilizadas são redundantes, de baixa qualidade ou parcialmente incompletas, o que leva a necessidade da utilização de ferramentas

bioinformáticas para melhoria da qualidade e análise destes resultados, além de uma técnica que facilite a investigação de resultados obtidos (ALVARENGA; KALAPOTHAKIS, 2012).

Desta forma, com o passar dos anos, uma técnica mais refinada, simples e que gera uma quantidade de dados significativamente maior que os gerados pelo método de Sanger foi elaborada: O sequenciamento *de novo*, utilizado neste projeto para realização da análise das metaloproteases da peçonha de aranhas *L. laeta* provenientes do Peru. Essa técnica proporciona um sequenciamento em larga escala, mais otimizado e de custo e tempo reduzidos (SCHUSTER, 2008).

A utilização do sequenciamento profundo (sequenciamento massivo de DNA) permite a utilização de tecnologias que possibilitam um sequenciamento em grande escala de fragmentos pequenos de DNA e cDNA, obtendo resultados com Gigabases (Gb) de sequências de determinado transcriptoma. Isso pode ocorrer devido a um paralelismo otimizado das reações de sequenciamento, permitindo uma maior produtividade de dados do que a metodologia anterior (SCHUSTER, 2008; GRADA; WEINBRECHT, 2013).

A partir deste momento foram criadas diversas plataformas de sequenciamento NGS, como a Illumina MiSeq, utilizada neste trabalho, um aparelho versátil que gasta em média 8 horas para a realização completa de um processo. Seu sequenciamento é baseado na síntese, em que uma fita de DNA se liga a adaptadores presentes na superfície da lâmina de sequenciamento e é amplificada por PCR em ponte, criando *clusters* que irão conter cópias do mesmo fragmento de DNA. Por fim, deoxinucleotídeos serão adicionados a lâmina permitindo a síntese de uma nova fita. Este processo de montagem de biblioteca ocorre de duas formas distintas: Paired-End ou Single-end, e cada sequenciamento pode gerar até 600 bilhões de *reads* (WANG; GERSTEIN; SNYDER, 2009).

Após o sequenciamento, deve ser realizado um processo de análise da qualidade das sequências geradas, pois as extremidades dos sequenciamentos podem comumente possuir uma qualidade inferior ao desejado. Para isto, utiliza-se programas como FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>) para análise da qualidade e o programa Trimmomatic (<http://www.usadellab.org/cms/?page=trimmomatic>) para limpeza das sequências de qualidade duvidosa. Após esta análise, segue-se para o processo de montagem das *reads* obtidas através de ferramentas bioinformáticas específicas. Existem diversos montadores disponíveis para análises transcriptômicas, como por exemplo o TopHat (<https://ccb.jhu.edu/software/tophat/index.shtml>), Cufflink (<http://cole-trapnell-lab.github.io/cufflinks/>) e o programa Trinity (*RNA-Seq assembly*)

(<https://github.com/trinityrnaseq/trinityrnaseq/wiki>), que realizam um mapeamento das *reads* contra o genoma de referência para obtenção de um consenso. Já para a montagem *de novo*, utilizam-se o Velvet-Oases (<https://www.ebi.ac.uk/~zerbino/oases/>) e novamente o Trinity, sendo este o citado como a ferramenta mais eficiente para montagens *de novo* de transcriptoma e por isso foi o escolhido como o montador a ser utilizado neste projeto. O programa utiliza o algoritmo de Grafo de *Bruijn*, que divide as leituras em k-mers, onde cada aresta do grafo corresponde a um k-mer da sequência, sendo considerado o padrão ouro da técnica (SURGET-GROBA; MONTOYA-BURGOS, 2010).

Seguidamente a montagem, os resultados necessitam ser analisados e alinhados, e para isto utilizam-se comumente o BLAST (<https://blast.ncbi.nlm.nih.gov/Blast.cgi>), *Basic Local Alignment Search Tool*, um programa que encontra regiões de similaridade entre as sequências e as compara com aquelas presentes em bancos de dados, além de calcular a significância estatística entre elas, ou o Diamond, que possui um alto rendimento em alinhamentos em até 20.000 vezes mais rápido que o BLAST e com uma alta sensibilidade. Isso permitirá com que o conteúdo analisado pela transcriptômica seja identificado por correlação a outros dados já depositados nos bancos de dados disponíveis.

Após a identificação das sequências, são possíveis diversas análises que contribuem para uma correta confirmação do reconhecimento das sequências, como por exemplo as análises físico-químicas, utilizando o programa ExPasy - ProtParam (<https://web.expasy.org/protparam/>), a identificação de peptídeos sinais através do ExPasy – SignalP (<http://www.cbs.dtu.dk/services/SignalP/>), além da possibilidade de realizar uma modelagem *in silico* das sequências escolhidas, como é o caso da utilização do programa Phyre2 (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>) ou SWISS-MODEL (<https://swissmodel.expasy.org/>), que utilizam a sequência primária de uma proteína para prever através de uma referência em banco de dados qual sua provável forma.

Todas estas análises possuem grande importância para um entendimento aprofundado das sequências que constituem parte da mistura proteica do veneno de *Loxosceles*, para permitir futuras buscas de melhores formas de identificação e tratamento do Loxoscelismo, que possui um diagnóstico difícil e tratamentos com poucos ensaios clínicos que comprovem eficácia (ISBISTER; FAN, 2011).

1.8. PREDIÇÃO DE EPITOPOS EM TOXINAS DE VENENOS E SUAS APLICAÇÕES

Existem diversos métodos terapêuticos descritos na literatura para o tratamento de Loxoscelismo, sendo os principais a utilização de antissépticos, anti-histamínicos, antibióticos, anti-inflamatórios, analgésicos, dapsona, esteroides, oxigenação hibernática, procedimentos cirúrgicos e a administração de antivenenos (DA SILVA et al., 2004; ISBISTER; FAN, 2011). Este último, no Brasil, é indicado quando se identifica um quadro de hemólise cutâneo-visceral, ou quando a forma cutânea é identificada nas primeiras 72 horas. Desta forma, o paciente de caso leve não recebe a administração do soro, o moderado recebe a administração de 5 ampolas e o paciente de caso grave recebe a administração de 10 ampolas (MINISTERIO DE SAUDE, 2001).

No Brasil, estes antivenenos são produzidos pelo Instituto Butantan, um soro anti-aracnídico pela mistura de venenos das aranhas *L. gaucho*, *Phoneutria nigriventer* e do escorpião *Tityus serrulatos*; e pelo Centro de Produção e Pesquisa de Imunobiológicos do Estado do Paraná (CPPI), um soro anti-loxoscélico contendo venenos de *L. gaucho*, *intermedia* e *laeta*. Ambos são produzidos a partir de anticorpos de soro de cavalos com a inoculação dos venenos totais (HOGAN; BARBARO; WINKEL, 2004).

A extração de veneno das aranhas *Loxosceles* é uma técnica laboriosa, em que comumente se utiliza o eletrochoque, obtendo de cada aranha a liberação de cerca de 50 µg de veneno, o que leva a necessidade de 1000 extrações para a realização de um ciclo completo de imunização em um cavalo. Após a imunização, 16 litros de sangue são retirados de cada cavalo, e apenas as hemácias são devolvidas a eles. Além disso, os cavalos produtores sofrem os efeitos causados pelo envenenamento, levando-os a uma expectativa de vida reduzida.

Devido aos problemas causados pela produção e administração do soro, é importante a realização da busca por novas técnicas para melhoria ou substituição da imunização comumente utilizada. Desta forma, nosso grupo vem trabalhando ao longo dos anos na utilização de epitopos para esta produção. Felicori e colaboradores em 2009 identificaram que a proteína dermonecrótica recombinante (rLiD1) do veneno de *L. intermedia* é capaz de induzir a produção de anticorpos que neutralizam eficientemente os efeitos tóxicos do Loxoscelismo (FELICORI et al., 2009). Já Lima e colaboradores em 2018 predisseram e caracterizaram uma proteína multiepitópica quimérica (rMEPLox) contendo epitopos das principais proteínas de aranhas do gênero *Loxosceles* (LIMA et al., 2018).

Diversas ferramentas bioinformáticas permitem a predição destes epitopos e usualmente utilizam redes neurais artificiais para interligar as informações disponibilizadas com as encontradas em bancos de dados.

2. JUSTIFICATIVA

O Loxoscelismo causado pela espécie *L. laeta* é um problema de saúde pública que causa inúmeros acidentes tanto no Brasil quanto no Peru. Em função disso, a caracterização dos compostos do veneno desta aranha permitirá um melhor entendimento dos processos fisiopatológicos do envenenamento que possibilitará o desenvolvimento e melhoria de métodos de prevenção e tratamento do Loxoscelismo.

Apesar de serem causados teoricamente pela mesma espécie, os quadros clínicos dos acidentes causados por *L. laeta* no Peru e no Brasil apresentam diferenças, sendo aqueles que ocorrem no Peru geralmente mais graves. Porém, não são claras as razões que levam a essa diferença, e por isso estudos são realizados na tentativa de distingui-los. Guimarães e colaboradores em 2013 identificaram em camundongos, que a dose letal mediana (DL50) do veneno de *L. laeta* (P) é de 1,21 mg/kg enquanto do veneno de *L. laeta* (B) é 2,18 mg/kg, e que o veneno peruano possui uma maior atividade de Fosfolipases D (GUIMARÃES et al., 2013). Nesse aspecto, a identificação da variabilidade entre os componentes dos venenos de uma mesma espécie torna-se relevante para obter informações que possam ajudar a correlacionar a gravidade e altos índices de letalidade dos casos atribuídos a *L. laeta* (P) em relação a *L. laeta* (B).

Aquino e colaboradores em 2016 confirmaram por análise transcriptômica a presença de Fosfolipases D no veneno de *L. laeta* (P) em maior quantidade e com diferenças significativas em relação as encontradas no veneno da *L. laeta* (B) (AQUINO; MARTINS; CHÁVEZ-OLORTEGUI, 2016). Seguindo esta linha de investigação, as metaloproteases do veneno de *L. laeta* (P), foram escolhidas como foco deste estudo, por serem uma das principais proteínas que auxiliam na dermonecrose e na disseminação de outras toxinas, além de serem pouco estudadas.

3. OBJETIVOS

3.1. OBJETIVO GERAL

Descrever as metaloproteases da glândula de veneno das *L. laeta* peruanas por análise transcriptômica e caracterizá-las utilizando ferramentas *in silico* e *in vitro*.

3.2. OBJETIVOS ESPECÍFICOS

- Obter as sequências dos transcritos presentes na glândula de veneno de *L. laeta* (P) pelo método de sequenciamento de nova geração;
- Realizar a montagem *de novo* do transcriptoma com as *reads* geradas;
- Selecionar as sequências similares as metaloproteases de *L. laeta* e *L. intermedia*;
- Caracterizar estruturas e funções das metaloproteases identificadas por ferramentas bioinformatizadas;
- Validar *in vitro* a presença de metaloproteases no veneno de *L. laeta* proveniente do Peru em comparação com o veneno de *L. laeta* proveniente do Brasil;
- Analisar e comparar as atividades fibrinogenolíticas e colagenolíticas *in vitro* dos venenos de *L. laeta* peruanos e brasileiros;
- Identificar e mapear os epitopos das metaloproteases identificadas na transcriptômica do veneno de *L. laeta* peruano;

4. MATERIAIS E MÉTODOS

4.1. ANIMAIS E VENENOS

Cinco glândulas de aranhas de *L. laeta* fêmeas foram coletadas no Peru e processadas no Laboratório de Biotecnologia e Marcadores Moleculares (ICB-UFMG).

4.2. SEQUENCIAMENTO DE NOVA GERAÇÃO – MISEQ ILLUMINA

A biblioteca foi construída utilizando o kit TruSeq™ RNA Sample Prep Kit v3 Set A (Illumina) a partir do protocolo fornecido pelo kit, quantificada e aplicada no cartucho de sequenciamento após a desnaturação.

A biblioteca foi sequenciada pelo sequenciador de nova geração MiSeq (Illumina) pela equipe do Prof. Dr. Evanguedes Kalapothakis, utilizando dois cartuchos de sequenciamento pela técnica de *paired-end*, gerando 600 ciclos (300 para *reads* 3´5´ e 300 para *reads* 5´3´), que resultou em um arquivo fastq contendo as bases e os valores de qualidade das *reads* brutas.

4.3. ANÁLISE DA QUALIDADE DAS SEQUÊNCIAS

A partir do programa FastQC - uma ferramenta de controle de qualidade para dados de sequências de alto rendimento - verificou-se a qualidade dos dados das sequências brutas provenientes do sequenciamento. Ele avaliou, dentre diversos outros parâmetros, a quantidade total de *reads*, o conteúdo de GC e comprimento variável das sequências, oportunizando a melhoria dos resultados, quando necessário (ANDREWS, 2010).

O arquivo resultante do FastQC apresenta um valor de qualidade (Q) para cada base sequenciada denominado *Phred quality score*. Estes valores se relacionam com a probabilidade de erro (P) da nomeação de uma base, de acordo com a fórmula $Q = -10 \log P$. Desta forma, um valor de $Q = 20$, corresponde a probabilidade de 1 base errada em 100 bases nomeadas, o que é considerado o valor mínimo de qualidade aceitável (EWING et al., 1998) (Tabela 3).

Score <i>Phred</i>	Probabilidade de chamada de base incorreta (P)	Chamada de base de precisão
10	1 em 10	90%
20	1 em 100	99%
30	1 em 1000	99,90%
40	1 em cada 10.000	99,99%
50	1 em 100.000	100,00%

Tabela 3: Índice de score *Phred*. O valor de *Phred* (Q) onde a probabilidade de erro de chamada de base é representada por (P) (EWING et al., 1998).

4.4. PROCESSAMENTO DAS *READS*

As *reads* passaram então por um processo de melhoria de qualidade, também conhecido como trimagem, pelo programa Trimmomatic por janela deslizante, uma ferramenta que realiza a remoção de trechos de bases e de adaptadores com baixa qualidade garantindo resultados flexíveis e eficientes (BOLGER; LOHSE; USADEL, 2014). A necessidade da realização da trimagem foi definida a partir do valor de qualidade (*Phred quality score*), em que o valor mínimo de qualidade aceitável foi definido em 30 e o tamanho das *reads* definidas a acima de 40 pb.

Após a trimagem, realizou-se uma segunda análise de qualidade das sequências com o programa FastQC, para visualização do resultado e confirmação de conclusão desta etapa.

4.5. MONTAGEM DAS *READS*

Uma vez que a análise da qualidade apresentou resultados adequados, realizou-se a montagem das *reads* sem a utilização de um genoma de referência (montagem *de novo*) com calda poli-a. O termo “sequenciamento *de novo*” deriva-se do latim e significa “desde o princípio”, pois refere-se aos métodos de determinação das sequências quando não há outras semelhantes disponíveis para uso como referência. Esta montagem foi realizada no servidor Sagarana do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais utilizando o Trinity (Figura 10), um programa baseado em três

algoritmos: *Inchworm*, que reúne o conjunto de dados em sequências únicas; *Chrysalis*, que infere os *clusters* e constrói grafos de *Brujin*. Estes grafos utilizam a abordagem em que as leituras são divididas em *k*-mers (*strings* de tamanho *k*), onde cada aresta do grafo corresponde a um *k*-mer e os nós do grafo representam *k*-mers = (*k*-1); e *Butterfly*, que analisa os grafos individualmente e reporta os transcritos resultantes a partir da posição das *reads*.

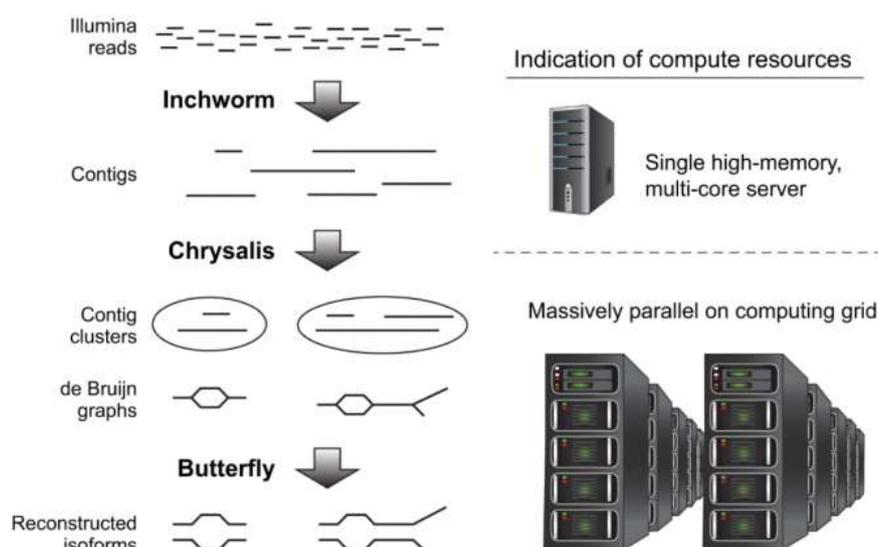


Figura 10: Visão geral da montagem e análise em Trinity. São apresentadas etapas sequenciais do Trinity (esquerda) e os recursos computacionais associados (direita) (Fonte: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3875132/pdf/nihms-537313.pdf>).

Durante a construção do grafo, as *reads* são transformadas em termos de tamanho fixo igual a *k* e são unidas quando possuem o mesmo valor de *k*-mer e a medida em que novas ocorrências de *k*-mers aparecem, a *read* tem seu peso incrementado. Para este projeto foram realizadas montagens com diferentes valores de *k*-mers (18 a 32) para definição do valor de *k*-mer mais adequado quando alinhado as metaloproteases através do programa CLC Genomics Workbench, onde também foram traduzidos.

4.6. SELEÇÃO DE SEQUÊNCIAS SIMILARES AS METALOPROTEASES

Após a montagem, os dados resultantes foram submetidos a um alinhamento de sequências utilizando a ferramenta BlastP no servidor Sagarana e no site do Centro Nacional de Informação Biotecnológica, NCBI (<http://www.ncbi.nlm.nih.gov>),

objetivando a localização de similaridade com metaloproteases. Para isto, utilizaram-se termos como parâmetros de busca: “*Loxosceles*” e “*metalloproteases*”, além de parâmetros de qualidade das sequências, como um tamanho mínimo de alinhamento de 250 aminoácidos (correspondendo ao tamanho médio de metaloproteases de *L. intermedia*), identidade igual ou superior a 50% e um valor de *e-value* máximo de 0,0001.

As sequências de melhores qualidades, que continham maior similaridade com as metaloproteases foram selecionadas e analisadas quanto a presença dos elementos essenciais de metaloproteases: Zinc-binding Signature (HEXXHXXGXXHE) e Met-turn (MXY). Além disso, para comparar as nove sequências em similaridade, identidade e porcentagem de gaps utilizou-se o programa Emboss Stretcher, que calcula um alinhamento global ideal de duas sequências usando programação de espaço linear.

4.7. CARACTERIZAÇÃO DAS METALOPROTEASES

Após identificar as prováveis sequências de metaloproteases da glândula de veneno de *L. laeta* (P) *in silico*, realizaram-se diversas análises de características destas sequências para melhor descreve-las estruturalmente e funcionalmente.

4.7.1. ALINHAMENTO DE SEQUÊNCIAS

Inicialmente realizou-se um alinhamento com as nove sequências de metaloproteases de *L. laeta* (LALP-LP – *Loxosceles astacin-like protease peruvian*) com as sequências disponíveis no banco de dados Uniprot (<http://www.uniprot.org/uniprot/>) das proteoformas de metaloproteases de *Loxosceles*: LALPs 1 a 5 e da astacina de *Astacus astacus* pelo programa Clustal Ômega (<https://www.ebi.ac.uk/Tools/msa/clustalo/>) que posteriormente foi analisado pelo programa Jalview (<http://www.jalview.org/>) para visualização e evidenciação de aminoácidos relevantes. Este possui diversas ferramentas que permitem evidenciar áreas importantes no alinhamento e avaliar por vários parâmetros informações relevantes, como por exemplo a predição de sítios de fosforilação, ordenação de alinhamentos e porcentagens de identidade. Além disso, ele disponibiliza inúmeras funções online como a predição de estruturas secundárias e análise de conservação dos aminoácidos (WATERHOUSE et al., 2009).

4.7.2. ANÁLISES FÍSICO-QUÍMICAS

Para a realização das análises físico-químicas, utilizou-se primeiramente o programa Expasy – ProtParam (<https://web.expasy.org/protparam/>), que permite a realização de uma análise de diversos parâmetros para cada sequência, como o número de aminoácidos, a massa molecular, o ponto isoelétrico, o número total de átomos e o N-terminal. Já para a separação e localização do peptídeo sinal das sequências utilizou-se o programa Expasy – SignalP (<http://www.cbs.dtu.dk/services/SignalP/>), baseado em redes neurais. E, através do programa NetNGlyc (<http://www.cbs.dtu.dk/services/NetNGlyc/>) realizou-se a predição dos sítios de N-glicosilação, que também utiliza redes neurais artificiais para examinar o contexto das sequências. Em seguida utilizou-se o banco de dados de domínios conservados (CDD) através do programa BlastP, em que foi possível realizar alinhamentos entre as sequências para a predição de domínios das metaloproteases, classificando-as em famílias utilizando modelos preditivos (assinaturas) e reduzindo a redundância dos resultados. Para prever os propeptídeos destas sequências utilizou-se o programa ProP (<http://www.cbs.dtu.dk/services/ProP/>), que prevê os locais de clivagem dos propeptídeos arginina e lisina em sequências de proteínas eucarióticas usando um conjunto de redes neurais.

4.7.3. PREDIÇÃO DE EPITOPOS

Utilizando redes neurais recorrentes (técnica de aprendizado de máquina), o servidor ABCPred (<http://crdd.osdd.net/raghava/abcpred/>) tem como objetivo prever os epitopos em uma sequência de aminoácidos utilizando padrões de comprimento fixo.

