

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Instituto de Ciências Biológicas
Programa de Pós-graduação em Bioquímica e Imunologia

Manuela Cristina Emiliano Ferreira

**ESTUDO DE REPERTÓRIO SOROLÓGICO DE ANTICORPOS DE CAVALO
CONTRA VENENO DE ARANHAS DO GÊNERO *LOXOSCELES* DE
IMPORTÂNCIA MÉDICA NO BRASIL**

Belo Horizonte
2022

Manuela Cristina Emiliano Ferreira

**ESTUDO DE REPERTÓRIO SOROLÓGICO DE ANTICORPOS DE CAVALO
CONTRA VENENO DE ARANHAS DO GÊNERO *LOXOSCELES* DE
IMPORTÂNCIA MÉDICA NO BRASIL**

Dissertação apresentada ao Programa de Pós-graduação em Bioquímica e Imunologia da Universidade Federal de Minas Gerais, como requisito parcial para a obtenção do grau de Mestra em Bioquímica e Imunologia.

Orientadora: Profa. Dra. Liza Figueiredo Felicori Vilela
Coorientadora: Dra. Marcella Nunes de Melo-Braga

**Belo Horizonte
2022**

043

Ferreira, Manuela Cristina Emiliano.

Estudo de repertório sorológico de anticorpos de cavalo contra veneno de aranhas do gênero *Loxosceles* de importância médica no Brasil [manuscrito] / Manuela Cristina Emiliano Ferreira. – 2022. 81 f. : il. ; 29,5 cm.

Orientador: Liza Figueiredo Felicori Vilela. Coorientador: Marcella Nunes de Melo-Braga.

Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Bioquímica e Imunologia.

1. Bioquímica e Imunologia. 2. Cavalos. 3. Imunoglobulinas. 4. Venenos de Aranha. I. Vilela, Liza Figueiredo Felicori. II. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. III. Título.


CDU: 577.1



ATA DA DEFESA DA DISSERTAÇÃO DE MESTRADO DE MANUELA CRISTINA EMILIANO FERREIRA. Aos trinta dias do mês de agosto de 2022 às 14:00 horas, reuniu-se de forma “on line” utilizando a plataforma “Zoom” e presencialmente na Sala 301 no CAD1, da Universidade Federal de Minas Gerais, a Comissão Examinadora da dissertação de Mestrado, indicada *ad referendum* do Colegiado do Curso, para julgar, em exame final, o trabalho intitulado "Estudo de repertório sorológico de anticorpos de cavalo contra veneno de aranhas do gênero *Loxosceles* de importância médica no Brasil", requisito final para a obtenção do grau de Mestre em Bioquímica e Imunologia, área de concentração: Bioquímica. Abrindo a sessão, a Presidente da Comissão, Profa. Liza Figueiredo Felicori Vilela, da Universidade Federal de Minas Gerais, após dar a conhecer aos presentes o teor das Normas Regulamentares do Trabalho Final, passou a palavra à candidata para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores, com a respectiva defesa da candidata. Logo após a Comissão se reuniu, sem a presença da candidata e do público, para julgamento e expedição do resultado final. Foram atribuídas as seguintes indicações: Dr. Alexandre Keiji Tashima (Universidade Federal de São Paulo - UNIFESP), aprovada; Dra. Mariana Torquato Quezado de Magalhães (Universidade Federal de Minas Gerais), aprovada; Dra. Marcella Nunes de Melo Braga - Coorientadora (Universidade Federal de Minas Gerais), aprovada; Dra. Liza Figueiredo Felicori Vilela - Orientadora (Universidade Federal de Minas Gerais), aprovada. Pelas indicações a candidata foi considerada:

- APROVADA
 REPROVADA


O resultado final foi comunicado publicamente à candidata pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente da Comissão encerrou a reunião e lavrou a presente Ata que será assinada por todos os membros participantes da Comissão Examinadora. Belo Horizonte, 30 de agosto de 2022.


 Dr. Alexandre Keiji Tashima (Universidade Federal de São Paulo - UNIFESP)

Dra. Mariana Torquato Quezado de Magalhães (UFMG)



Documento assinado digitalmente
 MARIANA TORQUATO QUEZADO DE MAGAL
 Data: 12/09/2022 11:29:43-0300
 Verifique em <https://verificador.itl.br>


 Dra. Marcella Nunes de Melo Braga - Coorientadora (UFMG)


 Dra. Liza Figueiredo Felicori Vilela - Orientadora (UFMG)

Agradecimentos

À CAPES pela bolsa de pesquisa concedida e aos apoios financeiros do CNPq e da FAPEMIG.

Ao Instituto de Ciências Biológicas e ao Programa de Pós Graduação em Bioquímica e Imunologia, agradeço pela oportunidade de ter desenvolvido o projeto de mestrado em um programa de excelência.

Agradeço com muito carinho a todos que contribuíram direta ou indiretamente para a construção desse trabalho:

À minha orientadora, Liza, por sua dedicação e sensibilidade e por compartilhar toda a sua empolgação com a pesquisa científica. À minha coorientadora, Marcella, por toda a parceria nos desafios da bancada, das análises, da escrita e por seu carinho e bom humor infinitos.

A todos os colegas de laboratório pela companhia na rotina de experimentos e por todas as contribuições nos momentos de discussão de artigos e resultados. À Marcele pela companhia nos dias estranhos de pandemia e por todas as conversas motivacionais. À Milene e ao Fábio pela ajuda nos desafios da programação. À Carlana por todas as trocas na descoberta dos caminhos para entender os anticorpos de cavalo. Aos técnicos do departamento pela disponibilidade em ajudar em diversos momentos, em especial ao Jamil e ao Elimar.

À toda minha família pelo apoio e respeito na escolha pelo caminho da pesquisa especialmente à minha mãe, Jorginho e Laura, por serem meu porto seguro nesse momento e em todos os outros. Às amigas que estiveram pertinho para os momentos de comemoração e de dúvida. Ao meu tio por todo o suporte para que eu pudesse me dedicar àquilo que gosto de fazer. Ao Clóvis por ser o melhor companheiro sempre e por trazer tantas coisas boas para os meus dias.

Resumo

Os acidentes com aranhas do gênero *Loxosceles* (ou aranhas-marrom) têm incidência crescente no Brasil e podem desencadear lesões dermonecroticas severas, falência renal e óbito. A soroterapia é o único tratamento específico à disposição e consiste em fragmentos de anticorpos de cavalos hiperimunizados com o veneno das aranhas. O uso dessa terapia pode gerar reações adversas e sua produção depende da extração do veneno das aranhas para imunização dos equinos com moléculas tóxicas que impactam na sobrevivência desses animais. Diversas tentativas têm sido exploradas para contornar essas limitações, mas o desconhecimento das características de anticorpos de cavalo limita a busca por novas soluções. Neste trabalho, o repertório de anticorpos sorológico de quatro cavalos hiperimunizados com veneno de *L. intermedia*, *L. gaucho* e *L. laeta* foi caracterizado quanto às classes e subclasses de imunoglobulinas e quanto às sequências dos anticorpos que o compõem. Plasmas equinos antes e após hiperimunização foram usados para quantificação das classes de imunoglobulinas e subclasses de IgG e em sequência para isolamento de IgG, obtenção de $F(ab')_2$, seleção de $F(ab')_2$ específicas contra *pool* de venenos das três espécies ou contra o principal componente do veneno, a fosfolipase D e digestão por tripsina. Os peptídeos trípticos de $F(ab')_2$ específicas foram analisados por LC-MS/MS e mapeados às sequências de cadeia pesada imunoglobulina do transcriptoma de PBMCs de cada animal, utilizando o CDRH3 como parâmetro para identificação. Quanto às classes de imunoglobulinas, a mediana da concentração de IgM foi de 0,63 e de IgG de 241,42 mg/mL nos plasmas hiperimunes. Antes do isolamento com proteína G, a subclasse IgG(T) foi encontrada com maior abundância nos plasmas e após o isolamento nota-se predomínio de IgGb. A partir dos dados de LC-MS/MS, foram identificadas 132 sequências da região variável da cadeia pesada específicas contra o *pool* de venenos, 19 contra rLiD1 e 12 sequências foram identificadas com as duas especificidades. Nessas sequências, os segmentos gênicos IGHV4-29, IGHD2-26 e IGHJ6 foram mais usados, os CDRH3 apresentaram maior frequência de resíduos carregados em relação ao repertório inespecífico e tinham comprimento mediano de 17 aminoácidos. O cálculo da abundância a partir da área sob a curva dos peptídeos revelou uma resposta polarizada com mais de 80% da abundância concentrada nos top 3 e top 5 clones de imunoglobulinas específicos contra o *pool* de venenos e rLiD1, respectivamente. A predição de paratopos destes anticorpos contra rLiD1 indica alta convergência, com quatro possíveis epítomos reconhecidos pelas 31 sequências. Pelo nosso conhecimento, esse é o primeiro trabalho a utilizar o Ig-Seq na caracterização de anticorpos equinos. A descoberta de sequências da região variável da cadeia pesada de imunoglobulinas e suas características podem ser aplicadas na produção e otimização de anticorpos sintéticos para uso terapêutico. Palavras-chave: imunoglobulinas; cavalos; veneno; proteômica; *Loxosceles*.

Abstract

The accidents caused by spiders of the *Loxosceles* genus (also called brown-spiders) are rising in incidence in Brazil and may cause severe dermonecrotic lesions, renal failure and death. Serotherapy is the only specific treatment available and is based on antibodies from horses hyperimmunized with the spiders' venom. The use of this therapy may cause adverse reactions and its production relies on the extraction of the venom from the spiders for the immunization of the horses, a process that impacts negatively on the life of the inoculated animals. Many attempts to overcome these limitations have been explored, but the lack of knowledge about the characteristics of horses' immunoglobulins limits the search for new solutions. In this work, the repertoire of serological antibodies of four horses hyperimmunized with venom from *L. intermedia*, *L. gaucho* and *L. laeta* was characterized relative to its classes and subclasses and to the sequences that compose them. Equine plasmas collected before and after the immunization were used for the quantification of immunoglobulin classes and IgG subclasses and, after that, for isolation of IgG, obtention of F(ab')₂, selection of F(ab')₂ fractions specific against the pool of venoms or against rLiD1 and trypsin digestion. Specific F(ab')₂ tryptic peptides were analyzed by LC-MS/MS and mapped to the immunoglobulin heavy chain sequences of the transcriptome of PBMCs of each animal, using CDRH3 as a parameter for identification. As for the immunoglobulin classes, a median concentration of 0.63 for IgM and of 241.42 mg/mL for IgG were found in hyperimmune plasmas. Before isolation using protein G, the IgG(T) subclass was found with greater abundance in plasmas and after isolation, IgGb predominated. From the LC-MS/MS data, 132 heavy chain variable region sequences specific against the pool of venoms were identified, as well as 19 against rLiD1 and 12 with both specificities. In these sequences, the IGHV4-29, IGHD2-26 and IGHJ6 gene segments were more used, the CDRH3 had a higher frequency of charged residues compared to the nonspecific repertoire and had a median length of 17 amino acids. Calculation of abundance from the area under the curve of the peptides revealed a polarized response with more than 80% of the abundance concentrated in the top 3 and top 5 clones of specific immunoglobulins against the pool of venoms and rLiD1, respectively. The prediction of paratopes against rLiD1 indicates high convergence, with four possible epitopes recognized by the 31 sequences. To our knowledge, this is the first work to use Ig-Seq for the characterization of equine antibodies. The discovery of sequences of the variable region of the heavy chain of immunoglobulins and its characteristics may be useful in the production and optimization of synthetic antibodies designed for therapeutical applications.

Keywords: immunoglobulins; horses; venom; proteomics; *Loxosceles*.

Lista de Figuras

Figura 1 - Evolução da lesão dermonecrotica de paciente acidentado com aranha-marrom.....	14
Figura 2 - Etapas de produção de soroterapia anti-veneno em equinos.....	16
Figura 3 - Estrutura de imunoglobulina.....	20
Figura 4 - Complementariedade das abordagens do Ig-Seq e do Rep-Seq para a descoberta do repertório sorológico de imunoglobulinas.....	26
Figura 5 - Etapas do sequenciamento <i>high throughput</i> e de proteômica para obtenção das imunoglobulinas antígeno-específicas.....	27
Figura 6 - Preparo das amostras para o Ig-Seq.....	34
Figura 7 - Fluxo de trabalho da construção do banco de sequências para busca de peptídeos encontrados na espectrometria de massa.....	37
Figura 8 - Identificação de sequências de imunoglobulina a partir de peptídeos detectados na espectrometria de massa.....	38
Figura 9 - Imunorreatividade dos plasmas equinos hiperimunizados com o <i>pool</i> de veneno de três espécies de <i>Loxosceles</i>	43
Figura 10 - Quantificação de IgM e IgG nos plasmas equinos.....	45
Figura 11 - Distribuição de subclasses de IgG nos plasmas equinos pré-ímmes.....	46
Figura 12 - Títulos das subclasses de IgG presentes nos plasma hiperímmes e na cromatografia de afinidade com proteína G.....	47
Figura 13 - Identificação de peptídeos trípticos nas amostras de frações F(ab') ₂ antígeno-específicas.....	50
Figura 14 - Peptídeos mapeados por região da cadeia variável do anticorpo.....	51
Figura 15 - Frequência de resíduos de lisina (K) e arginina (R) na posição 107 do CDRH3 e de peptídeos trípticos por região da fração variável do IGH.....	52
Figura 16 - Composição das 30 sequências da região variável de IGH com CDRH3 em comum encontradas na amostra 1H específica contra o <i>pool</i> de venenos.....	54
Figura 17 - Sequências da região variável da IGH identificadas a partir do CDRH3.....	54
Figura 18 - Distribuição do comprimento e dos aminoácidos das sequências de CDRH3 encontradas no Ig-Seq e no Rep-Seq.....	56
Figura 19 - Composição de aminoácidos por posição nos CDRH3 antígeno-específicos não redundantes.....	57
Figura 20 - Preferência dos segmentos gênicos IGH da cadeia variável dos anticorpos.....	59
Figura 21 - Frequência de hipermutação somática na região variável da cadeia pesada das imunoglobulinas encontradas nos diferentes repertórios estudados.....	60
Figura 22 - Frequência dos clonotipos das sequências da região variável IGH antígeno-específicas.....	61
Figura 23 - Similaridade de paratopos entre as sequências da região variável de IGH específicas contra a rLiD1.....	63
Figura 24 - Distribuição de resíduos de aminoácidos por posição das 24 sequências de paratopos similares contidas no cluster 1.....	64

Lista de tabelas

Tabela 1 - Filtros aplicados para construção do banco de sequências de cada amostra.....	48
Tabela 2 - Número de peptídeos, CDRH3 e sequências IGHV identificadas no Ig-Seq.....	53
Tabela 3 - Abundância de clonotipos no Ig-Seq e no Rep-Seq.....	62

Lista de abreviaturas e siglas

BCR: Receptor de célula B

BSA: Albumina sérica bovina

C: Celsius

CDR: Região determinante de complementaridade

CENADI: Centro Nacional para Distribuição e Armazenamento de Imunobiológicos

CEUA: Comissão de Ética no Uso de Animais

CMV: Citomegalovírus humano

CPPI: Centro de Pesquisa e Produção de Imunobiológicos

csv: Formato de arquivo separado por vírgulas (do inglês, *comma-separated values*)

D: Segmento de diversidade

Da: Dalton

DDA: Análise dependente de dados

EDTA: Ácido etilenodiamino tetra-acético

ELISA: Ensaio de Imunoabsorção Enzimática

F(ab')₂: Frações F(ab) de imunoglobulinas conectadas por ligação dissulfeto

Fab: Fração de ligação ao antígeno de imunoglobulina

Fc: Fração cristalizável de imunoglobulinas

FDR: Taxa de falsas de descoberta (do inglês, *false discovery rates*)

FR: Região de framework

FT: Não interagido (do inglês, *flow through*)

FUNED: Fundação Ezequiel Dias

FWHM: Largura a meia altura (do inglês, *full width at half maximum*)

HBV: Vírus da hepatite B

HCD: Colisão em alta energia (HCD, do inglês *higher-energy C-trap dissociation*)

HMS: Hipermutação somática

HPLC: Cromatografia líquida de alta eficiência (do inglês, *high performance liquid chromatography*)

I₂MS: Espectrometria de massa de íons individuais (do inglês, *individual ions mass spectrometry*)

IgA: Imunoglobulina de classe A

IgG: Imunoglobulina de classe G

IGH: Cadeia pesada de imunoglobulina

IGHD: Segmento gênico de diversidade da cadeia pesada de imunoglobulina

IGHJ: Segmento gênico de junção da cadeia pesada de imunoglobulina

IGHV: Segmento gênico variável da cadeia pesada de imunoglobulina ou região variável da cadeia pesada de imunoglobulina

IgM: Imunoglobulina de classe M

IMGT: *International ImMunoGeneTics Information System*

INCQS: Instituto Nacional de Controle de Qualidade em Saúde

J: Segmento gênico de junção

kDa: Quilodalton

LC: Cromatografia líquida

M: Molar

m/z: Relação massa/carga

mAb: Anticorpo monoclonal (do inglês, *monoclonal antibody*)

mg: Miligrama(s)

mL: Mililitro(s)

mM: Milimolar(es)

MS: Espectrometria de massa

ng: Nanograma(s)

NGS: Sequenciamento de nova geração (do inglês, *next generation sequencing*)

nm: Nanômetro(s)

OPD: o-fenilenodiamina

ORFs: Fase de leitura aberta (do inglês, *open reading frame*)

PDB: *Protein Data Bank*

PBS: Tampão fosfato-salino

PEP: Probabilidade de erro posterior

ppm: Parte por milhão

PR A/B: Receptor de progesterona A/B

PSM: *Peptide spectrum matches*

rLiD1: Fosfolipase 1 de *Loxosceles intermedia* recombinante

SaLox: Soro anti-loxoscélico

scFv: Fragmento variável de cadeia única (do inglês, *single-chain fragment variable*)

TA: Temperatura ambiente

TMB: 3,3',5,5'-Tetrametilbenzidina

V: Segmento gênico variável

VH: Região variável da cadeia pesada

µg: Micrograma(s)

µL: Microlitro(s)

Sumário

1. Introdução	13
1.1. Acidentes com aranhas do gênero <i>Loxosceles</i>	13
1.2. Soroterapia anti-veneno	15
1.3. Imunoglobulinas	18
1.3.1. Imunoglobulinas equinas.....	18
1.4. Estratégias para estudo de repertório de anticorpos.....	22
1.4.1. Avaliando o repertório de receptores de células B.....	20
1.4.2. Estudos de repertório baseados em proteômica <i>middle</i> ou <i>top-down</i>	21
1.4.3. Estudos de repertório baseados em proteômica <i>bottom-up</i>	22
1.4.3.1. Metodologia do Ig-Seq.....	24
2. Objetivos	30
2.1. Objetivo geral.	30
2.2. Objetivos específicos.....	30
3. Metodologia	31
3.1. Hiperimunização dos equinos e coleta dos plasmas.....	31
3.2. <i>Pool</i> de venenos de <i>Loxosceles</i> e rLiD1	31
3.3. Reatividade e quantificação de imunoglobulinas nos plasmas equinos	32
3.3.1. Reatividade dos plasmas contra <i>pool</i> de venenos e rLiD1.....	30
3.3.2. Quantificação absoluta das classes de imunoglobulinas.....	30
3.3.3. Quantificação relativa das subclasses de IgG.....	31
3.4. Obtenção de fragmentos trípticos de F(ab') ₂ específicos contra o veneno de <i>Loxosceles</i> e rLiD1.....	33
3.5. Obtenção de sequências de peptídeos trípticos derivados das frações F(ab') ₂ de anticorpos.....	34
3.6. Preparo do banco de sequências para análise proteômica.....	35
3.7. Análise dos dados de espectrometria de massa para identificação dos peptídeos correspondentes à região IGHV	37

3.8. Mapeamento dos peptídeos identificados aos CDRs das sequências de anticorpos equinos.....	37
3.9. Alinhamento e comparação de sequências	39
3.10. Determinação de clonotipos e cálculo de abundância a partir dos peptídeos	40
3.11. Agrupamento de sequências de VH contendo paratopos similares	41
4. Resultados	42
4.1. Reatividade dos plasmas equinos contra <i>pool</i> de venenos de <i>Loxosceles</i> ou rLiD1 .	42
4.2. Quantificação absoluta das classes de imunoglobulinas	44
4.3. Quantificação relativa das subclasses de IgG	45
4.4. Composição do banco de sequências para busca.....	48
4.5. Análise dos dados do Ig-Seq para identificação dos peptídeos mapeados à região variável de IGH.....	49
4.6. Análise das sequências da região variável de IGH identificadas a partir do CDRH3	52
4.7. Tamanho e composição dos CDRH3 identificados	55
4.8. Preferência de uso de genes V(D)J e frequência de hipermutação somática	58
4.9. Identificação de clonotipos e cálculo de abundância	60
4.10. Similaridade de paratopos das sequências específicas contra rLiD1	63
5. Discussão	66
6. Conclusão	71
7. Referências bibliográficas	72

1. Introdução

1.1. Acidentes com aranhas do gênero *Loxosceles*

Os envenenamentos causados por picada de aranha são eventos de notificação obrigatória do Sistema Único de Saúde (SUS) brasileiro cuja incidência vem crescendo no país desde o ano de 2015. No Brasil, o araneísmo é marcado por três gêneros de artrópodes de importância médica: *Latrodectus*, conhecido popularmente como viúva-negra, *Phoneutria* ou aranha-armadeira e *Loxosceles* ou aranha-marrom, grupo que concentra as espécies associadas a maior parte dos casos. *Loxosceles laeta*, *Loxosceles gaucho* e *Loxosceles intermedia* provocam em média 8.000 acidentes anuais contra média de 4.000 acidentes associados ao gênero *Phoneutria*, 129 atribuídos a *Latrodectus* e 158.656 casos de araneísmo cuja a espécie causadora não foi identificada. Os acidentes com *Loxosceles* se distribuem de forma heterogênea no Brasil: cerca de 60% dos casos são detectados na região sul e 30% na região sudeste (MINISTÉRIO DA SAÚDE, 2022).

A picada por aranha-marrom caracteriza-se por ser indolor e ocorrer com maior frequência pela compressão do animal contra o corpo do indivíduo. Os sinais e sintomas iniciam-se tipicamente algumas horas após o acidente e em conjunto compõem as síndromes do loxoscelismo (FUTRELL, 1992). No local da picada, podem ser observados inchaço, vermelhidão e coceira, além de úlcera que pode evoluir para lesão dermonecrotica, sinal característico do loxoscelismo cutâneo (Figura 1). Enquanto a forma cutânea afeta mais de 70% dos acidentados, estima-se que a incidência do loxoscelismo víscero-cutâneo varia entre 0,7%-27%, dependendo da localidade e da espécie de aranha-marrom causadora da picada (DE OLIVEIRA et al., 2005; GREMSKI et al., 2014). Na forma víscero-cutânea, além dos sintomas já citados são detectadas hemólise intravascular, trombocitopenia e coagulação intravascular disseminada que podem desencadear falência renal e óbito (FUTRELL, 1992; SWANSON; VETTER, 2006).

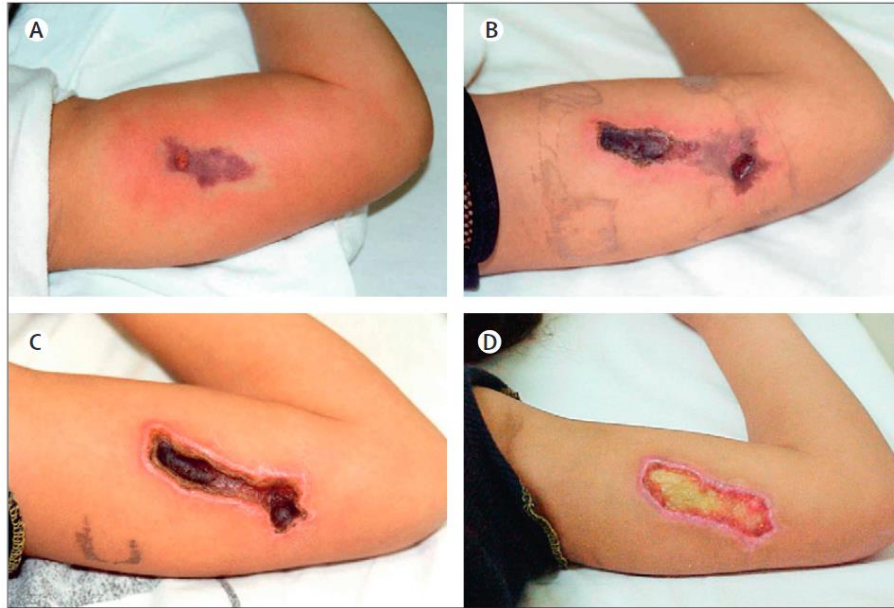


Figura 1 – Evolução da lesão dermonecrotica de paciente acidentado com aranha-marrom. Progressão de úlcera e necrose no local da picada nos dias 1 (A), 9 (B), 16 (C) e 25 (D) após o acidente. (ISBISTER; FAN, 2011).

Os sinais e sintomas são resultado da ação sinérgica dos diversos componentes dos venenos de *Loxosceles*. O estudo do transcriptoma da glândula de veneno de *L. intermedia* revela que as moléculas mais abundantes são da classe dos peptídeos inseticidas, seguidas pelas metaloproteases e pelas fosfolipases D (GREMSKI et al., 2010), enquanto no transcriptoma de *L. laeta*, as fosfolipases ocuparam a posição de maior abundância, seguidas pelas metaloproteases (FERNANDES-PEDROSA et al., 2008). Além dessas proteínas, serino-proteases, neurotoxinas e hialuronidases são encontradas em menor frequência. Dentre as classes de proteínas citadas, as fosfolipases D (também denominadas esfingomielinases D, toxinas dermonecroticas ou proteínas LoxTox) destacam-se pela alta imunogenicidade e pela capacidade de replicar os sinais do envenenamento induzindo o surgimento da lesão dermonecrotica e de alterações vasculares (CHAIM et al., 2011; GREMSKI et al., 2014). A imunização de animais com fosfolipase D gera soro anti-veneno capaz de proteger contra a evolução da lesão dermonecrotica, com proteção cruzada parcial contra o envenenamentos causados por espécies de *Loxosceles* distintas da espécie de origem da fosfolipase (DE ALMEIDA et al., 2008; DUARTE et al., 2015; FINGERMANN et al., 2020).

Os pacientes acidentados que passam a manifestar sintomas são diagnosticados principalmente pela avaliação clínica e epidemiológica (FUTRELL, 1992). Para os casos leves e moderados, apenas o manejo dos sintomas e o uso de anti-inflamatórios é recomendado. Para os casos graves, além dessas intervenções, preconiza-se o uso de soroterapia anti-veneno, o único tratamento específico para neutralizar os efeitos dos componentes do veneno a nível sistêmico a disposição (MINISTÉRIO DA SAÚDE, 2001).

