UNIVERSIDADE FEDERAL DE MINAS GERAIS INSTITUTO DE CIÊNCIAS BIOLÓGICAS, DEPARTAMENTO DE GENÉTICA, ECOLOGIA E EVOLUÇÃO, PROGRAMA DE PÓS-GRADUAÇÃO EM GENÉTICA

DISSERTAÇÃO

Abordagens de genômica subtrativa, vacinologia reversa e imunoinformática para predição de drogas e imunógenos contra *Mycoplasma pneumoniae*

ORIENTADO: THAÍS CRISTINA VILELA RODRIGUES

ORIENTADOR: VASCO ARISTON DE CARVALHO AZEVEDO

CO-ORIENTADORES: SIOMAR DE CASTRO SOARES E SANDEEP TIWARI

BELO HORIZONTE

MAIO – 2021

THAÍS CRISTINA VILELA RODRIGUES

Abordagens de genômica subtrativa, vacinologia reversa e imunoinformática para predição de drogas e imunógenos contra *Mycoplasma pneumoniae*

Dissertação apresentada ao Programa de Pós-graduação em Genética, do Instituto de Ciências Biológicas da Universidade Federal de Minas Gerais, como parte dos requerimentos para obtenção do título de mestre.

ORIENTADOR: DR. VASCO ARISTON DE CARVALHO AZEVEDO
COORIENTADORES: DR. SIOMAR DE CASTRO SOARES E DR. SANDEEP TIWARI

BELO HORIZONTE

MAIO – 2021

043 Rodrigues, Thaís Cristina Vilela.
Abordagens de genômica subtrativa, vacinologia reversa e imunoinformática para predição de drogas e imunógenos contra Mycoplasma pneumoniae [manuscrito] / Thaís Cristina Vilela Rodrigues. – 2021.
136 f. : il. ; 29,5 cm.

Orientador: Dr. Vasco Ariston de Carvalho Azevedo. Coorientadores: Dr. Siomar de Castro Soares e Dr. Sandeep Tiwari.
Dissertação (mestrado) – Universidade Federal de Minas Gerais, Instituto de Ciências Biológicas. Programa de Pós-Graduação em Genética.

1. Genética. 2. Pneumonia. 3. Mycoplasma pneumoniae. 4. Vacinas. 5. Epitopos Imunodominantes. I. Azevedo, Vasco Ariston de Carvalho. II. Soares, Siomar de Castro. III. Tiwari, Sandeep. IV. Universidade Federal de Minas Gerais. Instituto de Ciências Biológicas. V. Título.

CDU: 575

Ficha catalográfica elaborada por Fabiane C. M. Reis – CRB: 6/2680

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Instituto de Ciências Biológicas

Programa de Pós-Graduação em Genética

ATA DE DEFESA DE DISSERTAÇÃO / TESE

ATA DA DEFESA DE DISSERTAÇÃO	**308/2021** **entrada** **2º/2019** **CPF: 126.791.666-47**
THAÍS CRISTINA VILELA RODRIGUES	

Às treze horas do dia **04 de maio de 2021**, reuniu-se remotamente, devido ao isolamento social em razão da pandemia de COVID-19, a Comissão Examinadora de Dissertação, indicada pelo Colegiado do Programa, para julgar, em exame final, o trabalho intitulado: **"Abordagens de genômica subtrativa, vacinologia reversa e imunoinformática para predição de drogas e imunógenos contra Mycoplasma pneumoniae",** requisito para obtenção do grau de Mestre em **Genética.** Abrindo a sessão, o Presidente da Comissão, **Vasco Ariston de Carvalho Azevedo**, 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 de resultado final. Foram atribuídas as seguintes indicações:

Prof./Pesq.	**Instituição**	**CPF**	**Indicação**
Vasco Ariston de Carvalho Azevedo	UFMG	283.171.225-49	APROVADA
Siomar de Castro Soares	UFTM	056.951.826-11	APROVADA
Sandeep Tiwari	UFMG	020.977.186-04	APROVADA
Francisco Pereira Lobo	UFMG	012.273.736-94	APROVADA
Bruno Silva Andrade	UESB	001.755.685-61	APROVADA
Mateus Matiuzzi da Costa	UNIVASF	803.979.560-53	APROVADA

Pelas indicações, a candidata foi considerada: APROVADA

O resultado final foi comunicado publicamente à candidata pelo Presidente da Comissão. Nada mais havendo a tratar, o Presidente encerrou a reunião e lavrou a presente ATA, que será assinada por todos os membros participantes da Comissão Examinadora.

Belo Horizonte, 04 de maio de 2021.

Vasco Ariston de Carvalho Azevedo – Orientador (UFMG)

Siomar de Castro Soares – Coorientador (UFTM)

Sandeep Tiwari– Coorientador (UFMG)

Francisco Pereira Lobo (UFMG)Bruno Silva Andrade (UESB)

Matheus Matiuzzi da Costa (UNIVASF)

Assinatura dos membros da banca examinadora:

Documento assinado eletronicamente por **Vasco Ariston de Carvalho Azevedo**, **Professor do Magistério Superior**, em 05/05/2021, às 10:42, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Siomar de Castro Soares**, **Usuário Externo**, em 05/05/2021, às 10:48, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Bruno Silva Andrade**, **Usuário Externo**, em 05/05/2021, às 12:39, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Mateus Matiuzzi da Costa**, **Usuário Externo**, em 05/05/2021, às 13:48, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Sandeep Tiwari**, **Usuário Externo**, em 05/05/2021, às 16:21, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Francisco Pereira Lobo**, **Professor do Magistério Superior**, em 07/05/2021, às 11:01, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0704039** e o código CRC **067F9765**.

UNIVERSIDADE FEDERAL DE MINAS GERAIS

Instituto de Ciências Biológicas
Programa de Pós-Graduação em Genética

FOLHA DE APROVAÇÃO

"Abordagens de genômica subtrativa, vacinologia reversa e imunoinformática para predição de drogas e imunógenos contra Mycoplasma pneumoniae"

THAÍS CRISTINA VILELA RODRIGUES

Dissertação aprovada pela banca examinadora constituída pelos Professores:

Vasco Ariston de Carvalho Azevedo - Orientador
Universidade Federal de Minas Gerais

Siomar de Castro Soares - Coorientador
Universidade Federal do Triângulo Mineiro

Sandeep Tiwari
Universidade Federal de Minas Gerais

Francisco Pereira Lobo
Universidade Federal de Minas Gerais

Bruno Silva Andrade
Universidade Estadual do Sudoeste da Bahia

Mateus Matiuzzi da Costa Universidade Federal do Vale do São Francisco

Belo Horizonte, 04 de maio de 2021.

Documento assinado eletronicamente por **Vasco Ariston de Carvalho Azevedo**, **Professor doMagistério Superior**, em 05/05/2021, às 10:42, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Siomar de Castro Soares**, **Usuário Externo**, em 05/05/2021, às 10:48, conforme horário oficial de Brasília, com fundamento no art. 5º do Decretonº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Mateus Matiuzzi da Costa**, **Usuário Externo**, em 05/05/2021, às 11:25, conforme horário oficial de Brasília, com fundamento no art. 5º do Decretonº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Bruno Silva Andrade**, **Usuário Externo**, em 05/05/2021, às 12:39, conforme horário oficial de Brasília, com fundamento no art. 5º do Decretonº 10.543, de 13 de novembro de 2020.

Documento assinado eletronicamente por **Sandeep Tiwari**, **Usuário Externo**, em 05/05/2021, às16:21, conforme horário oficial de Brasília, com fundamento no art. 5º do Decreto nº 10.543, de13 de novembro de 2020.

Documento assinado eletronicamente por **Francisco Pereira Lobo**, **Professor do Magistério Superior**, em 07/05/2021, às 11:01, conforme horário oficial de Brasília, com fundamento no art.5º do Decreto nº 10.543, de 13 de novembro de 2020.

A autenticidade deste documento pode ser conferida no site https://sei.ufmg.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0704065** e o código CRC **4B8891A1**.

Referência: Processo nº 23072.222439/2021-86

Dedico esse trabalho aos meu pais, por todo apoio e trabalho dedicado à nossa família. Dedico também aos meus orientadores, pelo suporte e a todos pesquisadores brasileiros que independente das adversidades sustentam à ciência no nosso país.

AGRADECIMENTOS

Gostaria de agradecer a todos que contribuíram de alguma forma para minha formação e conclusão da minha dissertação.

Primeiramente agradeço à Deus por guiar e direcionar meus caminhos. Agraço aos meus pais, Fábio e Marta por tudo suporte durante toda minha vida, com apoio incondicional à nossa família. Ao meu irmão Lucas, que sempre está do meu lado e foi essencial nessa jornada. Agradeço a todos meus familiares pelo carinho e conforto.

Agradeço ao meu namorado Gabriel por sempre acreditar no meu potencial e me dar forças nos momentos mais difíceis. Agradeço a todos meus amigos, pela parceria e suporte.

Ao professor Dr. Siomar, agradeço a oportunidade de me aceitar em seu laboratório, pelo apoio e conhecimento a mim concedidos durante minha formação, desde a graduação. Ao prof. Dr. Vasco, agradeço por toda orientação e esforço para construir esse trabalho e por toda dedicação à equipe do LGCM. Gostaria de agradecer também ao Sandeep, pelo exemplo de postura científica e humanidade, e pelo trabalho realizado para conclusão desse processo.

Agradeço a todos do Laboratório de Genética Celular e Molecular, amigos que foram essenciais para meu crescimento pessoal e por toda experiência obtida no laboratório. Agradeço também a Pós-Graduação em Genética, todos professores e secretários, por todo conhecimento e estrutura organizados para fazer possível alcançar essa meta.

Por fim, agradeço o apoio dos órgãos de fomento, a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), a Fundação de Amparo à Pesquisa do Estado de Minas (FAPEMIG) e ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) que financiaram este projeto e sustentam a ciência brasileira.

A todos vocês, muito obrigada.

SUMÁRIO

LISTA DE FIGURAS ... 12
LISTA DE TABELAS ... 13
LISTA DE SIGLAS E ABREVIATURAS ... 14
RESUMO ... 17
ABSTRACT ... 18
I.PREFÁCIO ... 19
I.I Colaboradores ... 19
I.II Introdução geral ... 19
I.III Estrutura do manuscrito ... 20
1. INTRODUÇÃO ... 22
1.1MYCOPLASMA PNEUMONIAE ... 22
1.1.1 Mecanismos de patogenicidade ... 23
1.1.2 Resposta imune ao *M.pneumoniae* ... 24
1.1.3 *M.pneumoniae* e a Pneumonia Adquirida na Comunidade (PAC) ... 24
1.2. PNEUMONIA ... 26
1.2.1 Pneumonia adquirida em hospital ... 27
1.2.2 Pneumonia associada à ventilação ... 27
1.2.3 Pneumonia por aspiração ... 27
1.3 PNEUMONIA ADQUIRIDA NA COMUNIDADE ... 28
1.3.1 Agentes Etiológicos ... 28
1.3.2 Fatores de risco e manifestações clínicas ... 29
1.3.3 Diagnóstico ... 29
1.3.4 Abordagem terapêutica ... 30
1.4 VACINAS ... 31
1.4.1 Vacinologia Reversa ... 33
1.4.2 Imunoinformática ... 35
1.4.2 Imunologia e resposta a vacinas multi-epítopo ... 36
2. JUSTIFICATIVA ... 39
3. OBJETIVOS ... 40
3.1 Objetivo principal ... 40
3.2 Objetivos específicos ... 40
4. CAPITULO I ... 41
Reverse vaccinology and subtractive genomics reveal new therapeutic targets against Mycoplasma pneumoniae: a causative agent of pneumonia ... 41
CONCLUSÃO CAPÍTULO I ... 61
6. CAPÍTULO II ... 62
Prediction of multi-epitope vaccine for *Mycoplasma pneumoniae* infection applying immunoinformatics approaches ... 62
7. CONCLUSÃO CAPÍTULO II ... 97
8. DISCUSSÃO ... 98
9. CONCLUSÕES ... 101

10. REFERÊNCIAS 102
11. ANEXOS 113
11.1 MATERIAL SUPLEMENTAR REFERENTE AO CAPÍTULO I: 113
11.2 MATERIAL SUPLEMENTAR REFERENTE AO CAPÍTULO II: 120
11.3 PRODUÇÃO ACADÊMICA, *CURRICULUM VITAE* 128

LISTA DE FIGURAS

Figura 1. Componentes proteicos importantes na citoaderência e motilidade do *M.pneumoniae*.

Figura 2. Antimicrobianos utilizados no tratamento à infecções por *M.pneumoniae*

Figura 3. Diferentes tipos de vacinas

Figura 4. Vacinologia Reversa

Figura 5. Workflow projected with the methodologies used and the total number of proteins identified in each step

Figura 6. The (a) 3D and (b) 2D representation docking analyses of protein 30S ribosome-binding factor (WP_010874513.1) with compound ZINC04259381

Figura 7. The (a) 3D and (b) 2D representation docking analyses of protein division/cell-wall cluster transcriptional repressor MraZ (WP_010874670.1) with compound ZINC04235924.

Figura 8. The (a) 3D and (b) 2D representation docking analyses of protein dTIGR00282 family metallophosphoesterase (WP_010874705.1) with compound ZINC04259703.

Figura 9. The (a) 3D and (b) 2D representation docking analyses of hypothetical protein (WP_010874779.1) with compound ZINC05415832.

Figura 10. The (a) 3D and (b) 2D representation docking analyses of hypothetical protein (WP_014325598.1) with compound ZINC04236030.

Figura 11. Pathogenicity islands predicted using 15 strains of *M.pneumoniae* having *M.pneumoniae* C267 as reference and *M. Gallinarum* as non-pathogenic reference.

Figura 12. The overall workflow of the Methodology used in this study.

Figura 13. Multi-epitope vaccine design and structure prediction.

Figura 14. *Mycoplasma pneumoniae* multi-epitope vaccine tertiary structure with the location of B-cell conformational epitopes.

Figura 15. Molecular docking of the TLR2multi-epitope vaccine.

Figura 16. Molecular dynamics simulation plots of the ligand-receptor complex

Figura 17. *In silico* cloning through SnapGene (+) software

Figura 18. Immuno simulation results regarding innate immunity and cytokine production.

Figura 19. The Immuno simulation results regarding T lymphocytes.

Figura 20. Immuno simulation results regarding B cell population and immunoglobulins production.

LISTA DE TABELAS

Tabela 1. Localization of target proteins.

Tabela 2. Drug targets.

Tabela 3. Identified druggable pocket with its volume, surface area and drug score of each target protein obtained from DoGSiteScorer.

Tabela 4. Docking studies of drug-like molecules (compounds) from the ZINC database with five drug target proteins. The table shows the binding scores/affinity, number of hydrogen bonds and the residues of proteins interacting with the respective compounds.

Tabela 5. Vaccine target candidates for *M.pneumoniae* identified by Vaxign. TMHMM, transmembrane propeller prediction software.

Tabela 6. Final epitopes selected for vaccine construct, highlighting the overlap between CTL and HTL epitopes and each population coverage score.

LISTA DE SIGLAS E ABREVIATURAS

2D	Two dimensions
3D	Three-dimensional
ALA	Alanine
ANN	Artificial Neural Network
ASN	Asparagine
ASP	Aspartic acid
BATS	Blast Automatic Targeting for Structures
BLASTn	Nucleotide basic local alignment search tool
BLASTp	Protein basic local alignment search tool
CAI	Codon adaptation index
CAP	Community Acquired Pneumonia
CAPES	Coordenação de Aperfeiçoamento de Pessoal de Nível Superior
CARDS	Community-Acquired Respiratory Distress Syndrome
CDC	Centers for Disease Control and Prevention
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico
COVID	Corona Virus Disease
CTL	Cytotoxic T Lymphocytes
CYS	Cysteine
DA	Dalton
DCW	Division and cell wall
DEG	Database of Essential Genes
DNA	Deoxyribonucleic acid
DPOC	Doença pulmonar obstrutiva crônica
FAPEMIG	Fundação de Amparo à Pesquisa do Estado de Minas
GC	Guanine Cytosine
GIPSy	Genomic Island Prediction Software
GLN	Glutamine
GLY	Glycine
GRAVY	Grand average of hydropathicity
HIS	Histidine
HMM	Hidden Markov Model
HTL	Helper T Lymphocytes
ICB	Instituto de Ciências Biológicas
ICBN	Instituto de Ciências Biológicas e Naturais
IEDB	Immune Epitope Database

IFN	Interferon
IgG	Immunoglobulin G
IgM	Immunoglobulin M
IL	Interleukin
ILE	Isoleucine
Kd	kilo Dalton
LCMV	Lymphocytic choriomeningitis virus
LEU	Leucine
LGCM	Laboratório de Genética Celular e Molecular
MCL	Markov Cluster Algorithm
MD	Molecular dynamics
MHC	Major histocompatibility complex
NaCl	Sodium chloride
NCBI	National Center for Biotechnology Information
NK	Natural killer
nM	NanoMolar
PAC	Pneumonia adquirida na comunidade
PAH	Pneumonia adquirida em hospital
PAI	Pathogenic Island
PCR	Reação em cadeia da polimerase
PDB	Protein data bank
PDBQT	Protein Data Bank, Partial Charge (Q), Atom Type (T)
PHE	Phenylalanine
pI	Isoelectric point
PRO	Proline
Ps	Picoseconds
PSE	Putatively surface exposed
PSSM	Position-Specific Scoring Matrix
RCSB	PDB Research Collaboratory for Structural Bioinformatics of Protein Data Bank
RMSD	Root Mean Square Deviation
RMSF	Root Mean Square Fluctuation
rRNA	Ribossomal Ribonucleic acid
SDF	Simulation Description Format
SER	Serine
SMM	Scoring Matrix Method
SPC/E	Simple point-charge / eletric
SVM	Support vector machine

tBLAST Translated comparison nucleotide basic local alignment search tool
TGF Transforming growth factor
THR Threonine
TLR Toll-Like Receptor
tRNA Transporter Ribonucleic acid BRIG BLAST Ring Image Generator IF Initiation factor
TRP Tryptophan
TYR Tyrosine LYS Lysine GLU Glutamine ARG Arginine
UFMG Universidade Federal de Minas Gerais
UFTM Universidade Federal do Triângulo Mineiro.
UTI Unidades de tratamento intensivo
VAL Valine
VPP vacina polissacarídica pneumocócica

RESUMO

O *Mycoplasma pneumoniae* é uma bactéria com características únicas quanto a seus aspectos morfológicos e metabólicos, sendo um dos principais patógenos associados à pneumonia, que mata cerca de três milhões de pessoas anualmente, e que tem se tornado mais incidente nos últimos anos graças ao sucesso da vacina pneumocócica. As peculiaridades relacionadas ao *M.pneumoniae* corroboram com a resistência à certos tipos de antibióticos, além da evidente capacidade de resistir aos fármacos atuantes, ademais, esse patógeno é extremamente difícil de ser cultivado e trabalhado em laboratório. A carência por vacinas contra a pneumonia é uma questão de preocupação global visto o aumento de patógenos resistentes, custos hospitalares associados à doença e a alta taxa de mortalidade. Métodos de desenvolvimento de vacinas utilizando dados genômicos por meio de ferramentas de bioinformática, demonstram ser estratégias promissoras, principalmente para microorganismos de difícil cultivo como o *M.pneumoniae.* Dessa forma, por meio da Vacinologia Reversa e Imunoinformática, foi construído, *in silico*, uma vacina multi-epítopo de alta cobertura contra o *M.pneumoniae*. A vacina construída contém epítopos pertencentes á lipoproteínas e proteínas adesinas altamente imunogênicas, selecionadas criteriosamente a fim de pertencerem à todas 88 linhagens de *M.pneumoniae* estudadas e serem reconhecidos por moléculas do complexo de histocompatibilidade frequentes na população mundial. Testes estruturais demonstraram estabilidade e qualidade relacionada aos parâmetros físico-químicos. O *docking* molecular do complexo vacina-receptor do tipo Toll, juntamente com a simulação da dinâmica molecular demonstraram satisfatória interação com o sistema imune, além das avaliações de imunogenicidade e simulação imunológica. A vacina demonstrou ser segura e sem potencial alérgico. Dessa forma, essa prospecção contribui para o desenvolvimento de uma vacina e prevenção da pneumonia causada pelo *M.pneumoniae*, demonstrando uma estratégia alternativa para formulação de vacinas contra patógenos de difícil cultivo, corroborando para o entendimento da interação e dos mecanismos imunológicos decorrentes das infecções por *M.pneumoniae*.

Palavras-chave: Vacina baseada em epítopo, vacina, dinâmica molecular, simulação imune, pneumonia adquirida na comunidade.

ABSTRACT

Mycoplasma pneumoniae is a bacterium with unique characteristics in terms of its morphological and metabolic aspects, being one of the main pathogens associated with pneumonia, which kills about three million people annually, and which has become more incident in recent years thanks to the success pneumococcal vaccine. The peculiarities related to *M.pneumoniae* corroborate the resistance to certain types of antibiotics, in addition to the evident ability to resist active drugs, also, this pathogen is extremely difficult to be cultivated and used in the laboratory. The lack of pneumonia vaccines is a matter of global concern given the increase in resistant pathogens, hospital costs associated with the disease and the high mortality rate. Vaccine development methods using genomic data applying bioinformatics tools, have demonstrated to be promising strategies, mainly for microorganisms of difficult cultivation such as *M. pneumoniae*. Thus, through Reverse Vacinology and Immunoinformatics, a high coverage multi-epitope vaccine against *M.pneumoniae* was built *in silico*. The constructed vaccine contains epitopes belonging to highly immunogenic lipoproteins and adhesin proteins, carefully selected to belong to all 88 *M.pneumoniae* strains studied and to be recognized by molecules of the histocompatibility complex common in the world population. Structural tests have demonstrated stability and quality related to physical-chemical parameters. The molecular docking of the Toll-like vaccine-receptor complex, together with the simulation of molecular dynamics, demonstrated satisfactory interaction with the immune system, besides to evaluations of immunogenicity and immunological simulation. The vaccine has been shown to be safe and without allergic potential. Thus, this prospecting contributes to the development of a vaccine and prevention of pneumonia caused by *M.pneumoniae*, demonstrating an alternative strategy for formulating vaccines against pathogens that are difficult to cultivate furthermore, corroborating to the understanding of the interaction and the immunological mechanisms resulting from infections by *M.pneumoniae*.

Keywords: Epitope-based vaccine, vaccine, molecular dynamics, immune simulation, community acquired pneumonia

I.PREFÁCIO

I.I Colaboradores

O presente trabalho foi realizado no Laboratório de Genética Celular e Molecular (LGCM) do Instituto de Ciências Biológicas (ICB) da Universidade Federal de Minas Gerais (UFMG) em parceria com o Laboratório de Bioinformática do Instituto de Ciências Biológicas e Naturais (ICBN) da Universidade Federal do Triângulo Mineiro, contando com as seguintes colaborações:

- Dr.Carlo José Freire de Oliveira, Professor do Instituto de Ciências Biológicas e Naturais da Universidade Federal do Triângulo Mineiro,
- Dr. Debmayla Barh, pesquisador do Institute of Integrative Omics and Applied Biotechnology, Índia,
- Dr. Helioswilton Sales-Campos, Professor do departamento de Biociências e Tecnologia da Universidade Federal de Goiás, Goiânia,
- Dr. Preetam Gosh, Professor da Virginia Commonwealth University, Estados Unidos.

Este estudo também contou com o apoio financeiro das instituições: Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG), Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES)

I.II Introdução geral

O *Mycoplasma pneumoniae* é uma das principais bactérias responsáveis por causar a pneumonia adquirida na comunidade (PAC), uma parte considerável dessas infecções necessita de cuidados hospitalares, gerando custos financeiros altos, especialmente em países com dificuldades socioeconômicas.

A falta de métodos diagnósticos rápidos e específicos para detecção do *M. pneumoniae* corrobora com o início do tratamento empírico, o que contribui para a ampla resistência à certas classes de antimicrobianos que vem se tornado um problema cada vez mais preocupante. Esses fatos, juntamente com a falta de métodos de prevenção eficientes, sustentam a necessidade de se buscar por vacinas, que são uma das melhores estratégias para prevenir a infecção e desenvolvimento de doenças (CAO *et al.*, 2017; WIEMKEN *et al.*, 2020).

Com o avanço nas tecnologias de sequenciamento e análises genômicas, ferramentas de bioinformática que visam avaliar proteínas quanto seus potenciais imunogênicos vêm sendo desenvolvidas. A Vacinologia Reversa, permite a identificação, via sequências genômicas de uma ou mais espécies de interesse, de proteínas conservadas, sem homologia com o hospedeiro e com localização celular propícia à interação e reconhecimento pelos receptores do sistema imunológico. Após posterior avaliação quanto ao caráter imunogênico

dessas proteínas, os alvos candidatos têm melhores chances de resultar em desfechos promissores. Essa abordagem, além de contribuir para formulação de vacinas de forma mais rápida e com menor exigência de recursos financeiros, é crucial no desenvolvimento de imunógenos contra patógenos fastidiosos de difícil manipulação. O *M.pneumoniae* é um dos microorganismos de vida livre com capacidade de replicação com menor tamanho físico e genômico, implicando em laborioso trabalho para isolamento e cultivo da espécie (RAZIN et al., 1996). Dessa forma, metodologias *in silico* são grandes aliadas no processo de desenvolvimento de vacinas e drogas contra essa bactéria.

As ferramentas de imunoinformática, foram desenvolvidas com base na identificação e categorização de epítopos imunogênicos de proteínas de interesse. Os epítopos considerados com melhores condições de interação e indução de respostas imunes, são destinados à construção de proteínas quiméricas que são utilizadas em estudos de diagnóstico e de vacinas. No presente estudo, essas ferramentas são aplicadas para construção de um imunógeno e a capacidade de interação com o sistema imune e indução de respostas protetoras, bem como as propriedades físico-químicas da proteína construída são avaliadas detalhadamente (TOMAR; DE, 2014).

