



UNIVERSIDADE FEDERAL RURAL DO SEMI-ÁRIDO
PRÓ-REITORIA DE GRADUAÇÃO
DEPARTAMENTO DE BIOCÊNCIAS
CURSO DE GRADUAÇÃO EM BIOTECNOLOGIA

Afonso Hudson Martins Cordeiro Neto

**Predição e caracterização de epítomos imunogênicos para o desenvolvimento de
vacina contra a Linfadenite Caseosa**

MOSSORÓ
2024

Afonso Hudson Martins Cordeiro Neto

**Predição e caracterização de epítomos imunogênicos para o desenvolvimento de
vacina contra a Linfadenite Caseosa**

Monografia apresentada a Universidade
Federal Rural do Semi-Árido como requisito
para obtenção do título de Bacharel em
Biotecnologia.

Orientador: Isadora Louise Alvez da Costa
Ribeiro Quintans, Prof. Dr.

Co-orientador: Silvestre Brilhante Bezerra,
Prof. Dr.

MOSSORÓ
2024

©Todos os direitos estão reservados à Universidade Federal Rural do Semi-Árido. O conteúdo desta obra é de inteira responsabilidade do (a) autor (a), sendo o mesmo, passível de sanções administrativas ou penais, caso sejam infringidas as leis que regulamentam a Propriedade Intelectual, respectivamente, Patentes: Lei nº 9.279/1996, e Direitos Autorais: Lei nº 9.610/1998. O conteúdo desta obra tornar-se-á de domínio público após a data de defesa e homologação da sua respectiva ata, exceto as pesquisas que estejam vinculadas ao processo de patenteamento. Esta investigação será base literária para novas pesquisas, desde que a obra e seu (a) respectivo (a) autor (a) seja devidamente citado e mencionado os seus créditos bibliográficos.

Ficha catalográfica elaborada pelo Sistema de Bibliotecas
da Universidade Federal Rural do Semi-Árido, com os dados fornecidos pelo(a) autor(a)

NN469 Neto, Afonso Hudson Martins Cordeiro.
p Predição e caracterização de epítomos
 imunogênicos para o desenvolvimento de vacina
 contra a Linfadenite Caseosa / Afonso Hudson
 Martins Cordeiro Neto. - 2024.
 48 f. : il.

 Orientadora: Isadora Louise Alvez da Costa
 Ribeiro Quintans.
 Coorientador: Silvestre Brilhante Bezerra.
 Monografia (graduação) - Universidade Federal
 Rural do Semi-árido, Curso de Biotecnologia,
 2024.

 1. Corynebacterium pseudotuberculosis. 2.
 linfócitos. 3. bioinformática. 4. proteínas. 5.
 docking molecular. I. Quintans, Isadora Louise
 Alvez da Costa Ribeiro, orient. II. Bezerra,
 Silvestre Brilhante, co-orient. III. Título.

Setor de Informação e Referência

Bibliotecária KEINA CRISTINA SANTOS SOUSA E SILVA

O serviço de Geração Automática de Ficha Catalográfica para Trabalhos de Conclusão de Curso (TCC's) foi desenvolvido pelo Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo (USP) e gentilmente cedido para o Sistema de Bibliotecas da Universidade Federal Rural do Semi-Árido (SISBI-UFERSA), sendo customizado pela Superintendência de Tecnologia da Informação e Comunicação (SUTIC) sob orientação dos bibliotecários da instituição para ser adaptado às necessidades dos alunos dos Cursos de Graduação e Programas de Pós-Graduação da Universidade.

Afonso Hudson Martins Cordeiro Neto

**Predição e caracterização de epítomos imunogênicos para o desenvolvimento de
vacina contra a Linfadenite Caseosa**

Monografia apresentada a Universidade
Federal Rural do Semi-Árido como requisito
para obtenção do título de Bacharel em
Biotecnologia.

Defendida em: _19_ / _04_ / _2024_.

BANCA EXAMINADORA

Documento assinado digitalmente



ISADORA LOUISE ALVES DA COSTA RIBEIRO QU

Data: 25/04/2024 09:14:15-0300

Verifique em <https://validar.iti.gov.br>

Isadora Louise Alves da Costa Ribeiro Quintans, Prof. Dr. (UFERSA)

Presidente

Documento assinado digitalmente



FREDERICO SCHMITT KREMER

Data: 26/04/2024 14:05:09-0300

Verifique em <https://validar.iti.gov.br>

Fredérico Schmitt Kremer, Prof. Dr. (UFPEL)

Membro Examinador

Documento assinado digitalmente



TALLYSON NOGUEIRA BARBOSA

Data: 25/04/2024 20:04:36-0300

Verifique em <https://validar.iti.gov.br>

Tallyson Nogueira Barbosa, Prof. Dr. (UFPEL)

Membro Examinador

AGRADECIMENTOS

Agradeço os meus pais, que me deram todo o apoio necessário para conclusão deste trabalho

Agradeço a Orientadora e a Universidade Federal Rural do Semiárido por me fornecer o suporte financeiro e as orientações no qual tornou possível a conclusão dessa monografia

Agradeço a Banca Examinadora por ser compreensiva, competente e atenciosa nas correções e orientações para trabalhos futuros

Agradeço aos meus amigos e amigas Gabriel, Dantas, Leandro, Victor, Kleiton, Mateus, Alexandre, Marcus, Lucas, Lara, Lenise, Mariana, Deisy, Karina, Vanessa e Raissa, por se divertirem comigo, me ouvirem e me corrigirem quando necessário, foi fundamental para meu desenvolvimento pessoal e profissional.

Agradeço a minha namorada, Lara Gabriela, por me mostrar que com foco, muito amor e perseverança consigo fazer qualquer coisa.

RESUMO

A Linfadenite Caseosa (LC) é uma doença que afeta caprinos e ovinos. Ela se caracteriza pela formação de abscessos nos linfonodos superficiais e internos, bem como em órgãos como pulmões, baço, rins, fígado e sistema nervoso central. O organismo causador da LC é a bactéria *Corynebacterium pseudotuberculosis*, um bacilo gram-positivo, intracelular facultativo, imóvel, não esporulado e com capacidade fermentativa. Linfadenite Caseosa (LC) é uma doença que afeta caprinos e ovinos. Ela se caracteriza pela formação de abscessos nos linfonodos superficiais e internos, bem como em órgãos como pulmões, baço, rins, fígado e sistema nervoso central. A prospecção de novos epítomos com potencial antigênico continua sendo crucial, visto que uma vacina eficaz ainda não foi desenvolvida. A bioinformática oferece oportunidades para a busca de respostas por meio de algoritmos computacionais, sem altos investimentos financeiros, ampliando as ferramentas disponíveis para a comunidade científica. Assim o objetivo do trabalho é prever *in silico* epítomos imunogênicos para linfócitos B, TCD4⁺ e TCD8⁺ por meio de bioinformática. Para isso foi realizada a prospecção de proteínas da membrana plasmática, da parede celular e proteínas extracelulares a partir do proteoma da

C. pseudotuberculosis recuperado de bancos de dados biológicos, com posterior análise nos softwares SignalP, modelagem no ColabFold e predição nos softwares ElliPro, BepiPred e NetMHCpan 4.1. Os epítomos preditos foram selecionados com base no grau de estabilidade com ProtParam, antigenicidade com o VaxiJen, alergenicidade com o AllerTop e toxicidade com o ToxinPred. Utilizando o programa ColabFold os epítomos selecionados foram modelados e tiveram sua afinidade de se ligar com as moléculas do sistema imune avaliadas por meio do Docking Molecular realizado no software PyRx. Aplicando as ferramentas de bioinformática apresentadas no estudo foi possível prever um total de 19 epítomos antigênicos que interagem com os receptores do sistema imune, além de estáveis, atóxicos e hipoalergênicos, sendo que, dentre eles, 5 eram linfócitos B, 8 eram de linfócitos T-MHC I e 6 eram de linfócitos T-MHC II. Esses epítomos podem formular uma vacina multi-epitopo para a LC.

Palavras-chave: *Corynebacterium pseudotuberculosis*; linfócitos; bioinformática; proteínas; docking molecular.

ABSTRACT

Caseous Lymphadenitis (CL) is a disease that affects goats and sheep. It is characterized by the formation of abscesses in the superficial and internal lymph nodes, as well as in organs such as the lungs, spleen, kidneys, liver and central nervous system. The causative organism of CL is the bacterium *Corynebacterium pseudotuberculosis*, a gram-positive, facultative intracellular, immobile, non-sporulated bacillus with fermentative capacity. Caseous lymphadenitis (CL) is a disease that affects goats and sheep. It is characterized by the formation of abscesses in the superficial and internal lymph nodes, as well as in organs such as the lungs, spleen, kidneys, liver and central nervous system. The search for new epitopes with antigenic potential remains crucial, since an effective vaccine has not yet been developed. Bioinformatics offers opportunities to search for answers using computer algorithms, without high financial investment, expanding the tools available to the scientific community. The aim of this study is to predict *in silico* immunogenic epitopes for B, TCD4 and TCD8 lymphocytes using bioinformatics. To this end, proteins from the plasma membrane, cell wall and extracellular proteins were prospected from the proteome of *C. pseudotuberculosis* retrieved from biological databases, with subsequent analysis in SignalP software, modeling in ColabFold and prediction in ElliPro, Bepipret and NetMHCpan 4.1 software. The predicted epitopes were selected based on their degree of stability with ProtParam, antigenicity with VaxiJe, allergenicity with AllerTop and toxicity with ToxinPred. Using the ColabFold program, the selected epitopes were modeled. Using the bioinformatics tools presented in the study, it was possible to predict a total of 19 antigenic epitopes that interact with immune system receptors, as well as being stable, non-toxic and hypoallergenic. Of these, 5 were from B lymphocytes, 8 were from T- MHC I lymphocytes and 6 were from T-MHC II lymphocytes. These epitopes could formulate a multi-epitope vaccine for CL.

Keywords: *Corynebacterium pseudotuberculosis*; lymphocytes; bioinformatics; proteins; molecular docking;.

LISTA DE FIGURAS

Figura 1 - Fluxograma da metodologia que será aplicada na predição dos epítomos.....	28
Figura 2 - Diagrama de Venn com informações das análises do TOPCONS. T omíniosnrios transmembrana, SP = peptídeo sinal.....	32
Figura 3 – Docking Molecular entre epítomos preditos (Coloridos) e molécula MHC classe I (marrom). (A) Epítopo D9Q9T5. (B) Epítopo D9QA26. (C) Epítopo D9QAE3. (D) Epítopo D9QAU0.....	38
Figura 4 – Docking Molecular entre epítomos preditos (Coloridos) e molécula MHC classe I (marrom). (A) Epítopo D9QBB8. (B) Epítopo D9QC81. (C) Epítopo D9QCC6. (D) Epítopo D9QER2.....	39
Figura 5 - Docking Molecular entre epítomos preditos (Coloridos) e molécula MHC classe II (marrom). (A) Epítopo D9Q9Y1. (B) Epítopo D9Q982. (C) Epítopo D9QA85. (D) Epítopo D9QB77.....	40
Figura 6 - Docking Molecular entre epítomos preditos (Coloridos) e molécula MHC classe II (marrom). (A) Epítopo D9QD82. (B) Epítopo D9QD10.....	40

LISTA DE TABELAS

Tabela 1 - Número de proteínas com localização celular de interesse selecionadas com base nas informações do Gene Ontology.....	33
Tabela 2 - Epítopos de linfócitos B selecionados. Células com pontos indicam dados abaixo da média. Células em branco indicam dados não recuperados.....	34
Tabela 3 - Epítopos de linfócitos T MHCI selecionados. Células com pontos indicam dados abaixo da média. Células em branco indicam dados não recuperados.....	35
Tabela 4 - Epítopos de linfócitos T MHCII selecionados. Células com pontos indicam dados abaixo da média. Células em branco indicam dados não recuperados.....	36
Tabela 5 - Dados do docking molecular de epítopos contra MHCI realizado no PyRx.....	38
Tabela 6 - Dados do docking molecular de epítopos contra MHCII realizado no PyRx.....	39

LISTA DE ABREVIATURAS E SIGLAS

ATP	Adenosina Trifosfato
<i>C. pseudotuberculosis</i>	<i>Corynebacterium pseudotuberculosis</i>
LC	Linfadenite Caseosa
MHC	Complexo principal de histocompatibilidade
NCBI	National Center for Biotechnology Information
PDB	Protein Data Bank
SP	Peptídeo Sinal
TM	Transmenbrana

Sumário

1	INTRODUÇÃO.....	16
2	FUNDAMENTAÇÃO TEÓRICA.....	18
2.1	A Caprinovincultura e o problema da Linfadenite Caseosa.....	18
2.2	Linfadenite Caseosa.....	19
2.3	Agente etiológico.....	20
2.4	Resposta imunológica a doença.....	21
2.5	Tratamentos disponíveis para LC.....	22
2.6	Ferramentas de Bioinformática.....	23
3	HIPÓTESES.....	25
4	JUSTIFICATIVA.....	26
5	OBJETIVOS DA PESQUISA.....	27
5.1	OBJETIVO GERAL.....	27
5.2	OBJETIVOS ESPECÍFICOS.....	27
6	MATERIAL E MÉTODOS.....	28
6.1	Recuperação e seleção das sequências.....	28
6.2	Modelagem das Sequências.....	28
6.3	Predição dos epítomos.....	29
6.3.1	Seleção dos epítomos de linfócitos B.....	29
6.3.2	Seleção dos epítomos de linfócitos T-MHC I.....	29
6.3.3	Seleção dos epítomos de linfócitos T-MHC II.....	30
6.4	Triagem dos epítomos preditos.....	30
6.4.1	ProtParam.....	30
6.4.2	VaxiJen 2.0.....	31
6.4.3	AllerTop.....	31
6.4.4	ToxinPred.....	31

6.4.5	Validação <i>in silico</i>	31
7	RESULTADOS.....	32
7.1	Recuperação e seleção das sequencias.....	32
7.2	Modelagem das Sequencias.....	33
7.3	Predição dos Epítomos.....	34
7.3.1	Linfócitos B.....	34
7.3.2	Linfócitos T MHC-I.....	35
7.3.3	Linfócitos T MHC-II.....	35
7.4	Triagem dos epítomos.....	36
7.4.1	ProtParam.....	36
7.4.2	VaxiJen.....	37
7.4.3	AllerTop e ToxinPred.....	37
7.5	Docking Molecular.....	37
7.5.1	Linfócitos T MCHI.....	37
7.5.2	Linfócitos T MCHII.....	39
8	DISCUSSÃO.....	41
9	CONSIDERAÇÕES FINAIS.....	44
10	REFERÊNCIAS.....	45
	ANEXO A – TABELA COM OS DADOS FÍSICO-QUÍMICOS DO EPÍTOPOS COLETADOS DOS SOFTWARES DE TRIAGEM.....	52

1 INTRODUÇÃO

A Linfadenite Caseosa (LC) é uma doença que afeta caprinos e ovinos e é caracterizada por formar abscessos nos linfonodos superficiais e internos, além de órgãos, tais como pulmões, baço, rins, fígado e sistema nervoso central (SOUZA et al., 2021). A LC tem alcance global, mas muitas vezes é subnotificada, pois há países que não relatam a presença de casos (GUIMARÃES et al., 2011). Isto afeta as estatísticas e consequentemente as medidas para a prevenção e controle da doença (GUIMARÃES et al., 2011).