Uma rede neural artificial é um paradigma de processamento de informações inspirados na forma da estrutura paralela interconectada de informações do processo cerebral de mamíferos. Elas são coleções de modelos matemáticos e se baseiam em analogias de aprendizagem biológica adaptativa. Desta forma, o servidor ABCPred utiliza uma rede neural recorrente parcial, ou rede Jordan, com uma única camada oculta. As redes possuem comprimentos de janela opcionais de 10, 12, 14, 16, 18 e 20 e com 35 resíduos em uma camada oculta. Já a saída alvo consiste em um único número binário (1 para epitopo e 0 para não epitopo) (SAHA; RAGHAVA, 2006). Para a predição dos epitopos das sequências de LALP-LPs identificadas na análise transcriptômica deste

projeto foram utilizadas as opções padrões do sistema, sendo o limiar de 0,51 e a janela de previsão em 16. Além disso, foram utilizados filtros de sobreposição.

Outro programa também utilizado se denomina BepiPred-2.0: *Sequential B-Cell Epitope Predictor* (<http://www.cbs.dtu.dk/services/BepiPred/index.php>). Ele prediz epitopos de células B a partir de uma sequência proteica, utilizando um algoritmo de Floresta Aleatória, treinando epitopos e aminoácidos não-epitopos determinados a partir de estruturas cristalinas.

Foram escolhidas e analisadas em ambos os programas as três sequências mais expressivas (LALP-LP3, LALP-LP4 e LALP-LP8), e seus resultados comparados. Estas mesmas sequências foram posteriormente selecionadas para a produção de três membranas de celulose para mapeamento de epitopos pela técnica de SPOT.

4.7.4. FILOGENIA

A partir da utilização dos programas Jalview e Mega 7.0 também foi possível realizar a construção de árvores filogenéticas para representar a hipótese sobre a ancestralidade evolutiva das 15 sequências (LALPs 1 a 5, astacina de *Astacus astacus* e LALP-LPs 1 a 9), utilizando o modelo de máxima verossimelhança (maximum-likelihood), que a partir de um conjunto de dados em um modelo estatístico estima valores para os diferentes parâmetros do modelo com 500 réplicas (WATERHOUSE et al., 2009).

As árvores são calculadas com base em uma medida de similaridade entre cada par de sequências no alinhamento: Sendo PID a porcentagem de identidade entre as duas sequências em cada posição alinhada e BLOSUM62 a opção que usa a matriz de substituição para calcular uma soma de pontuações para os pares de resíduos em cada posição alinhada, utilizada para marcar alinhamentos entre sequências proteicas evolutivamente divergentes (WATERHOUSE et al., 2009).

4.7.5. MODELAGEM

A predição e modelagem de uma estrutura proteica exige a utilização de uma ferramenta de alta capacidade de detecção de homologias e construção de modelos 3D. Por este motivo, o programa Phyre2 (*Protein Fold Recognition Server*) (<http://www.sbg.bio.ic.ac.uk/phyre2/html/page.cgi?id=index>), foi escolhido para a modelagem das metaloproteases de *L. laeta* (P) identificadas neste projeto, utilizando como referência a Astacina de *Astacus astacus*. Este programa possui um conjunto de ferramentas que prevê e analisa as estruturas proteicas, utilizando uma interface simples e intuitiva. A partir da análise, é possível construir os modelos de proteínas em 3D, prever sítios de ligação e analisar o efeito de variantes de aminoácidos (KELLY et al., 2015).

Após a modelagem das metaloproteases pelo programa Phyre2, é importante analisar o resultado gerado no programa PyMOL (*The PhMOL Molecular Graphics System*, Versão 2.0 Schrodinger, LLC - <https://pymol.org/2/>), um sistema de visualização estrutural molecular, que permitirá a análise e sobreposição das sequências em conjunto.

4.8. VALIDAÇÃO DA PRESENÇA DE METALLOPROTEASES *IN VITRO*

4.8.1. QUANTIFICAÇÃO DAS PROTEÍNAS DOS VENENOS

Para realizar as próximas metodologias e validar os dados identificados *in silico*, foram realizadas as dosagens das concentrações totais das proteínas dos venenos de *L. laeta* brasileiras e peruanas pelo método de Lowry, descrito em 1951 por Oliver H. Lowry, utilizando uma curva padrão de BSA. Foram utilizadas as soluções do kit DC Protein Assay-BIORAD e o ensaio foi realizado em placa de 96 poços com duplicatas. Quatro poços foram definidos como branco para controle do ensaio, recebendo apenas as soluções do kit. As reações foram mensuradas em leitor de ELISA a 655nm e as absorbâncias resultantes foram adicionadas a um cálculo após realização de uma curva a partir das concentrações do padrão de BSA.

4.8.2. SDS-PAGE E WESTERN BLOTTING

Para analisar o conteúdo das proteínas nos venenos, foram realizadas separações em geis de poliacrilamida com 10 µg de veneno de *L. laeta* do Brasil e do Peru, em condições

não redutoras, além de utilizar-se o veneno de *L. intermedia* como controle. Um gel foi então corado com prata enquanto o outro foi utilizado para a realização de *western blotting*, em que se utilizou membrana de nitrocelulose com tampão de transferência em banho de gelo a 24V *overnight* e uma hora a 48V. A membrana foi então corada com Ponceau para certificação de transferência do gel para a membrana, e posteriormente lavada com água destilada e bloqueada com solução de bloqueio por uma hora e então lavadas com PBS-T 0,05% três vezes por cinco minutos. O soro anti recombinante de metaloprotease de *L. intermedia* de coelho foi diluído em PBS (1:400) e colocado em contato com a membrana por uma hora e trinta minutos e posteriormente foram realizadas mais três lavagens com PBS-T. O conjugado foi então diluído em PBS (1:2500) e colocado em contato com a membrana por uma hora e após esta fase foram realizadas outras três lavagens com PBS-T e depois lavada com PBS duas vezes por cinco minutos, e por fim a solução de revelação foi adicionada. Esta solução ficou em contato com a membrana por aproximadamente 15 minutos para revelação das bandas.

4.8.3. ENSAIOS ENZIMÁTICOS

Foram realizados ensaios *in vitro* para verificar a presença das atividades proteolíticas de metaloproteases no veneno de *L. laeta* (P) e avaliar as diferenças destas atividades frente ao veneno de *L. laeta* (B).

4.8.3.1. ATIVIDADES FIBRINOGENOLÍTICAS

Em 1998, Feitosa e colaboradores identificaram no veneno de *L. intermedia* uma proteoforma de metaloprotease, denominada Loxolisina A, que apresenta atividade fibrinogenolítica (FEITOSA et al., 1998). Por este motivo foi realizado o experimento que avalia a atividade de digestão do fibrinogênio pelos venenos de *L. laeta* brasileira e peruana.

De cada veneno (*L. laeta* (B) e *L. laeta* (P)) foram adicionados separadamente 3 µg em 50 µL de solução de fibrinogênio bovino (25mM TRIS-HCl + 0,15M NaCl com pH de 7.4), completou-se o volume de 100 µL com PBS e incubou a 37°C por 16 horas. Ao fim deste tempo uma alíquota de 50 µL da amostra foi retirada e adicionou-se 50 µL de solução desnaturante (10M ureia, 4%β-mercaptoethanol, 4% SDS) em temperatura ambiente por mais 16 horas. Por fim uma alíquota de 20 µL da amostra foi acrescida de

20 µL de tampão em condições redutoras após ser fervido por 5 minutos e 35 µL dessa solução final foi aplicada em um gel SDS-PAGE 10%. O gel ao final foi corado com comassie blue para visualização das bandas.

Um segundo ensaio de atividade fibrinogenolítica (tempo-dependente) foi realizado utilizando as mesmas amostras e condições descritas anteriormente, porém foram feitas soluções com um volume final três vezes maior de fibrinogênio bovino com o veneno de *L. laeta* (B) e igualmente para o veneno de *L. laeta* (P), para que fossem retiradas alíquotas de 25 µL nas seguintes horas de tempos de incubação: 1, 2, 4, 6, 8, 10, 12 e 16 horas. Cada alíquota recebeu 25 µL de solução desnaturante e aplicados 35 µL de cada em dois géis SDS-PAGE 10%, além de acrescentar ao final uma banda com EDTA.

4.8.3.2. ENSAIO DE GELATINASE/COLAGENASE

No mesmo trabalho em que Feitosa e colaboradores em 1998 avaliaram a atividade fibrinogenolítica, mostraram que uma segunda proteoforma de metaloprotease denominada Loxolisina B no veneno de *L. intermedia* possui atividade gelatinolítica, em gel de SDS-PAGE (FEITOSA et al., 1998). Por esta razão realizou-se este experimento para os venenos de *L. laeta* brasileira e peruana a partir do kit *EnzChek Gelatinase/Colagenase*, mensurando por fluorescência a capacidade dos venenos em degradar gelatina. Utilizou-se para este experimento a metodologia disponibilizada no manual do kit, onde o controle positivo de colagenase é composto por enzima colagenase isolada de *Clostridium histolyticum*. Foram utilizadas concentrações diferentes dos venenos: 0,37 µg, 0,75 µg, 1,5 µg, 3 µg, 6 µg e 12 µg, cada ensaio foi feito em duplicata e uma replicata do experimento foi realizada, totalizando quatro replicatas.

4.8.4. MAPEAMENTO DE EPITOPOS PELO MÉTODO DE SPOT

A síntese das membranas de celulose foi realizada pelo Prof. Ricardo Andrez Machado de Ávila na Universidade Federal do Paraná (UFPR), através do aparelho de síntese automática ResPep SL (IntavisAG, Bioanalytical Instruments, Alemanha), onde pentadecapeptídeos sobrepostos por três resíduos cobrindo as três principais sequências de metaloproteases da glândula de veneno de *L. laeta* (P) identificadas pela análise transcriptômica realizada neste trabalho (LALP-LP3, LALP-LP4 e LALP-LP8) foram sintetizadas seguindo o protocolo descrito por Laune e colaboradores em 2002,

depositando um volume de 0,6 µl de cada aminoácido, permitindo a obtenção de aproximadamente 50 nanomoles de peptídeo por ponto de membrana.

A membrana de celulose é transformada para disponibilizar grupamentos amino para o acoplamento de aminoácidos pela esterificação de uma Fluorenilmetiloxycarbonil (FMOC)-βAla-OH às funções hidroxila disponíveis na celulose, permitindo o suporte se tornar funcional e afasta o peptídeo do suporte, garantindo maior mobilidade e manutenção de sua conformação.

Para avaliar a reatividade de anticorpos frente as membranas de SPOT, foi utilizado soro anti-metaloprotease de *L. intermedia* (LALP-1) gentilmente cedido pelo Prof. Silvio Sanchez Veiga (UFPR). As membranas foram lavadas com solução fosfato salina (PBS pH 7,4) e bloqueadas com TBS 3% e Tween 0,1% *overnight*. Elas foram então posteriormente lavadas e incubadas com o soro por 120 minutos (diluição 1:500) e novamente lavadas com TBS 3% e Tween 0,1% e a ligação dos anticorpos foi detectada utilizando anticorpos anti-IgG de coelho secundários ligados a fosfatase (diluição 1:2.000/SIGMA). Após a adição do substrato BCIP-MTT em 0,4% de MgCl₂ em tampão citrato salina (CBS pH 7,0) as regiões reativas das membranas foram visualizadas, ao formarem precipitados azulados sobre os spots que contem peptídeos que foram reconhecidos pelos anticorpos do soro testado durante 30 minutos, sob agitação. As membranas foram então escaneadas para serem posteriormente analisadas utilizando o programa ImageJ. As imagens escaneadas das membranas foram modificadas para escala de cores em cinza pelo programa para a quantificação de pixels de cada SPOT.

Através da técnica de SPOT, Lima e colaboradores em 2018 conseguiram identificar as proteínas mais reativas do veneno de *Loxosceles* ao serem reconhecidos pelo soro anti rMEPLox produzido em coelhos e produziram uma proteína multiepitópica quimérica (rMEPLox) contendo Hialuronidase, SmaseD e Metaloprotease. Esta proteína conferiu proteção à dermonecrose a um animal vacinado causada pela mordida da aranha *L. intermedia*, além de não ser tóxica nem letal a este animal (LIMA et al., 2018). Na busca de correlacionar os epitopos identificados neste trabalho com os presentes em rMEPLox e também encontrar diferenças importantes entre eles realizou-se a comparação dos epitopos correspondentes à metaloproteases de rMEPLox com os epitopos reativos identificados no experimento anterior (soro anti-metaloprotease de *L. intermedia*) através de um segundo SPOT utilizando as três membranas de metaloproteases do veneno de *L. laeta* (P) produzidas neste trabalho com o soro anti rMEPLox (diluição 1:100) e também por uma comparação direta dos aminoácidos por alinhamento utilizando o programa

Jalview e uma modelagem das estruturas tridimensionais marcados os epitopos pelo programa PyMOL.

Como controle deste experimento, o mesmo ensaio foi realizado nas três membranas com soro de coelhos que receberam apenas adjuvante de PBS (soro pré-imune) (diluição 1:100).

Para serem reutilizadas em outros ensaios, as membranas foram submetidas ao tratamento de regeneração após cada experimento, para remoção dos complexos moleculares precipitados sobre os peptídeos. Inicialmente são realizadas lavagens com mercaptoetanol), reagente B (etanol/água/ácido acético nas proporções 50:40:10 vol/vol/vol) e ao fim com metanol, para secagem da membrana.

5. RESULTADOS E DISCUSSÃO

5.1. ANÁLISE DA QUALIDADE DAS SEQUÊNCIAS

Os arquivos disponibilizados pelo sequenciamento foram denominados “L_laeta_S1_L001_R1_001.fastq” e “L_laeta_S1_L001_R2_001.fastq”, sendo R2 o resultado do sequenciamento reverso. Foram obtidas 11.870.865 *reads*, com comprimento das seqüências variando entre 35 a 301 bases e com um conteúdo de GC de 41% ao serem analisados pelo programa FastQC (Figura 11).

Modica e colaboradores em 2015 realizam a análise transcriptômica de metaloproteases do caracol *Columbraria reticulata*

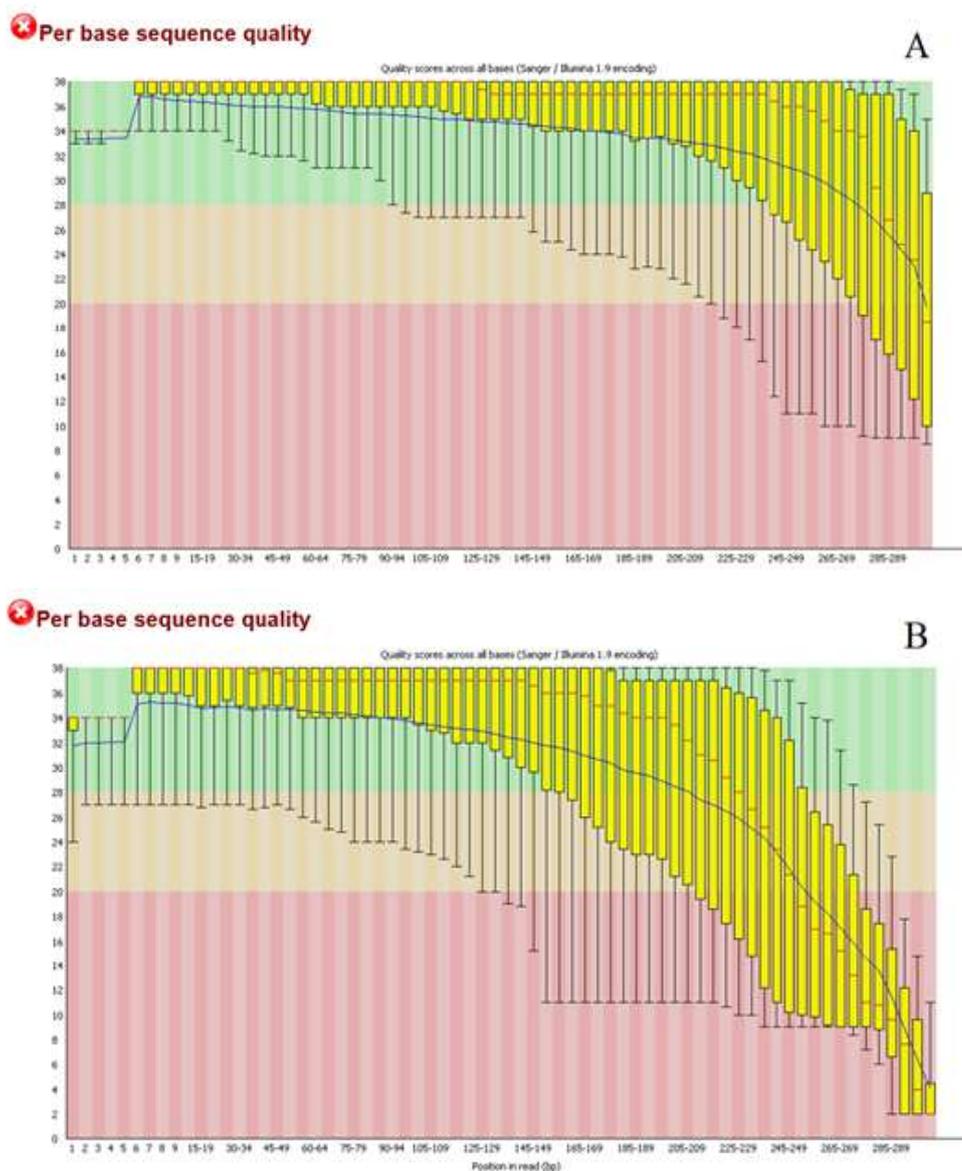


Figura 11: Valores de qualidade Q de Phred por base de reads brutas em (A) R1 e (B) R2. Análise das reads brutas analisadas pelo programa FastQC, onde apresentam-se os valores médios de qualidade das bases. (A) L-laeta_S1_L001_R1_001.fastq. (B) L-laeta_S1_L001_R2_001.fastq.

Os gráficos emitidos pelo programa FastQC apresentam resultados importantes a serem analisados, onde o eixo x representa cada base do sequenciamento e o eixo y, o índice de qualidade. As colunas amarelas verticais correspondem à abrangência de interquartil dos valores de qualidade para cada base e as linhas em preto acima e abaixo dessas colunas correspondem a *outliers*. Dentro de cada barra amarela observa-se uma linha em azul, que indica a flutuação do valor médio da qualidade e uma linha em vermelho, que se relaciona ao valor mediano de qualidade de cada base.

Como comumente em sequenciamentos realizados na plataforma Illumina, a qualidade dos resultados demonstrou a necessidade de melhoria das sequências, pois foi observada uma queda de qualidade das *reads* na porção final e, além disso, o programa FastQC não consegue identificar quais as possíveis fontes contaminantes em um sequenciamento *de novo* (ENDRULLAT et al., 2016; ZHOU; SU; NING, 2014).

5.2. PROCESSAMENTO DAS READS

Vista a necessidade de melhoria das sequências, foi realizado o processo de trimagem utilizando o programa Trimmomatic, sendo obtidas 10.289.208 *reads* no arquivo R1 e 10.203.328 *reads* no arquivo R2, com qualidade média acima de 40:20 (Tamanho:*Phred*). Este resultado foi então novamente analisado utilizando o programa FastQC, onde observou-se o valor de qualidade final *Phred* de no máximo 38, garantindo uma probabilidade de chamada de base correta maior que 99,90% (Figura 12).

Apesar de em outros projetos comumente existir a necessidade de filtragens subsequentes devida a flutuação de valor de qualidade das últimas bases apresentarem um valor a baixo de 20, como foi o caso de Aquino e colaboradores em 2016, neste projeto não houve a necessidade da realização deste passo, pois a qualidade de todas as bases se localizaram dentro do limite de *Phred* aceitável (AQUINO; MARTINS; CHÁVEZ-OLORTEGUI, 2016).

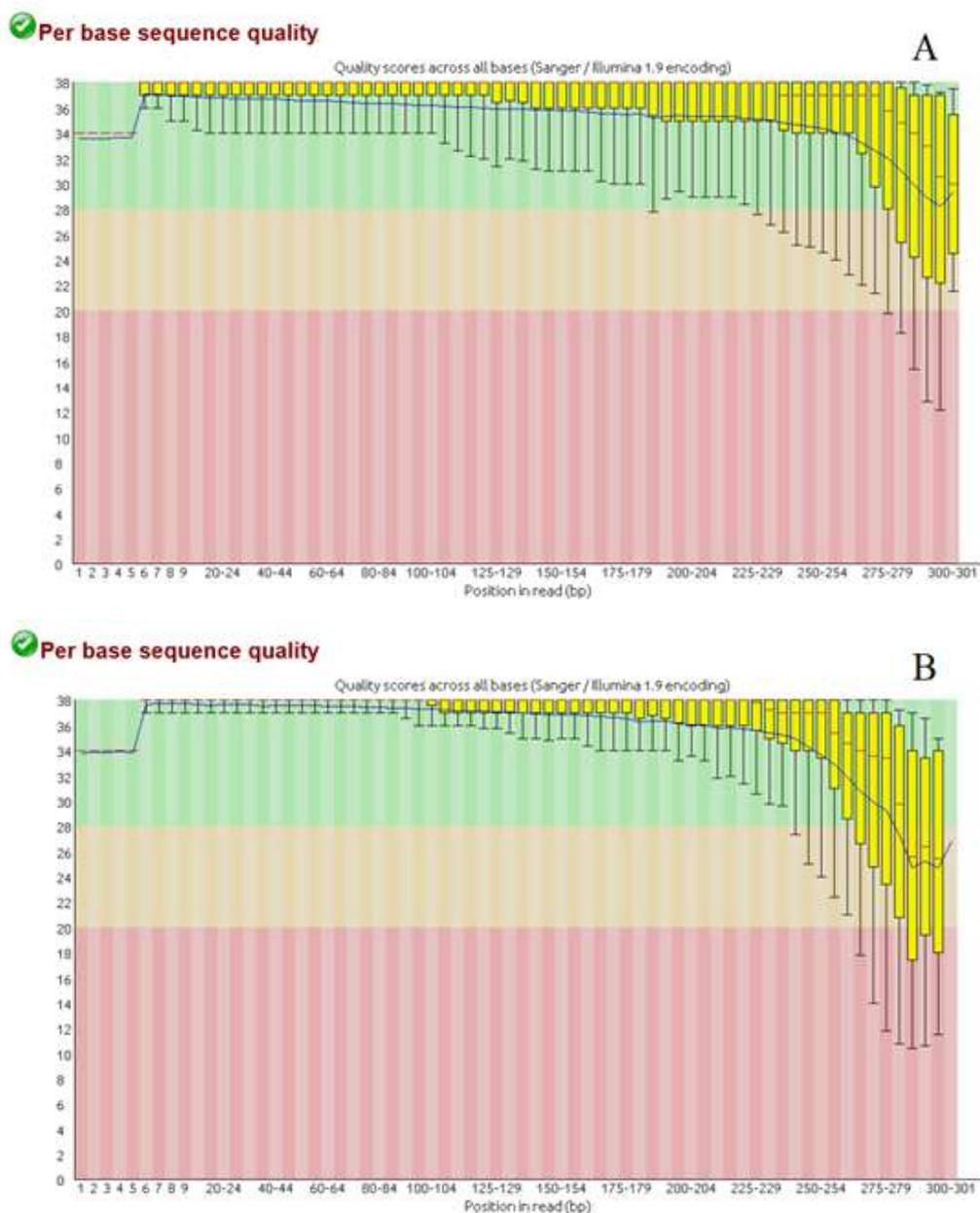


Figura 12: Valores de qualidade Q de Phred por base de reads após trimagem em (A) R1 e (B) R2. Análise das reads após trimagem analisadas pelo programa FastQC, onde apresentam-se os valores médios de qualidade das bases. (A) L-laeta_S1_L001_R1_001.fastq. (B) L-laeta_S1_L001_R2_001.fastq.