O grupo responsável por desenvolver o presente trabalho vem aplicando a abordagem de vacinologia reversa extensivamente, com diversos patógenos de relevância clínica, como contra *Treponema pallidum*, (ARAÚJO et al., 2019; KUMAR JAISWAL et al., 2017), *Mycobacterium lepromatosis* (JAISWA et al., 2021) e *Haemophilus ducreyi* (SAROM et al., 2018).

Juntamente com a vacinologia reversa, diversos trabalhos utilizando a imunoinformática vem sendo desenvolvidos pela equipe responsável por este trabalho (SANCHES *et al.*, 2021; SANTOS *et al.*, 2013; TOSTA *et al.*, 2021).

As perspectivas futuras são baseadas no aprimoramento da metodologia, à medida que programas e algoritmos robustos são desenvolvidos, com posterior produção dessas proteínas quiméricas, validação da estabilidade estrutural e análise *in vivo* da capacidade de induzir respostas imunes protetivas.

I.III Estrutura do manuscrito

Esta dissertação está dividida em introdução e dois capítulos baseados em artigos de pesquisa, como descrito a seguir:

a. Na introdução, está presente a descrição do patógeno abordado no estudo, o *Mycoplasma pneumoniae*, informações morfológicas e características genômicas são abordadas, bem como os fatores determinantes na patogenicidade, resposta imune e complicações ao causar a pneumonia. Posteriormente é abordada uma visão geral da pneumonia, bem como as principais classes existentes e patógenos associados, seguido da

descrição detalhada da pneumonia adquirida na comunidade, com informações epidemiológicas, fisiopatológicas, fatores de risco, diagnóstico e tratamento. Por fim há uma breve descrição sobre as vacinas e os tipos existentes, com foco nas abordagens tratadas nesta prospecção, vacinologia reversa e imunoinformática.

b. O capítulo I se refere ao artigo publicado, em que abordagens de vacinologia reversa foram aplicadas à 88 genomas de *M.pneumoniae,* a fim determinar potenciais alvos vacinais e de drogas contra essa bactéria. Ferramentas de genômica comparativa foram utilizadas para encontrar o genoma central (core) do patógeno e a genômica subtrativa foi empregada para determinar os genes não homólogos ao hospedeiro (*core-non-host*). As proteínas selecionadas através dessas análises foram caracterizadas quanto à localização subcelular, essencialidade e imunogenicidade para os alvos de vacinais, enquanto os alvos de drogas foram submetidos à ancoragem, *docking* molecular, contra um banco de dados de moléculas com potencial ao uso como drogáveis. 8 candidátos à alvos vacinais e 5 potenciais alvos de drogas foram encontrados.

c. O capítulo II trata-se do artigo submetido, já revisado, em que proteínas encontradas por vacinologia reversa ou selecionadas da literatura pelo alto potencial imunogênico, foram submetidas a análises de imunoinformática, com o intuito de construir uma proteína quimérica de alta cobertura e com características imunogênicas contra o *M.pnemoniae*. A proteína construída foi submetida a análises de alergenicidade, toxicidade e homologia com o hospedeiro. A proteína construída foi avaliada quanto as suas características estruturais e submetida ao docking molecular contra receptores do sistema imunológico. A simulação imune *in silico* e a clonagem imune *in silico* também foram realizadas, ambas com resultados satisfatórios.

Ao final de cada capítulo foi apresentada uma breve conclusão e posteriormente, a bibliografia e sessão de anexos com materiais suplementares referentes à cada capítulo, juntamente com o *curriculum vitae*.

1. INTRODUÇÃO

1.1 MYCOPLASMA PNEUMONIAE

Os *micoplasmas* pertencem à uma classe dos *Molicutes,* um grupo procarionte separado dos demais por agregar organismos que não possuem parece celular (RAZIN et al., 1996). Essas bactérias singulares, são consideradas as menores com capacidade de autorreplicação, de forma que são compostas apenas das organelas e estruturas essenciais para seu desenvolvimento, o que também confere um genoma consideravelmente pequeno. A hipótese de que os *micoplasmas* evoluíram de bactérias Gram-positivas é sustentada por diversos achados genéticos obtidos através de estudos filogenéticos (ROGERS et al., 1985). Durante esse processo evolutivo, os *micoplasmas* sofreram uma perda genética significante, mas simultaneamente, adquiriram capacidades metabólicas e variações proteicas capazes de permitir o escape do sistema imune hospedeiro e manter seu desenvolvimento na forma parasitária (RAZIN; YOGEV; NAOT, 1998). Existem mais de 200 espécies de *micoplasmas* conhecidas, dentre elas, o *M.pneumoniae* é um patógeno parasitário exclusivo de humanos e é a espécie mais estudada por sua incidência nos casos de pneumonia (WAITES; TALKINGTON, 2004).

O tamanho do genoma da espécie *M.pneumoniae* é de cerca de 820 kb, que codifica 700 operons, o que é mais curto do que uma grande parte das bactérias conhecidas (HIMMELREICH *et al.*, 1996). Uma das proteínas mais importantes para a genotipagem é a proteína de superfície adesina P1, que divide as cepas de *M.pneumoniae* em tipo 1 e tipo 2 (DORIGO-ZETSMA *et al.*, 2001). Estudos comparando genomas de sequência dos dois tipos, demonstraram que mesmo com a alta similaridade entre as cepas (> 99% de identidade), os tipos diferem em cerca de 1.500 polimorfismos de nucleotídeo único (SNP) e variantes *indel.* O mesmo estudo também revela que o genoma de *M.pneumoniae* se manteve estável ao longo do tempo e em diferentes regiões do mundo, o que demonstra baixa taxa de transferência gênica (XIAO *et al.*, 2015). Essas informações foram sustentadas por outros estudos que também não encontraram evidências de plasmídeos ou fagos, demonstrando a estabilidade genômica entre as cepas, com poucos SNPs e a clonalidade entre as linhagens de *M.pneumoniae* (LEE *et al.*, 2019; XIAO *et al.*, 2015).

O *M.pneumoniae* não apresenta parece celular, sua membrana é composta em grande parte por lipídeos e proteínas diversas. Dentre essas proteínas, destacam-se as lipoproteínas, as quais já foram identificados mais de 50 tipos, elas possuem além da alta capacidade imunomodulatória, funções específicas de transporte, citoaderência e síntese de subunidades proteicas (CHRISTODOULIDES et al., 2018). Ademais, as lipoproteínas auxiliam também na evasão imunológica através de variação antigênica e formação de uma “capa micoplasmática” evitando a interação com anticorpos (CHAMBAUD; WRÓBLEWSKI; BLANCHARD, 1999).

1.1.1 Mecanismos de patogenicidade

Os mecanismos patogênicos associados ao *M.pneumoniae* se devem a diversos fatores de virulência. (1) Visto que o *M.pneumoniae* não apresenta parede celular, a capacidade de adesão as células respiratórias hospedeiras é mediada por um complexo de proteínas adesinas, sendo as principais as adesinas P1 e a P30, que além da citoaderência também apresentam propriedades imunogênicas (DRASBEK *et al*., [S.d.]) (Figura 1). (2) O processo inflamatório é mediado de forma intensa pelas lipoproteínas membranares, pela produção de metabólitos reativos de oxigênio e algumas enzimas (CHOI et al., 2012; GROSSHENNIG et al., 2016); (3) ao mesmo tempo, nucleases e proteínas de ligação à imunoglobulina como a Ibpm, auxiliam no escape imunológico (BLÖTZ et al., 2020; YAMAMOTO et al., 2017). (4) A Toxina da Síndrome do Desconforto Respiratório Adquirido na Comunidade (Toxina CARDS) é um importante agente citotóxico, também existem proteínas associadas à motilidade como à P24, P30, P41 e P65 (HASSELBRING; KRAUSE, 2007; HASSELBRING; SHEPPARD; KRAUSE, 2012; KANNAN et al., 2014).

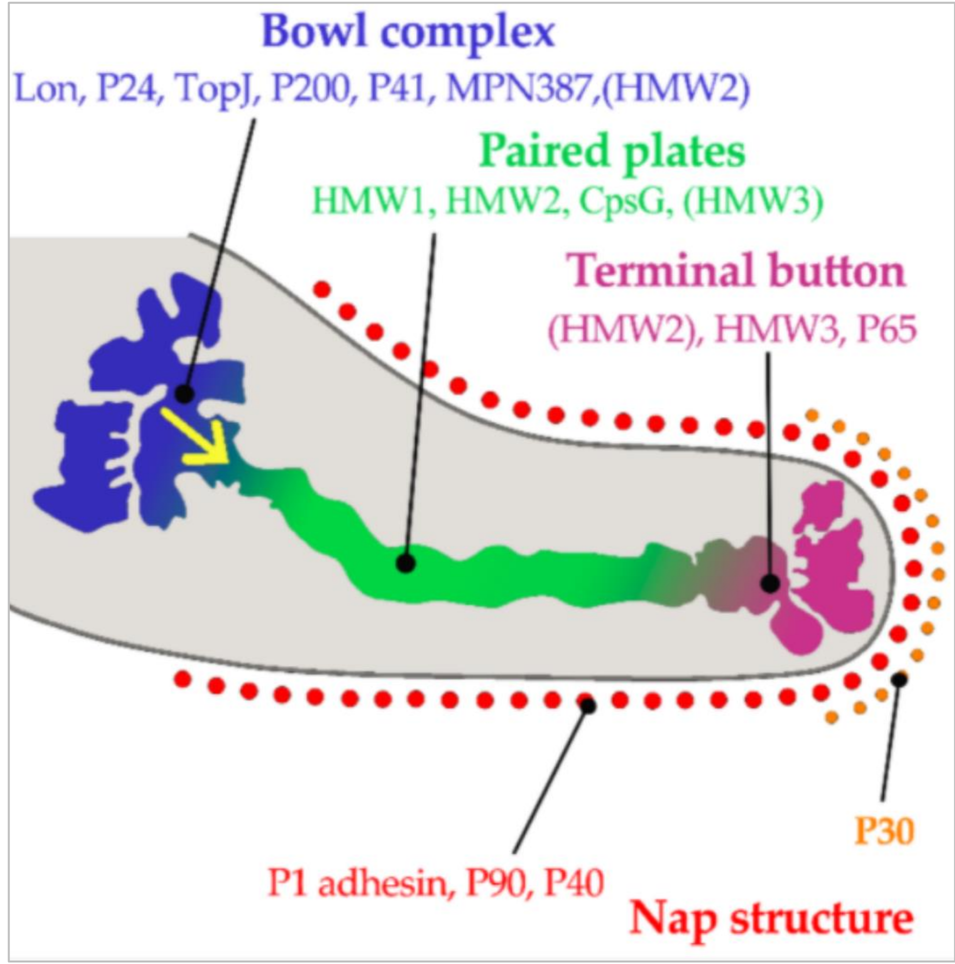


Figura 1. Componentes proteicos importantes na citoaderência e motilidade do *M. pneumoniae*. A imagem retrata o complexo de proteínas de alto peso molecular HMW, que fazem parte de complexos proteicos que transmitem a força gerada dos complexos de *Bowl* para o complexo *Paired*, até o *Terminal button* e a estrutura NAP. A estrutura NAP consiste no conjunto de proteínas de adesão à

superfície que juntamente com a adesina P30 permitem a fixação da bactéria ao epitélio hospedeiro e a integração intracelular.
Fonte: Zhulin et all., 2021.

1.1.2 Resposta imune ao *M.pneumoniae*

A colonização do trato respiratório pelo *M.pneumoniae* leva a ativação e infiltração pulmonar de linfócitos T CD4+ e células B, ocasionando uma amplificação de resposta mediada pelas citocinas pró-inflamatórias e imunoglobilinas. Anticorpos como IgG, IgM e IgA são produzidos em altas titulações e a produção de IgE também é relatada (OPITZ et al., 1997; STELMACH et al., 2005). O perfil de resposta predominante é o Th2 nas respostas ao *M.pneumoniae*, embora estudos também relatam produção de citocinas características de resposta Th1 (KOH et al., 2001).

Grande parte do potencial inflamatório induzido pelo *M.pneumoniae* se dá pelos fatores de virulência que são altamente imunogênicos. Os glicolipídeos de membrana, ativam mecanismos de resposta imune inata não específicos mediados por anticorpos, alexina, fagócitos e componentes do sistema complemento (HE et al., 2016). As lipoproteínas, também são cruciais nesse processo, favorecendo a imunidade mediada por células através do recrutamento neutrofílico e induzindo a produção de citocinas inflamatórias como a IL-2, IL-6 e TNF-α (GARCIA et al., 1998). A interação com receptores do sistema imune inato como os do tipo Toll-like Receptor (TLR) 2(CACCIOTTO et al., 2016), TLR-4 (SHIMIZU et al., 2014) e TLR-6 (TAKEUCHI et al., 2002) também compõem uma parte importante no reconhecimento antigênico, eles reconhecem padrões associados à patógenos, pertencentes em grande parte às lipoproteínas e induzem a liberação de citocinas, quimiocinas, fazendo parte também do recrutamento celular (ANDREWS et al., 2013; SHIMIZU, 2016). Os receptores intracelulares também são importantes mediadores do reconhecimento de padrões moleculares associados ao *M.pneumoniae* e ativam citocinas pró-inflamatórias como IL-1β (SUGIYAMA et al., 2016). Processos de autofagia também são relatados como indutores de respostas inflamatórias, com recrutamento de macrófagos e produção de citocinas, no qual a citoaderência é um fator crucial (SHIMIZU et al., 2014). As proteínas envolvidas no processo de adesão e fixação as células epiteliais respiratórias, como as adesinas P1, também são potentes estimuladoras da liberação de citocinas como IL-4 pelos mastócitos e da precoce produção de anticorpos (HOEK et al., 2005; HU et al., 1983).

1.1.3 *M.pneumoniae* e a Pneumonia Adquirida na Comunidade (PAC)

O *M.pneumoniae* é um dos principais agentes causadores da PAC, com taxas de incidência que variam de 10 a 40% (M. WINCHELL, 2013). As infecções causadas por esse patógeno podem ser endêmicas e ocasionar surtos de tempos em tempos, com índices de até 50% dos casos de PAC causadas por *M. pneumoniae* (FOY *et al.*, 1973;

HAMMERSCHLAG, 2001). Esse tipo de infecção atinge mais crianças e adultos jovens, porém, esse número pode ser ainda maior devido à falta de protocolos de diagnóstico microbiológico específico, surtos não detectados e ao fato de parte significativa dos infectados apresentar apenas sintomas leves (WAITES *et al.*, 2017; WINCHELL, 2013). Os surtos relatados ao *M.pneumoniae* ocorreram em lares de idosos, universidades e escolas, onde uma parte significativa dos pacientes adquiriram pneumonia agressiva, com manifestações neurológicas e cutâneas, nas quais as taxas de mortalidade atingiram 13% em um desses surtos ocorrido em Nebraska (BAJANTRI; VENKATRAM; DIAZ-FUENTES, 2018; HASTINGS *et al.*, 2015; RALSTON, 1979; WALLER *et al.*, 2014).

As características físicas auxiliam no diagnóstico de infecções por *M. pneumoniae,* como as condições pulmonares, entretanto esses achados clínicos e radiológicos não são específicos. Dessa forma a determinação microbiológica é extremamente importante na decisão terapêutica. Porém, a espécie é extremamente difícil de ser cultivada em laboratório. Por não possui parede celular, a visualização pelo Gram não é possível, então exames como Reação em cadeia da polimerase (PCR), cultura e sorologia são indicadas para o diagnóstico, mas também apresentam desafios quando os resultados não são concordantes (WAITES *et al.*, 2017). Além disso, o *M.pneumoniae* é um organismo fastidioso e requer culturas específicas, sendo bactéria de crescimento lento que torna o processo mais difícil. Dessa forma, o padrão ouro é uma sorologia quádrupla que mede os títulos de anticorpos ao longo do tempo (BAJANTRI; VENKATRAM; DIAZ-FUENTES, 2018).

O tratamento contra o *M.pneumoniae* é um desafio complexo. Por não produzir compostos da parede celular como os peptídeosglicanos, as classes de antimicrobianos que atuam nesta estrutura, como os β-lactâmicos, fosfomicinas e glicopeptídeos, não podem ser usadas no tratamento e opções como sulfonamidas, trimetoprim, rifampicina, polimixinas e ácido nalidíxico também não demonstraram eficácia terapêutica (WAITES et al., 2017). Devido a isso, aqueles antimicrobianos cujo mecanismo de ação contemple a síntese de proteínas ou que inibam a replicação do DNA são as escolhas mais utilizadas, como macrolídeos, cetolídeos tetraciclinas, estreptograminas e fluoroquinolonas (Figura 2) (WAITES; TALKINGTON, 2004). Os macrolídeos são os antimicrobianos de primeira linha usados para tratar essas infecções, no entanto, desde 2000, cepas resistentes a essa classe de antibióticos têm sido cada vez mais isoladas em todo o mundo (MOROZUMI; TAKAHASHI; UBUKATA, 2010). As infecções disseminadas por *M.pneumoniae* resistente a macrolídeos (MRMP) estão relacionadas à pneumonia grave com sintomas severos e evidências radiológicas mais agressivas (HSIEH *et al.*, 2012; KOGA *et al.*,2012).

As infecções causadas por *M. pneumoniae* geram principalmente sintomas leves como febre, dor de cabeça, dor de garganta, dor no peito e tosse que pode ocorrer com a produção de muco. A PAC causada por *M.pneumoniae* também é conhecida como “pneumonia

ambulante" devido às suas manifestações clínicas leves. Porém, as infecções podem se agravar causando pneumonia agressiva com sintomas de asma, encefalite, anemia hemolítica e até disfunção renal ("*Mycoplasma pneumoniae* | Home | CDC", [S.d.]). A doença pulmonar obstrutiva crônica (DPOC) também está relacionada às infecções por *M. pneumoniae* (LIEBERMAN *et al.*, 2002).

Antimicrobiano	Mecanismos de ação	Estrutura básica
MACROLÍDEOS	Ligam-se ao ribossomo impedindo a síntese proteica	ERITROMICINA
CETOLÍDEOS	Pertencem ao grupo dos macrolídeos e se ligam ao ribossomo impedindo a síntese proteica	TELITROMICINA
TETRACICLINAS	Ligam-se ao ribossomo impedindo a síntese proteica	
ESTREPTOGRAMINAS	Pertencem ao grupo dos macrolídeos e ligam-se ao sítio P do ribossomo impedindo a síntese proteica	QUINOPRISTINA
FLUOROQUINOLONAS	Inibidoras da enzima DNA topoisomerase, inibindo a replicação do DNA	CIPROFLOXACINA

Figura 2. Antimicrobianos utilizados no tratamento à infecções por *M.pneumoniae*
A imagem retrata os principais antimicrobianos utilizados no tratamento à infecções por *M.pneumoniae,* uma breve explicação do mecanismo de ação e sua respectiva estrutura química.
Fonte: Dados obtidos de WAITES e TALKINGTON, 2004.

1.2. PNEUMONIA

A pneumonia é uma infecção pulmonar aguda que pode ser causada por diversos microrganismos como vírus, bactérias e fungos. Essa doença é responsável por uma enorme morbidade e mortalidade, levando também a problemas socioeconômicos. A pneumonia é considerada uma das quatro principais causas de morte no mundo e também está no ranking das internações ("The top 10 causes of death", [S.d.]). Ela pode ser classificada de diferentes

formas, mas as principais são: Pneumonia Adquirida na Comunidade (PAC) referente a infecções adquiridas fora do hospital e Pneumonia Adquirida em Hospital (PAH) que se relaciona com as infecções adquiridas após 48 horas da internação hospitalar (LANKS; MUSANI; HSIA, 2019).

Os microrganismos causadores da pneumonia podem ser organizados em duas classes, organismos típicos, que se referem a aqueles corados pelos métodos de Gram e cultivados utilizando meios de cultura padrão; e os atípicos, que possuem diagnóstico que não se enquadra nos convencionais. As principais bactérias típicas *são Streptococcus pneumoniae*, *Staphylococcus aureus*, estreptococos do Grupo A, *Moraxella catarrhallis*, bactérias aeróbias Gram-negativas como *Klebsiella pneumoniae* e *Pseudomonas aeruginosa* e alguns anaeróbios. Os microorganismos atípicos comumente causadores da pneumonia são: *Mycoplasma pneumoniae*, *Legionella* e *Chlamydia pneumoniae* (RAMYA, 2014). Entre os vírus, Influenza e o coronavírus são os grandes responsáveis por causar a pneumonia (“Causas da pneumonia | CDC”, [S.d.]).

1.2.1 Pneumonia adquirida em hospital

Se trata das infecções pulmonares adquiridas nas unidades de tratamento intensivo (UTI). As principais bactérias associadas são *Staphylococcus aureus, Pseudomonas aeruginosa*, espécies de *Klebsiella* spp*, Escherichia coli, Enterobacter* e *Acinetobacter*. Uma grande preocupação a respeito dessas infecções, é que devido ao alto nível de debilidade dos pacientes nessas condições, o tratamento empírico é o mais utilizado, muitos casos sendo prescritos até 3 classes de antimicrobianos de amplo espectro, como β-lactâmicos, com aminoglicosídeos ou com fluoroquinolonas o que agrava ainda mais a resistência à antibióticos que vem sendo desenvolvida por diversas bactérias (MARTIN-LOECHES; RODRIGUEZ; TORRES, 2018; LYNCH, 2001).

1.2.2 Pneumonia associada à ventilação

Essa classificação de pneumonia, que se enquadra dentro da PAH, consiste das infeções adquiridas entre 48-72 horas após a intubação. O índice de pacientes que necessitam de suporte ventilatório varia de 8 a 28%, com taxas de mortalidade alarmantes. Os principais organismos responsáveis por esse tipo de pneumonia são *S. aureus* e *P. aeruginosa* (CHASTRE; FAGON, 2002).

1.2.3 Pneumonia por aspiração

Pneumonia causada por aspiração, é outro termo utilizado para caracterizar infeções pulmonares causadas após aspiração de grandes volumes de conteúdo gastrointestinal ou oriundos da orofaringe. Esse tipo de aspirado normalmente contém muitas bactérias, grande

parte delas anaeróbicas, provenientes do trato gastrointestinal superior e orofaringe que levam a infecções. Esse tipo de pneumonia apresenta fisiopatologia diferente da PAC e PAH, o que condiz com diferentes protocolos de tratamento (DIBARDINO; WUNDERINK, 2015).

1.3 PNEUMONIA ADQUIRIDA NA COMUNIDADE

A PAC é caracterizada como uma doença grave que atinge principalmente menores de 5 anos de idade e idosos portadores de comorbidades, além disso é a mais comum das pneumonias (JAIN; WILLIAMS; *et al.*, 2015; PAHAL; SHARMA, 2018). Outro aspecto importante é que essa morbidade frequentemente requer cuidados hospitalares de custo elevado e embora não seja uma doença restrita à países menos desenvolvidos, esse fato corrobora com a mortalidade intensificada em regiões com problemas econômicos, demonstrando a expressiva importância das ações políticas e socioeconômicas para tratamento e prevenção da doença (SULTANA *et al.*, 2019). A incidência da PAC provavelmente é subestimada devido aos casos não notificados ou diagnosticados ao fato de que na maioria dos casos as infecções são brandas, dispensando a procura dos serviços hospitalares. (ANGUS *et al.*, 2001). Entretanto, apenas nos Estados Unidos, mais de 4 milhões de pessoas são atendidas por ano com PAC, levando a 1,5 milhões de hospitalizações (RAMIREZ *et al.*, 2017).

1.3.1 Agentes Etiológicos

Algumas bactérias são comumente introduzidas no trato respiratório inferior por aspiração normal como bacilos Gram-negativos, *S.pneumoniae* e *H.influenzae.* Vírus e bactérias intracelulares, como *M. pneumoniae*, *C.pneumoniae* e *Legionella spp*, podem ser inalados para o trato respiratório inferior em micropartículas. Na maioria dos casos, o sistema imunológico inibe a progressão para pneumonia, porém, dependendo da frequência ou volume de bactérias aspiradas, virulência da cepa, fatores de risco e condições imunológicas, a aspiração pode levar à pneumonia (MICHAEL SCHELD; MANDELL, 1991).

O principal microrganismo responsável pela PAC é o *S.pneumoniae*, sendo isolado em cerca de 25% dos casos, causando pneumonias leves a graves (SAID *et al.*, 2013). Vírus e outras espécies como *M. pneumoniae* e *C.pneumoniae* causam infecções leves em geral, mas podem desenvolver sintomas mais agressivos. Os organismos relacionados à evolução grave da doença são os vírus como *influenza* e espécies de bactérias como *S.aureus*, *Legionella* e *H.influenzae*. Bactérias Gram-positivas como *S.aureus* também estão relacionadas com a PAC assim como bactérias Gram-negativas, sendo as mais conhecidas *Pseudomonas aeruginosa*, *Klebsiella pneumoniae* e *Enterobacter spp* (GROSSO; FAMIGLIETTI; LUNA, 2015).

1.3.2 Fatores de risco e manifestações clínicas

Comorbidades e idade avançada são fatores que aumentam o risco de desenvolver pneumonia. As principais comorbidades consistem em problemas respiratórios crônicos, doenças cardiovasculares e renais. Hábitos diários como fumo, uso constante de álcool, desnutrição e falta de cuidado odontológico podem aumentar o potencial para desenvolver pneumonia. Além disso, variações genéticas foram determinadas como fatores de risco para pneumonia, como as encontradas no gene FER do proto-oncogene da proteína tirosina-quinase. Essa variante está relacionada a um menor risco de morte por pneumonia, pois a proteína codificada por esse gene está associada à adesão, quimiotaxia e migração celular (RAUTANEN *et al.*, 2015).

A infecção do trato respiratório inferior que leva à PAC causa tosse, febre, dispneia, pleurisia e produção de escarro como os sintomas mais comuns. Nesse ínterim, a doença pode levar a uma pneumonia atípica, principalmente em pessoas de idade avançada, com manifestações não respiratórias como diarreia, mialgia, confusão e mal-estar, que também aumentam as taxas de mortalidade (BISCEVIC-TOKIC; TOKIC; MUSANOVIC, 2013; NONIKOV, 2001).