O organismo causador da LC é a bactéria *Corynebacterium pseudotuberculosis*, um bacilo gram-positivo, intracelular facultativo, imóvel, não esporulado e com capacidade fermentativa (BRUM et al., 2017; ZHOU et al., 2019). Os principais mecanismos de virulência desta bactéria são: um lipídio leucotóxico de superfície, que a protege da fagocitose, e a exotoxina fosfolipase D, que atua nas células endoteliais, causando aumento da permeabilidade dos vasos sanguíneos e linfáticos, facilitando desta maneira a invasão bacteriana (SOUZA, 2021).

A maioria das vacinas atuais para a LC se baseia em um único antígeno, ou em alguns poucos antígenos (MICHEL-TODÓ et al., 2019). A utilização de múltiplos antígenos alvo é essencial para o desenvolvimento de vacinas mais eficazes contra parasitas complexos (MICHEL-TODÓ et al., 2019). Portanto, a prospecção de novos epítomos com potencial antigênico contra a LC continua sendo crucial, visto que uma vacina eficaz ainda não foi desenvolvida.

Com o surgimento da bioinformática, novas oportunidades se abriram para prospecção de novos alvos terapêuticos, verificação da compatibilidade entre moléculas de interesse, e criação de mapas e vacinas gênicas (DE MACEDO JUNIOR, 2019). O uso da bioinformática amplia as ferramentas e recursos disponíveis ao meio científico, permitindo uma busca por respostas de questões biológicas através de algoritmos computacionais (DE MACEDO JUNIOR, 2019). Os programas de bioinformática fazem o alinhamento de sequências de nucleotídeos ou aminoácidos, e geram inferências biológicas, tais como: determinação de estrutura de proteínas e função de domínios proteicos, comparações entre genomas e filogenia, baseadas no seu grau de similaridade entre as sequências comparadas (DE MACEDO JUNIOR, 2019).

Além disso, as análises *in silico* não demandam altos investimentos financeiros. Neste contexto, vários estudos têm sido empregados para prospecção de epítomos com potencial imunogênico de muitos microorganismos utilizando ferramentas de bioinformática (MICHEL-

TODÓ et al., 2019). Por exemplo, pesquisas realizadas com *Trypanosoma cruzi*, causador da doença de chagas, estabeleceram uma coleção de 30 epítomos preditos *in silico* para células T e B com potencial imunogênico (MICHEL-TODÓ et al., 2019). Estes resultados poderão ser utilizados para auxiliar no design de uma vacina para doença de chagas. Ainda, MEI et al. (2022) realizaram a predição de epítomos de células T visando o design *in silico* de uma vacina para o COVID-19, utilizando o banco de dados do NCBI (National Center for Biotechnological Information). Este estudo resultou na identificação de 4 epítomos únicos de células T não alergênicos (MEI et al., 2022).

Com relação à LC, MARQUES et al. (2019) realizaram a análise do proteoma de *C. pseudotuberculosis*, com posterior predição de *in silico* de epítomos lineares de células B com potencial antigênico. Neste estudo, foram identificados cerca de 76 epítomos com potencial imunogênico contra a LC, que foram disponibilizados para comunidade científica, e poderão ser testados em ensaios *in vitro* (MARQUES et al., 2019). Ainda, o genoma completamente sequenciado da bactéria *C. pseudotuberculosis* strain MEX 29 (COSTA et al., 2016) está disponível atualmente no NCBI, bem como um total de 2.129 genes (NCBI, Acesso em 21 de agosto de 2022). Também está disponível o proteoma completo da bactéria *C. pseudotuberculosis* no Uniprot (UNIPROT, Acesso em 21 de agosto de 2022) e 1.555 estruturas tridimensionais no PDB (Protein Data Bank) (PROTEIN DATA BANK, Acesso em 21 de Agosto de 2022). Sendo assim, este trabalho visa prever *in silico* epítomos imunogênicos *C. pseudotuberculosis* para o desenvolvimento de uma vacina contra a Linfadenite Caseosa. Para isto serão utilizados os dados já disponíveis em bancos de dados como NCBI, Uniprot e PDB, combinado com o auxílio de softwares para caracterização, predição e modelagem de moléculas.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 A Caprinovinocultura e o problema da Linfadenite Caseosa

A caprinocultura e a ovinocultura são atividades econômicas praticadas no mundo todo, principalmente em regiões semiáridas, já que caprinos e ovinos são animais mais resilientes às diferentes condições climáticas, quando comparados a bovinos, por exemplo, que requerem mais alimento e temperaturas mais amenas para produzirem bem (SÁ et al., 2018; NAVARINI et al., 2009). Segundo MARTINS et al. (2016), o mundo possuía um rebanho estimado de 1,06 bilhão de animais em 2014, onde a maior parte está nos países em desenvolvimento. Se tratando de ovinos, a população mundial era da ordem de 1,2 bilhão, no mesmo período, e que também se concentram em países em desenvolvimento.

No Brasil, o rebanho de 13.789.345 milhões de ovinos e 8.260.607 milhões de caprinos, o que nos dá o posto do país com o 22º rebanho mundial de caprinos e o 18º de ovinos, que se concentra principalmente na região nordeste, onde essa atividade é fundamental para pequenos produtores, pois suprem a carência de proteínas e necessitam de baixo investimento inicial, tanto para aquisição dos animais, quanto para infraestrutura (SILVA et al., 2021a; MARTINS et al., 2016). A carne, leite e couro são os principais produtos deste setor comercial (SÁ et al., 2018). Dados da produção mundial de carne caprina e ovina apontam para um total de 5,4 e 8,6 milhões de toneladas em 2013, respectivamente (MARTINS et al., 2016).

Com relação às perdas econômicas, a LC tem um grande impacto, principalmente quando há uma alta incidência regional, devido a condenação da carcaça, desvalorização do couro e da lã, gastos no tratamento. Uma pesquisa realizada na Austrália apontou uma perda que varia de \$12 a 15 milhões nos abatedouros de caprinos e ovinos. Com relação a produção de lã, o prejuízo pode chegar a \$17 milhões por ano para indústrias australianas (ALVES et al., 2007). Outra pesquisa realizada por BARNABÉ (2018), que objetivou caracterizar e dimensionar o prejuízo financeiro associado à LC por condenação na linha de abate no Abatedouro Municipal da cidade de Patos (PB), revelou uma perda econômica total de R\$ 56.525,57, o que representa uma queda 6,09% da receita. Utilizando esses dados para medir as perdas do estado da Paraíba, que possui um rebanho estimado 545.994, chegamos a uma perda de R\$2.332.081,57. Para a região Nordeste tem-se um prejuízo de 32.720.894,74, já a nível nacional, o estrago pode chegar a R\$ 35.257.364,56 (BARNABÉ, 2018).

Os prejuízos da LC a caprino e ovinocultura estão ligados a alguns aspectos da sua patogenicidade. Primeiramente, os animais acometidos com a forma superficial da doença ficam debilitados para executar atividades básicas como pastejar, mastigar, regurgitar e

amamentar as crias (ALVES, 1988). Além disso, a LC afeta a produção de lã e desvaloriza a pele, podendo reduzir seu valor em até 40% com um único nódulo (Figueiredo et al., 1982). Já na sua forma visceral, a LC é assintomática e pode ocasionar problemas mais sérios, como debilitação do animal e disseminação dos abscessos, causando problemas reprodutivos, condenação da carcaça e morte prematura (ALVES, 1988). Quando o rebanho apresenta elevada incidência da LC, os animais apresentam caquexia progressiva, anemia, hiperplasia dos linfonodos superficiais e dispneia (SÁ et al., 2018). Além disso, em ovelhas pode ocorrer a migração para linfonodos mamários, causando queda na produção de leite (SÁ et al., 2018).

Além dos impactos socioeconômicos e a saúde animal, o controle da LC também é importante para prevenção de spillover (migração entre hospedeiros) para humanos, pois Faccioli-Martins et al. (2014) apontou essa doença como uma zoonose emergente relacionada ao consumo de alimento contaminado e contato com animais infectados.

2.2 Linfadenite Caseosa

A Linfadenite Caseosa (LC), é uma das principais doenças que acometem ovinos e caprinos, está presente em todos os países que possuem uma população significativa desses animais, tais como: China, Índia, e Austrália, (ALVES et al., 2007). A LC é uma doença infectocontagiosa de caráter crônico, que tem como principal característica a formação de processos inflamatórios dos linfonodos, com o surgimento de nódulos cutâneos e viscerais, de aspecto caseoso e esbranquiçado (SÁ et al., 2018). Os linfonodos afetados com mais frequência pela LC são os pré-protídeos e pré-escapulares, podendo acometer também linfonodos internos como os mediastínicos, broquiáticos e pulmonares, levando a problemas respiratórios como tosse crônica (ALVES et al., 2007; RIBEIRO, 1997).

A transmissão da LC ocorre principalmente por meio de ferimentos na pele, o que indica que a vegetação predominante na região nordeste (Caatinga), composta principalmente por plantas baixas e com espinhos, pode ajudar da dispersão da doença, devido a ferimentos superficiais causados nos animais quando se deslocam por essa vegetação. (UNANIAN et al. 1985). A disseminação do agente etiológico no ambiente ocorre através da ruptura dos granulomas, com sua possível liberação em aerossóis gerados por nódulos pulmonares (COSTA 2002; RIET-CORREA et al., 2011). A alta capacidade de sobrevivência do agente etiológico da LC no ambiente, e por longos períodos, contribui para a incidência e disseminação da doença (ALAVEZ 1988). Outros possíveis métodos de disseminação da doença são brigas entre os

animais, marcação, castração, forrageio em local contaminado, vacinação e tatuagens (ALVES et al., 2007). A transmissão para humanos dificilmente acontece e está relacionada com pessoas que trabalham diretamente com os animais ou em frigoríficos (ALVES et al., 2007). Os principais sintomas relatados são de fadiga, dor muscular, febre, fígado aumentado e macio e linfadenopatia localizada (ALVES et al., 2007). As portas de entrada do agente em humanos são geralmente o consumo de leite contaminado, consumo de carne crua contaminada, contato com fezes do animal infectado e lesões na pele (ALVES et al., 2007). Na maioria dos casos descritos o tratamento é realizado com antibióticos e remoção cirúrgica dos linfonodos afetados (ALVES et al., 2007).

2.3 Agente etiológico

A *C. pseudotuberculosis*, bactéria causadora da LC, foi descrita primeiramente por Nocard em 1885, a partir de um linfonodo infectado de um bovino, e posteriormente por Pierez em 1891, por meio de um isolamento bacteriano de abscesso renal de um ovino (TORRES, 2011). Essa bactéria pertence ao gênero *Corynebacterium*, o qual abrange bactérias patogênicas e não patogênicas, e ao grupo dos Actinomycetes, no qual estão incluídos os gêneros *Mycobacterium*, *Nocardia* e *Rhodococcus* (DORELLA et al., 2006). Várias espécies de grande importância médica, veterinária e biotecnologia pertencem a este grupo, que compartilha algumas características, tais como: alto conteúdo de G-C no DNA e uma parede celular composta principalmente por peptidoglicano, arabinogalactano e ácido micólico (DORELLA et al., 2006).

Há dois biotipos de *C. pseudotuberculosis*: a biovar ovis e a biovar equis, que se distinguem pela presença ou ausência de atividade na enzima nitrato redutase (que reduz nitrato a nitrito), seus hospedeiros e, consequentemente, ao perfil da doença (ALMEIDA et al., 2017). A biovar ovis, quanto a atividade da enzima, é considerada nitrato redutase negativa, e quanto ao hospedeiro, afeta caprinos, ovinos e bovinos, podendo evoluir para a forma visceral ou externa, afetando rins, pulmões, fígado e outros órgãos (BAIRD; FONTAINE, 2007; SONGER et al. 1988). Já a biovar equis, quanto a atividade da enzima, é considerada nitrato redutase positiva, e quanto ao hospedeiro, afeta alguns equinos, bovinos e camelos, causando linfangite ulcerativa, mastite e aborto em equinos e lesões viscerais e mastite em bovinos (YERUHAM et al., 2004; BARAÚNA et al., 2017; MOUSSA et al., 2016).

As bactérias dependem de fatores de virulência para conseguir infectar um hospedeiro, que são genes e moléculas responsáveis por permitir a invasão, colonização, sobrevivência e

multiplicação no organismo hospedeiro (TAUCH; BURKOVSKI, 2015). Alguns fatores de virulência da *C. pseudotuberculosis* já descritos são lipídios tóxicos da parede celular (ácido corinemicólico), e as exotoxinas fosfolipase D e diftérica (IONEDA, 1993; SONGER et al., 1990; SELIM et al., 2015). O lipídio tóxico de membrana, especificamente o ácido corinemicólico, não só é citotóxico, como facilita bastante a sobrevivência da *C. pseudotuberculosis* à fagocitose pelos leucócitos (JOLLY 1966; SONGER 1990). A fosfolipase D é caracterizada por catalisar a reação de hidrólise da esfingomiélin e da lisofosfatidilcolina, pode agir no endotélio vascular do hospedeiro aumentando a sua permeabilidade, facilitando a circulação da bactéria pelo corpo do hospedeiro, além disso já foi citada com um fator importante para a morte dos macrófagos (MCNAMARA et al., 1994; MCKEAN et al., 2007).

Já a toxina diftérica foi encontrada em isolados de búfalos em 2015 por SELIM et al., e sua principal atividade é inibir a síntese proteica na célula hospedeira e, conseqüentemente, levar ao processo de apoptose (MURPHY, 2011). Outro fator de virulência importante são os pili, definidos por estruturas proteicas de adesão bacteriana que têm um importante papel na infecção e multiplicação no hospedeiro (MANDLIK et al., 2008). Dois operons de pili da *C. pseudotuberculosis* já foram descritos, são eles o spaA (spaA-spaB-spaC) e spaD (spaD-spaE-spaF), onde SpaA e SpaD são pilis de eixo, SpaB e SpaE são pilis de base e SpaC e spaF são pilis de ponta (TROST et al., 2010).

Apesar de já terem sido iniciadas pesquisas de predição de epítomos *C. pseudotuberculosis* utilizando bioinformática, ainda faltam dados sobre a predição de epítomos de *C. pseudotuberculosis* para os linfócitos TCD4 e TCD8, bem como com epítomos conformacionais de linfócitos B. Estes tipos celulares são responsáveis por mediar a resposta celular e humoral contra bactérias (ABBAS et al., 2015).

2.4 Resposta imunológica a doença

O processo infeccioso da LC se inicia no local de entrada do agente etiológico, no qual há recrutamento de linfócitos e drenagem dos linfonodos. Após isso há o desenvolvimento de granuloma e posteriormente sua maturação e persistência no organismo (PÉPIN et al., 1997). Após o agente infeccioso (*C. pseudotuberculosis*) ser fagocitado por células de defesa do hospedeiro, ocorre a formação do fagolisossomo: uma fusão do fagossomo com o lisossomo (BATEY, 1986). Porém a *C. pseudotuberculosis* é resistente ao baixo pH e às enzimas do fagolisossomo, sobrevivendo no fagócito por até 48 horas (STEFANÍSKA et al., 2010). Além disso, esta bactéria é capaz de se multiplicar dentro do fagócito e posteriormente se libertar, causando sua morte (BATEY, 1986; ALVES et al., 2007).