5.3. MONTAGEM DAS *READS*

Para escolha do método adequado na realização de uma montagem de transcriptoma deve-se considerar a existência de um genoma de referência da espécie. Neste trabalho há ausência de uma referência, logo foi necessária a montagem do transcriptoma *de novo* (GRABHERR, M. G. et al., 2011). Foram realizadas 10 montagens e dentre elas, o

programa CLC Genomics Workbench selecionou a de k-mer 22 como de melhor correspondência e de maior expressividade, apresentando um total de 143.237 contigs.

Não existe uma fórmula explícita para a escolha do melhor k-mer, porém é possível calcular alguns limites com base no tamanho e cobertura estimados. Apesar disso, tais estimativas não levam em consideração o impacto de uma repetição ou taxa de erro de leituras, por isso o k-mer é comumente escolhido baseando-se em experiências anteriores com conjuntos de dados semelhantes ao analisado, ou utilizando histogramas de abundância para valores putativos de k-mer e compara-los (KELLY et al., 2015; MARÇAIS; KINGSFORD, 2011). A escolha do tamanho k-mer tem resultados distintos das sequências, logo é importante escolher a dimensão apropriada dos k-mers para criar um equilíbrio nas montagens (ZERBINO, D.R.; BIRNEY, 2008).

5.4. SELEÇÃO DE SEQUÊNCIAS SIMILARES AS METALOPROTEASES

A partir do alinhamento utilizando a ferramenta BlastP foram selecionadas 10 sequências que possuíam os parâmetros utilizados nas buscas, possuindo os maiores valores de identidade com LALP1 (77% - 84%) e menores *e-values*.

A identidade corresponde a extensão em que duas sequências (nucleotídeo ou aminoácido) possuam os mesmos resíduos nas mesmas posições em um alinhamento, geralmente expresso com uma porcentagem. Já o *e-value*, ou valor esperado (E), é um parâmetro que descreve o número de ocorrências ao acaso que se pode esperar ao pesquisar um banco de dados de um tamanho específico. Quanto menor o valor de E, ou quanto mais próximo de zero, mais significativa é a correspondência, mas apesar disso é importante considerar que alinhamentos curtos virtualmente idênticos possuem valores de E relativamente altos devido ao comprimento da sequência. Além disso, é possível utilizar o *e-value* para criação de um limite de significância para os resultados (OSTELL, 2002).

Estas sequências foram então alinhadas pelo programa Jalview para localização dos elementos essenciais das metaloproteases: Zinc-binding Signature (HEXXHXXGXXHE) e Met-turn (MXY), e uma das sequências foi rejeitada e excluída pois não possuía o elemento essencial Met-turn, restando ao final nove sequências de possíveis proteoformas de metaloproteases do veneno de *L. laeta* (P) (Tabela 4). Além disso, as similaridades entre as nove sequências foram definidas pelo programa Emboss Stretcher (Figura 13).

SEQUÊNCIA	IDENTIDADE	E-VALUE
LALP-LP1	77%	8e-113
LALP-LP2	75%	3e-111
LALP-LP3	81%	8e-121
LALP-LP4	75%	2e-112
LALP-LP5	75%	7e-112
LALP-LP6	74%	2e-111
LALP-LP7	77%	1e-115
LALP-LP8	73%	1e-105
LALP-LP9	75%	2e-112

Tabela 4: Sequências de possíveis metaloproteases de *L. laeta* (P) selecionadas a partir de BlastP e alinhadas por Jalview em relação à LALP1.

Similarity between LALP-LPs - Emboss Stretcher

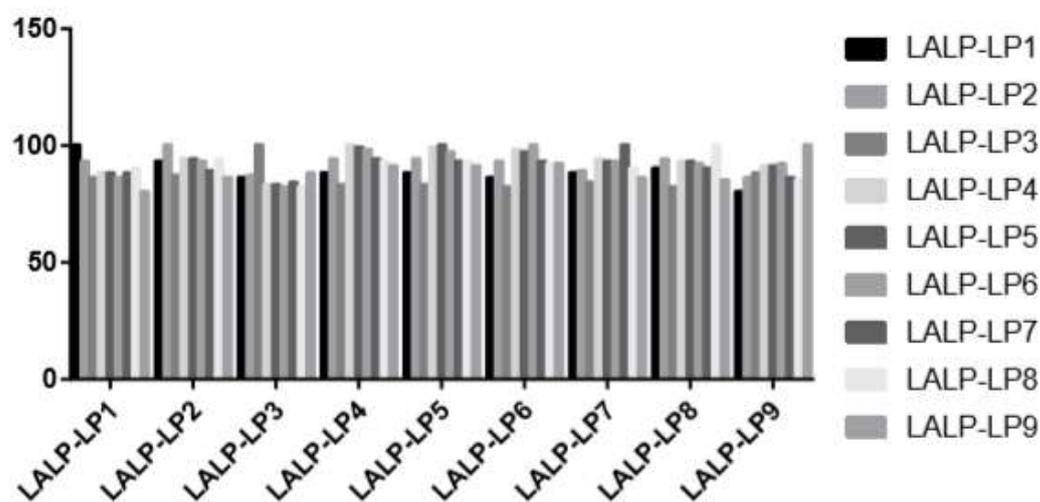


Figura 13: Similaridade entre nove LALP-LPs pelo programa Emboss Stretcher.

5.5. CARACTERIZAÇÃO DAS METALOPROTEASES

5.5.1. ALINHAMENTO DE SEQUÊNCIAS

O alinhamento realizado pelo Clustal Ômega foi analisado pelo Jalview para marcação de aminoácidos relevantes (Figura 14), onde é possível observar a presença de peptídeos sinais em todas as sequências, com exclusão da LALP4 que por ser um fragmento não possui peptídeo sinal. Identificou-se também a presença de propeptídeos em *Astacus astacus* e LALPs de *L. intermedia*, e a partir do programa ProP para as nove LALP-LPs é possível perceber que as mesmas possuem propeptídeos semelhantes as LALPs 1 a 3. A localização das pontes dissulfeto em *Astacus astacus* e nas LALPs que permite prevê-las nas LALP-LPs por semelhança. A presença dos elementos essenciais Metal binding e Met-turn, além do sítio ativo em todas as 15 sequências e os sítios de N-glicosilação de acordo com a localização identificada no NetNGlyc.

A partir do alinhamento pode-se observar um *gap* repetitivo em todas as LALP-LPs, com exceção de LALP-LP1, localizado aproximadamente a partir do aminoácido de número 70, marcado em vermelho na Figura 14. Isso pode significar uma possível diferença que se relaciona com o elevado número de casos fatais causados por *L. laeta* no Peru quando comparado aos outros países que possuem casos de Loxoscelismo por *L. laeta*.

5.5.2. ANÁLISES FÍSICO-QUÍMICAS

A análise realizada pelo programa ExPASy – ProtParam permitiu uma comparação entre as nove sequências de metaloproteases de *L. laeta* (P) identificadas neste projeto com as metaloprotease de *L. intermedia* (LALP1) e *L. laeta* (B) (LALP4). É possível observar que as nove sequências de LALP-LPs possuem resultados físico-químicos semelhantes, com um número de aminoácidos entre 184 e 196, uma massa molecular média de 21667,69, pontos isoeletrônicos próximos a 5 e número de átomos próximos aos 3000, enquanto a LALP1 possui resultados mais elevados por possuir uma sequência de maior tamanho (264 aminoácidos), e a LALP4, por ser um fragmento, possui os menores resultados analisados (114 aminoácidos) (Tabela 5).

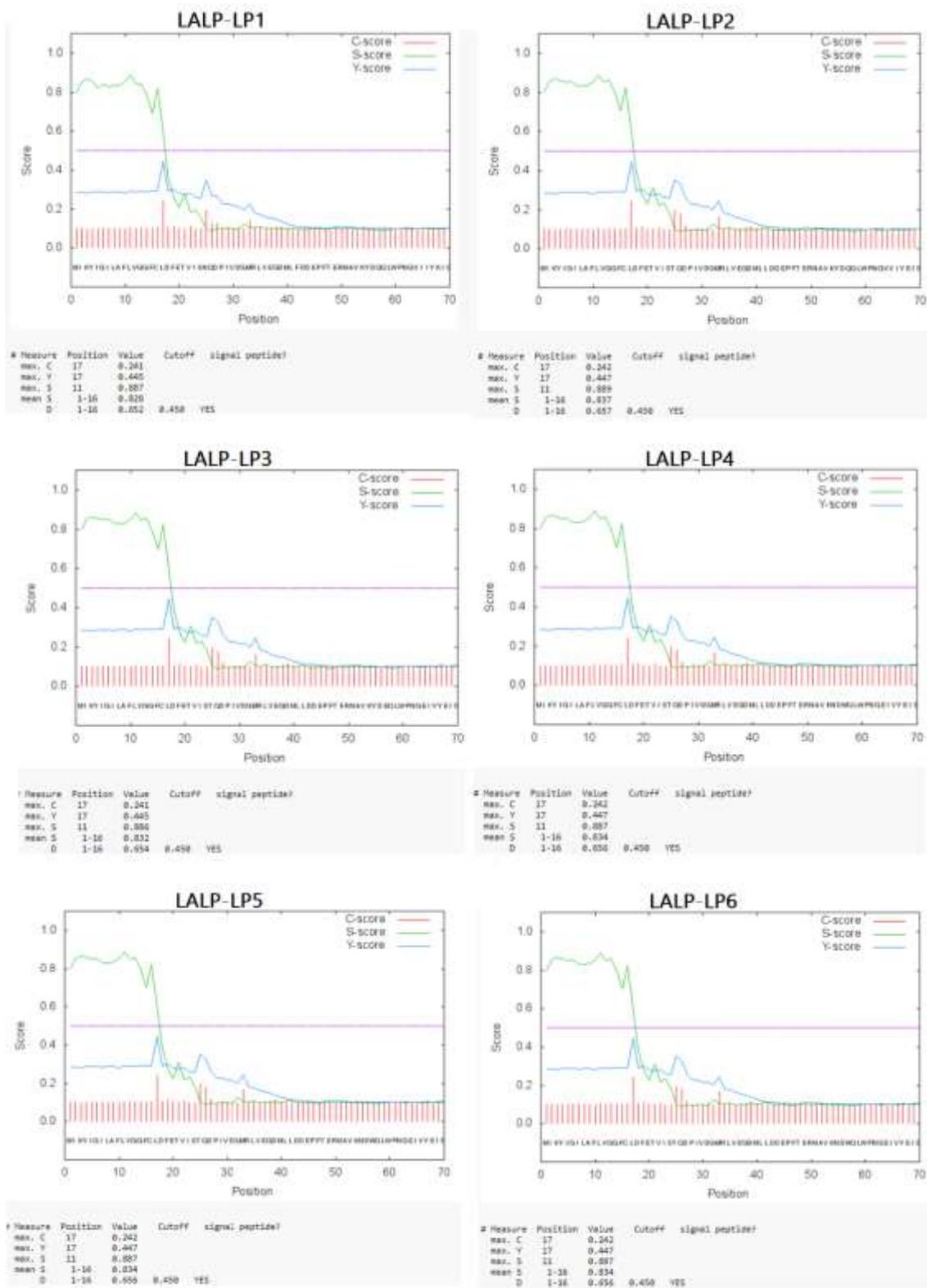
Desta forma, é importante observar que as LALP-LPs terão provavelmente maiores semelhanças com a LALP1 do que com a LALP4 nos resultados que se seguem, pois, uma sequência que possui apenas um fragmento disponível nos bancos de dados mundiais perde muitas informações relevantes a proteína e seu conteúdo representará um resultado de baixa identidade quando comparada com outras sequências semelhantes.

SEQÜENCIA	Nº DE AA	MASSA MOLECULAR	PONTO ISOELÉTRICO	Nº DE ÁTOMOS	N-TERMINAL
LALP-LP1	189	21622.43	5.35	3008	M (Met)
LALP-LP2	184	21053.80	5.12	2928	M (Met)
LALP-LP3	193	22342.11	4.91	3091	M (Met)
LALP-LP4	189	21685.45	5.14	3007	M (Met)
LALP-LP5	189	21597.33	5.14	2996	M (Met)
LALP-LP6	188	21520.29	5.05	2988	M (Met)
LALP-LP7	196	22466.52	5.81	3132	M (Met)
LALP-LP8	185	21065.09	5.87	2951	M (Met)
LALP-LP9	188	21656.27	4.87	2990	M (Met)
LALP1	264	30384.50	6.25	4224	M (Met)
LALP4	114	13048.48	5.62	1794	R (Arg)

Tabela 5: Análise das seqüências de metaloproteases de *L. laeta* (P), *L. intermedia* (LALP1) e *L. laeta* (B) (LALP4) pelo Expasy - Protparam.

A identificação de todos os componentes das nove seqüências por análise transcriptômica e compara-los com os das LALPs já descritos nos bancos de dados mundiais disponibilizaram informações que confirmam essas seqüências como proteoformas de metaloproteases. Trevisan-Silva e colaboradores em 2013 identificaram em spots resultantes de atividades gelatinolíticas com metaloproteases de *L. laeta* valores de massa molecular de aproximadamente entre 20 e 30kDa e de ponto isoelétrico entre 6.00 e 9.00, ambos valores próximos aos encontrados neste trabalho (TREVISAN-SILVA et al., 2013).

Ao utilizar o programa Expasy – SignalP, identificou-se os peptídeos sinais das mesmas seqüências analisadas na Tabela 5. Observa-se que as nove LALP-LPs possuem os mesmos aminoácidos definidos como peptídeos sinais que LALP1, além de apresentarem a mesma localização nas seqüências (do primeiro ao décimo sexto aminoácido). Já a LALP4, por ser um fragmento, não possui peptídeo sinal identificável, pois o mesmo provavelmente se encontraria em uma área de aminoácidos que não foi disponibilizada nos bancos de dados (Figura 15). Estes peptídeos sinais podem ser identificados no alinhamento das seqüências na Figura 13 coloridos em laranja.



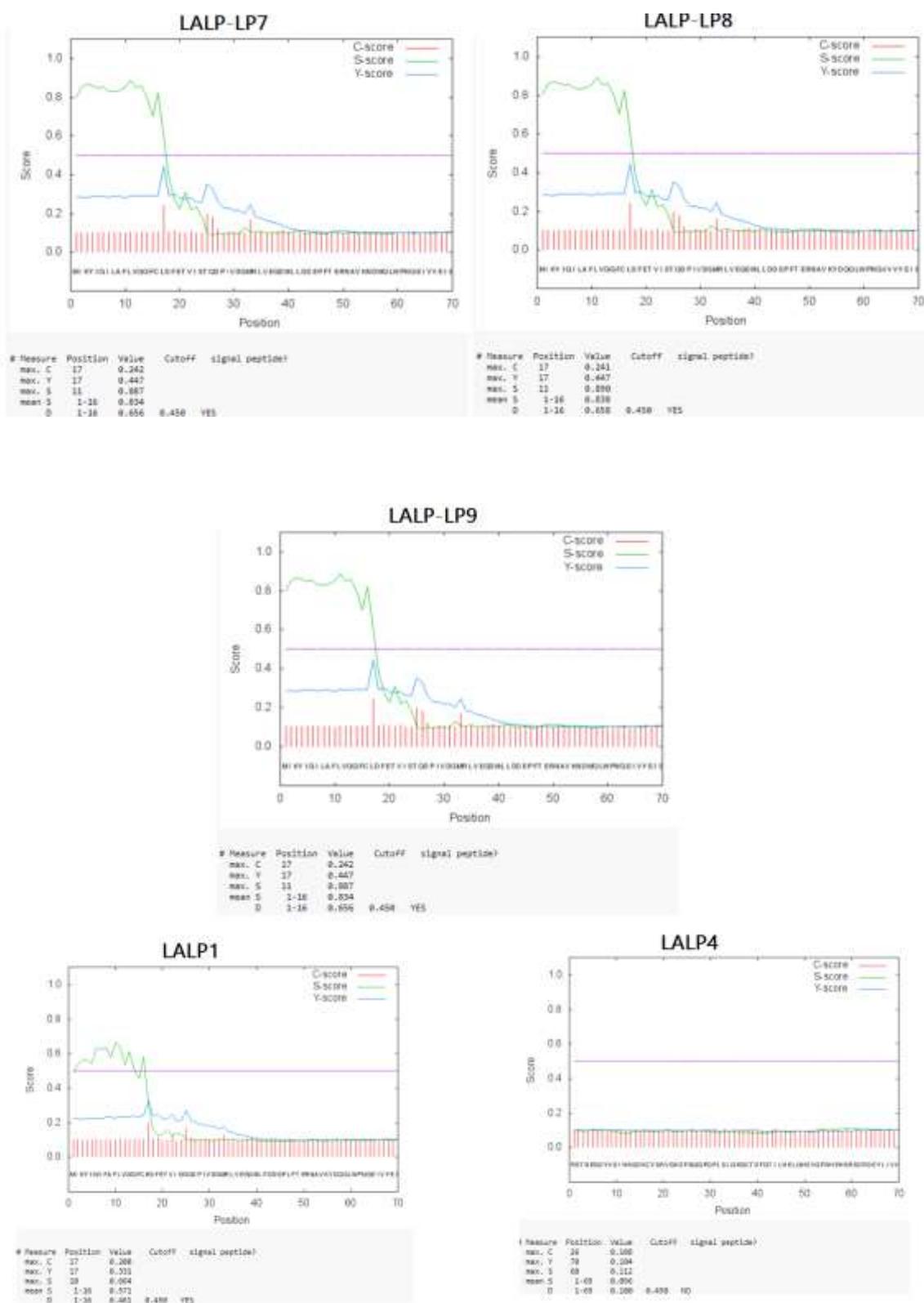


Figura 15: Análise das sequências de metaloproteases de *L. laeta* (P) LALP-LPs, pelo Expsy - SignalP.

A glicosilação é considerada uma das modificações pós-traducionais mais complexas encontradas em proteínas. As cadeias glicídicas podem afetar a estabilidade,

atividade e enovelamento biológico, além de possuírem um papel importante em processos biológicos fundamentais, como a transdução de sinais, adesão celular e endocitose (SPIRO, 2002). Os carboidratos podem ser ancorados à cadeia polipeptídica de várias formas, como por uma ligação amida em cadeiras laterais de asparagina (N-glicosilação), sendo o mais frequente em glicoproteínas.

Através do programa NetNGlyc foi possível predizer os sítios de N-glicosilação das nove sequências LALP-LPs (Figura 16). Ela identifica a presença de sequências consenso de N-glicosilação em estruturas primárias e utiliza redes neurais para analisar a possibilidade de glicosilação na proteína, marcando-os em vermelho. Desta forma foi possível observar que todas as sequências analisadas possuíam grande proximidade nos resultados de identificação destes sítios, com exclusão das sequências LALP-LP3 e LALP-LP9. Estes sítios de N-Glicosilação podem ser identificados no alinhamento das sequências na Figura 13 coloridos em cinza (Tabela 6).