1.3.3 Diagnóstico

Com o avanço da tecnologia, o diagnóstico é um desafio no PAC, pois o patógeno não é identificado em cerca de 50% dos casos (JAIN; SELF; *et al.*, 2015). Com a disseminação da resistência antimicrobiana, a definição específica precoce do patógeno para rápida decisão terapêutica contribui tanto para um bom prognóstico do paciente quanto para retardar o aumento da resistência.

A radiografia e a tomografia computadorizada (TC) podem definir o comprometimento pulmonar, assim como os infiltrados pulmonares. No entanto, ambas as opções têm suas limitações. A radiografia tem menos sensibilidade que a TC, mas esta última, por mais que seja o padrão ouro, expõe o paciente à radiação e não pode ser realizada no leito de internação. Outra abordagem é a ultrassonografia, que tem melhor sensibilidade que a radiografia e pode ser realizada à beira do leito com os resultados avaliados no momento do exame (XIROUCHAKI *et al.*, 2011).

Os exames laboratoriais para pacientes com pneumonia são os de rotina, nos quais exames de sangue podem indicar infecções agudas por leucocitose e os testes bioquímicos realizados com o soro podem indicar a associação da doença com outros órgãos, como rins ou fígado. Outros resultados fornecem informações sobre o tipo de infecção PAC e podem auxiliar no direcionamento do tratamento. Por exemplo, a detecção de procalcitonina indica respostas à infecção bacteriana, enquanto a falta dessa proteína sérica é um sinal para infecções virais (LANKS; MUSANI; HSIA, 2019; SELF *et al.*, 2017).

O diagnóstico microbiológico, pode ser desafiador em muitos casos, mas tem sua importância na determinação do tratamento e deve ser realizado de acordo com um protocolo que considera a condição do paciente. Portanto, de acordo com os achados clínicos e epidemiológicos de um paciente, esse tipo de investigação é altamente recomendável, pois pode modular o início do tratamento. Com a identificação do patógeno, o tratamento que era empírico pode ser alterado quanto à dosagem, espectro da mesma classe, o que também diminui a pressão para resistência aos antibióticos. Algumas diretrizes indicam ampla cobertura empírica como primeira conduta, pois a redução ou a escolha errada do antibiótico pode aumentar o risco de mortalidade. No entanto, a seleção do antibiótico com base no diagnóstico pode reduzir as reações adversas, o custo do tratamento e o potencial de induzir resistência aos antibióticos (ARANCIBIA *et al.*, 2000; MANDELL *et al.*, 2007).

A coloração por Gram e cultura são indicadas para todos os pacientes com escarro purulento, entretanto, a cultura de pacientes sem esse sintoma ou outra amostra do trato respiratório inferior pode ser ineficaz devido à grande possibilidade de cultivo não patogênico. As hemoculturas são destinadas a pacientes de gravidade moderada ou alta, devido às dificuldades de isolamento e de crescimento a partir desse tipo de amostra. Também podem ser realizados testes específicos para identificar algumas bactérias como *S. pneumoniae* e *Legionella pneumophila*. Com o teste do antígeno urinário, essa espécie causadora de pneumonia pode ser identificada, mas não substitui as culturas (LANKS; MUSANI; HSIA, 2019).

A reação em cadeia da polimerase (PCR) é um teste muito comum usado para identificar patógenos no escarro como a possibilidade de detectar uma variedade de espécies de vírus e bactérias. O teste é altamente sensível e específico, porém não diferencia os patógenos do trato respiratório superior do inferior. De fato, o uso de PCR juntamente com técnicas de cultura pode melhorar a capacidade de detecção do patógeno (JOHANSSON *et al.*, 2010; LANKS; HSIA, 2019). O principal problema é que em mais da metade dos casos de PAC o patógeno causador é mal diagnosticado, devido à falta de exames específicos, protocolos laboratoriais avançados ou existência de equipamentos e recursos para confirmação via PCR (VAN WERKHOVEN; HUIJTS, 2018).

1.3.4 Abordagem terapêutica

Os antimicrobianos são a linha de frente do tratamento da PAC. Em geral, as respostas são adequadas, no entanto, os tratamentos empíricos podem levar a falhas na resposta à terapia escolhida e o paciente pode desenvolver pneumonia progressiva. Essa condição é caracterizada por deterioração clínica nas primeiras 24 horas após o tratamento, com alterações notáveis na radiografia e complicações vitais (ARANCIBIA *et al.*, 2000; MENÉNDEZ; PERPIÑÁ; TORRES, 2003). A definição do tratamento não está apenas

relacionada a ser empírico ou baseado em testes microbiológicos, mas também segue outras orientações quanto à farmacocinética, custo, espectro e perfil de segurança (FILE, 2003). O tratamento empírico recomendado para pacientes com PAC é o uso de antimicrobianos de amplo espectro, com eficácia contra os principais agentes causadores da doença, por exemplo, amoxicilina e macrolídeos. Além disso, as instruções diferem quanto à possibilidade de patógenos multirresistentes como *S. aureus* resistente à meticilina (MRSA) e *Pseudomonas aeruginosa*, com a prescrição de β-lactâmicos ou fluoroquinolonas (METLAY *et al.*, 2019). O tratamento dirigido por patógenos pode ser aplicado após o diagnóstico específico, porém pode levar até 72 horas para ser determinado.

1.4 VACINAS

É um consenso que as vacinas previnem milhões de mortes anualmente e reduzem a mortalidade e morbidade da população vacinada. Com os programas globais de vacinação, a varíola foi erradicada e a prevenção de outras doenças como a pneumonia é extremamente importante. No entanto, atualmente existem apenas duas vacinas implementadas contra pneumonia, mas ainda sem uma concordância geral sobre a eficácia, elas são: a vacina polissacarídica pneumocócica (VPP) e a vacina contra influenza, ambas estão disponíveis e são recomendadas pelo CDC (Centers for Disease Control and Prevention) ("General Recommendations on Immunization", [S.d.]).

Vacinas podem ser definidas como um produto biológico, que contém antígenos oriundos de um patógeno ou sintetizados para seus componentes. Esses antígenos devem, de forma segura, induzir respostas imunes protetivas contra futura exposição ao determinado patógeno ou à uma doença (POLLARD; BIJKER, 2021). As vacinas vivas têm potencial replicativo no hospedeiro e com isso, grande poder de induzir respostas imunes de eficácia. Entretanto, em pacientes imunocomprometidos, essa replicação pode se dar de forma exacerbada, o que faz com que o uso dessas vacinas seja limitado nestes indivíduos (RUBIN *et al.*, 2014). Vacinas contra sarampo, caxumba e rubéola são exemplos de vacinas vivas. Para equilibrar a capacidade replicativa que auxilia no desenvolvimento de resposta imune adequada, pode haver a atenuação do patógeno para impedir sua replicação descontrolada sendo necessário um alto controle e entendimento do organismo, além de alguns tipos de vacinas necessitarem de doses extras (MILLIGAN *et al.*, 2018).

As vacinas de antígenos não vivos podem ser compostas pelo organismo inteiro morto, como a vacina da poliomielite, mas também podem ser compostas por frações dos organismos como proteínas recombinantes (vacinas de subunidade) ou polissacarídeos como as vacinas da hepatite B e a vacina pneumocócica respectivamente. As vacinas formuladas a partir de toxinas, como as contra tétano e difteria, também se enquadram nessa classificação (POLLARD; BIJKER, 2021).

Dentre as vacinas de subunidade, aquelas baseadas em epítopos, têm demonstrado serem opções promissoras para prevenir doenças. Os epítopos são partes de um antígeno capazes de serem reconhecidos por receptores do sistema imune e através dessa interação, induzir imunidade protetiva contra determinado patógeno (PARVIZPOUR *et al.*, 2020).

Com as tecnologias de sequenciamento de próxima geração, a acessibilidade a sequências genômicas de vários bancos de dados por exemplo - NCBI (AGARWALA *et al.*, 2018), GenBank (CLARK *et al.*, 2016), EuPathDB (WARRENFELTZ *et al.*, 2018), WormBase (HOWE *et al.*, 2012) e *Virus Pathogen Database and Analysis Resource* (ViPR) (PICKETT *et al.*, 2012) orientaram de forma notável a identificação de alvos de vacinas e novos candidatos a medicamentos. A área da genômica abrange técnicas *in silico* de alto rendimento como Vacinologia reversa e Imunoinformática para triagem e identificação de alvos vacinais (BAH *et al.*, 2018).

TIPO DE VACINA		VACINAS LICENCIADAS USANDO ESSA TECNOLOGIA
Viva atenuada ou inativada		Sarampo, caxumba, rubeola, febre amarela, influenza, polio, febre tifoide, rotavirus, BCG, varicella zooster
Organismo completo morto		Coqueluche, polio, influenza, encefalite japonesa, Hepatite A, raiva
Toxoide		Tétano e difteria
Subunidade (proteína purificada, Proteína recombinante, Polissacarídeo, peptídeo)		Coqueluche, influenza, hepatite B, meningococco, pneumococo, febre tifóide, hepatite A
Partícula *Vírus-like*		Papilomavírus humano
Vesícula de membrana externa	Antígeno do patógeno; Membrana externa de Gram-negativo	Meeningocócos do grupo B
Conjugada proteína-polissacarídeo	Polissacarídeo; Proteína carreadora	*Haemophilus influenzae* tipo B, pneumococo, meningococo, tifóide
Vetor viral	Vetor viral; Gene do patógeno; Genes do Vetor viral	*Ebola*
Vacina de DNA	DNA; RNA; Capa Lipídica	SARS-CoV-2
Vetor bacteriano	Gene do patógeno; Vetor Bacteriano	Experimental
Células apresentadoras de antígenos	Antígeno do patógeno; MHC	Experimental

Figura 3. Diferentes tipos de vacinas

A imagem mostra as diferentes classes de vacinas existentes, os exemplos de vacinas já disponíveis no mercado para cada uma das classes e as classes de vacinas que ainda estão em fase de teste. Fonte: Adaptado de Pollard e Bijker., 2020.

1.4.1 Vacinologia Reversa

Com os avanços nos métodos de sequenciamento e análise de genomas, foi determinado a “vacinômica”, que utiliza ferramentas de bioinformática aplicadas à farmacogenômica e geraram uma massiva quantidade de dados (HU et al., 2004). A partir dessas informações, surgiu a Vacinologia Reversa, que consiste no processo de descoberta de um antígeno, começando com a interrogação de todo o repertório antigênico do patógeno

a partir de dados genômicos. Essa abordagem permite a análise de todo repertório proteico do patógeno de interesse, permitindo a avaliação de cada uma delas em relação ao potencial imunogênico para assim determinar quais os meljores candidatos vacinais (RAPPUOLI, 2000). O desenvolvimento convencional da vacina se baseia principalmente nos conceitos de Pasteur, com isolamento, manipulação e injeção do agente causador da doença e avaliação das condições para produção da vacina, que acarreta à uma alta demanda de recursos financeiros e trabalho laboratorial que pode levar anos (DEL TORDELLO; RAPPUOLI; DELANY, 2017). Por outro lado, a abordagem da vacinologia reversa não necessita necessariamente que os patógenos sejam cultivados em laboratório, já que existe a possibilidade de trabalhar com os dados já sequenciados (RAPPUOLI, 2000). A vacinologia reversa (Figura 4) usa uma triagem *in silico* de todo o genoma de patógenos específicos, para determinar os genes adequados ou proteínas candidatas para melhores alvos vacinais. Primeiramente, através de genômica comparativa, as proteínas do genoma central, aquelas conservadas em todas as linhagens estudadas, são determinadas. Posteriormente, através da genômica subtrativa, são excluídas todas as proteínas do genoma central, *core*, que apresentam similaridade com o proteoma do hospedeiro humano, sobrando apenas o *core-non-host,* correspondendo ao grupo de proteínas conservadas, mas que não apresenta relação com proteínas humanas (RINAUDO et al., 2009). Por fim, essas proteínas são avaliadas quanto a localização subcelular, aquelas consideradas de membrana, expostas a superfície ou secretadas são submetidas a análises quanto seu potencial imunogênico. Essa abordagem permite o reconhecimento não só de antígenos já estudados por métodos anteriores mas também de novos antígenos com aspectos diferentes e que podem atuar de maneira eficaz na proteção imunológica (KANAMPALLIWAR, 2020). Outras aplicações se dão a partir da utilização dessa estratégia para construir vacinas capazes de atuar contra um gênero de interesse ou contra diferentes patógenos. Após a identificação dos alvos mais adequados, estes podem ser expressos e testados *in vitro* ou *in vivo* para validar as propriedades imunogênicas (SEIB; ZHAO; RAPPUOLI, 2012). Rino Rappuoli foi o primeiro a usar o termo vacinologia reversa (RAPPUOLI, 2001). Ele desenvolveu um trabalho com *Neisseria meningitides* serogrupo B (MenB), uma bactéria Gram- negativa responsável por 50% dos casos de meningite meningocócica globalmente, foi o primeiro patógeno a ser submetido à esta técnica para desenvolvimento de vacinas (KELLY; RAPPUOLI, 2005). A partir desses avanços na genômica, proteômica e imunoinformática, ferramentas dessas abordagens vêm sendo consecutivamente aplicadas no diagnóstico, na terapêutica e no desenvolvimento de vacinas.

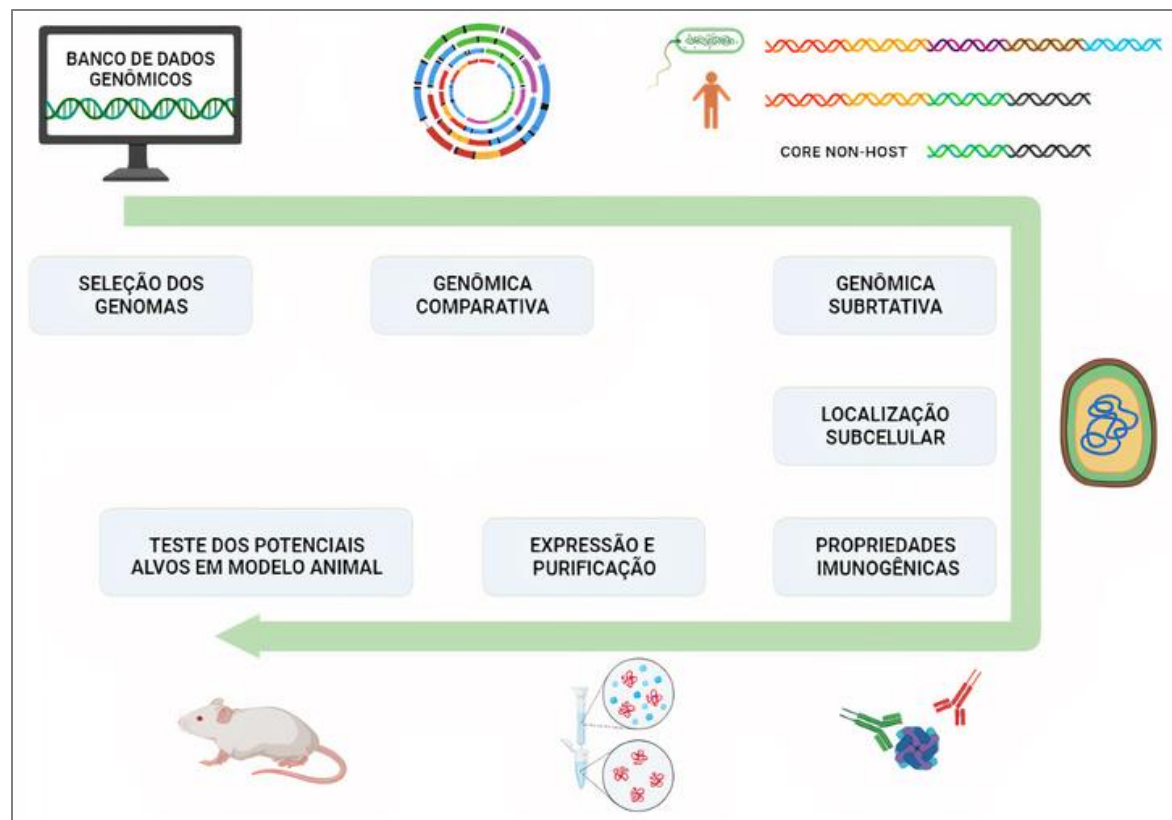


Figura 4. Vacinologia Reversa
A imagem abrange os principais passos da Vacinologia Reversa, desde a obtenção dos dados genômicos aos testes em modelos animais.
Fonte: Imagem gerada com BioRender (biorender.com)

1.4.2 Imunoinformática

A ampla produção de dados relacionados ao sistema imunológico, seus componentes e interações gerou uma demanda de softwares específicos para analisar esse tipo de material. Esses fatos corroboraram para o surgimento da imunoinformática, que se constitui de ferramentas e algoritmos para, entre outros, distinguir epítopos ligantes de linfócitos T e B, de epítopos não ligantes (BACKERT; KOHLBACHER, 2015). Dentre as grandes subáreas dessa abordagem, se destacam a predição de epítopos para design de proteínas quiméricas com potencial de serem usadas em testes diagnósticos e vacinas. (TOMAR; DE, 2010).

As aplicações da imunoinformática no desenvolvimento de vacinas têm aguçado a construção de vacinas multi-epitópo, que possuem grande potencial de enfrentar os atuais desafios para desenvolvimento de vacinas contra patógenos de importância clínica. As vacinas multi-epítopos baseadas em abordagens de imunoinformática, são compostas por uma proteína quimérica formada por epítopos selecionados por diversas técnicas e geralmente um adjuvante, conectados por peptídeos de ligação, ou *linkers* (NEGAHDARIPOUR *et al.*, 2018). Os peptídeos de ligação, além de conectarem os epítopos, auxiliam na manutenção da atividade biológica (IKEBE et al., 1998), no dobramento e estabilidade proteica, além de

dar assistência no processo de expressão (CHEN; ZARO; SHEN, 2013) e em certos casos pode melhorar a indução de respostas imunes (HU et al., 2004).

Dentre as vantagens sobre as vacinas de epítopo único e clássicas, estão por exemplo: 1) O acúmulo de vários epítopos identificados a partir de antígenos específicos, 2) vários receptores de células T podem identificar determinantes nas vacinas a partir das várias moléculas de MHC presentes no hospedeiro, 3) essas vacinas podem ser confortavelmente moduladas para aumentar a imunogenicidade, 4) também podem ser construídas para ativar respostas imunológicas mediadas por anticorpos e células, na medida em que seus epítopos de linfócitos T auxiliares são sobrepostos aos epítopos de células T citotóxicas e epítopos de células B, e 5) antígenos indesejados podem ser removidos para reduzir a possibilidade de consequências adversas ou respostas imunológicas que provavelmente causam efeitos colaterais (OLI *et al.*, 2020).

Através desta abordagem, a fusão de epítopos específicos para induzir respostas imunes eficazes implica no desenvolvimento de respostas imunes direcionadas, levando a um risco menor de induzir reações adversas. Além disso, não há possibilidade de integração genética como nas vacinas de DNA e nenhum risco de reversão para a forma virulenta como nas vacinas atenuadas. Outros benefícios potenciais importantes sobre vacinas multi-epítopos estão relacionados à produção, armazenamento e transporte. A possibilidade de produzir essas vacinas na forma liofilizada facilita o manejo e custos para sintetizar peptídeos também estão diminuindo com o avanço das tecnologias (MOYLE; TOTH, 2008; PURCELL; MCCLUSKEY; ROSSJOHN, 2007).

O grande desafio das vacinas de epítopos está relacionado à imunogenicidade. Peptídeos sintéticos tendem a ser menos imunogênicos devido ao seu baixo tamanho e quantidade de determinantes imunogênicos portanto, para administrar esse viés, são utilizados adjuvantes para melhorar a capacidade imunogênica (AZMI *et al.*, 2014; LEI *et al.*, 2019). Apesar de ainda não haver vacina multi-epítopo disponível no mercado contra pneumonia, devido às dificuldades de estabilidade, entrega e quanto à variabilidade imunogenética humana, novas abordagens para contornar essas características estão em estudo (PURCELL; MCCLUSKEY; ROSSJOHN, 2007).

Com os avanços tecnológicos no desenvolvimento no geral de vacinas, tornou-se mais acessível a busca por imunógenos multi-epítopos a partir de antígenos de proteínas pertencentes a patógenos de difícil cultivo. Esses fatos, contribuíram para a determinação desse trabalho.

1.4.2 Imunologia e resposta a vacinas multi-epítopo

As respostas imunológicas são divididas em sistema inato e sistema adaptativo. A resposta imune humoral e a resposta imune citotóxica compreendem os dois braços do

sistema adaptativo. O primeiro contato com as proteínas estranhas é feito pelo sistema inato, além disso as células responsáveis por este contato têm uma enorme lacuna na apresentação de antígenos ao sistema adaptativo. (TURVEY; BROIDE, 2010).

Todas as células nucleadas podem processar proteínas intracelulares a fim de apresentar epítopos através do complexo principal de histocompatibilidade classe I para células CTL. Células apresentadoras de antígenos especializadas (APCs), como macrófagos e células dendríticas, pertencem à imunidade inata e têm um papel importante para preencher a lacuna entre a imunidade inata e adaptativa. Essas células processam proteínas extracelulares e apresentam os epítopos para células T auxiliares (ou TCD4 +) por meio do complexo principal de histocompatibilidade de classe II. O reconhecimento das moléculas de MHC mediadas pelos linfócitos T auxiliam na propagação de respostas imunes específicas contra o antígeno e no controle adicional da infecção (KOTSIAS; CEBRIAN; ALLOATTI, 2019).

As células CTL atuam destruindo as células infectadas diretamente. As células B são as principais responsáveis pela atividade humoral devido à sua produção de anticorpos. As células T helper (Th) ou CD4 + participam tanto da produção de anticorpos quanto da proliferação de CTL, pois produzem citocinas que modulam e melhoram as respostas imunológicas para células B completas reconhecerem antígenos por meio de receptores de células B (BCRs). A ativação das células B depende do reconhecimento direto do receptor BCR com os epítopos do patógeno e por meio de co-estimuladoras nas células Th e B. Eles também respondem a citocinas secretadas para células Th e expressam receptores para citocinas que auxiliam na proliferação de células B (JONES, 2005; LEI *et al.*, 2019).

As vacinas multi-epítopos são projetadas com um conjunto selecionado de epítopos pertencentes a um ou mais patógenos. A proteína quimérica construída é composta por epítopos de células B, epítopos de células T auxiliares (HTL) e epítopos de linfócitos T citotóxicos (CTL) (LEI *et al.*, 2019).

A fim de prever os alvos potenciais da vacina, diferentes abordagens de vacinologia reversa foram testadas. Um deles utilizou 88 genomas de *M. pneumoniae* para identificar proteínas de superfície não similares as proteínas do hospedeiro, com características de essencialidade e potencial de antigenicidade para selecionar as melhores vacinas candidatas (VILELA RODRIGUES *et al.*, 2019).

Usando abordagens imunoinformáticas, Unni e colaboradores 2019, desenvolveram uma vacina de peptídeo baseada em epítopos contra *M. pneumoniae*, usando 12 proteínas de membrana e 5 proteínas de citoaderência como conjunto de dados. Os epítopos previstos foram submetidos à análise de interações com moléculas HLA por meio dos escores de energia de ligação e os três melhores epítopos foram direcionados à simulação dinâmica para garantir essa afinidade (UNNI *et al.*, 2019).

Outro estudo com o mesmo aspecto foi proposto por Mahmood e colegas, no entanto, eles usaram o proteoma de uma cepa M129 de *M. pneumoniae* para selecionar as proteínas não homólogas com propensão antigênica como alvos da vacina. Eles previram os epítopos e construíram uma vacina candidata de múltiplos epítopos com um adjuvante. O docking molecular com o receptor Toll-like 4 e simulação dinâmica também foram realizados (MAHMOOD *et al.*, 2021).

Devido a esses achados promissores, uma vacina para prevenir infecções por *M. pneumoniae* é necessária, além disso, está vacina também pode avançar na prevenção da asma e PAC (BAJANTRI; VENKATRAM; DIAZ-FUENTES, 2018).

2. JUSTIFICATIVA

Diante da atual preocupação mundial acerca dos dados epidemiológicos referentes a PAC, principalmente devido ao crescimento da resistência antimicrobiana e as altas taxas de mortalidade, a necessidade de medidas de prevenção eficazes se torna mais evidente. Nesse contexto, o *Mycoplasma pneumoniae* é um dos principais patógenos associados e tem sido detectado de forma mais expressiva nos últimos anos, possivelmente devido ao sucesso da vacina pneumocócica e pelos avanços nos métodos diagnósticos.

Ainda não existe nenhuma vacina disponível para prevenção de infecções por *M.pneumoniae* ademais, são restritos os métodos de tratamento para essa bactéria devido às suas características morfofisiológicas e a existência de muitas linhagens resistentes a certas classes de antibacterianos. Essas peculiaridades no metabolismo e a arquitetura estrutural do *M.pneumoniae* também são obstáculos significativos no cultivo e manuseio da bactéria em laboratório. Sendo assim, abordagens *in silico* demonstram ser alternativas viáveis e de eficácia para busca de novos alvos vacinais contra patógenos fastidiosos. Essa abordagem além de agilizar o processo para determinação de uma vacina, diminui os recursos financeiros necessários para seu desenvolvimento reduzindo também de forma significativa, a quantidade de animais submetidos aos experimentos.

Portanto, o presente estudo visou determinar através de ferramentas de bioinformática, proteínas candidatas à alvos de drogas, vacinais e um imunógeno multi-epítopo de ampla cobertura para prevenir infecções causadas pelo *M. pneumoniae*.

3. OBJETIVOS

3.1 Objetivo principal

Predizer por meio de Vacinologia Reversa potenciais proteínas candidatas à alvos vacinais e de drogas e desenvolver *in silico* um imunógeno multi-epítopo de alta cobertura contra *Mycoplasma pneumoniae* através de ferramentas de imunoinformática.