Como consequência da multiplicação e persistência da bactéria, o hospedeiro tenta isolar a infecção por meio dos granulomas, que são caracterizados pelo encapsulamento das células infectadas em um processo complexo que envolve, tanto a resposta humoral, quanto celular (PAULE et al., 2003). Estudos indicam que as células do sistema imune presentes nos granulomas são, em sua maioria, macrófagos expressando o MHC II e linfócitos T CD4 e T CD8 em uma proporção de 3,5:1; já os linfócitos B e os granulócitos estão presentes em menor quantidade (ELLIS, 1988). Um estudo imunohistoquímico de granulomas experimentalmente induzidos em ovelhas revelou que, em lesões imaturas, a presença de linfócitos T CD4 foi mais elevada, já lesões maduras apresentam linfócitos T CD8 e linfócitos T $\gamma\delta$ em sua maioria; ainda foi revelado uma variabilidade grande nas proporções de células T e macrófagos em granulomas no mesmo estágio de maturação (PÉPIN, 1994). Esses dados sugerem uma composição celular variável nos granulomas dependendo da persistência e eliminação do agente (PÉPIN, 1994).

2.5 Tratamentos disponíveis para LC

O tratamento disponível para LC é pouco eficiente e se baseia na drenagem do abscesso formado pela infecção bacteriana, na cauterização química, e no tratamento com antibiótico (MARQUES, 2019). Este tratamento possui custo elevado, e é muitas vezes ineficaz, pois as bactérias ficam alojadas dentro de granulomas, que são de difícil acesso ao antibiótico (MARQUES, 2019). Além do tratamento, a prevenção pode ser realizada através de vacinas. Porém, as vacinas atualmente disponíveis para o tratamento da LC oferecem apenas proteção parcial, que por muitas vezes é insuficiente (PINHO et al., 2021)

A eliminação da LC é um desafio multifatorial. Estudos indicam que a *C. pseudotuberculosis* é sensível aos antibióticos em laboratório. Porém, *in vivo* essa bactéria apresenta baixa reação aos medicamentos (GUIMARÃES et al., 2009) o que dificulta o tratamento, pois o remédio tem dificuldade em penetrar nos granulomas e alcançar o foco da infecção (ALMEIDA, 2019). Além disso, o microrganismo possui imunidade e pode permanecer no meio ambiente por até 8 meses, em diferentes climas (LATIF et al., 2017). Detectar criaturas contaminadas por meio de manifestações clínicas ou exames de sangue também é um impasse (GUIMARÃES et al., 2009).

Algumas vacinas de primeira geração foram desenvolvidas ao longo dos anos. Como vacinas de toxóides multiuso, formuladas a partir da fosfolipase D inativada, associada a antígenos de patógenos do gênero *Clostridium* (PINHO et al., 2021). Estas vacinas têm níveis

de segurança duvidosos, com efeitos colaterais como lesão ou abscesso no local da injeção, febre, letargia e redução da produção de leite (PINHO et al., 2021).

Ainda, a deficiência nos Programas Nacionais de Sanidade de Ovinos e Caprinos é evidente para o controle da doença em nível nacional (MARTINS et al., 2021). Isso resulta em dificuldades de cumprir o tratamento e, conseqüentemente, a erradicação da doença no país. O desenho de testes diagnósticos e vacinas eficazes na prevenção e controle da LC devem estar fundamentados na predição de epítomos com potencial antigênico para a doença.

2.6 Ferramentas de Bioinformática

Com o crescimento das ciências ômicas vem o desenvolvimento da vacinologia reversa, método que visa identificar alvos terapêuticos a partir do genoma ou proteoma do patógeno. Essa abordagem inicialmente dispensa cultura *in vitro* e experimentos laboratoriais, exigindo apenas os dados de sequenciamentos proteicos e/ou genéticos do microrganismo. São especialmente úteis para o desenvolvimento de vacinas proteicas ou de subunidades, pois permitem a identificação *in silico* de antígenos imunogênicos e alvos de drogas. Assim, além de reduzirem drasticamente o tempo e gasto econômico, garantem maior estabilidade e segurança durante o desenvolvimento de vacinas (NAZ et al., 2020).

Outrossim, essas as vacinas de subunidades contêm componentes imunogênicos específicos contra o patógeno, o que potencializa o desenvolvimento de uma imunização mais ampla e eficiente, isso pode auxiliar no problema de cobertura vacinal caso ocorram mutações (RAHMAN et al., 2020). Ademais, vacinas que utilizam epítomos são responsáveis por desencadear respostas celulares e humorais mais eficientes (BIBI et al., 2021).

Primeiramente, esse método foi utilizado para prospectar alvos terapêuticos e desenvolver vacina contra o sorogrupo B da bactéria *Neisseria meningitidis* (MenB) (PIZZA et al., 2000), deste então, tem sido utilizada para a identificação de numerosos candidatos de vacina contra patógenos bacterianos, incluindo *Mycoplasma pneumoniae* (RODRIGUES et al., 2019), *Pseudomonas aeruginosa* (SOLANKI, 2019), *Mycobacterium tuberculosis* (MONTERRUBIO-LÓPEZ, 2015) e *Acinetobacter baumannii* (SOLANKI, 2018). Ainda, mais recentemente, essa abordagem foi utilizada para busca de candidatos à fármacos e desenvolvimento de vacinas contra o vírus SARS-CoV2 (MA et al., 2021), o que foi fundamental para salvar inúmeras vidas devido a rápida resposta ao vírus na pandemia da COVID-19.

Segundo RAOUFI et al. (2019) os softwares Bepipred (Larsen J.E. et al, 2006) e ElliPro (PONOMARENKO et al., 2008) são os mais utilizados na predição de epítomos

conformacionais virais e bacterianos para linfócitos B. O software ElliPro prediz os epítomos com base na estrutura 3D da proteína por meio de um sistema de pontuação, onde a estrutura da proteína é comparada a vários elipsoides, de modo que o elipsoide com índice de protrusão (PI) de 0,9 inclua 90% dos resíduos, o de 0,8 inclua 80% dos resíduos e assim por diante. Dessa forma cada resíduo recebe um valor PI com base no centro de massa do resíduo situado fora do maior elipsoide possível; com isso, os resíduos com maior pontuação serão aqueles mais acessíveis aos receptores (PONOMARENKO et al., 2008). O software Bepipred faz a predição baseado na sequência de aminoácidos da proteína, para isso utiliza um algoritmo que combina o modelo oculto de Markov com um dos melhores métodos de escala de propensão (LARSEM J. E. et al, 2006).

Para a predição de epítomos capazes de se ligar ao MHC, os pesquisadores YURINA (2022) e RAOUFI et al. (2019) NetMHCpan (REYNISSON et al., 2020) como um dos mais utilizados e precisos para antígenos bacterianos. O NetMHCpan utiliza redes neurais artificiais treinadas com uma combinação de mais de 850.000 peptídeos quantitativos de afinidade de ligação e ligantes purificados por espectrometria de massa para predição de epítomos exclusivos para MHC classe I (REYNISSON et al., 2020). Para a predição específica dos epítomos para o MHC classe II, a Base de Dados de Epítomos Imunodominantes (do inglês, The Immune Epitope Database, IEDB) recomenda utilizar o método Consensus, trata-se de uma combinação de três dos quatro modelos preditivos disponíveis na plataforma (netMHCII-2.3, netMHCII-2.2, netMHCII-1.1 e Sturniolo), que tenta usar o melhor método de acordo com a proteína de entrada. Caso não seja encontrado um modelo adequado dentre os quatro para essa proteína, é utilizado o modelo NetMHCIIpan (VITA et al., 2019). Esse método é baseado em um estudo em larga escala, com os desempenhos das previsões de lições ao MHC II composto por mais de

50.000 afinidades de ligações e comparações de abordagens pan-específicas (VITA et al., 2019).

Juntas essas ferramentas serão úteis para prospectar os epítomos para linfócitos B, T CD8 e TCD4 presentes nas proteínas de *C. pseudotuberculosis* recuperadas de banco de dados e artigos, para posteriormente modelá-los e verificar sua compatibilidade com ambas as classes do MHC.

3 HIPÓTESES

Utilizando os softwares de bioinformática descritos é possível prever e selecionar epítopos estáveis, atóxicos e hipoalérgico, a partir do proteoma da *C. Pseudotuberculosis*, que tem afinidade com os complexos MHC I e MHC II, bem como os linfócitos B.

4 JUSTIFICATIVA

A LC é uma doença causada pela *C. pseudotuberculosis* com potencial zoonótico que compromete a caprinovinocultura de muitos países, incluindo o Brasil, com destaque para os pequenos produtores da região nordeste, que muitas vezes dependem dessa atividade para suprir suas necessidades básicas. Quando acometidos pela LC, os animais ficam debilitados, impactando negativamente a produção de leite, carne, pele, lã e a reprodução desses animais, comprometendo o rendimento dos produtores. Esses fatores associados com a falta de tratamentos eficazes, vacinas eficientes e diagnósticos precisos, justificam a necessidade de estudos voltados para prospecção de potenciais alvos terapêuticos contra esse agente etiológico, visando o controle e posterior erradicação da doença. Neste contexto, tendo em vista que os experimentos laboratoriais requerem reagentes caros e, muitas vezes, difíceis de encontrar, as ferramentas de bioinformática são imprescindíveis para identificar, analisar e selecionar esses possíveis alvos terapêuticos dentro de um grande conjunto de moléculas, assim barateando os custos de pesquisa *in vitro*.

5 OBJETIVOS DA PESQUISA

5.1 OBJETIVO GERAL

Predizer *in silico* epítomos imunogênicos para linfócitos B, TCD4 e TCD8 por meio de bioinformática, fornecendo dados que poderão ser utilizados no desenvolvimento de futuras vacinas contra a Linfadenite Caseosa.

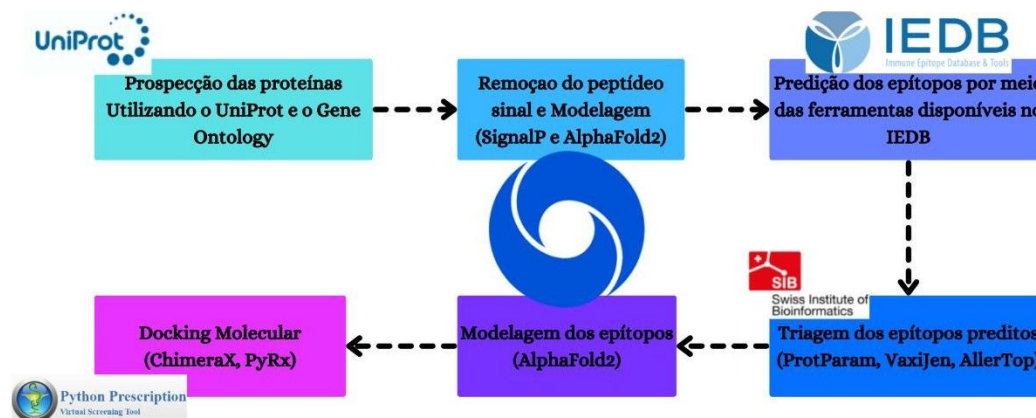
5.2 OBJETIVOS ESPECÍFICOS

- Recuperar proteoma da bactéria C.
- Gerar modelos de proteínas com boa qualidade
- Gerar dados de epítomos que auxiliem em elaboração da vacina;
- Gerar modelos de epítomos de boa qualidade
- Desenvolver scripts em python que auxiliem no tratamento dos dados
- Desenvolver planilhas de dados físico-químicos de epítomos
- Obter epítomos com boa afinidade e estabilidade aos receptores MHCI e II.

6 MATERIAL E MÉTODOS

O passo a passo da metodologia realizada neste estudo está descrito na Figura 1, e foi adaptada de RODRÍGUES-DOMÍNGUEZ et al. (2023) e SILVA et al., (2021b).

Figura 1- Fluxograma da metodologia que será aplicada na predição dos epítomos



Fonte: Autoria Própria

6.1 Recuperação e seleção das sequências

Primeiramente foi realizada a prospecção de proteínas da membrana plasmática, da parede celular e proteínas extracelulares a partir do proteoma da *C. pseudotuberculosis*, através de revisão da literatura e mineração de dados nos bancos NCBI, PDB, UNIPROT.

6.2 Modelagem das Sequencias

Primeiramente foi removido o peptídeo sinal das proteínas selecionadas com a finalidade de otimizar o processo de modelagem tridimensional. Para a modelagem, foi utilizado o software ColabFold v1.5.5 (MIRDITA, M. et al., 2022) disponível no Google Colab. Os modelos selecionados foram os que atingiram um índice de qualificação maior ou igual a 90 no IDDT indicado pelo próprio AlphaFold2. Bem como um total de pelo menos 90% dos resíduos em regiões mais favoráveis e a ausência de resíduos em regiões não permitidas do gráfico de Ramachandran, realizado pela ferramenta PROCHECK (LASKOWSKI R. A. et al., 1993) disponível na plataforma SAVES.

6.3 Predição dos epítomos

As sequências selecionadas tiveram seus peptídeos sinais removidos com o software SignalP, e em seguida seus epítomos foram preditos utilizando os modelos ElliPro (PONOMARENKO et al., 2008) (epítomos lineares para linfócitos B), Bepipret 1.0 (LARSEN J. E. et al, 2006) (epítomos lineares para linfócitos B), NetMHCpan 4.1 (epítomos para MHC classe I e II) (REYNISSON et al., 2020).

6.3.1 Seleção dos epítomos de linfócitos B

A qualidade dos epítomos preditos foi determinada por meio do score fornecido por cada software. Foram selecionados epítomos com resíduos semelhantes e com boa qualificação nos dois softwares utilizados. No caso do ElliPro (PONOMARENKO et al., 2008) foram selecionados epítomos com score acima de 0.6. No programa Bepipret 1.0 (LARSEN J. E. et al, 2006) foram selecionados epítomos com resíduos de score superior a 0.1. Esses valores foram baseados em informações fornecidas na própria documentação do programa. Também foi feita uma predição da inserção da proteína na membrana utilizando a ferramenta Structure Assessment Project do software SWISS-MODEL (WATERHOUSE, Andrew et al., 2018) Essa predição foi utilizada para ter certeza de que o epítomo selecionado está na região extracelular ou intermembranar.

6.3.2 Seleção dos epítomos de linfócitos T-MHC I

A predição de epítomos dos linfócitos T-MHC I se deu com o software NetMHCpan 4.1 (REYNISSON et al., 2020) EL, disponível na plataforma IEDB (Immune Epitope Database). Como não está disponível alelos de ovinos ou caprinos, foram utilizados alelos de bovinos devido a maior semelhança filogenética entre as espécies. Os alelos utilizados na predição foram BoLa-1:02801, BoLa-2:01601, BoLa-2:02201, BoLa-3:03601, BoLa-3:00101, BoLa-3:01001, BoLa-3:01101 e BoLa-6:01501, que segundo estudos, são alelos mais presentes na população de gado brasileira. Foi utilizado um número de resíduos de 8 a 10 devido a capacidade de reconhecimento do MHC classe I está nessa faixa. Os epítomos selecionados serão aqueles que obtiverem um Score (índice de probabilidade da sequência se ligar ao receptor) acima de 0.5 e um Percentile Rank (força de ligação com o receptor) abaixo de 1.