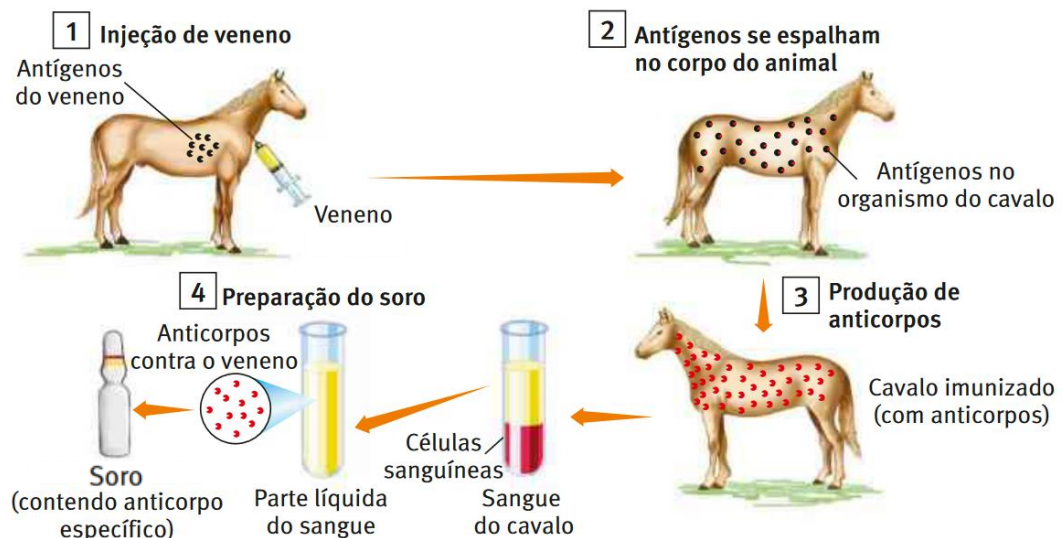
1.2. Soroterapia anti-veneno

A soroterapia vem sendo usada como tratamento para o envenenamento desde 1894, ano do desenvolvimento do soro antiofídico por Césaire Phisalix e Albert Calmette (BOCHNER, 2016). No Brasil, essa inovação foi explorada por Vital Brazil Mineiro da Campanha, que iniciou em 1898 os estudos para a produção de soros específicos contra o veneno de serpentes brasileiras (para revisão SQUAIELLA-BAPTISTÃO et al., 2018). Atualmente no país, os soroterápicos, são gerados tipicamente a partir do sangue de cavalos hiperimunizados com o antígeno contra o qual a proteção é desejada (MIRANDA, 2021).

Atualmente, quatro instituições públicas produzem os soros anti-venenos que são distribuídos nacionalmente: o Instituto Butantan, o Centro de Pesquisa e Produção de Imunobiológicos (CPPI), o Instituto Vital Brazil e a Fundação Ezequiel Dias (FUNED) (MIRANDA et al., 2021). Os imunobiológicos produzidos incluem terapias para os principais animais peçonhentos do país como as serpentes dos gêneros *Bothrops* (soro antibotrópico), *Lachesis* (soro antibotrópico e antilaquéutico), *Crotalus* (soro anticrotálico ou soro antibotrópico e anticrotálico) e *Micrurus* (soro anti-elapídico), todos produzidos pelo Butantan; escorpiões dos gêneros *Tityus* (soros anti-escorpiônicos do Butantan ou da FUNED) e aranhas dos gêneros *Latrodectus* (soro anti-latrodético, do Instituto Vital Brazil), *Phoneutria* (soro anti-aracnídico do Instituto Butantan, gerado contra o veneno de *Loxosceles* spp., *Phoneutria* spp. e *Tityus* spp.) e *Loxosceles* (soro anti-loxoscélico do CPPI, gerado contra os veneno de *L. intermedia*, *L. gaucho* e *L. laeta*) (MINISTÉRIO DA SAÚDE, 2001). As ampolas produzidas têm sua qualidade assegurada pelo Instituto Nacional de Controle de Qualidade em Saúde (INCQS) e são distribuídas às secretárias de saúde estaduais por intermédio do Centro Nacional para Distribuição e Armazenamento de

Imunobiológicos (CENADI) de acordo com as demandas regionais (GUTIÉRREZ et al., 2009).

A produção dos soros anti-venenos tem como principais etapas: i) hiperimunização de equinos com veneno e adjuvante, ii) sangria, iii) fracionamento do plasma, iv) digestão por pepsina ou papaína e v) precipitação por ácido caprílico ou sulfato de amônio (Figura 2). A partir desses passos são obtidas as frações $F(ab')_2$ das imunoglobulinas equinas encontradas no plasma que serão administradas a pacientes com manifestações moderadas ou graves de loxoscelismo.



Fonte: Governo do Rio de Janeiro, Secretaria de Saúde. Instituto Vital Brazil. Disponível em: <http://www.vitalbrazil.rj.gov.br/etapas_producao.html>. Acesso em: mar. 2016.

Figura 2 – Etapas de produção de soroterapia anti-veneno em equinos. As etapas de produção do soro anti-veneno são exemplificadas. O veneno de interesse é inoculado no animal, anticorpos são produzidos em resposta a componentes do veneno após diversos ciclos de imunização, o sangue é coletado e dele é separado o plasma que é ainda processado para a obtenção do soro contendo anticorpos específicos (SILVA JÚNIOR; SASSON; CALDINI JÚNIOR, 2016).

As desvantagens na produção e no uso de anti-venenos contra o loxoscelismo são constantemente discutidas. O protocolo utilizado baseia-se na extração de venenos de pequenos artrópodes, processo que demanda a ordenha de um grande número de aranhas para que seja obtido veneno suficiente para os diversos ciclos de hiperimunização. A produção é apontada ainda como um processo caro, fator que limita sua utilização em países em desenvolvimento (THEAKSTON; WARRELL; GRIFFITHS, 2003). Os cavalos, que são a espécie mais utilizada na produção por sua tolerância à exposição a antígenos tóxicos e pelo potencial de extração de grandes

volumes de sangue, têm prejuízo na qualidade e expectativa de vida e podem desenvolver abscessos e inflamação nos locais de inoculação do veneno, além de anemia e alterações em outros parâmetros hematológicos pelas sangrias recorrentes (MIRANDA et al., 2022). A eficiência dos soros anti-loxoscélico nunca foi avaliada através de estudos clínicos randomizados (FINGERMANN et al., 2020). As conclusões geradas a partir dos dados disponíveis indicam que a efetividade é circunstancial: seus efeitos são mais notáveis contra o loxoscelismo víscero-cutâneo e quando a intervenção é realizada poucas horas após o acidente (PAULI et al., 2006). Quanto aos sinais cutâneos, investigações em cobaias indicam que os anti-venenos contribuem para a redução da lesão dermonecrótica quando administrados em até 48 horas após o envenenamento (PAULI et al., 2006, 2009). Por fim, eventos adversos precoces ou tardios são relatados em pacientes tratados com a soroterapia com incidência que varia de 6,5% (SEZERINO et al., 1998) a 20% (MÁLAQUE et al., 2002).

Algumas ideias já foram exploradas para tratar as limitações do processo de produção de anti-venenos. O uso de proteínas recombinantes para a imunização solucionaria a dificuldade de extração do veneno de aranha-marrom. Assim, foram testadas formas recombinantes de fosfolipases D de *L. laeta* e *L. intermedia* que foram efetivas para gerar soros neutralizantes a partir da imunização de equinos (DE ALMEIDA et al., 2008; DUARTE et al., 2015). Apesar de efetivas na imunização, os componentes mantinham sua atividade enzimática e, portanto, ainda eram danosos aos cavalos. O desenvolvimento subsequente foi a criação de proteínas multiepitópicas que tivessem a antigenicidade original e toxicidade reduzida. São exemplos a rCpLi (MENDES et al., 2013; FIGUEIREDO et al., 2014), proteína atóxica e capaz de elicitar anticorpos neutralizantes em coelhos e em cavalos quando combinada a etapas de hiperimunização com veneno total, e Lil (SOUZA et al., 2018), cujo o soro foi parcialmente eficaz na neutralização dos efeitos tóxicos do veneno *in vitro*, ambas resultando da combinação de três diferentes epítomos da LiD1. A partir de epítomos de metaloproteases, fosfolipase D e hialuronidase do veneno de *Loxosceles* foi gerada a rMEPLox (LIMA et al., 2018), também atóxica e que teve sua capacidade de geração de anticorpos neutralizantes em coelhos comprovada *in vitro* e *in vivo*. O uso de enzimas recombinantes, sejam elas similares as formas selvagens ou resultantes da reorganização de epítomos, é vantajoso, mas não elimina a necessidade de animais para a produção de anti-venenos e não apontam melhorias na ocorrência de reações adversas a soroterapia.

Alternativas à soroterapia também foram propostas na forma de produção de anticorpos monoclonais contra algum dos componentes do veneno. Guilherme, Fernandes e Barbaro, em 2001, geraram quatro anticorpos monoclonais contra o veneno de *L. gaucho*, dentre os quais o MoALg₁, que reconheceu o veneno dessa espécie com alta afinidade e apresentou ainda discreta reatividade contra o veneno das outras espécies de importância médica. Através da imunização com o veneno de *L. intermedia*, foi gerado o anticorpo denominado LimAb(7), que foi capaz de neutralizar os efeitos tóxicos do veneno dessa espécie (ALVARENGA et al., 2003). Dias-Lopes e colaboradores (2014) relataram a obtenção de um monoclonal (mAb16) reativo contra LiD1 e capaz de reconhecer também frações do veneno de *L. intermedia*, *L. laeta* e *L. gaucho*. Posteriormente, o LimAb(7) teve seus domínios variáveis humanizados, e foi expresso na forma de fragmento de anticorpo de cadeia única (scFv, *single chain antibody fragment*) (KARIM-SILVA et al., 2020). A forma humanizada de LimAb(7) demonstrou reatividade contra o veneno de *L. intermedia* e neutralização *in vitro* da atividade hemolítica do veneno dessa espécie. Os anticorpos monoclonais relatados até o momento apresentam limitações na reatividade contra múltiplas espécies de *Loxosceles* e a busca por anticorpos com menor potencial de respostas adversas ainda é pouco explorada.

Apesar da corrida por inovações, ainda há um descompasso entre as tecnologias pesquisadas e a realidade da produção de anti-venenos (MIRANDA et al., 2021). Novos estudos são necessários para sanar a escassez de informações sobre os anticorpos de cavalo e sobre as particularidades dos animais hiperimunizados, conhecimentos de grande importância para continuar o desenvolvimento da soroterapia equina em suas diversas formas.

1.3. Imunoglobulinas

Os soros anti-venenos consistem nas frações F(ab')₂ de imunoglobulinas séricas de cavalos hiperimunizados com as peçonhas. Assim, o entendimento de sua eficácia e limitações perpassa pelo estudo aprofundado das imunoglobulinas que o compõem.

As imunoglobulinas são os receptores da resposta imune adaptativa de células B que têm, entre outras funções, papel no reconhecimento e neutralização de

antígenos. Estruturalmente, tratam-se de glicoproteínas quaternárias compostas por um par de cadeias leves e um par de cadeias pesadas, que possuem seus domínios variáveis (V_L para cadeia leve e V_H para cadeia pesada) e constantes (C_L para cadeia leve e C_H para cadeia pesada) que podem variar em número de acordo com a classe da imunoglobulina. As cadeias pesadas são unidas por ligações dissulfeto na região de dobradiça que marca ainda a divisão entre a região Fc (do inglês, *Fragment crystallizable region*) que contém apenas domínios constantes e a região Fab (do inglês *Fragment antigen-binding*) que contém um domínio constante e o domínio variável (INBAR; HOCHMAN; GIVOL, 1972; KABAT; WU; BILOFSKY, 1976; TE WU; KABAT, 1970). Os domínios constantes são codificados pelos segmentos gênicos IGC enquanto os domínios variáveis resultam da combinação dos segmentos de variedade (V), diversidade (D) e junção (J) para a cadeia pesada e V e J para a cadeia leve. O domínio variável caracteriza-se por apresentar quatro regiões conservadas estruturais (FR, do inglês *framework*) que flanqueiam as três regiões hipervariáveis (CDR, do inglês *complementarity determinant region*) normalmente envolvidas na interação com o antígeno (Figura 3)(CHOTHIA; LESK, 1987; LITMAN; ANDERSON; RAST, 1999).

A variedade das imunoglobulinas é resultado de uma série de processos de diversificação que garantem a capacidade do organismo de reconhecer qualquer antígeno apresentado. A diversificação ocorre pela variedade alélica e recombinação dos segmentos V(D)J e C, pelo pareamento das cadeias leve e pesada, pela adição e remoção de nucleotídeos, pela hipermutação somática e pela introdução de modificações pós-traducionais (GEORGIU et al., 2014; KOERS et al., 2019). Como resultado desses mecanismos de variabilidade, estima-se que humanos possuam, no sangue, uma diversidade de imunoglobulinas da ordem de 10^{11} a 10^{18} (BOYD; JOSHI, 2014).

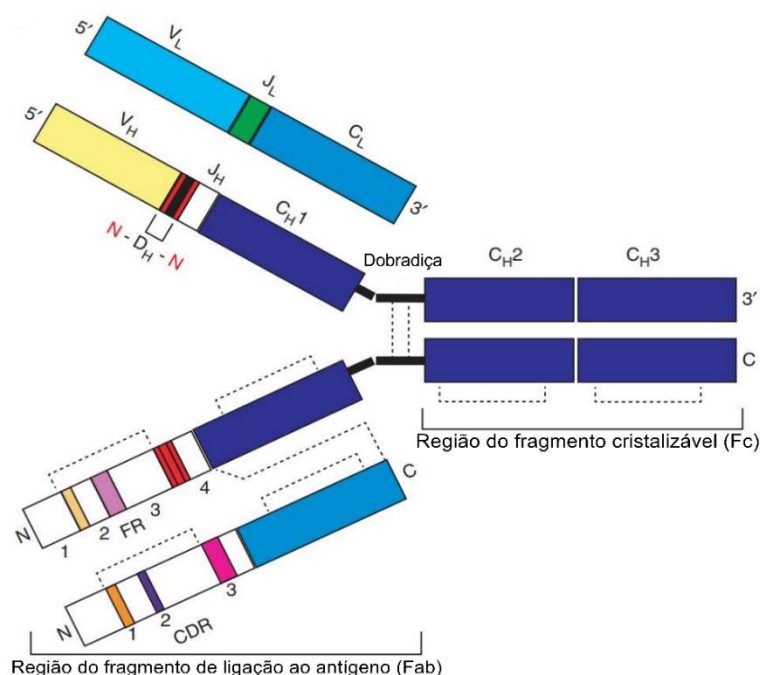


Figura 3 – Estrutura de imunoglobulina. As imunoglobulinas são formadas por um par de cadeia leve e um par de cadeia pesada e são subdivididas nos domínios constantes das regiões Fc e Fab e nos domínios variáveis da Fab. N: extremidade N-terminal. C: extremidade C-terminal. V_H: região da cadeia pesada codificada pelo segmento gênico V. D_H: região da cadeia pesada codificada pelo segmento gênico D. J_H: região da cadeia pesada codificada pelo segmento gênico J. V_L: região da cadeia leve codificada pelo segmento gênico V. J_L: região da cadeia leve codificada pelo segmento gênico J. C_H: domínios constantes da cadeia pesada. C_L: domínio constante da cadeia leve. FR: região estrutural ou *framework*. CDR: região determinante de complementariedade. As linhas tracejadas indicam as ligações dissulfeto (GEORGIU et al., 2014).

1.3.1. Imunoglobulinas equinas

Inicialmente, as cadeias pesadas de imunoglobulinas equinas foram classificadas como IgA, IgM, IgGa, IgGb, IgGc, IgG(T) e IgG(B) (ou AI, imunoglobulina agregadora) baseado em ensaios sorológicos e eletroforéticos (MCGUIRE; CRAWFORD; HENSON, 1973; MONTGOMERY, 1973). A partir de análises mais recentes, onze genes funcionais foram identificados em cavalos e a distribuição de classes foi atualizada: um segmento gênico foi encontrado para as cadeias alfa (IgA), delta (IgD), épsilon (IgE), mi (IgM) e sete para cadeia gamma (IgG1-IgG7). A equivalência entre as duas denominações é importante pois os anticorpos monoclonais contra as subclasses de IgG estabelecidos a partir da classificação

antiga continuam disponíveis comercialmente. Assim, estabeleceu-se que IgGa corresponde aos segmentos gênicos IgG1 e IgG2; IgGb aos segmentos IgG4 e IgG7; IgGc ao IgG6 e IgG(T) aos segmentos IgG3 e IgG5 (WAGNER, 2006; WALTHER et al., 2015). Quanto às cadeias leves, os segmentos gênicos kappa (IgK) e lambda (IgL) são encontrados nos cavalos, com predomínio de cadeias lambda (ARUN; BREUER; HERMANNNS, 1996; WALTHER et al., 2015).

As características funcionais das diferentes classes são similares à de outras espécies. As IgMs são típicas do período agudo de infecções e atuam na resposta imune humoral primária. As IgEs são encontradas em menor abundância no soro e são relevantes nas reações de hipersensibilidade do tipo I e na imunidade contra parasitas. Apesar da identificação do segmento gênico que a codifica e de sua expressão, as imunoglobulinas de classe D equinas ainda não foram detectadas no soro e sua função permanece desconhecida (WAGNER et al., 2004). Dentre as subclasses de IgG, que são mais numerosas em cavalos do que em outras espécies, IgG4 e IgG7 já foram associadas ao combate de patógenos intracelulares enquanto IgG3 e IgG5 são detectadas em resposta a patógenos extracelulares (LEWIS; WAGNER; WOOF, 2008; KEGGAN et al., 2013). IgG1, IgG3, IgG4, IgG5 e IgG7 interagem com receptores de Fc de células efetoras da resposta imune enquanto IgG1, IgG3, IgG4 e IgG7 foram capazes de se ligar a C1q e ativar a via clássica do sistema complemento (LEWIS; WAGNER; WOOF, 2008).

Além das avaliações sorológicas e funcionais das cadeias constantes já mencionadas, a caracterização da região variável de imunoglobulinas equinas já foi alvo de estudos baseados no sequenciamento de transcriptoma (ALMAGRO et al., 2006; TALLMADGE et al., 2013). Os achados indicam que os segmentos gênicos IGHV equinos são marcados pela alta frequência de pseudogenes, ou seja, segmentos não expressos pela presença de códons de parada e/ou alterações na fase de leitura, além de alguns ORFs (*Open Reading frame*), sequências em que alterações são detectadas nas regiões de *splicing* ou de sinal de recombinação ou ainda que não apresentam correlatos em nenhuma outra espécie (LEFRANC, 1998). Assim, para a cadeia pesada são conhecidos: 21 segmentos gênicos V funcionais, 44 segmentos D e 9 segmentos J, além de 74 pseudogenes e 9 ORFs do segmento gênico V. Para cadeia leve lambda (IGL), são relatados 29 segmentos V funcionais, 112 pseudogenes, 5 ORFs e 7 genes J. Para cadeia leve kappa (IGK), 16 genes V

funcionais, 47 pseudogenes, 2 ORFs e 5 segmentos J funcionais e 1 pseudogene já foram descritos (LEFRANC, 2001).

O trabalho mais recente de Manso e colaboradores (2019) baseado em sequenciamento de nova geração evidenciou que, assim como outras espécies, os cavalos possuem preferências no uso dos segmentos gênicos de imunoglobulina, com predomínio das famílias IGHV4, IGHJ6 e IGHD2 (nomenclatura do International ImMunoGeneTics Information System – IMGT). O uso de aminoácidos no CDRH3 foi similar ao encontrado em humanos e camundongos. Porém, quando os resíduos são avaliados por posição, os equinos apresentam menor frequência de resíduos carregados nas posições 106 e 116 que são determinantes para a formação de uma alça/dobra no CDRH3 (conformação *kinked*). Assim, a conformação estendida dessa região possivelmente é predominante em cavalos. Outra característica determinada foi o tamanho mediana do CDRH3, que foi de 15 aminoácidos. Apesar deste avanço no conhecimento do repertório de anticorpos equinos, pelo nosso conhecimento, nenhum trabalho até agora, caracterizou repertório de anticorpos equinos antígeno-específicos por metodologias em larga escala.

1.4. Estratégias para estudo de repertório de anticorpos

1.4.1. Avaliando o repertório de receptores de células B

De forma a entender de uma maneira sistêmica a resposta imune humoral, metodologias para determinar o repertório de anticorpos codificados pelas células B avançaram rapidamente (BENICHOU et al., 2012; GEORGIU et al., 2014). A maior parte dos estudos para caracterização de repertório de anticorpos baseiam-se no isolamento de células B a partir do PBMC (sigla em inglês para *peripheral blood mononuclear cells*), da medula ou órgãos linfoides secundários, seguidas do isolamento de DNA ou mRNA dessas células, produção do amplicon das regiões variáveis da imunoglobulina e sequenciamento de ácidos nucleicos. As numerosas sequências identificadas são agrupadas pela similaridade da região de CDR3 e dos segmentos V e J utilizados, gerando conjuntos que representariam anticorpos originados de um mesmo clone de célula B, denominados clones ou clonotipos (HERSHBERG; LUNING PRAK, 2015). Trabalhos baseados no transcriptoma dos

receptores de células B (denominados BCR-Seq ou Rep-Seq) podem contribuir para a descoberta de anticorpos contra antígenos de interesse, como contra dengue (PARAMESWARAN et al., 2013), influenza (JIANG et al., 2013) ou HIV (WILLIAMS et al., 2017), para o entendimento e diagnóstico de distúrbios de células B, como no caso das leucemias (LOGAN et al., 2011) para o desenvolvimento de vacinas, entre outras aplicações (GEORGIU et al., 2014).

Apesar de relevantes, os estudos dos receptores de células B a partir de mRNA ou DNA geram dados de distribuição e abundância de anticorpos que não refletem as características de imunoglobulinas secretadas encontradas nos repertórios antígeno-específicos (CHEN et al., 2017; CHEUNG et al., 2012; BONISSONE et al., 2020). Até o momento, a espectrometria de massa (MS, do inglês *mass spectrometry*) tem sido a técnica mais apropriada para a investigação de conjuntos complexos de anticorpos secretados, ao permitir a elucidação da diversidade de sequências do repertório a partir de frações de especificidade conhecida. Esse método, denominado Ig-Seq, vem sendo padronizado nos últimos dez anos para a exploração de amostras policlonais (IONOV; LEE, 2022).

Os estudos baseados em Ig-Seq têm se dedicado a explorar frações do repertório total de anticorpos de amostras como plasma ou soro, líquido cefalorraquidiano, biópsias intestinais, entre outras possibilidades. Os subconjuntos explorados são resultado da seleção de anticorpos com reatividade contra um ou vários antígenos de interesse ou ainda do isolamento preferencial de uma determinada classe ou subclasse de imunoglobulina. Esses estudos baseiam-se do preparo de imunoglobulinas ou F(ab')₂ intactas, no caso do *top-down*, ou no uso de proteases que geram peptídeos proteolíticos mais longos, no caso do *middle-down*, em comparação aos peptídeos menores obtidos a partir da digestão por tripsina ou outras enzimas utilizadas na estratégia *bottom-up*.

1.4.2. Estudos de repertório baseados em proteômica *middle* ou *top-down*

No contexto de estudo de repertório de anticorpos, as abordagens *top-down* ou *middle-down* vêm sendo empregadas com maior frequência na investigação de distúrbios proliferativos de células B com produção excessiva de imunoglobulinas monoclonais. Barnidge e colaboradores (2014) demonstraram a diferenciação dos

perfis policlonais de cadeias kappa e lambda encontradas no soro de indivíduos saudáveis e também a realizaram a identificação de picos oligoclonais ou monoclonais em indivíduos com hiper- ou hipogamaglobulinemia. Os resultados foram obtidos pelo uso da estratégia *top-down*, de execução rápida e simples, segundo os autores. Apesar da elevada resolução, a metodologia não foi suficiente para a deconvolução das sequências policlonais que compunham cada um dos perfis.

Essa limitação pode ser parcialmente solucionada a partir da combinação de múltiplas estratégias de tratamento com proteases, de separação e fragmentação dos peptídeos e de obtenção de sequências. No trabalho de Wang e colaboradores de 2019, foram utilizadas cromatografia líquida de ultra eficiência (UPLC, *ultra-high-pressure liquid chromatography*) acoplada à MS e a combinação de diferentes métodos de fragmentação para obter as sequências de uma mistura de doze anticorpos monoclonais com alta cobertura. O método também foi testado na análise de soros de pacientes com lúpus eritematoso sistêmico e permitiu a descoberta de 47 sequências de cadeia leve e 16 sequências de cadeia pesada únicas que geram um perfil de repertório distinto dos pacientes saudáveis utilizados como controle.

Em outra abordagem, as imunoglobulinas dos plasmas de pacientes convalescentes para COVID-19 foram detectadas e deconvoluídas a partir da espectrometria de massa de íons-individuais (I²MS, do inglês individual *ions mass spectrometry*). Nesse método, a carga de cada íon detectado no MS/MS é definida e utilizada para processamentos que permitem a atribuição da relação m/z daquele íon. As transformações matemáticas permitiram a obtenção dos parâmetros de título do íon e grau de clonalidade, utilizados para traçar perfis distintos de indivíduos com COVID-19 leve, severa e indivíduos vacinados contra a doença. Assim, são investigadas características de abundância e heterogeneidade do repertório antígeno-específico sem o sequenciamento de células B (MELANI et al., 2022).

Os exemplos demonstram que o uso de *top* ou *middle-down* na investigação de repertórios de anticorpos é promissor, mas seu poder resolutivo permanece limitado diante da complexa composição das amostras. Os desenvolvimentos mais recentes que incluem a combinação de abordagens ou a espectrometria de massa de íons individuais, indicam que esses métodos em breve podem ser consolidados como protagonistas nesse objetivo ou como auxiliares às técnicas baseadas em *bottom-up* já utilizadas.

1.4.3. Estudos de repertório baseados em proteômica *bottom-up*

A proteômica baseada em *bottom-up* é utilizada na maior parte dos trabalhos baseados em espectrometria de massa, inclusive nos estudos de imunoglobulinas. As publicações dedicadas a explorar algumas frações de repertórios de imunoglobulinas a partir do *bottom-up* foram recentemente revisadas por Snapkov e colaboradores (2021) e por Ionov e Lee (2022). Os trabalhos de Sato e o de Cheung e colaboradores, ambos realizados em 2012, foram pioneiros no uso da proteômica *bottom-up* baseada em espectrometria de massa acoplada à cromatografia líquida (LC-MS/MS) junto ao sequenciamento de nova geração (NGS) de transcriptoma para a descoberta de anticorpos sorológicos antígeno-específicos, uma combinação do BCR-Seq/Rep-Seq e do Ig-Seq (Figura 4). A ideia da complementariedade das abordagens é de grande relevância para a investigação dos peptídeos de imunoglobulinas com maior cobertura e profundidade, já que a variação da composição dos repertórios individuais e a escassez de sequências de anticorpos em repositórios públicos impede ou limita muito o uso de bancos de dados centralizados.

Nesse primeiro uso da técnica, foram avaliados o repertório de anticorpos humanos com reatividade contra o vírus da hepatite B (HBV) de um indivíduo vacinado, os anticorpos reativos contra o citomegalovírus humano (CMV) de um indivíduo infectado naturalmente e ainda o repertório de coelhos hiperimunizados com os receptores de progesterona humana A e B (PR A/B). No trabalho de Sato, a partir do uso de múltiplas proteases no preparo para o *bottom-up* e do piro sequenciamento para geração do banco de sequências de referência, foram obtidas múltiplas sequências de imunoglobulinas das quais 30 cadeias leves e 24 cadeias pesadas foram selecionadas para expressão em todas os possíveis pareamentos gerando 37 pares de cadeias de imunoglobulinas específicos contra o vírus da hepatite B. Para o CMV, 15 anticorpos reativos foram encontrados. No trabalho de Cheung, oito cadeias leves e dez cadeias pesadas foram selecionadas para expressão e geraram seis clonotipos específicos contra o PR A/B. Nos dois estudos, foi alcançada a descoberta de anticorpos específicos de alta afinidade pelo antígeno selecionado.

As duas investigações iniciais serviram como prova de conceito para esse complexo conjunto de técnicas, cujos detalhes vêm sendo aperfeiçoados e ajustados a diferentes aplicações. A partir de 2013, o grupo liderado pelo Dr. George Georgiou,

da Universidade do Texas, iniciou uma série de estudos que contribuíram para a padronização e evolução do Ig-Seq.