3.2 Objetivos específicos

- Selecionar os genes do genoma central e que não estejam presentes no hospedeiro (*core-non-host*)
- Caracterizar as proteínas codificadas pelo *core-non-host* de acordo com a localização subcelular
- Avaliar as proteínas de superfície e extracelulares quanto seu potencial imunogênico
- Realizar a predição de ilhas de patogenicidade e verificar se alguma das proteínas candidatas se encontra dentro dessas ilhas
- Identificar epítopos específicos de células B, CTL e HTL
- Avaliar os epítopos quanto a antigenicidade, potencial tóxico e alérgico, cobertura populacional e imunogenicidade
- Construir a sequência de aminoácidos da proteína quimérica do imunógeno e realizar a modelagem estrutural
- Avaliar a interação da vacina com o sistema imune do hospedeiro através de ferramentas de bioinformática

4. CAPITULO I

Reverse vaccinology and subtractive genomics reveal new therapeutic targets against Mycoplasma pneumoniae: a causative agent of pneumonia

Thaís Cristina Vilela Rodrigues, Arun Kumar Jaiswal, Alissa de Sarom, Letícia de Castro Oliveira, Carlo José Freire Oliveira, Preetam Ghosh, Sandeep Tiwari, Fábio Malcher Miranda, Leandro de Jesus Benevides, Vasco Ariston de Carvalho Azevedo and Siomar de Castro Soares

O *Mycoplasma pneumoniae* é um patógeno comumente associado à CAP. Devido às particularidades morfológicas e ao aumento da resistência aos antimicrobianos, as opções para tratar esse tipo de infecção são cada vez menores. Dessa forma, métodos de prevenção e tratamento demonstram ser opções necessárias de se buscar, para esse tipo de infeção. A fim de prever novos alvos de vacinas e de drogas contra *M. pneumoniae*, ferramentas *in silico* da abordagem de vacinologia reversa foram utilizadas. Os genomas de 88 cepas de *M.pneumoniae* foram submetidos a diversas análises como a genômica comparativa, para encontrar genes do genoma central (core genome), mas que não apresentam similaridade com o genoma humano. As proteínas selecionadas por meio da análise, possibilitaram avaliar sua capacidade de ligação ao complexo MHC, entre outros, encontrando 8 alvos com potencial imunogênico e cinco prováveis alvos de drogas.

ROYAL SOCIETY OPEN SCIENCE

royalsocietypublishing.org/journal/rsos

Research

Cite this article: Vilela Rodrigues TC *et al*. 2019 Reverse vaccinology and subtractive genomics reveal new therapeutic targets against *Mycoplasma pneumoniae*: a causative agent of pneumonia. *R. Soc. open sci.* **6**: 190907. http://dx.doi.org/10.1098/rsos.190907

Received: 17 May 2019
Accepted: 4 July 2019

Subject Category:
Genetics and genomics

Subject Areas:
bioinformatics/genomics/immunology

Keywords:
vaccinology, *Mycoplasma pneumoniae*, bioinformatics, genomic, molecular docking, pneumonia

Author for correspondence:
Siomar de Castro Soares
e-mail: siomars@gmail.com

†These authors contributed equally to this study.

Electronic supplementary material is available online at https://dx.doi.org/10.6084/m9.figshare.c.4593944.

THE ROYAL SOCIETY PUBLISHING

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

Reverse vaccinology and subtractive genomics reveal new therapeutic targets against *Mycoplasma pneumoniae*: a causative agent of pneumonia

Thaís Cristina Vilela Rodrigues[1,†], Arun Kumar Jaiswal[1,2,†], Alissa de Sarom[1], Letícia de Castro Oliveira[1,2], Carlo José Freire Oliveira[1], Preetam Ghosh[3], Sandeep Tiwari[2], Fábio Malcher Miranda[2], Leandro de Jesus Benevides[4], Vasco Ariston de Carvalho Azevedo[2] and Siomar de Castro Soares[1]

[1]Department of Microbiology, Immunology and Parasitology, Federal University of Triângulo Mineiro, Minas Gerais, Brazil
[2]Department of Genetics, Ecology and Evolution, Federal University of Minas Gerais, Minas Gerais, Brazil
[3]Department of Computer Science, Virginia Commonwealth University, Richmond, VA 23284, USA
[4]Bioinformatics Laboratory - LABINFO, National Laboratory of Scientific Computation - LNCC/MCTI, Rio de Janeiro, Brazil

TCVR, 0000-0002-2048-5522; AdS, 0000-0002-3146-7784; LdCO, 0000-0002-2036-4456; PG, 0000-0003-3880-5886; ST, 0000-0002-8554-1660; FMM, 0000-0002-6823-5995

Pneumonia is an infectious disease caused by bacteria, viruses or fungi that results in millions of deaths globally. Despite the existence of prophylactic methods against some of the major pathogens of the disease, there is no efficient prophylaxis against atypical agents such as *Mycoplasma pneumoniae*, a bacterium associated with cases of community-acquired pneumonia. Because of the morphological peculiarity of *M. pneumoniae*, which leads to an increased resistance to antibiotics, studies that prospectively investigate the development of vaccines and drug targets appear to be one of the best ways forward. Hence, in this paper, bioinformatics tools were used

© 2019 The Authors. Published by the Royal Society under the terms of the Creative Commons Attribution License http://creativecommons.org/licenses/by/4.0/, which permits unrestricted use, provided the original author and source are credited.

for vaccine and pharmacological prediction. We conducted comparative genomic analysis on the genomes of 88 *M. pneumoniae* strains, as opposed to a reverse vaccinology analysis, in relation to the capacity of *M. pneumoniae* proteins to bind to the major histocompatibility complex, revealing seven targets with immunogenic potential. Predictive cytoplasmic proteins were tested as potential drug targets by studying their structures in relation to other proteins, metabolic pathways and molecular anchorage, which identified five possible drug targets. These findings are a valuable addition to the development of vaccines and the selection of new *in vivo* drug targets that may contribute to further elucidating the molecular basis of *M. pneumoniae*–host interactions.

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

1. Introduction

The genus *Mycoplasma* belongs to the class Mollicutes, which are bacteria without a cell wall. Through comparative genomics and phylogenetic analysis, it has been suggested that these bacteria probably originated from Gram-positive ancestors [1–3]. Bacteria of the genus *Mycoplasma* are the smallest known microorganisms in cell and genomic size with a capacity for self-replication [4]. Owing to the fact that the *Mycoplasma* genome comprises fewer than 1000 genes, its metabolic capacities are reduced, thus it requires specific cellular compounds for its survival. For their growth in culture media, these microorganisms require singular components such as sterols that protect them against their osmotic fragility. The absence of a cell wall in the genus *Mycoplasma* makes it difficult to classify these microorganisms as cocci or bacilli; furthermore, this characteristic confers a natural resistance to β-lactams and impairs Gram staining [5,6]. Among the 16 *Mycoplasma* species that infect humans, six of them are pathogenic and, of these, *Mycoplasma pneumoniae* is the one with the highest clinical significance [7,8].

Mycoplasma pneumoniae, the main causative agent of community-acquired pneumonia (CAP), has a genome of approximately 800 000 base pairs that encodes approximately 700 different proteins [9]. Studies have demonstrated a prevalence rate of CAP for *M. pneumoniae* that can reach 40% among confirmed cases [10]; however, it is still extremely difficult to make a specific diagnosis in asymptomatic cases. CAP is an acute lung infection responsible for high morbidity and mortality rates and tends to be contracted by individuals outside the healthcare system [11]. In most cases, the pathogen is not identified and the diagnosis is sometimes based only on clinical signs; consequently, the treatment is not specific, and therapeutic intervention may compromise the patient's life and even contribute to the development of antibiotic-resistant bacteria [12].

Because of the absence of the cell wall, infections caused by *M. pneumoniae* are treated using antibiotics from the macrolide, tetracycline and quinolone classes [13,14]. It is known that a mutation in the 23S rRNA gene conferred resistance to macrolides in several strains, reducing treatment options to only two classes of antibiotics. Asia is the continent with the highest rate of resistance, where about 90% of the strains of this bacterium have already demonstrated this mutation. This high rate was due to the frequent and indiscriminate antibiotic use in this region. Cases of resistance in Europe and America have also been reported and the proportion of resistant strains has been growing progressively. Thus, there is a need to promote not only restrictions for the use of macrolides, but also advancement of research for the development of new drugs and prophylactic methods [15–17]. There are currently vaccines against *Streptococcus pneumoniae* and *Haemophilus influenzae*, the most common causative agents of pneumonia. On the other hand, for atypical pathogens that cause this disease, such as *M. pneumoniae* and *Chlamydia pneumoniae*, studies are still needed to achieve the same level of progress [18]. Such studies can be extremely important since, according to the World Health Organization [19], pneumonia vaccines prevent the deaths of almost 2 million children per year [20].

With the advancement of genome sequencing technologies, the number of microorganism species with completed genome sequences has increased rapidly, adding thousands of new sequenced genomes to databases and providing material for numerous types of studies, including the prediction of new drug targets and vaccines, through approaches such as comparative and subtractive genomics [21–23]. Another approach is reverse vaccinology, which is also based on the use of the genomic sequence of a given microorganism for *in silico* analysis. Using this bioinformatics tool, it is possible to investigate, in several ways, all of the proteins that can be produced by the bacterium and evaluate their ability to induce an adaptive immune response or to bind to drugs [24,25]. Reverse vaccinology optimizes the prediction of drug and vaccine targets, especially for microorganisms that are difficult to grow in the laboratory, such as intracellular bacteria including *M. pneumoniae*. In addition, reverse vaccinology allows the simultaneous analysis of targets in multiple genomes, which is important for the predicted targets to achieve greater

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

coverage among lineages [26]. The use of reverse vaccinology gained prominence in 1995 with the publication of the complete genome of the bacterium *H. influenzae* [27]. Subsequently, this methodology was used for the screening of antigens to prepare vaccines against *Neisseria meningitidis* [28,29], *Acinetobacter baumannii* [30], *Streptococcus agalactiae* [31], human cytomegalovirus, respiratory syncytial virus, human immunodeficiency virus, influenza and dengue virus [32,33], among other pathogenic agents.

Given the relevance of the problem, the objective of this work was to predict drug and vaccine targets through bioinformatics tools, by selecting those that act against all 88 *M. pneumoniae* strains whose genomes are already deposited in GenBank. This study will facilitate future *in vitro* and *in vivo* tests for the production of drugs and prophylactic targets against a species of bacteria with high clinical relevance for CAP, especially among children.

2. Methodology

2.1. Genomes

The 88 genomes of *M. pneumoniae* strains available in the GenBank database were downloaded through the National Center for Biotechnology Information (NCBI) for the bioinformatics analysis. For this, both the complete and incomplete downloaded genomes were converted to the FASTA format.

2.2. Identification of conserved proteins of *M. pneumoniae* and subtractive genomics

The FASTA format files containing the amino acid sequences were submitted to the software OrthoFinder under its default parameters. The algorithm developed for this software performs calculations based on searches through BLAST and the MCL clustering algorithm to identify the regions of homology, thereby generating the orthogroups with the protein sequences. Subsequently, in-house scripts were employed to classify genes into three groups: core genes, which represent those present in all studied strains; shared genes, which are present in some, but not all, strains; and the singletons, which are strain-specific genes present in only one strain [34]. With the amino acid sequences (faa), a BLASTp was performed against the proteins of the human genome, also using the OrthoFinder, to identify the proteins belonging to the bacterium *M. pneumoniae* that have no homology with those from the host. This stage is called subtractive genomics and was essential to avoid the selection of drug targets or vaccines without protective effect or even those that may cause autoimmunity [35].

2.3. Characterization and prediction of the subcellular location of proteins of *M. pneumoniae*

To verify the importance of each of the identified proteins, we used the Database of Essential Genes (DEG), which includes all essential bacterial and eukaryotic gene records [36]. This online platform contains information on genes from bacteria, archaea and eukaryotes, responsible for the production of several proteins, as well as data from non-coding RNAs (http://www.essentialgene.org/). Only proteins considered essential to the microorganism were used in the prediction of candidate vaccine antigens and drug targets. SurfG+ is a software that predicts the subcellular localization of the proteins of interest. The prediction consists of identifying peptide signal, retention signals, transmembrane helices and secretion pathways to classify proteins as secreted, PSE (putatively exposed to the surface) and membrane proteins. Among the identified proteins, the cytoplasmic proteins were subjected to an analysis for potential drug targets, because of their involvement in the basic survival processes of the organism, while proteins characterized as membrane, PSE and secreted were directed to reverse vaccinology analysis, since they are the first proteins to come into contact with the immune response of the host [37].

2.4. Selection of drug targets and druggability analysis

The MHOLline program was used to model three-dimensional (3D) cytoplasmic proteins. This software combines other programs such as HMMTOP, BLAST, BATS, MODELLER and PROCHECK to analyse and classify potential drug targets according to their structural quality. BLAST performs a random search against protein databases (PDBs) and provides three-dimensional structures of the targets. The BATS (Blast Automatic Targeting for Structures) program selected the proteins in which the comparative modelling technique was applied and rated the models in seven groups according to quality (from 'very high' to 'very low'). Three-dimensional models and global alignment were

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

produced by the MODELLER program and evaluated for stereochemical quality through PROCHECK. To complement the process, transmembrane helix topology studies were performed by the software HMMTOP. The BATS program organized the BLAST output files into four groups—G0, G1, G2 and G3—following the criteria: G0, non-aligned sequence; G1, $E > 10 \times 10^{-5}$ or identity < 15%; G2, $E \leq 10 \times 10^{-5}$, identity ≥ 25% and length variation index (LVI) ≤ 0.7; G3, $E \leq 10 \times 10^{-5}$, identity ≤ 15% to <25% or LVI > 0.7 [38]. The 3D structures with identity < 25%, corresponding to groups G1 and G3, did not fit into the comparative modelling technique of the MHOLline program, and, thus, only the G2 group sequences were submitted to the next stages of docking.

Furthermore, for the druggability analyses, the final lists of drug target proteins were subjected to DoGSiteScorer. The DoGSiteScorer is a web-based automated pocket detection and analysis tool for calculating the druggability of protein cavities. For each detected cavity, the tool returns the pocket residues and a druggability score ranging from 0 to 1. Values closer to 1 indicate highly druggable protein cavities, i.e. the predicted cavities are likely to bind ligands with high affinity [39]. The druggable cavity for each target with value greater than 0.8 was used for the docking analysis.

2.5. Ligand library preparation and docking analysis

The ligand library of ZINC drug-like molecules (Natural Product and its derivatives) was downloaded from the ZINC database [40,41]. The 5008 ligands obtained in .SDF format were then converted into .PDB by using the OpenBabel (v. 2.4.1) tool [42]. After converting the file into .PDB format, the Gasteiger atomic partial charges were assigned to convert all the ligand compounds to the PDBQT format by using the prepare_ligand4.py script on the terminal. Furthermore, for the docking analysis, the final identified drug target proteins' 3D structures were examined and converted to the required PDBQT format using the AutoDockTools MGL tool (v. 1.5.4) [43]. A grid box parameter for each target comprising the residues of the DoGSiteScorer [44] druggable pocket with drug score greater than 0.8 was used for virtual screening of the ligand using AutoDock Vina [45]. The top 10 ranked ligand molecules were identified by virtual screening using the Python script topmolecule.py. Furthermore, the flexible docking was performed with the identified top 10 molecules, keeping the residues obtained from DoGSiteScorer for each target. The 3D poses of docked molecules were analysed in Chimera [46], whereas Pose View was used for two-dimensional (2D) representation [47].

2.6. Selection of vaccine targets

In order to test the adhesion and binding capacity to major histocompatibility complex (MHC) class I and class II, all membrane, secreted and PSE protein targets of *M. pneumoniae* were submitted to the Vaxign tool, a system based on genomic features for predicting vaccine targets in the reverse vaccinology platform. In this software, we used default parameters except for subcellular localization and transmembrane helices that were already predicted by means of SurfG+. This system has tools to identify the subcellular localization of the product of the studied sequences, analyses the transmembrane helices and is able to exclude the sequences present in non-pathogenic strains. SPAAN is a program with sensitivity of 89% and specificity of 100% that evaluates the adhesion capacity of the targets, establishing a cut-off of 0.51. The prediction of MHC-I- and MHC-II-binding epitopes is performed by Vaxitope, which searches the Immune Epitope Database (IEDB) and calculates the affinity of each molecule [48].

Sequences of the 46 PSE, secreted or membrane proteins were submitted to this platform in the FASTA format for analysis of antigenic properties. This presents the option 'Dynamic Vaxign Analysis', which is configured as the desired parameter for the prediction based on the binding capacity to MCH-I and MCH-II. Thus, proteins theoretically arising from the 88 genomes with adhesion capacity greater than 0.51 were considered immunogenic and selected for further analysis [49].

The sequences of the proteins with good MHC-binding capabilities were subsequently subjected to B-cell epitope prediction analysis to verify their ability to develop humoral immune responses. For this, we used the IEDB with a threshold of 0.5. In the platform, it is possible to analyse the proteins of interest to find the main epitopes and the value of each residue [50].

2.7. Analysis of proteins of interest and their interactions

To understand the metabolic interactions of the proteins of interest, we used the STRING program with default parameters, which shows the specific interactions between the proteins of *M. pneumoniae* and those present in its database, allowing the pathways' functional activities to be understood in greater

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

detail. For each protein–protein interaction, a score is generated. These scores represent the confidence interval ranging from 0 to 1, with 1 being the highest probability of the interaction being true. In addition to the STRING platform, other platforms were also used to contribute and reinforce the identification of the functions and metabolic pathways of proteins of *M. pneumoniae* [51].

In summary, we used the Universal Protein Resource (UniProt), which is a protein sequence and annotation database [52]. Proteins that had signal peptide were directed to the secretory pathway and were identified using the SignalP program that located the cleavage sites of each signal peptide [53]. To predict transmembrane helices, we submitted the amino acid sequences of each *M. pneumoniae* protein to the TMHMM server, which predicted the topology of these proteins by the Markov method [54].

To find out whether any of the proteins had already been tested for drug targets in previous studies, DrugBank searches were performed. DrugBank is an online database that contains information about drugs, their binding targets, interactions with other drugs, and their relationships with metabolism, gene expression and protein. Potential drugs being tested in clinical trials are also found on this platform [55].

It is well known that the use of antibiotics affects the human microbiota and is associated with immunological and metabolic alterations detrimental to the normal functioning of the organism [56]. To determine whether the proteins investigated in this study are also part of the metabolism of some of the bacteria most commonly found in the intestinal microbiota, BLASTp was performed through NCBI. Each potential drug target was submitted to the platform and compared with the bacterial protein sequences of the genera *Bacillus*, *Lactobacillus* and *Streptococcus*, which are some of the major genera found in the gut [57].

2.8. Analysis of genomic similarities and phylogenetic reconstruction

To compare the 88 genomes studied and to understand the differences present in each strain of *M. pneumoniae* that could enable future identification of reference genomes for pathogenicity island prediction, we used the Gegenees [58] tool; this tool fragments genomes at predefined sizes and makes an alignment of all fragments against all using the tools BLASTn, tBLAST and FASTA. With the data from this alignment, a heat map is generated that demonstrates the similarity between the lines and that ranges from 0% to 100%. The results from the Gegenees software were exported in the 'Nexus' format for later phylogenetic reconstruction using the software SplitsTree4, by using the neighbour-joining method [59].

2.9. Prediction of genomic islands

Prediction of genomic islands was carried out in order to identify the existence of potential drug and vaccine targets within these islands. Genomic Island Prediction Software (GIPSy) was used for the prediction of genomic islands that were classified as follows: (i) pathogenicity islands, which contain virulence factors; (ii) metabolic islands, with genes related to proteins important for metabolic pathways; (iii) resistance islands, which have genes involved in the processes of resistance to antibiotics; and (iv) symbiotic islands, with genes coding for proteins that allow the symbiotic interaction of the bacterium with the host. The characteristics analysed for predicting whether a given genome region is a genomic island were: deviations in genome signature (GC content and codon usage); the presence of transposase genes, high concentrations of virulence factors, genes related to antibiotic resistance, metabolic pathways and symbioses for pathogenicity, resistance, and metabolic and symbiotic islands, respectively; the presence of insertion sequences or flanking tRNA genes; and size ranging from 6 to 200 kb [60].

The genomes used in this step were selected from the results of the phylogenetic analysis. The phylogenetically closest lineages according to the SplitsTree [56] program were organized into clusters and, from that result, 15 genomes were selected for prediction analyses of genomic islands. The genome of the species *Mycoplasma gallinarum*, which is phylogenetically close to the species *M. pneumoniae* but is not pathogenic to humans, was selected as a reference in predicting the islands. The results obtained by GIPSy were later plotted in a circular figure using the software BRIG [61,62].

3. Results

The key steps for target identification, the methodologies used and the total number of proteins described in each step are summarized in the workflow of figure 1.

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

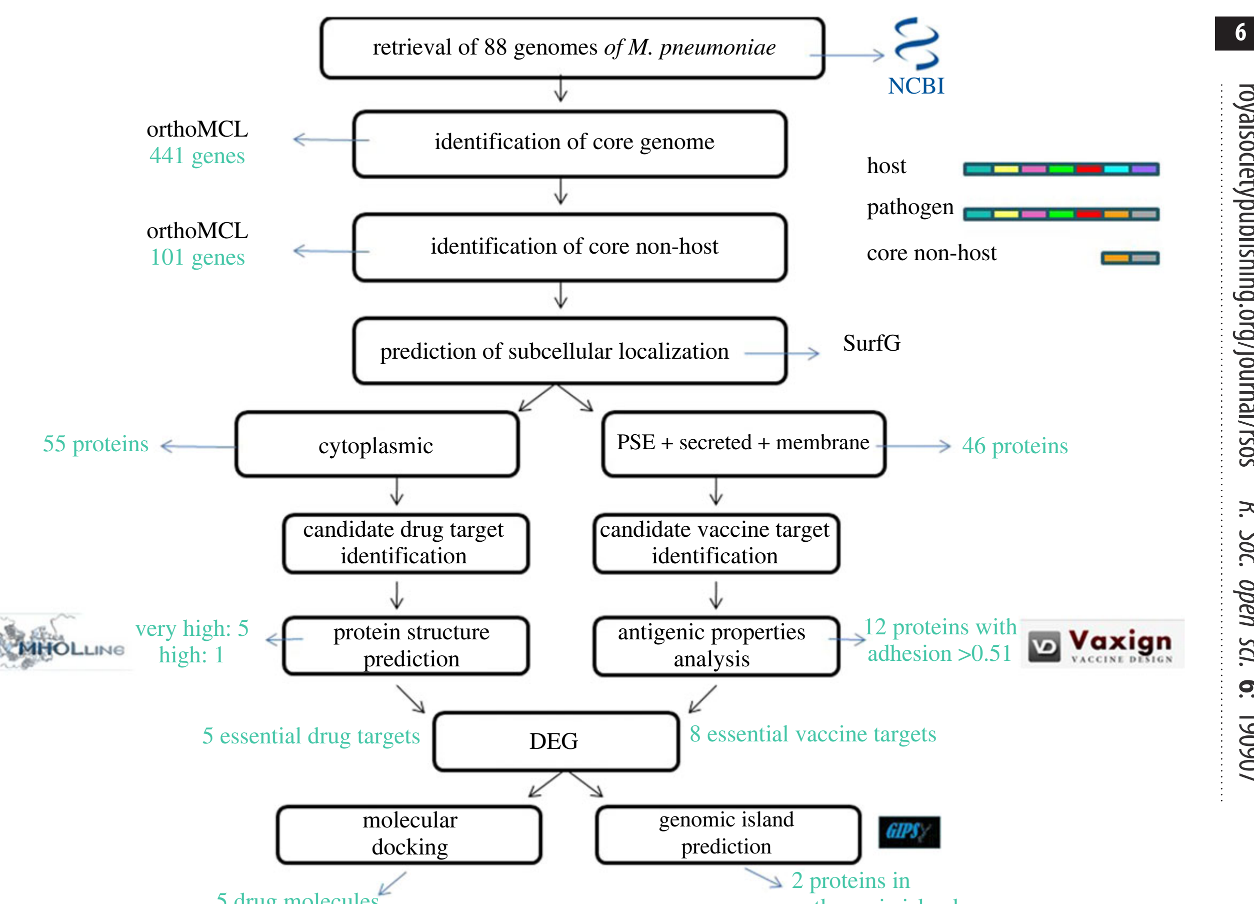


Figure 1. Workflow projected with the methodologies used and the total number of proteins identified in each step.

3.1. Identification of *M. pneumoniae* conserved proteins and subtractive genomics

Using the software OrthoFinder, we found 441 genes belonging to the core, 289 shared genes and 50 singletons. After the subtractive genomic analysis with these core genes, the number was reduced from 441 to only 101 potential targets.

3.2. Localization of target proteins

As described previously, for the prediction of protein localization we employed the software SurfG+. From the 101 targets, 55 proteins were predicted as cytoplasmic and directed to drug targeting. The other 46 proteins, considered PSE, secreted or from membrane were directed to analyses for vaccine targets (table 1).

3.3. Drug target identification and druggability analysis

The proteins predicted as cytoplasmic and essential for the bacteria are frequently considered good candidates for drug targets [35]. Thus, the protein sequences predicted as cytoplasmic were submitted to MHOLline, which used the HMMTOP, BLAST, BATS, MODELLER and PROCHECK software to predict 3D modelling. Based on the drug target analysis, only those proteins from the G2 group ($E \leq 10 \times 10^{-5}$, identity $\geq 25\%$ and LVI ≤ 0.7) were selected. Five proteins with very high classification were identified within the program criteria and one with high potential (electronic supplementary material, file S1). The other proteins with lower level of quality were discarded. These six proteins were also submitted to the DEG and only five of them were considered vital for *M. pneumoniae*, following the criteria of bit score of 100 and *E*-value with a cut-off of 1×10^{-4} (table 2).