LALP-LP1

```
MIKYIGILAFVLVGGFCLDFETVISNQOPIVDGMRLVEGDMLFDDEPFTERNAVKYDQQLWPNGVIIYKISQALRQYAQRI 80
REAMKSNEDENDYVDIHRGPKCYSRVGKSKGGPQPLSLGKGCTEFGTILHELGH5VGF5SHEHSRIDRDEFLIVHKENILN 160
GSQHDFKKKLENETRTIGDFDYDSIMLYG
..... 80
.....N..... 160
.....N..... 240
```

LALP-LP2

```
MIKYIGILAFVLVGGFCLDFETVISTQDPIVDGMRLVEGDMLLDDEPFTERNAVKYDQQLWPNGVVIYEISQSLRPFKQRI 80
REAMRSYEDIHKGDKCYSRVGSFQGGPQPLSLGKGCTEFGTILHELGH5VGF5SHEHSRVDRDEFLIVHEENILNGSQHD 160
FKKKLENETRTIGDFDYDSIMLYG
..... 80
.....N..... 160
.....N..... 240
```

LALP-LP3

```
MIKYIGILAFVLVGGFCLDFETVISTQDPIVDGMRLVEGDMLLDDEPFTERNAVKYDEQLWPNGEIVYEISQALRPFKQRI 80
REAMRSYEDTTCIEKDYVNIHKGDRCYSRVGSFNGGEQPLSLATGCTEFGTILHELGH5VGF5DHEHSRSDRDEYLIHK 160
ENNVATKEHNFDKLENETRTIGDFDYDSIMLYG
..... 80
.....N..... 160
.....N..... 240
```

LALP-LP4

MIKYIGILAFVLVGGFCLDFETVISTQDPIVDGMRLVEGDMLLDDEPFTERNAVKNWQLWPNGEIVYEISQALRPFKQRI	80
REAMRSYEDHKGDCKYSRVGKSFQGGPQPLSLGKGCTEFGTILHELGHVSVGFSHEHSRVDRDEFLLVHEENILNGSQHDF	160
KKKLENETRTIGDFDYDSIMLYGPIYAFSR	
.....	80
.....N.....	160
.....N.....	240

LALP-LP5

MIKYIGILAFVLVGGFCLDFETVISTQDPIVDGMRLVEGDMLLDDEPFTERNAVKNWQLWPNGEIVYEISQALRPFKQRI	80
REAMRSYEDHKGDCKYSRVGKSFQGGPQPLSLGKGCTEFGTILHELGHVSVGFSHEHSRVDRDEFLLVHEENILNGSQHDF	160
KKKLENETRTIGDFDYDSIMLYGPIYASSK	
.....	80
.....N.....	160
.....N.....	240

LALP-LP6

MIKYIGILAFVLVGGFCLDFETVISTQDPIVDGMRLVEGDMLLDDEPFTERNAVKNWQLWPNGEIVYEISQALRPFKQRI	80
REAMRSYEEKGDCKYSRVGKSFQGGPQPLSLGKGCTEFGTILHELGHVSVGFSHEHSRVDRDEFLLVHEENLLNGSQHDFKK	160
KLENETRTIGDFDYDSIMLYGPIYAFSKD	
.....	80
.....N.....	160
.....N.....	240

LALP-LP7

MIKYIGILAFVLVGGFCLDFETVISTQDPIVDGMRLVEGDMLLDDEPFTERNAVKNWQLWPNGEIVYEISQALRPVKQRI	80
REAMRSYEDTTCIKFRHRGPKCYSRVGKSKGGPQPLSLGKGCTEYGTILHELGHVSVGFSHEHSRIDRDEFLLVHKENILN	160
GSQHDFKKKLENETRTIGDFDYDSIMLYGPIYAFSKD	
.....	80
.....N.....	160
.....N.....	240

LALP-LP8

MIKYIGILAFVLVGGFCLDFETVISTQDPIVDGMRLVEGDMLLDDEPFTERNAVKNWQLWPNGVYVEISQALRPFKQRI	80
REAMRSVDIHRGPKCYSRVGKSKGGPQPLSLGKGCTEFGTILHELGHVSVGFSHEHSRIDRDEFLLVHKENILNGSPDFKK	160
KLENETRTIGYFDYDSIMLYGPIYAF	
.....	80
.....N.....	160
.....N.....	240

LALP-LP9

MIKYIGILAFVLVGGFCLDFETVISTQDPIVDGMRLVEGDMLLDDEPFTERNAVKNWQLWPNGEIVYEISQALRPFKQRI	80
REAMRSYEQGDRCYSRVGSFNNGEQPLSLATGCTEFGTILHELGHVSVGFDHEHSRSDRDEYLIHKENVVATKEHNFDK	160
LWENRTIGNFDYDSIMLYGPIYAFSTD	
.....	80
.....N.....	160
.....N.....	240

Figura 16: Predição dos sítios de N-glicosilação de metaloproteases de *L. laeta* (P) (LALP-LPs) pelo NetNGlyc.

LALP-LPs	N-Glyc	Score
LALP-LP1	NGSQ	0.5919
LALP-LP2	NGSQ	0.5953
LALP-LP3	NRTR	0.6797
LALP-LP4	NGSQ	0.6002
LALP-LP5	NGSQ	0.6002
LALP-LP6	NGSQ	0.6073
LALP-LP7	NGSQ	0.5974
LALP-LP8	NGSP	0.1313
LALP-LP9	NRTR	0.6863

Tabela 6: Predição dos sítios de N-glicosilação de metaloproteases de *L. laeta* (P) (LALP-LPs) e score pelo NetNGlyc.

Veiga e colaboradores em 1999 identificaram a influência da glicosilação na atividade dos componentes do veneno de *L. intermedia*, levando a diminuição sensivelmente da sua atividade. Em todas as nove proteoformas de metaloproteases identificadas neste trabalho foram encontrados sítios de N-Glicosilação, o que pode colaborar com a informação de que estas proteínas auxiliam na dermonecrose (VEIGA et al., 1999).

A partir do programa BlastP foram feitas as predições de domínios das nove sequências LALP-LP, além das LALP1 e LALP4. Observa-se que as LALP-LPs possuem domínios idênticos aos presentes em LALP1 e LALP4, e que as LALP-LP3 e LALP-LP9 possuem uma adição dos domínios cd0479 e cd04277, presente apenas em LALP4 (Tabela 7), que são zinco dependentes, presentes na subfamília serralisina, importantes em mecanismos de translocação e na ligação de substratos.

Sequência	Domínio	Sequência	Domínio	Sequência	Domínio	Sequência	Domínio	Sequência	Domínio	Sequência	Domínio
LALP-LP1	pfam01400	LALP-LP2	pfam01400	LALP-LP3	cd04280	LALP-LP4	pfam01400	LALP-LP5	cd04280	LALP-LP6	cd04280
	cd04280		cd04280		pfam01400		pfam01400				
	cd04281		cd04281		cd04283		cd04281				
	cd04283		cd04283		cd04281		cd04283				
	cd04282		smart00235		smart00235		cd04282				
	smart00235		cd04282		cd04282		smart00235				
LALP-LP7	cd04282	LALP-LP8	cd04282	LALP-LP9	cd04282	LALP1	cd04282	LALP4 (FRAGM.)	cd04282		cd04282
	pfam01400		pfam01400		pfam01400		pfam01400				
	cd04281		cd04281		cd04283		cd04283				
	cd04283		cd04283		cd04281		cd04281				
	smart00235		smart00235		cd04282		smart00235				
	cd04268		cd04268		smart00235		cd04268				
cd00203	cd00203	cd00203	cd00203								
cd04327	cd04327	cd04268	cd04327								
cd04268	cd04268	cd04327	cd04268								
		cd04279	cd04279								
		cd04277	cd04277								

Tabela 7: Predição de domínios de metaloproteases de *L. laeta* (P), *L. intermedia* (LALP1) e *L. laeta* (B) (LALP4) pelo BlastP.

Com a utilização do programa ProP, os propeptídeos das sequências de metaloproteases de *L. laeta* (P) foram identificados e ao serem comparados com os propeptídeos de LALP1 observou-se uma identidade praticamente absoluta entre eles, o que corrobora a veracidade deste resultado (Tabela 8). Estes propeptídeos podem ser identificados no alinhamento das sequências na Figura 14 coloridos em rosa.

SEQUÊNCIA	PROPEPTÍDEO
LALP-LP1	D-F-E-T-V-I-S-N-Q-D-P-I-VG-M-R-L-V-E-G-D-M-L-F-D-D-E-P-F-T-E-R
LALP-LP2	D-F-E-T-V-I-S-Q-D-P-I-VG-M-R-L-V-E-G-D-M-L-L-D-D-E-P-F-T-E-R
LALP-LP3	D-F-E-T-V-I-S-Q-D-P-I-VG-M-R-L-V-E-G-D-M-L-L-D-D-E-P-F-T-E-R
LALP-LP4	D-F-E-T-V-I-S-Q-D-P-I-VG-M-R-L-V-E-G-D-M-L-L-D-D-E-P-F-T-E-R
LALP-LP5	D-F-E-T-V-I-S-Q-D-P-I-VG-M-R-L-V-E-G-D-M-L-L-D-D-E-P-F-T-E-R
LALP-LP6	D-F-E-T-V-I-S-Q-D-P-I-VG-M-R-L-V-E-G-D-M-L-L-D-D-E-P-F-T-E-R
LALP-LP7	D-F-E-T-V-I-S-Q-D-P-I-VG-M-R-L-V-E-G-D-M-L-L-D-D-E-P-F-T-E-R
LALP-LP8	D-F-E-T-V-I-S-Q-D-P-I-VG-M-R-L-V-E-G-D-M-L-L-D-D-E-P-F-T-E-R
LALP-LP9	D-F-E-T-V-I-S-Q-D-P-I-VG-M-R-L-V-E-G-D-M-L-L-D-D-E-P-F-T-E-R

Tabela 8: Predição de propeptídeos de metaloproteases de *L. laeta* (P) pelo ProP.

5.5.3. PREDIÇÃO DE EPITOPOS

A partir da predição dos epítopos utilizando o servidor ABCPred nota-se a presença dos epítopos de todas as nove sequências selecionadas como metaloproteases da glândula de veneno de *L. laeta* (P) (LALP-LPs). Abaixo são descritos os resultados das três sequências mais expressivas por representatividade que posteriormente foram utilizadas para análise *in vitro* utilizando de mapeamento de epítopos em membrana de celulose pelo método de SPOT, permitindo a validação dos resultados a seguir (Figura 17, 18 e 19).

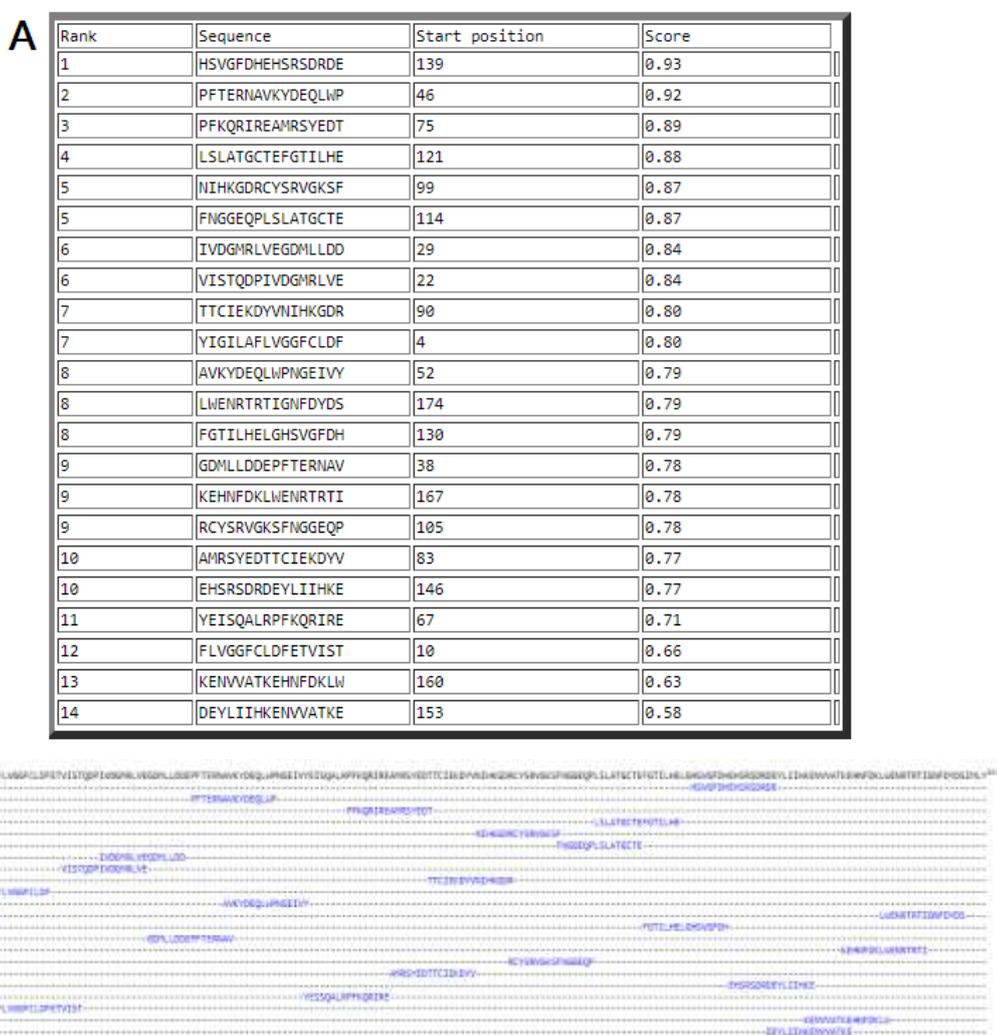


Figura 17: Predição de epítopos de LALP-LP3 utilizando ABCPred. (A) Tabela com epítopos. (B) Sobreposição dos epítopos em alinhamento.

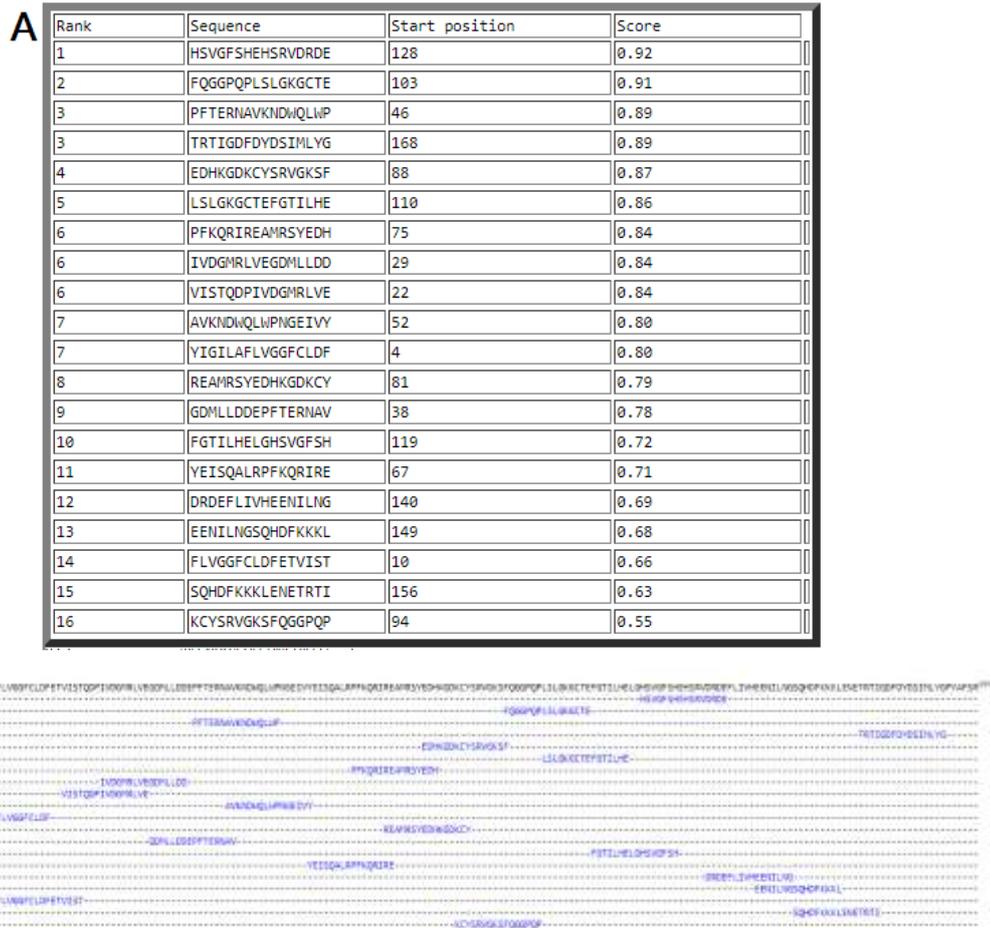


Figura 18: Predição de epitopos de LALP-LP4 utilizando ABCPred. (A) Tabela com epitopos. (B) Sobreposição dos epitopos em alinhamento.

A

Rank	Sequence	Start position	Score
1	PFTERNAVKYDQQLWP	46	0.91
1	HSVGF SHEHSRIDRDE	127	0.91
2	SKGGPQLSLGKGCTE	102	0.89
3	TRTIGYFDYDSIMLYG	166	0.87
4	RGPKCYSRVGKSKGGP	91	0.86
4	LSLGKGCTEFGTILHE	109	0.86
5	IVDGMRLVEGDM LDD	29	0.84
5	VISTQDPIVDGMRLVE	22	0.84
6	EAMRSVDIHRGPKCYS	82	0.83
7	HKENILNGSPDFK KKL	147	0.81
8	YIGILAF LVGGFCLDF	4	0.80
9	AVKYDQQLWPNGVVVY	52	0.78
9	GMMLDDEPFTERNAV	38	0.78
10	EHSRIDRDEF LIVHKE	134	0.75
11	FGTILHELGH SVGF SH	118	0.72
12	YEISQALRPFKQRI RE	67	0.71
13	KKLENETR TIGYFDYD	160	0.70
14	FLVGGFCLDFETVIST	10	0.66
15	DEF LIVHKENILNGSP	141	0.64
16	GSPDFK KLENETR TI	154	0.62



Figura 19: Predição de epitopos de LALP-LP8 utilizando ABCPred. (A) Tabela com epitopos. (B) Sobreposição dos epitopos em alinhamento.

Foram identificados 14 epitopos em LALP-LP3 e 16 nas outras duas proteoformas (LALP-LP4 e LALP-LP8) e, nas três membranas, os resultados possuem grande semelhança tanto de conteúdo quanto de posição de epitopos.

Após a utilização do segundo programa de predição de epitopos, BepiPred-2.0, novos resultados foram disponibilizados evidenciados na Figura 20. Foi utilizado o valor 0,5 como limite de epitopo, que representa a especificidade e sensibilidade limítrofes.

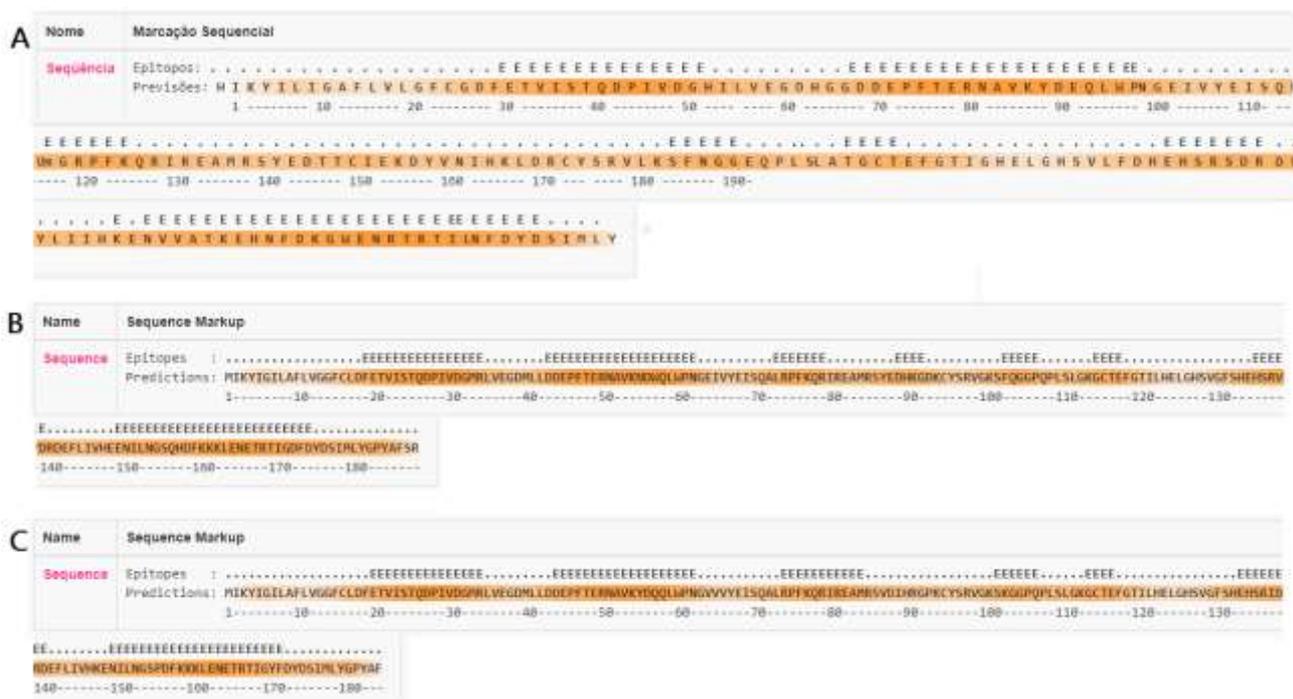


Figura 20: Predição de epitopos de LALP-LP3, LALP-LP4 e LALP-LP8 utilizando BepiPred-2.0. (A) LALP-LP3. (B) LALP-LP4. (C) LALP-LP8.

Observou-se que o programa BepiPred-2.0 identificou epitopos em menores quantidades que o programa ABCPred, porem com maior quantidade de aminoácidos em cada epitopo. Apesar disso, os epitopos de ambos os programas possuem similaridades

que corroboram com uma possível escolha que ajude no desenho de imunógenos para possíveis abordagens terapêuticas e vacinais, como Lima e colaboradores em 2018, que desenvolveram o anticorpo anti-rMEPLox a partir do reconhecimento do mesmo em epitopos das principais toxinas loxoscélicas, e garantindo proteção contra a dermonecrose causada pela mordida da aranha *Loxosceles* (LIMA et al., 2018).

5.5.4. FILOGENIA

A partir da construção das árvores filogenéticas utilizando os métodos de Máxima verossimelhança (Figura 21) observa-se que todas as proteínas possuem uma proximidade considerável, e que se formam dois grandes grupos: Um com as nove LALP-LPs e outro com as LALPs e astacina de *Astacus astacus*. Apesar disso, é possível observar que, LALP1 e LALP4 estão mais próximas das LALP-LPs do que das outras sequências.

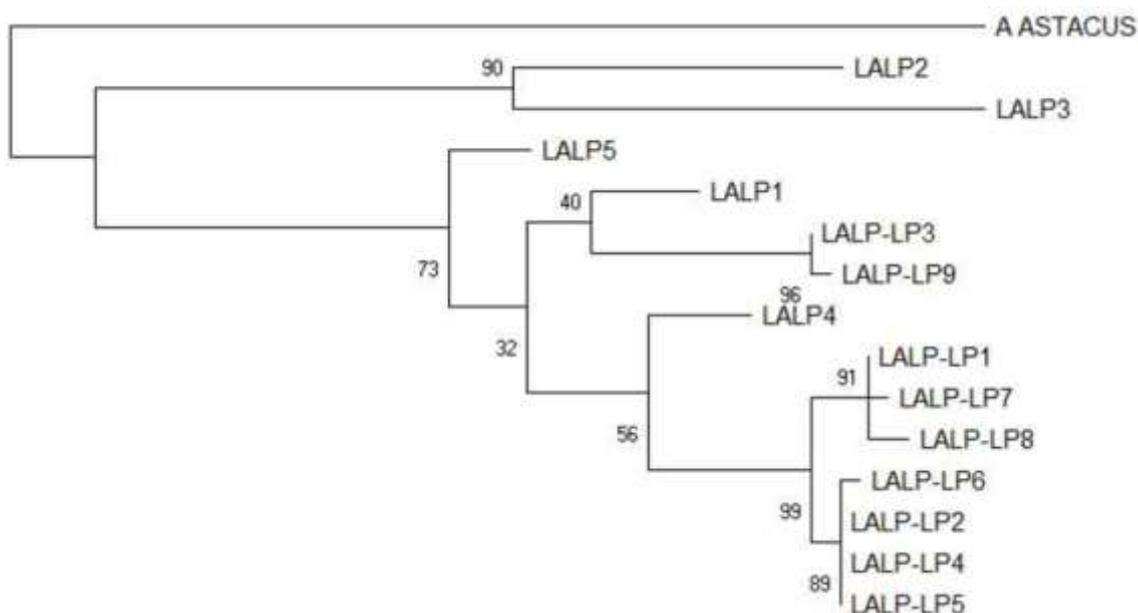


Figura 21: Árvore filogenética de máxima verossimelhança de astacina de *Astacus astacus*, LALPs 1 a 3 (metaloproteases de *L. intermedia*) LALP4 (metaloprotease de *L. laeta* (B)) e metaloproteases de *L. laeta* (P) (LALP-LP) pelo Mega 7.0.