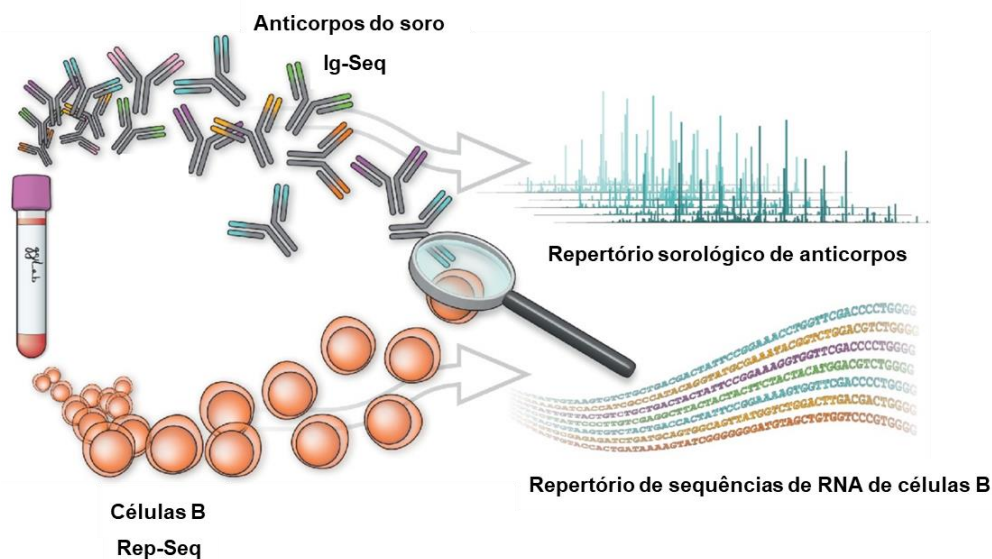


Figura 4 – Complementariedade das abordagens do Ig-Seq e do Rep-Seq para a descoberta do repertório sorológico de imunoglobulinas. A partir de amostras de sangue do organismo a abordagem multi-ômica é realizada para a descoberta do repertório sorológico de anticorpos. A partir de células B são obtidas as seqüências de RNA para compor os dados de Rep-Seq ou BCR-Seq e partindo dos anticorpos circulantes do soro é definido o repertório sorológico elucidado a partir das seqüências de referência do transcriptoma do próprio organismo. Adaptado de Lavinder et al. (2014).

A combinação dessas abordagens também foi utilizada na investigação do repertório sorológico até nove meses após a vacinação contra toxina tetanoide (LAVINDER et al., 2014). Neste estudo, foram descobertos cerca de 100 clonotipos de anticorpos (seqüências possivelmente originadas de um mesmo clone de célula B), dos quais 3 correspondiam a cerca de 40% da resposta específica. A expressão de algumas das seqüências identificadas foi seguida pela determinação dos coeficientes de dissociação para indicar quais imunoglobulinas possivelmente estariam associados a neutralização da toxina (LAVINDER et al., 2014). Lee e colaboradores (2016) avaliaram anticorpos no plasma de indivíduos antes e após a vacinação contra Influenza e evidenciaram a presença de diversos clones com reatividade cruzada contra diferentes hemaglutininas virais. Mais recentemente a abordagem foi utilizada para a descoberta das características dos epítomos e da

convergência dos clones de IgG em indivíduos convalescentes para a infecção por SARS-CoV-2 (VOSS et al., 2021).

1.4.3.1. Metodologia do Ig-Seq

Na publicação de Wine e colaboradores de 2013, o método do Ig-Seq foi aplicado ao estudo de imunoglobulinas de coelho que haviam sido imunizados com a hemocianina do molusco gastrópode *Concholepas concholepas*. Para investigar o repertório sorológico foram realizados o isolamento de IgG a partir de cromatografia de afinidade com proteína A, a digestão dos anticorpos isolados por pepsina, o enriquecimento das frações F(ab')₂ específicas contra elementos do veneno usado para imunização e por fim a desnaturação, redução, alquilação e digestão por tripsina dos fragmentos antígeno-específicos seguida da espectrometria de massa (Figura 5). Paralelamente, foi executada a amplificação e o sequenciamento de regiões de imunoglobulina do mRNA de células B periféricas (PBCs) e de células da medula óssea. Os dados do sequenciamento foram usados para compor o banco de seqüências personalizado empregado para a busca dos espectros de massa. Uma vez que as cadeias leves e pesadas foram identificadas, os pares de seqüências mais abundantes foram expressos e avaliados quanto a sua afinidade no reconhecimento antigênico.

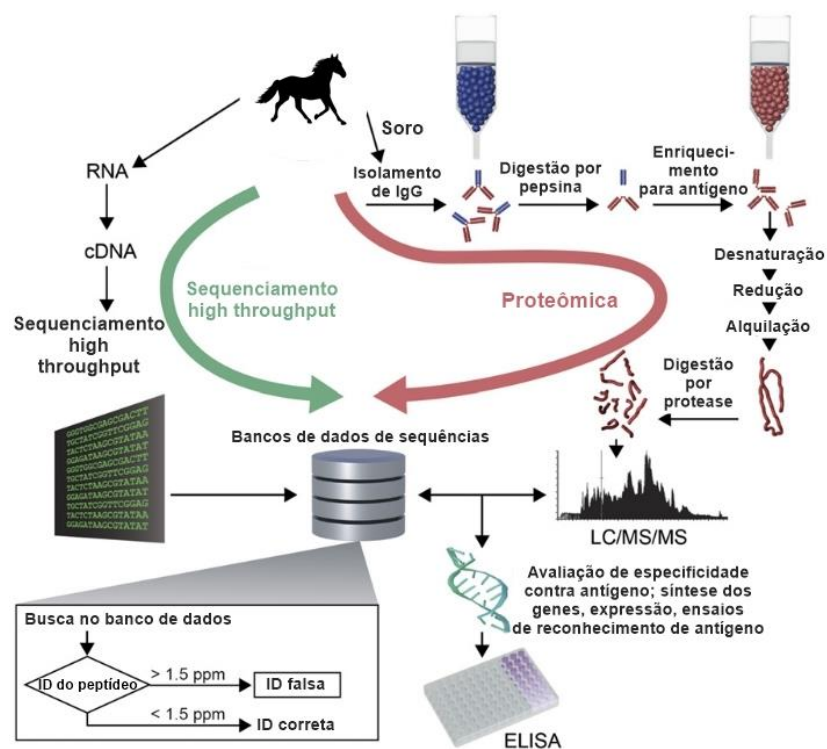


Figura 5 – Etapas do sequenciamento *high throughput* e de proteômica para obtenção das imunoglobulinas antígeno-específicas. Detalhes dos processos paralelos do Ig-Seq (representado como Proteômica) e do Rep-Seq (representado como sequenciamento *high throughput*), terminando com a identificação dos peptídeos pela combinação dos dados das duas abordagens. São ilustradas ainda as possíveis etapas posteriores a descoberta de sequências, que incluem a expressão dos anticorpos e os ensaios de reconhecimento do antígeno de interesse. Adaptado de Wine e colaboradores (2013).

Em comparação aos trabalhos iniciais de Sato e Cheung, foram propostos alguns parâmetros importantes para a descoberta de sequências com maior confiabilidade, como menor tolerância de desvio de massa ($\pm 1,5$ ppm) e de taxa de falsa descoberta (FDR, do inglês *false discovery rates*) ($< 1\%$). Além do eluato da cromatografia de afinidade executada para enriquecimento dos anticorpos específicos para o antígeno de interesse, a fração não interagida também foi submetida a análise proteômica a fim de filtrar os peptídeos identificados apenas nas frações específicas.

A deconvolução da diversidade de anticorpos contidos no repertório baseia-se no uso da região de CDR3 para a identificação das sequências. O CDR3 resulta da junção dos segmentos gênicos V e J na cadeia leve (CDRL3) e dos genes V(D)J na cadeia pesada (CDRH3), estando suscetível a variabilidade gerada pela adição e remoção de nucleotídeos e pela hipermutação somática. Essas características fazem com que a região esteja frequentemente envolvida na definição da especificidade ao antígeno (XU; DAVIS, 2000). No Ig-Seq, os peptídeos mapeados à região de CDRH3 e com sobreposição suficiente (de quatro resíduos ou mais) são considerados para a identificação das cadeias de anticorpos e são denominados peptídeos informativos.

As etapas iniciais de análise dos dados da LC-MS/MS *bottom-up* usada nesses estudos caracterizam-se pela aquisição dos dados no modo dependente de dados (DDA, do inglês *data-dependent analysis*) e posteriormente pelo uso de ferramentas de busca como o SEQUEST (ENG; MCCORMACK; YATES, 1994), Mascot (PERKINS et al., 1999) ou MS Amanda (DORFER et al., 2014), para mapear os peptídeos obtidos nos bancos de sequências personalizados gerados no Rep-Seq. Na maior parte dos trabalhos, ferramentas tradicionais de análise dos dados de proteômica como o MaxQuant (COX; MANN, 2008), PEAKS (MA et al., 2003) ou Proteome Discoverer (Thermo Scientific) são utilizadas apesar da existência de alternativas específicas para os estudos de imunoglobulinas como o PASA (AVRAM et al., 2021) ou o PEAKS

AB 2.0 (TRAN et al., 2016). Como consequência, os passos posteriores de mapeamento dos peptídeos identificados às diferentes regiões das imunoglobulinas vêm sendo executados através de *scripts in house* frequentemente desenvolvidos em Python.

Além das aplicações já citadas, o Ig-Seq também já foi utilizado para: i) a caracterização de autoanticorpos de pacientes com pênfigo, uma doença rara que acomete a pele e mucosas com a formação de bolhas (CHEN et al., 2017), ii) para a comparação dos repertórios sorológico e intestinal de IgAs de indivíduos com doença celíaca (IVERSEN et al., 2017), iii) para avaliação de bandas oligoclonais encontradas no sangue e no fluido cefalorraquidiano de pacientes com esclerose múltipla (TOMESCU-BACIU et al., 2019), iv) para a identificação de *nanobodies* de camelídeos e seus epítomos (XIANG et al., 2021), dentre outras. O conjunto de estudos demonstra a versatilidade da técnica para a investigação de repertórios de diferentes classes de imunoglobulinas, de múltiplos compartimentos corporais e provenientes de variados organismos.

A metodologia de investigação de repertório de anticorpos sorológicos ainda não foi aplicada a amostras equinas ou ao estudo de anti-venenos. No contexto do loxoscelismo, a descoberta das características dos anticorpos específicos contra componentes do veneno seria de grande contribuição para o aprimoramento do soro anti-veneno já utilizado ou para a busca de novas terapias baseadas em anticorpos específicos policlonais.

2. Objetivos

2.1. Objetivo geral

Identificar e caracterizar anticorpos séricos de cavalos hiperimunizados que sejam reativos contra componentes do veneno de aranhas do gênero *Loxosceles* de importância médica no Brasil.

2.2. Objetivos específicos

- a. Avaliar a reatividade dos plasmas equinos frente ao veneno de *Loxosceles* e a fosfolipase D recombinante de *L. intermedia* (rLiD1);
- b. Quantificar as diferentes classes de imunoglobulinas encontradas nos plasmas;
- c. Determinar a abundância relativa das diferentes subclasses de IgG nos plasmas equinos;
- d. Identificar e caracterizar as sequências dos fragmentos das cadeias pesadas das frações F(ab')₂ antígeno-específicos através de análise proteômica e bioinformática;
- e. Avaliar a similaridade dos possíveis sítios de interação com antígeno das sequências específicas contra rLiD1.

3. Metodologia

3.1. Hiperimunização dos equinos e coleta dos plasmas

Os experimentos foram realizados com amostras provenientes de quatro cavalos machos, adultos, de idades entre 5 e 9 anos, fornecidas gentilmente pelo Centro de Pesquisa e Produção de Imunobiológicos (CPPI). O delineamento experimental foi aprovado pela Comissão de Ética no Uso de Animais da Universidade Federal de Minas Gerais (CEUA - UFMG) sob o número de protocolo 190/2018. Para obtenção dos plasmas, os animais foram submetidos a coleta de 10-30 mL de sangue em tubos BD Vacutainer® com EDTA, imediatamente antes da primeira imunização e sete dias após a última etapa do processo de hiperimunização com o *pool* contendo quantidades equivalentes dos venenos de *L. intermedia*, *L. gaucho* e *L. laeta*. O protocolo de hiperimunização foi composto pelas seguintes etapas:

- a. 1ª imunização (dia 0): 300 µg de *pool* de venenos solubilizado em 1,200 mL de solução salina e 1,5 mL de Adjuvante Completo de Freund.
- b. 2ª imunização (dia 10): 450 µg de *pool* de venenos solubilizado em 1,050 mL de solução salina e 1,5 mL de Adjuvante Completo de Freund.
- c. 3ª imunização (dia 21): 750 µg de *pool* de venenos solubilizado em 2,250 mL de solução salina.
- d. 4ª imunização (dia 30): 750 µg de *pool* de venenos solubilizado em 2,250 mL de solução salina.

As amostras de plasma recebidas foram aliquotadas em micro tubos *protein low binding* de 2,0 mL e armazenadas a -20° C. As amostras coletadas antes da imunização foram denominadas pré-ímmunes e passaram a ser identificadas por 1P, 2P, 3P e 4P. Aquelas coletadas sete dias após a última imunização foram designadas hiperímmunes e foram identificadas como 1H, 2H, 3H e 4H.

3.2. *Pool* de venenos de *Loxosceles* e rLiD1

Assim como os plasmas equinos, os venenos liofilizados de *L. gaucho*, *L. laeta* e *L. intermedia* foram fornecidos pelo CPPI. A coleta desse material foi realizada a partir de eletrochoque de 15V no cefalotórax das aranhas. Para compor o *pool* utilizado nos ensaios de imunoabsorção enzimática (ELISA), quantidades

equivalentes dos três venenos foram combinadas e armazenadas a -20°C até o momento de sua utilização. A forma recombinante da fosfolipase D de *L. intermedia* (rLiD1), também utilizada no ensaio de ELISA, foi produzida anteriormente no laboratório como descrito no trabalho de Felicori e colaboradores (2006).

3.3. Reatividade e quantificação de imunoglobulinas nos plasmas equinos

3.3.1. Reatividade dos plasmas contra *pool* de venenos e rLiD1

A presença de anticorpos reativos contra o *pool* de venenos ou contra a rLiD1 nos plasmas equinos foi investigada através de ELISA. Placas de 96 poços (Corning® 96 Well EIA/RIA Assay Microplate) foram sensibilizadas com o *pool* dos três venenos (5 $\mu\text{g}/\text{mL}$, 100 μL , 4°C , *overnight*) em tampão carbonato (pH 9,6, 62,5 mM) e bloqueadas com leite Molico (3% em PBS 1X, 100 μL , 37°C , 1h). Após o bloqueio, foram adicionadas diluições seriadas de 1.000 a 128.000 vezes dos plasmas pré ou hiperimunes, dos quatro cavalos, em duplicata (100 μL , 37°C , 1h). Em sequência, adicionou-se o conjugado (*anti-horse IgG peroxidase (whole molecule) antibody produced in rabbit*, Sigma-Aldrich) diluído 6.000 vezes em tampão de incubação (PBS-Tween® 20 0,05% com 0,1% de leite Molico). Após a incubação por 1 hora a 37°C , o experimento foi revelado utilizando-se solução de 0,33 mg/ml o-fenilenodiamina (OPD, Sigma-Aldrich) em tampão citrato pH 5,2 com 0,04% de H_2O_2 (100 μL , TA, 30 minutos, ao abrigo de luz). Entre as diferentes etapas, os poços foram lavados por três vezes com PBS-Tween® 20 0,01%. A reação foi interrompida com 25 μL de H_2SO_4 2N e a absorbância no comprimento de onda de 490 nm foi determinada usando o equipamento Multiskan™ Sky (Thermo Scientific).

3.3.2. Quantificação absoluta das classes de imunoglobulinas

A fim de caracterizar a distribuição das diferentes classes de imunoglobulinas no plasma equino, foi realizada a quantificação de IgM, IgG e IgA nas amostras pré e hiperimunes utilizando os kits: Horse IgM ELISA Kit (ab190532), Horse IgG ELISA Kit (ab190531) e Horse IgA ELISA Kit (ab190530) da Abcam. O protocolo foi executado de acordo com as recomendações do fabricante. Em resumo: nos poços pré-tratados,

foram pipetados 100 µL do padrão nas diluições recomendadas para a curva de calibração, 100 µL dos 8 plasmas nas diluições de 1:200.000 para IgG e 1:10.000 para IgM e IgA, em duplicata, e aguardou-se a incubação por 30 minutos à temperatura ambiente (TA). Os poços foram lavados com o diluente por quatro vezes, adicionou-se 100 µL do conjugado e a placa foi incubada por 20 minutos à TA. Após nova lavagem, o teste foi revelado usando TMB (3,3',5,5'-tetrametilbenzidina) (100 µL, TA, 10 minutos) e, após adição da solução de parada, a absorbância foi determinada no comprimento de onda de 450 nm no Multiskan™ Sky (Thermo Scientific).

3.3.3. Quantificação relativa das subclasses de IgG

Para a caracterização das imunoglobulinas dos plasmas equinos, foi realizada ainda a avaliação de três das subclasses de IgG (IgGa, IgGb e IgG(T)) antígeno-específicas nos plasmas pré-ímmunes e hiperímmunes, de acordo com o protocolo descrito por Halassy e colaboradores em 2019. Placas de 96 poços foram sensibilizadas com 1 µg/mL do *pool* dos três venenos de *Loxosceles* em PBS 1X. Após a incubação *overnight*, os poços foram bloqueados com BSA (Sigma, 0,5% em PBS 1X) por 2 horas a 37° C. Em sequência, foram adicionadas as diluições seriadas das amostras: para os plasmas hiperímmunes de 1:5.000 a 1:640.000 e para os plasmas pré-ímmunes de 1:2.500 a 1:80.000 (incubação à TA, *overnight*). No dia seguinte, foi realizada a adição dos anticorpos anti-subclasse de IgG equina nas diluições de 1:1.000 de IgGa (clone CVS48, Bio-Rad) e IgGb (clone CVS39, Bio-Rad) e de 1:4000 de IgG(T) (Polyclonal anti-IgG(T), Bio-Rad). Após mais duas horas de incubação à 37° C, adicionou-se o conjugado *anti-mouse* (Sigma-Aldrich) nos poços marcados com IgGa ou IgGb e *anti-goat* (Sigma-Aldrich) nos poços marcados com IgG(T) em diluição 1:5.000 (2 horas, 37° C). Entre as diferentes etapas, os poços foram lavados por três vezes com PBS-Tween® 20 0,01%. Por fim, o ensaio foi revelado e a absorbância final foi determinada como descrito no item 3.3.1.

3.4. Obtenção de fragmentos trípticos de F(ab')₂ específicos contra o veneno de *Loxosceles* e rLiD1

As amostras de plasma de cavalos hiperímmunizados foram preparadas para a descoberta de sequências de imunoglobulinas pelo Ig-Seq através de uma série de

experimentos discutidos na monografia de Ferreira (2019) que não serão o alvo deste trabalho. Resumidamente, as etapas realizadas foram: i) isolamento de IgG por cromatografia de afinidade com proteína G, ii) clivagem por pepsina para obtenção da fração $F(ab')_2$, iii) enriquecimento dos $F(ab')_2$ específicos para o antígeno por cromatografia de afinidade com *pool* de venenos ou com a rLiD1 e iv) obtenção dos peptídeos por clivagem com tripsina para subsequente análise por espectrometria de massa (Figura 6).

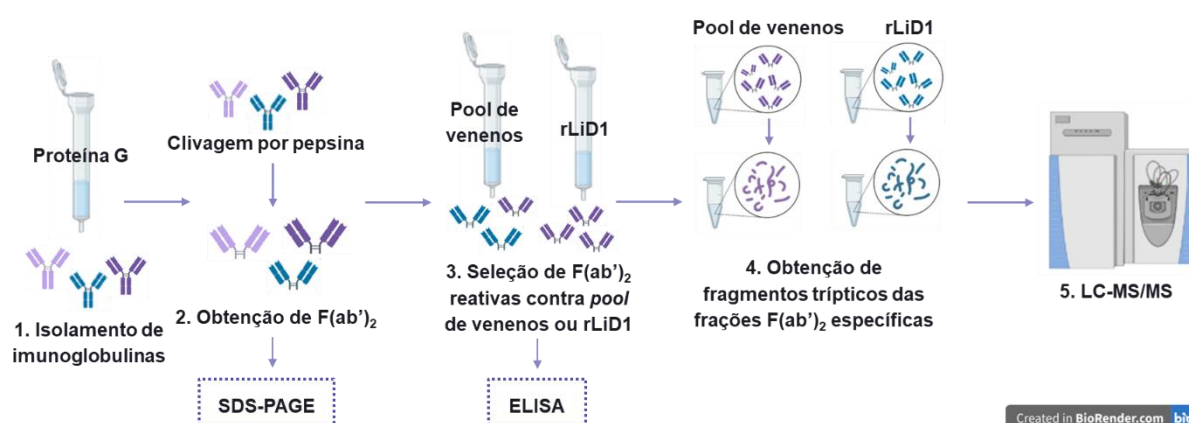


Figura 6 – Preparo das amostras para o Ig-Seq. Etapas realizadas para o preparo das quatro amostras de plasma equino hiperimune: 1) Cromatografia de afinidade com proteína G para obtenção das imunoglobulinas presentes no plasma; 2) Clivagem proteolítica das imunoglobulinas isoladas com pepsina para obtenção das frações $F(ab')_2$ dos anticorpos; 3) Cromatografia de afinidade com o *pool* de venenos ou com rLiD1 para obtenção dos anticorpos com reatividade contra as proteínas imobilizadas; 4) Tratamento das frações $F(ab')_2$ com reatividade específica com ditioneitol (DTT) para redução das pontes dissulfeto, iodoacetamida para alquilação dos grupos tióis, seguido da digestão por tripsina. Os fragmentos trípticos foram dessalinizados e submetidos a cromatografia líquida acoplada à espectrometria de massa.

Adicionalmente, no presente trabalho, foi realizada a quantificação relativa das subclasses de IgG do eluído e *flow-through* (FT) da cromatografia de afinidade com proteína G conforme descrito no tópico 3.3.3 com diluições seriadas de 1:2.000 a 1:256.000.

3.5. Obtenção de seqüências de peptídeos trípticos derivados das frações $F(ab')_2$ de anticorpos

As quatro amostras de plasma equino hiperimune preparadas segundo a descrição do tópico anterior foram submetidas às análises por espectrometria de massa em colaboração com o Prof. Martin R. Larsen do Departamento de Bioquímica e Biologia Molecular da *University of Southern Denmark* (SDU).

Os fragmentos trípticos dessalinizados foram ressuspensos em 0,1% de ácido fórmico e injetados no sistema de cromatografia líquida de alta precisão (HPLC) nano fluxo EASYnLC (Thermo Scientific). O sistema conta com duas colunas de C18: uma pré coluna com 3 cm de comprimento, 100 µm de diâmetro interno e 5 µm de diâmetro das partículas e uma coluna analítica para separação dos peptídeos com 17 cm de comprimento, 75 µm de diâmetro interno e diâmetro das partículas igual a 3 µm. O fracionamento foi realizado com o solvente A (0,1% de ácido fórmico) e o solvente B (90% de acetonitrila/0,1% de ácido fórmico) utilizando o gradiente cromatográfico em porcentagem do solvente B: 3-8% por 7 min; 8-17% por 32 min; 17-20% por 25 min; 20-28% por 20 min; 28-45% por 20 min; 45-95% por 13 minutos, com o fluxo mantido em 300 nL/min.

Os peptídeos eluídos da cromatografia líquida foram gradualmente ionizados por *nanoelectrospray* e direcionados para o equipamento Q Exactive HF-X Hybrid Quadrupole-Orbitrap Mass Spectrometer (Thermo Fisher). Os íons dos peptídeos foram resolvidos no Orbitrap, em polaridade positiva, na faixa de massa de 350 a 1600 m/z, com uma resolução de 120.000 FWHM (do inglês, *full width at half maximum*). O modo de aquisição dependente de dados (DDA) foi utilizado, com a seleção dos 20 íons mais intensos a cada MS para fragmentação através de dissociação por colisão em alta energia (HCD, do inglês *higher-energy C-trap dissociation*), com uma energia de colisão normalizada de 28%. Para a seleção dos Top 20 íons, foi considerado o *threshold* de intensidade de 1×10^4 e a janela de isolamento de 1,2 m/z. Os fragmentos gerados foram resolvidos no Orbitrap, na faixa de massa de 200 a 2000 m/z, com resolução de 15.000 e incluídos na lista de exclusão dinâmica por 30s.

3.6. Preparo do banco de sequências para análise proteômica

De acordo com o descrito na introdução, o método de estudo de repertório de anticorpos utilizado no presente trabalho baseia-se no emprego de bancos de sequências construídos a partir do transcriptoma de imunoglobulinas do mesmo

indivíduo cujo plasma ou soro está sendo investigado. Assim, amostras de células mononucleares do sangue periférico dos mesmos cavalos hiperimunizados por colaboradores do CPPI foram utilizadas para o sequenciamento de nova geração na plataforma da Illumina, seguindo o protocolo descrito por Navas e colaboradores (2022). Os dados de sequenciamento obtidos para a região variável da cadeia pesada de imunoglobulina foram processados de acordo com *pipeline* descrito nessa mesma referência.

Resumidamente, os dados brutos do sequenciamento foram manipulados usando o pRESTO (VANDER HEIDEN et al., 2014), uma ferramenta para o processamento de dados de sequenciamentos *high-throughput* de repertórios de imunoglobulinas. Nessa etapa, os dados das *paired-end reads* são organizados em sequências completas da região variável do IGH. A partir desse resultado, as sequências únicas foram mapeadas aos genes de imunoglobulinas equinas com a ferramenta HighV-Quest do IMGT (ALAMYAR et al., 2010). Foram definidas as regiões V, D e J, CDRs e FRs, assim como a frequência de mutação em toda a extensão da região variável, processo denominado anotação.

No presente estudo, os arquivos em formato .csv gerados a partir da anotação contendo os dados dos genes V, D e J utilizados, assim como as sequências de aminoácidos de cada uma das regiões do anticorpo foram filtrados através de *scripts* em Python. As sequências duplicadas foram colapsadas e foram removidas sequências com dados faltantes ou caracteres inadequados e as que foram identificadas por uma única *read*. Após a aplicação desse filtro, foram gerados arquivos finais em formato FASTA contendo sequências não redundantes da região variável da IGH expandidas até o primeiro sítio de clivagem por tripsina da região constante (adição do trecho ASTTAPK ao final de cada sequência) (Tabela 1).

O conjunto de sequências da região variável IGH filtradas e específicas de cada cavalo foi manualmente concatenado às 247 sequências de contaminantes comuns da proteômica (fasta obtido no *software* MaxQuant contendo sequências de proteínas como queratina, tripsina, albumina, etc.) e às sequências de proteínas equinas (Uniprot *Equus caballus* – 50.370 sequências obtidas em 21/12/2021) e proteínas loxoscélicas (Uniprot *Loxosceles* spp. – 726 sequências obtidas em 21/12/2021). O arquivo FASTA resultante foi denominado “banco de sequências completos para busca” e foi utilizado para a análise dos resultados da espectrometria de massa no *software* Proteome Discoverer™ (Thermo) (Figura 7).