The five proteins identified as potential candidates for drug action were ribosome-binding factor A (WP_010874513.1), division/cell-wall cluster transcriptional repressor MraZ (WP_010874670.1), dTIGR00282 family metallophosphoesterase (WP_010874705.1) and the hypothetical proteins WP_010874779.1 and WP_014325598.1.

Ribosome-binding factor A (WP_010874513.1) is one of the most important bacterial proteins that assist in the late stages of maturation of the 30S ribosomal subunit. Furthermore, such a protein is

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

Table 1. Localization of target proteins.

location	number of proteins
cytoplasmic	55
PSE	15
secreted	3
membrane	28
total	101

essential for the efficient processing of 16S rRNA and may interact with the 5′-terminal helix region of 16S rRNA (116 aa). Its metabolic interactions occur through interactions with the following proteins: (i) translation initiation factor (IF); (ii) phenylalanine–tRNA ligase; (iii) translation elongation factor; and (iv) ribosomal protein S15. Ribosomal protein S15 is one of the major rRNA-binding proteins that binds directly to the 16S rRNA, where it assists in the assembly of the 30S subunit platform by ligating and joining several helices of 16S rRNA [63]. Aminoglycosides are antibiotics that bind to the 30S ribosomal subunit, causing base modifications that consequently modify codon reading by interfering with mRNA translation [64]. Thus, although no ribosome-binding factor A (WP_010874513.1) studies have been found as a potential drug target, it is believed that interfering with its binding to the 30S subunit may interfere with its structural function and lead to transcriptional errors that may affect bacterial protein synthesis (electronic supplementary material, files S2 and S3).

Division/cell-wall cluster transcriptional repressor MraZ (WP_010874670.1) is a DNA-binding transcription factor, which interacts with (i) HrcA transcription repressor, which is a negative regulator of class I heat shock genes (operons grpE-dnaK-dnaJ and groELS); (ii) protein B of segregation and condensation, which acts on the chromosomes during cell division; (iii) IF-3, which binds to the 30S ribosomal subunit, increasing the availability of those subunits in which the initiation of the protein synthesis begins; (iv) chaperone proteins that prevent aggregation of stress-depleted proteins; and (v) protein RecA, which can catalyse the hydrolysis of ATP in the presence of single-stranded DNA (electronic supplementary material, files S4 and S5).

WP_010874705.1 is a protein of the metallophosphoesterase family, which is related to DNA repair (electronic supplementary material, files S6 and S7) [65]. WP_010874779.1 is a hypothetical protein whose function remains incompletely elucidated, but, according to STRING's predictions, it participates in interactions with chromosomal segregation proteins, carrier proteins and endonucleases (electronic supplementary material, files S7 and S9). Finally, WP_014325598.1 is also a hypothetical protein. It is related not only to the folding and transport of proteins, but also to tRNA ligase of threonine and arginine in addition to the tRNA responsible for thiamine synthesis (electronic supplementary material, files S10 and S11).

When comparing these proteins with the proteome of a group of bacteria present in the intestinal microbiota (*Bacillus/Lactobacillus/Streptococcus* group) through BLAST NCBI, we observed that three of the five potential drug targets present a protein profile similar to those of the database. 30S ribosome-binding factor showed 28% identity with the protein 30S ribosome-binding factor RbfA, which is present in *Lactobacillus sanfranciscensis*, and 24% identity with the same protein in *Lactobacillus pantheris*. The transcriptional regulator MraZ showed identity of about 40% with a series of *Bacillus* species, a result similar to that found through the BLAST analysis of the protein dTIGR00282 from the metallophosphoesterase family.

3.4. Molecular docking and virtual screening

Natural products have played important roles in recent drug development, where an enormous number of natural product-derived compounds in various stages of clinical development were highlighted [66]. For each target protein, 5008 drug-like compounds (Natural Product and its derivatives) were screened from the ZINC database. The top 10 compounds obtained by means of the AutoDock Vina binding affinity score (electronic supplementary material, file S12) were further used for flexible docking analysis with the residues of the most druggable cavity identified by DoGSiteScorer (table 4). As a result, the predicted protein–ligand interactions for best ligands with each target are displayed in

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

Table 2. Drug targets.

target	ID	name	gene UniProt	length (aa)	molecular weight (Da) UniProt	structural quality MHOLline	biological process
1	WP_010874513.1	ribosome-binding factor A	rbfA	116	13 389	very high	maturation of the functional nucleus of the 30S ribosomal subunit
2	WP_010874670.1	transcriptional regulator MraZ	MraZ	141	16 335	very high	division/cell-wall cluster transcriptional repressor MraZ
3	WP_010874705.1	dTIGR00282 family metallophosphoesterase	MPNE_0406	281	31 431	very high	metal ion binding
4	WP_010874779.1	hypothetical protein MPN423	MPN_423	129	14 939	very high	hydrolase activity, metal ion binding
5	WP_014325598.1	hypothetical protein	MPN_555	193	22 434	very high	protein folding protein transport

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

Table 3. Identified druggable pocket with its volume, surface area and drug score of each target protein obtained from DoGSiteScorer.

protein name	volume (Å^3)	surface area (Å^2)	drug score	residues
30S ribosome-binding factor (WP_010874513.1)	1125.38	1672.38	0.82	TYR1, LYS5, LYS6, GLU7, ARG8, LEU9, GLU10, ASN11, ASP12, ILE13, ILE14, LEU16, ILE17, ASN18, VAL21, VAL30, LYS31, THR32, GLY33, HIS34, VAL35, THR36, HIS37, VAL38, LYS39, LEU40, ASP42, ASP43, LEU44, VAL47, VAL49, LEU51, VAL63, PHE66, ASN67, ALA69, LYS70, PHE73, VAL76, LEU77, ASN80, ILE89, HIS90, PHE91
division/cell-wall cluster transcriptional repressor MraZ (WP_010874670.1)	395.39	672.27	0.76	ASN33, ARG34, GLY35, PHE36, GLU37, ASN38, CYS39, LEU40, GLU41, TYR51, LEU68, LEU71, ILE72, ASP72, ASP96, ALA97, ILE106, GLN108, HIS111, GLU113, TRP115, TYR120, TYR123, LEU124
dTIGR00282 family metallophosphoesterase (WP_010874705.1)	177.28	311.54	0.31	LYS49, ASN71, HIS72, TRP74, PHE75, PHE99, LEU130, PRO131, PHE132
hypothetical protein (WP_010874779.1)	423.81	585.63	0.66	PHE62, SER66, VAL69, VAL86, LYS87, CYS89, CYS90, PHE93, TYR94, LEU97, PHE100, ILE101, LEU104, TYR115, LEU119, GLY120, PHE123, GLY124, VAL125
hypothetical protein (WP_014325598.1)	568.26	839.56	0.81	LYS45, GLU130, ILE131, THR132, VAL135, VAL139, ILE140, TYR143, TYR144, GLU145, THR147, ASN148, TYR154, VAL164, ALA167, LEU168, GLU171, ARG172, LEU175

Table 4. Docking studies of drug-like molecules (compounds) from the ZINC database with five drug target proteins. The table shows the binding scores/affinity, number of hydrogen bonds and the residues of proteins interacting with the respective compounds.

ZINC compound ID	AutoDock Vina binding affinity	no. of H-bond/residues
30S ribosome-binding factor (WP_010874513.1)		
ZINC04259381	−10.5	3/ASN18, ARG15
division/cell-wall cluster transcriptional repressor MraZ (WP_010874670.1)		
ZINC04235924	−10.2	1/ARG34
dTIGR00282 family metallophosphoesterase (WP_010874705.1)		
ZINC04259703	−8.9	3/LYS49, ASN71
hypothetical protein (WP_010874779.1)		
ZINC05415832	−11.1	1/PHE93
hypothetical protein (WP_014325598.1)		
ZINC04236030	−10.3	2/LYS45, TYR154

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

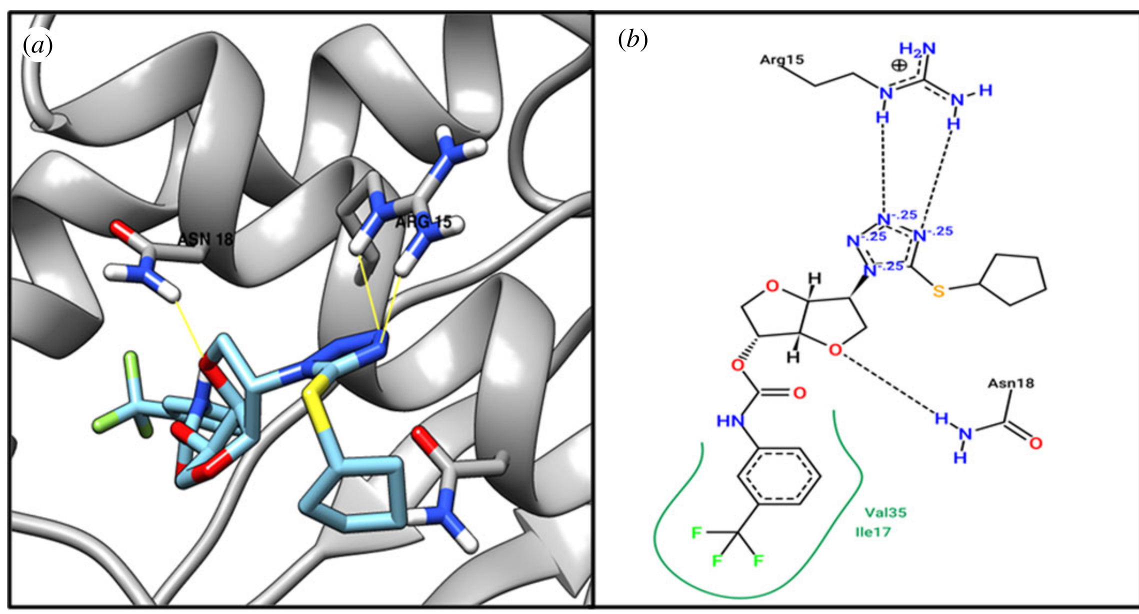


Figure 2. The (*a*) 3D and (*b*) 2D representation docking analyses of protein 30S ribosome-binding factor (WP_010874513.1) with compound ZINC04259381.

table 4, with ZINC database compound ID, AutoDock Vina binding affinity for the selected ligands as well as interactions of hydrogen bonds with the targets' residues involved in the interaction.

Based on structural comparison with a crystallographic structure of 30S ribosome-binding factor (WP_010874513.1) template (PDB ID: 1pa4) (putative ribosomal protein), we performed active site identification analysis with DoGSiteScorer [44], an online tool for active site residues (table 3). By performing virtual screening of 5008 drug-like molecules, we identified the top 10 molecules (electronic supplementary material, file S12); then flexible docking was performed on the identified top 10 molecules to find interactions with residues of the 30S ribosome-binding factor (WP_010874513.1) protein. We found that compound ZINC04259381 interacts with the active residue ASN18 from our active site identification analysis (table 4). Figure 2 shows the 3D and 2D representations of compound ZINC04259381.

Target MraZ (cell-wall cluster transcriptional repressor) protein, which is a transcription factor of *Escherichia coli*, regulates its own operon, also known as the division and cell wall (DCW) cluster [67]; active residue ARG34 evidenced an interaction with compound ZINC04235924 (table 4), while figure 3 shows the 3D and 2D representations of compound ZINC04235924. The target dTIGR00282 family metallophosphoesterase showed interaction with compound ZINC04259703 and interacts with the active residues LYS49 and ASN71 (table 4). Figure 4 shows the 3D and 2D representations of compound ZINC04259703. The hypothetical protein WP_010874779.1 showed interaction with compound ZINC05415832, by interacting with residue PHE93 (table 4). Figure 5 displays the 3D and 2D molecular representations of compound ZINC05415832. The hypothetical protein WP_014325598 evidenced interaction with compound ZINC04236030, by interacting with residues LYS45 and TYR154. Figure 6 depicts the 3D and 2D molecular representations of compound ZINC04236030.

3.5. Vaccine targets

From the 46 proteins predicted as membrane, PSE or secreted and whose structures were evaluated positively for adhesion capacity to MHC-I and MHC-II, 12 were noted with probability higher than 0.51 and considered good targets. They were also submitted to the DEG database, which indicated that eight of them were considered essential for *M. pneumoniae*. Among the eight potential vaccine targets found through these analyses, seven were lipoproteins and three of these belong to a specific group of membrane proteins of *M. pneumoniae*, with a lipid binding site and characterized as a membrane anchor (table 5).

The pro-lipoprotein diacylglyceryl transferase (WP_010874581.1) is an enzyme that catalyses the first step in the biogenesis of lipoproteins. It transfers the *n*-acyl diglyceride group into an N-terminal cysteine of the membrane lipoproteins. It is also an integral membrane protein that participates in a number of interactions; for example, the signal peptidase protein II that catalyses the removal of peptides that

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

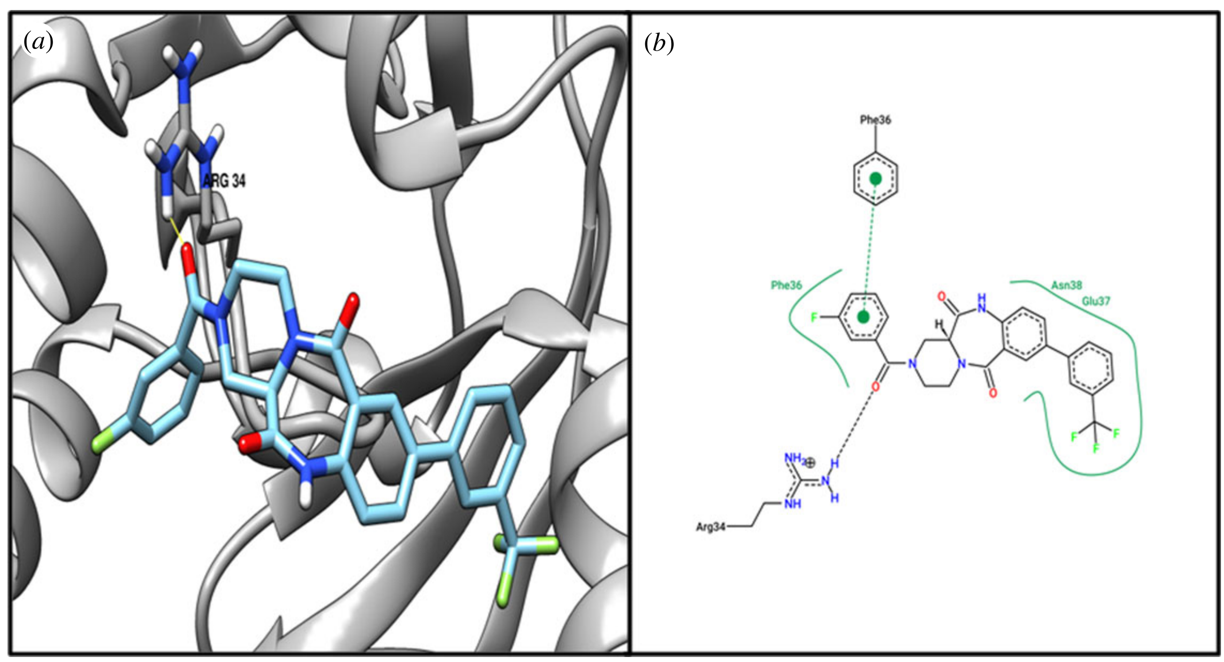


Figure 3. The (*a*) 3D and (*b*) 2D representation docking analyses of protein division/cell-wall cluster transcriptional repressor MraZ (WP_010874670.1) with compound ZINC04235924.

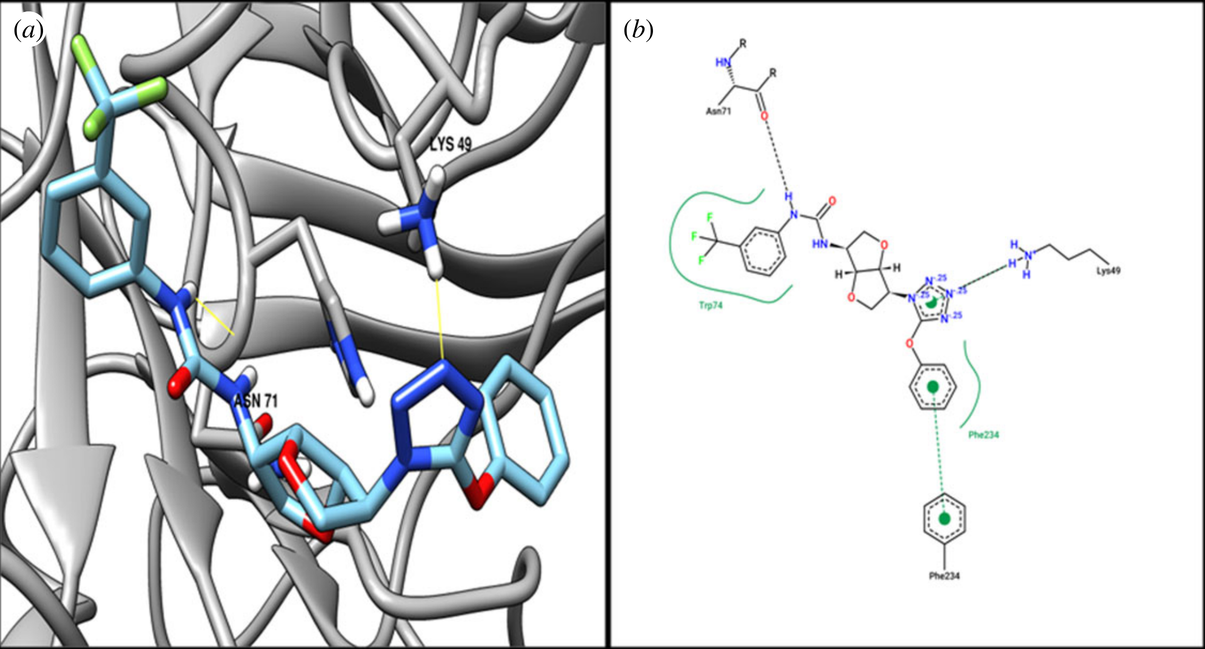


Figure 4. The (*a*) 3D and (*b*) 2D representation docking analyses of protein dTIGR00282 family metallophosphoesterase (WP_010874705.1) with compound ZINC04259703.

signal pro-lipoproteins. In addition, it is also involved with proteins that present DNA repair properties. The targets found by Vaxign with better MHC adhesion capacity were the proteins WP_014325486.1, with an adhesion index of 0.618, and WP_010874862.1, with index of 0.667. These two lipoproteins, which are predicted to belong to the cytoplasmic membrane or exposed to the surface, include 793 amino acids in their composition. The WP_010874581.1 protein showed seven transmembrane domains while the WP_014325660.1 protein showed one domain. The other proteins displayed no predicted domains through TMHMM (table 5).

The candidate proteins for vaccine targets were subjected to the antigenic prediction of B-cell epitopes. For each protein, the number of peptides with ability to induce the humoral immune response was predicted. We found 19 epitopes on the WP_010874862.1 protein and 16 epitopes on the WP_014574866.1 protein. Epitopes with fewer than seven amino acids were discarded from the study because they are considered too small to induce immunogenicity (electronic supplementary material, files S13–S20).

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

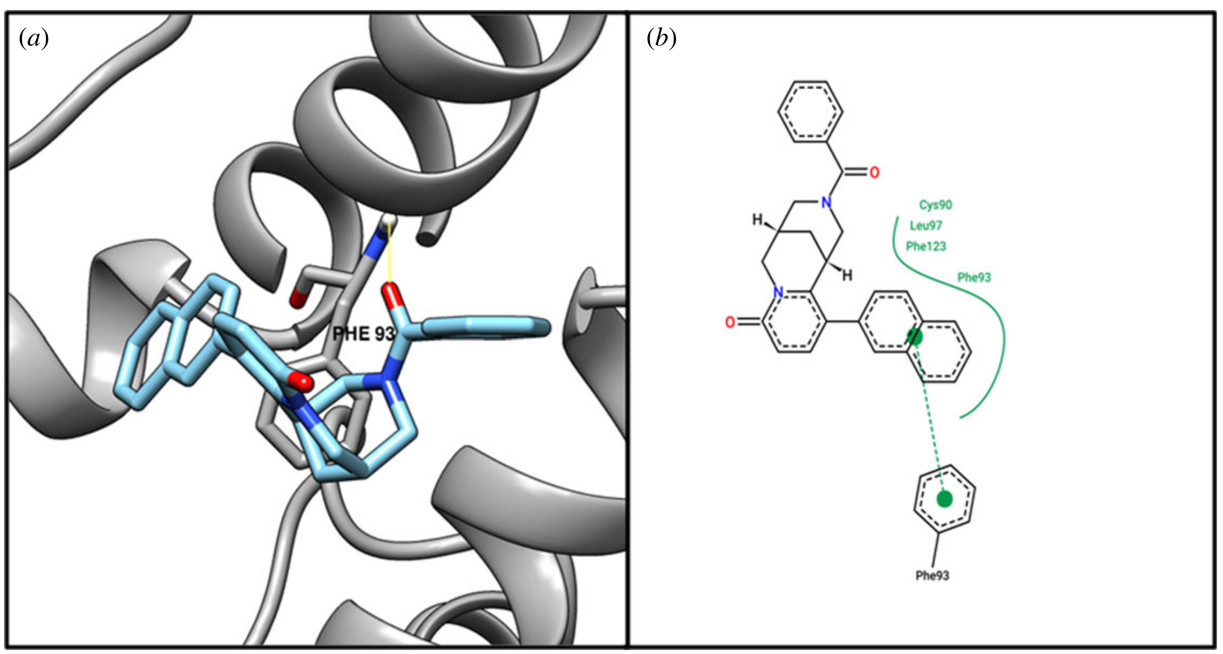


Figure 5. The (*a*) 3D and (*b*) 2D representation docking analyses of hypothetical protein (WP_010874779.1) with compound ZINC05415832.

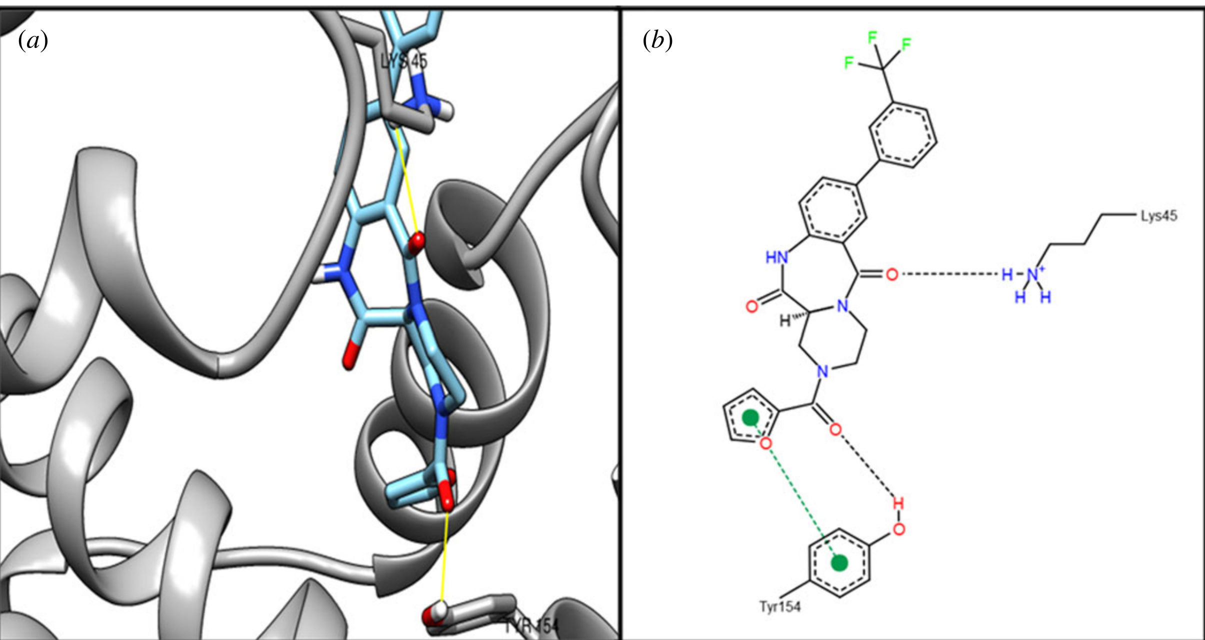


Figure 6. The (*a*) 3D and (*b*) 2D representation docking analyses of hypothetical protein (WP_014325598.1) with compound ZINC04236030.

3.6. Analysis of genomic similarities and phylogenetic reconstruction

The 88 genomes studied showed a high similarity. The heat map generated by the software Gegenees presented colours ranging from green (high similarity) to red (low similarity). Most genomes presented approximately 99% similarity, with the lowest being 95%. In the phylogenetic reconstruction performed through the software SplitsTree4, we can note the formation of seven clusters of *M. pneumoniae* genomes organized according to their phylogenetic characteristics (electronic supplementary material, files S21 and S22).