Dantas e colaboradores em 2016 utilizaram o software MEGA 6 que também utiliza o algoritmo de Clustal W para criar uma árvore filogenética por agrupamento de vizinhos para avaliar a similaridade entre as sequências de Fosfolipases D do veneno de *L. similis* dos contigs resultantes de um sequenciamento com outras sequências de

Fosfolipases D da mesma aranha presentes no banco de dados do GenBank, utilizando um valor confiável de bootstrap de 70. Desta forma, eles identificaram que as sequências pertencentes ao mesmo grupo podem apresentar atividade similar (DANTAS et al., 2016). Da mesma forma foi possível observar uma clara separação das sequências por representatividade de LALP-LPs em três grupos: LALP-LP3 e LALP-LP9; LALP-LP1, LALP-LP7 e LALP-LP8; LALP-LP2, LALP-LP4, LALP-LP5 e LALP-LP6, logo um representante de cada grupo foi posteriormente escolhido para comporem as três membranas de celulose para a realização do mapeamento de epitopos pelo método de SPOT.

5.5.5. MODELAGEM

Utilizou-se o programa Phyre2 para a predição e modelagem 3D das nove metaloproteases de *L. laeta* (P), e o resultado disponibilizado foi analisado pelo programa PyMOL, onde inicialmente realizou-se a marcação dos elementos essenciais das metaloproteases (zinc-binding signature e met-turn) em amarelo (Figura 22). É visível que, da mesma forma que as sequências, as proteoformas são muito similares entre si e a partir da sobreposição das mesmas é possível observar pequenas diferenças em alguns aminoácidos (Figura 23). Além disso, uma sobreposição de LALP-LPs com LALP1 permitiu a identificação do gap visualizado no alinhamento (Figura 24).

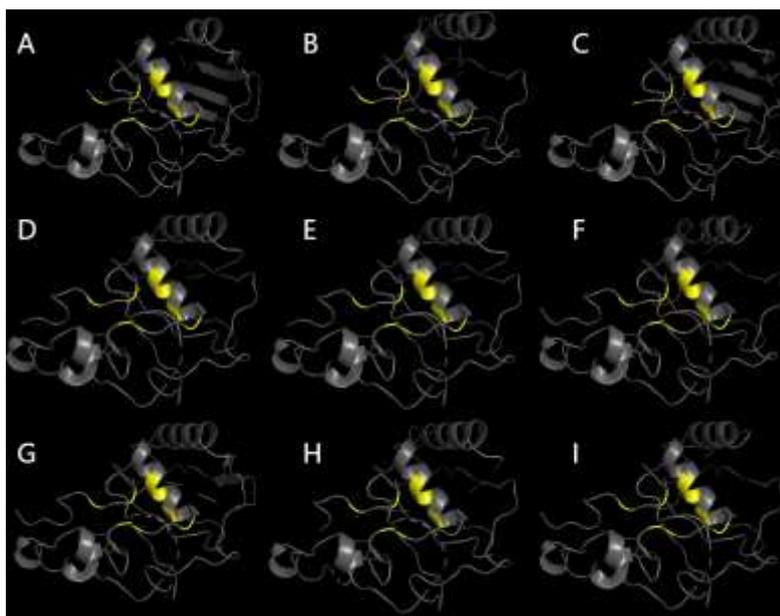


Figura 22: Modelagem das metaloproteases identificadas por transcriptômica utilizando o programa Phyre2 e análise dos resultados feita pelo programa PyMOL. (A) LALP-LP1; (B) LALP-LP2; (C)

LALP-LP3; (D) LALP-LP4; (E) LALP-LP5; (F) LALP-LP6; (G) LALP-LP7; (H) LALP-LP8; (I) LALP-LP9; (Amarelo) Elementos essenciais de metaloproteases (HEXXXHXXGXXHE e MXY).

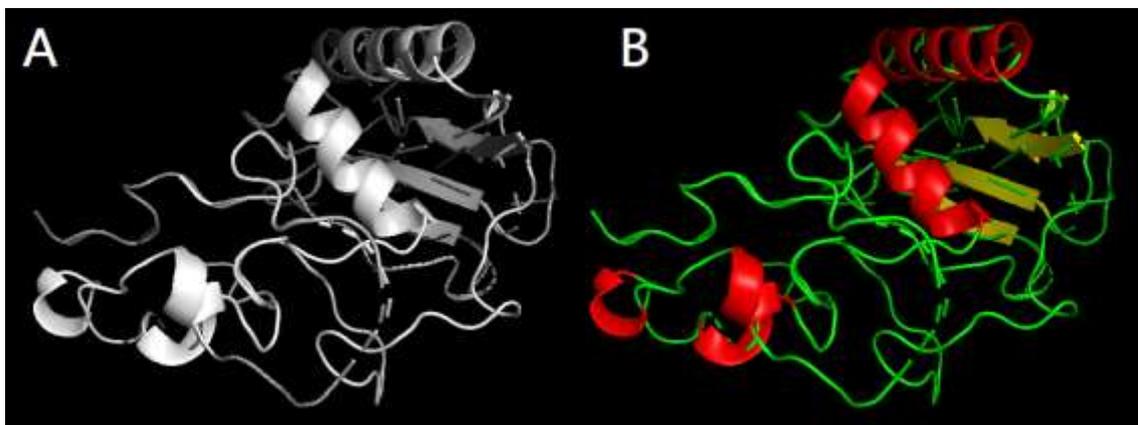


Figura 23: Modelagem das nove metaloproteases identificadas por transcriptômica utilizando o programa Phyre2 e sobrepostas pelo programa PyMOL. A: Sobreposição em tons de cinza; B: Sobreposição identificando alças em verde, folhas beta em amarelo e alfa hélices em vermelho.

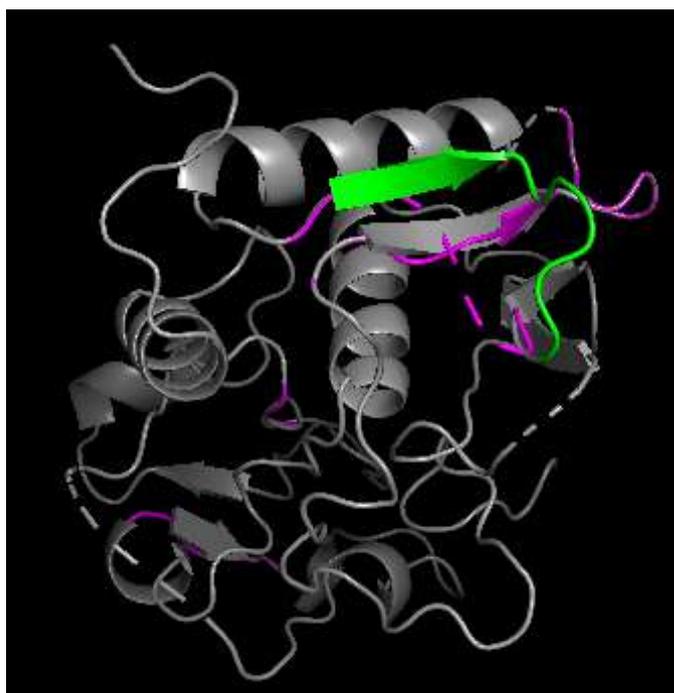


Figura 24: Modelagem de LALP1 e LALP-LP3 utilizando o programa Phyre2 sobrepostas pelo programa PyMOL. (Cinza) LALP1; (Rosa) LALP-LP3; (Verde) Gap de LALP-LPs que está presente em LALP1.

Através da marcação dos elementos essenciais das nove sequências de metaloproteases é possível observar exatamente onde provavelmente ocorre a interação do sítio catalítico com o íon zinco (Figura 22 – amarelo). Morgon e colaboradores em

2016 realizaram a predição da estrutura tridimensional de LALP3 utilizando o programa SWISS-MODEL para predizer e o programa Open Astex Viewer tool para visualização do modelo, onde foi possível localizar o íon zinco interagindo com o sítio catalítico das enzimas e os locais de possíveis N-glicosilação (MORGON et al., 2016).

5.6. VALIDAÇÃO DA PRESENÇA DE METALLOPROTEASES *IN VITRO*

5.6.1. SDS-PAGE E *WESTERN BLOTTING*

Com a finalidade de verificar a presença das proteínas nos venenos de *L. laeta* provenientes do Brasil e do Peru, para validar os resultados identificados por transcriptômica, foram realizadas as técnicas de SDS-PAGE e *Western Blotting*.

Na Figura 25 é possível observar um gel de SDS-PAGE 12,5% corado com prata que utiliza como controle o veneno de *L. intermedia* e como amostras os venenos de *L. laeta* brasileira e peruana, onde é possível observar a partir do padrão de peso molecular bandas próximas a 24kDa, supondo então a presença de metaloproteases, que comumente possuem peso molecular de 29kDa, além da presença de uma banda a mais nesta mesma região.

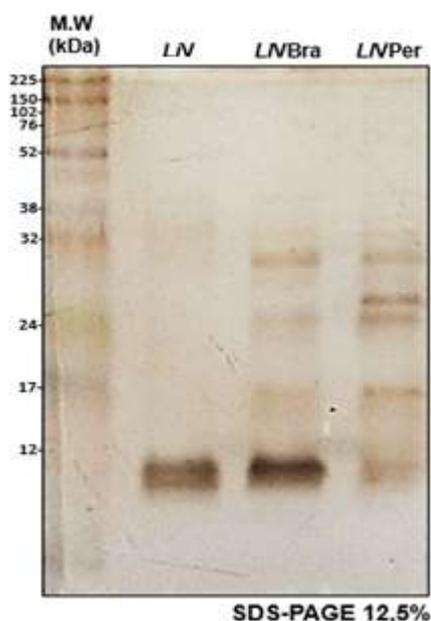


Figura 25: SDS-PAGE de venenos de *L. intermedia*, *L. laeta* brasileira e peruana. A primeira canaleta possui os padrões de massa molecular, a segunda possui veneno de *L. intermedia*, a terceira possui veneno de *L. laeta* (B) e a última canaleta possui veneno de *L. laeta* (P).

Os resultados do *western blotting* estão apresentados na Figura 26, onde estão apresentados consecutivamente a membrana corada com Ponceau e esta mesma membrana após a finalização do *western blotting*, onde é possível observar uma acentuação das bandas próximas ao peso molecular de 24kDa, demonstrando que os anticorpos anti-metaloproteases se ligaram nas bandas em que havíamos suspeitado serem referentes às metaloproteases, presumindo assim a presença destas proteínas.

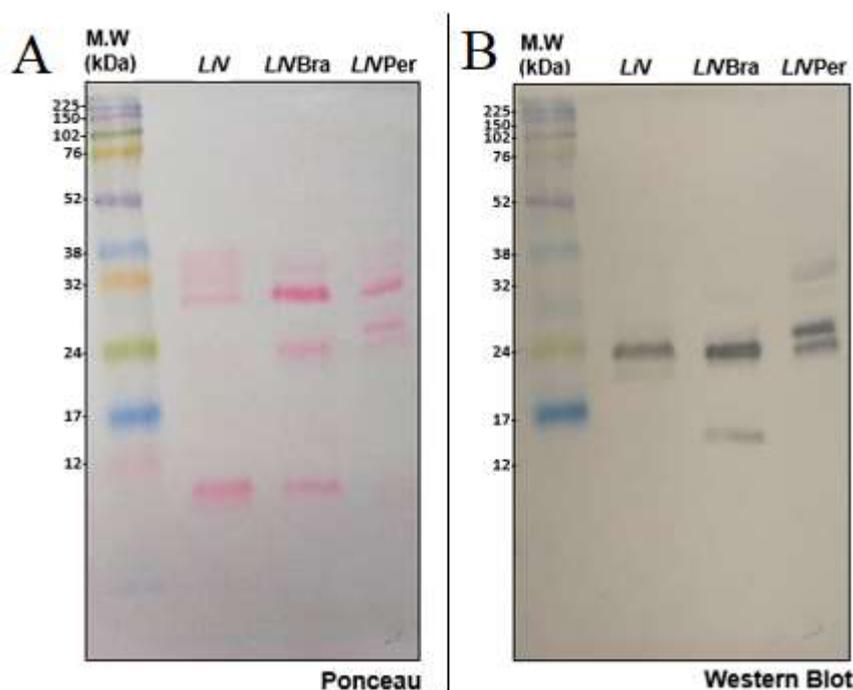


Figura 26: *Western blotting* de venenos de *L. intermedia*, *L. laeta* brasileira e peruana. A) Corado com Ponceau e B) Após finalização do *western blotting*. A primeira canaleta possui os padrões de massa molecular, a segunda possui veneno de *L. intermedia*, a terceira possui veneno de *L. laeta* (B) e a última canaleta possui veneno de *L. laeta* (P).

5.6.2. ENSAIOS ENZIMÁTICOS

5.6.2.1 ATIVIDADES FIBRINOGENOLÍTICAS

O fibrinogênio bovino é composto por três subunidades: $A\alpha$, $B\beta$ e γ , sendo possível observar na Figura 27 que a atividade fibrinogenolítica dos venenos de *L. laeta* brasileiro e peruano degradam a subunidade $A\alpha$ (seta).

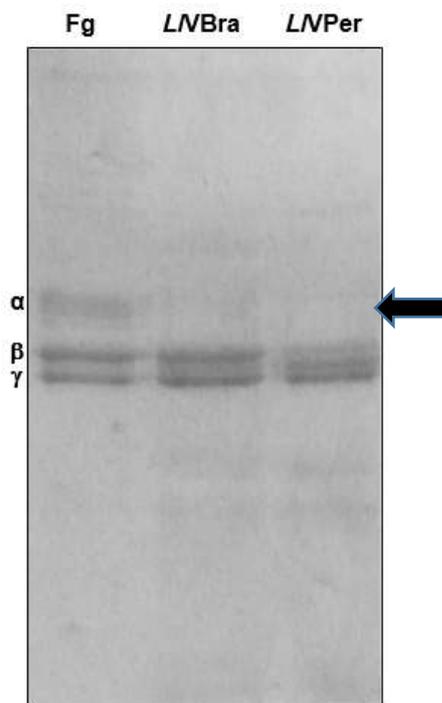


Figura 27: Atividade fibrinogenolítica de *L. laeta* brasileira e peruana. A primeira canaleta possui fibrinogênio, a segunda possui fibrinogênio incubado com veneno de *L. laeta* (B) e a última canaleta possui fibrinogênio incubado com veneno de *L. laeta* (P).

Em outro experimento (Figura 28), em que a atividade fibrinogenolítica foi realizada condicionada a tempos específicos (tempo-dependente) é possível observar que o veneno de *L. laeta* (P) degrada a subunidade A α com em um menor tempo que o veneno de *L. laeta* (B), como observado no tempo de 10 horas, onde as bandas foram quantificadas utilizando o programa ImageJ, em que a densidade da banda α em *L. laeta* (P) era de 1%, enquanto em *L. laeta* (B) ainda era de 10%, sendo possível presumir que aparentemente existe uma diferença de degradação do fibrinogênio entre os dois venenos que necessita de confirmação.

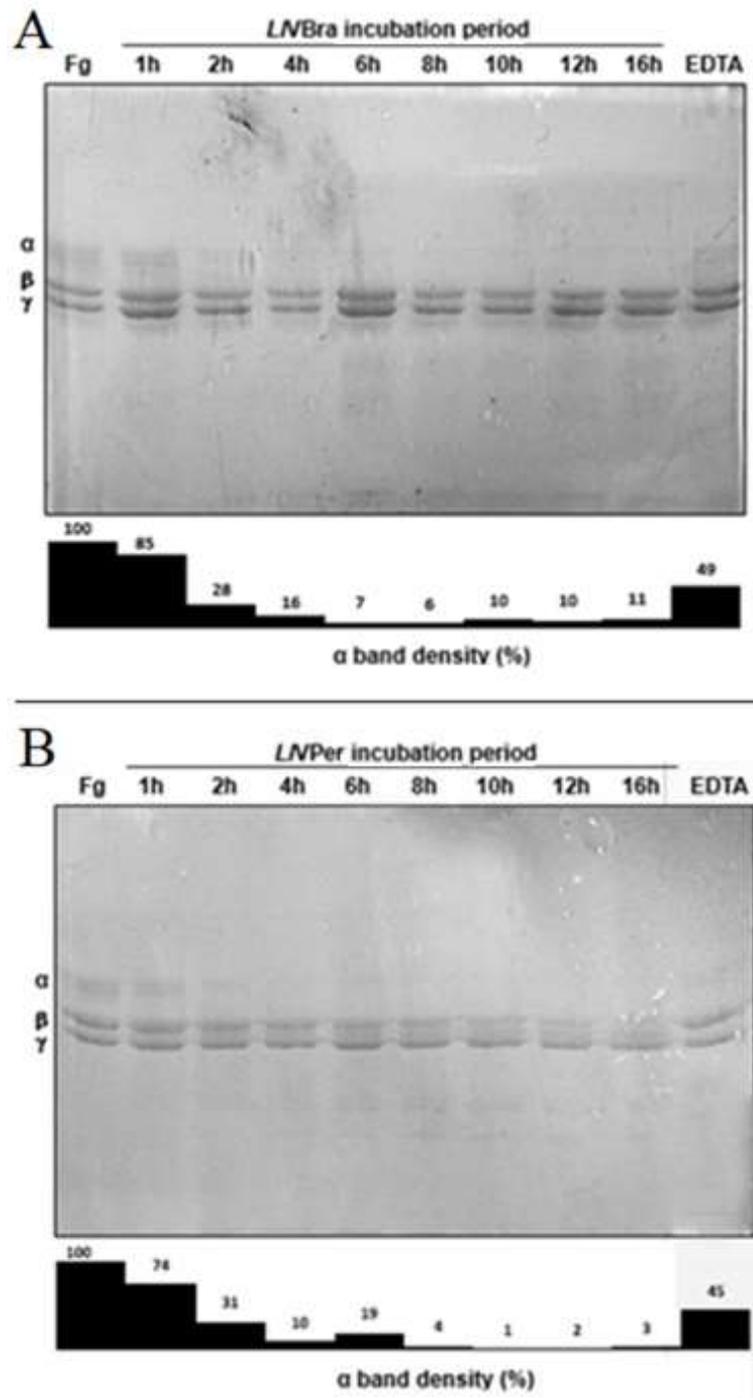


Figura 28: Atividade fibrinogenolítica tempo-dependente de *L. laeta* brasileira (A) e peruana (B). A primeira canaleta possui fibrinogênio, da segunda a nona canaleta possuem venenos de *L. laeta* brasileira e peruana acrescidas de fibrinogênio, e a última canaleta possui apenas EDTA.

5.6.2.2 ENSAIO DE GELATINASE/COLAGENASE

A partir dos resultados obtidos pelo ensaio utilizando o kit *EnzChek Gelatinase/Colagenase* é possível observar que em todas as concentrações o veneno de

L. laeta (P) possui uma maior porcentagem de atividade da colagenase que o veneno de *L. laeta* (B) (Figura 29).

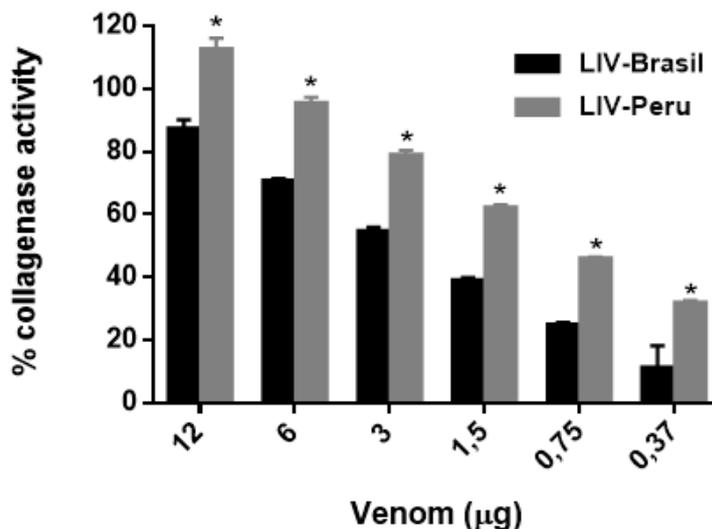
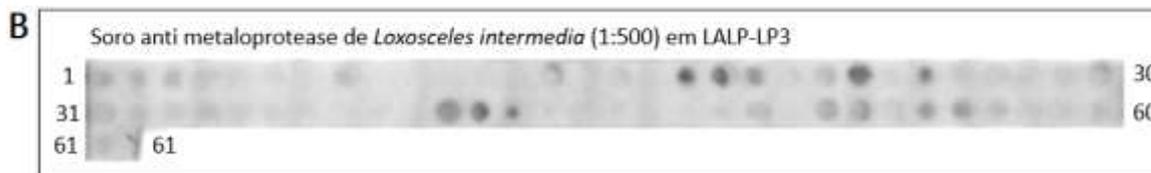


Figura 29: Gráfico de ensaio de gelatinase/colagenase de venenos de *L. laeta* brasileira e peruana. Porcentagem de atividade da colagenase em relação a concentração dos venenos.