6.3.3 Seleção dos epítomos de linfócitos T-MHC II

A predição de epítomos dos linfócitos T-MHC II se deu com o software NetMHCIIpan 4.1EL (REYNISSON et al., 2020), disponível na plataforma IEDB (Immune Epitope Database). Devido a falta de alelos ovinos, caprinos e bovinos nesse software, foram selecionados os alelos murinos para a predição de epítomos, sendo eles: H2-IAb, H2-Aq, H2-Ak, H2-As, H2-Ad, H2-Au, H2-Eq, H2-Eu. Foi utilizado um número de resíduos de 12 a 18 devido a capacidade de reconhecimento do MHC classe II se encontrar nessa faixa. Os epítomos selecionados foram aqueles que obtiveram um Score (índice de probabilidade da sequência se ligar ao receptor) acima de 0.5 e um Percentile Rank (força de ligação com o receptor) abaixo de 1.

6.4 Triagem dos epítomos preditos

Os epítomos preditos foram selecionados com base no grau de estabilidade com ProtParam (GASTEIGER, 2005), antigenicidade com o VaxiJen, (DOYTCHINOVA, 2007), alergenicidade com o AllerTop, (DIMITROV, 2014) e toxicidade com o ToxinPred, (GUPTA, 2015). Outros dados obtidos com o ProtParam foram características físico-químicas, tais como: massa molecular, ponto isoelétrico, grau estabilidade, tempo de meia vida in vitro e in vivo, índice alifático, grau de hidrofobicidade (GASTEIGER, 2005).

6.4.1 ProtParam

Esse software, disponível na plataforma Expasy.org, foi utilizado para estimar dados úteis sobre as características físico-químicas dos polipeptídios com base em suas sequências de aminoácidos. Os dados dessas características podem ser úteis na elaboração da vacina. A triagem foi feita baseada no índice de instabilidade, esse índice mostra se a proteína tem uma tendência a ser estável ou instável, o threshold utilizado será 40, indicado pelo próprio programa (GASTEIGER, 2005).

6.4.2 VaxiJen 2.0

O VaxiJen utiliza a sequência da proteína e seu organismo de origem para prever se uma proteína dispara alguma resposta imune no organismo humano. Foram selecionados os peptídeos que possuíam OPPA (Overall Prediction for the Protective Antigen) acima de 0.5, indicando que o mesmo possui antigenicidade. Esse número foi baseado na própria documentação do programa (DOYTCHINOVA 2007).

6.4.3 AllerTop

Esse programa utiliza de modelos matemáticos para prever se uma sequência de aminoácidos é capaz de disparar uma reação alérgica (DIMITROV, 2014).

6.4.4 ToxinPred

O ToxinPred se baseia na sequência de aminoácidos para prever se um peptídeo é tóxico para o ser humano. (GUPTA, 2015)

6.4.5 Validação *in silico*

Utilizando o programa ColabFold v1.5.5 (MIRDITA, M. et al., 2022) os epítomos preditos foram modelados, e tiveram sua afinidade de se ligar com as moléculas do sistema imune avaliadas por meio do Docking Molecular realizado no software PyRx, utilizando as moléculas dos MHC I e MHC II de *Ovis aries* recuperados do UniProt (IDs: A0A2P9DU17 e A0A1Y5U0V2 respectivamente) como receptores. Os dados de qualidade fornecidos pelo programa são: Binding Affinity(kcal/mol), que indica a energia livre da ligação com o receptor, RMSD lower bound e RMSD Upper bound que indica se a posição dos átomos na molécula é favorável. (TORCHALA et al., 2013).

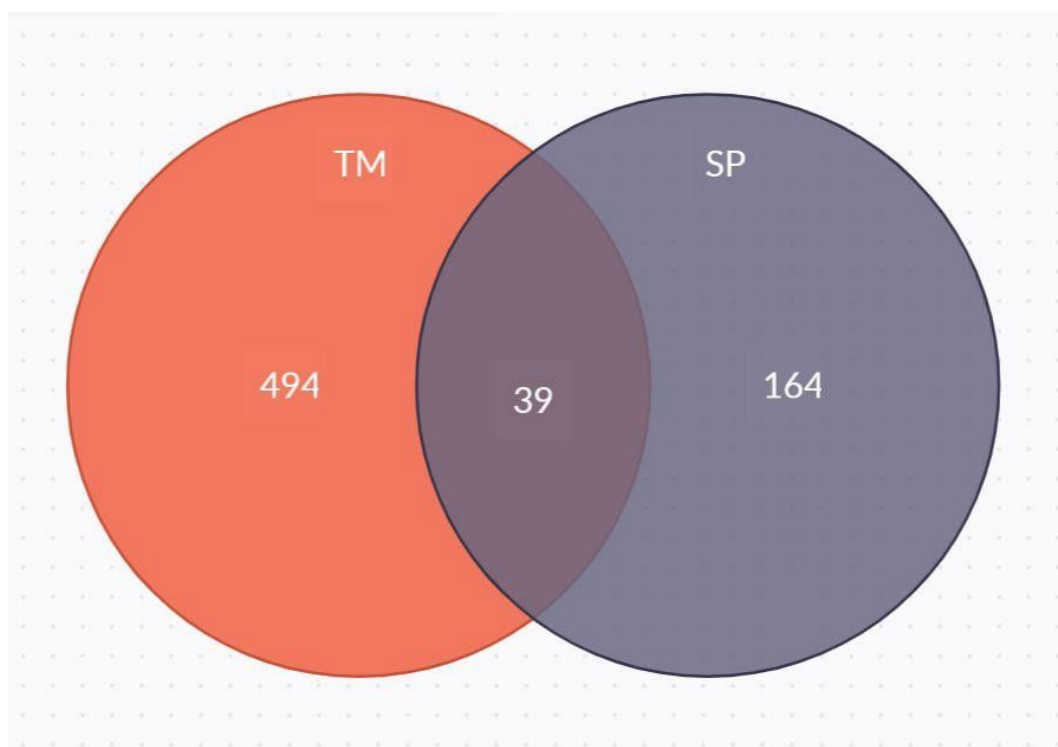
7 RESULTADOS

7.1 Recuperação e seleção das sequencias

Primeiramente foi feita a recuperação do proteoma completo da *Corynebacterium pseudotuberculosis* (strain C231), disponível no UniProt, o que resultou em um arquivo fasta com 2002 sequencias de proteínas (disponível em: <https://drive.google.com/file/d/17JbGKnqBrGp93hmf1RM-B3-dUdeSWAaL/view?usp=sharing>).

Em seguida esse arquivo foi submetido a uma triagem no software TOPCON, esse software faz uma predição de domínios transmembrana TM e peptídeos sinais (SP) contidos nas sequências de aminoácidos das proteínas, com a finalidade de identificar a localização subcelular das proteínas. Com isso, foram obtidas um total de 697 proteínas que possuem domínios transmembranas ou/e peptídeos sinais, sendo que 39 possuem TM e SP, 494 possuem apenas TM e 164 possuem apenas SP (Figura 2).

Figura 2 - Diagrama de Venn com informações das análises do TOPCONS. T omíniosnios transmembrana, SP = peptídeo sinal.



Fonte: autoria própria

Subsequentemente, as 697 proteínas preditas como extracelulares ou de membrana foram submetidas a uma pesquisa no UniProt utilizando seu ID. Com isso tiveram sua localização subcelular determinada utilizando as anotações do Gene Ontology, o parâmetro observado foi o Cellular Component, que determina em qual região a proteína se encontra. Assim, foi confirmado que 212 proteínas atingiram os parâmetros de localização, sendo: 151 localizadas na membrana plasmática, 31 são proteínas secretadas, 16 pertencem ao ATP-Binding-Casset (um complexo proteico de transporte transmembranar de solutos para dentro e fora da célula), 8 são proteínas extracelulares, 3 estão no espaço periplasmático (região entre a membrana plasmática e a parede celular) e 3 pertencem ao complexo TAT (Complexo de transporte de proteínas dobradas para fora da célula ou entre organelas e o citoplasma) (Tabela 1). As demais proteínas não possuíam anotação no Gene Ontology e foram excluídas.

A tabela completa gerada com os dados das proteínas selecionadas está disponível em: <https://docs.google.com/spreadsheets/d/17oLFtOicSNE4j5KflbvBOJjVWVfRIV3k/edit?usp=sharing&ouid=103272525517299048048&rtpof=true&sd=true>.

Tabela 1 Número de proteínas com localização celular de interesse selecionadas com base nas informações do Gene Ontology

Proteínas de interesse	
ATP-Binding-Casset =	16
plasma membrane	151
extracellular region	8
Periplasmatic Space =	3
Secreted Protein=	31
TAT complex =	3
total=	212

Fonte: Autoria Própria

7.2 Modelagem das Sequencias

Das 212 sequencias que foram submetidas a modelagem, 37 atingiram o padrão de qualidade estabelecido na metodologia (6.2), para o gráfico de Ramachandran e IDDT, foram utilizadas para a predição de epítomos. Algumas foram excluídas pela ausência de informação do IDDT geral, onde não foi possível identificar a qualidade da proteína.

7.3 Predição dos Epítomos

7.3.1 Linfócitos B

Das 37 sequencias selecionadas na etapa anterior, 32 apresentaram epítomos que atendiam os padrões de qualidade estabelecidos na metodologia (6.3.1), para linfócitos B. Sendo 6 epítomos com 15 resíduos, 4 com 14 resíduos, 6 com 13 resíduos, 3 com 12 resíduos, 6 com 11 resíduos, 2 com 10 resíduos e 5 com 9. Esses epítomos seguiram para a triagem físico-química. Dados disponíveis na tabela 2

Tabela 2 Epítomos de linfócitos B selecionados. Celulas com pontos indicam dados abaixo da média. Celulas em branco indicam dados não recuperados.

Protein_ID	Elipore_Score	Bepipred	Resid_ID	Epitop_LinearSeq	Num_resid
A0A6D2LEJ1_86256	0.724		424-439	EEEDETEVESTADAS	15
D9Q9A5_1921a	0.742	0.524-0.765	123-135	VDFGQAEAAARSLG	13
D9Q9L7_fa3b0	0.887	0.566-1.928	30-43	AVADESAPTSTWGE	14
D9Q9S2_bb395	0.822	0.734-2.179	225-239	HNDYGPYTDGTGNPS	15
D9Q9T5_3013c	0.799	0.414-2.303	270-282	KPTTPEDQASSIN	13
D9Q9Y1_a5967	0.731	0.284-(-1.396)	177-187	GPDARAAQGVVM	11
D9Q982_97ff8	0.805	0.394-1.135	157-167	YVQAARNAGAS	11
D9QA26_3c1b4	0.832	0.436-1.341	213-224	RASQKIPRPTAR	12
D9QA80_8a90d	0.805	0.502-1.381	210-224	KDLDKDEEFQERLKD	15
D9QA85_8ebd8	0.809	0.376-1.026	205-217	LSTGESVSKAIGA	13
D9QAE3_7a5ac*	0.82	0.573-1.483	210-222	VSAPGKKEKYIPT	13
D9QAU0_1a18c	0.734	0.404-1.634	36-47	SFSYSPPAPIQM	12
D9QAY3_b9a95	0.901	0.262-1.029	420-428	VGEGQTARF	9
D9QAZ6_05ca1	0.862	0.483-1.320	185-197	DIRGSRRTSTAG	13
D9QB77_abef5	0.756	0.512-0.565	470-480	CWDGTDVPVTRR	11
D9QB91_2753b	0.74	0.184-1.453	8~16	YLPSPQGV	9
D9QBB8_d7faa	0.866	0.527-1.907	284-294	KKISAPHSSQD	11
D9QBL2_d6336	0.812	0.421-1.435	1~15	MSRNTESSRGSL	12
D9QBU0_6b62f	...	..	..	..	..
D9QBU1_98faa	0.866	0.356-1.484	267~277	RDLTDPRSTRR	11
D9QC81_017cd	...				
D9QCC6_a596a	0.671	0.437-1.159	1~14	CGEKKQQVAGSQWQ	14
D9QCC6_a596a	0.672	0.575-1.804	59-73	SFSAEGKQDVAPQDA	15
D9QCI7_5fb02	0.826	0.008-0.97	89-97	RAEPPCSVI	9
D9QCI8_5d925	...	...	...		
D9QD06_59a73	0.764	0.396-1.223	214-223	MPETEKTLTD	10
D9QD82_eefde	...			...	3
D9QDC3_cc76b	0.731	0.250-1.306	33-41	SPADPARLA	9
D9QDF0_97491	0.623	0.394-1.315	120-128	ARPEAGNIS	9
D9QDG9_fcdee	..0..	...	..	..	..
D9QDI0_8c9ea	0.763	0.422-1.285	110-123	AGVASTDHDIVDAA	14
D9QDM4_40196	0.657	0.470-2.426	163-177	NSEAE DRPNPKGLK	15
D9QDM4_40196	0.697	0.498-1.684	126-135	AEASDQDSYV	10
D9QE05_d6df5	0.709	0.540-1.322	85-99	VQSISTGSFSGGPF	14
D9QED2_17dfd	0.699	0.413-1.983	34~48	RSNYAGRPPVPESR	15
D9QEE1_b738c	0.851	0.405-1.402	28-40	QPVLQVPSDSSVG	13
D9QEK3_b1922	...	...	...		...
D9QEL1_940e4	0,86	0,428-1,332	270-280	GRRTEKSDASC	11
D9QER2_b0015	...	...	...	...	...

7.3.2 Linfócitos T MHC-I

Das 37 sequencias selecionadas na etapa anterior, 35 apresentaram epítomos que atendiam os padrões de qualidade estabelecidos na metodologia (6.3.2), para Linfócitos T MHC-I. O Alelo que apresentou maior porcentagem de epítomos de boa qualidade foi o BoLA-3:01101, com 85,7% dos epítomos totais. Sendo que 33 epítomos possuem 9 resíduos e 2 possuem 8 resíduos. Esses epítomos seguiram para triagem físico-química. Dados disponíveis na tabela 3

Tabela 3 Epítomos de linfócitos T MHC-I selecionados. Células com pontos indicam dados abaixo da média. Células em branco indicam dados não recuperados.