3.7. Prediction of genomic islands

In order to predict genomic islands, we selected 15 genomes belonging to the different clusters observed in the phylogenetic analyses. A genome of the species *M. gallinarum* was used as a reference for this

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

Table 5. Vaccine target candidates for *M. pneumoniae* identified by Vaxign. TMHMM, transmembrane propeller prediction software.

target	ID	name	location SurfG+	adhesin probability	no. predicted epitopes	TMHMM	protein length (aa)	SignalP	gene	molecular weight (DA) UniProt
1	WP_010874999.1	*Mycoplasma* specific lipoprotein, type 3	PSE	0.529	5	0	279	yes 25–26	MPN_642	31 287
2	WP_014574866.1	hypothetical lipoprotein	PSE/outer membrane	0.557	16	0	524	no	MPN_084	59 553
3	WP_010874581.1	pro-lipoprotein diacylglyceryl transferase	cytoplasmic membrane	0.578	9	7	389	no	MPN_XXX (lgt)	44 596
4	WP_010874862.1	uncharacterized lipoprotein MPN_506	PSE/cytoplasmic membrane	0.618	19	0	793	yes 24–25	MPN_506	87 494
5	WP_014325486.1	uncharacterized protein	PSE/outer membrane	0.667	16	0	793	yes 24–25	MPNE_0422	87 951
6	WP_014325517.1	uncharacterized lipoprotein MPN_408	PSE	0.606	15	0	760	yes 28–29	MPN_408	83 344
7	WP_014325659.1	uncharacterized lipoprotein MG440	PSE	0.536	7	0	277	yes 26–27	MPN_646	31 097
8	WP_014325660.1	uncharacterized lipoprotein MG439 homologue 1	extracellular	0.543	5	1	290	yes 28–29	MPN_647	31 823

14
royalsocietypublishing.org/journal/rsos *R. Soc. open sci.* **6**: 190907

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

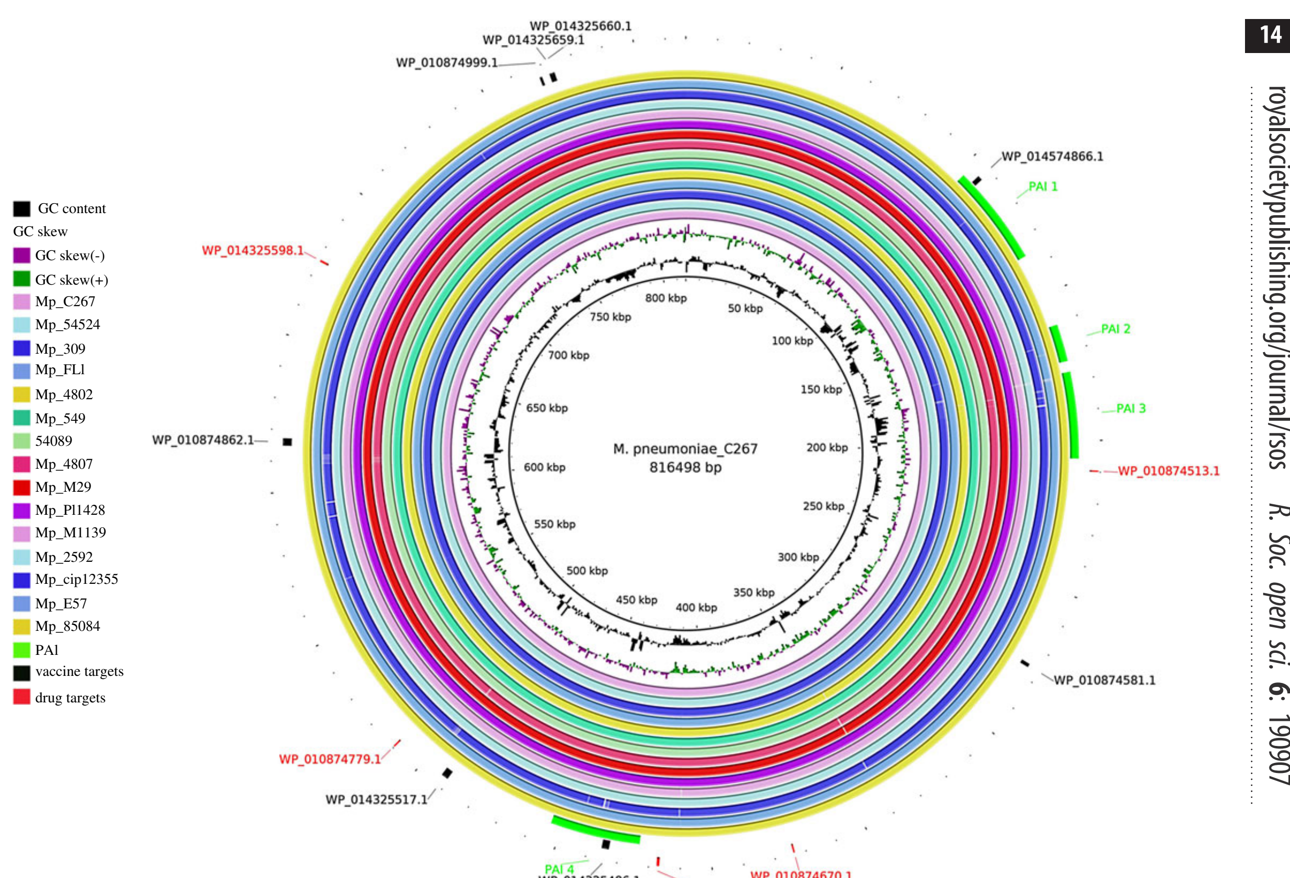


Figure 7. Pathogenicity islands predicted using 15 strains of *M. pneumoniae* having *M. pneumoniae* C267 as reference and *M. Gallinarum* as non-pathogenic reference.

prediction, given that this species is not pathogenic for humans. Four pathogenicity islands common to 15 genomes were predicted. These islands were common to all strains of *M. pneumoniae* tested and are very similar, as seen in the image generated by BRIG. For example, PAI1 is present in all strains and exhibits minimal deletions only in the CIP12355 and MP4807 strains. Among the vaccine targets, the WP_014574866.1 protein was predicted within PAI1 and the WP_014325486.1 protein within PAI4. The other proteins indicated as vaccine targets were not found in regions of pathogenicity islands. No potential drug targets were predicted among the genomic islands (figure 7).

4. Discussion

Mycoplasma pneumoniae is the main pathogen for pneumonia in children and, while the diagnosis is often limited, its incidence and number of antimicrobial-resistant cases have increased worldwide [68,69]. The scenario gets even worse if we consider the persistent absence of prophylactic methods and the fact that there are few treatment options for acute and chronic infections by this pathogen [18,70]. In the present study, comparative genomics and reverse vaccinology of 88 *M. pneumoniae* genomes were carried out in an attempt to predict vaccine and drug targets that could be tested in the near future in order to solve this public health problem. In the present study, employing the software OrthoFinder, 441 proteins belonging to the core genome were found and these targets were common to all 88 strains analysed throughout all coding sequences. Through subtractive genomics, we identified the proteins homologous to human proteins and removed them from the study, so that the selected targets act only against *M. pneumoniae*, preventing possible adverse reactions. After this filtration, 101 proteins remained. In order to evaluate the 101 proteins for their usefulness as drug targets or vaccines, we defined the subcellular location as the main parameter. Of the 101 proteins, 44 proteins considered membrane, PSE or secreted were, therefore, selected for the analysis of vaccine targets. The other 55 were analysed for their ability to act as drug targets. Finally, using reverse vaccinology, we found seven proteins with high potential for vaccine use and five with great potential for drug targeting.

The immune system can identify molecules foreign to our body and generate an immune response acquired through the interaction of these molecules with the MHC present in antigen-presenting cells

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

such as dendritic cells and macrophages. At the MHC-binding site, these foreign antigens are presented to $CD4^+$ and $CD8^+$ T cells in a process called antigen presentation, a process essential for the activation of the adaptive immune response and generation of the differential pattern of $CD4^+$ and $CD8^+$ T-cell immunity [71,72]. B cells are also involved in this immune response because they are also important antigen-presenting cells whose presentation is essential for the development of humoral immunity. The potential antigenic targets found in this study have not yet been tested *in vivo* for their ability to induce one or another pattern of immune response. Thus, further studies must be performed in order to test these targets and ascertain whether the best humoral and/or cellular immune response pattern is induced [73,74].

In the present study, 12 proteins were found capable of adhering to MHC-I and MHC-II with an index greater than 0.51, which means that they may induce either cellular or humeral adaptive immune responses [75]. As mentioned previously, these proteins are found either in the membrane, PSE or secreted by *M. pneumoniae* and, therefore, are the first to come into contact with host defences, given that *M. pneumoniae* has no cell wall and its location leaves them more exposed to the extracellular environment, which in turn facilitates recognition and specific memory immune responses [76,77]. Eight of the 12 proteins tested for *M. pneumoniae* essentiality were found to be strictly essential according to the software DEG.

From the amino acid sequences of these proteins, we predicted their B-cell epitopes. These epitopes can be recognized by the immune system, thus contributing to the development of immunity through the production of antibodies. Thus, in the present study, all eight proteins presented regions with B-cell interaction capabilities and could be used as vaccine targets [50].

Of these, only two proteins showed transmembrane domains. WP_010874581.1, belonging to the cytoplasmic membrane, presented seven transmembrane helices while the PSE protein WP_014325660.1 presented a predicted domain. These helices cross the outer membrane several times, forming loops where the epitopes are exposed, with precisely organized amino acids, thus enabling contact with the immune system [78]. However, proteins with more than one transmembrane helix in their structure hamper purification in assays for vaccine production [79].

Cytoplasmic proteins act in the maintenance of cell survival, and, for this reason, all 55 sequences of the 55 cytoplasmic proteins found were analysed for their potentials as drug targets. For this, the MHOLline tool was used, which identifies the set of 3D models of proteins on the core non-host. We found six proteins with $E \leq 10 \times 10^{-5}$, identity $\geq 25\%$ and LVI ≤ 0.7, criteria used to verify the significance in modelling. We prioritized the five proteins considered essential by the software DEG for further study because, if the target interferes with some vital metabolic pathway of the bacteria, the effectiveness of the possible drug that will come into contact with this protein will probably be greater.

The first protein found with potential as a drug target was WP_010874513.1 (ribosome-binding factor A) and this molecule is essential in the processing of 16S rRNA. The protein WP_010874670.1 (division / cell-wall cluster transcriptional repressor MraZ) interacts with components of cell division; this aspect when analysed with regard to drug targets is interesting since it is a vital cellular process. The third predicted protein, WP_010874705.1, from the metallophosphoesterase family, may also be considered a potential target since it relates to DNA repair, another very important process for cell integrity. The functions of the other two proteins found as potential drug targets remain incompletely elucidated. WP_010874779.1 is a hypothetical protein that according to STRING's predictions interacts with endonucleases, carrier proteins and proteins of the chromosomes, a property that may be theoretically important. The last protein, WP_014325598.1, is also a hypothetical protein related to the folding and transport of protein. These activities are intrinsically related to many important functions of any cell and, therefore, changes in these activities may compromise all cell biology.

The software AutoDock Vina was used for docking analysis. The five protein targets 30S ribosome-binding factor (WP_010874513.1), division/cell-wall cluster transcriptional repressor MraZ (WP_010874670.1), dTIGR00282 family metallophosphoesterase (WP_010874705.1), hypothetical protein (WP_010874779.1) and hypothetical protein (WP_014325598.1) were tested for their efficacy in binding to natural compounds obtained from the ZINC database that can act as drugs. Lower levels of energy and other parameters are related to greater interaction capacities [80]; so, we found 50 ligands with high druggability and, for each protein target, one of those ligands was selected to verify the structural interaction. ZINC04259381, ZINC04235924, ZINC04259703, ZINC05415832 and ZINC04236030 are the identified compounds with high affinity to bind the proteins. ZINC04259381 is the molecule with the best affinity score and binds to 30S ribosome-binding factor (WP_010874513.1), a target that is involved with RNA processing; and any alteration in this pathway can lead to cell

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

death. Therefore, ZINC04259381 is identified as the best drug candidate in our analysis and both identified molecules could be considered a candidate for antimicrobial chemotherapy in future studies for the development of drugs against pneumonia caused by *M. pneumoniae*.

The phylogenomic analysis was performed through two software programs, Gegenees and SplitsTree. This evaluation demonstrated the relationships and differences between the strains and modifications that normally occur during the evolution of microorganisms, including bacteria. These data reveal that, despite the differences between the strains of *M. pneumoniae,* visible in the phylogenetic tree generated by SplitsTree, the genomes were very similar. This proved to be very important given that the main goal of this work was to find potential targets for drugs and vaccines that could act against all strains of *M. pneumoniae*. This high level of genetic similarity is seen as a result of a degenerative evolution process, in which losses of genomic regions occurred over time, leaving only those genes essential for the species [81,82].

To understand the relationship of *M. pneumoniae* with eukaryotic cells, as well as their evolution, the studies on pathogenicity islands and virulence factors were essential. This information has already been shown to be important for the development of new methods of treatment and vaccination against bacteria [83]. From the methodology, we found four pathogenicity islands present in all 15 strains used. Comparison of the genomes performed via the software BRIG showed that there are few regions of deletion between the genomes, which indicates a small difference between the islands. This fact suggests that these islands already existed in the ancestral species that gave rise to *M. pneumoniae*.

Of all the targets that were detected in this research study, only the proteins WP_014574866.1 and WP_014325486.1 were found on pathogenicity islands. In this reverse vaccinology approach, previous studies reported that proteins associated with pathogenicity islands are considered to be excellent vaccine targets [22]. The protein WP_014325486.1 is PSE. This protein does not have a well-elucidated structure or functions, but it does have the highest capacity to bind to MCH-I and MHC-II, according to the results generated through Vaxign; it was also the protein with the highest number of predicted epitopes capable of activating B cells and developing the humoral immune response. All these characteristics reinforce the inference that this may be a good candidate for vaccines. Therefore, we believe that *in vivo* experiments should follow this direction and these targets should be tested in the near future.

5. Conclusion

CAP causes millions of deaths worldwide, an outcome that could be avoided with the development of appropriate prophylactic and treatment methods. In the present study, 88 genomes deposited in the NCBI database were employed to predict *in silico* proteins that can be used as vaccine targets or targets for new drugs. Through reverse vaccinology and subtractive genomic approaches, seven proteins with potential to induce immune responses were predicted as vaccine targets for protection against different strains of *M. pneumoniae*, a bacterium responsible for most of the infections that lead to pneumonia. Since treatment for this type of infection is limited because of the bacterium's high resistance to antibiotics, the genomes were also submitted to comparative genomic analysis that identified five possible drug targets. These targets were compared with different databases through molecular docking and should be subjected to future analyses as potential therapeutic resources.

Taking all the data together, we can assert that the current work presents great relevance to world health for finding new therapeutic targets for pneumonia due to *M. pneumoniae* infection. These targets could be quickly tested on new vaccine formulations and drug tests identified, representing a breakthrough in the area. In addition, further studies should also be performed on the other bacterial species that cause pneumonia in order to find new methods of treating the disease.

Data accessibility. The authors declare free access to the data obtained with this work. The information on how to obtain these data is indicated in the article. All the software used in the work are publicly available, as well as the genomes used for the analyses in general. Preliminary results of the gene screening have been made available; images and tables resulting from the analyses of the relationships between the lineages, molecular docking and genomic islands are also available in the electronic supplementary material along with other results obtained by the programs used during the development of the work. These files contain information on protein interactions, BLAST with the DEG and more descriptive results of epitope prediction. Access to these data can contribute to a better understanding of how it was possible to arrive at the final result and provides details on each step.

Authors' contribuitions. T.C.V.R.: carried out the download and genomes processing, carried out sequence alignments and data analyses, participated in the design of the study and drafted the manuscript. A.d.S., L.d.C.O., L.d.J.B.: carried out the download and genomes processing, carried out sequence alignments and data analyses, participated in the design

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

of the study. A.K.J., S.T., F.M.M.: participated in molecular docking analyses. C.J.F.O., P.G., V.A.d.C.A.: participated in the review of the article, contributing with suggestions and criticisms for approval. S.d.C.S.: conceived of the study, designed the study and coordinated the study.
Competing interests. We declare we have no competing interests.
Funding. This work was supported by the Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG, grant no. APQ-01323-15), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

1. Razin S, Yogev D, Naot Y. 1998 Molecular biology and pathogenicity of mycoplasmas. *Microbiol. Mol. Biol. Rev.* **62**, 1094–1156.
2. de Crécy-Lagard V, Marck C, Brochier-Armanet C, Grosjean H. 2007 Comparative RNomics and Modomics in Mollicutes: prediction of gene function and evolutionary implications. *IUBMB Life* **59**, 634–658. (doi:10.1080/15216540701604632)
3. Schnee C, Schulsse S, Hotzel H, Ayling RD, Nicholas RAJ, Schubert E, Heller M, Ehricht R, Sachse K. 2012 A novel rapid DNA microarray assay enables identification of 37 *Mycoplasma* species and highlights multiple mycoplasma infections. *PLoS ONE* **7**, e33237. (doi:10.1371/journal.pone.0033237)
4. Wilson MH, Collier AM. 1976 Ultrastructural study of *Mycoplasma pneumoniae* in organ culture. *J. Bacteriol.* **125**, 332–339.
5. Kumar S. 2018 *Mycoplasma pneumoniae*: a significant but underrated pathogen in paediatric community-acquired lower respiratory tract infections. *Ind. J. Med. Res.* **147**, 23–31. (doi:10.4103/ijmr.IJMR_1582_16)
6. Baron S. 1996 *Medical microbiology*. Galveston, TX: University of Texas Medical Branch at Galveston.
7. Waites KB, Crabb DM, Bing X, Duffy LB. 2003 *In vitro* susceptibilities to and bactericidal activities of garenoxacin (BMS-284756) and other antimicrobial agents against human mycoplasmas and ureaplasmas. *Antimicrob. Agents Chemother.* **47**, 161–165. (doi:10.1128/AAC.47.1.161-165.2003)
8. Waites KB, Talkington DF. 2004 *Mycoplasma pneumoniae* and its role as a human pathogen. *Clin. Microbiol. Rev.* **17**, 697–728. (doi:10.1128/CMR.17.4.697-728.2004)
9. Himmelreich R, Hilbert H, Plagens H, Pirkl E, Li B-C, Herrmann R. 1996 Complete sequence analysis of the genome of the bacterium *Mycoplasma pneumoniae*. *Nucleic Acids Res.* **24**, 4420–4449. (doi:10.1093/nar/24.22.4420)
10. Bajantri B, Venkatram S, Diaz-Fuentes G. 2018 *Mycoplasma pneumoniae*: a potentially severe infection. *J. Clin. Med. Res.* **10**, 535. (doi:10.14740/jocmr3421w)
11. Musher DM, Thorner AR. 2014 Community-acquired pneumonia. *N. Engl. J. Med.* **371**, 1619–1628. (doi:10.1056/NEJMra1312885)
12. Wang K, Gill P, Perera R, Thomson A, Mant D, Harnden A. 2012 Clinical symptoms and signs for the diagnosis of *Mycoplasma pneumoniae* in children and adolescents with community-acquired pneumonia. *Cochrane Database Syst. Rev.* **10**, CD009175. (doi:10.1002/14651858.cd009175)
13. Saraya T *et al.* 2014 Novel aspects on the pathogenesis of *Mycoplasma pneumoniae* pneumonia and therapeutic implications. *Front. Microbiol.* **5**, 410. (doi:10.3389/fmicb.2014.00410)
14. Ma Y-J *et al.* 2015 Clinical and epidemiological characteristics in children with community-acquired mycoplasma pneumonia in Taiwan: a nationwide surveillance. *J. Microbiol. Immunol. Infect.* **48**, 632–638. (doi:10.1016/j.jmii.2014.08.003)
15. Okada T, Morozumi M, Tajima T, Hasegawa M, Sakata H, Ohnari S, Chiba N, Iwata S, Ubukata K. 2012 Rapid effectiveness of minocycline or doxycycline against macrolide-resistant *Mycoplasma pneumoniae* infection in a 2011 outbreak among Japanese children. *Clin. Infect. Dis.* **55**, 1642–1649. (doi:10.1093/cid/cis784)
16. Zheng X *et al.* 2015 Macrolide-resistant *Mycoplasma pneumoniae*, United States. *Emerg. Infect. Dis.* **21**, 1470–1472. (doi:10.3201/eid2108.150273)
17. Cao B, Qu J-X, Yin Y-D, Van EJ. 2017 Overview of antimicrobial options for *Mycoplasma pneumoniae* pneumonia: focus on macrolide resistance. *Clin. Respir. J.* **11**, 419–429. (doi:10.1111/crj.12379)
18. Nascimento-Carvalho CMC. 2001 Etiology of childhood community acquired pneumonia and its implications for vaccination. *Braz. J. Infect. Dis.* **5**, 87–97. (doi:10.1590/S1413-86702001000200007)
19. Rudan I, Boschi-Pinto C, Biloglav Z, Mulholland K, Campbell H. 2008 Epidemiology and etiology of childhood pneumonia. *Bull. World Health Organ.* **86**, 408–416. (doi:10.2471/BLT.07.048769)
20. Madhi SA, Levine OS, Hajjeh R, Mansoor OD, Cherian T. 2008 Vaccines to prevent pneumonia and improve child survival. *Bull. World Health Organ.* **86**, 365–372. (doi:10.2471/BLT.07.044503)
21. Rappuoli R. 2000 Reverse vaccinology. *Curr. Opin Microbiol.* **3**, 445–450. (doi:10.1016/S1369-5274(00)00119-3)
22. Holmes RK. 2000 Biology and molecular epidemiology of diphtheria toxin and the *tox* gene. *J. Infect. Dis.* **181**, S156–S167. (doi:10.1086/315554)
23. Soares SC *et al.* 2012 PIPS: pathogenicity island prediction software. *PLoS ONE* **7**, e30848. (doi:10.1371/journal.pone.0030848)
24. Sachdeva G, Kumar K, Jain P, Ramachandran S. 2005 SPAAN: a software program for prediction of adhesins and adhesin-like proteins using neural networks. *Bioinformatics* **21**, 483–491. (doi:10.1093/bioinformatics/bti028)
25. Muzzi A, Masignani V, Rappuoli R. 2007 The pan-genome: towards a knowledge-based discovery of novel targets for vaccines and antibacterials. *Drug Discov. Today.* **12**, 429–439. (doi:10.1016/j.drudis.2007.04.008)
26. Santos A, Ali A, Barbosa E, Silva A, Miyoshi A, Barh D, Azevedo V. 2011 The lioab journal regular issue the reverse vaccinology—a contextual overview. *IIOABJ* **2**, 8–15.
27. Fleischmann RD *et al.* 1995 Whole-genome random sequencing and assembly of *Haemophilus influenzae* Rd. *Science* **269**, 496–512. (doi:10.1126/science.7542800)
28. Serruto D, Bottomley MJ, Ram S, Giuliani MM, Rappuoli R. 2012 The new multicomponent vaccine against meningococcal serogroup B, 4CMenB: immunological, functional and structural characterization of the antigens. *Vaccine* **30**, B87–B97. (doi:10.1016/j.vaccine.2012.01.033)
29. Giuliani MM *et al.* 2006 A universal vaccine for serogroup B meningococcus. *Proc. Natl Acad. Sci. USA* **103**, 10 834–10 839. (doi:10.1073/pnas.0603940103)
30. Chiang M-H, Sung W-C, Lien S-P, Chen Y-Z, Lo AF, Huang J-H, Kuo S-C, Chong P. 2015 Identification of novel vaccine candidates against *Acinetobacter baumannii* using reverse vaccinology. *Hum. Vaccin. Immunother.* **11**, 1065–1073. (doi:10.1080/21645515.2015.1010910)
31. Maione D *et al.* 2005 Identification of a universal group B *Streptococcus* vaccine by multiple genome screen. *Science* **309**, 148–150. (doi:10.1126/science.1109869)
32. Rappuoli R, Bottomley MJ, D'Oro U, Finco O, De Gregorio E. 2016 Reverse vaccinology 2.0: human immunology instructs vaccine antigen design. *J. Exp. Med.* **213**, 469–481. (doi:10.1084/jem.20151960)
33. Burton DR. 2017 What are the most powerful immunogen design vaccine strategies? *Cold Spring Harb. Perspect. Biol.* **9**, a030262. (doi:10.1101/cshperspect.a030262)
34. Emms DM, Kelly S. 2015 OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. *Genome Biol.* **16**, 157. (doi:10.1186/s13059-015-0721-2)
35. Mondal SI, Ferdous S, Jewel NA, Akter A, Mahmud Z, Islam MM, Afrin T, Karim N. 2015 Identification of potential drug targets by subtractive genome analysis of *Escherichia coli* 0157:H7: an *in silico* approach. *Adv. Appl. Bioinform. Chem.* **8**, 49–63. (doi:10.2147/AABC.S88522)