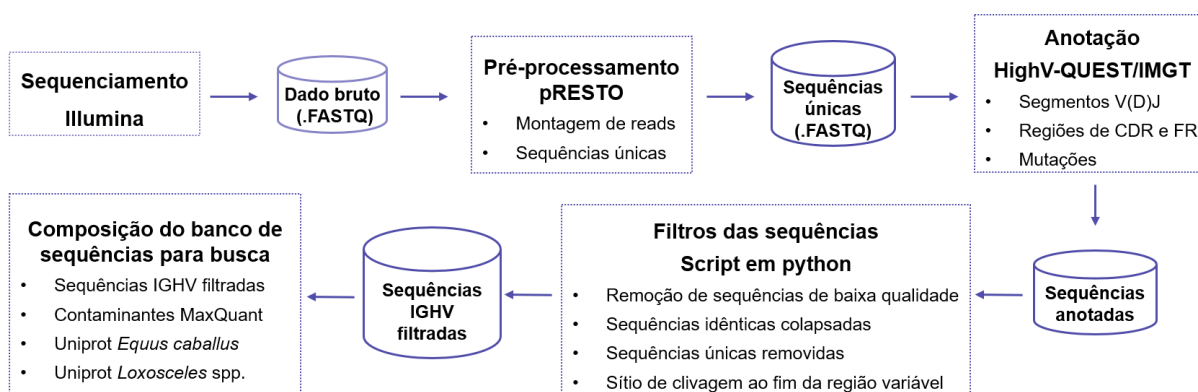


Figura 7 – c. Representação esquemática das etapas descritas para construção dos bancos de sequências utilizados na análise proteômica. O sequenciamento e anotação das sequências foram frutos do trabalho de doutorado da aluna Carlena Tahina Navas de Reyes.

3.7. Análise dos dados de espectrometria de massa para identificação dos peptídeos correspondentes à região IGHV

Como mencionado na seção 3.6, os dados brutos obtidos na espectrometria de massa foram confrontados com os bancos de dados completos para a identificação de peptídeos de anticorpos específicos contra componentes do veneno de *Loxosceles* ou contra rLiD1. Essa etapa foi executada no software Proteome Discoverer™ (Thermo Scientific) na versão 2.4.1.15 com auxílio da ferramenta de busca SEQUEST HT. Os principais parâmetros utilizados na busca foram: uso do módulo de recalibração de massas, clivagem das proteínas por tripsina, tolerância de até duas *missed cleavages*, intervalo de massa do íon precursor variando de 350 a 8000 Da, comprimento do peptídeo variando de 6 a 150 resíduos e tolerância de massa do íon precursor de 10 ppm e do íon fragmento de 0,05 Da. Foi definida como modificação fixa a carbamidometilação da cisteína e como modificações variáveis a oxidação da metionina e a acetilação na extremidade N-terminal da proteína. Somente os peptídeos assinalados com FDR inferior a 1%, classificados como de alta confiança, foram utilizados para as análises posteriores.

3.8. Mapeamento dos peptídeos identificados aos CDRs das sequências de anticorpos equinos

Os resultados exportados do Proteome Discoverer fornecem informações sobre todos os peptídeos cuja a sequência foi identificada no banco personalizado completo, o que pode incluir contaminantes, proteínas de cavalo ou de aranha-marrom. Para avaliar exclusivamente as sequências de anticorpos foi necessário filtrar apenas os peptídeos mapeados às sequências de imunoglobulinas obtidas no Rep-Seq. Porém, partindo desse filtro inicial, um mesmo peptídeo teria o potencial de ser mapeado a diversas sequências distintas, dado que essas moléculas possuem regiões altamente conservadas, como ilustrado na Figura 8. Para contornar essa particularidade, definiu-se que apenas peptídeos mapeados ao CDRH3 de anticorpos seriam considerados para a identificação das sequências.

	FR1	CDR1	FR2	CDR2	FR3	CDR3	FR4		
1.	QVQLK	ESGPGLVK	PSQTLSLTCTVS	GFSASNYG	VDWVVRQAPGKGLEFVGG	LINGRDA	KYTP(...)EDTAVYYC	AGGTR DSSYAYSR NSDETY	WGQGILVTVSS
2.	QVQLK	ESGPGLVK	PSQTLSLTCTVS	GLSLNINV	MGWVVRQAPGKLEFVGA	TYTSA	KYNP(...)EDTAVYYC	AGGSVIWYDR TSYPDSSGIK	Y WGQGILVTVSS
3.	QVQLK	ESGPGLVK	PSQTLSLTCTVS	GLNLNNA	VGWVVRQAPGKLEFVGA	MYTSA	KYNP(...)EDTAVYYC	VGGSVIWYGR TSYPDSSGIK	Y WGQGILVTVSS
4.	QVQLK	ESGPGLVK	PSQTLSLTCTVS	GLSLSSRS	AGWVVRQAPGKLEYVAS	IDSGGGA	NYNP(...)EDTAVYYC	SGR DDGDGTFYYTR TTTY	WGQGIPVTVSS
5.	QVQLK	ESGPGLVK	PSQTLSLTCTVS	GFSLSYD	VGWVVRQAPGKLELVGG	IAGSGSV	TYNP(...)EDTAVYFC	ARGR RNDYAYR VFDD	WGQGLVTVSS

Exemplo de peptídeo tríptico encontrado em várias sequências com distintos CDR3

Exemplos de peptídeos trípticos informativos

Figura 8 – Identificação de sequências de imunoglobulina a partir de peptídeos detectados na espectrometria de massa. São ilustradas cinco diferentes sequências da região variável da IGH de imunoglobulinas que compartilham trechos das regiões de *framework* (FR), mas possuem diferentes composições em seus CDRs. No FR1, é destacado nas cinco sequências um peptídeo tríptico comum que não contribuiu para a identificação de um único IGHV. Na região de CDR3, são destacados os cinco diferentes peptídeos trípticos informativos que permitiram a identificação de cada uma das sequências listadas.

Através de *scripts* em Python, cada um dos bancos de sequências de imunoglobulinas personalizados do Rep-Seq foi inicialmente tratado para remoção de sequências que continham leucina ou isoleucina na mesma posição (e, portanto, poderiam ser idênticas). Os dados tratados foram submetidos à digestão *in silico* por tripsina, sendo toleradas até duas *missed cleavages*, assim como nos dados experimentais. Os peptídeos trípticos teóricos foram mapeados aos diferentes trechos da região variável da cadeia pesada dos anticorpos. Além disso, os dados relacionados aos PSMs (do inglês, *peptide spectrum matches*) exportados do Proteome Discoverer foram preparados, selecionando-se as colunas necessárias: sequência do peptídeo, ID da proteína a qual o peptídeo foi mapeado, área do precursor, cálculo da probabilidade de erro posterior (PEP), XCorr, desvio de massa

em ppm e *first scan* (primeiro MS de identificação do peptídeo). As linhas foram ordenadas de modo a apresentar os dados de *first scan* e PEP em ordem crescente e o XCorr em ordem decrescente. As duplicatas de *first scan* foram removidas, mantendo-se as linhas de maior qualidade de acordo com os valores de PEP e XCorr. Na coluna contendo sequências dos peptídeos, os resíduos de isoleucina foram substituídos por leucina e os peptídeos idênticos após a substituição foram colapsados e contados (dado adicionado a nova coluna PSM) e ainda tiveram calculados seus dados de desvio de massa médio (soma dos dados de desvio de massa dividido pelo nº de PSMs) e de valor área total. O conjunto de peptídeos gerados *in silico* a partir do banco de sequências do Rep-Seq foi confrontado com os peptídeos filtrados encontrados experimentalmente, permitindo definir quais peptídeos de quais sequências da região variável de IGHV foram identificados no proteoma de imunoglobulinas de cada uma das amostras. Apenas as sequências que continham peptídeos mapeados ao CDRH3 foram consideradas como possíveis IGHVs antígeno-específicas.

Porém, os peptídeos mapeados ao CDRH3 dos anticorpos foram submetidos ainda a alguns filtros que garantem a confiabilidade da sua detecção: os peptídeos deveriam ter sido identificados nas duas replicatas e ter sobreposição de no mínimo quatro resíduos com a região de CDRH3. Os fragmentos que cumpriam todos esses critérios foram designados peptídeos informativos, de forma similar a denominação usada em outros trabalhos (LEE et al., 2016; WINE et al., 2013). Por fim, os IGHVs que continham os peptídeos de CDRH3 que satisfaziam aos filtros citados foram considerados para as análises posteriores.

3.9. Alinhamento e comparação de sequências

As sequências da região variável da IGH identificadas por peptídeos mapeados ao seu CDRH3 foram comparadas para avaliação das características típicas de cada amostra ou de cada especificidade. Para isso, as sequências foram alinhadas considerando-se a numeração determinada pelo International ImMunoGeneTics information system (IMGT) que fixa resíduos conservados das sequências nas mesmas posições: cisteína na posição 23, triptofano na posição 41, aminoácido hidrofóbico na posição 89, cisteína na posição 104 e triptofano ou fenilalanina na

posição 118 (LEFRANC et al., 2003). A partir da demarcação desses aminoácidos são definidos os seguintes intervalos para as regiões de CDR e FR:

- a. *Framework 1 (FR1)*: posições 1 a 26
- b. CDR1: posições 27 a 38
- c. *Framework 2 (FR2)*: posições 39 a 55
- d. CDR2: posições 56 a 65
- e. *Framework 3 (FR3)*: posições 66 a 104
- f. CDR3: posições 105 a 117
- g. *Framework 4 (FR4)*: posições 118 a 129

Os alinhamentos das regiões variáveis de IGH completas ou dos trechos de CDRH3 foram realizados através da ferramenta Yvis (CARVALHO; MOLINA; FELICORI, 2019). As compilações dos achados de comprimento e composição de CDRH3, uso de segmentos V(D)J e frequência de hipermutação foram realizadas a partir de *scripts* em Python e os gráficos foram gerados no GraphPad Prism 5.00 (GraphPad Software, Inc., San Diego California USA).

3.10. Determinação de clonotipos e cálculo de abundância a partir dos peptídeos

A totalidade de sequências da região variável do IGH encontradas no Rep-Seq para cada uma das amostras de plasma equino foi agrupada em clonotipos de forma hierárquica utilizando o *software* YClon (GERVÁSIO; FERREIRA; FELICORI, 2022) desenvolvido pelo nosso grupo de pesquisa. Nessa ferramenta, as sequências de nucleotídeos contendo mesmos segmentos gênicos V e J são agrupadas e as regiões de CDR3 são vetorizadas e comparadas de acordo com uma matriz de similaridade. A informação dos clonotipos foi recuperada para cada uma das sequências identificadas no Ig-Seq e considerada para a estimativa de abundância.

A abundância foi determinada em função da área sob a curva do pico do íon precursor que originou o peptídeo identificado. O cálculo consiste na razão entre a soma da área de todos os peptídeos de CDRH3 mapeados a um clonotipo sobre a soma das áreas de todos os peptídeos de CDRH3 mapeados naquela amostra. O

dado de abundância calculado foi utilizado para avaliar a distribuição relativa dos clonotipos dentro de cada amostra.

3.11. Agrupamento de sequências de VH contendo paratopos similares

Para avaliar funcionalmente *in silico* as sequências da região variável de imunoglobulinas encontradas, utilizamos a ferramenta computacional Ab-Ligity (WONG et al., 2021). Nesse software, a estrutura tridimensional de cada região variável é predita e as diferentes sequências têm suas estruturas comparadas. Estruturas que apresentam resíduos de aminoácidos com as mesmas características físico-químicas na mesma posição são apontadas como similares e ranqueadas quanto a potencial de reconhecimento de um epítopo em comum. Com esse método, as sequências de anticorpos com especificidade contra a rLiD1 foram avaliadas quanto a sua similaridade de sítio de interação com o antígeno.

De acordo com as instruções de uso do software, as sequências da região variável da IGH específicas contra rLiD1 tiveram suas estruturas tridimensionais modeladas no ABodyBuilder (LEEM et al., 2016). Os arquivos em formato .pdb gerados foram utilizados para a predição de paratopos e de similaridade de interação com o epítopo através da execução dos *scripts* do Ab-Ligity. Como sugerido no artigo que descreve a ferramenta, o *threshold* de 0,1 foi aplicado para definir as sequências que apresentavam sítios de interação similar.

4. Resultados

4.1. Reatividade dos plasmas equinos contra *pool* de venenos de *Loxosceles* ou rLiD1

As quatro amostras de plasma de cavalos hiperimunizados com o *pool* de venenos de três espécies de *Loxosceles* fornecidas pelo CPPI foram testadas quanto a sua reatividade contra o *pool* de venenos de *Loxosceles* e contra a rLiD1 através de ELISA direto. Foram avaliadas ainda alíquotas de dois lotes dos soros anti-loxoscélicos (SaLox) e um *pool* das amostras de plasmas equinos antes da hiperimunização (*pool* pré-imune).

Como esperado, as amostras coletadas antes da hiperimunização apresentaram absorvância próxima a zero contra os dois tipos de antígenos testados (Figuras 9A e 9B). Frente ao *pool* de venenos de *Loxosceles*, os plasmas dos quatro cavalos imunizados apresentaram titulação média de 60.022, valor superior ao título médio de 48.936 encontrado para os lotes de SaLox testados (Figura 9C). Ainda em relação ao veneno, as amostras apresentaram maior variabilidade no título calculado: os plasmas dos cavalos 2H e 4H possuem maior título de anticorpos na absorvância de 1,0 do que os cavalos 1H e 3H. Quando testadas contra a forma recombinante de LiD1, os plasmas apresentaram reatividade e titulação similar, com valor médio de 46.472 (Figura 9B e 9C). De forma geral, os resultados indicam que os plasmas hiperimunes possuem imunoglobulinas reativas contra componentes do *pool* de venenos de *Loxosceles* ou contra a rLiD1 com título médio comparável à do soro anti-veneno utilizado como terapia. Porém, é importante ressaltar que os títulos encontrados resultam da comparação de imunoglobulinas intactas do plasma equino e frações F(ab')₂ presentes no SaLox, características que podem ter impactado na variação das absorvâncias detectadas nos dois tipos de amostra.

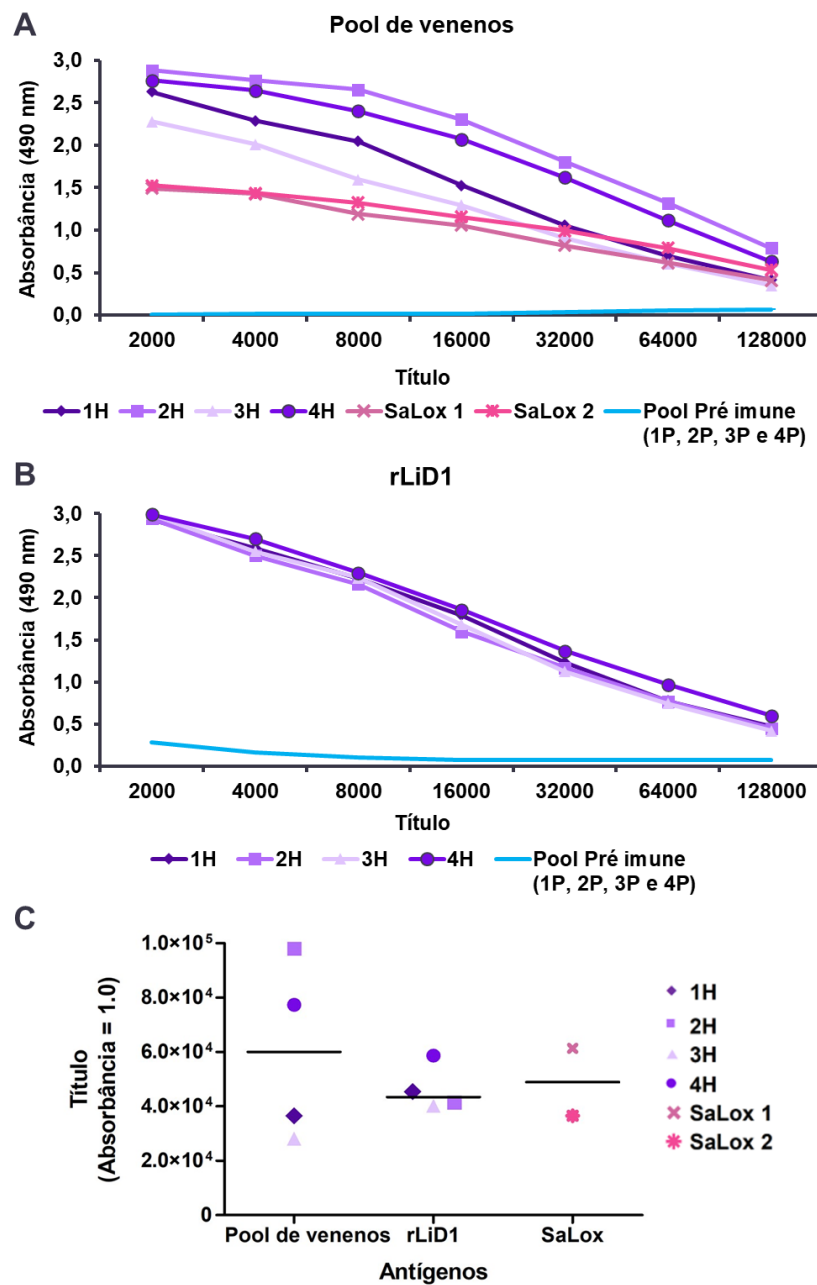


Figura 9 – Imunorreatividade dos plasmas equinos hiperimunizados com o *pool* de veneno de três espécies de *Loxosceles*. Avaliação da reatividade dos plasmas dos cavalos após a hiperimunização (1H, 2H, 3H e 4H) contra o *pool* de venenos de *Loxosceles* (A) e contra a rLiD1 (B). O ensaio foi revelado com anticorpo contra IgG equina conjugado a peroxidase em diluição 1:6000 na absorbância de 490 nm. O anti-veneno SaLox foi utilizado como controle positivo e o *pool* dos plasmas dos cavalos antes da imunização (pré-imune, P) como controle negativo. (C) Título de anticorpos calculado para a diluição dos plasmas imunizados que obteve absorbância igual a 1,0 frente ao *pool* de venenos de *Loxosceles* ou rLiD1 ou título de SaLox com absorbância igual a 1,0 frente ao *pool* de venenos.

4.2. Quantificação absoluta das classes de imunoglobulinas

Na próxima etapa de caracterização do repertório, as amostras dos plasmas antes e após a hiperimunização foram usadas para quantificação de IgM, IgG e IgA utilizando kits comerciais. Para as três classes de imunoglobulinas, as curvas padrão foram traçadas de acordo com as recomendações do fabricante.

Na quantificação de IgM, anticorpo relacionado à resposta imune primária, as amostras 1 e 4 apresentaram concentrações similares dessa imunoglobulina antes e após a hiperimunização (Figura 10A). A amostra 2 apresentou concentração similar à 1 e 4 antes da imunização, porém sofreu redução significativa desse valor após o processo. A amostra 3 na sua forma pré-imune apresentou a maior dosagem de IgM ($1,77 \pm 0,01$ mg/mL) e dosagem significativamente menor dessa imunoglobulina na forma hiperimune. As medianas das concentrações encontradas foram de 0,63 mg/mL para as amostras hiperimunes e de 0,61 mg/mL para as amostras pré-imunes.

Na quantificação de IgG, imunoglobulina relacionada à resposta imune secundária, as menores concentrações foram encontradas nas amostras pré-imunes 1 e 3 que tiveram aumento significativo dessa imunoglobulina após a hiperimunização (Figura 10B). Na amostra 2, observa-se uma alta dosagem de IgG no plasma pré-imune ($160,94 \pm 8,21$ mg/mL) e esse valor sofre pequena redução na dosagem no plasma hiperimune ($148,26 \pm 18,80$ mg/mL). A amostra 3 tem concentração intermediária em relação as outras amostras pré-imunes ($22,52 \pm 2,57$ mg/mL) e passa por um aumento na quantidade de IgG após a hiperimunização ($171,04 \pm 16,08$ mg/mL). As amostras de plasma equino hiperimunizado tiveram a concentração mediana de IgG de 142,21 mg/mL, enquanto as amostras coletadas antes da imunização tiveram a mediana de 44,04 mg/mL. Considerando apenas os achados de quantificação de IgM e IgG, foi possível observar um perfil de desenvolvimento de resposta imune secundária em substituição a resposta primária nas amostras dos plasmas avaliados.

Para a quantificação de IgA, a curva padrão foi traçada nos mesmos moldes das outras classes de imunoglobulina, porém as concentrações das amostras não puderam ser interpoladas por estarem abaixo do valor mínimo da curva, que é de 15,63 ng/mL. Esse resultado condiz com a característica de baixa concentração, ou até mesmo ausência, das imunoglobulinas de classe A na composição de soros anti-veneno.

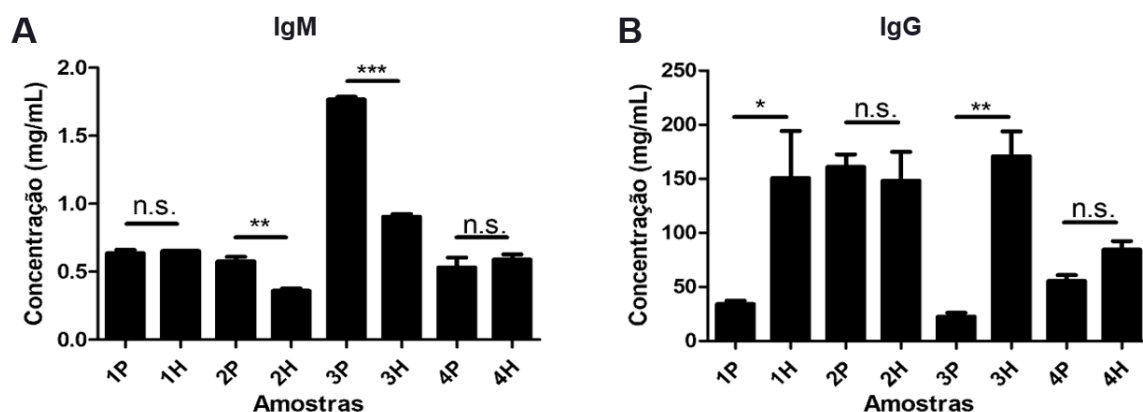


Figura 10 – Quantificação de IgM e IgG nos plasmas equinos. Concentrações de IgM (A) e IgG (B) obtidas a partir de interpolação da curva padrão para cada uma das amostras que foram avaliadas em duas replicatas. Os plasmas pré-ímmunes de cada cavalo são representados como 1P, 2P, 3P e 4P, e os plasmas hiperímmunes como 1H, 2H, 3H e 4H. Cada amostra foi analisada em duplicata.

4.3. Quantificação relativa das subclasses de IgG

Novamente, um ensaio de ELISA baseado na sensibilização com o *pool* de venenos de *Loxosceles* foi utilizado para comparar a reatividade das diferentes subclasses de IgG. Foram testadas amostras dos plasmas antes e após a hiperimunização, assim como amostras das imunoglobulinas isoladas a partir da cromatografia de afinidade com proteína G e da fração não interagida (FT, do inglês *flow through*) dessa mesma cromatografia.

Nas amostras de plasmas pré-ímmunes, as absorvâncias dos anticorpos antígeno-específicos das classes IgGa e IgGb foram quase nulas, com valores inferiores a 0,1 em todas as diluições. Para IgG(T), foram encontradas absorvâncias similares para as amostras 1P, 2P e 3P, mas com valores consideravelmente inferiores aos das mesmas amostras após a hiperimunização (Figura 11).

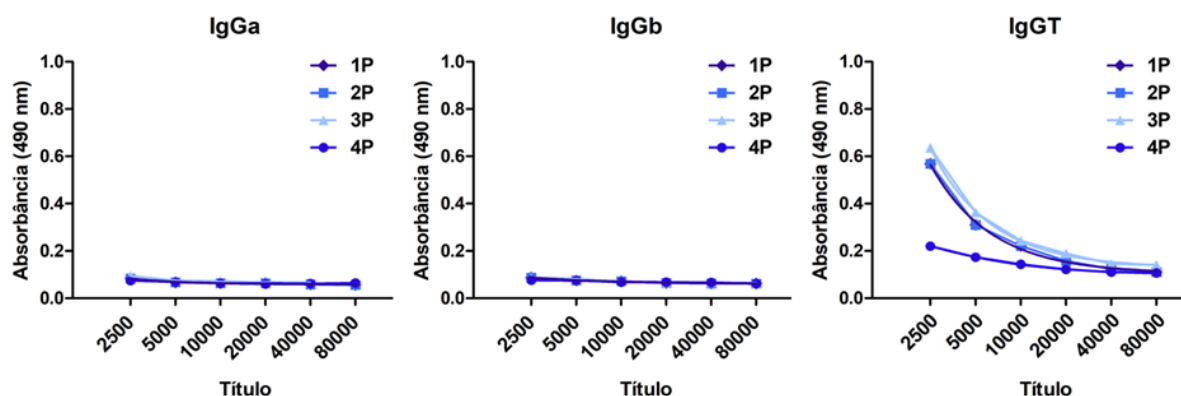


Figura 11 – Distribuição de subclasses de IgG nos plasmas equinos pré-ímunos.

Absorbâncias obtidas para as subclasses IgGa, IgGb e IgG(T) nas amostras de plasmas dos quatro cavalos antes da hiperimunização (1P, 2P, 3P, 4P) no comprimento de onda de 490 nm. ELISA indireto com a placa sensibilizada com o *pool* de venenos. As amostras foram testadas em duplicatas em diluição seriada de 2.500 a 80.000 vezes (os pontos indicam a absorbância média para cada diluição). Os ensaios foram realizados com conjugados anti-mouse para IgGa e IgGb e anti-goat para IgG(T) e revelados com OPD.

Nas amostras hiperímunes, observa-se maior titulação média de IgG(T), seguida por IgGb e por último IgGa (Figura 12A), achado similar a composição de outros anti-venenos investigados na literatura (FERNANDES et al., 1997). A amostra 2H destaca-se com maior título de IgG(T) e IgGa no valor de absorbância selecionado e o segundo maior valor de IgGb, enquanto as amostras 1H e 3H possuem valores abaixo da média dos três subtipos de IgG. Em relação às amostras de IgG isolada, a maior titulação média é encontrada na subclasse IgGb, seguida por IgGa e por último IgG(T) (Figura 12B). As amostras 2H e 4H destacam-se com títulos acima da média na subclasse IgGb. Nas amostras de FT, predomina a subclasse IgG(T) enquanto IgGa e IgGb mantêm títulos baixos e com médias similares (Figura 12B). A amostra 3H não teve o título de IgGb em seu FT determinado, por apresentar absorbância inferior a 1,0 em todas as diluições.