Protein ID	Allele	Start	End	Length	Peptide	Score	Percentile Rank
A0A6D2LEJ1_86256	BoLA-1:02801	513	521	9	AQLSGKVSL	0.7078	0.01
D9Q9A5_1921a	BoLA-3:01101	167	175	9	KNTTIASVI	0.733968	0.02
D9Q9L7_fa3b0	BoLA-3:01101	258	266	9	LGIGLPPHI	0.519827	0.1
D9Q9S2_bb395	BoLA-3:01101	322	329	8	KSYETIFL	0.529928	0.1
D9Q9T5_3013c	BoLA-3:01101	133	141	9	AGFGAPIAI	0.648336	0.03
D9Q9Y1_a5967	BoLA-3:01101	308	315	8	IAAPPQEI	0.668564	0.03
D9Q982_97ff8	BoLA-3:01101	136	144	9	VGSPYAARV	0.505448	0.11
D9QA26_3c1b4	BoLA-3:01101	210	218	9	IGARASQKI	0.776903	0.01
D9QA80_8a90d	BoLA-3:01101	245	253	9	KSAYRATYI	0.80258	0.01
D9QA85_8ebd8	BoLA-3:03601	239	247	9	SAAGPIGFL	0.555169	0.01
D9QAE3_7a5ac*	BoLA-3:01101	406	414	9	AVFNHPFLI	0.617311	0.05
D9QAU0_1a18c	BoLA-3:01101	160	168	9	KSALFPLVI	0.864628	0.01
D9QAY3_b9a95	BoLA-3:01101	56	64	9	TGFPVNTIL	0.551362	0.08
D9QAZ6_05ca1	BoLA-3:01101	115	123	9	KSKPLTTF	0.637406	0.04
D9QB77_abef5	BoLA-3:01101	66	74	9	ISAFKAFTI	0.639334	0.03
D9QB91_2753b	BoLA-3:01101	158	166	9	TSVPWALEI	0.652618	0.03
D9QBB8_d7faa	BoLA-3:01101	198	206	9	GSAHSALLI	0.677899	0.02
D9QBL2_d6336	BoLA-3:01101	61	69	9	SSSMLPILI	0.798827	0.01
D9QBU0_6b62f	BoLA-3:01101	83	91	9	TSSTDISSI	0.622876	0.04
D9QBU1_98faa	BoLA-3:01001	35	43	9	TPYDPLRVM	0.658695	0.01
D9QC81_017cd	BoLA-3:01101	287	295	9	ASSAYTLKV	0.659997	0.03
D9QCC6_a596a	BoLA-3:03601	129	137	9	SVDAPAIRL	0.497861	0.02
D9QCI7_5fb02	BoLA-3:01101	89	97	9	RAEPPCSVI	0.522446	0.1
D9QCI8_5d925	...	...	...	...	...	...	...
D9QD06_59a73	BoLA-3:01101	351	359	9	AGFANISSI	0.574614	0.03
D9QD82_eefde	BoLA-3:01101	260	268	9	ASPHIAQRI	0.701826	0.02
D9QDC3_cc76b	BoLA-3:01101	95	103	9	KAFPIITLQL	0.839926	0.01
D9QDF0_97491	BoLA-3:01101	127	135	9	ISSRLLTI	0.810458	0.01
D9QDG9_fcdee	BoLA-3:01101	336	344	9	RTIAHPTEI	0.904348	0.01
D9QDI0_8c9ea	BoLA-3:01101	70	78	9	RALPTLGLL	0.537584	0.09
D9QDM4_40196	BoLA-3:01101	98	106	9	SALPSSLQV	0.689284	0.02
D9QE05_d6df5	BoLA-3:01101	128	136	9	RAAAPHARI	0.87136	0.01
D9QED2_17dfd	BoLA-3:01101	90	98	9	INSQISYRI	0.58171	0.07
D9QEE1_b738c	BoLA-3:01101	366	374	9	RAMAASVTV	0.634039	0.04
D9QEK3_b1922	BoLA-3:01101	80	88	9	ITAFGPLTL	0.504236	0.11
D9QEL1_940e4	..	...	...	...		...	...
D9QER2_b0015	BoLA-3:01101	147	155	9	ISFALPLTI	0.808356	0.01

7.3.3 Linfócitos T MHC-II

Das 37 sequencias selecionadas na etapa anterior, 36 apresentaram epítomos que atendiam os padrões de qualidade estabelecidos na metodologia (6.3.3), para os Linfócitos T MHC-II. O

Alelo que apresentou maior porcentagem de epítomos de boa qualidade foi o h2-IAb, com 52.77% dos epítomos totais. Sendo que 1 epítomo possuía 16 resíduos, 8 epítomos possuíam 15 resíduos, 11 possuíam 14 resíduos e 16 possuíam 13 resíduos. Esses epítomos seguiram para triagem físico-química. Dados disponíveis na tabela 4

Tabela 4 Epítomos de linfócitos T MHCII selecionados. Células com pontos indicam dados abaixo da média. Células em branco indicam dados não recuperados.

Protein ID	Allele	Start	End	Length	Peptide	Score	Percentile Rank
A0A6D2LEJ1	H2-IAb	92	104	13	ITQLIGPKAVASY	0.8946	0.05
D9Q9A5_192	H2-IAb	142	154	13	FNKIIFPQAIRAA	0.7434	0.25
D9Q9L7_fa3b	H2-IAd	205	2018	14	VKNEEFVTAARAVG	0.9443	0.01
D9Q9S2_bb3b	H2-IAb	47	59	13	NPFVAAALSAGASD	0.6752	0.37
D9Q9T5_301b	H2-IAb	132	144	13	LAGFGAPIAIVAA	0.9180	0.02
D9Q9Y1_a59b	H2-IAb	172	185	14	FLVFKGPDARAAQG	0.9675	0.01
D9Q982_97fb	H2-IAb	21	34	14	AQLFQLPTPNAAG	0.9571	0.02
D9QA26_3c1f	H2-IAd	144	156	13	AGPAITVYAQVSR	0.7103	0.29
D9QA80_8a9b	H2-IAd	309	321	13	AGKVASTTVFKAG	0.6777	0.37
D9QA85_8ebf	H2-IAd	209	223	15	ESVSKAIGANPALVR	0.8347	0.21
D9QAE3_7a5f	H2-IAd	203	217	15	QRLAEGLV SAPGKKE	0.7922	0.38
D9QAU0_1a1f	H2-IAb	35	48	14	RSFSYSPPAPIQMP	0.9089	0.11
D9QAY3_b9af	H2-IAb	12	27	16	GDGKTIAPGAVVAPEE	0.9378	0.1
D9QAZ6_05cf	H2-IAb	158	170	13	NIPFLGPDVVKLV	0.8223	0.12
D9QB77_abef	H2-IAd	60	73	14	RTDEQGISAFKFT	0.9478	0.01
D9QB91_275f	H2-IAd	183	195	13	SNGEVIATVHPTF	0.6808	0.37
D9QBB8_d7ff	H2-IAb	158	171	14	KGLPLSPSASLAAE	0.5850	0.98
D9QBL2_d63f	H2-IAb	246	258	13	LPILGSPKVMAN	0.8357	0.11
D9QBU0_6b6f	H2-IAd	32	48	17	PGNPAEVALGVTATPEA	0.6255	0.8
D9QBU1_98ff	H2-IAu	228	242	15	GRMLQSAQSSLATAP	0.5855	0.52
D9QC81_017f	H2-IAb	282	294	13	DIYNAASSAYTLK	0.6830	0.35
D9QCC6_a59f	H2-IAb	127	140	14	GDSVDAPAIRLVVK	0.6375	0.74
D9QCI7_5fb0	H2-IAd	148	161	14	IDAEIAAIAQYEQR	0.7752	0.4
D9QCI8_5d9f	..	..	..	..	..	..	..
D9QD06_59af	H2-IAu	242	256	15	GRGAMSGGQIAIVG	0.5261	0.86
D9QD82_eefb	H2-IAb	53	65	13	VFSWRAPRIVTAL	0.8892	0.05
D9QDC3_cc7f	H2-IAb	186	198	13	NNLFLPAVALAVP	0.5760	0.6
D9QDF0_974f	H2-IAu	114	128	15	AYAIVSARPEAGNIS	0.6890	0.14
D9QDG9_fce0	H2-IAb	90	102	13	VWQFRLPRTLLAG	0.5775	0.6
D9QDI0	H2-IAb	14	26	13	ASQWWGPTGYLAR	0.7751	0.16
D9QDM4_401f	H2-IAb	83	95	13	DTTARLPEAVYAE	0.6308	0.42
D9QE05_d6df	H2-IAb	156	170	15	IPNITLPVPAGAMER	0.7589	0.45
D9QED2_17df	H2-IAd	289	303	15	KNDIHGVVAVAAFGA	0.8433	0.19
D9QEE1_b73f	H2-IAb	153	165	13	KNIFRNPRPYIAH	0.9297	0.02
D9QEK3_b19f	H2-IAu	125	138	14	PGVIEAARSMGASP	0.6044	0.23
D9QEL1_940f	H2-IAd	236	249	14	GLWITGVSARISQP	0.7956	0.31
D9QER2_b00f	H2-IAb	235	249	15	PSAFEFPYAAIMAG	0.7687	0.44

7.4 Triagem dos epítomos

7.4.1 ProtParam

Foram coletados dados sobre o peso molecular, ponto isoelétrico teórico, coeficientes de extinção, meia-vida estimada, índice de instabilidade, índice alifático e GRAVY dos epítomos preditos anteriormente. Além da triagem realizada com base no grau de estabilidade,

os demais dados serão úteis na elaboração da vacina. Um total de 56 epítopos atenderam o padrão de estabilidade estabelecido na metodologia (6.4) e seguiram para a análise posterior, sendo que 12 eram de linfócitos B, 25 eram de linfócitos T-MHC I e 21 eram de linfócitos T-MHC II.

7.4.2 VaxiJen

Após os testes de antigenicidade foi constatado um total de 36 epítopos apresentaram antigênicos sendo que 8 eram linfócitos B, 13 eram de linfócitos T-MHC I e 15 eram de linfócitos T-MHC II, esses epítopos seguiram para análise de alergenicidade e toxicidade.

7.4.3 AllerTop e ToxinPred

No teste de alergenicidade um total 19 epítopos foram preditos como não alergênicos, sendo que, dentre eles, 5 eram linfócitos B, 8 eram de linfócitos T-MHC I e 6 eram de linfócitos T-MHC II. Já a análise de toxicidade não acusou nenhum epítipo como tóxico, então esses epítopos seguiram para a modelagem. Os dados detalhados estão disponíveis no anexo A e em melhor qualidade por meio do link: <https://encurtador.com.br/gio09>

7.5 Docking Molecular

O preparo das moléculas para o docking molecular foi realizado no software ChimeraX, o docking foi realizado no software PyRx. O docking foi realizado com 14 epítopos, exceto os de linfócitos B, que possuem um mecanismo diferenciado de ativação.

7.5.1 Linfócitos T MCHI

Todos os 8 epítopos de Linfócitos T MCHI se ligaram muito bem ao receptor. Todos os parâmetros estabelecidos na metodologia (6.5) foram positivos. Energia de ligação negativa, RMSD Upper e lower zerados, mostrando alta afinidade e estabilidade na ligação dos epítopos com o receptor celular. Dados disponíveis na tabela 5. Figuras 2 e 3 demonstram a posição dos epítopos no receptor.

Tabela 5 Dados do docking molecular de epítomos contra MHC I realizado no PyRx

Ligand	Binding Affinity (kcal/mol)	Mode	RMSD lower bound >	RMSD upper bound
AF-A0AZP9DU17-F1-model_v4_prep_PREP_D9Q9T5_3013c	-8.0	0	RMSD lower bound	0.0
AF-A0AZP9DU17-F1-model_v4_prep_prep_D9QA26_3c1b4j	-6.3	0	0.0	0.0
AF-A0AZP9DU17-F1-model_v4_prep_prep_D9QAE3_7a5ac	-7.5	0	0.0	0.0
AF-A0AZP9DU17-F1-model_v4_prep_prep_D9QAU0_1a18c	-7.4	0	0.0	0.0
AF-A0AZP9DU17-F1-model_v4_prep_prep_D9Q888_d7faa	-7.2	0	0.0	0.0
AF-A0AZP9DU17-F1-model_v4_prep_prep_D9QC81_017cdj	-7.7	0	0.0	0.0
AF-A0AZP9DU17-F1-model_v4_prep_prep_D9QCC5_a596a	-8.5	0	0.0	0.0
AF-A0AZP9DU17-F1-model_v4_prep_prep_D9QER2_b0015	-7.3	0	0.0	0.0

Figura 3 – Docking Molecular entre epítomos preditos (Coloridos) e molécula MHC classe I (marrom). (A) Epítomo D9Q9T5. (B) Epítomo D9QA26. (C) Epítomo D9QAE3. (D) Epítomo D9QAU0