36. Zhang R, Ou H-Y, Zhang C-T. 2004 DEG: a database of essential genes. *Nucleic Acids Res.* **32** (Database issue), D271–D272. (doi:10.1093/nar/gkh024)
37. Barinov A, Loux V, Hammani A, Nicolas P, Langella P, Ehrlich D, Maguin E, Van De Guchte M. 2009 Prediction of surface exposed proteins in *Streptococcus pyogenes*, with a potential application to other Gram-positive bacteria. *Proteomics* **9**, 61–73. (doi:10.1002/pmic.200800195)
38. Capriles PV, Guimarães AC, Otto TD, Miranda AB, Dardenne LE, Degrave WM. 2010 Structural modelling and comparative analysis of homologous, analogous and specific proteins from *Trypanosoma cruzi* versus *Homo sapiens*: putative drug targets for Chagas' disease treatment. *BMC Genomics* **11**, 610. (doi:10.1186/1471-2164-11-610)
39. Jamal SB *et al.* 2017 An integrative *in-silico* approach for therapeutic target identification in the human pathogen *Corynebacterium diphtheriae*. *PLoS ONE* **12**, e0186401. (doi:10.1371/journal.pone.0186401)
40. Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. 2012 ZINC: a free tool to discover chemistry for biology. *J. Chem. Inf. Model.* **52**, 1757–1768. (doi:10.1021/ci3001277)
41. Irwin JJ, Shoichet BK. 2005 ZINC—a free database of commercially available compounds for virtual screening. *J. Chem. Inf. Model.* **45**, 177–182. (doi:10.1021/ci049714+)
42. O'Boyle NM, Banck M, James CA, Morley C, Vandermeersch T, Hutchison GR. 2011 Open babel: an open chemical toolbox. *J. Cheminform.* **3**, 33. (doi:10.1186/1758-2946-3-33)
43. Morris GM, Huey R, Lindstrom W, Sanner MF, Belew RK, Goodsell DS, Olson AJ. 2009 AutoDock4 and AutoDockTools4: automated docking with selective receptor flexibility. *J. Comput. Chem.* **30**, 2785–2791. (doi:10.1002/jcc.21256)
44. Volkamer A, Kuhn D, Rippmann F, Rarey M. 2012 DoGSiteScorer: a web server for automatic binding site prediction, analysis and druggability assessment. *Bioinformatics* **28**, 2074–2075. (doi:10.1093/bioinformatics/bts310)
45. Trott O, Olson AJ. 2010 AutoDock Vina: improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **31**, 455–461. (doi:10.1002/jcc.21334)
46. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE. 2004 UCSF Chimera? A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612. (doi:10.1002/jcc.20084)
47. Stierand K, Maass PC, Rarey M. 2006 Molecular complexes at a glance: automated generation of two-dimensional complex diagrams. *Bioinformatics* **22**, 1710–1716. (doi:10.1093/bioinformatics/btl150)
48. He Y, Xiang Z, Mobley HLT. 2010 Vaxign: The first web-based vaccine design program for reverse vaccinology and applications for vaccine development. *J. Biomed. Biotechnol.* **2010**, 297505. (doi:10.1155/2010/297505)
49. He Y *et al.* 2014 Updates on the web-based VIOLIN vaccine database and analysis system. *Nucleic Acids Res.* **42**, D1124–D1132. (doi:10.1093/nar/gkt1133)
50. Vita R, Mahajan S, Overton JA, Dhanda SK, Martini S, Cantrell JR, Wheeler DK, Sette A, Peters B. 2019 The Immune Epitope Database (IEDB): 2018 update. *Nucleic Acids Res.* **47**, D339–D343. (doi:10.1093/nar/gky1006)
51. Szklarczyk D *et al.* 2019 STRING v11: protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613. (doi:10.1093/nar/gky1131)
52. UniProt Consortium TU. 2008 The universal protein resource (UniProt). *Nucleic Acids Res.* **36** (Database issue), D190–D195. (doi:10.1093/nar/gkm895)
53. Petersen TN, Brunak S, von Heijne G, Nielsen H. 2011 SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* **8**, 785–786. (doi:10.1038/nmeth.1701)
54. Krogh A, Larsson B, von Heijne G, Sonnhammer EL. 2001 Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J. Mol. Biol.* **305**, 567–580. (doi:10.1006/jmbi.2000.4315)
55. Wishart DS *et al.* 2018 DrugBank 5.0: a major update to the DrugBank database for 2018. *Nucleic Acids Res.* **46**(D1), D1074–D1082. (doi:10.1093/nar/gkx1037)
56. Langdon A, Crook N, Dantas G. 2016 The effects of antibiotics on the microbiome throughout development and alternative approaches for therapeutic modulation. *Genome Med.* **8**, 39. (doi:10.1186/s13073-016-0294-z)
57. Lievin-Le Moal V, Servin AL. 2014 Anti-infective activities of *Lactobacillus* strains in the human intestinal microbiota: from probiotics to gastrointestinal anti-infectious biotherapeutic agents. *Clin. Microbiol. Rev.* **27**, 167–199. (doi:10.1128/CMR.00080-13)
58. Ågren J, Sundström A, Håfström T, Segerman B. 2012 Gegenees: fragmented alignment of multiple genomes for determining phylogenomic distances and genetic signatures unique for specified target groups. *PLoS ONE* **7**, e39107. (doi:10.1371/journal.pone.0039107)
59. Huson DH, Bryant D. 2006 Application of phylogenetic networks in evolutionary studies. *Mol. Biol. Evol.* **23**, 254–267. (doi:10.1093/molbev/msj030)
60. Soares SC *et al.* 2016 GIPSy: genomic island prediction software. *J. Biotechnol.* **232**, 2–11. (doi:10.1016/j.jbiotec.2015.09.008)
61. Huson DH. 1998 SplitsTree: analyzing and visualizing evolutionary data. *Bioinformatics* **14**, 68–73. (doi:10.1093/bioinformatics/14.1.68)
62. Alikhan N-F, Petty NK, Ben Zakour NL, Beatson SA. 2011 BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* **12**, 402. (doi:10.1186/1471-2164-12-402)
63. López-Alonso JP *et al.* 2017 RsgA couples the maturation state of the 30S ribosomal decoding center to activation of its GTPase pocket. *Nucleic Acids Res.* **45**, 6945–6959. (doi:10.1093/nar/gkx324)
64. Mehta R, Champney WS. 2002 30S ribosomal subunit assembly is a target for inhibition by aminoglycosides in *Escherichia coli*. *Antimicrob. Agents Chemother.* **46**, 1546–1549. (doi:10.1128/AAC.46.5.1546-1549.2002)
65. Durán AAA. 2018 Caracterização fenotípica de linhagens mutantes das RNA helicases DEAD-box *de Caulobacter crescentus* em condições de baixa temperatura, [São Paulo]: Biblioteca Digital de Teses e Dissertações da Universidade de São Paulo. See http://www.teses.usp.br/teses/disponiveis/42/42132/tde-22022018-160033/ (cited 9 October 2018).
66. Gyawali R, Ibrahim SA. 2014 Natural products as antimicrobial agents. *Food Control* **46**, 412–429. (doi:10.1016/j.foodcont.2014.05.047)
67. Fisunov GY, Evsyutina DV, Semashko TA, Arzamasov AA, Manuvera VA, Letarov AV, Govorun VM. 2016 Binding site of MraZ transcription factor in Mollicutes. *Biochimie* **125**, 59–65. (doi:10.1016/j.biochi.2016.02.016)
68. Del Valle-Mendoza J *et al.* 2017 Molecular etiological profile of atypical bacterial pathogens, viruses and coinfections among infants and children with community acquired pneumonia admitted to a national hospital in Lima, Peru. *BMC Res. Notes* **10**, 688. (doi:10.1186/s13104-017-3000-3)
69. Yang T-I, Chang T-H, Lu C-Y, Chen J-M, Lee P-I, Huang L-M, Chang LY. 2019 *Mycoplasma pneumoniae* in pediatric patients: do macrolide-resistance and/or delayed treatment matter? *J. Microbiol. Immunol. Infect.* **52**, 329–335. (doi:10.1016/j.jmii.2018.09.009)
70. Bébéar C, Pereyre S, Peuchant O. 2011 *Mycoplasma pneumoniae*: susceptibility and resistance to antibiotics. *Future Microbiol.* **6**, 423–431. (doi:10.2217/fmb.11.18)
71. Kappler JW, Staerz U, White J, Marrack PC. 1988 Self-tolerance eliminates T cells specific for Mls-modified products of the major histocompatibility complex. *Nature* **332**, 35–40. (doi:10.1038/332035a0)
72. Klein J, Figueroa F. 1986 Evolution of the major histocompatibility complex. *Crit. Rev. Immunol.* **6**, 295–386.
73. Fagarasan S, Honjo T. 2000 T-independent immune response: new aspects of B cell biology. *Science* **290**, 89–92. (doi:10.1126/science.290.5489.89)
74. den Haan JM, Lehar SM, Bevan MJ. 2000 $Cd8^+$ but not $Cd8^-$ dendritic cells cross-prime cytotoxic T cells *in vivo*. *J. Exp. Med.* **192**, 1685–1696. (doi:10.1084/jem.192.12.1685)
75. Solanki V, Tiwari V. 2018 Subtractive proteomics to identify novel drug targets and reverse vaccinology for the development of chimeric vaccine against *Acinetobacter baumannii*. *Sci. Rep.* **8**, 9044. (doi:10.1038/s41598-018-26689-7)
76. Varela JN, Izidoro Jr MS, de Hollanda LM, Lancellotti M. 2012 Membrane protein as novel targets for vaccine production in *Haemophilus influenzae* and *Neisseria meningitid*. *J. Vaccines Vaccin.* **3**, 152. (doi:10.4172/2157-7560.1000152)
77. Kumar Jaiswal A, Tiwari S, Jamal SB, Barh D, Azevedo V, Soares SC. 2017 An in silico identification of common putative vaccine candidates against *Treponema pallidum*: a

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

reverse vaccinology and subtractive genomics based approach. *Int. J. Mol. Sci.* **18**, 402. (doi:10.3390/ijms18020402)

78. Rappuoli R. 2001 Reverse vaccinology, a genome-based approach to vaccine development. *Vaccine* **19**, 2688–2691. (doi:10.1016/S0264-410X(00)00554-5)
79. Meunier M, Guyard-Nicodème M, Hirchaud E, Parra A, Chemaly M, Dory D. 2016 Identification of novel vaccine candidates against campylobacter through reverse vaccinology. *J. Immunol. Res.* **2016**, 5715790. (doi:10.1155/2016/5715790)
80. Thomsen R, Christensen MH. 2006 MolDock: a new technique for high-accuracy molecular docking. *J. Med. Chem.* **49**, 3315–3321. (doi:10.1021/jm051197e)
81. Pettersson B, Uhlen M, Johansson K-E. 1996 Phylogeny of some mycoplasmas from ruminants based on 16S rRNA sequences and definition of a new cluster within the Hominis group. *Int. J. Syst. Bacteriol.* **46**, 1093–1098. (doi:10.1099/00207713-46-4-1093)
82. Siezen RJ *et al.* 2011 Genome-scale diversity and niche adaptation analysis of *Lactococcus lactis* by comparative genome hybridization using multi-strain arrays. *Microb. Biotechnol.* **4**, 383–402. (doi:10.1111/j.1751-7915.2011.00247.x)
83. Naz A, Awan FM, Obaid A, Muhammad SA, Paracha RZ, Ahmad J, Ali A. 2015 Identification of putative vaccine candidates against *Helicobacter pylori* exploiting exoproteome and secretome: a reverse vaccinology based approach. *Infect. Genet. Evol.* **32**, 280–291. (doi:10.1016/j.meegid.2015.03.027)

Downloaded from https://royalsocietypublishing.org/ on 30 May 2021

CONCLUSÃO CAPÍTULO I

A pneumonia adquirida na comunidade causa milhões de mortes em todo o mundo, fato que poderia ser evitado com o desenvolvimento de métodos profiláticos e de tratamento adequados. No presente estudo, 88 genomas depositados no banco de dados do NCBI foram usados para predição, *in silico*, de proteínas que poderiam ser usadas como alvos de vacina ou de novas drogas. Através da vacinologia reversa e abordagens de genômicas comparativa e subtrativa, 8 proteínas com potencial para induzir respostas imunes foram preditas como alvos vacinais, para proteção contra uma ampla gama de linhagens de *M. pneumoniae*, uma bactéria responsável pela maioria das infecções atípicas que levam à pneumonia. Como o tratamento para esse tipo de infecção é limitado devido à alta resistência a antibióticos, os genomas também foram submetidos a análises de genômica comparativa e 5 possíveis alvos de drogas foram identificados. Esses alvos foram comparados a diferentes bancos de dados por meio de ancoramento, *docking* molecular, e podem ser submetidos a análises futuras para serem utilizados como recurso terapêutico.

Tomando todos os dados juntos, este trabalho apresenta relevância ao determinar novos potenciais alvos terapêuticos e vacinais contra pneumonia causada por *M. pneumoniae*. Esses alvos de drogas e candidatos a vacinas podem ser futuramente validados em novas formulações de vacinas e testes de drogas identificadas, representando um avanço prevenção e tratamento da PAC.

6. CAPÍTULO II

Prediction of multi-epitope vaccine for *Mycoplasma pneumoniae* infection applying immunoinformatics approaches

Thaís Cristina Vilela Rodrigues, Arun Kumar Jaiswal, Marcela Rezende Lemes, Marcos Vinícius da Silva, Helioswilton Sales-Campos, Luiz Carlos Júnior Alcântara, Sthephane Fraga de Oliveira Tosta, Rodrigo Bentes Kato, Vasco Ariston de Carvalho Azevedo, Sandeep Tiwari, Siomar de Castro Soares,

Com a predição dos alvos vacinais por meio do trabalho anterior, a etapa seguinte proposta foi a construção de uma vacina multi-epítopos contra o *M.pneumoniae*. Este processo consiste em avaliar proteínas em busca de determinantes imunogênicos para a construção de proteínas quiméricas imunogênicas. Nesse contexto, proteínas determinadas pela vacinologia reversa e outras proteínas imunogênicas selecionadas na literatura foram submetidas à predição dos epítopos MHC-I, MHC-II e B. Os epítopos que se enquadraram na cobertura populacional, afinidade de ligação e outros critérios de avaliação, foram combinados com ligantes de peptídeos e um adjuvante para finalizar a construção da vacina multi-epítopos. A estrutura da proteína foi predita e submetida ao ancoramento, *docking* molecular contra um receptor de resposta imune inata, seguido de simulação dinâmica. Posteriormente, clonagem *in silico* e simulação imune *in silico* foram realizadas para avaliar o potencial da vacina.

International Journal of Biological Macromolecules

Prediction of multi-epitope vaccine for Mycoplasma pneumoniae infection applying immunoinformatics approaches

--Manuscript Draft--

Manuscript Number:	
Article Type:	Research Paper
Section/Category:	Proteins and Nucleic acids
Keywords:	Mycoplasma pneumoniae; multi-epitope vaccine; immunoinformatics
Corresponding Author:	Sandeep Tiwari, PhD.,MSc. Universidade Federal de Minas Gerais BRAZIL
First Author:	Thaís Cristina Vilela Rodrigues
Order of Authors:	Thaís Cristina Vilela Rodrigues
	Arun Kumar Jaiswal
	Marcela Rezende Lemes
	Marcos Vinícius da Silva
	Helioswilton Sales-Campos
	Luiz Carlos Júnior Alcântara
	Sthephane Fraga de Oliveira Tosta
	Rodrigo Bentes Kato
	Debmalya Barh
	Vasco Ariston de Carvalho Azevedo
	Sandeep Tiwari, PhD.,MSc.
	Siomar de Castro Soares
Abstract:	Globally Pneumonia is a severe health problem, being the cause of death of over one million people annually. Among the main microorganisms leading to pneumonia, Mycoplasma pneumoniae is one of the most common, with a significant increase in cases during the last years. This study aims to identify epitopes with immunological foreseen properties through immunoinformatics approaches to construct a multi- epitope vaccine against M. pneumoniae . Our multi-epitope vaccine was constructed by CTL, HTL and B epitopes belonging to core proteins previously identified as vaccine candidates through reverse vaccinology approaches. The epitopes were filtered according to their immunogenicity, population coverage and among other criteria. The final 16 CTL and 13 HTL epitopes were joined with peptide linkers and the heat-labile enterotoxin from Escherichia coli was used as an adjuvant. The vaccine was considered stable, nontoxic, non-allergenic, non-host homologous, and with antigenic and immunogenic properties. The docking of the vaccine with the Toll-Like Receptor 2 was performed, followed by the dynamic simulation. In silico cloning was tested in an expression vector with positive results. We constructed an effective multi-epitope vaccine candidate through immunoinformatics approaches that could contribute to the prevention of pneumonia on a massive scale.
Suggested Reviewers:	Kenneth Lundstrom PanTherapeutics, CH1095 Lutry, Switzerland, lundstromkenneth@gmail.com

	Chandana Mohanty chandanamohanty@gmail.com
	George D. Vavougios gvavougyios@uth.gr

Powered by Editorial Manager® and ProduXion Manager® from Aries Systems Corporation

Prediction of multi-epitope vaccine for *Mycoplasma pneumoniae* infection applying immunoinformatics approaches

Thaís Cristina Vilela Rodrigues[1], Arun Kumar Jaiswal[1], Marcela Rezende Lemes[2], Marcos Vinícius da Silva[2], Helioswilton Sales-Campos[3], Luiz Carlos Júnior Alcântara[4], Sthephane Fraga de Oliveira Tosta[1], Rodrigo Bentes Kato[1], Debmalya Barh[5],Vasco Ariston de Carvalho Azevedo[1], Sandeep Tiwari[1*], Siomar de Castro Soares[2*],

[1]PG Program in Bioinformatics, Institute of Biological Sciences, Federal University of Minas Gerais (UFMG), Belo Horizonte 31270-901, MG, Brazil

[2]Department of Immunology, Microbiology and Parasitology, Institute of Biological Science and Natural Sciences, Federal University of Triângulo Mineiro (UFTM), Uberaba 38025-180, MG, Brazil

[3]Institute of Tropical Pathology and Public Health, Federal University of Goias (UFG), Goiânia 74605-050, GO, Goiás, Brazil

[4]Oswaldo Cruz Fundation, Oswaldo Cruz Institute

[5]Centre for Genomics and Applied Gene Technology, Institute of Integrative Omics and Applied Biotechnology (IIOAB), Nonakuri, Purba Medinipur, West Bengal-721172, India

[*]Correspondence author: Sandeep Tiwari: sandip_sbtbi@yahoo.com & Siomar de Castro Soares: siomars@gmail.com

ABSTRACT

Pneumonia is a globally severe health problem, being the cause of death of over one million people annually. *Mycoplasma pneumoniae* is one of the most common pathogens, with a significant increase in cases during the last years. This study aims to identify epitopes with immunological foreseen properties through immunoinformatics approaches to construct a multi-epitope vaccine against *M.pneumoniae*. Our multi-epitope vaccine was constructed by CTL, HTL and B epitopes belonging to core proteins previously identified as vaccine candidates through reverse vaccinology approaches. The epitopes were filtered according to their immunogenicity, population coverage and among other criteria. The final 16 CTL and 13 HTL epitopes were joined with peptide linkers and the heat-labile enterotoxin from *Escherichia coli* was used as an adjuvant. Those epitopes have high affinity with immunological receptors, are recognized by frequent receptors in the world population and belong to proteins previously described for having high immunogenic properties. The vaccine was considered stable, nontoxic, non-allergenic, non-host homologous, and with antigenic and immunogenic properties. The docking of the vaccine with the Toll-Like Receptor 2 was performed, followed by the dynamic simulation demonstrating good immunogenic potential. We provide, through immunoinformatics approaches, a less laborious alternative to vaccine against *M.pneumoniae*, without the need to grow the bacteria, that with validation tests could contribute to the prevention of pneumonia on a massive scale.

Keywords: *Mycoplasma pneumoniae*, multi-epitope vaccine, immunoinformatics, chimeric antigen, epitope vaccine

1. INTRODUCTION

Community-acquired pneumonia (CAP) is an acute lower tract infection affecting infants under 5 years old and elderly people with comorbidities [1,2]. CAP incidence is probably underestimated because of the non-notification or misdiagnosis [3]. Further, this morbidity commonly requires hospital care with high costs. Although this disease is not restricted to the less developed countries, the mortality rate is intensified in regions with economic issues, demonstrating the expressive importance of political and socioeconomic actions in disease prevention and treatment [4,5]. With the pneumococcal vaccine's success, the number of people developing CAP associated with *Streptococcus pneumoniae* infection has significantly decreased. However, the number of cases associated with *Chlamydia pneumoniae* and *Mycoplasma pneumoniae* infection has drastically increased, even in coinfection with COVID-19 [6–8]. *M.pneumoniae* infection is one of the major causes of morbidity and mortality worldwide and has been associated with 3 million deaths annually [9,10].

M.pneumoniae is pleomorphic because it cannot produce protein constituents of the cell wall [11,12]. These organisms are difficultly identified by conventional culture methods and are naturally resistant to the effects of antibiotics towards cell wall. This bacteria has a high rate of lineages resistant to macrolide, which was one of the most important therapeutic options [13,14], thus suggesting the importance of developing vaccines.

There are two lineages of *M.pneumoniae*, Type 1 and Type 2, differentiated by the existing polymorphisms between the adhesin proteins P1 [15] and together with P30 adhesin has a crucial role in *M.pneumoniae*'s virulence [16]. Those adhesins are key proteins in cytoadherence to the respiratory epithelium and activation of immune responses [16,17]. After adhesion, these *Mycoplasmas* can penetrate in cells and cause mild symptoms as is known as walking pneumonia; however, it can vary from nasal congestion and pharyngitis to pneumonia with the need for hospital services due to the severity of pulmonary commitment (Centers for Disease Control and Prevention) [18]. There is also a serious possibility to develop chronic infections, mainly in asthmatic patients [19,20].

Computational methods of reverse vaccinology and immunoinformatics have demonstrated to be an excellent strategy to minimize time and resources in vaccine development, moreover, those strategies prove to be crucial alternatives in the development of vaccines against difficult-to-cultivate pathogens such as *M.pneumoniae*. In this context, multi-epitope vaccines are constituted by judiciously selected epitopes, which can induce both cellular and humoral immune responses. The benefits of searching for these new vaccines candidates

are: the possibilities to direct the immune response to conserved regions of the pathogen and the ability to induce strong immunological response with no adverse reactions, which is more frequently observed with attenuated or inactivated vaccines. [21]. This fact supports the choice of these approaches, as several attempts to develop vaccines against *M.pneumoniae* have failed due to the exacerbation of immune responses which led to the disease associated with the vaccines [22]. Multi-epitope vaccines are safer and have uncomplicated pharmaceutical development [23–25]. Some studies have successfully used this platform for bacterial infections such as *Staphylococcus aureus* [26] and *Helicobacter pylori* [27]. Also, multi-epitope vaccines have been tested against viral infections such as influenza, yellow fever [28,29] and lymphocytic choriomeningitis virus (LCMV) [30].

To the best of our knowledge, only one vaccine platform, using attenuated strains of *M. pneumoniae* has been tested [31], which was not effective to protect mice from developing infection. On other hand, in human, an inactivated vaccine demonstrated a moderate protective effect with some adverse reactions [32]. With the diversity of the lineages, hardness in diagnosing, and the alarming increase of macrolide-resistant strains, the development of new strategies to prevent and control CAP caused by *M.pneumoniae* is of paramount importance. Hence, based on the magnitude of CAP caused by *M.pneumoniae* and the fact that there is no vaccine available for this pathogen, we seek to determine a multi-epitope vaccine against it.

2. METHODOLOGY

2.1 Data selection

For carrying out the antigenicity analysis and *in silico* construction of the vaccine, sequences of eight proteins previously determined through reverse vaccinology [33] were selected from the core genome of 88 genomes of *M.pneumoniae*. Core proteins have an elevated potential for being immunogenic due to its high expression level [34]. Further, these proteins were analyzed for similarity with human proteins, essentiality, and affinity with CTL/HTL and to the ability to induce specific B cell antibodies. Two other proteins, P1 and P30 adhesins, considered to be highly immunogenic by *in vitro/in vivo* studies, were also selected for this study. The amino acid sequence of each of these proteins was retrieved from NCBI (Nacional Center for Biotechnology Information) in FASTA format and used as a dataset for the analyzes. As there are two types of *M.pneumoniae* regarding P1 adhesins, we select the two P1 proteins responsible for this differentiation. One from M129 lineage (protein WP_010874498.1), which characterizes the lineages of type 1, and other protein from FH lineage (WP_053344075.1),

representing type 2. Hence, the initial analysis was performed with 11 protein sequences (**Supplementary** t**able 1**).

2.2 Prediction of epitopes that binds MHC-I alleles

To predict the putative epitopes that bind MHC-I alleles, and therefore, are able to activate cytotoxic T lymphocytes (CTL), two different approaches and platforms were used to improve the selected epitope's confidence. The Immune Epitope Database and Analysis Resource (IEDB) is a robust database that contains thousands of high and low-affinity epitopes used on training to enhance the accuracy of the predictor [35,36]. The MHC-I binding prediction tool was used for epitopes identification. As our study aimed to design *in silico* a vaccine that reaches the largest and more diverse population worldwide, the set of 27 alleles with high frequency in the global population observed were selected for epitope predictions of 9 amino acid residues. [37]. For the prediction, default parameters, which are the most suitable, were applied because they combine Artificial Neural Network (ANN), Scoring Matrix Method (SMM), and combinatorial library. Only the peptides with percentile rank smaller than 1%, which are the ones with the more significant probability to be immunogenic, were chosen.

Additionally, we applied an integrated epitope prediction method that asses the binding affinity and antigenic processing and transportation. The NETCTL-1.2 uses of ANN and SMM to make the predictions. The default parameters were applied to select the epitopes: 0.15 and 0.05 to weight on C terminal cleavage and weight on transport efficiency respectively, and 0.75 was the threshold for epitope identification [38,39]. All the available allele supertypes tested in the previously step that were also used for this prediction.

2.3 Prediction of HTL epitopes that binds MHC-II alleles

For exogenous epitopes of MHC-II, and therefore, able to activate helper T-lymphocytes (HTL) two predictors of high-quality were used to identify the best epitopes [40,41], the IEDB tool [35] in association with the results obtained from NetMHCII 2.3 Server [42]. Different the previous analysis with CTL, the HTL epitopes contain 13 to 25 amino acid residues due to the MHC cleft size formed by the alpha and beta chains. Thus, we set as standard 15 residues for this prediction. In this step, were selected only IEDB epitopes with percentile rank <3% and IC50 <1000 nM as indicated by the predictor. The IC50 is the metric used to determine the epitopes' affinity with the MHC and as indicated by the predictor [43]. Meanwhile, the NetMHCII 2.3 Server [44] uses Artificial Neural Network (ANN) with diverse epitope databases to enhance the data training and realize the predictions. The epitopes selected was those considered strong binders with the threshold 2.

2.4 Prediction of B cell epitopes

Linear epitopes that interact with B cell receptors were selected through ABCpred [45,46] server, which makes predictions through ANN. The platform is commonly in those approaches and have high accuracy [47]. As it is possible to set the epitope size, we adopt 16 residues for having shown better accuracy and the threshold of 0.51 [48–50].

2.5 Filtering

As there are two types of P1 adhesin (one from each lineage) and the study aims to develop a vaccine with wide coverage, only epitopes within the two P1 adhesin types were kept on the analyses. An *in-house* python script was applied to determine the epitopes shared among these proteins. We applied the scrip for each cell that can recognize it, CTL, HTL, or B lymphocytes.

Afterwards, the same script was used to find predicted epitopes with high confidence among the 11 proteins, which was predicted by the two methods chosen. We searched for overlapping epitopes of different methods for each protein separately and each cell type, CTL, HTL, and B. Finally, to find epitopes capable of inducing both humoral and cellular responses, we applied the same script and only the epitopes with overlap between CTL with B or HTL with B were selected for further analysis. The complete workflows of epitopes prediction and selection steps are shown in **Figure 1**.