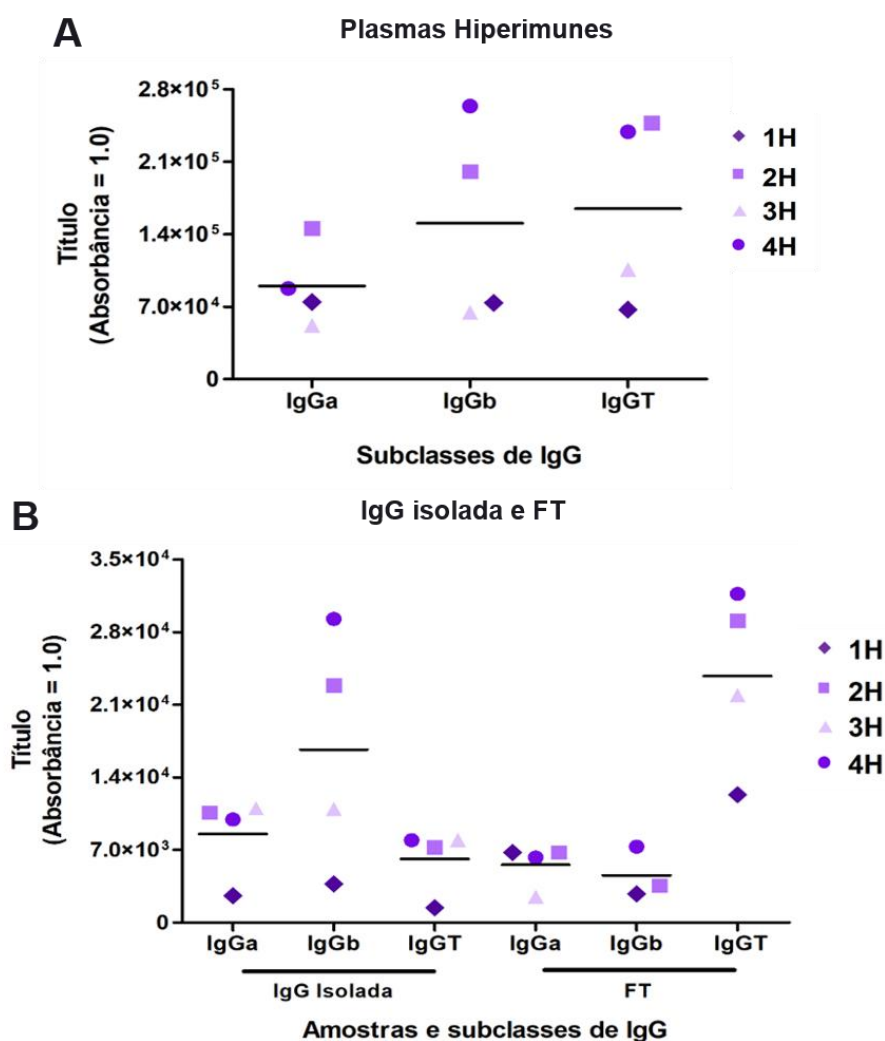


Figura 12 – Títulos das subclasses de IgG presentes nos plasmas hiperimunes e na cromatografia de afinidade com proteína G. Títulos das subclasses IgGa, IgGb e IgG(T) calculados para a diluição dos plasmas hiperimunes (A) e das IgGs isolada e FTs da coluna de afinidade com proteína G (B), para obter absorvância igual a 1,0 no comprimento de onda de 490 nm. Para cada uma das subclasses é destacada a média do título das quatro amostras. Os títulos foram calculados a partir das curvas geradas por ELISA indireto com a placa sensibilizada com o *pool* de venenos. As amostras de plasma e IgG isolada foram testadas em duplicatas, em diluição seriada de 5.000 a 640.000 vezes e as amostras de FT em diluições de 2.000 a 256.000 vezes. Os ensaios foram revelados com conjugados *anti-mouse* para IgGa e IgGb e *anti-goat* para IgG(T). FT: *flow through*, não interagido.

Os dados expostos na Figura 12 indicam que o enriquecimento na cromatografia de afinidade com proteína G privilegia as subclasses IgGb e IgGa, em detrimento da IgG(T), que era encontrada em maior abundância nos plasmas equinos. Como consequência, nota-se o elevado título de IgG(T) no *flow through* da

cromatografia. A perda dessa subclasse pode impactar os resultados subsequentes baseados nas amostras de IgG isolada.

4.4. Composição do banco de sequências personalizado para busca

Para iniciar as etapas de identificação dos peptídeos obtidos a partir da espectrometria de massa, os dados de sequências anotadas do transcriptoma de imunoglobulinas foram filtrados para remoção de linhas com dados incompletos e caracteres inadequados (denominadas sequências de baixa qualidade), sequências identificadas por uma única *read* e colapso das sequências em replicatas. Em média, para cada cavalo foram anotadas 481.140 ± 176.230 sequências das quais cerca de 0,05% foram removidas pela baixa qualidade. As sequências anotadas já haviam passado pela exclusão de replicatas em sequências de nucleotídeos, porém, a partir da tradução para resíduos de aminoácidos, em média 160.000 sequências por amostra foram encontradas em replicatas. Por fim, sequências que apareciam uma única vez na anotação foram também removidas para reduzir a dimensão dos dados, gerando conjuntos finais com, em média, 29.241 sequências por amostra (Tabela 1).

Tabela 1 – Filtros aplicados para construção do banco de sequências de cada amostra

Amostra	Sequências anotadas	Sequências de qualidade	Sequências em replicatas colapsadas	Sequências colapsadas e de abundância > 1
1H	317.894	316.580	240.639	17.381
2H	653.023	649.005	415.163	38.822
3H	340.916	339.457	226.612	21.946
4H	612.728	607.982	408.314	38.815
Média e desvio padrão	481.140 ± 176.230	478.256 ± 174.536	322.682 ± 103.031	29.241 ± 11.215

4.5. Análise dos dados do Ig-Seq para identificação dos peptídeos mapeados à região variável de IGH

Para a próxima etapa, os plasmas hiperimunes foram preparados para a identificação dos anticorpos antígeno-específicos através da espectrometria de massa. Para isso, as amostras foram submetidas a cromatografia de afinidade com proteína G, clivagem por pepsina, enriquecimento antígeno-específico contra *pool* de venenos ou rLiD1 e clivagem por tripsina. Em sequência, os peptídeos trípticos foram analisados por LC-MS/MS. Os espectros de massa gerados foram confrontados contra os bancos de dados personalizados de cada cavalo contendo as sequências IGHV e proteínas contaminantes descritas na seção 3.6.

Poucos peptídeos identificados foram mapeados a proteínas de *Loxosceles* depositadas no Uniprot. Nas amostras com especificidade contra rLiD1, todas as identificações desse subgrupo correspondiam a proteínas descritas como fosfolipase D, esfingomielinase D ou proteína LoxTox. Esse achado possivelmente indica que a cromatografia de afinidade com rLiD1 foi a fonte das proteínas de *Loxosceles* identificadas. Nas amostras específicas contra o *pool* de venenos, além das fosfolipasas D, foram identificadas metaloproteases nas amostras 1H e 2H, proteína *knottin* (peptídeo inseticida) nas amostras 1H e 4H e um fragmento de histona nas amostras 1H e 3H.

Por outro lado, foram identificados um elevado número de peptídeos de proteínas contaminantes, com uma média de 555 peptídeos por amostra (Figura 13). Nas amostras específicas contra rLiD1 ou contra o *pool* de venenos, em média 24% e 17%, dos peptídeos identificados correspondiam a contaminantes, respectivamente. Apesar das precauções na realização dos experimentos e do uso de tripsina com baixa taxa de autólise, queratina humana e a tripsina porcina usada para o *bottom-up* foram encontradas em todas as amostras.

Dentre os peptídeos mapeados a proteínas equinas presentes no UniProt, em média 49% e 38% foram identificados como imunoglobulinas dessa espécie nas amostras específicas contra o *pool* de venenos e rLiD1, respectivamente. Foram identificados trechos de sequências de cadeia pesada gama e cadeia leve lambda de cavalos. Além disso, foram encontrados peptídeos mapeados a proteínas comuns do

sangue como albumina, alfa-2-macroglobulina, fibrinogênio, apolipoproteínas, entre outras.

Com exceção da amostra 1H com especificidade contra rLiD1, a maior parte dos peptídeos identificados foi mapeado às sequências da região variável do IGH dos próprios cavalos avaliados (Figura 13). As amostras 3H e 4H apresentaram maior número de peptídeos de IGH identificados contra os dois tipos de antígenos, assim como maior número total de identificações. Esse achado reforça a importância da construção do banco de sequências personalizado para cada organismo testado, dada a variabilidade das cadeias das imunoglobulinas e a escassez dessas sequências nos bancos públicos de dados de proteínas.

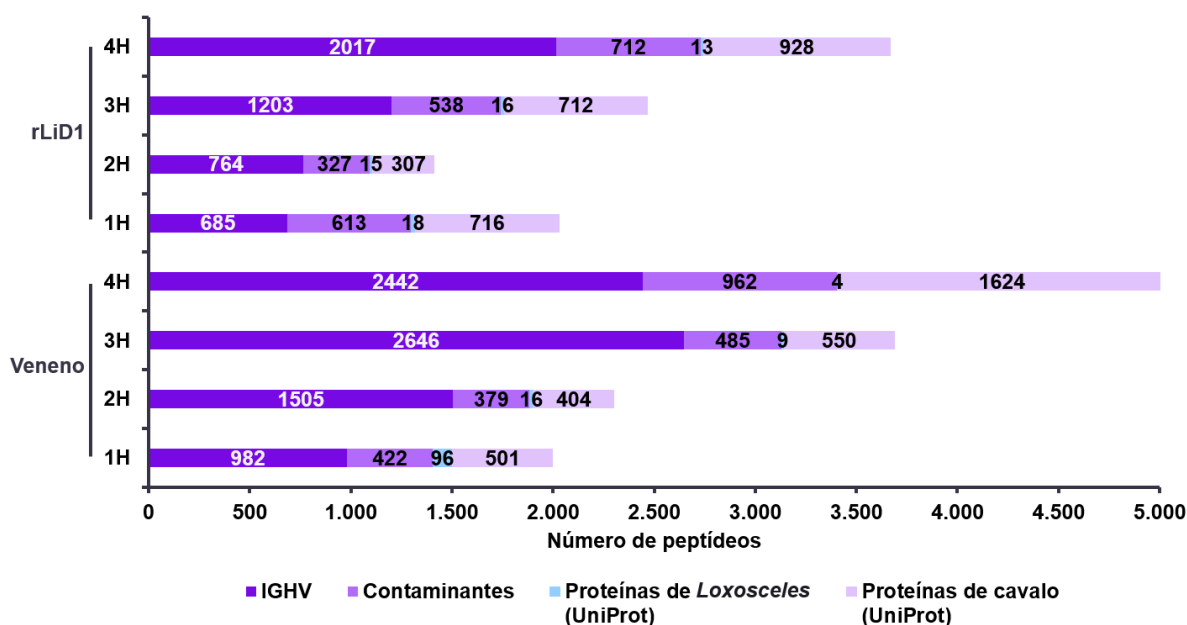


Figura 13 – Identificação de peptídeos trípticos nas amostras de frações $F(ab')_2$ antígeno-específicas. Número de peptídeos trípticos não redundantes mapeados a cada um dos grupos de proteínas presentes no banco de sequências personalizado. São eles: sequências da região variável do IGH específicas para cada cavalo (Rep-Seq), contaminantes (MaxQuant), proteínas de *Loxosceles* e de cavalo presentes no UniProt.

Em sequência, os peptídeos identificados da região variável da IGH foram mapeados quanto a sua região específica no anticorpo. As amostras 3H e 4H apresentaram o maior número de peptídeos em todas as regiões de imunoglobulina. A região de CDRH2 apresentou o maior número de peptídeos, com uma média de 515 por amostra, seguida pelo CDRH1, com média de 143 peptídeos e, por fim, a região de *framework* com média de 77 peptídeos. A região de CDRH3 foi escolhida como

parâmetro para identificação das sequências de anticorpos pela sua variabilidade, como feito em trabalhos anteriores baseados nos mesmos métodos. Porém, essa foi a região com menor número de peptídeos identificados em todas as amostras (Figura 14).

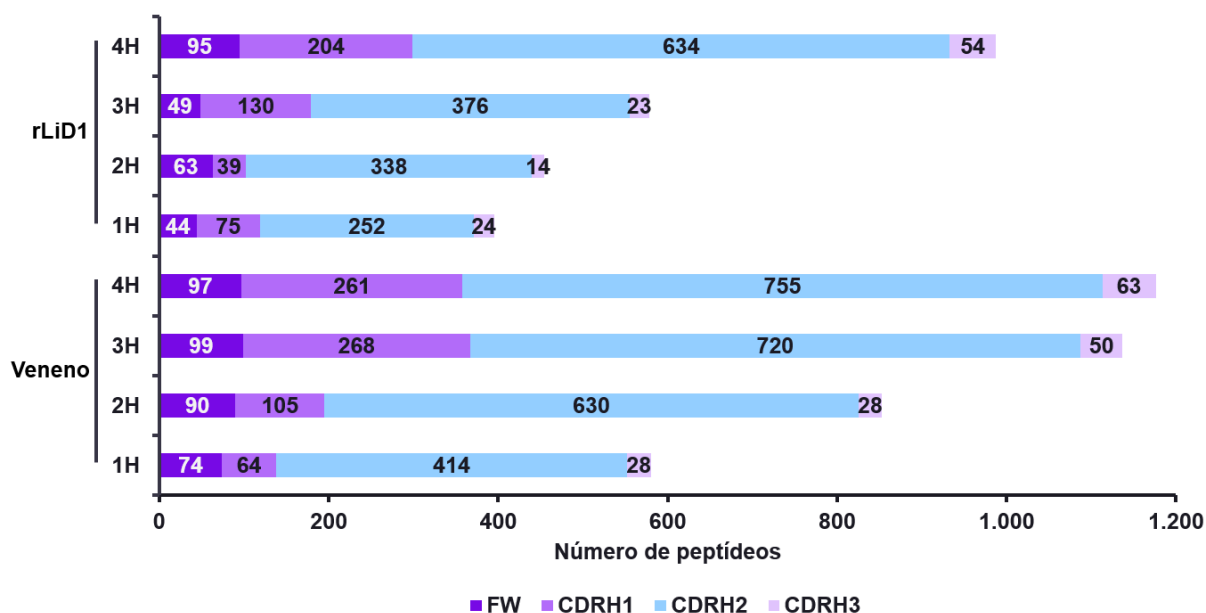


Figura 14 – Peptídeos mapeados por região da cadeia variável do anticorpo. Para cada amostra e cada antígeno são ilustradas as contagens de peptídeos mapeados as regiões determinantes de complementariedade (CDR) 1, 2 ou 3 ou a região estrutural de *framework* (FW).

A frequência de sítios de clivagem por tripsina, enzima utilizada no preparo para a espectrometria de massa, foi avaliada como uma possível causa da baixa identificação de resíduos de CDRH3. De fato, a presença de sítios de clivagem por tripsina (resíduos de lisina ou arginina) na posição 107 do CDRH3 é baixa e não representa um sítio âncora para obtenção de fragmentos de CDRH3, como proposto em trabalhos realizados com outras espécies (Figura 15A) (WINE et al., 2013). Em média, o sítio tríptico na posição 107 é encontrado em 17,04% das sequências dos bancos de dados, com maior frequência no cavalo 1H (28,09% das sequências) e menor frequência no cavalo 3H (9,57% das sequências). De forma similar aos achados experimentais, os peptídeos trípticos gerados *in silico* são mais frequentemente mapeados a região de CDRH2 (Figura 15B). Porém, na digestão teórica, os fragmentos de CDRH3 aparecem como a segunda região mais abundante em contraposição aos achados experimentais.

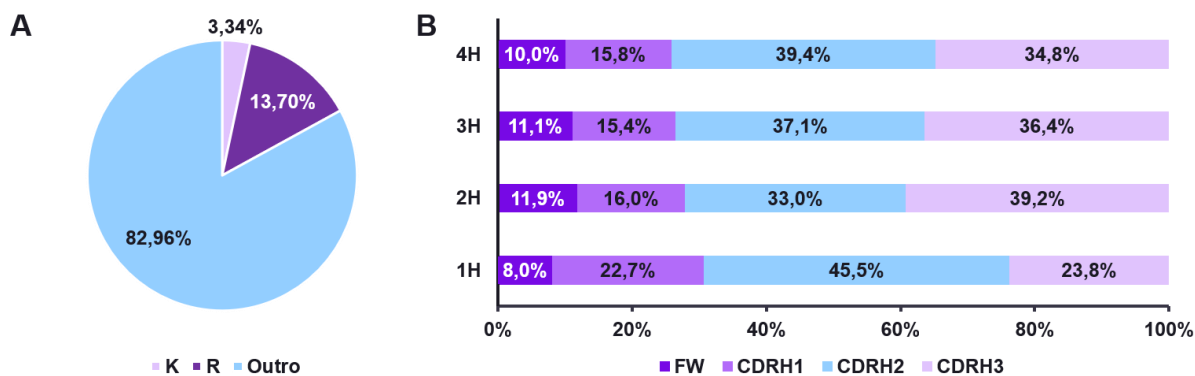


Figura 15 – Frequência de resíduos de lisina (K) e arginina (R) na posição 107 do CDRH3 e de peptídeos trípticos por região da fração variável do IGH. A ocupação da posição 107 do CDRH3, um sítio conservado de clivagem por tripsina em outras espécies, foi avaliado na totalidade de sequências encontradas do Rep-Seq equino (A). Foi avaliada ainda a distribuição de peptídeos trípticos mapeados às diferentes regiões dos anticorpos através da digestão *in silico* das sequências do Rep-Seq (B).

4.6. Análise das sequências da região variável de IGH identificadas a partir do CDRH3

Como detalhado na metodologia, os peptídeos mapeados a CDRH3 e que atendiam a critérios específicos de qualidade foram utilizados para identificação de sequências IGHV. Foram aceitas apenas sequências encontradas nas duas replicatas experimentais e que possuíam sobreposição de no mínimo quatro resíduos entre o peptídeo encontrado e a região de CDRH3. Os fragmentos qualificados foram então mapeados a todos possíveis CDRH3 e a todas possíveis sequências IGHV que continham dado CDRH3. Nesse processo, um único peptídeo poderia ser mapeado a diferentes CDRH3 (que em alguns casos apresentavam pequenas variações em suas sequências) e um mesmo CDRH3 poderia ser encontrado em mais de uma sequência da região variável da IGH do banco de sequências. Os dados do número de peptídeos e de CDRH3s e IGHVs aos quais os peptídeos foram mapeados são demonstrados na Tabela 2.

A amostra 2H apresentou menor número de peptídeos, CDRH3 e IGHV identificados para os dois tipos de antígenos. As amostras 3H e 4H com especificidade ao *pool* de venenos apresentaram distribuição similar dos três tipos de achados (Tabela 2). Na amostra 1H, uma sequência de CDRH3 (SGRDDGDGTFYYTRTTTY)

foi mapeada a partir de dois peptídeos trípticos: DDGDGTFYYTR e TTTYWGQGIPVTVSSASTTAPK, sendo o fragmento DDGDGTFYYTR encontrado nas amostras específicas contra o *pool* de venenos e contra a rLiD1.

Tabela 2 – Número de peptídeos, CDRH3 e sequências IGHV identificadas no Ig-Seq

Antígeno	Amostra	Peptídeos	CDRH3	Sequências IGHV
Veneno	1H	9	18	54
	2H	6	7	26
	3H	12	17	32
	4H	13	13	32
Média e Desvio Padrão		10 ± 3,16	13,75 ± 4,99	36,25 ± 12,82
rLiD1	1H	8	8	11
	2H	1	1	2
	3H	6	7	10
	4H	5	5	8
Média e Desvio Padrão		5 ± 2,94	5,25 ± 3,10	7,75 ± 4,03

A amostra 1H (*pool* de venenos) destacou-se por apresentar um peptídeo tríptico (TSYPDDSGIK) mapeado à 8 diferentes CDRH3. Dentre eles, o fragmento AGGSVIWYDRTSYPDDSGIKY foi encontrado em 30 diferentes sequências da região variável da IGH. O conjunto de sequências que compartilham esse CDRH3 diferenciam-se principalmente por variações na região de CDRH1, mas também há diversidade no CDRH2 e nas regiões de *framework* (Figura 16). Essa peculiaridade coloca a amostra 1 com o maior número de sequências IGH e CDRH3 apesar do menor número de peptídeos encontrados. Esse achado surpreende ao colocar o CDRH3, região com maior potencial de variabilidade dos anticorpos, como elemento constante em diversas sequências que se diferenciam a partir de outras regiões de menor variabilidade teórica. No caso das sequências específicas à rLiD1, as amostras 1, 3 e 4 tiveram números similares, enquanto a amostra 2 apresentou números muito inferiores de IGHV identificados.

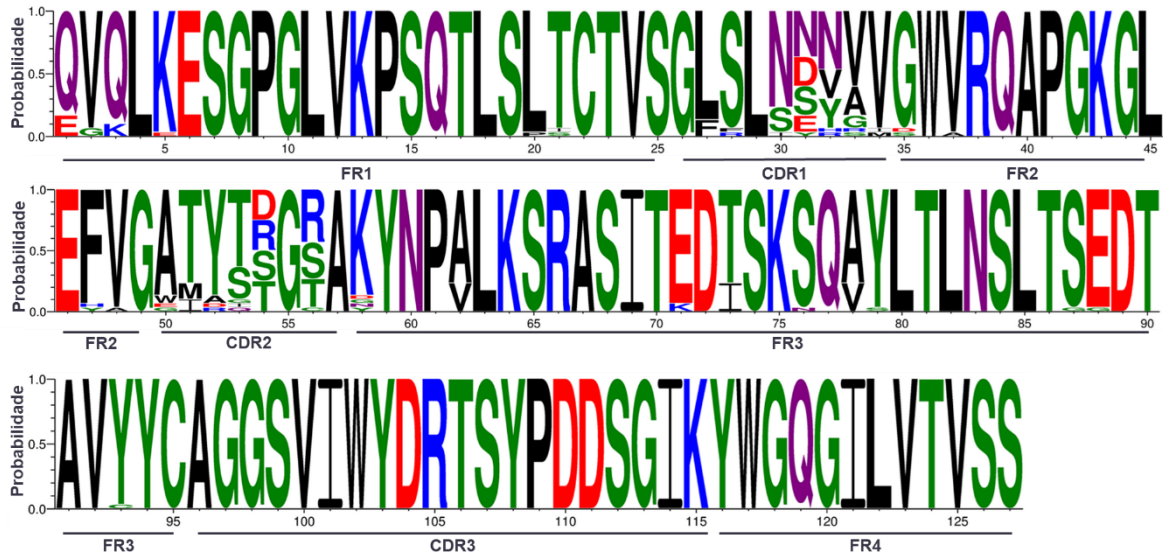


Figura 16 – Composição das 30 seqüências da região variável de IGH com CDRH3 em comum encontradas na amostra 1H específica contra o *pool* de venenos. Os gaps gerados por posições não ocupadas em nenhuma das seqüências foram removidos gerando alteração nas posições canônicas das regiões do anticorpo. Diferentes cores dos aminoácidos representam suas classes. Verde: polares, roxo: neutros, azul: carregados positivamente, vermelho: carregados negativamente, preto: hidrofóbicos. Figura gerada em <http://weblogo.threeplusone.com/>.

Considerando as 163 seqüências únicas da região variável de IGH identificadas a partir do CDRH3, 81,1% são exclusivas aos componentes do veneno de *Loxosceles*, 11,6% são exclusivas à rLiD1 e 7,3% são comuns entre os dois antígenos (Figura 17A). A amostra 1H possui o maior número de seqüências exclusivas com as duas especificidades, enquanto a amostra 4H possui o maior número de seqüências comuns entre os dois antígenos (Figura 17B).

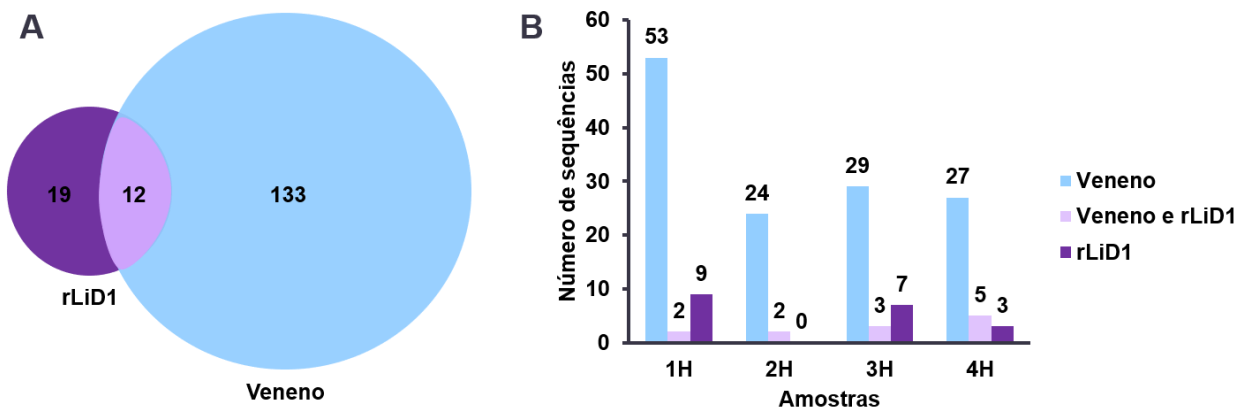


Figura 17 – Sequências da região variável da IGH identificadas a partir do CDRH3. Distribuição do número geral de sequências específicas e comuns a cada antígeno (A) (*pool* de veneno de *Loxosceles* e a rLiD1) e por amostras (B).

Em seguida, as sequências de CDRH3 antígeno-específicas foram caracterizadas quanto ao seu comprimento e composição, e as sequências de IGVH antígeno-específicas foram caracterizadas quanto a frequência gênica e hipermutação. Essas características foram comparadas aos das sequências totais identificadas no transcriptoma e serão descritas nas seções posteriores.

4.7. Tamanho e composição dos CDRH3 identificados

A frequência relativa dos diferentes tamanhos de CDRH3 são ilustradas na Figura 18A. Os comprimentos de CDRH3 identificados no Ig-Seq variam entre 5 a 24 aminoácidos, enquanto os comprimentos do Rep-Seq variam entre 2 a 36 aminoácidos. As principais diferenças são observadas nos comprimentos de 21, 18 e 16 resíduos de aminoácidos que são mais frequentes no Ig-Seq e nos comprimentos de 20, 14 e 11 resíduos que foram mais frequentes no Rep-Seq. Quando separadas por antígeno, as sequências específicas contra rLiD1 possuem maior frequência de CDRH3 com comprimento de 16 resíduos enquanto as específicas contra o *pool* de venenos possuem mais sequências com tamanhos de 21, 18 e 15 resíduos. A mediana dos comprimentos obtidos foi de 17 resíduos para o repertório sorológico e de 16 para o repertório de BCR.

Na Figura 18B, são expressas as frequências relativas de cada aminoácido que compõe os CDRH3 identificados no Ig-Seq e no Rep-Seq. Os resíduos de aminoácidos mais frequentes em ambos os repertórios são glicina e tirosina. Os resíduos que aparecem com menor frequência como metionina, cisteína, glutamina, histidina, prolina e ácido glutâmico possuem distribuição semelhante nos dois tipos de repertório. As maiores variações são notadas nos resíduos carregados de arginina e ácido aspártico, que são mais frequentes no repertório sorológico e de tirosina, que tem maior presença no repertório de BCR.