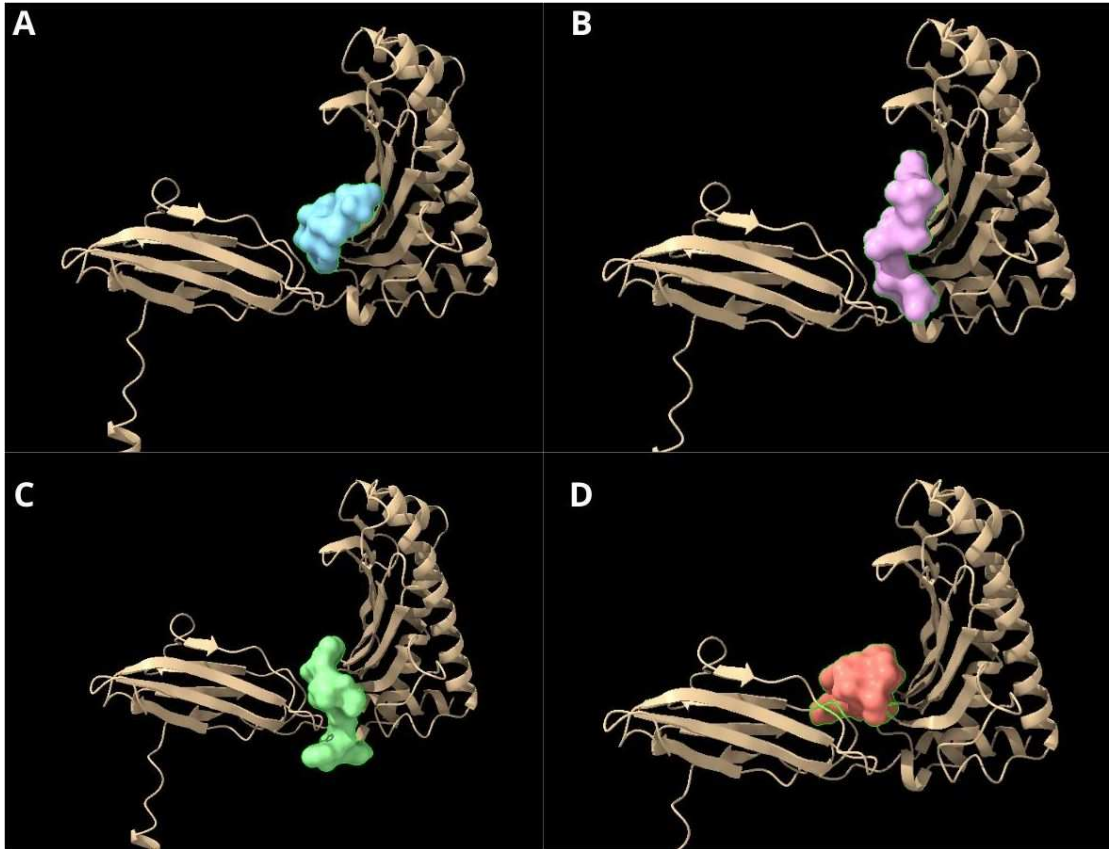
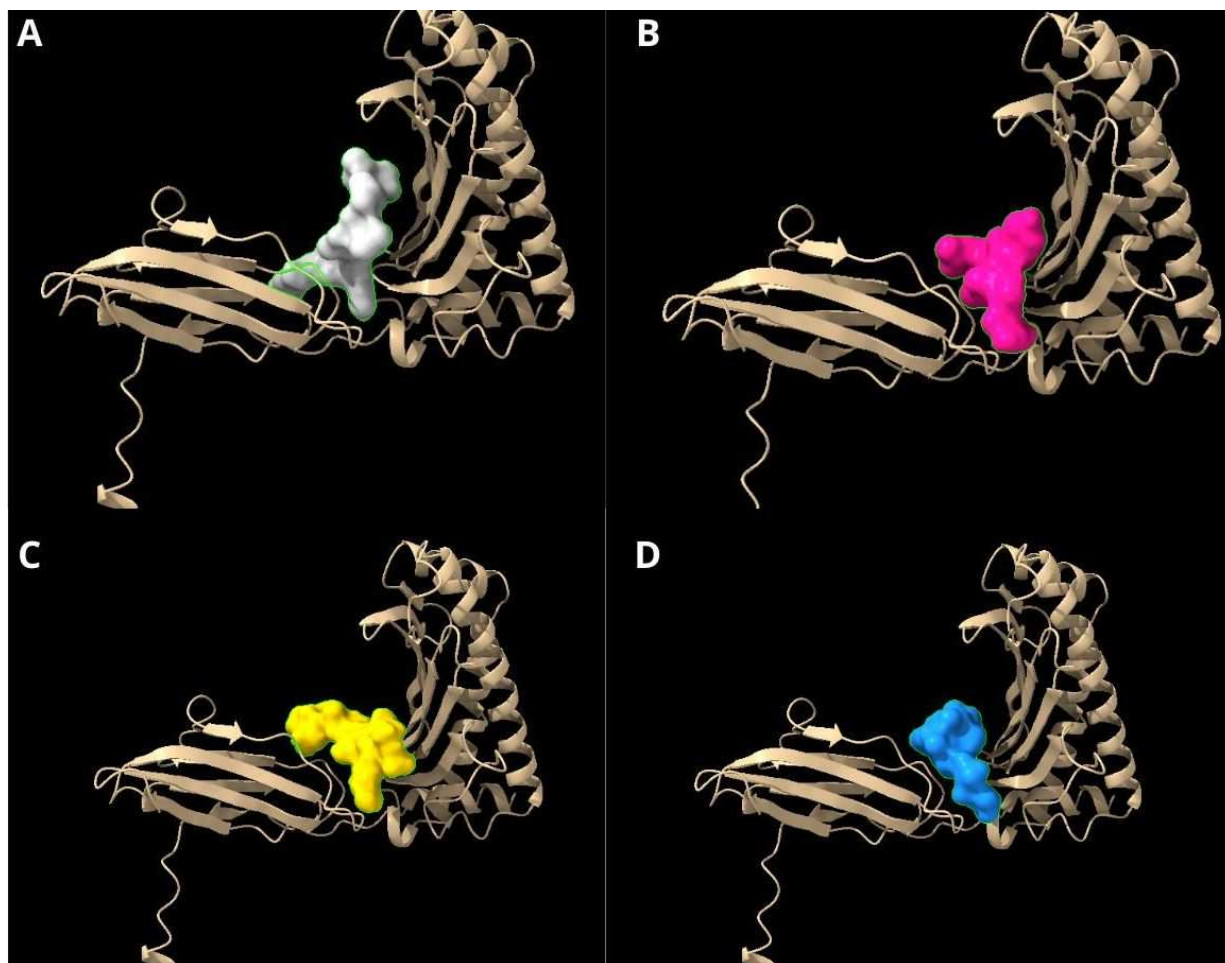


Figura 4 – Docking Molecular entre epítomos preditos (Coloridos) e molécula MHC classe I (marrom). (A) Epítopo D9QBB8. (B) Epítopo D9QC81. (C) Epítopo D9QCC6. (D) Epítopo D9QER2



7.5.2 Linfócitos T MCHII

Todos os 6 epítomos de Linfócitos T MCHII se ligaram muito bem ao receptor. Todos os parâmetros estabelecidos na metodologia (6.5) foram positivos. Energia de ligação negativa, RMSD Upper e lower zerados, mostrando alta afinidade e estabilidade na ligação dos epítomos com o receptor celular. Dados disponíveis na tabela 6. Figuras 4 e 5 demonstram a posição dos epítomos no receptor.

Tabela 6 Dados do docking molecular de epítomos contra MHCII realizado no PyRx

Ligand	Binding Affinity (kcal/mol)	Mode »	RMSD lower bound	RMSD upper bound
AF-ADA1YSU0V2-F1-model_y4_prep_prep_D9Q9Y1_a5967_I	-6.8	0	0.0	0.0
AF-ADA1YSU0V2-F1-model_y4_prep_prep_D9Q882_97f8_I	-6.8	0	0.0	0.0
AF-ADA1YSU0V2-F1-model_y4_prep_prep_D9Q485_8ebd8_I	-11.2	0	0.0	0.0
AF-ADA1YSU0V2-F1-model_y4_prep_prep_D9Q877_abef5_I	-9.2	0	0.0	0.0
AF-ADA1YSU0V2-F1-model_y4_prep_prep_D9Q882_eefde_I	-6.8	0	0.0	0.0
AF-ADA1YSU0V2-F1-model_y4_prep_prep_D9Q010_8c3ea_I	-7.4	0	0.0	0.0

Figura 5 - Docking Molecular entre epítomos preditos (Coloridos) e molécula MHC classe II (marrom). (A) Epítopo D9Q9Y1. (B) Epítopo D9Q982. (C) Epítopo D9QA85. (D) Epítopo D9QB77

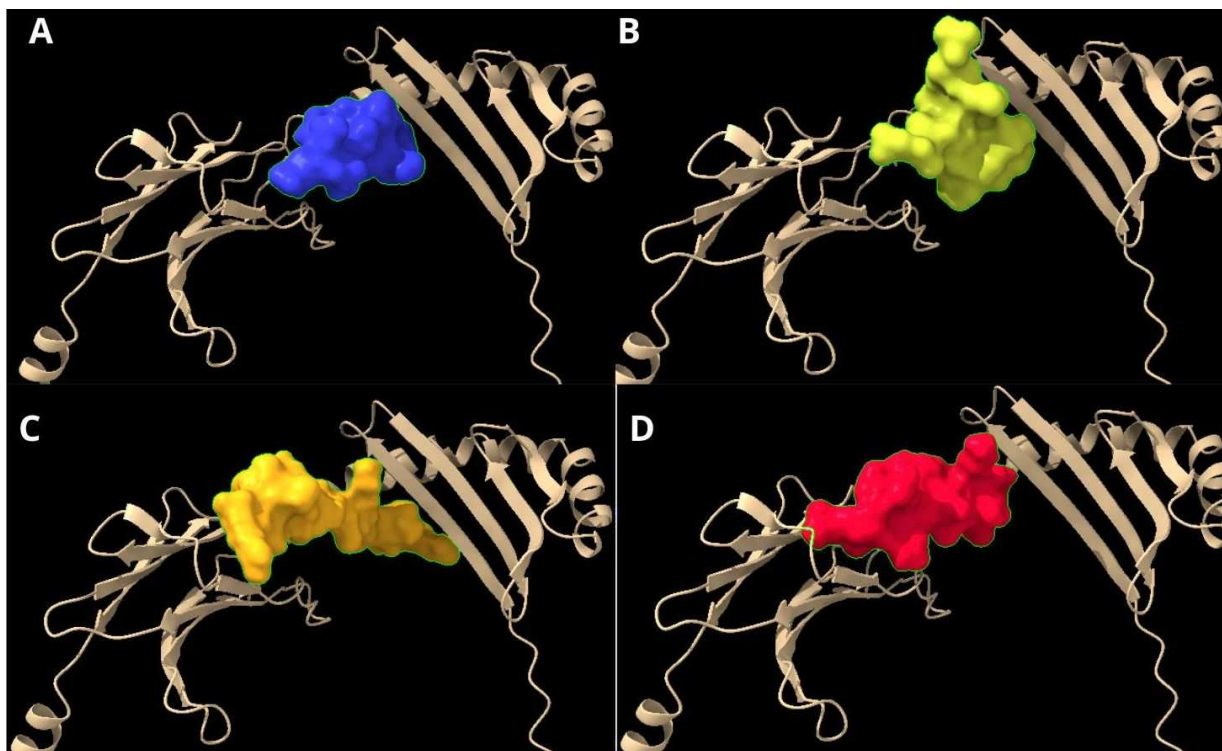
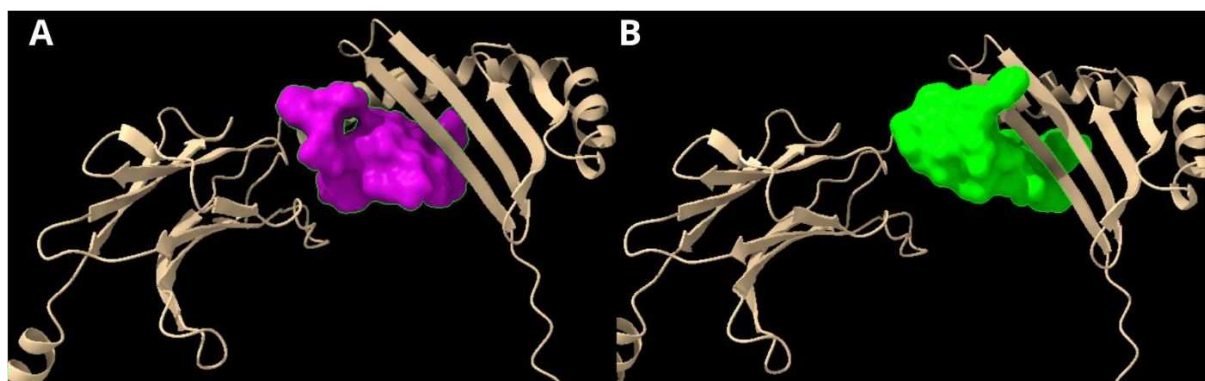


Figura 6- Docking Molecular entre epítomos preditos (Coloridos) e molécula MHC classe II (marrom). (A) Epítopo D9QD82. (B) Epítopo D9QD10.



8 DISCUSSÃO

Os resultados mostraram que os melhores epítomos preditos foram D9QCC6 e D9QA85 dos receptores MHCI e MHCII, respectivamente, por apresentarem menor energia livre e, conseqüentemente, uma ligação mais estável com o receptor. Dos 5 epítomos de linfócitos B (D9Q9L7, D9Q9Y1, D9Q982, D9QAE3 e D9QDM4), os melhores foram D9Q9L7 e D9QDM4 por apresentarem melhores pontuações nos softwares de predição, bem como boa pontuação no teste de antigenicidade e alto tempo de meia vida. Conforme mostrado, todos os epítomos de linfócitos B estão em regiões expostas da proteína e da célula, facilitando o reconhecimento pelo sistema imune. Com relação a similaridade entre os epítomos, os epítomos D9Q9Y1 e D9QAE3 de linfócitos B e T MHCII se mostram semelhantes em alguns pontos, possuindo regiões sobrepostas, como também encontrado por RODRÍGUEZ-DOMÍNGUEZ et al. (2023).

Aplicando uma metodologia parecida, MARQUES et al., (2019), fez uma predição de epítomos para linfócitos B a partir do proteoma da *C. pseudotuberculosis*. Como resultado obteve 315 proteínas extracelulares, com uma quantidade ainda maior não revelada de proteínas de membrana, uma quantidade bem superior a encontrada no presente trabalho. Isso foi consequência, em parte, do uso do software o CELLO2GO, que pode ser mais efetivo do que as anotações do GeneOntology para identificação da localização subcelular. Entretanto, apesar de mais proteínas, o autor obteve 4 epítomos de linfócitos B, um a menos dos obtido no presente trabalho, isso pode ser consequência da metodologia baseada em regiões conservadas nas cepas de *C. pseudotuberculosis*.

Além disso, DROPPA-ALMEIDA (2020) utilizou uma metodologia diferente para selecionar antígenos de *C. pseudotuberculosis*. A partir do genoma da bactéria foi selecionada apenas as proteínas extracelulares pelo software Vaxign para posterior caracterização e predição de epítomos pelo IEDB. Foi predito um total de 22 proteínas extracelulares, um número bem inferior aos dois trabalhos apresentados na atual discussão, indicando que esse software pode estar desatualizado ou pode não seja indicado para o genoma bacteriano. Ademais, após a caracterização, foram obtidos 7 epítomos capazes de ativar o sistema imunológico com posterior utilização do software Vaccine CAD que é capaz de melhorar a entrega e ativação da resposta imune. Porém não fica claro quais são esses epítomos, de quais proteínas saíram, suas sequências e muito menos os receptores no qual esses epítomos se ligam, assim dificultando a comparação entre os trabalhos.

Os epítomos preditos nesse trabalho podem enriquecer os bancos de dados de epítomos, que são repositórios que agregam epítomos de agentes patogênicos. Esses bancos têm sido cada vez mais consultados na identificação de antígenos para uso em vacinas, diagnósticos e terapias imunológicas. A ausência de vacinas e diagnósticos sorológicos para a LC torna crucial a análise de novos alvos. Por isso, essas proteínas foram analisadas *in silico*, permitindo a identificação dos epítomos imunodominantes. A seleção desses epítomos visou incluir sequências atóxicas, hipoalergênicas e estáveis, e fundamentalmente, epítomos capazes de se ligar aos receptores das células do sistema imune do hospedeiro. Essas características são fundamentais para a transformação das células T e B virgens em linfócitos efetores específicos para o patógeno, para garantir a segurança do animal que possivelmente receberá o imunizante, e, por conseguinte, garantir a produção de anticorpos (DROPPA-ALMEIDA et al., 2018).

9 CONSIDERAÇÕES FINAIS

Com isso, com as ferramentas de bioinformática apresentadas no estudo foi possível prever um total de 19 epítomos antigênicos que interagem com os receptores do sistema imune, além de estáveis, atóxicos e hipoalergênicos, sendo que, dentre eles, 5 eram linfócitos B, 8 eram de linfócitos T-MHC I e 6 eram de linfócitos T-MHC II. Esses epítomos podem formular uma vacina multi-epitopo para a LC, porém estudos adicionais *in vitro* e *in vivo* com os epítomos preditos são de extrema necessidade para confirmar a eficácia dos epítomos.

Os resultados indicam que a combinação das ferramentas bioinformática representam um meio valioso e de grande relevância na identificação de epítomos como possíveis alvos de antigenicidade, com base em proteomas já sequenciados. O sequenciamento de proteomas de patógenos possibilitou a identificação e análise dos principais antígenos associados a doenças de forma mais rápida e econômica.