5.6.3. MAPEAMENTO DE EPITOPOS PELO MÉTODO DE SPOT

O primeiro experimento utilizou o soro anti-metaloprotease de *L. intermedia* (diluição 1:500) para avaliação de reatividade de anticorpos frente as membranas. As imagens das membranas de LALP-LP3 (Figura 30), LALP-LP4 (Figura 31) e LALP-LP8 (Figura 32) foram analisadas pelo programa ImageJ, que realiza a quantificação de pixels pela área pré-determinada de um SPOT e um gráfico foi plotado para cada membrana afim de demonstrar as regiões de reação de acordo com o resultado do programa ImageJ.

Spot #	Peptideo						
1	M-I-K-Y-I-G-I-L-A-F-L-V-G-G-F	17	E-R-N-A-V-K-F-D-E-Q-L-W-P-N-G	33	V-V-N-I-H-K-G-D-R-C-Y-S-R-V-G	49	H-E-H-S-R-S-D-R-Q-E-Y-L-I-I-H
2	Y-I-G-I-L-A-F-L-V-G-G-F-C-L-D	18	A-V-K-Y-D-E-Q-L-W-P-N-G-E-V	34	H-H-G-D-R-C-Y-S-R-V-G-K-S-F	50	S-R-S-D-R-D-E-Y-L-I-I-H-K-E-N
3	I-L-A-F-L-V-G-G-F-C-L-D-F-F-T	19	Y-D-E-Q-L-W-P-N-G-E-I-V-Y-E-I	35	G-D-R-C-Y-S-R-V-G-K-S-E-N-G-G	51	D-R-D-E-Y-L-I-I-H-K-E-N-V-V-A
4	F-L-V-G-G-F-C-L-D-F-F-T-V-I-S	20	Q-L-W-P-N-G-E-I-V-Y-E-I-S-Q-A	36	C-Y-S-R-V-G-K-S-F-N-G-G-E-Q-P	52	E-Y-L-I-I-H-K-E-N-V-V-A-T-K-E
5	G-G-F-C-L-D-F-F-T-V-I-S-T-Q-D	21	P-N-G-E-I-V-Y-E-I-S-Q-A-I-R-F	37	R-V-G-K-S-F-N-G-G-E-Q-P-L-S-L	53	I-I-H-H-E-N-V-V-A-T-K-E-H-N-F
6	C-L-D-F-F-T-V-I-S-T-Q-D-P-I-V	22	E-I-V-Y-E-I-S-Q-A-L-R-P-F-K-Q	38	K-S-F-N-G-G-E-Q-P-L-S-L-A-T-G	54	W-E-N-V-V-A-T-K-E-H-N-F-D-K-I
7	F-E-T-V-I-S-T-Q-D-P-I-V-G-G-M	23	Y-E-I-S-Q-A-L-R-P-F-K-Q-R-I-E	39	N-G-G-E-Q-P-L-S-L-A-T-G-C-T-E	55	V-V-A-T-K-E-H-N-F-D-K-L-W-E-N
8	V-I-S-T-Q-D-P-I-V-D-G-M-R-L-V	24	S-Q-A-L-R-P-F-K-Q-R-I-E-A-M	40	E-Q-P-L-S-L-A-T-G-C-T-E-F-G-T	56	T-K-E-H-N-F-D-K-L-W-E-N-R-T-R
9	T-Q-D-P-I-V-D-G-M-R-L-V-E-G-D	25	L-R-P-F-K-Q-R-I-R-E-A-M-R-S-F	41	L-S-L-A-T-G-C-T-E-F-G-T-I-I-H	57	H-N-F-D-K-L-W-E-N-R-T-R-T-I-G
10	P-I-V-D-G-M-R-L-V-E-G-D-M-L-L	26	F-E-Q-R-I-R-E-A-M-R-S-Y-E-D-T	42	A-T-G-C-T-E-F-G-T-I-I-H-E-L-G	58	D-K-L-W-E-N-R-T-R-T-I-G-N-F-D
11	D-G-M-R-L-V-E-G-D-M-L-L-D-D-E	27	R-I-R-E-A-M-R-S-Y-E-D-T-T-C-I	43	C-T-E-F-G-T-I-I-H-E-L-G-H-S-V	59	W-E-N-R-T-R-T-I-G-N-F-D-Y-D-S
12	R-L-V-E-G-D-M-L-L-D-D-E-P-F-T	28	E-A-M-R-S-Y-E-D-T-T-C-I-E-K-D	44	F-G-T-I-I-H-E-L-G-H-S-V-G-F-D	60	R-T-R-T-I-G-N-F-D-Y-D-S-I-M-L
13	E-G-D-M-L-L-D-D-E-P-F-T-E-R-N	29	R-S-Y-E-D-T-T-C-I-E-E-B-Y-V-N	45	I-L-H-E-L-G-H-S-V-G-F-D-H-E-N	61	T-R-T-I-G-N-F-D-Y-D-S-I-M-L-V
14	M-L-L-D-D-E-P-F-T-E-R-N-A-V-K	30	E-D-T-T-C-I-E-K-D-Y-V-N-I-H-K	46	E-L-G-H-S-V-G-F-D-H-E-H-S-R-S		
15	D-D-E-P-F-T-E-R-N-A-V-K-Y-D-E	31	T-C-S-E-K-D-Y-V-N-I-H-K-G-D-R	47	H-S-V-G-F-D-H-E-H-S-R-S-D-R-D		
16	P-F-T-E-R-N-A-V-K-Y-D-E-Q-L-W	32	E-K-D-Y-V-N-I-H-K-G-D-R-C-Y-S	48	G-F-D-H-E-H-S-R-S-D-R-D-E-Y-L		



C LALP-LP3 com soro anti metaloprotease de *Loxosceles intermedia* (1:500)

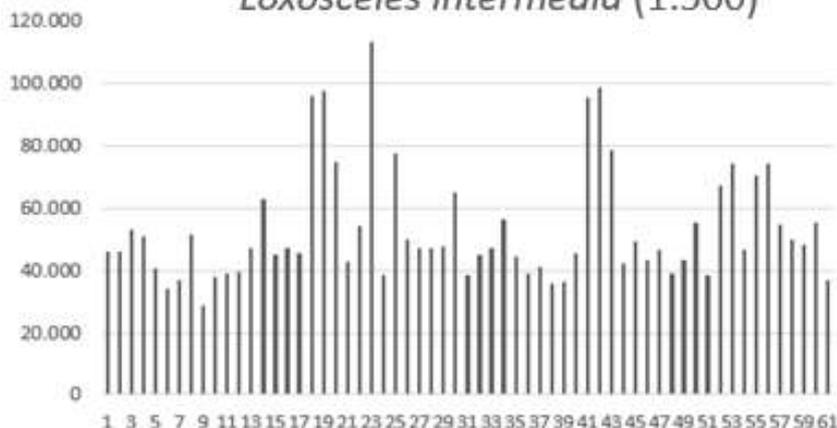
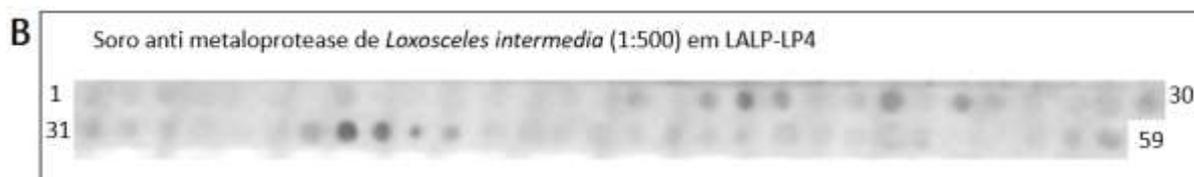


Figura 30: Mapeamento de epitopos de LALP-LP3 com soro anti-metaloprotease de *L. intermedia*. (A) Sequências dos pentadecapeptídeos correspondentes a cada SPOT sintetizado na membrana. (B) Membrana contendo peptídeos derivados da sequência primária de LALP-LP3 testada com soro anti-metaloprotease de *L. intermedia* (diluição 1:500). (C) Gráfico contendo a quantificação de coloração dos spots medida pelo programa ImageJ.

A

Spot #	Peptídeo						
1	M-I-K-Y-I-D-I-L-A-F-L-V-G-G-F	17	E-R-N-A-V-K-N-D-W-Q-L-W-P-N-Q	33	S-R-V-D-K-S-F-D-G-G-P-Q-P-L-S	49	L-I-V-H-E-E-N-I-L-N-G-S-Q-H-D
2	Y-I-G-I-L-A-F-L-V-G-G-F-C-L-D	18	A-V-K-N-D-W-Q-L-W-P-N-G-E-I-V	34	G-K-S-F-Q-G-G-P-Q-P-L-S-L-G-K	50	H-E-E-N-I-L-N-G-S-Q-H-D-F-K-K
3	I-L-A-F-L-V-G-G-F-C-L-D-F-E-T	19	N-D-W-Q-L-W-P-N-G-E-I-V-Y-E-I	35	F-Q-G-G-P-Q-P-L-S-L-G-F-G-C-T	51	N-I-L-N-G-S-Q-H-D-F-K-K-L-E-E
4	F-L-V-G-G-F-C-L-D-F-E-T-V-I-S	20	Q-L-W-P-N-G-E-I-V-Y-E-I-S-Q-A	36	G-P-Q-P-L-S-L-G-K-G-C-T-E-F-G	52	N-G-S-Q-H-D-F-K-K-L-E-N-E-T
5	G-G-F-C-L-D-F-E-T-V-I-S-T-Q-D	21	P-N-G-E-I-V-Y-E-I-S-Q-A-L-R-P	37	F-L-S-L-G-K-G-C-T-E-F-G-T-I-L	53	Q-H-D-F-K-K-L-E-N-E-T-R-T-I
6	C-L-D-F-E-T-V-I-S-T-Q-D-P-I-V	22	E-I-V-Y-E-I-S-Q-A-L-R-P-F-K-Q	38	L-G-K-G-C-T-E-F-G-T-I-L-H-E-L	54	F-K-K-L-E-N-E-T-R-T-I-G-D-F
7	F-E-T-V-I-S-T-Q-D-P-I-V-D-G-M	23	Y-E-I-S-Q-A-L-R-P-F-K-Q-R-I-R	39	G-C-T-E-F-G-T-I-L-H-E-L-G-H-S	55	K-L-E-N-E-T-R-T-I-G-D-F-D-Y-D
8	V-I-S-T-Q-D-P-I-V-D-G-M-R-L-V	24	S-Q-A-L-R-P-F-K-Q-R-I-R-E-A-M	40	E-F-G-T-I-L-H-E-L-G-H-S-V-G-F	56	N-E-T-R-T-I-G-D-F-O-Y-D-S-I-M
9	T-Q-D-P-I-V-D-G-M-R-L-V-E-G-Q	25	L-R-P-F-K-Q-R-I-R-E-A-M-R-S-Y	41	T-I-L-H-E-L-G-H-S-V-G-F-S-H-E	57	R-T-I-G-D-F-D-Y-D-S-I-M-L-Y-G
10	P-I-V-D-G-M-R-L-V-E-G-D-M-L-L	26	F-K-Q-R-I-R-E-A-M-R-S-Y-E-D-H	42	H-E-L-G-H-S-V-G-F-S-H-E-H-S-R	58	G-O-F-D-Y-D-S-I-M-L-Y-G-P-Y-A
11	D-G-M-R-L-V-E-G-D-M-L-L-D-D-E	27	R-I-R-E-A-M-R-S-Y-E-D-H-K-G-Q	43	G-H-S-V-G-F-S-H-E-H-S-R-V-D-R	59	D-Y-D-S-I-M-L-Y-G-P-Y-A-F-S-R
12	R-L-V-E-G-D-M-L-L-D-D-E-P-F-T	28	E-A-M-R-S-Y-E-D-H-K-G-D-K-C-Y	44	V-G-F-S-H-E-H-S-R-V-D-R-D-E-F		
13	E-G-D-M-L-L-D-D-E-P-F-T-E-R-N	29	R-S-Y-E-D-H-K-G-D-K-C-Y-S-R-V	45	S-H-E-H-S-R-V-D-R-D-E-F-L-I-V		
14	M-L-L-D-D-E-P-F-T-E-R-N-A-V-K	30	E-D-H-K-G-D-K-C-Y-S-R-V-G-K-S	46	H-S-R-V-D-R-D-E-F-L-I-V-H-E-E		
15	D-D-E-P-F-T-E-R-N-A-V-K-N-D-W	31	K-G-D-K-C-Y-S-R-V-G-K-S-F-Q-G	47	V-D-R-D-E-F-L-I-V-H-E-E-N-I-L		
16	P-F-T-E-R-N-A-V-K-N-D-W-Q-L-W	32	K-C-Y-S-R-V-G-K-S-F-Q-G-G-P-Q	48	D-E-F-L-I-V-H-E-E-N-I-L-N-G-S		



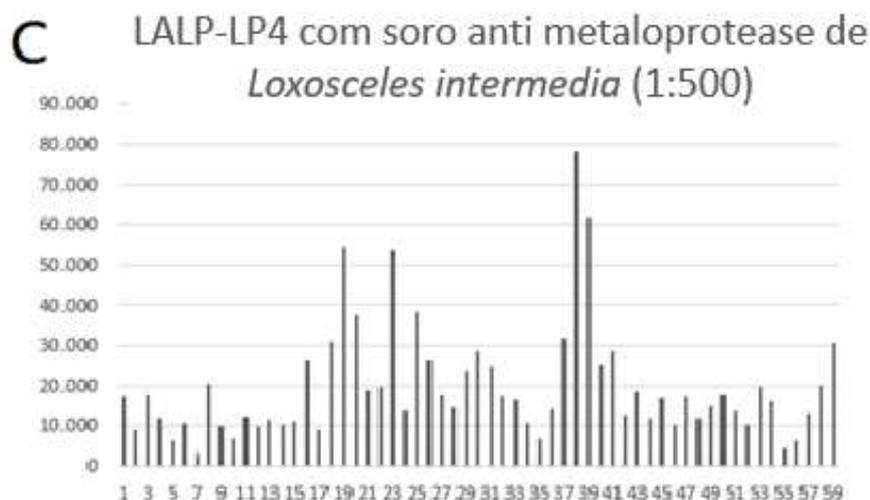
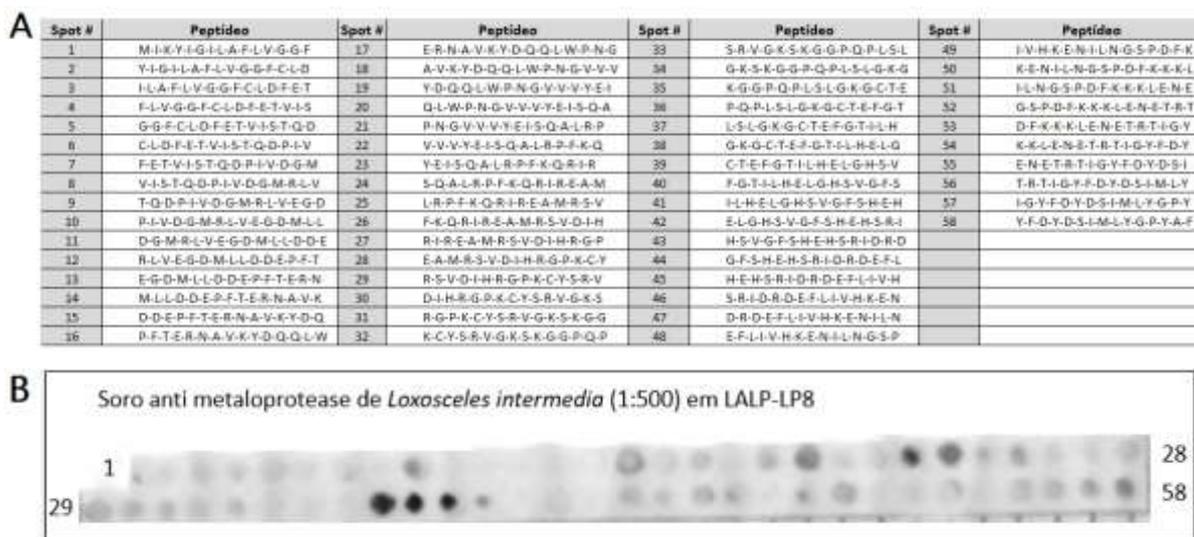


Figura 31: Mapeamento de epitopos de LALP-LP4 com soro anti-metaloprotease de *L. intermedia*. (A) Sequências dos pentadecapeptídeos correspondentes a cada SPOT sintetizado na membrana. (B) Membrana contendo peptídeos derivados da sequência primária de LALP-LP4 testada com soro anti-metaloprotease de *L. intermedia* (diluição 1:500). (C) Gráfico contendo a quantificação de coloração dos spots medida pelo programa ImageJ.



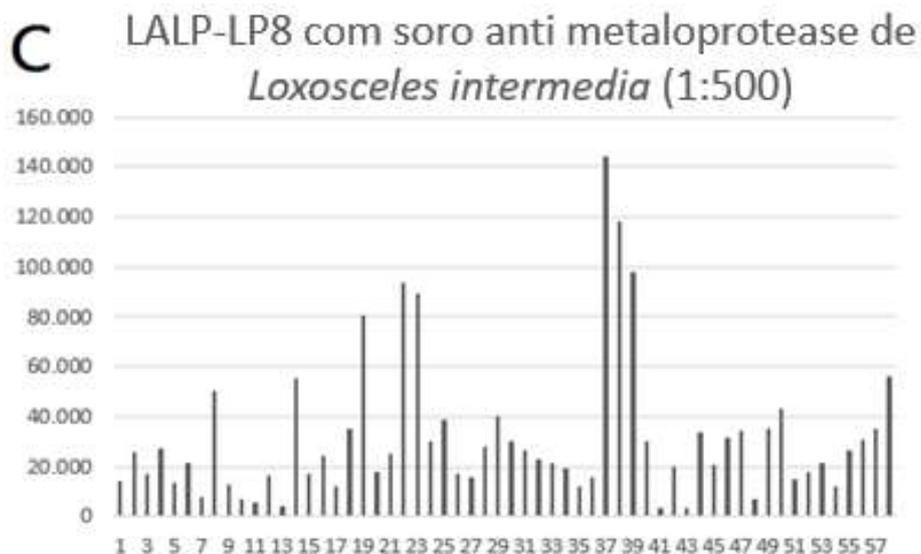
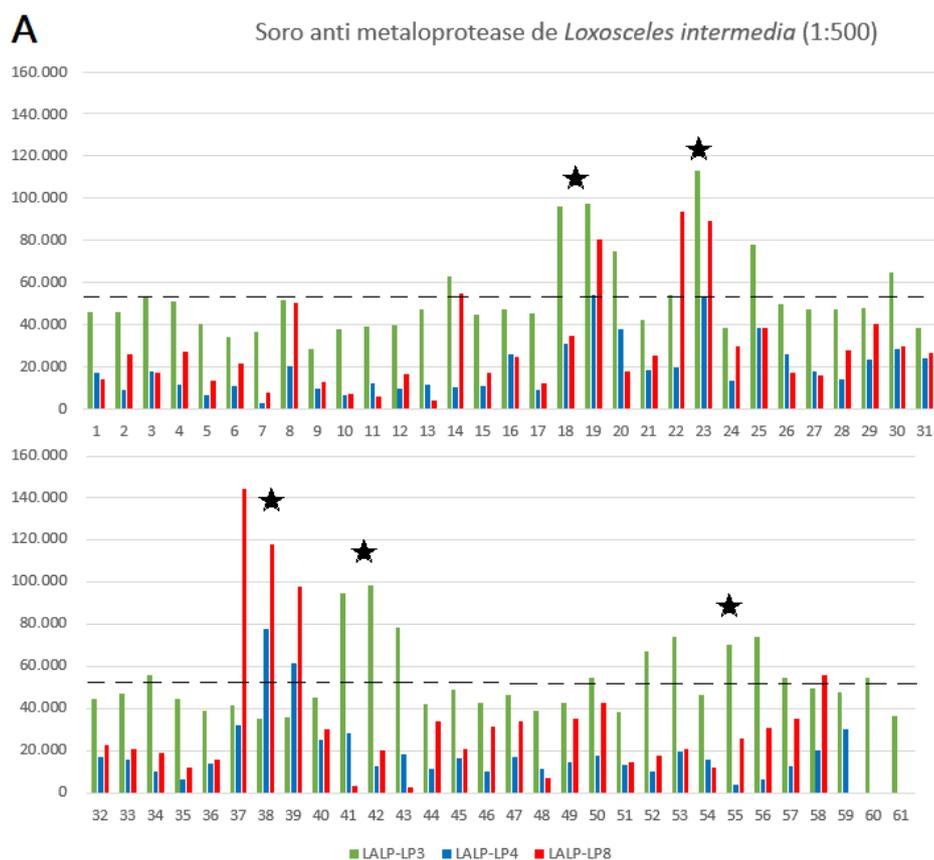


Figura 32: Mapeamento de epitopos de LALP-LP8 com soro anti-metaloprotease de *L. intermedia*. (A) Sequências dos pentadecapeptídeos correspondentes a cada SPOT sintetizado na membrana. (B) Membrana contendo peptídeos derivados da sequência primária de LALP-LP8 testada com soro anti-metaloprotease de *L. intermedia* (diluição 1:500). (C) Gráfico contendo a quantificação de coloração dos spots medida pelo programa ImageJ.

Com os resultados da análise realizada com o programa ImageJ, um gráfico sobrepondo o mapeamento dos epitopos das três membranas com soro anti-metaloprotease de *L. intermedia* também foi realizado. Os spots com reatividade acima de 60.000 pixels/área foram considerados os principais epitopos e evidenciados por estrelas pretas (Figura 33). Quatro áreas foram escolhidas como as de localização dos principais epitopos, sendo LALP-LP3 responsável pela reatividade de três áreas, LALP-LP4 de uma área e LALP-LP8 de três áreas.