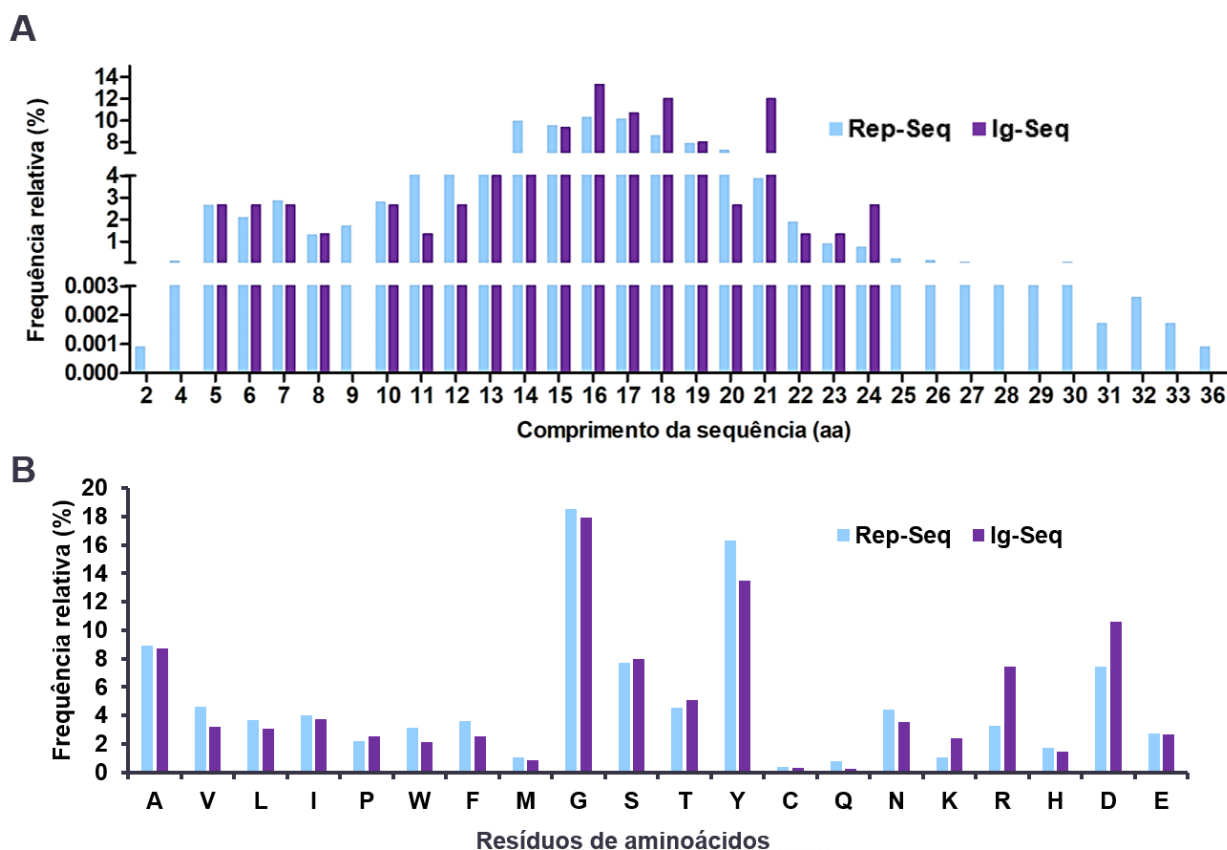


Figura 18 – Distribuição do comprimento e dos aminoácidos das sequências de CDRH3 encontradas no Ig-Seq e no Rep-Seq. (A) Frequência relativa dos comprimentos das sequências únicas de CDRH3, em número de aminoácidos (aa). (B) Frequência relativa dos aminoácidos presentes na região de CDRH3. Os aminoácidos estão agrupados de acordo com sua classe.

A frequência dos aminoácidos nas diferentes posições do CDRH3 também foi avaliada através da construção de WebLogos para as sequências específicas contra rLiD1, *pool* de venenos e para as sequências totais do Rep-seq removendo as redundâncias. Devido à grande variação do comprimento dessa região, foram comparados apenas os sete resíduos iniciais e os seis finais dos CDRH3 alinhados, de acordo com a classificação do IMGT.

Uma alta frequência de alanina (A), glicina (G) e tirosina (Y) foi observada nas posições 105, 106 e 117, respectivamente, em todos os repertórios (Figura 19). Na antepenúltima posição é mantido o predomínio de isoleucina e aminoácidos hidrofóbicos. Nos CDRH3 específicos contra rLiD1, é notável a maior frequência de ácido aspártico (D) nas posições 107 a 109 além da presença dos resíduos carregados positivamente arginina (R) e lisina (K) na posição 113 e dos resíduos

carregados negativamente ácido aspártico e ácido glutâmico (E) na posição 116. Nas sequências específicas contra o *pool* de venenos, o ácido aspártico é mais presente na posição 112 e a arginina na posição 113 em comparação ao Rep-Seq. A posição 116 também apresenta alta probabilidade desses resíduos carregados. A arginina desponta ainda com maior frequência nas sequências encontradas contra o *pool* de venenos na posição 108. De modo geral, podemos observar a preferência por aminoácidos carregados em diferentes posições do CDRH3 do repertório sorológico específico contra os antígenos investigados.

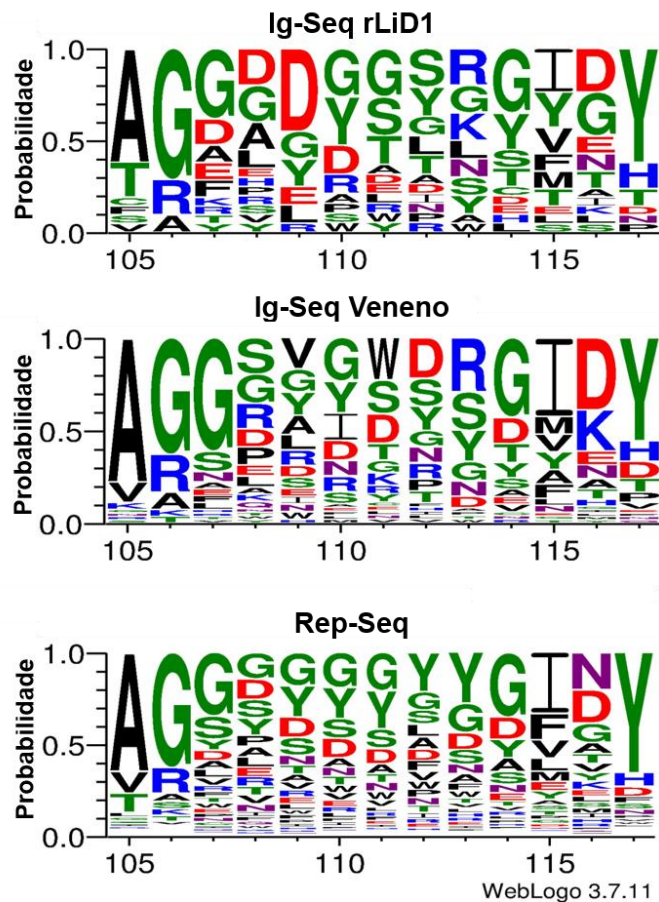


Figura 19 – Composição de aminoácidos por posição nos CDRH3 antígeno-específicos não redundantes. Weblogo indicando a probabilidade da presença de resíduos de aminoácidos em cada posição dos CDRH3 não redundantes do repertório Ig-Seq específico contra rLiD1 (A), *pool* de venenos de *Loxosceles* (B) e do repertório BCR total (Rep-Seq) (C). As figuras contêm os sete resíduos iniciais e os seis finais dos CDRH3 encontrados, sendo omitidas as posições intermediárias encontradas apenas em CDRH3 mais longos. Posições compostas por letras mais estreitas possuem menor probabilidade de estarem ocupadas por qualquer resíduo. Diferentes cores dos aminoácidos representam suas classes. Verde:

polares, roxo: neutros, azul: carregados positivamente, vermelho: carregados negativamente, preto: hidrofóbicos. Figura gerada em <http://weblogo.threeplusone.com/>.

4.8. Preferência de uso de genes V(D)J e frequência de hipermutação somática

Em uma última comparação, foram avaliadas as frequências dos segmentos gênicos V, D e J nas imunoglobulinas encontradas no repertório Ig-Seq e as obtidas no repertório resultante do Rep-Seq. De forma geral, as frequências do uso gênico são similares nos dois repertórios. No caso do segmento gênico V, uma alta predominância da família IGHV4 foi observada, com maior frequência alternada dos segmentos gênicos IGHV4-21 e IGHV4-29 (Figura 20A) entre os repertórios. É importante ressaltar que somente 6 dos 17 segmentos gênicos do Rep-Seq foram identificados no repertório específico do Ig-Seq. Quanto aos genes D e J, os segmentos IGHD2-26 (Figura 20C) e IGHJ6-1 (Figura 20B) foram os mais frequentes nos dois repertórios avaliados, sendo observada frequência consideravelmente maior de IGHD2-26 no Ig-Seq em comparação ao Rep-Seq. Parte dos genes V e D de baixa frequência no Rep-Seq não foram encontrados no repertório sorológico. As combinações de genes V(D)J encontradas com maior frequência no Ig-Seq foram IGHV4-21, IGHD2-26, IGHJ6-1 e IGHV4-29, IGHD2-26, IGHJ6-1, ambas correspondendo a 11% do total de sequências encontradas.

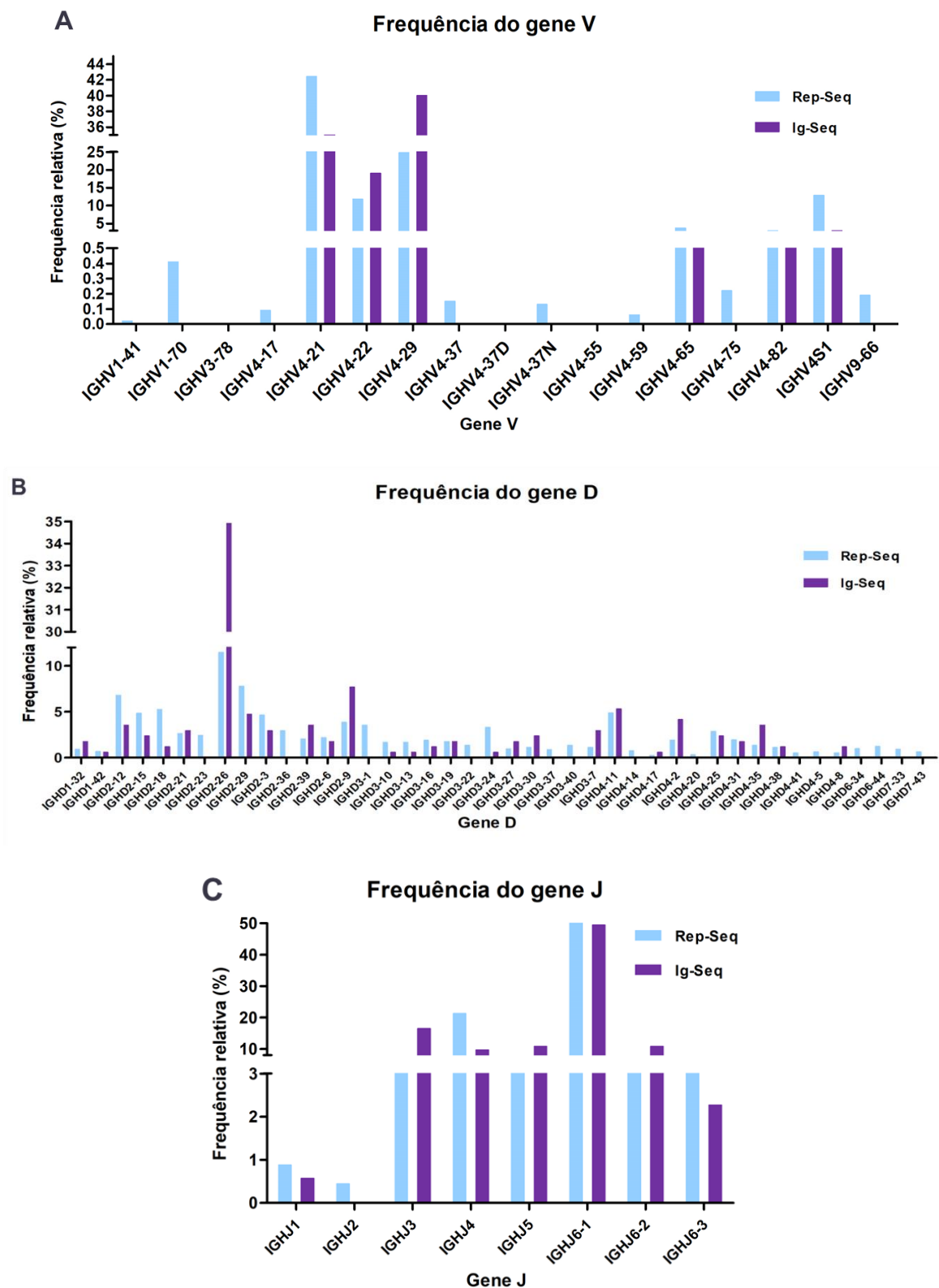


Figura 20 – Preferência dos segmentos gênicos IGH da cadeia variável dos anticorpos. Frequência relativa do uso dos segmentos gênicos variável - V (A), diversidade - D (B) e junção - J (C) das imunoglobulinas encontradas no repertório sorológico (Ig-Seq) e no repertório BCR (Rep-Seq).

A taxa de hipermutação somática (HMS) foi calculada em função do número de alterações nos nucleotídeos das sequências encontradas em relação aos genes *germline* atribuídos a essas sequências. A distribuição dessa taxa nos repertórios é demonstrada na Figura 21. A mediana encontrada foi similar no Ig-Seq com as duas especificidades e no Rep-Seq. Apesar do menor número de achados, as sequências específicas contra rLiD1 apresentam maior variação na taxa de HMS quando comparadas às sequências específicas contra o *pool* de venenos. Somente no repertório descoberto a partir do transcriptoma são encontradas sequências com altas taxa de hipermutação que variam até 50%.

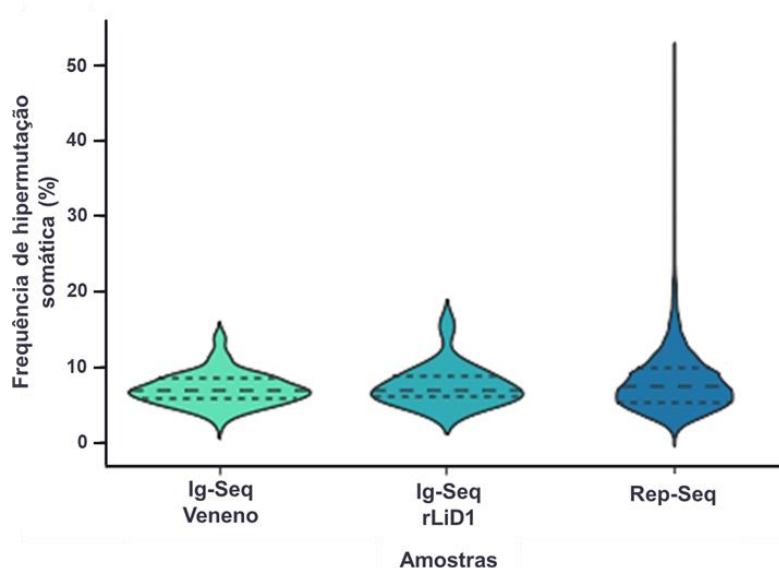


Figura 21 – Frequência de hipermutação somática na região variável da cadeia pesada das imunoglobulinas encontradas nos diferentes repertórios estudados. A distribuição das frequências foi calculada para as sequências específicas contra o *pool* de venenos, contra rLiD1 e no repertório do transcriptoma completo. A linha pontilhada central indica a mediana e as linhas com pontilhado fino indicam o primeiro e o terceiro quartis dos dados.

4.9. Identificação de clonotipos e cálculo de abundância

O conjunto de sequências encontradas no sequenciamento de células B (Rep-Seq) foi organizado em clonotipos a partir do *software* YClon (GERVÁSIO; FERREIRA; FELICORI, 2022), desenvolvido pelo nosso grupo de pesquisa. Os clonotipos das sequências da região variável do IGH antígeno-específicas identificadas foram recuperados a partir da clusterização do Rep-Seq e tiveram sua

abundância relativa calculada pela soma da área sob a curva dos peptídeos trípticos mapeados.

Para as imunoglobulinas específicas contra rLiD1, 8 clonotipos foram encontrados para a amostra 1H, 1 para a amostra 2H, 7 para a amostra 3H e 5 para a amostra 4H. A soma dos três clonotipos mais abundantes corresponde a mais de 80% das sequências encontradas em todas as amostras com essa especificidade. Na amostra 2H, em que apenas duas sequências foram identificadas, um único clonotipo equivale a 100% da resposta contra rLiD1 (Figura 22A).

Quanto as sequências específicas contra o *pool* de venenos, 18 clonotipos foram encontrados na amostra 1H, 7 na amostra 2H, 17 na amostra 3H e 14 na amostra 4H, sendo somente os 10 mais abundantes representados na Figura 22B. As amostras 3H e 4H apresentam maior distribuição da abundância entre os diferentes clonotipos. Por outro lado, as amostras 2H e 4H apresentam maior concentração da abundância em único clonotipo, com o top clone correspondendo a 58% da resposta específica detectada. Em média, os top 5 clones das amostras contra o *pool* de venenos representam 80% dos clonotipos encontrados.

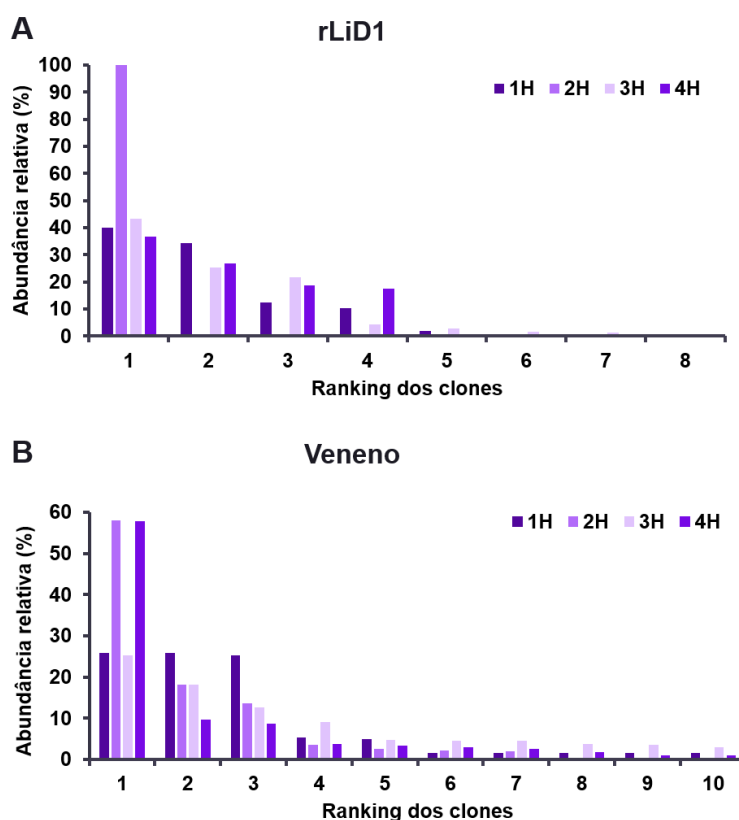


Figura 22 - Frequência dos clonotipos das sequências da região variável IGH antígeno-específicas. A abundância relativa de cada clonotipo das quatro amostras foi calculada e

expressa em percentual para os clones encontrados nas amostras contra rLiD1 (A) e os top 10 clones encontrados nas amostras contra *pool* de venenos de *Loxosceles* (B).

O ranking de abundância gerado a partir das informações de área sob a curva dos peptídeos foi comparado ao ranking gerado a partir do número de *reads* encontradas no Rep-Seq (Tabela 3). Nas amostras específicas contra rLiD1, os clones identificados no Ig-Seq não representam os clones mais abundantes do Rep-Seq. Para os clonotipos específicos para rLiD1, destacam-se apenas os clones de posição 4 e 5 da amostra 1H, 1, 2, 4 e 7 da amostra 3H e 2 da amostra 4H que figuram dentre os top 1.000 clones identificados no Rep-Seq. Nas amostras específicas contra o *pool* de venenos, o segundo e o quinto clones mais abundantes do Rep-Seq foram identificados no Ig-Seq para o cavalo 1H. Nos outros animais, os dois clones mais abundantes do Ig-Seq estão no top 1.000 clones do Rep-Seq.

Nas amostras 3H e 4H, os clones mais abundantes identificados para o *pool* de venenos foram também encontrados nas amostras contra rLiD1, indicando a possível especificidade dessas sequências. As evidências encontradas indicam que não foi observada correlação entre a abundância de transcritos de imunoglobulina totais e a abundância de imunoglobulinas secretadas antígeno-específicas.

Tabela 3 – Abundância de clonotipos no Ig-Seq e no Rep-Seq

Antígenos	rLiD1				Veneno				
	Amostras	1H	2H	3H	4H	1H	2H	3H	4H
Ranking Ig-Seq									
					Ranking Rep-Seq				
1	3.176	4.272	911	14.394	5	297	911	191	
2	3.552		891	191	2	321	561	131	
3	11.825		17.572	20.375	14.197	1.408	17.572	132	
4	684		949	9.671	3.552	907	8.329	20.375	
5	547		11.158	12.352	48	13.251	1.272	43.269	
6	7.497		2.427		3.926	4.272	12.476	3.727	
7	14.197		343		1.541	882	21.767	28.182	
8	3.271				14.846		2.759	636	
9					2.720		18.563	9.671	
10					7.641		87	5.053	

4.10. Similaridade de paratopos das sequências específicas contra rLiD1

A caracterização das sequências encontradas no Ig-Seq foi finalizada com a avaliação funcional das 31 sequências da região variável de IGH específicas contra a rLiD1. Junto às sequências descobertas foi incluído o anticorpo monoclonal mab16 que tem especificidade contra a mesma proteína e tem seu epítopo conhecido (Dias-Lopes et al, 2014). Como descrito na metodologia, essas sequências tiveram suas estruturas tridimensionais modeladas e comparadas quanto à probabilidade de interação com um mesmo epítopo da proteína rLiD1 utilizando o *software* Ab-Ligity (WONG et al., 2021). Os resultados são demonstrados na Figura 23 na forma de clusters de sequências que possuem sítios de interação similares de acordo com a predição realizada.

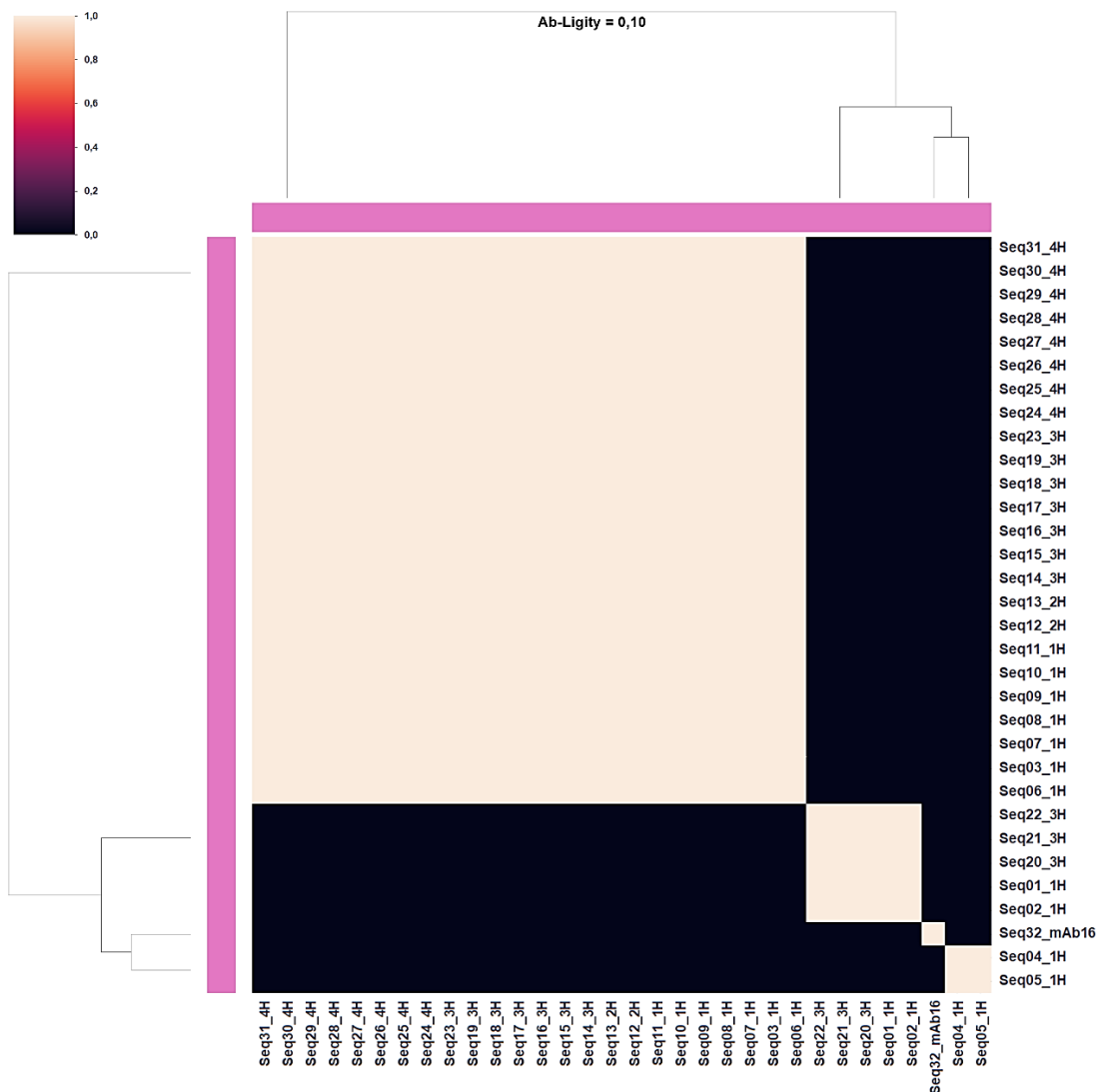


Figura 23 – Similaridade de paratopos entre as sequências da região variável de IGH específicas contra a rLiD1. Trinta e uma sequências descobertas nas amostras dos quatro

cavalos foram avaliadas junto ao anticorpo monoclonal mAb16 no *software* Ab-Ligity. A figura destaca os pares de seqüências com *score* de similaridade de epítipo superior a 0,1 em bege claro, valor de *threshold* estabelecido na publicação de referência.

Foram gerados três clusters contendo as seqüências específicas contra a rLiD1 e um cluster contendo apenas o mAb16, que, portanto, não apresenta sítio de interação similar a nenhuma das seqüências IGHV encontradas.

O maior cluster abarca 24 das 32 seqüências, unindo elementos de todas as quatro amostras. A ferramenta utilizada não informa a localização dos paratopos preditos e o alinhamento dos resíduos de aminoácidos dessas seqüências não revelou sítios similares dos CDRs que estariam atuando no reconhecimento do mesmo epítipo (Figura 24). Na avaliação do comprimento do CDRH3 e do uso de segmentos V, D e J também não são notadas evidências de similaridade que deem suporte a convergência de paratopos predita para esse grupo de seqüências. O cluster intermediário inclui as duas amostras que contêm o maior número de seqüências contra rLiD1 e o menor cluster possui apenas seqüências do cavalo 1H.

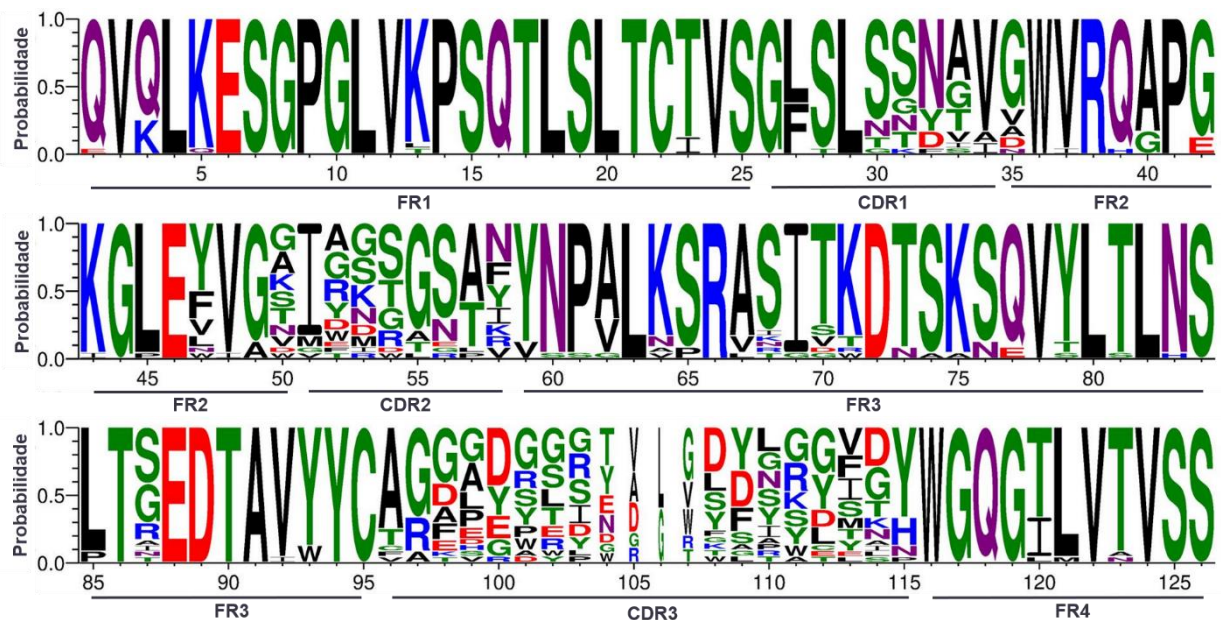


Figura 24 – Distribuição de resíduos de aminoácidos por posição das 24 seqüências de paratopos similares contidas no cluster 1. As seqüências da região variável do IGH agrupadas no cluster 1 foram alinhadas e o WebLogo construído para avaliar os resíduos mais frequentes em cada posição. Os gaps gerados por posições não ocupadas em nenhuma das seqüências foram removidos gerando alteração nas posições canônicas das regiões do anticorpo. Posições compostas por letras mais estreitas possuem menor probabilidade de

estarem ocupadas por qualquer resíduo. Diferentes cores dos aminoácidos representam suas classes. Verde: polares, roxo: neutros, azul: carregados positivamente, vermelho: carregados negativamente, preto: hidrofóbicos. Figura gerada em <http://weblogo.threeplusone.com/>.