10 REFERÊNCIAS¹

- ABBAS, A.K.; LICHTMAN, A.H.; PILLAI, S. *Imunologia Celular e Molecular*. 8ª Edição. Elsevier, 2015
- ALMEIDA, J. V. F. C. Estudo farmacocinético de nanopartículas poliméricas de cloxacilina administradas em cabras Toggenburg naturalmente infectadas por *Corynebacterium pseudotuberculosis*. 2019. 44 p. Dissertação (Mestrado em Ciências Veterinárias) - Universidade Federal de Lavras, Lavras, 2019.
- ALMEIDA, S. et al. Exploration of Nitrate Reductase Metabolic Pathway in *Corynebacterium pseudotuberculosis*. **International Journal of Genomics**, v. 2017, p. 1–12, 2017.
- ALVES, F. S. F. **Immunokinetics of goats with *C. ovis* vaccination and infection**. 1988. Tese (Mestrado em Patologia Comparada). University of California, Davis, 1988.
- ALVES, F. S. F.; SANTIAGO, L. B.; PINHEIRO, R. R. Linfadenite Caseosa: o Estado da Arte. **Embrapa Caprinos**, p. 7–25, 2007
- BAIRD, G. J.; FONTAINE, M. C. *Corynebacterium pseudotuberculosis* and its Role in Ovine Caseous Lymphadenitis. **Journal of Comparative Pathology**, v. 137, n. 4, p. 179–210, nov. 2007.
- BARNABÉ, N. N. DA C. **AVALIAÇÃO DE PERDA ECONÔMICA POR LINFADENITE CASEOSA EM CAPRINOS ABATIDOS NO SEMIÁRIDO BRASILEIRO**. 2018. Tese (Mestrado). Programa de PósGraduação em Ciência Animal. Universidade Federal de Campina Grande, Patos, 2018. Disponível em: <http://dspace.sti.ufcg.edu.br:8080/jspui/handle/riufcg/8858>. Acesso em: 18/04/2024
- BARAÚNA, R. A. et al. Assessing the Genotypic Differences between Strains of *Corynebacterium pseudotuberculosis* biovar equi through Comparative Genomics. **PLOS ONE**, v. 12, n. 1, p. e0170676, 26 jan. 2017.
- BASTOS, B. L. et al. *Corynebacterium pseudotuberculosis*: Immunological Responses in Animal Models and Zoonotic Potential. **Journal of Clinical & Cellular Immunology**, v. 01, n. S4, 2012.
- BATEY, R. G. Pathogenesis of caseous lymphadenitis in sheep and goats. **Australian Velerinary Journal**, v. 63, p. 269–272, 1986.
- BIBI, S. et al. *In silico* analysis of epitope-based vaccine candidate against tuberculosis using reverse vaccinology. **Scientific Reports**, v. 11, n. 1, p. 1–16, 2021.
- BRUM, Alexandre Antunes; REZENDE, Andrea de Fatima Silva; BRILHANTE, Francisco Silvestre; *et al.* Recombinant esterase from *Corynebacterium pseudotuberculosis* in DNA and
-

subunit recombinant vaccines partially protects mice against challenge. **Journal of Medical Microbiology**, v. 66, n. 5, p. 635–642, 2017.

COSTA, L. F. M. *Corynebacterium pseudotuberculosis*, o agente etiológico da linfadenite caseosa em caprinos. **Revista de Ciências Médicas e Biológicas**, Salvador, v. 1, n. 1, p. 105-115, nov. 2002.

PINHO, Rodrigo Barros de; SILVA, Mara Thais de Oliveira; BEZERRA, Francisco Silvestre Brilhante; *et al.* Vaccines for caseous lymphadenitis: up-to-date and forward-looking strategies. **Applied Microbiology and Biotechnology**, v. 105, n. 6, p. 2287–2296, 2021.

DIMITROV, Ivan *et al.* AllerTOP v. 2—a server for *in silico* prediction of allergens. **Journal of molecular modeling**, v. 20, n. 6, p. 1-6, 2014.

DORELLA, F. A. *et al.* *Corynebacterium pseudotuberculosis*: microbiology, biochemical properties, pathogenesis and molecular studies of virulence. **Veterinary Research**, v. 37, n. 2, p. 201–218, mar. 2006.

DROPPA-ALMEIDA, D.; FRANCESCHI, E.; PADILHA, F. F. Immune-informatic analysis and design of peptide vaccine from multi-epitopes against *Corynebacterium pseudotuberculosis*. **Bioinformatics and Biology Insights**, v. 12, 2018

DROPPA-ALMEIDA, Daniela *et al.* USO DA IMUNOINFORMÁTICA PARA DESENHO VACINAL BASEADA EM PROTEÍNAS E EPÍTOPOS DE *CORYNEBACTERIUM PSEUDOTUBERCULOSIS*. **SEMPESq-Semana de Pesquisa da Unit-Alagoas**, n. 8, 2020.

DOYTCHINOVA, Irini A.; FLOWER, Darren R. VaxiJen: a server for prediction of protective antigens, tumour antigens and subunit vaccines. **BMC bioinformatics**, v. 8, n. 1, p. 1-7, 2007.

ELLIS, J. A. Immunophenotype of pulmonary cellular infiltrates in sheep with visceral caseous lymphadenitis. *Veterinary Pathology*, v. 25, n. 5, p. 362-368, 1988.

EMBRAPA, Caprinos e Ovinos. **Linfadenite Caseosa (LC)**. Portal Embrapa. Disponível em: <<https://www.embrapa.br/cim-inteligencia-e-mercado-de-caprinos-e-ovinos/zoossanitario-linfadenite>>. Acesso em: 19 Aug. 2022.

FACCIOLI-MARTINS, P. Y.; ALVES, F. S. F.; PINHEIRO, R. R. Linfadenite Caseosa: perspectivas no diagnóstico, tratamento e controle. Sobral: **EMBRAPA Caprinos e Ovinos**, 2014, 71 p. (Documentos, 113).

GUIMARÃES, A. S.; SEYFFERT, N.; GOUVEIA, A. M. G.; LAGE, A. P.; PORTELA, R. W. D.; MEYER, R.; AZEVEDO, V. A. C.; CARMO, F. B.; CRUZ, J. C. M.; HEINEMANN, M. B.

Linfadenite caseosa em rebanhos ovinos no estado de Minas Gerais, Brasil: Prevalência e informações de manejo. *Ciência Animal Brasileira*. Goiânia, v. 1, p. 597–602, 2009.

GUIMARÃES, A. D. S. *et al.* Lymphadenitis : Epidemiology, Diagnosis, and Control. **HIOAB J.**, v. 2, n. 2, p. 33–43, 2011.

FIGUEIREDO, E. A. P.; SHELTON, M.; PANT, K. P. Goats skins. In: INTERNATIONAL CONFERENCE ON GOAT PRODUCTION AND DISEASE, **Proceedings - Scottsdale: Dairy Goat Journal**, v. 3 p.488-490. Tucson. 1982.

GASTEIGER, Elisabeth et al. Protein identification and analysis tools on the ExPASy server. **The proteomics protocols handbook**, p. 571-607, 2005.

GUPTA, Sudheer et al. Peptide toxicity prediction. In: **Computational peptidology**. Humana Press, New York, NY, 2015. p. 143-157.

HATTOTUWAGAMA, Channa K. et al. Quantitative online prediction of peptide binding to the major histocompatibility complex. *Journal of Molecular Graphics and Modelling*, v. 22, n. 3, p. 195-207, 2004.

IONEDA, T. Chromatographic and mass spectrometric characterization of 3 -O-benzoylmethyl ester derivates of mycolic acid fractions from *Corynebacterium pseudotuberculosis*, *C.diphtheriae* and *Rhodococcus rhodochrous*. **Chem. Phys. Lipids.**, n. 65, p. 93-101, 1993.

JOLLY, R. D. Some observations on surface lipids of virulent and attenuated strains of *Corynebacterium ovis*. **J. Appl. Bact.**, v. 29, n. 1, p. 189 -196, 1966.

KRINGELUM, Jens Vindahl et al. Reliable B cell epitope predictions: impacts of method development and improved benchmarking. **PLoS computational biology**. v. 8, n. 12, p. e1002829, 2012.

LASKOWSKI R A, MacArthur M W, Moss D S, Thornton J M (1993). PROCHECK - a program to check the stereochemical quality of protein structures. *J. App. Cryst.*, 26, 283-291.

LATIF, N. A. A.; ABBA, Y.; JESSE, F. F. A.; CHUNG, E. L. T.; ZAMRI-SAA, M.; SAHAREE, A. A.; ZAKARIA, Z.; HARON, A. W.; MOHD-LILA, M. A.i. Histopathological assessment of chronic *Corynebacterium pseudotuberculosis* infection in the reproductive tract and iliac lymph node of Katjang does. **Comparative Clinical Pathology**, v. 26, n. 1, p. 147–154. 2017. doi:10.1007/s00580-016-2357-3.

MA, Lifei; LI, Huiyang; LAN, Jinping; HAO, Xiuqing; LIU, Huiying; WANG, Xiaoman; HUANG, Yong. Comprehensive analyses of bioinformatics applications in the fight against COVID-19 pandemic. *Computational Biology And Chemistry*, [S.L.], v. 95, n. 0, p. 107599, dez. 2021. Elsevier BV.
<http://dx.doi.org/10.1016/j.compbiolchem.2021.107599>

MACEDO JUNIOR, Lafayette Modesto de; MELO, Tuane Ferreira; PECONICK, Ana Paula. Predição *in silico* de epítomos antigênicos para produção de uma vacina humana contra leishmaniose visceral. **Scire Salutis**, v. 9, n. 1, p. 62–71, 2019.

MARQUES, Andressa Souza; BASTOS, Bruno Lopes; RAYNAL FILHO, José Tadeu; *et al.* Identificação *in silico* de potenciais alvos antigênicos de *Corynebacterium pseudotuberculosis*. **Pubvet**, v. 13, n. 8, p. 1–7, 2019.

MANDLIK, A. et al. Pili in Gram-positive bacteria: assembly, involvement in colonization and biofilm development.

Trends in microbiology. v. 16, n. 1, p. 33–40, 2008.

MARTINS, A. M.; CARNEIRO, A. S.; DINIZ, L. G.; DALL'ACQUA, P. C.; BEZZERRIL, J. E.; SOUSA, F. E.;

MCNAMARA, P. J.; CUEVAS, W. A.; SONGER, J. G. Toxic phospholipases D of *Corynebacterium pseudotuberculosis*, *C. ulcerans* and *Arcanobacterium haemolyticum*: cloning and sequence homology. **Gene**, n. 156, p. 113 -118, 1995.

MCKEAN, S. C.; DAVIES, J. K.; MOORE, R. J. Expression of phospholipase D, the major virulence factor of *Corynebacterium pseudotuberculosis*, is regulated by multiple 33 environmental factors and plays a role in macrophage death. **Microbiology**. v. 153, n. 7, p. 2203–2211, 2007.

MEI, X. et al. Computer-Based Immunoinformatic Analysis to Predict Candidate T-Cell Epitopes for SARS-CoV-2 Vaccine Design. **Frontiers in Immunology**, v. 13, 30 mar. 2022.

MEI, Xueyin; GU, Pan; SHEN, Chuanlai; *et al.* Computer-Based Immunoinformatic Analysis to Predict Candidate TCell Epitopes for SARS-CoV-2 Vaccine Design. **Frontiers in Immunology**, v. 13, 2022.

MONTERRUBIO-LÓPEZ, G. P.; GONZÁLEZ-Y-MERCHAND, J. A.; RIBASAPARICIO, R. M. Identification of novel potential vaccine candidates against tuberculosis based on reverse vaccinology. *BioMed Research International*, v. 2015, p. 11–14, 2015.

MOUSSA, I. M. et al. Vaccination against *Corynebacterium pseudotuberculosis* infections controlling caseous lymphadenitis (CLA) and oedematousskin disease. **Saudi Journal of Biological Sciences**, v. 23, n. 6, p. 718–723. 2016.

MICHEL-TODÓ, Lucas; RECHE, Pedro Antonio; BIGEY, Pascal; *et al.* *In silico* Design of an Epitope-Based Vaccine Ensemble for Chagas Disease. **Frontiers in Immunology**, v. 10, 2019. MIRDITA, M., Schütze, K., Moriwaki, Y. et al. ColabFold: making protein folding accessible to all. *Nat Methods* 19, 679–682 (2022). <https://doi.org/10.1038/s41592-022-01488-1>

MURPHY, J. R. Mechanism of Diphtheria Toxin Catalytic Domain Delivery to the Eukaryotic Cell Cytosol and the Cellular Factors that Directly Participate in the Process. **Toxins**. v. 3, n. 12, p. 294–308. 2011.

NAVARINI, Franciele C. et al. Conforto térmico de bovinos da raça nelore a pasto sob diferentes condições de sombreamento e a pleno sol. **Engenharia Agrícola**, v. 29, p. 508-517, 2009.

NAZ, Anam et al. Reverse vaccinology and drug target identification through pan-genomics. In: **Pan-genomics: Applications, challenges, and future prospects**. Academic Press, 2020. p. 317-333. Disponível em: <https://doi.org/10.1016/B978-0-12-817076-2.00016-0> Acesso em: 23/04/2024

NCBI. **corynebacterium pseudotuberculosis**[orgn] - Gene. Disponível em: <<https://www.ncbi.nlm.nih.gov/gene/?term=corynebacterium+pseudotuberculosis%5Borgn%5D>>. Acesso em: 21 Aug. 2022.

PAULA, E. M. N.; MARTINI, A. C. Linfadenite caseosa intestinal em ovino. **Acta Scientiae Veterinariae**. v. 49, p. 1-4, 2021. DOI: 10.22456/1679-9216.105839.

PAULE, Bruno Jean Adrien et al. Experimental *Corynebacterium pseudotuberculosis* primary infection in goats: kinetics of IgG and interferon- γ production, IgG avidity and antigen recognition by Western blotting. **Veterinary immunology and immunopathology**, v. 96, n. 3- 4, p. 129-139, 2003.

PONOMARENKO, Julia et al. ElliPro: a new structure-based tool for the prediction of antibody epitopes. **BMC bioinformatics**, v. 9, n. 1, p. 1-8, 2008.

PROTEIN DATA BANK. *Corynebacterium pseudotuberculosis* Pdb. Disponível em: <encurtador.com.br/anpX8>. Acesso em: 21 Aug. 2022.

PÉPIN, Michel et al. Cellular composition of *Corynebacterium pseudotuberculosis* pyogranulomas in sheep. **Journal of Leukocyte Biology**, v. 56, n. 5, p. 666-670, 1994.

PÉPIN M., SEOWN H.F., CORNER L., ROTHEN J.S., Hodgson A.L., et al. Cytokine gene expression in sheep following experimental infection with various strains of *Corynebacterium pseudotuberculosis* differing in virulence. **Vet Res** 28. p. 149-163. 1997

PIZZA, M. et al. Identification of vaccine candidates against serogroup B meningococcus by whole-genome sequencing. *Science*, v. 287, n. 5459, p. 1816– 1820, 2000

RAHMAN, M. S. et al. Epitope-based chimeric peptide vaccine design against S, M and e proteins of SARS-CoV-2, the etiologic agent of COVID-19 pandemic: An *in silico* approach. *PeerJ*, v. 8, 2020.