B

Membrana	Spot #	Peptídeo
LALP-LP3	18	A-V-K-Y-D-E-Q-L-W-P-N-G-E-I-V
	19	Y-D-E-Q-L-W-P-N-G-E-I-V-Y-E-I
	23	Y-E-I-S-Q-A-L-R-P-F-K-Q-R-I-R
	41	L-S-L-A-T-G-C-T-E-F-G-T-I-L-H
	42	A-T-G-C-T-E-F-G-T-I-L-H-E-L-G
	43	C-T-E-F-G-T-I-L-H-E-L-G-H-S-V
	52	E-Y-L-I-I-H-K-E-N-V-V-A-T-K-E
	53	I-I-H-K-E-N-V-V-A-T-K-E-H-N-F
	55	V-V-A-T-K-E-H-N-F-D-K-L-W-E-N
56	T-K-E-H-N-F-D-K-L-W-E-N-R-T-R	
LALP-LP4	38	L-G-K-G-C-T-E-F-G-T-I-L-H-E-L
	39	G-C-T-E-F-G-T-I-L-H-E-L-G-H-S
LALP-LP8	19	Y-D-Q-Q-L-W-P-N-G-V-V-Y-E-I
	22	V-V-V-Y-E-I-S-Q-A-L-R-P-F-K-Q
	23	Y-E-I-S-Q-A-L-R-P-F-K-Q-R-I-R
	37	L-S-L-G-K-G-C-T-E-F-G-T-I-L-H
	38	G-K-G-C-T-E-F-G-T-I-L-H-E-L-G
	39	C-T-E-F-G-T-I-L-H-E-L-G-H-S-V
58	Y-F-D-Y-D-S-I-M-L-Y-G-P-Y-A-F	

Figura 33: Gráfico com a quantificação de coloração dos spots medido de LALP-LP3, LALP-LP4 e LALP-LP8 LP3 com soro anti-metaloprotease de *L. intermedia* medidos por ImageJ. (A) Gráfico com quantificação de spots. Os que possuíam reatividade acima de 55.000 pixels/área foram considerados os principais epitopos e foram evidenciados por estrelas pretas. (B) Sequências dos pentadecapeptídeos marcados que possuíam reatividade em LALP-LP3, LALP-LP4 e LALP-LP8.

Os epitopos de metaloproteases de rMEPLox foram então alinhados com as sequências completas das três principais LALP-LPs identificadas neste trabalho utilizando o programa Jalview (Figura 34), onde foi possível observar uma conservação de aminoácidos de rMEPLox_Ep1 (SLGRGCTDFGTILHE) e rMEPLox_Ep2 (ENNTRTIGPFDYDSIMLYGAY) com epitopos encontrados nas três membranas de LALP-LP, marcados em vermelho e azul, sequencialmente. Já o ultimo epitopo - rMEPLox_Ep3 – não possui identidade suficiente com as outras sequências, logo o mesmo não pode ser considerado presente nas LALP-LPs, marcado em verde.

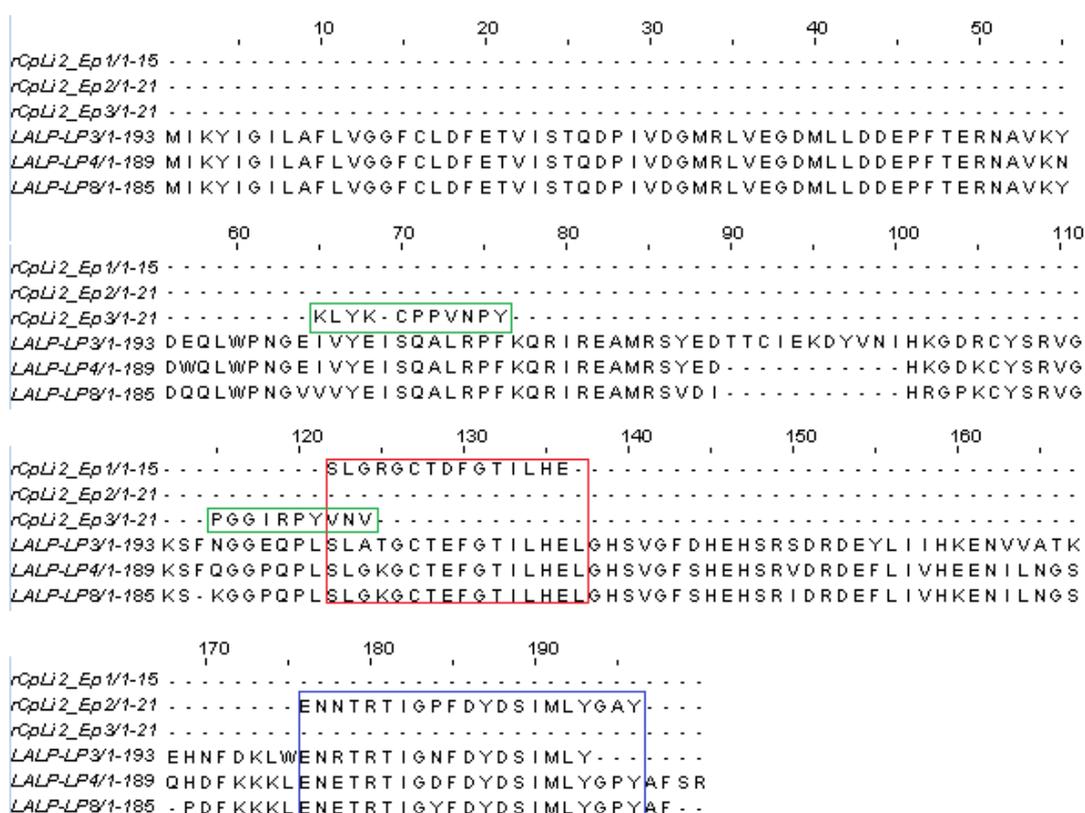


Figura 34: Alinhamento dos epitopos de metaloproteases de rMEPLox com sequências completas de LALP-LP3, LALP-LP4 e LALP-LP8. (Vermelho e Azul) Epitopos identificados em rMEPLox que também podem ser encontrados em LALP-LPs. (Verde) Epitopo de rMEPLox que não pode ser encontrado nas membranas de LALP-LP.

Após os alinhamentos, realizou-se um segundo ensaio de SPOT utilizando as três membranas de metaloproteases do veneno de *L. laeta* (P) LALP-LP3 (Figura 35), LALP-LP4 (Figura 36) e LALP-LP8 (Figura 37) com o soro anti rMEPLox disponível em nosso laboratório (diluição 1:100), afim de confirmar os resultados identificados nos

alinhamentos. As imagens das membranas foram analisadas pelo programa ImageJ e um gráfico foi plotado para cada membrana afim de demonstrar as regiões de reação.

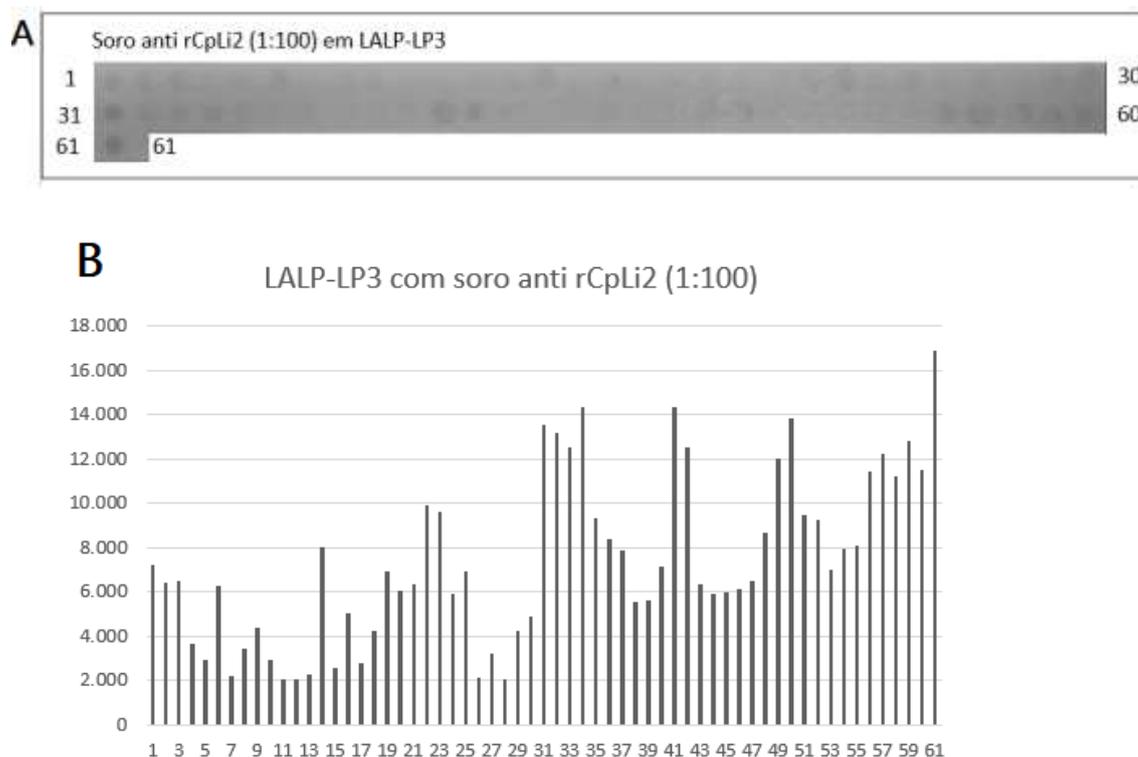


Figura 35: Mapeamento de epítomos de LALP-LP3 com soro anti rMEPLox (Quimera) de *L. intermedia*. (A) Membrana contendo peptídeos derivados da sequência primária de LALP-LP3 testada com soro anti rMEPLox (Quimera) de *L. intermedia* (diluição 1:100). (C) Gráfico contendo a quantificação de coloração dos spots medida pelo programa ImageJ.



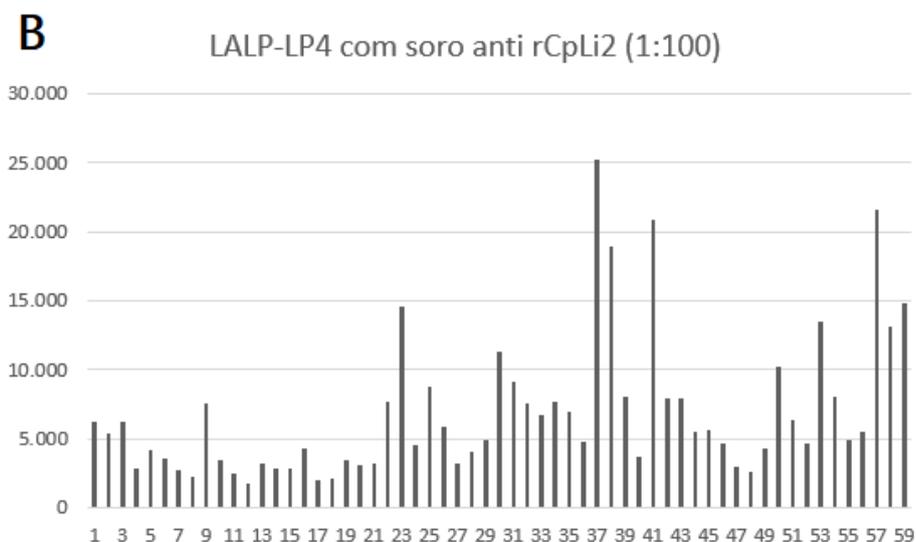


Figura 36: Mapeamento de epítomos de LALP-LP4 com soro anti rMEPLox (Quimera) de *L. intermedia*. (A) Membrana contendo peptídeos derivados da sequência primária de LALP-LP4 testada com soro anti rMEPLox (Quimera) de *L. intermedia* (diluição 1:100). (C) Gráfico contendo a quantificação de coloração dos spots medida pelo programa ImageJ.

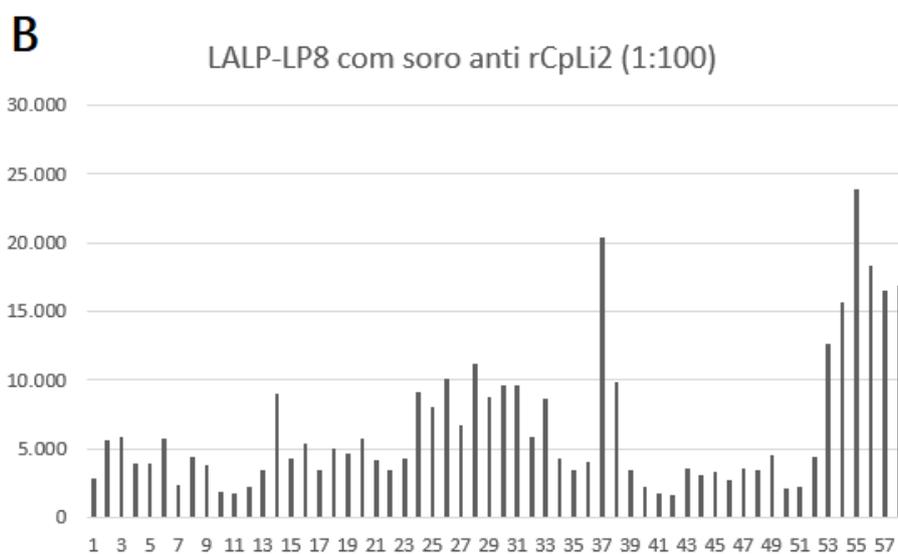
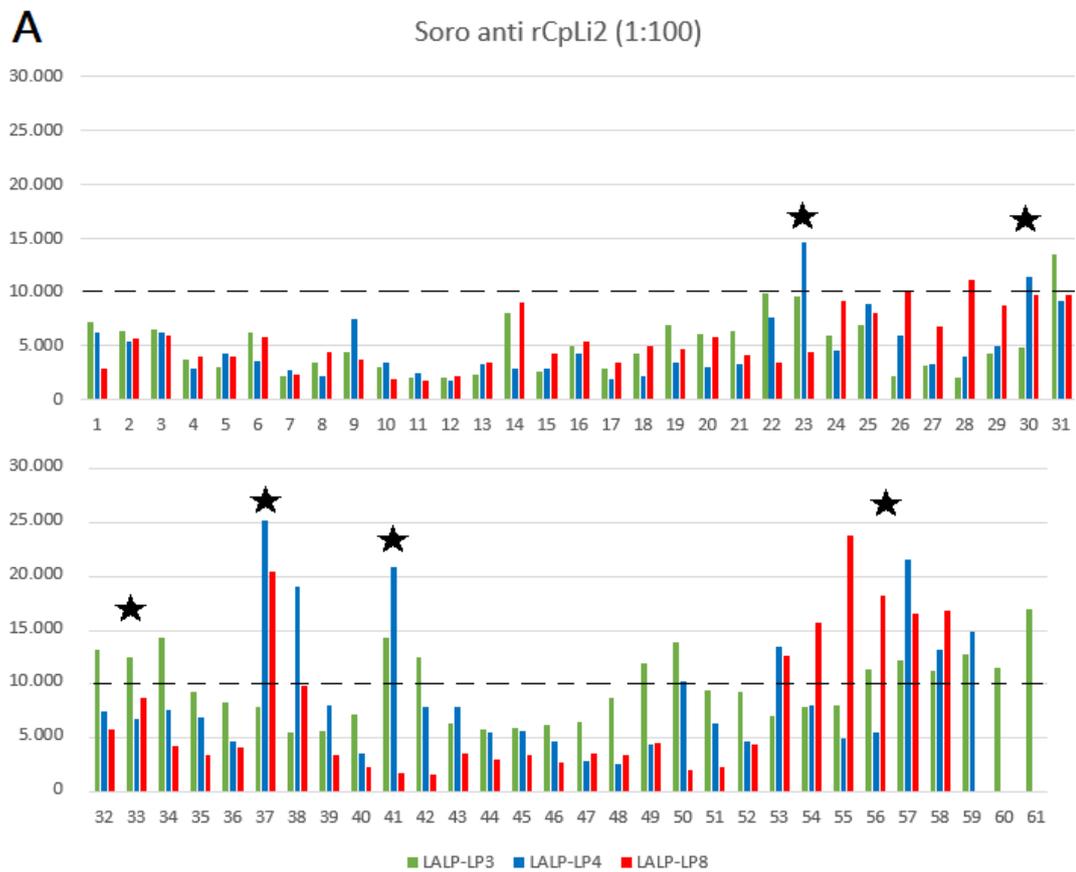


Figura 37: Mapeamento de epítomos de LALP-LP8 com soro anti rMEPLox (Quimera) de *L. intermedia*. (A) Membrana contendo peptídeos derivados da sequência primária de LALP-LP8 testada com soro anti rMEPLox (Quimera) de *L. intermedia* (diluição 1:100). (C) Gráfico contendo a quantificação de coloração dos spots medida pelo programa ImageJ.

Com os resultados da análise realizada com o programa ImageJ, um gráfico sobrepondo o mapeamento dos epitopos das três membranas com soro anti rMEPLox de *L. intermedia* também foi realizado. Os spots com reatividade acima de 10.000 pixels/área foram considerados os principais epitopos e evidenciados por estrelas pretas (Figura 38).



B

Membrana	Spot #	Peptídeo	Membrana	Spot #	Peptídeo
LALP-LP3	31	T-C-I-E-K-D-Y-V-N-I-H-K-G-D-R	LALP-LP8	28	E-A-M-R-S-V-D-I-H-R-G-P-K-C-Y
	32	E-K-D-Y-V-N-I-H-K-G-D-R-C-Y-S		37	L-S-L-G-K-G-C-T-E-F-G-T-I-L-H
	33	Y-V-N-I-H-K-G-D-R-C-Y-S-R-V-G		53	D-F-K-K-K-L-E-N-E-T-R-T-I-G-Y
	34	I-H-K-G-D-R-C-Y-S-R-V-G-K-S-F		54	K-K-L-E-N-E-T-R-T-I-G-Y-F-D-Y
	41	L-S-L-A-T-G-C-T-E-F-G-T-I-L-H		55	E-N-E-T-R-T-I-G-Y-F-D-Y-D-S-I
	42	A-T-G-C-T-E-F-G-T-I-L-H-E-L-G		56	T-R-T-I-G-Y-F-D-Y-D-S-I-M-L-Y
	49	H-E-H-S-R-S-D-R-D-E-Y-L-I-I-H		57	I-G-Y-F-D-Y-D-S-I-M-L-Y-G-P-Y
	50	S-R-S-D-R-D-E-Y-L-I-I-H-K-E-N		58	Y-F-D-Y-D-S-I-M-L-Y-G-P-Y-A-F
	56	T-K-E-H-N-F-D-K-L-W-E-N-R-T-R			
	57	H-N-F-D-K-L-W-E-N-R-T-R-T-I-G			
	58	D-K-L-W-E-N-R-T-R-T-I-G-N-F-D			
	59	W-E-N-R-T-R-T-I-G-N-F-D-Y-D-S			
	LALP-LP4	60	R-T-R-T-I-G-N-F-D-Y-D-S-I-M-L		
61		T-R-T-I-G-N-F-D-Y-D-S-I-M-L-Y			
23		Y-E-I-S-Q-A-L-R-P-F-K-Q-R-I-R			
30		E-D-H-K-G-D-K-C-Y-S-R-V-G-K-S			
37		P-L-S-L-G-K-G-C-T-E-F-G-T-I-L			
38		L-G-K-G-C-T-E-F-G-T-I-L-H-E-L			
41		T-I-L-H-E-L-G-H-S-V-G-F-S-H-E			
53		Q-H-D-F-K-K-K-L-E-N-E-T-R-T-I			
57		R-T-I-G-D-F-D-Y-D-S-I-M-L-Y-G			
58		G-D-F-D-Y-D-S-I-M-L-Y-G-P-Y-A			
59	D-Y-D-S-I-M-L-Y-G-P-Y-A-F-S-R				

Figura 38: Gráfico com a quantificação de coloração dos spots medido de LALP-LP3, LALP-LP4 e LALP-LP8 LP3 com soro anti rMEPLox de *L. intermedia* medidos por ImageJ. (A) Gráfico com quantificação de spots. Os que possuíam reatividade acima de 10.000 pixels/área foram considerados os principais epitopos e foram evidenciados por estrelas pretas. (B) Sequências dos pentadecapeptídeos marcados que possuíam reatividade em LALP-LP3, LALP-LP4 e LALP-LP8.

A proteína quimérica rMEPLox produzida por Lima e colaboradores em 2018 foi capaz de neutralizar a ação de metaloproteases e da ação dermonecrótica do veneno de *L. intermedia*, permitindo assim confirmar que os epitopos presentes na mesma são de fato epitopos importantes para ação contra o veneno desta aranha (LIMA et al., 2018).

Ao realizar o ensaio de SPOT com as três membranas LALP-LPs com o soro anti metaloprotease de *L. intermedia* foi possível identificar possíveis epitopos que, ao serem comparados com os três epitopos de rMEPLox em um alinhamento, permitiu detectar dois epitopos com grandes similaridades. Desta forma, para confirmar a identidade destes epitopos, um segundo ensaio de SPOT foi realizado, desta vez utilizando o soro anti rMEPLox, que constatou a importância destes dois epitopos previamente assinalados pelo alinhamento. Estes resultados iniciais podem colaborar para uma identificação precisa de epitopos de metaloproteases importantes no veneno de *Loxosceles*, e desta forma auxiliar na busca de novos tratamentos e métodos de prevenção.

Para o controle destes experimentos, a mesma metodologia foi utilizada para as três membranas com soro de coelhos que receberam apenas adjuvante de PBS (soro pré-

imune) (diluição 1:100), não sendo capaz de reconhecer nenhuma região das membranas (Figura 39).