Em resumo, o resultado indica alta convergência dos epítomos reconhecidos pela maior parte das sequências identificadas no conjunto das quatro amostras testadas, ainda que não seja possível apontar os resíduos de aminoácidos ou as características que sejam determinantes para associar tais sequências ao reconhecimento de um epítomo em comum.

5. Discussão

Os derivados de plasmas equinos hiperimunes são utilizados na forma de soroterápicos contra o envenenamento há quase 130 anos. No caso do loxoscelismo, os soros permanecem como o único tratamento específico e eficaz na neutralização dos componentes do veneno. Apesar dessa relevância, nenhum estudo havia se dedicado a explorar o repertório de sequências de imunoglobulinas específicas contra os venenos de *Loxosceles* utilizados para imunização até o momento.

Nesse trabalho, as imunoglobulinas de plasmas equinos hiperimunizados foram caracterizadas a partir de abordagens sorológicas e proteômicas. Inicialmente, testes de ELISA foram utilizados para determinar as concentrações de IgM, IgG e IgA e em sequência, para avaliar a distribuição das subclasses de IgG nos plasmas. Nas amostras testadas, foram obtidas as concentrações medianas de IgM de 0,61 mg/mL para as amostras pré-imunes e de 0,63 mg/mL para as amostras hiperimunes, valores que se encaixam nos intervalos de 0,25-0,86 mg/mL (DE CAMARGO et al., 2009) ou de 0,63-1,43 mg/mL (PERKINS et al., 2003) já relatados para cavalos adultos saudáveis. Essa observação e a semelhança das concentrações medianas encontradas nos estados pré e hiperimunes, corroboram com o dado da baixa relevância da IgM na neutralização dos componentes do veneno (FERNANDES; TAKEHARA; MOTA, 1991). Para IgG, foram obtidas as concentrações medianas de 44,04 mg/mL para as amostras pré-imunes e de 142,21 mg/mL para as hiperimunes, sendo a concentração dos animais hiperimunes 4,7 vezes maior que o limite superior do intervalo de 13,72–30,32 mg/mL já descrito por Perkins e colaboradores (2003). A variação é esperada uma vez que estão sendo comparadas dosagens de animais hiperimunizados e animais saudáveis. É importante considerar ainda que o momento de coleta das amostras utilizadas pode não representar o pico da produção das imunoglobulinas avaliadas e que os animais já haviam passado por imunizações vacinais. Além disso, Halassy e colaboradores (2019) demonstram que a quantificação adequada de IgG em soros equinos hiperimunes depende do uso de calibrador/padrão com distribuição similar de subclasses dessa imunoglobulina, fator que não foi considerado nas dosagens realizadas nesse trabalho.

Em cavalos saudáveis, a subclasse de IgG mais abundante é IgGb (IgG4 e IgG7), seguida por IgG(T) (IgG3 e IgG5), IgGa (IgG1) e então IgGc (IgG6) (SHEORAN

et al., 2000). Por outro lado, no plasma de animais hiperimunizados, IgG(T) foi encontrada em maior abundância (FERNANDES et al., 1997) e mostrou-se essencial para a neutralização dos efeitos do veneno de *Bothrops jararaca* in vivo (FERNANDES; TAKEHARA; MOTA, 1991; FERNANDES et al., 1997, 2000). Nas amostras equinas, antes e após a hiperimunização, IgG(T) foi encontrada com absorvâncias superiores às demais subclasses. O uso da cromatografia de afinidade com proteína G favoreceu o enriquecimento da subclasse IgGb (IgG4) em detrimento da IgGT (IgG3 e IgG5) que foi encontrada em maior abundância nos plasmas hiperimunes. O achado é corroborado pelos estudos de Sheoran e Holmes (1996) e Lewis, Wagner e Woof (2008), que determinam a seguinte reatividade da proteína G contra IgGs equinas: interação forte com IgG1 (IgGa), IgG4 e IgG7 (IgGb); interação moderada com IgG3 (IgG(T)); interação fraca com IgG6 (IgGc) e nenhuma interação com IgG5 (IgG(T)). A cromatografia de afinidade é apontada como o método de isolamento de IgG que provoca maior alteração na distribuição de subclasses equinas em comparação às cromatografias de troca aniônica ou catiônica e às precipitações por sulfato de amônio ou ácido caprílico, apesar de ser vantajosa na eliminação de outras proteínas do plasma (LUKAČEVIĆ et al., 2020). A perda considerável de IgG(T) após a cromatografia de afinidade com proteína G, pode ter implicado na identificação de menor número de peptídeos da região variável de IGH no presente trabalho e pode ter contribuído para a baixa descoberta de clonotipos de maior abundância no Rep-Seq. O uso das precipitações tradicionalmente utilizadas na produção de anti-venenos ou a seleção das imunoglobulinas apenas a partir de sua afinidade contra o antígeno, poderiam contribuir para a obtenção de um retrato mais fiel do repertório de anticorpos específicos do soro anti-loxoscélico.

Para a caracterização proteômica, a metodologia do Ig-Seq foi aplicada pela primeira vez ao estudo de anticorpos de cavalo. Em comparação a trabalhos anteriores, um menor número de clonotipos foi encontrado usando essa abordagem. Lavinder e colaboradores (2014) relatam a descoberta de cerca de 100 clonotipos de anticorpos contra a toxina tetanoide vacinal, Lee e colaboradores (2016) reportam entre 40 e 147 clonotipos por amostra após vacinação contra o vírus Influenza e, nesse trabalho, uma média de 14 clonotipos por amostra foram encontrados com especificidade contra o *pool* de venenos. Em primeiro lugar, a busca por peptídeos que caracterizem um único clone de imunoglobulina em meio ao universo de fragmentos de regiões mais conservadas e abundantes dos anticorpos tem como

resultado esperado a descoberta de um número restrito de sequências. Ainda assim, a baixa identificação de sequências a partir de peptídeos mapeados ao CDRH3 pode ser justificada por outros fatores passíveis de otimização no protocolo experimental. No trabalho de Xiang e colaboradores (2021), a baixa frequência de sítios trípticos no repertório transcrito de lhamas motivou o uso de quimiotripsina para obtenção de maior cobertura das sequências de imunoglobulina. De forma similar às lhamas, os cavalos apresentaram menor frequência de resíduos de arginina e lisina do que outras espécies (em nosso banco de sequências e no trabalho de Manso e colaboradores, 2019). A predição de digestão *in silico* para a escolha de outras proteases possivelmente aumentaria o número de identificações do presente trabalho. O trabalho de Wine e colaboradores (2013) demonstra a frequência de resíduos de lisina ou arginina de 91% na posição 107 do CDRH3 no banco de sequências personalizados de coelhos contra 17% nas sequências de cavalo dos bancos de cavalos hiperimunizados com veneno de *Loxosceles*. Outras hipóteses geradas para incrementar a identificação de sequências incluem: a inclusão de peptídeos assinalados a uma única imunoglobulina e mapeados a outros CDRs, o uso da espectrometria de massa *targeted* para a busca de peptídeos encontrados em apenas uma das replicatas e a utilização de uma lista de exclusão contendo peptídeos de regiões conservadas de imunoglobulinas (regiões constantes ou de *framework*) ou de contaminantes comuns, cuja intensidade pode concorrer com a de peptídeos informativos de CDR3 nas etapas de DDA.

As características de CDRH3 equinos e do uso dos genes de imunoglobulinas já foram avaliadas em estudos baseados em transcriptoma (MANSO et al., 2019). A frequência relativa de aminoácidos observada nessa região nos anticorpos antígeno-específicos assemelha-se a desses trabalhos: glicina, tirosina, ácido aspártico, alanina e serina são os cinco resíduos mais abundantes. A manutenção da abundância dos resíduos de tirosina e glicina é comprovadamente necessária para a geração, migração e maturação e de células B (IPPOLITO et al., 2006). Apesar dessa similaridade, os resíduos de ácido aspártico, arginina e lisina, aparecem com frequência superior no Ig-Seq, indicando que esses resíduos possivelmente foram inseridos em eventos de hipermutação somática e/ou são de maior relevância em anticorpos específicos contra componentes do veneno de *Loxosceles*. Alguns estudos indicam que maior frequência de arginina e aminoácidos carregados positivamente é encontrada em anticorpos inespecíficos ou autor reativos (BIRTALAN et al., 2008),

porém, resíduos com essas características ainda podem ter papel significativo na interação com epítomos não-próprios quando adicionados durante a hipermutação somática em pontos específicos da sequência.

Os genes V(D)J utilizados com maior frequência também corroboram os resultados de trabalhos anteriores nos quais as sequências com IGHV4-29, IGHD2-26 e IGHJ6-1 são as mais abundantes (MANSO et al., 2019; TALLMADGE et al., 2013). Genes não detectados no Ig-Seq como IGHV4-37, IGHV9-66, IGHV1-70 também foram encontrados com menor frequência nos outros trabalhos citados. Quanto ao comprimento do CDRH3, a mediana das sequências antígeno-específicas detectadas foi maior do que encontrada por Manso e colaboradores, que foi de 14 para um cavalo e 15 para outro e do que a encontrada por Navas e colaboradores, que foi de 16 resíduos, contra a mediana de 17 observada no Ig-Seq, resultado que se assemelha ao comprimento encontrado com maior frequência em humanos. As taxas de hipermutação somática observadas também são similares aos trabalhos citados. A característica de predominância de poucos clones na resposta antígeno-específica foi observada em condições de hiperimunização em modelo animal (WINE et al., 2013) e nas respostas a antígenos vacinais em humanos (LAVINDER et al., 2014; LEE et al., 2016) ou a infecções naturais (VOSS et al., 2021). De forma similar, nesse trabalho, os top 3 ou top 5 clones representam, em média, 80% da resposta imune detectada específica contra rLiD1 ou contra o *pool* de venenos, respectivamente. Os clones de maior abundância no Ig-Seq, porém, não correspondiam aos mais frequentes no Rep-Seq, mais uma vez, como já relatado por Cheung (2012) e Chen (2017) e colaboradores. Bonissone (2020) ressalta ainda que mesmo com aumento da profundidade do sequenciamento, alguns dos anticorpos séricos de maior reatividade contra o antígeno não são identificados no transcriptoma de células B. Estima-se que apenas 10 a 15% das IgGs nos soros terapêuticos sejam específicas contra o veneno de interesse (GUTIÉRREZ et al., 2017), informação que pode justificar parcialmente a falta de correlação entre os dois repertórios.

A similaridade funcional das sequências identificadas contra rLiD1 foi avaliada computacionalmente considerando as características dos resíduos que compõem o possível paratopo de cada anticorpo. O Ab-Ligity realiza a predição de paratopos similares a partir de modelos estruturais de cadeia leve e pesada ou apenas da cadeia pesada, baseados em referências do PDB (Protein Data Bank). As estruturas de imunoglobulinas mais abundantes nesse repositório são de humanos ou

camundongos. As diferenças entre os anticorpos dessas espécies e os equinos pode reduzir a confiabilidade das estruturas propostas para essas moléculas, o que é dificultado também pela ausência de dados da cadeia leve. Ainda assim, os resultados apontam alta convergência dos paratopos da maior parte das sequências encontradas, sugerindo uma imunodominância de um epítipo particular de rLiD1, ainda não identificado, que foi capaz de induzir 24 dos 31 anticorpos (sequências de IGHV) identificadas neste trabalho.

O conjunto de achados indica a aplicabilidade do Ig-Seq para a descoberta de repertório de imunoglobulinas equinas, assim como possíveis pontos a serem aprimorados nessa metodologia. A metodologia selecionada foi bem sucedida na identificação da cadeia pesada de anticorpos reativos contra o *pool* de venenos de *Loxosceles* ou contra a fosfolipase D recombinante de *L. intermedia*. A partir das sequências do transcriptoma da cadeia leve, será buscada a identificação dos pares de cadeias leve e pesadas que poderão ser utilizados para a constituição de soros policlonais produzidos *in vitro*. O conhecimento gerado por essa abordagem, pode contribuir para o avanço da busca por alternativas mais vantajosas aos soros anti-veneno convencionais.

6. Conclusão

i) A quantificação de isotipos de imunoglobulina revelou concentrações de IgM que variavam de 0,53 a 1,77 mg/mL nas amostras pré-ímmunes e de 0,36 a 0,90 mg/mL nas amostras hiperímmunes e de IgG de 22,52 a 160,94 mg/mL nas amostras pré-ímmunes e de 84,54 a 171,04 mg/mL nas amostras hiperímmunes.

iii) A avaliação de subclasses indicou predomínio de IgG(T) específica contra o *pool* de venenos nos plasmas hiperímmunes. IgGb foi encontrada como segunda subclasse mais abundante seguida por IgGa. Após o isolamento por cromatografia com proteína G, foi encontrado maior título de IgGb específicas, seguido por IgG(T) e IgGa.

v) Neste trabalho foi realizada, pela primeira vez, a identificação de anticorpos (sequências de IGHV) de cavalos presentes no soro anti-loxoscélico reativos contra os venenos de *L. intermedia*, *L. gaucho* e *L. laeta* e contra a principal fosfolipase D do veneno de *L. intermedia*, LiD1, em sua forma recombinante. A metodologia do Ig-Seq permitiu a identificação de 163 sequências da região variável da cadeia pesada de imunoglobulinas equinas específicas contra componentes do veneno das espécies de *Loxosceles* de importância médica no Brasil.

vii) Os CDRH3 antígeno-específicos encontrados possuem a mediana de comprimento de 17 resíduos e maior frequência de resíduos carregados em comparação ao repertório de transcriptoma. Nas sequências encontradas os segmentos gênicos IGHV4-29, IGHD2-26 e IGHJ6-1 foram os mais utilizados.

viii) Os cálculos de abundância indicaram predomínio de, em média, 5 clones nas sequências específicas contra o *pool* de venenos e 3 clones nas específicas contra rLiD1 que juntos correspondiam a mais de 80% da abundância de todos os peptídeos identificados.

ix) Foram identificados três possíveis sítios de interação para as 31 sequências específicas contra rLiD1, com 24 das sequências compartilhando paratopos similares.

7. Referências bibliográficas

1. ALMAGRO, Juan C.; MARTINEZ, Laisel; SMITH, Sylvia L.; ALAGON, Alejandro; ESTEVEZ, Judith; PANIAGUA, Jorge. Analysis of the horse VH repertoire and comparison with the human IGHV germline genes, and sheep, cattle and pig VH sequences. **Molecular Immunology**, [S. l.], v. 43, n. 11, p. 1836–1845, 2006. DOI: 10.1016/J.MOLIMM.2005.10.017. Acesso em: 25 jun. 2022.
2. ALVARENGA, Larissa M.; MARTINS, Maria S.; MOURA, Juliana F.; KALAPOTHAKIS, Evanguedes; OLIVEIRA, Júlio C.; MANGILI, Oldemir C.; GRANIER, Claude; CHÁVEZ-OLÓRTEGUI, Carlos. Production of monoclonal antibodies capable of neutralizing dermonecrotic activity of *Loxosceles intermedia* spider venom and their use in a specific immunometric assay. **Toxicon**, [S. l.], v. 42, n. 7, p. 725–731, 2003. DOI: 10.1016/j.toxicon.2003.09.006. Acesso em: 11 nov. 2019.
3. ARUN, S. S.; BREUER, W.; HERMANN, W. Immunohistochemical examination of light-chain expression (λ/κ ratio) in canine, feline, equine, bovine and porcine plasma cells. **Journal of Veterinary Medicine Series A: Physiology Pathology Clinical Medicine**, [S. l.], v. 43, n. 9, p. 573–576, 1996. DOI: 10.1111/J.1439-0442.1996.TB00489.X. Acesso em: 27 jun. 2022.
4. AVRAM, Oren; KIGEL, Aya; VAISMAN-MENTESH, Anna; KLIGSBERG, Sharon; ROSENSTEIN, Shai; DROR, Yael; PUPKO, Tal; WINE, Yariv. PASA: Proteomic analysis of serum antibodies web server. **PLoS Computational Biology**, [S. l.], v. 17, n. 1, 2021. DOI: 10.1371/JOURNAL.PCBI.1008607. Disponível em: /pmc/articles/PMC7861515/. Acesso em: 18 abr. 2022.
5. BARNIDGE, D. R. et al. Phenotyping polyclonal kappa and lambda light chain molecular mass distributions in patient serum using mass spectrometry. **J. Proteome Res.**, [S. l.], v. 13, n. 11, p. 5198–5205, 2014. DOI: 10.1021/pr5005967. Acesso em: 16 out. 2021.
6. BENICHO, Jennifer; BEN-HAMO, Rotem; LOUZOUN, Yoram; EFRONI, Sol. Rep-Seq: uncovering the immunological repertoire through next-generation sequencing. **Immunology**, [S. l.], v. 135, n. 3, p. 183–191, 2012. DOI: 10.1111/J.1365-2567.2011.03527.X. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1111/j.1365-2567.2011.03527.x>. Acesso em: 28 jun. 2022.
7. BOCHNER, Rosany. Paths to the discovery of antivenom serotherapy in France. **Journal of Venomous Animals and Toxins Including Tropical Diseases**, [S. l.], v. 22, n. 1, 2016. DOI: 10.1186/S40409-016-0074-7. Acesso em: 14 jun. 2022.

8. BONISSONE, Stefano R.; LIMA, Thiago; HARRIS, Katherine; DAVIDSON, Laura; AVANZINO, Brian; TRINKLEIN, Nathan; CASTELLANA, Natalie; PATEL, Anand. Serum proteomics expands on high-affinity antibodies in immunized rabbits than deep B-cell repertoire sequencing alone | Enhanced Reader. **BioRxiv**, [S. l.], 2020. . Acesso em: 19 jul. 2022.
9. BOYD, Scott D.; JOSHI, Shilpa A. High-Throughput DNA Sequencing Analysis of Antibody Repertoires. **Microbiology Spectrum**, [S. l.], v. 2, n. 5, 2014. DOI: 10.1128/MICROBIOLSPEC.AID-0017-2014. Disponível em: <https://journals.asm.org/doi/abs/10.1128/microbiolspec.AID-0017-2014>. Acesso em: 25 jun. 2022.
10. BRASIL, Ministério da Saúde. **Manual de Diagnóstico e Tratamento de Acidentes por Animais Peçonhentos**. 2. ed. Brasília: Fundação Nacional de Saúde, 2001. . Acesso em: 7 jul. 2022.
11. CARVALHO, Milene B.; MOLINA, Franck; FELICORI, Liza F. Yvis: Antibody high-density alignment visualization and analysis platform with an integrated database. **Nucleic Acids Research**, [S. l.], 2019. DOI: 10.1093/nar/gkz387.
12. CHAIM, Olga M. et al. Phospholipase-D activity and inflammatory response induced by brown spider dermonecrotic toxin: Endothelial cell membrane phospholipids as targets for toxicity. **Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids**, [S. l.], v. 1811, n. 2, p. 84–96, 2011. DOI: 10.1016/J.BBALIP.2010.11.005. Acesso em: 29 jun. 2022.
13. CHEN, Jing et al. Proteomic Analysis of Pemphigus Autoantibodies Indicates a Larger, More Diverse, and More Dynamic Repertoire than Determined by B Cell Genetics. **Cell Reports**, [S. l.], v. 18, n. 1, p. 237–247, 2017. DOI: 10.1016/J.CELREP.2016.12.013. Acesso em: 1 jun. 2022.
14. CHEUNG, Wan Cheung et al. A proteomics approach for the identification and cloning of monoclonal antibodies from serum. **Nature Biotechnology** 2012 30:5, [S. l.], v. 30, n. 5, p. 447–452, 2012. DOI: 10.1038/nbt.2167. Disponível em: <https://www.nature.com/articles/nbt.2167>. Acesso em: 31 maio. 2022.
15. CHOTHIA, Cyrus; LESK, Arthur M. Canonical structures for the hypervariable regions of immunoglobulins. **Journal of molecular biology**, [S. l.], v. 196, n. 4, p. 901–917, 1987. DOI: 10.1016/0022-2836(87)90412-8. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/3681981/>. Acesso em: 22 jan. 2024.
16. COX, Jürgen; MANN, Matthias. MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. **Nature Biotechnology** 2008 26:12, [S. l.], v. 26, n. 12, p. 1367–1372,

2008. DOI: 10.1038/nbt.1511. Disponível em: <https://www.nature.com/articles/nbt.1511>. Acesso em: 29 jul. 2022.

17. DE ALMEIDA, Daniel Manzoni; FERNANDES-PEDROSA, Matheus De F.; GONÇALVES DE ANDRADE, Rute M.; MARCELINO, José Roberto; GONDO-HIGASHI, Hisako; JUNQUEIRA DE AZEVEDO, Inácio De L. M.; HO, Paulo Lee; VAN DEN BERG, Carmen; TAMBOURGI, Denise V. A New Anti-loxoscelic Serum Produced Against Recombinant Sphingomyelinase D: Results of Preclinical Trials. **The American Journal of Tropical Medicine and Hygiene**, [S. l.], v. 79, n. 3, p. 463–470, 2008. DOI: 10.4269/AJTMH.2008.79.463. Disponível em: <https://www.ajtmh.org/view/journals/tpmd/79/3/article-p463.xml>. Acesso em: 11 jul. 2022.

18. DE OLIVEIRA, Kátia C.; GONÇALVES DE ANDRADE, Rute M.; PIAZZA, Roxane M. F.; FERREIRA, Jorge M. C.; VAN DEN BERG, C. W.; TAMBOURGI, Denise V. Variations in Loxosceles spider venom composition and toxicity contribute to the severity of envenomation. **Toxicon**, [S. l.], v. 45, n. 4, p. 421–429, 2005. DOI: 10.1016/J.TOXICON.2004.08.022. Acesso em: 29 dez. 2021.

19. DIAS-LOPES, C. et al. Generation and molecular characterization of a monoclonal antibody reactive with conserved epitope in sphingomyelinases D from Loxosceles spider venoms. **Vaccine**, [S. l.], v. 32, p. 2086–2092, 2014. DOI: 10.1016/j.vaccine.2014.02.012. Disponível em: <http://dx.doi.org/10.1016/j.vaccine.2014.02.012>. Acesso em: 14 jun. 2022.

20. DORFER, Viktoria; PICHLER, Peter; STRANZL, Thomas; STADLMANN, Johannes; TAUS, Thomas; WINKLER, Stephan; MECHTLER, Karl. MS Amanda, a universal identification algorithm optimized for high accuracy tandem mass spectra. **Journal of Proteome Research**, [S. l.], v. 13, n. 8, p. 3679–3684, 2014. DOI: 10.1021/PR500202E. Acesso em: 29 jul. 2022.

21. DUARTE, C. G.; BONILLA, C.; GUIMARÃES, G.; MACHADO DE AVILA, R. A.; MENDES, T. M.; SILVA, W.; TINTAYA, B.; YARLEQUE, A.; CHÁVEZ-OLÓRTEGUI, C. Anti-loxoscelic horse serum produced against a recombinant dermonecrotic protein of Brazilian Loxosceles intermedia spider neutralize lethal effects of Loxosceles laeta venom from Peru. **Toxicon**, [S. l.], v. 93, p. 37–40, 2015. DOI: 10.1016/j.toxicon.2014.10.023.

22. ENG, Jimmy K.; MCCORMACK, Ashley L.; YATES, John R. An approach to correlate tandem mass spectral data of peptides with amino acid sequences in a protein database. **Journal of the American Society for Mass Spectrometry**, [S. l.], v. 5, n. 11, p. 976–989, 1994. DOI: 10.1016/1044-0305(94)80016-2. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/24226387/>. Acesso em: 9 ago. 2022.

23. FELICORI, L.; ARAUJO, S. C.; MACHADO DE ÁVILA, R. A.; SANCHEZ, E. F.; GRANIER, C.; KALAPOTHAKIS, E.; CHÁVEZ-OLÓRTEGUI, C. Functional characterization and epitope analysis of a recombinant dermonecrotic protein from *Loxosceles intermedia* spider. **Toxicon**, [S. l.], v. 48, n. 5, p. 509–519, 2006. DOI: 10.1016/J.TOXICON.2006.06.019. Acesso em: 12 maio. 2022.
24. FERNANDES, I.; TAKEHARA, H. A.; MOTA, I. Isolation of IgGT from hyperimmune horse anti-snake venom serum: Its protective ability. **Toxicon**, [S. l.], v. 29, n. 11, p. 1373–1379, 1991. DOI: 10.1016/0041-0101(91)90124-A. Acesso em: 6 jul. 2022.
25. FERNANDES, I.; TAKEHARA, H. A.; SANTOS, A. C. R.; CORMONT, F.; LATINNE, D.; BAZIN, H.; MOTA, I. Neutralization of bothropic and crotalic venom toxic activities by IgG(T) and IgGa subclasses isolated from immune horse serum. **Toxicon**, [S. l.], v. 35, n. 6, p. 931–936, 1997. DOI: 10.1016/S0041-0101(96)00177-8. Acesso em: 6 jul. 2022.
26. FERNANDES-PEDROSA, Matheus de F.; JUNQUEIRA-DE-AZEVEDO, Inácio de L. M.; GONÇALVES-DE-ANDRADE, Rute M.; KOBASHI, Leonardo S.; ALMEIDA, Diego D.; HO, Paulo L.; TAMBOURGI, Denise V. Transcriptome analysis of *Loxosceles laeta* (Araneae, Sicariidae) spider venomous gland using expressed sequence tags. **BMC Genomics**, [S. l.], v. 9, p. 279, 2008. DOI: 10.1186/1471-2164-9-279. Disponível em: /pmc/articles/PMC2440767/. Acesso em: 7 jul. 2022.
27. FIGUEIREDO, Luís F. M.; DIAS-LOPES, Camila; ALVARENGA, Larissa M.; MENDES, Thais M.; MACHADO-DE-ÁVILA, Ricardo A.; MCCORMACK, Jessica; MINOZZO, João C.; KALAPOTHAKIS, Evanguedes; CHÁVEZ-OLÓRTEGUI, Carlos. Innovative immunization protocols using chimeric recombinant protein for the production of polyspecific loxoscelic antivenom in horses. **Toxicon**, [S. l.], v. 86, p. 59–67, 2014. DOI: 10.1016/J.TOXICON.2014.05.007. Disponível em: <https://reader.elsevier.com/reader/sd/pii/S0041010114001330?token=17449ED4EA8B6AA7AAB2E79A29055A1894FC5557EAB1AA46867AB3AB432ED04DD41A943B8B42699C34AF86E1D73C8FBD&originRegion=us-east-1&originCreation=20220810125544>. Acesso em: 8 ago. 2022.
28. FINGERMANN, Matías; DE ROODT, Adolfo Rafael; CASCONI, Osvaldo; MIRANDA, María Victoria. Biotechnological potential of Phospholipase D for *Loxosceles* antivenom development. **Toxicon: X**, [S. l.], v. 6, p. 100036, 2020. DOI: 10.1016/J.TOXCX.2020.100036. Acesso em: 2 jan. 2022.
29. FUTRELL, J. M. Loxoscelism. **American Journal of the Medical Sciences**, [S. l.], v. 304, n. 4, p. 261–267, 1992. DOI: 10.1097/00000441-199210000-00008. Acesso em: 14 jun. 2022.