RAOUFI, E. et al. Epitope Prediction by Novel Immunoinformatics Approach: A State-of-the-art Review. International. **Journal of Peptide Research and Therapeutics** Springer, , 1 jun. 2020

REYNISSON, Birkir et al. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. **Nucleic acids research**, v. 48, n. W1, p. W449-W454, 2020.

RIET-CORREA, F., Rivero, B. R. C., & Correa, G. R. Plantas que causam alterações mecânicas ou traumáticas em ruminantes e equinos, com ênfase em *Stipa* spp.(Gramineae). **Pesquisa Veterinária Brasileira**, v. 31, p. 516-520, 2011.

RIBEIRO, S. D. A.. Caprinocultura: Criação Racional de Caprinos. **São Paulo: Nobel**. 318p. 1997

RODRIGUES, T. C. V. et al. Reverse vaccinology and subtractive genomics reveal new therapeutic targets against *Mycoplasma pneumoniae*: A causative agent of pneumonia. *Royal Society Open Science*, v. 6, n. 7, 2019

RODRÍGUEZ-DOMÍNGUEZ, Maria Carla et al. Bioinformatic Approach of B and T Cell Epitopes of PLD and CP40 Proteins of *Corynebacterium pseudotuberculosis* ovis Mexican Isolate 2J-L towards a Peptide-Based Vaccine. **International Journal of Molecular Sciences**, v. 25, n. 1, p. 270, 2023.

SÁ, M. DA C. A. DE et al. Linfadenite caseosa em caprinos e ovinos: Revisão. **Pubvet**, v. 12, n. 11, p. 1–13, nov. 2018.

SANTOS, E. M. S.; SANTOS, H. O.; CANGUSSU, A. R.; COSTA, K. S.; DIAS, I. S. Antigens of *Corynebacterium pseudotuberculosis* with promising potential for caseous lymphadenitis vaccine development: a literature review. *Caderno de Ciências Agrárias*. v. 8, n. 2, p. 90-99, 2016

SANTOS, F. R. F.; SANTANA, E. S.; DROPPA-ALMEIDA, D. Desenho vacinal por imunoinformática contra HCV: Caracterização química e predição de epítomos de células T. **Research, Society and Development**, v. 10, n. 16, 6 dez. 2021.

SILVA, J. M. DA et al. EVOLUÇÃO DO REBANHO EFETIVO E PRINCIPAIS CARACTERÍSTICAS DA OVINOCAPRINOCULTURA NO ESTADO DO PARÁ. **International Journal Education and Teaching**, 2 fev. 2021a.

SILVA, M. T. DE O. et al. *In silico* analyses and design of a chimeric protein containing epitopes of SpaC, PknG, NanH, and SodC proteins for the control of caseous lymphadenitis. **Applied Microbiology and Biotechnology**. v. 105, n. 21–22, p. 8277–8286, 1 nov. 2021b.

SOLANKI, V.; TIWARI, V. Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against *Acinetobacter baumannii*. *Scientific Reports*, v. 8, n. 1, p. 1–19, 2018

SOLANKI, V.; TIWARI, M.; TIWARI, V. Prioritization of potential vaccine targets using comparative proteomics and designing of the chimeric multi-epitope vaccine against *Pseudomonas aeruginosa*. *Scientific Reports*, v. 9, n. 1, p. 1–19, 2019

SONGER, J. G. Cloning and expression of the phospholipase D gene of *Corynebacterium pseudotuberculosis* in *Escherichia coli*. **Infect. Immun.**, v. 58, n. 1, p. 131 - 136, 1990.

SONGER, J. G. et al. Biochemical and genetic characterization of *Corynebacterium pseudotuberculosis*. *American Journal of Veterinary Research*, v. 49, n. 2, p. 223–226, 1988.

SOUZA, M.F.; CARVALHO, A.Q.; GARINO JR, F.; et al. Linfadenite caseosa em ovinos deslanados abatidos em um frigorífico da Paraíba. **Pesquisa Veterinária Brasileira**, v. 31, n. 3, p. 224–230, 2011. STEFAŃSKA, I. et al. Survival of *Corynebacterium pseudotuberculosis* within macrophages and induction of phagocytes death. **Polish Journal of Veterinary Sciences**, v. 13, n. 1, p. 143–149, 2010.

TAUCH, A.; BURKOVSKI, A. Molecular armory or niche factors: virulence determinants of *Corynebacterium* species. *FEMS Microbiology Letters*, v. 67, n. 2. Out. 2015.

TORCHALA, Mieczyslaw et al. SwarmDock: a server for flexible protein–protein docking. **Bioinformatics**, v. 29, n. 6, p. 807-809, 2013.

TORRES, L. DE F. C. Estudos das características fenotípicas e genotípicas de amostras de *Corynebacterium pseudotuberculosis*. Universidade do Estado do Rio de Janeiro. Tese (Mestrado). Rio de Janeiro. 2011.

TROST, E. et al. The complete genome sequence of *Corynebacterium pseudotuberculosis* FRC41 isolated from a 12year-old girl with necrotizing lymphadenitis reveals insights into gene-regulatory networks contributing to virulence. **BMC genomics**, v. 11, n. 1, p. 728, 2010.

UNANIAN, M. M.; FELICIANO SILVA, A. E. D.; PANT, K. P. Abscesses and caseous lymphadenitis in goats in tropical semi-arid North-east Brazil. *Tropical Animal Health and Production*, v. 17, n. 1, p. 57-62, Fev. 1985.

UNIPROT. **Corynebacterium pseudotuberculosis**. UniProtKB. Disponível em: <<https://www.uniprot.org/uniprotkb?query=Corynebacterium%20pseudotuberculosis>>. Acesso em: 21 Aug. 2022. VITA, Randi et al. The immune epitope database (IEDB): 2018 update. **Nucleic acids research**, v. 47, n. D1, p. D339D343, 2019.

WEBB, Benjamin; SALI, Andrej. Comparative protein structure modeling using MODELLER. **Current protocols in bioinformatics**, v. 54, n. 1, p. 5.6. 1-5.6. 37, 2016.

WATERHOUSE, Andrew et al. SWISS-MODEL: homology modelling of protein structures and complexes. **Nucleic acids research**, v. 46, n. W1, p. W296-W303, 2018.

YERUHAM, I. et al. A herd level analysis of a *Corynebacterium pseudotuberculosis* outbreak in a dairy cattle herd. **Veterinary Dermatology**, v. 15, n. 5, p. 315–320, 2004.

YURINA, V.; ADIANINGSIH, O. R. Predicting epitopes for vaccine development using bioinformatics tools. **Therapeutic Advances in Vaccines and Immunotherapy**, v. 10, 2022.

ANEXO A – TABELA COM OS DADOS FÍSICO-QUÍMICOS DO EPÍTOPOS COLETADOS DOS SOFTWARES DE TRIAGEM

Protein ID	Molecular weight	Theoretical pI	Enrichment factor	ProteinM13			Vachin13			Ionized pI	Ionized MW	Ionized pI	Ionized MW
				Estimated	Intensity	Aliphatic index	Score	Definition	Score				
AA0401E1_86256	1640.54	3.27	...	1100s (td)	85.66	32.67	-1.547						
DQ04A_1921a	1320.42	4.37	...	100 hours	25.67	75.38	-0.031	0.1531					
DQ04T_16300	1420.45	3.57	5500	>20 hours	7.21	42.14	-0.486	0.5717	non-allergen Non-Toxic				
DQ05C_3035c	1394.57	4.20	2980	10 min (f)	7.74	0.00	-1.760	0.6647	allergen				
DQ05T_3013c	1387.47	4.37	...	ns/3 min	46.43	37.69	-1.369						
DQ05T_3067	1072.20	5.84	...	>20 hours	2.24	53.64	-0.218	1.4791	non-allergen Non-Toxic				
DQ06C_978b	1107.19	8.75	1490	ns/10 min	12.59	62.73	-0.236	0.5731	non-allergen Non-Toxic				
DQ06C_3c14	1380.62	12.30	...	ns/2 min	25.51	49.17	-1.458	0.4188					
DQ06T_8004d	1908.05	4.39	...	ns/3 min	32.59	52.00	-2.253	0.3456					
DQ06T_8008	1219.56	6.00	...	ns/3 min	5.18	0.469	0.6063	allergen					
DQ06T_768c*	1417.67	9.53	1490	>20 hours	15.50	60.00	-0.854	0.6230	non-allergen Non-Toxic				
DQ06T_1a18c	1324.51	5.24	1490	>20 hours	9.93	40.83	-0.083						
DQ06T_8965	964.05	5.97	...	>20 hours	-8.92	43.33	-0.467	1.6287	allergen				
DQ06T_05a1	1482.56	12.00	...	ns/3 min (f)	105.64	37.69	-1.462						
DQ06T_4a55	1305.43	5.95	5500	>20 hours	84.76	26.56	-1.236						
DQ06T_2753b	957.09	5.52	1490	ns/10 min	14.31	75.56	-0.311						
DQ06T_871a	1197.31	8.60	...	ns/3 min	71.41	44.55	-1.427						
DQ06T_2635c	1324.43	9.35	...	>20 hours	122.17	32.50	-1.217						
DQ06T_06c2	...	...	...	...	...	...	...	...	...	...	...	...	...
DQ06T_961a	1572.51	11.52	...	ns/2 min (f)	94.25	35.45	-2.273						
DQ06T_017d	1576.75	8.20	5500	>20 hours	42.94	27.86	-1.379						
DQ06T_859a	1359.61	4.03	...	>20 hours	70.82	39.33	-0.840						
DQ06T_5802	971.14	5.99	0	ns/2 min (f)	43.31	86.67	0.111						
DQ06T_58025	1164.29	4.14	...	>20 hours	89.58	39.00	-1.240						
DQ06T_8e1a	...	...	...	...	...	...	...	...	...	...	...	...	...
DQ06T_6c70b	897.00	5.55	...	19 hours	91.08	76.67	-0.311						
DQ06T_97191	913.99	6.05	...	>20 hours	88.99	65.56	-0.689						
DQ06T_4e2e	...	...	...	...	...	...	...	...	...	...	...	...	...
DQ06T_829a	1329.34	3.93	...	>20 hours	-12.19	70.00	-0.050	1.1067	allergen				
DQ06T_40196	1611.73	6.18	...	ns/3 min (f)	21.49	32.67	-1.933	0.8321	non-allergen Non-Toxic				
DQ06T_40196	1084.06	3.49	1490	>20 hours	47.52	49.00	-0.910						
DQ06T_4d45	1370.48	5.49	...	>20 hours	57.18	48.57	0.293						
DQ06T_17d4	1741.93	11.54	1490	ns/2 min (f)	181.23	26.00	-1.807						
DQ06T_6738c	1312.44	3.80	...	ns/10 min	110.87	96.92	-0.008						
DQ06T_31922	...	...	...	...	...	...	...	...	...	...	...	...	...
DQ06T_19404	1209.30	8.22	0	>20 hours	108.31	9.09	-1.664						
DQ06T_00015	...	...	...	...	...	...	...	...	...	...	...	...	...
Protein ID	Molecular weight	Theoretical pI	Enrichment factor	ProteinM13			Vachin13			Ionized pI	Ionized MW	Ionized pI	Ionized MW
				Estimated	Intensity	Aliphatic index	Score	Definition	Score				
AA0401E1_86256	1640.54	3.27	...	1100s (td)	85.66	32.67	-1.547						
DQ04A_1921a	1320.42	4.37	...	100 hours	25.67	75.38	-0.031	0.1531					
DQ04T_16300	1420.45	3.57	5500	>20 hours	7.21	42.14	-0.486	0.5717	non-allergen Non-Toxic				
DQ05C_3035c	1394.57	4.20	2980	10 min (f)	7.74	0.00	-1.760	0.6647	allergen				
DQ05T_3013c	1387.47	4.37	...	ns/3 min	46.43	37.69	-1.369						
DQ05T_3067	1072.20	5.84	...	>20 hours	2.24	53.64	-0.218	1.4791	non-allergen Non-Toxic				
DQ06C_978b	1107.19	8.75	1490	ns/10 min	12.59	62.73	-0.236	0.5731	non-allergen Non-Toxic				
DQ06C_3c14	1380.62	12.30	...	ns/2 min	25.51	49.17	-1.458	0.4188					
DQ06T_8004d	1908.05	4.39	...	ns/3 min	32.59	52.00	-2.253	0.3456					
DQ06T_8008	1219.56	6.00	...	ns/3 min	5.18	0.469	0.6063	allergen					
DQ06T_768c*	1417.67	9.53	1490	>20 hours	15.50	60.00	-0.854	0.6230	non-allergen Non-Toxic				
DQ06T_1a18c	1324.51	5.24	1490	>20 hours	9.93	40.83	-0.083						
DQ06T_8965	964.05	5.97	...	>20 hours	-8.92	43.33	-0.467	1.6287	allergen				
DQ06T_05a1	1482.56	12.00	...	ns/3 min (f)	105.64	37.69	-1.462						
DQ06T_4a55	1305.43	5.95	5500	>20 hours	84.76	26.56	-1.236						
DQ06T_2753b	957.09	5.52	1490	ns/10 min	14.31	75.56	-0.311						
DQ06T_871a	1197.31	8.60	...	ns/3 min	71.41	44.55	-1.427						
DQ06T_2635c	1324.43	9.35	...	>20 hours	122.17	32.50	-1.217						
DQ06T_06c2	...	...	...	...	...	...	...	...	...	...	...	...	...
DQ06T_961a	1572.51	11.52	...	ns/2 min (f)	94.25	35.45	-2.273						
DQ06T_017d	1576.75	8.20	5500	>20 hours	42.94	27.86	-1.379						
DQ06T_859a	1359.61	4.03	...	>20 hours	70.82	39.33	-0.840						
DQ06T_5802	971.14	5.99	0	ns/2 min (f)	43.31	86.67	0.111						
DQ06T_58025	1164.29	4.14	...	>20 hours	89.58	39.00	-1.240						
DQ06T_8e1a	...	...	...	...	...	...	...	...	...	...	...	...	...
DQ06T_6c70b	897.00	5.55	...	>20 hours	92.15	9.80	-	106.53	108.89	-0.111			
DQ06T_97191	913.99	6.05	...	ns/3 min	66.51	141.11	0.778						
DQ06T_4e2e	...	...	...	>20 hours	75.09	173.33	1.000						
DQ06T_829a	1329.34	3.93	...	>20 hours	39.06	97.78	-0.378	-0.3466					
DQ06T_40196	1611.73	6.18	...	ns/2 min	30.29	184.44	1.089	0.0945					
DQ06T_40196	1084.06	3.49	1490	>20 hours	100.93	130.00	0.678						
DQ06T_4d45	1370.48	5.49	...	ns/2 min	30.29	87.78	-0.233	0.0926	allergen				
DQ06T_17d4	1741.93	11.54	1490	ns/2 min (f)	11.50	130.00	-0.100	1.7826	allergen				
DQ06T_6738c	1312.44	3.80	...	ns/2 min	13.17	97.78	1.078	0.6591	allergen				
DQ06T_31922	...	...	...	>20 hours	8.89	141.11	1.478	-0.0950					
DQ06T_19404	1209.30	8.22	0	>20 hours	70.32	184.44	2.011	1.759	non-allergen Non-Toxic				
DQ06T_00015	...	...	...	...	...	...	...	...	...	...	...	...	...