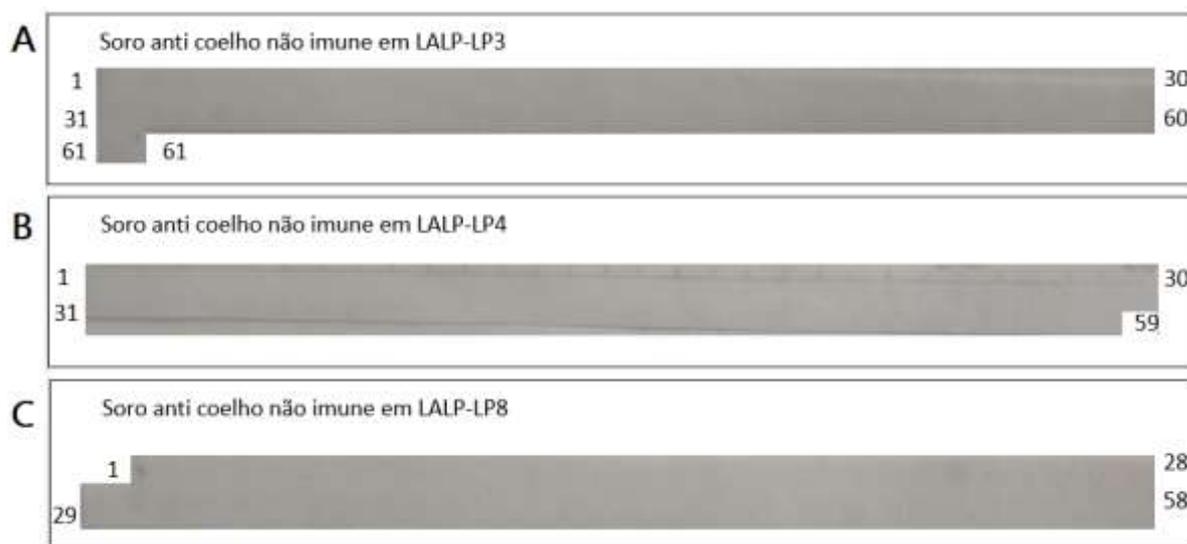


Figura 39: Controle negativo das membranas LALP-LP3, LALP-LP4 e LALP-LP8 no ensaio de SPOT. Membranas contendo peptídeos derivados das sequências (A) LALP-LP3, (B) LALP-LP4 e (C) LALP-LP8 resultantes da análise bioinformática deste projeto testadas com soro de coelhos não imunes.

Realizou-se então um alinhamento através do programa Jalview para a realização de uma comparação dos epitopos preditos tanto nos programas quanto nas membranas de spot (Figura 40). É possível observar que os dois epitopos de metaloproteases presentes tanto nas LALP-LP quanto na proteína quimera de Lima e colaboradores em 2018 foram também identificados pelos programas de predição de epitopos, confirmando-os como potenciais epitopos de metaloproteases do veneno de *L. laeta* (P).

Com os resultados deste alinhamento, modelagens tridimensionais foram realizadas das três LALP-LPs presentes nas membranas com marcações dos possíveis epitopos resultantes dos experimentos de SPOT com soro anti metaloprotease de *L. intermedia* (Figura 41) e soro anti quimera (Figura 42) utilizando o programa Jalview, onde novamente é possível observar uma identidade entre os epitopos dos dois ensaios, o que corrobora a confirmação de potenciais epitopos.

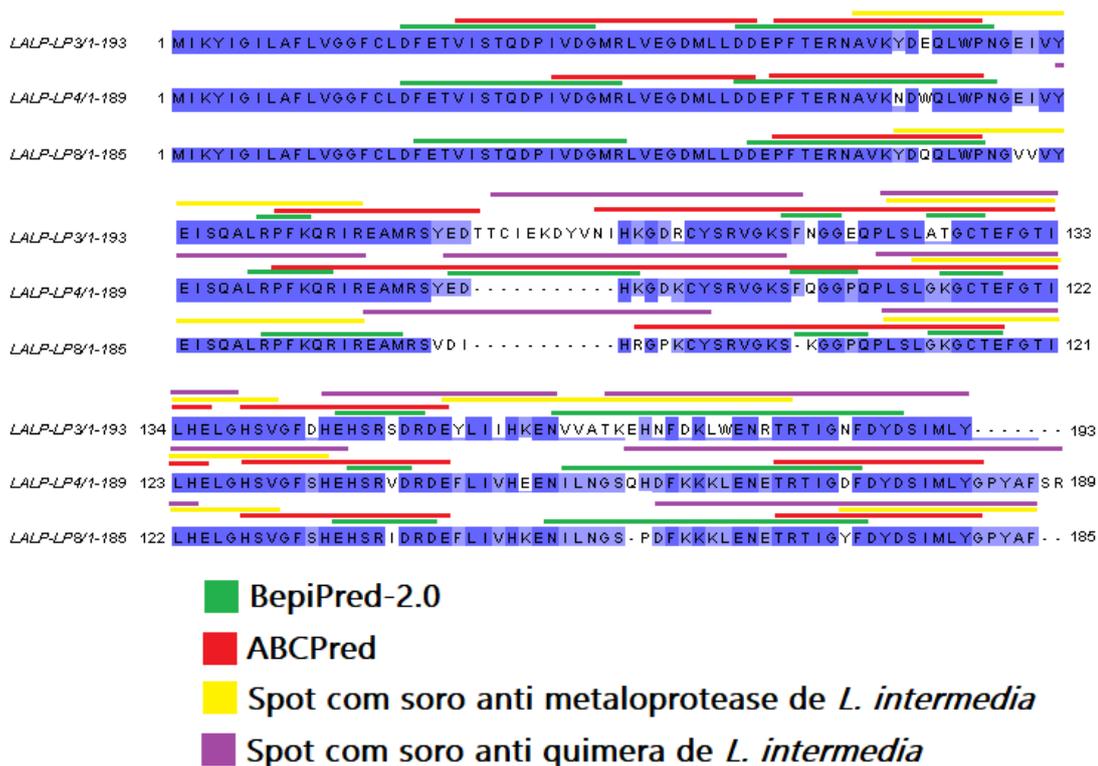


Figura 40: Alinhamento das três sequências de LALP-LPs com as marcações de prováveis epítopos identificados através de técnicas bioinformáticas e spot utilizando o programa Jalview.

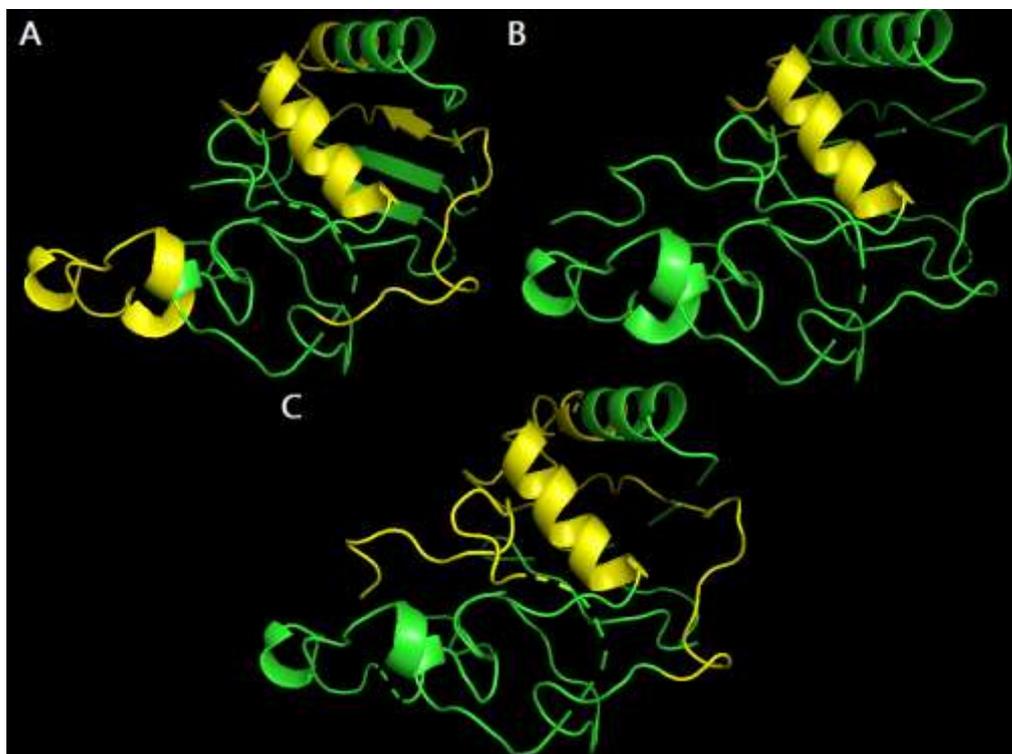


Figura 41: Modelagem das LALP-LPs identificadas por transcriptômica com marcação dos epítopos identificados na membrana de SPOT com soro anti metaloprotease de *L. intermedia* utilizando o programa PyMOL. (A) LALP-LP3; (B) LALP-LP4; (C) LALP-LP8.

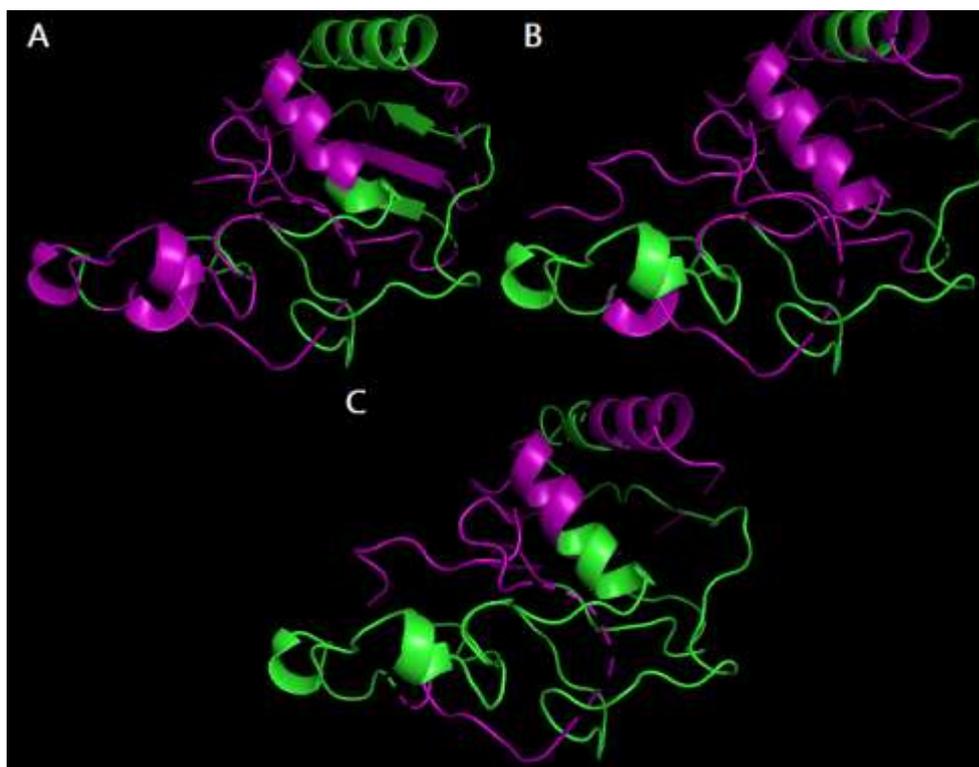


Figura 42: Modelagem das LALP-LPs identificadas por transcriptômica com marcação dos epitopos identificados na membrana de SPOT com soro anti quimera de *L. intermedia* utilizando o programa PyMOL. (A) LALP-LP3; (B) LALP-LP4; (C) LALP-LP8.

6. CONCLUSÕES

A anotação dos contigs identificados permitiu a determinação de nove proteoformas completas no transcriptoma de *L. laeta* peruana com similaridade a outras metaloproteases já identificadas em aranhas do gênero *Loxosceles*.

Os resultados obtidos sugerem que as estruturas primárias identificadas nas metaloproteases dos venenos de *L. laeta* (P) possuem diferenças importantes em relação as sequências das metaloproteases dos venenos brasileiros *L. laeta* (LALP4) e *L. intermedia* (LALP1).

A técnica de SPOT foi eficiente para mapear epítopos lineares células B das metaloproteases de *Loxosceles laeta* peruana.

7. PERSPECTIVAS

- Identificação e comparação das outras proteínas presentes na glândula de veneno de *L. laeta* peruana.
- Depositar as sequências identificadas neste trabalho em banco de dados.
- Produzir antivenenos utilizando como antígenos peptídeos sintéticos correspondentes aos epitopos.
- Realizar estudos de taxonomia molecular nas diferentes aranhas *L. laeta* da Argentina, Brasil, Chile e Peru.
- Validar a expressão por PCR das LALP-LPs e identificar o códon de parada.
- Buscar epitopos conservados em metalloproteases de outras espécies de *Loxosceles*.

8. BIBLIOGRAFIA

- ABREU, L. C. DE; BERNARDE, P. S.; MOTA DA SILVA, A. Acidentes Com Animais Peçonhentos No Brasil Por Sexo E Idade. **Journal of Human Growth and Development**, v. 25, n. 1, p. 54–62, 2015.
- ADAMS, M. J. et al. The structure of 6-phosphogluconate dehydrogenase refined at 2.5 Å resolution. **Acta Cryst**, v. B47, p. 817–820, 1991.
- ALVARENGA, É. R. DE; KALAPOTHAKIS, E. **Transcriptoma da glândula de veneno do escorpião *Tityus serrulatus* no contexto de catálogos de transcritos de quelicerados**. [s.l.: s.n.].
- ANDREWS, S. **FastQC: a quality control tool for high throughput sequence data**.
- AQUINO, P. A.; MARTINS, A. P. V.; CHÁVEZ-OLORTEGUI, C. **Diversidade molecular de Fosfolipases-D da peçonha da aranha *Loxosceles laeta* peruana revelada por sequenciamento de nova geração e análise transcriptômica**. [s.l.: s.n.].
- BODE, W.; GOMIS-RÜTH, F. X.; STÖCKLER, W. Astacins, serralysins, snake venom and matrix metalloproteinases exhibit identical zinc-binding environments (HEXXHXXGXXH and Met-turn) and topologies and should be grouped into a common family, the “metzincins”. **FEBS Letters**, v. 331, n. 1–2, p. 134–140, 1993.
- BOLGER, A. M.; LOHSE, M.; USADEL, B. Trimmomatic: A flexible trimmer for Illumina sequence data. **Bioinformatics**, v. 30, n. 15, p. 2114–2120, 2014.
- CARVALHO, L. S.; BONALDO, A. B.; BRESOVIT, A. D. first record family Cithaerionidae Gnaphosoidea) wor orld T he f ir st r ecor d of the f amil y Cithaer onidae (Ar aneae , Gna phosoidea) to the new w or ld. **Revista Brasileira de Zoologia**, v. 24, n. Ibsp 70223, p. 512–514, 2007.
- CHAIM, O. M. et al. Brown spider (*Loxosceles* genus) venom toxins: Tools for biological purposes. **Toxins**, v. 3, n. 3, p. 309–344, 2011.
- CHATZAKI, M. et al. Cutaneous loxoscelism caused by *Loxosceles similis* venom and neutralization capacity of its specific antivenom. **Toxicon**, v. 60, n. 1, p. 21–30, 2012.
- DA SILVA, P. et al. Brown spiders and loxoscelism. **Toxicon**, v. 44, p. 693–709, 2004.
- DANTAS, A. E. et al. Description of Loxtox protein family and identification of a new group of Phospholipases D from *Loxosceles similis* venom gland. **Toxicon**, v. 120, p. 97–106, 2016.
- ENDRULLAT, C. et al. Applied & Translational Genomics Standardization and quality management in next-generation sequencing. **ATG**, v. 10, p. 2–9, 2016.

- EWING, B. et al. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. **Genome Research**, v. 8, n. New York, p. 175–185, 1998.
- FEITOSA, L. et al. Detection and characterization of metalloproteinases with gelatinolytic, fibronectinolytic and fibrinogenolytic activities in Brown spider (*Loxosceles intermedia*) venom. **Toxicon**, v. 36, n. 7, p. 1039–1051, 1998.
- FELICORI, L. et al. An in vivo protective response against toxic effects of the dermonecrotic protein from *Loxosceles intermedia* spider venom elicited by synthetic epitopes. **Vaccine**, v. 27, n. 31, p. 4201–8, 2009.
- FERNANDES-PEDROSA, M. DE F. et al. Transcriptome analysis of *Loxosceles laeta* (Araneae, Sicariidae) spider venomous gland using expressed sequence tags. **BMC Genomics**, v. 9, p. 1–12, 2008.
- FOELIX, R. Biology of Spiders. **OUP USA/GEorg Thieme Verlag**, p. 330, 1996.
- FUTRELL, J. M. Loxoscelism. **Am. J. Med. Sci**, v. 304, p. 261–267, 1992.
- GONÇALVES-DE-ANDRADE, R. M.; VILARINHO TAMBOURGI, D. First record on *Loxosceles laeta* (Nicolet, 1849) (Araneae, Sicariidae) in the West Zone of São Paulo City, São Paulo, Brazil, and considerations regarding its geographic distribution. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 36, n. 3, p. 425–426, 2003.
- GUIMARÃES, G. et al. Biochemical and immunological characteristics of Peruvian *Loxosceles laeta* spider venom: Neutralization of its toxic effects by anti-loxoscelic antivenoms. **Toxicon**, v. 70, p. 90–97, 2013.
- HERNAN, S.; ZAVALA, A. **Panorama Epidemiológico del Loxoscelismo en el Peru** *Revista Peruana de Medicina Experimental y Salud Pública*, 1997.
- HOGAN, C. J.; BARBARO, K. C.; WINKEL, K. Loxoscelism: Old obstacles, new directions. **Annals of Emergency Medicine**, v. 44, n. 6, p. 608–624, 2004.
- ISBISTER, G. K.; FAN, H. W. Spider bite. **Lancet**, v. 378, n. 9808, p. 2039–2047, 2011.
- JUNIOR, E.; POLEGATTO, C. **Aspectos evolutivos em araneomorphae: arthropoda arachnida**. [s.l.] Centro Universitário Claretiano, 2008.
- KELLY, L. A. et al. The Phyre2 web portal for protein modelling, prediction, and analysis. **Nature Protocols**, v. 10, n. 6, p. 845–858, 2015.
- LIMA, S. DE A. et al. Recombinant protein containing B-cell epitopes of different *Loxosceles* spider toxins generates neutralizing antibodies in immunized rabbits. **Frontiers in Immunology**, v. 9, n. APR, 2018.
- MAGUINÃ-VARGAS, C. et al. Enfermedades por artrópodos. Parte I: Loxoscelismo

- cutáneo y cutáneo-visceral en el Perú. **Dermatología Peruana**, v. 14, p. 134–139, 2004.
- MAGUIÑA-VARGAS C. **Loxoscelismo viscerohemolítico**.
- MARÇAIS, G.; KINGSFORD, C. A fast, lock-free approach for efficient parallel counting of occurrences of k-mers. **Bioinformatics**, v. 27, n. 6, p. 764–770, 2011.
- MINISTERIO DE SAUDE, B. **Manual de Diagnóstico e Tratamento de Acidentes por Animais Peçonhentos**. [s.l.: s.n.].
- MODICA, M. V. et al. The venomous cocktail of the vampire snail *Colubraria reticulata* (Mollusca, Gastropoda). **BMC Genomics**, v. 16, n. 1, p. 1–21, 2015.
- MÖHRLÉN, F. et al. Evolution of astacin-like metalloproteases in animals and their function in development. **Evolution and Development**, v. 8, n. 2, p. 223–231, 2006.
- MORGON, A. M. et al. Expression and immunological cross-reactivity of LALP3, a novel astacin-like metalloprotease from brown spider (*Loxosceles intermedia*) venom. **Biochimie**, v. 128–129, p. 8–19, 2016.
- OLIVEIRA, S.; SILVA, C. **Levantamento da artropodofauna de um fragmento de mata e de uma pastagem, localizados próximo ao campus da universidade do estado de Mato Grosso, Tangará da Serra - MT**. [s.l.: s.n.].
- OSTELL, J. **The NCBI Handbook**.
- RIBEIRO, M. F. et al. Pattern of inflammatory response to *Loxosceles intermedia* venom in distinct mouse strains: A key element to understand skin lesions and dermonecrosis by poisoning. **Toxicon**, v. 96, p. 10–23, 2015.
- RUPPERT, E.; BARNES, D. **Zoologia dos Invertebrados**. São Paulo: Ed. Roca, v. 6 ed, p. 1028, 1996.
- SAHA, S.; RAGHAVA, G. P. S. Prediction of Continuous B-cell Epitopes in an Antigen Using Recurrent Neural Network. **Proteins**, v. 65, n. 1, p. 40–48, 2006.
- SILVEIRA, A. L. First synanthropic record of *Loxosceles laeta* (Nicolet, 1849) (Araneae, Sicariidae) in the municipality of Rio de Janeiro, State of Rio de Janeiro. **Revista da Sociedade Brasileira de Medicina Tropical**, v. 42, n. 6, p. 723–726, 2009.
- SPIRO, R. G. Protein Glycosylation: nature, distribution, enzymatic formation and disease implications of glycopeptide bonds. **Glycobiology**, v. 12, n. 4, p. 43R–56R, 2002.
- SURGET-GROBA, Y.; MONTOYA-BURGOS, J. I. Optimization of de novo transcriptome assembly from next-generation sequencing data. **Genome Research**, v. 20, n. 10, p. 1432–1440, 2010.
- SWANSON, D. L.; VETTER, R. S. Loxoscelism. **Clinics in Dermatology**, v. 24, n. 3,

p. 213–221, 2006.

TREVISAN-SILVA, D. et al. Astacin-like metalloproteases are a gene family of toxins present in the venom of different species of the brown spider (genus *Loxosceles*).

Biochimie, v. 92, n. 1, p. 21–32, 2010.

TREVISAN-SILVA, D. et al. Differential metalloprotease content and activity of three *Loxosceles* spider venoms revealed using two-dimensional electrophoresis approaches.

Toxicon, v. 76, p. 11–22, 2013.

VEIGA, S. S. et al. Oligosaccharide residues of *Loxosceles intermedia* (brown spider) venom proteins: dependence on glycosilation for dermonecrotic activity. **Toxicon**, v. 37, p. 587–607, 1999.

VETTER, R. S.; VISSCHER, P. K.; D, P. Bites and stings of medically important venomous arthropods. **International Journal of Dermatology**, v. 37, p. 481–496, 1998.

WANG, Z.; GERSTEIN, M.; SNYDER, M. RNA-Seq: a revolutionary tool for transcriptomics Zhong. **Nat Rev Genet**, v. 10, n. 1, p. 57–63, 2009.

WATERHOUSE, A. M. et al. Jalview version 2: A Multiple Sequence Alignment and Analysis Workbench. **Bioinformatics**, v. 25, n. 9, p. 1189–1191, 2009.

WEN, F. H. **Evolução Loxoscelismo cutâneo.**

ZERBINO, D.R.; BIRNEY, E. Velvet: algorithms for de novo short read assembly using de Bruijn graphs. **Genome Research**, v. 18, n. New York, p. 821–829, 2008.

ZHOU, Q.; SU, X.; NING, K. Assessment of quality control approaches for metagenomic data analysis. **Scientific Reports**, v. 4, n. 4, p. 1–11, 2014.