30. GEORGIU, George; IPPOLITO, Gregory C.; BEAUSANG, John; BUSSE, Christian E.; WARDEMANN, Hedda; QUAKE, Stephen R. The promise and challenge of high-throughput sequencing of the antibody repertoire. **Nature Biotechnology**, [S. l.], v. 32, n. 2, p. 158–168, 2014. DOI: 10.1038/nbt.2782. Disponível em: <http://dx.doi.org/10.1038/nbt.2782>.
31. GERVÁSIO, João; FERREIRA, Alice; FELICORI, Liza F. Yclon: Ultrafast clustering of B cell clones from high-throughput immunoglobulin repertoire sequencing data. **bioRxiv**, [S. l.], p. 2022.02.17.480909, 2022. DOI: 10.1101/2022.02.17.480909. Disponível em: <https://www.biorxiv.org/content/10.1101/2022.02.17.480909v1>. Acesso em: 19 jun. 2022.
32. GREMSKI, Luiza Helena et al. A novel expression profile of the *Loxosceles intermedia* spider venomous gland revealed by transcriptome analysis. **Molecular BioSystems**, [S. l.], v. 6, n. 12, p. 2403–2416, 2010. DOI: 10.1039/C004118A. Acesso em: 24 jun. 2022.
33. GREMSKI, Luiza Helena et al. Recent advances in the understanding of brown spider venoms: From the biology of spiders to the molecular mechanisms of toxins. **Toxicon**, [S. l.], v. 83, p. 91–120, 2014. DOI: 10.1016/J.TOXICON.2014.02.023. Acesso em: 27 jun. 2022.
34. GUILHERME, Patrícia; FERNANDES, Irene; BARBARO, Katia Cristina. Neutralization of dermonecrotic and lethal activities and differences among 32–35 kDa toxins of medically important *Loxosceles* spider venoms in Brazil revealed by monoclonal antibodies. **Toxicon**, [S. l.], v. 39, n. 9, p. 1333–1342, 2001. DOI: 10.1016/S0041-0101(01)00085-X. Acesso em: 11 nov. 2019.
35. GUTIÉRREZ, José María; FAN, Hui Wen; SILVERA, Carmen L. M.; ANGULO, Yamileth. Stability, distribution and use of antivenoms for snakebite envenomation in Latin America: Report of a workshop. **Toxicon**, [S. l.], v. 53, p. 625–630, 2009. DOI: 10.1016/j.toxicon.2009.01.020. Acesso em: 14 jun. 2022.
36. HALASSY, Beata; KURTOVIĆ, Tihana; LANG BALIJA, Maja; BRGLES, Marija; TUNJIĆ, Monika; SVIBEN, Dora. Concept of sample-specific correction of immunoassay results for precise and accurate IgG quantification in horse plasma. **Journal of Pharmaceutical and Biomedical Analysis**, [S. l.], v. 164, p. 276–282, 2019. DOI: 10.1016/J.JPBA.2018.10.020. Disponível em: <https://doi.org/10.1016/j.jpba.2018.10.020>. Acesso em: 27 mar. 2022.
37. HERSHBERG, Uri; LUNING PRAK, Eline T. The analysis of clonal expansions in normal and autoimmune B cell repertoires. **Philosophical Transactions of the Royal Society B: Biological Sciences**, [S. l.], v. 370, n. 1676, 2015. DOI: 10.1098/RSTB.2014.0239. Acesso em: 10 ago. 2022.

38. INBAR, D.; HOCHMAN, J.; GIVOL, D. Localization of antibody-combining sites within the variable portions of heavy and light chains. **Proceedings of the National Academy of Sciences of the United States of America**, [S. l.], v. 69, n. 9, p. 2659–2662, 1972. DOI: 10.1073/PNAS.69.9.2659. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/4560694/>. Acesso em: 22 jan. 2024.
39. IONOV, Steven; LEE, Jiwon. An Immunoproteomic Survey of the Antibody Landscape: Insights and Opportunities Revealed by Serological Repertoire Profiling. **Frontiers in Immunology**, [S. l.], v. 13, 2022. DOI: 10.3389/FIMMU.2022.832533. Disponível em: </pmc/articles/PMC8843944/>. Acesso em: 2 mar. 2022.
40. ISBISTER, Geoffrey K.; FAN, Hui Wen. Spider bite. **The Lancet**, [S. l.], v. 378, n. 9808, p. 2039–2047, 2011. DOI: 10.1016/S0140-6736(10)62230-1. Acesso em: 23 jun. 2022.
41. IVERSEN, Rasmus; SNIR, Omri; STENSLAND, Maria; KROLL, José E.; STEINSBØ, Øyvind; KORPONAY-SZABÓ, Ilma R.; LUNDIN, Knut E. A.; DE SOUZA, Gustavo A.; SOLLID, Ludvig M. Strong Clonal Relatedness between Serum and Gut IgA despite Different Plasma Cell Origins. **Cell Reports**, [S. l.], v. 20, n. 10, p. 2357–2367, 2017. DOI: 10.1016/J.CELREP.2017.08.036. Acesso em: 1 jun. 2022.
42. JIANG, Ning et al. Lineage Structure of the Human Antibody Repertoire in Response to Influenza Vaccination. **Science translational medicine**, [S. l.], v. 5, n. 171, p. 171ra19, 2013. DOI: 10.1126/SCITRANSLMED.3004794. Disponível em: </pmc/articles/PMC3699344/>. Acesso em: 9 ago. 2022.
43. KABAT, E. A.; WU, T. T.; BILOFSKY, H. Some correlations between specificity and sequence of the first complementary determining segments of human kappa light chains. **Proceedings of the National Academy of Sciences of the United States of America**, [S. l.], v. 73, n. 12, p. 4471–4473, 1976. DOI: 10.1073/pnas.73.12.4471. Acesso em: 22 jan. 2024.
44. KARIM-SILVA, Sabrina et al. Loxoscelism: Advances and Challenges in the Design of Antibody Fragments with Therapeutic Potential. **Toxins 2020, Vol. 12, Page 256**, [S. l.], v. 12, n. 4, p. 256, 2020. DOI: 10.3390/TOXINS12040256. Disponível em: <https://www.mdpi.com/2072-6651/12/4/256/htm>. Acesso em: 30 jun. 2022.
45. KEGGAN, Alison; FREER, Heather; ROLLINS, Alicia; WAGNER, Bettina. Production of seven monoclonal equine immunoglobulins isotyped by multiplex analysis. **Veterinary Immunology and Immunopathology**, [S. l.], v. 153, n. 3–4, p. 187–193, 2013. DOI: 10.1016/J.VETIMM.2013.02.010. Acesso em: 28 jun. 2022.
46. KOERS, Jana; DERKSEN, Ninotska I. L.; OOIJEVAAR-DE HEER, Pleuni; NOTA, Benjamin; VAN DE BOVENKAMP, Fleur S.; VIDARSSON, Gestur; RISPENS, Theo. Biased N -Glycosylation Site Distribution and Acquisition across the Antibody V

Region during B Cell Maturation . **The Journal of Immunology**, [S. l.], 2019. DOI: 10.4049/jimmunol.1801622.

47. LAVINDER, Jason J. et al. Identification and characterization of the constituent human serum antibodies elicited by vaccination. **Proceedings of the National Academy of Sciences of the United States of America**, [S. l.], v. 111, n. 6, p. 2259–2264, 2014. DOI: 10.1073/pnas.1317793111.

48. LEE, Jiwon et al. Molecular-level analysis of the serum antibody repertoire in young adults before and after seasonal influenza vaccination. **Nature Medicine** **22:12**, [S. l.], v. 22, n. 12, p. 1456–1464, 2016. DOI: 10.1038/nm.4224. Disponível em: <https://www.nature.com/articles/nm.4224>. Acesso em: 8 maio. 2022.

49. LEEM, Jinwoo; DUNBAR, James; GEORGES, Guy; SHI, Jiye; DEANE, Charlotte M. ABodyBuilder: Automated antibody structure prediction with data-driven accuracy estimation. **mAbs**, [S. l.], v. 8, n. 7, p. 1259–1268, 2016. DOI: 10.1080/19420862.2016.1205773/SUPPL_FILE/KMAB_A_1205773_SM1023.PDF. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/19420862.2016.1205773>. Acesso em: 3 maio. 2022.

50. LEFRANC, M. P. IMGT, the international ImMunoGeneTics database. **Nucleic Acids Research**, [S. l.], v. 29, n. 1, p. 207–209, 2001. DOI: 10.1093/NAR/29.1.207. Disponível em: <https://academic.oup.com/nar/article/29/1/207/1116025>. Acesso em: 9 ago. 2022.

51. LEFRANC, Marie Paule; POMMIÉ, Christelle; RUIZ, Manuel; GIUDICELLI, Véronique; FOULQUIER, Elodie; TRUONG, Lisa; THOUVENIN-CONTET, Valérie; LEFRANC, Gérard. IMGT unique numbering for immunoglobulin and T cell receptor variable domains and Ig superfamily V-like domains. **Developmental and Comparative Immunology**, [S. l.], v. 27, n. 1, p. 55–77, 2003. DOI: 10.1016/S0145-305X(02)00039-3. Acesso em: 29 jun. 2022.

52. LEFRANC, Marie-Paule. IMGT Locus on Focus A New Section of Experimental and Clinical Immunogenetics. **Exp Clin Immunogenet**, [S. l.], v. 15, p. 1–7, 1998. Disponível em: <http://imgt.cnusc.fr:8104>. Acesso em: 11 jul. 2022.

53. LEWIS, Melanie J.; WAGNER, Bettina; WOOF, Jenny M. The different effector function capabilities of the seven equine IgG subclasses have implications for vaccine strategies. **Molecular Immunology**, [S. l.], v. 45, n. 3, p. 818–827, 2008. DOI: 10.1016/j.molimm.2007.06.158.

54. LIMA, Sabrina de Almeida et al. Recombinant protein containing B-cell epitopes of different *Loxosceles* spider toxins generates neutralizing antibodies in immunized rabbits. **Frontiers in Immunology**, [S. l.], v. 9, n. APR, 2018. DOI: 10.3389/FIMMU.2018.00653/FULL. Acesso em: 26 jun. 2022.

55. LITMAN, Gary W.; ANDERSON, Michele K.; RAST, Jonathan P. Evolution of antigen binding receptors. **Annual Review of Immunology**, [S. l.], v. 17, p. 109–147, 1999. DOI: 10.1146/ANNUREV.IMMUNOL.17.1.109. Acesso em: 6 jul. 2022.
56. LOGAN, Aaron C. et al. High-throughput VDJ sequencing for quantification of minimal residual disease in chronic lymphocytic leukemia and immune reconstitution assessment. **Proceedings of the National Academy of Sciences of the United States of America**, [S. l.], v. 108, n. 52, p. 21194–21199, 2011. DOI: 10.1073/PNAS.1118357109/-/DCSUPPLEMENTAL. Disponível em: /pmc/articles/PMC3248502/. Acesso em: 9 ago. 2022.
57. MA, Bin; ZHANG, Kaizhong; HENDRIE, Christopher; LIANG, Chengzhi; LI, Ming; DOHERTY-KIRBY, Amanda; LAJOIE, Gilles. PEAKS: powerful software for peptide de novo sequencing by tandem mass spectrometry. **Rapid communications in mass spectrometry: RCM**, [S. l.], v. 17, n. 20, p. 2337–2342, 2003. DOI: 10.1002/RCM.1196. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/14558135/>. Acesso em: 9 ago. 2022.
58. MÁLAQUE, Ceila Mariana Sant’Ana; CASTRO-VALENCIA, Jaime Enrique; CARDOSO, João Luiz Costa; FRANÇA, Francisco Oscar de Siqueira; BARBARO, Kátia Cristina; FAN, Hui Wen. **CLINICAL AND EPIDEMIOLOGICAL FEATURES OF DEFINITIVE AND PRESUMED LOXOSCELISM IN SÃO PAULO, BRAZIL** *Rev. Inst. Med. trop. S. Paulo*. [s.l.: s.n.]. Disponível em: <https://www.scielo.br/j/rimtsp/a/TQRQWjhDPwcQ3JKkTh4Wfbw/?format=pdf&lang=en>. Acesso em: 6 jul. 2022.
59. MCGUIRE, T. C.; CRAWFORD, T. B.; HENSON, J. B. The Isolation, Characterisation and Functional Properties of Equine Immunoglobulin Classes and Subclasses1. **Equine Infectious Diseases**, [S. l.], p. 364–381, 1973. DOI: 10.1159/000393549. Acesso em: 25 jun. 2022.
60. MELANI, Rafael D. et al. Next-Generation Serology by Mass Spectrometry: Readout of the SARS-CoV-2 Antibody Repertoire. **Journal of Proteome Research**, [S. l.], v. 21, n. 1, p. 274–288, 2022. DOI: 10.1021/ACS.JPROTEOME.1C00882/ASSET/IMAGES/LARGE/PR1C00882_0004.JPEG. Disponível em: <https://pubs.acs.org/doi/full/10.1021/acs.jproteome.1c00882>. Acesso em: 28 jun. 2022.
61. MENDES, T. M. et al. Generation and characterization of a recombinant chimeric protein (rCpLi) consisting of B-cell epitopes of a dermonecrotic protein from *Loxosceles intermedia* spider venom. **Vaccine**, [S. l.], v. 31, n. 25, p. 2749–2755, 2013. DOI: 10.1016/J.VACCINE.2013.03.048. Acesso em: 26 jun. 2022.

62. MIRANDA, Ana Luísa Soares De; ANTUNES, Bruno Cesar; MINOZZO, João Carlos; LIMA, Sabrina de Almeida; BOTELHO, Ana Flávia Machado; CAMPOS, Marco Túlio Gomes; CHÁVEZ-OLÓRTEGUI, Carlos Delfin; SOTO-BLANCO, Benito. Clinical Effects of the Immunization Protocol Using *Loxosceles* Venom in Naïve Horses. **Toxins**, [S. l.], v. 14, n. 5, p. 338, 2022. DOI: 10.3390/TOXINS14050338. Disponível em: /pmc/articles/PMC9148105/. Acesso em: 2 jun. 2022.
63. MIRANDA, Ana Luísa Soares De; GUERRA-DUARTE, Clara; LIMA, Sabrina de Almeida; CHÁVEZ-OLÓRTEGUI, Carlos; SOTO-BLANCO, Benito. History, challenges and perspectives on *Loxosceles* (brown spiders) antivenom production in Brazil. **Toxicon**, [S. l.], v. 192, p. 40–45, 2021. DOI: 10.1016/J.TOXICON.2021.01.004. Acesso em: 15 fev. 2022.
64. MONTGOMERY, P. C. Molecular Aspects of Equine Antibodies. **Equine Infectious Diseases**, [S. l.], p. 343–363, 1973. DOI: 10.1159/000393548. Acesso em: 25 jun. 2022.
65. NAVAS, Carlana et al. THE MAJOR ROLE OF JUNCTIONAL DIVERSITY IN THE HORSE ANTIBODY REPERTOIRE. **bioRxiv**, [S. l.], p. 2022.06.20.496904, 2022. DOI: 10.1101/2022.06.20.496904. Disponível em: <https://www.biorxiv.org/content/10.1101/2022.06.20.496904v1>. Acesso em: 28 jul. 2022.
66. PARAMESWARAN, Poornima et al. Convergent antibody signatures in human dengue. **Cell host & microbe**, [S. l.], v. 13, n. 6, p. 691, 2013. DOI: 10.1016/J.CHOM.2013.05.008. Disponível em: /pmc/articles/PMC4136508/. Acesso em: 9 ago. 2022.
67. PAULI, Isolete; MINOZZO, João Carlos; HENRIQUE DA SILVA, Paulo; CHAIM, Olga Meiri; VEIGA, Silvio Sanches. Analysis of therapeutic benefits of antivenin at different time intervals after experimental envenomation in rabbits by venom of the brown spider (*Loxosceles intermedia*). **Toxicon**, [S. l.], v. 53, n. 6, p. 660–671, 2009. DOI: 10.1016/J.TOXICON.2009.01.033. Acesso em: 5 ago. 2022.
68. PAULI, Isolete; PUKA, Juliana; GUBERT, Ida Cristina; MINOZZO, João Carlos. The efficacy of antivenom in loxoscelism treatment. **Toxicon**, [S. l.], v. 48, n. 2, p. 123–137, 2006. DOI: 10.1016/J.TOXICON.2006.05.005. Acesso em: 30 jul. 2022.
69. PERKINS, David N.; PAPPIN, Darryl J. C.; CREASY, David M.; COTTRELL, John S. Probability-based protein identification by searching sequence databases using mass spectrometry data. **Electrophoresis**, [S. l.], v. 20, p. 3551–3567, 1999. DOI: [https://doi.org/10.1002/\(SICI\)1522-2683\(19991201\)20:18<3551::AID-ELPS3551>3.0.CO;2-2](https://doi.org/10.1002/(SICI)1522-2683(19991201)20:18<3551::AID-ELPS3551>3.0.CO;2-2). Disponível em: [https://sci-hub.se/10.1002/\(sici\)1522-2683\(19991201\)20:18%3C3551::aid-elps3551%3E3.0.co;2-2](https://sci-hub.se/10.1002/(sici)1522-2683(19991201)20:18%3C3551::aid-elps3551%3E3.0.co;2-2). Acesso em: 29 jul. 2022.

70. SATO, Shuji et al. Proteomics-directed cloning of circulating antiviral human monoclonal antibodies. **Nature Biotechnology** 2012 30:11, [S. l.], v. 30, n. 11, p. 1039–1043, 2012. DOI: 10.1038/nbt.2406. Disponível em: <https://www.nature.com/articles/nbt.2406>. Acesso em: 1 jun. 2022.
71. SEZERINO, Ulisses Marion et al. A clinical and epidemiological study of Loxosceles spider envenoming in Santa Catarina, Brazil. **Transactions of the Royal Society of Tropical Medicine and Hygiene**, [S. l.], v. 92, n. 5, p. 546–548, 1998. DOI: 10.1016/S0035-9203(98)90909-9. Acesso em: 3 jul. 2022.
72. SILVA JÚNIOR, César; SASSON, Sezar; CALDINI JÚNIOR, Nelson. **BIOLOGIA_CSC_V3_PNLD2018_Capa_PR.pdf | Enhanced Reader**. [s.l.: s.n.]. v. 11. Acesso em: 28 jun. 2022.
73. SNAPKOV, Igor; CHERNIGOVSKAYA, Maria; SINITCYN, Pavel; QUÝ, Khang Lê; NYMAN, Tuula A.; GREIFF, Victor. Progress and challenges in mass spectrometry-based analysis of antibody repertoires. **Trends in Biotechnology**, [S. l.], v. 0, n. 0, 2021. DOI: 10.1016/J.TIBTECH.2021.08.006. Disponível em: <http://www.cell.com/article/S0167779921001931/fulltext>. Acesso em: 26 set. 2021.
74. SOUZA, Natália Alves; DIAS-LOPES, Camila; MATOSO, Ítalo Hugo Gonçalves; DE OLIVEIRA, Camila Franco Batista; CHÁVEZ-OLORTEGUI, Carlos Delfin; MINOZZO, João Carlos; FELICORI, Liza F. Immunoprotection elicited in rabbit by a chimeric protein containing B-cell epitopes of Sphingomyelinases D from Loxosceles spp. spiders. **Vaccine**, [S. l.], v. 36, n. 48, p. 7324–7330, 2018. DOI: 10.1016/j.vaccine.2018.10.035. Acesso em: 16 jun. 2021.
75. SQUAIELLA-BAPTISTÃO, Carla Cristina; SANT'ANNA, Osvaldo Augusto; MARCELINO, José Roberto; TAMBOURGI, Denise V. The history of antivenoms development: Beyond Calmette and Vital Brazil. **Toxicon**, [S. l.], v. 150, p. 86–95, 2018. DOI: 10.1016/J.TOXICON.2018.05.008. Acesso em: 26 jun. 2022.
76. SWANSON, David L.; VETTER, Richard S. Loxoscelism. **Clinics in Dermatology**, [S. l.], v. 24, n. 3, p. 213–221, 2006. DOI: 10.1016/J.CLINDERMATOL.2005.11.006. Acesso em: 14 jun. 2022.
77. TALLMADGE, Rebecca L.; TSENG, Chia T.; KING, Rebecca A.; FELIPPE, M. Julia B. Developmental progression of equine immunoglobulin heavy chain variable region diversity. **Developmental and comparative immunology**, [S. l.], v. 41, n. 1, p. 33, 2013. DOI: 10.1016/J.DCI.2013.03.020. Disponível em: [/pmc/articles/PMC3672396/](https://pubmed.ncbi.nlm.nih.gov/23672396/). Acesso em: 6 jul. 2022.
78. TE WU, Tai; KABAT, Elvin A. An analysis of the sequences of the variable regions of Bence Jones proteins and myeloma light chains and their implications for antibody complementarity. **The Journal of experimental medicine**, [S. l.], v. 132, n. 2,

- p. 211–250, 1970. DOI: 10.1084/JEM.132.2.211. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/5508247/>. Acesso em: 22 jan. 2024.
79. THEAKSTON, R. D. G.; WARRELL, D. A.; GRIFFITHS, E. Report of a WHO workshop on the standardization and control of antivenoms. **Toxicon**, [S. l.], v. 41, n. 5, p. 541–557, 2003. DOI: 10.1016/S0041-0101(02)00393-8. Acesso em: 27 jun. 2022.
80. TOMESCU-BACIU, Alina; JOHANSEN, Jorunn N.; HOLMØY, Trygve; GREIFF, Victor; STENSLAND, Maria; DE SOUZA, Gustavo Antonio; VARTDAL, Frode; LOSSIUS, Andreas. Persistence of intrathecal oligoclonal B cells and IgG in multiple sclerosis. **Journal of Neuroimmunology**, [S. l.], v. 333, p. 576966, 2019. DOI: 10.1016/J.JNEUROIM.2019.576966. Acesso em: 1 jun. 2022.
81. TRAN, Ngoc Hieu; RAHMAN, M. Ziaur; HE, Lin; XIN, Lei; SHAN, Baozhen; LI, Ming. Complete De Novo Assembly of Monoclonal Antibody Sequences. **Scientific Reports** 2016 6:1, [S. l.], v. 6, n. 1, p. 1–10, 2016. DOI: 10.1038/srep31730. Disponível em: <https://www.nature.com/articles/srep31730>. Acesso em: 19 jun. 2022.
82. VANDER HEIDEN, Jason A.; YAARI, Gur; UDUMAN, Mohamed; STERN, Joel N. H.; O'CONNOR, Kevin C.; HAFLE, David A.; VIGNEAULT, Francois; KLEINSTEIN, Steven H. pRESTO: a toolkit for processing high-throughput sequencing raw reads of lymphocyte receptor repertoires. **Bioinformatics**, [S. l.], v. 30, n. 13, p. 1930, 2014. DOI: 10.1093/BIOINFORMATICS/BTU138. Disponível em: [/pmc/articles/PMC4071206/](https://pubmed.ncbi.nlm.nih.gov/24911111/). Acesso em: 18 abr. 2022.
83. VOSS, William N. et al. Prevalent, protective, and convergent IgG recognition of SARS-CoV-2 non-RBD spike epitopes. **Science**, [S. l.], v. 372, p. 1108–1112, 2021. DOI: 10.1126/science.abg5268. Disponível em: <https://www.sciencemag.org/lookup/doi/10.1126/science.abg5268>. Acesso em: 3 maio. 2021.
84. WAGNER, Bettina. Immunoglobulins and immunoglobulin genes of the horse. **Developmental & Comparative Immunology**, [S. l.], v. 30, n. 1–2, p. 155–164, 2006. DOI: 10.1016/J.DCI.2005.06.008. Acesso em: 25 jun. 2022.
85. WAGNER, Bettina; MILLER, Donald C.; LEAR, Teri L.; ANTCZAK, Douglas F. The Complete Map of the Ig Heavy Chain Constant Gene Region Reveals Evidence for Seven IgG Isotypes and for IgD in the Horse. **J Immunol**, [S. l.], v. 173, p. 3230–3242, 2004. DOI: 10.4049/jimmunol.173.5.3230. Disponível em: <http://www.jimmunol.org/content/173/5/3230><http://www.jimmunol.org/content/173/5/3230>0.full#ref-list-1. Acesso em: 28 jun. 2022.
86. WALTHER, Stefanie; RUSITZKA, Tamara V.; DIESTERBECK, Ulrike S.; CZERNY, Claus Peter. Equine immunoglobulins and organization of immunoglobulin

genes. **Developmental & Comparative Immunology**, [S. l.], v. 53, n. 2, p. 303–319, 2015. DOI: 10.1016/J.DCI.2015.07.017. Acesso em: 28 jun. 2022.

87. WANG, Zhe; LIU, Xiaowen; MUTHER, Jennifer; JAMES, Judith A.; SMITH, Kenneth; WU, Si. Top-down Mass Spectrometry Analysis of Human Serum Autoantibody Antigen-Binding Fragments. **Scientific Reports**, [S. l.], v. 9, n. 1, 2019. DOI: 10.1038/S41598-018-38380-Y. Acesso em: 9 jun. 2022.

88. WILLIAMS, La Tonya D. et al. Potent and broad HIV-neutralizing antibodies in memory B cells and plasma. **Science Immunology**, [S. l.], v. 2, n. 7, 2017. DOI: 10.1126/SCIIMMUNOL.AAL2200. Acesso em: 8 ago. 2022.

89. WINE, Yariv et al. Molecular deconvolution of the monoclonal antibodies that comprise the polyclonal serum response. **Proceedings of the National Academy of Sciences of the United States of America**, [S. l.], v. 110, n. 8, p. 2993–2998, 2013. DOI: 10.1073/PNAS.1213737110/SUPPL_FILE/SAPP.PDF. Acesso em: 8 maio. 2022.

90. WONG, Wing Ki; ROBINSON, Sarah A.; BUJOTZEK, Alexander; GEORGES, Guy; LEWIS, Alan P.; SHI, Jiye; SNOWDEN, James; TADDESE, Bruck; DEANE, Charlotte M. Ab-Ligity: identifying sequence-dissimilar antibodies that bind to the same epitope. **mAbs**, [S. l.], v. 13, n. 1, 2021. DOI: 10.1080/19420862.2021.1873478. Disponível em: /pmc/articles/PMC7833755/. Acesso em: 29 mar. 2022.

91. XIANG, Yufei; SANG, Zhe; BITTON, Lirane; XU, Jianquan; LIU, Yang; SCHNEIDMAN-DUHOVNY, Dina; SHI, Yi. Integrative proteomics identifies thousands of distinct, multi-epitope, and high-affinity nanobodies. **Cell Systems**, [S. l.], v. 12, n. 3, p. 220- 234.e9, 2021. DOI: 10.1016/J.CELS.2021.01.003. Acesso em: 1 jun. 2022.

92. XU, John L.; DAVIS, Mark M. Diversity in the CDR3 Region of VH Is Sufficient for Most Antibody Specificities. **Immunity**, [S. l.], v. 13, n. 1, p. 37–45, 2000. DOI: 10.1016/S1074-7613(00)00006-6. Disponível em: <http://www.cell.com/article/S1074761300000066/fulltext>. Acesso em: 26 jun. 2022.