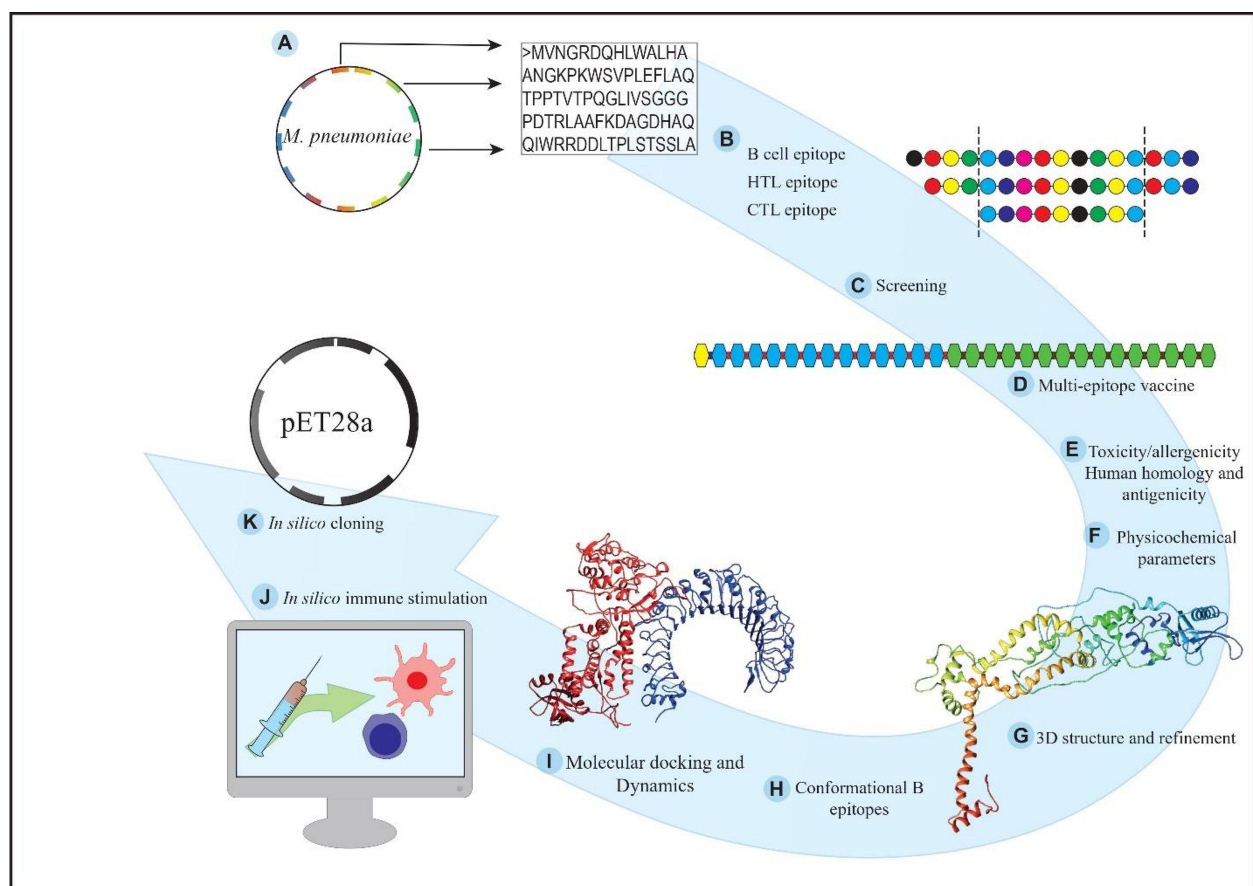


Figure 1: The overall workflow of the Methodology used in this study.

To select the number of epitopes to define the final structure carefully, we performed some filtering (**Supplementary Figure 2**):

a. It was established as a new criteria the IC50 determined through IEDB results, it was kept the overlapping HTL epitopes with B with IC50 of up to 50nM [51].
b. The CTL epitope tool, Immunogenicity of the IEDB, was used to rank the epitopes found up to that stage and perform a new screening [52] with a cut-off of 0.1, which indicates a better probability of these epitopes being immunogenic.
c. Select only CTL and HTL overlapping epitopes.
d. The remaining epitopes with only 1 or 2 different amino acid residues were filtered according to the criteria: less population coverage is excluded (according to IEDB tool), if this criterion does not apply, the percentile rank or IC50 score is evaluated, and we exclude the epitope with the highest score.

2.6 Construction of multi-epitope vaccine sequence

The epitopes that comply with all criteria determined were merged with linker peptides to construct the final sequence of multi-epitope. These linker peptides assist in protein folding and processing. For CTL epitopes, linker AAY was used, and HTL epitopes were merged with the GPGPG linker [21]. These linkers play an important role in stability and protein folding, in addition its assisting in biological activity such as antigenic presentation [53]. As adjuvant the heat-labile enterotoxin B chain precursor from *Escherichia coli* (GenBank: ALO79813.1) [54,55] was chosen to enhance the immune activation, followed by the peptide linker EAAAK. Three linear sequences of amino acids were built, always starting with the adjuvant and later changing the order of the amino acids, in order to verify which sequence would result in the best structural quality.

2.7 Antigenicity and IFN-γ inducing epitopes prediction

The vaccine's final structure was analyzed given some crucial aspects related to the induction of immune responses, allergenicity, toxicity, and solubility using tools commonly used for this type of work. First, it was used the platform VaxiJen, which asses the antigenic capacity of an amino acid sequence through the automatic cross-covariance method. Thus, analyzing the physico-chemical properties and predicting the ability to induce immune responses without the need to do alignments [56]. The search for epitopes that able to induce IFN-γ production with subsequent consequently activation of TCD4+ (HTL) and TCD8+ (CTL)

lymphocytes was performed with the IFNepitope predictor, which uses a SVM hybrid method based on motifs to perform the prediction [57].

2.8 Determination of allergenicity and toxicity potential of a chimeric protein

The chimeric protein sequence was submitted with default parameters to Toxinpred which uses a hybrid prediction method through machine learning and a quantitative matrix, combined with analyzing the physico-chemical properties for determination of toxicity epitopes [58]. Besides, the Allertop v. 2.0 server was used to asses allergenicity based on the different information obtained through the amino acid chain structure [59].

2.9 Physico-chemical properties and host homology analyses

Molecular mass, theoretical pI, extinction coefficient, aliphatic index, grand average of hydropathicity (GRAVY), estimated half-life for three model organisms (*Escherichia coli*, yeast, and mammal cells), and the instability index are the features analyzed through the final vaccine sequence using ProtParam tool [60]. Besides, the solubility index was also assessed by Protein-sol, which evaluates several particularities based on *E.coli* expression data. Autoimmunity reactions are generated by proteins similar to human proteins. To verify this similarity, a BlastP was performed with the chimeric protein vaccine against the human proteome.

2.10 Secondary structure prediction

In order to determine the secondary structure of the multi-epitope vaccine, due to its good performance and accessibility [61], it was used PSIPRED, an online tool that applies complex ANN and Position-Specific Scoring Matrix (PSSM) to predict the structure and generate the pictures, and to mensurate the percentage of β-strand, helix, and coil, was assessed by RaptorX [62,63].

2.11 Tertiary structure and refinement

For tertiary structure prediction, we use the Phyre 2 intensive model, RaptorX and I-TASSER servers, to select the best structure, as they has demonstrated high accuracy in structural prediction [64]. Phyre 2 intensive method comprises the multiple alignments of the sequence of interest with homologous sequences followed by the secondary structure's prediction with the PSIPRED. Then, information from these two steps was combined to determine a hidden Markov model. A search for this model is performed in an HMM database of known structure proteins, and the models with the best scores are used to determine the modeling and error correction [65]. RaptorX predicts 3D structure proteins using multiple-template threading (MTT) and scoring methods to indicate the quality of models predicted [63]. Finally, I-TASSER constructs the model with an iterative method based on templates according

to fragment assembly simulations with further refinement [66]. In order to verify the enhanced local and global structure of the multi-epitope vaccine, methods for refinement of amino acid side chains using light and aggressive relaxation were applied through the GalaxyRefiner [67]. To check the refined structure's quality, we used PROCHECK to generate the Ramachandran graph to compare the multi-epitope vaccine structure with the amino acid residues' geometry resulting from high-quality structures.

2.12 Prediction of conformational B cell epitopes

Conformational epitopes are indispensable in stimulating B cells effectively immune responses. Thus, the multi-epitope vaccine's refined structure was submitted to ElliPro web-based tool to predict these discontinuous epitopes. Default parameters were kept with 0.5 as a minimum score and 6 as maximum Angstrom distance [68].

2.13 Molecular docking between the chimeric protein and the TLR-2 receptor

The Toll-Like Receptor 2 (TLR2) was determined as one of the most important on the recognition and immune response against *M.pneumoniae* [69]. Thus, we performed the retrieval of two structures from RCSB PDB referring to that receptor, PDB IDs: 2z7x and 6NIG. Both structures were edited with Chimera software, where water molecules and the ligands PCJ, MAN, BMA, NDG, and NAG were removed. The 2z7x structure contained chains referring to other molecules, which were also excluded [70]. The molecular docking was performed with the multi-epitope vaccine, and the two models of TLR-2 were selected to verify the interaction between the chimeric protein and the TLR-2. The Swarmdock server was chosen to address the global and flexible docking, seeking the lowest energy conformations [71]. The server was chosen due to its results regarding to CAPRI assessment and uncomplicated access [72]. Hydrogen bonds and hydrophobic interactions were determined through LigPlot+ [73].

2.14 Molecular dynamic simulation of the receptor-ligand complex

To better understand the microscopic structural properties of the interaction between the multi-epitope vaccine and the TLR-2, the molecular dynamic simulation was performed with the Gromacs v5.0 program. The full system simulation was performed with the CHARMM27 all-atom force field. The models were solvated with an SPC/E water model in a cubic periodic box with a 1.5 nm distance from the edge of the complex atoms, and NaCl ions were introduced to neutralize the system. Besides, energy minimization was executed to ensure the quality of the system's geometry and vouch that there are no steric clashes, for this was applied the steepest descent algorithm. In the equilibration phase, the temperature was increased up to 300 K and pressure up to 1 bar during 100 ps. In production time, we used 10 ns for trajectory analysis.

The models obtained at the end of the production time were validated and taken for structural analysis.

2.15 *In silico* cloning

In silico cloning was performed to verify the capacity of cloning and expression of the multi-epitope vaccine in an appropriate expression vector. For this, we needed to adapt the codon of our peptide vaccine, according to the *E. coli* expression system's codon usage. So, we used the JCat tool for reverse translation. With the cDNA sequence, the codon optimization for *E. coli* k12 was performed, and it returned the codon adaptation index (CAI), which must have a score higher than 0.8 and the GC content rate should be between 30-70%. Furthermore, To clone the final vaccine constructs optimized gene sequence in *E. coli* pET28a(+) vector, *Blpi*, and *BamHI* restriction sites were introduced. Finally, the optimized sequence was inserted into the pET28a(+) vector using the SnapGene tool to ensure vaccine expression.

2.16 Immuno simulation

To better describe the immune response outlined by the multi-epitope vaccine, we conduct an immune stimulation through the widely used in immunoinformatics, C-ImmSim server [74]. The *in silico* method uses PSSM for epitope prediction and machine learning to assess interactions. The model also simulates the anatomical regions where important events of immune response occur: bone marrow, lymphoid and myeloid cell production; the thymus where the autoreactivity process happens and the tertiary lymphatic organ region the antigenic presentation to describe the immunogenic profile. For simulation, three injections containing 1000 vaccine proteins each were given at intervals of four weeks. Time steps are set at 1, 84, and 168 (each time step is 8 hours in real life, and time step 1 is the time of injection =0). The incremental steps have been modified to 1050, and the other parameters were kept as default. In order to check the efficacy of the multi-epitope vaccine construct, only adjuvant sequence and vaccine construct without adjuvant were also used as simulating injections maintaining the other parameters described above.

3. RESULTS

3.1 Prediction of CTL epitopes

We obtained 2063 epitopes with high frequency in the global population for all eleven proteins analyzed with the IEDB server's which can be recognized by the

MHC-I alleles. To improve the confidence of the chosen epitopes, we also submit these proteins to the NetCTL 1.2 server; the searches were performed with the allelic supertypes: A1,

A2, A3, A24, A26, B7, B8, B44, and B58, as they fit in the population incidence and with a threshold set at 0.75, more than 1300 epitopes were found in total.

3.2 Prediction of HTL and B cell epitopes

Following the previous proposal, the MHC-II epitopes with 15 amino acid residues were predicted through IEDB and NetMHCII 2.3 server. It was predicted 1623 epitopes that bind the alleles most common in the overall population according to IEDB and 1404 strong binder epitopes through NetMHCII 2.3 server. For this last prediction, we chose all the alleles available in the tool that was described in Greenbaum's work [37].

The ABCpred server predicted 718 epitopes with the length of 16 mer and scored higher than 0.51 for all the eleven proteins according to the ability to interact with B lymphocytes receptors.

3.3 Identification of epitope shared among P1 type FH and typed M129 proteins and selecting CTL and HTL epitopes with better confidence.

As the aim of the study is to utilize only epitopes that belong to all known *M. pneumoniae* lineages, we applied an in house python script to choose only those epitopes common between M129 and FH P1 proteins, and it was performed for each protein from B, CTL, and HTL cells for each predictor separately. Only these epitopes were kept on the next steps as belonging to P1 protein. In the next steps, we analyze epitopes from 10 proteins.

Aiming at selecting epitopes with improved confidence, we applied the same script to find the confidence epitopes, those that were predicted by the two methods for each category, CTL and HTL, for each protein at a time. It was found 862 epitopes predicted both by IEDB as NetCTL 1.2 server and 879 HTL epitopes predicted by IEDB and NetMHCII 2.3 server.

The total number of B cell epitopes predicted, excluding the ones that are not shared between the two types of P1 proteins, was 555, and they were used for further analysis.

3.4 Search for overlap between humoral and cellular epitopes and screening

We found 642 B and CTL overlapping epitopes and 135 epitopes that bind both to HTL and B cell receptors through the sliding window method. Those epitopes were submitted to the next screening methods.

a) The proteins WP_010874809 and WP_014325659 did not show overlapping epitopes that fit in the new IC50 threshold, thus, all their epitopes (CTL/B and HTL/B) were deleted. At this step, 61 HTL/B epitopes remain.

b) According to the immunogenicity screening, 182 CTL/B epitopes with a score higher than 0.1 kept on the analyses.

c) Those remaining epitopes were submitted to search for CTL and HTL overlapping epitopes. Then, we kept 26 MHC-I epitopes with overlapping sequences with 24 MHC-II epitopes that belong to 6 proteins.

After excluding the similar epitopes according to the aforementioned rules, 16 CTL/B and 13 HTL/B epitopes that belong to 5 proteins were considered the better qualified to construct the multi-epitope vaccine (**Table 1**).

Table 1: Final epitopes selected for vaccine construct, highlighting the overlap between CTL and HTL epitopes and each population coverage score.

	PROTEIN	HTL	POP. COVERAGE	% RANK	CTL	POP. COVERAGE	% RANK
1	WP_053344075/ WP_010874498	**DQIDFNRLFTHPV**TD	36.46%	0.87	DQIDFNRLF	53.55 %	0.1
					FNRLFTHPV	10.55%	0.1
		TMLVYD**QYIPLFIDI**	96.21%	0.41	QYIPLFIDI	40.52 %	0.04
2	WP_010874581	**AEVYGAALV**GSKNDT	74.36%	0.24	AEVYGAALV	22.86%	0.06
		FKY**NTSIVFSAL**LIL	7.15%	0.14	NTSIVFSAL	0.12%	0.16
		FR**IPLFLIESFFNT**I	98.27%	0.02	FLIESFFNT	42.84%	0.03
					IPLFLIESF	39.42 %	0.01
		GTIGFSYFLATGIVR	88.84%	0.27	GTIGFSYFL	61.37%	0.06
		GYFWTYFFLYGWYKV	96.59%	0.03	GYFWTYFFL	39.43 %	0.14
					WTYFFLYGW	16.45 %	0.08
		KVAAFFTTLFMKDRT	97.48%	0.01	VAAFFTTLF	55.61 %	0.27
3	WP_010874999	DKAK**FTADFIFSY**TP	94.3%	0.25	FTADFIFSY	54.84 %	0.03
4	WP_014325517	YTFKAS**VFENWNELL**	97.48%	0.24	VFENWNELL	26.18 %	0.75
		KQLI**SSNGLYINK**GG	6.69%	0.41	SSNGLYINK	43.03%	0.01
		KDS**IAALNTIYK**KYK	4.47%	1.5	IAALNTIYK	38.86 %	0.2
5	WP_014574866	LKRANF**QTDFTINRI**	27.97%	1.2	QTDFTINRI	19.61%	0.48

3.5 Multi-epitope vaccine sequence constructs

We have used the heat-labile enterotoxin from *E. coli* as an adjuvant to construct the multi-epitope vaccine, then the peptide linker EAAAK. The final epitopes with excellent

immunological foreseen properties were selected from the five proteins used in the vaccine construction and were joined together by appropriated linkers. A total of 13 HTL epitopes linked by GPGPG linkers and 16 CTL epitopes linked through AAY linkers were identified from five proteins out of 11 (**Figure 2A**).

3.6 Tertiary structure prediction, refinement and structure properties assessment

The best model obtained was modeled with Phyre2 from the second sequence built. The Ramachandran plot show that 74.8% of residues are in most favored regions, 18.5% in additional allowed regions, 4.0% of residues in generously allowed regions and 2.7% in disallowed regions (**Figure 2D**). Then, the refinement was carried out by the GalaxyWeb server to improve structure quality. The Ramachandran plot of the best structure had the followed scores: 87.3% of residues in most favored regions, 9.8% in additional allowed regions, 0.6% of residues in generously allowed regions and 2.3% in disallowed regions (**Figure 2E**).

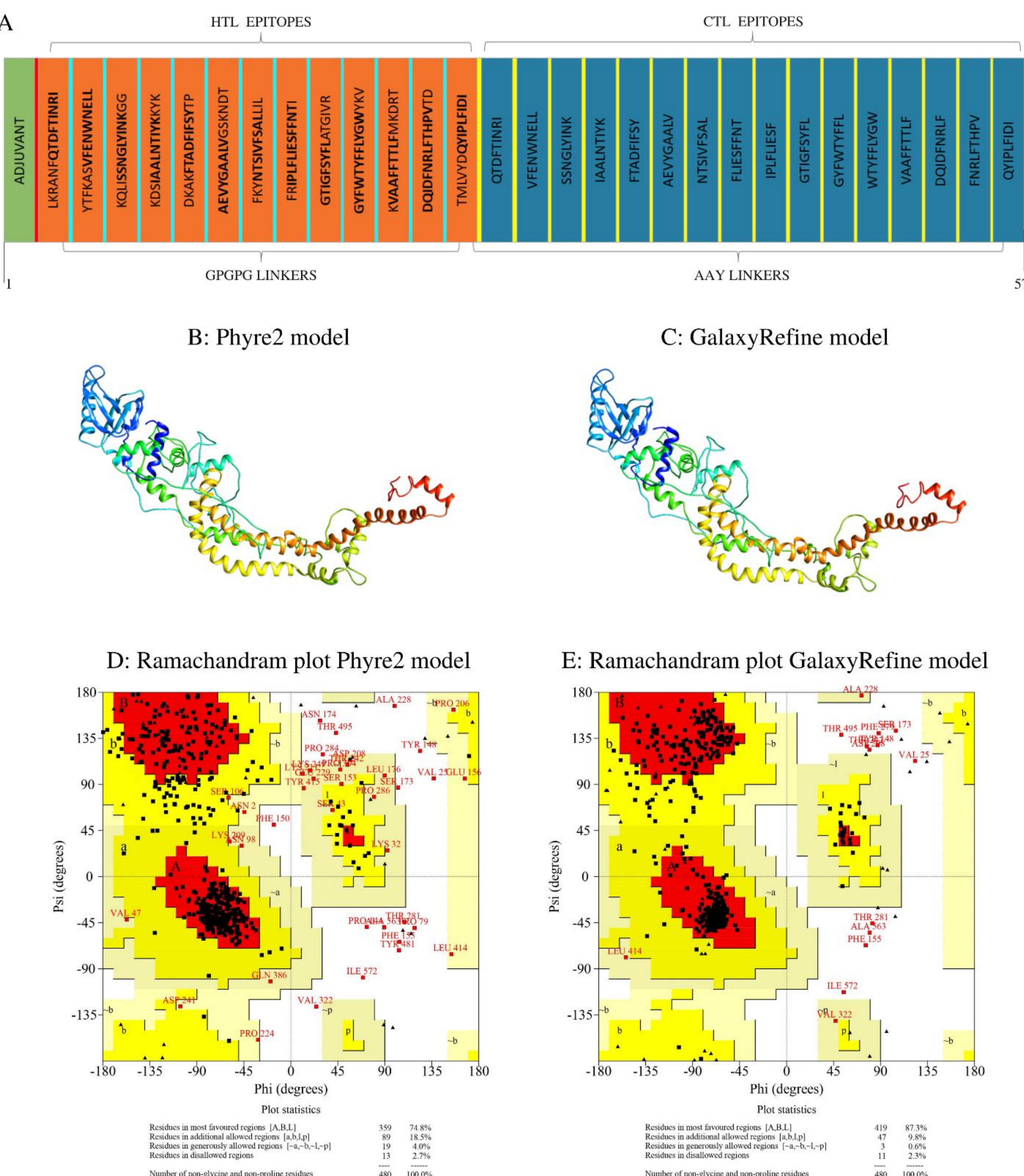


Figure 2: **Multi-epitope vaccine design and structure prediction**.

(A) The final vaccine construct with 574 amino acid residues. The heat-labile enterotoxin from *E. coli* as adjuvant shown in light green connected through the linker EAAAK in red. The HTL epitopes are shown in orange connected through the linker GPGPG in cyan, and CTL epitopes in blue connected through the linker AAY in yellow.

(B) Tertiary representation of the multi-epitope vaccine structure modeled by the Phyre2 server. (C) Represents the refined tertiary structure obtained from the GalaxyRefine server. The

Ramachandram graph represents the validation of the multi-epitope vaccine's tertiary structure using the Phyre2 (D) server; 74.8% of the amino acid residues were predicted in most favored regions. After the refinement using Galaxy algorithms (E), the refined model 4 presents the best score, with 87.3% of residues in the most favored regions.

3.7 Secondary structure prediction

The amino acid sequence of the multi-epitope vaccine was submitted to PSIPRED to investigate secondary structural properties, and it was observed that among the 574 residues of the sequence with 46% helix, 9% of β-sheet, and 43% of the loop (**Supplementary Figure 1**).

3.8 Host homology

BlastP was performed with the multi-epitope vaccine against the human proteome (taxid: 9606); no significant similarity was founded.

3.9 Predicted epitopes capable to inducing IFN- γ production, antigenicity, and linear and conformational B epitopes prediction

155 epitopes with positive and negative scores were predicted by the SVM method of the IFNepitope tool using the determined multi-epitope vaccine sequence. Among these, 57 epitopes of size 15mer had a score greater than 1 and considered more capable of inducing this cytokine production (**Supplementary Table 2**).

The multi-epitope vaccine's final structure was submitted to Ellipro. 3 linear and 4 conformational epitopes with scores above 0.7 were predicted (**Supplementary Table 3 & Figure 3**). The vaccine is probably antigenic, according to VaxiJen. The score obtained was 0.5274.

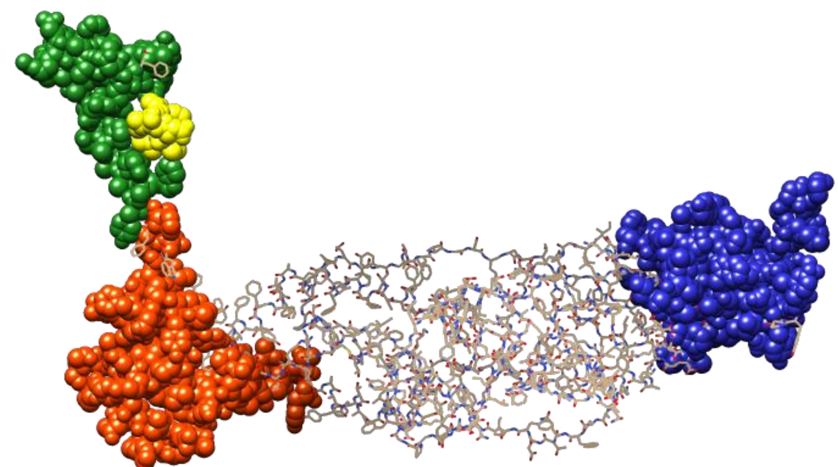


Figure 3: *Mycoplasma pneumoniae* multi-epitope vaccine tertiary structure with the location of B-cell conformational epitopes.

The predicted conformational B-cell epitopes with a PI score of 0.947 (green), 0.931 (yellow), 0.718 (orange), and 0.778 (blue) visualized in the 3D structure of the vaccine candidate.

3.10 Toxicity and allergenicity assess

According to the AllerTOP and ToxinPred tools, the multi-epitope vaccine's chimeric protein sequence has shown no prospect of being allergenic or toxic to the human organism

3.11 Physico-chemical properties

The vaccine's molecular mass is 62903.98 (62.91Kd), its isoelectric point (pI) is 8.78, showing behavior in a basic environment. The instability index II, which is related to the multi-epitope vaccine's stability, is 12.76, characterizing it as stable. The half-life in mammalian reticulocytes is 30 hours *in vitro*, > 20 hours in yeasts *in vivo*, and > 10 in *E. coli*, *in vivo*. The aliphatic index associated with protein stability in the face of temperature changes was 82.60, with high indexes indicating stability. The great average of the hydropathy value (GRAVY) is 0.202, and positive values indicate hydrophobicity. According to Protein-Sol, the predicted solubility was 0.383. This score was lower than the solubility threshold, 0.45, which is related to *E. coli* solubilities. Scores above have higher predicted solubility than the average. Further, the pI, according to this, the predictor was 9.09 [75].

3.12 Molecular docking of the multi-epitope vaccine and TLR2 receptor

We evaluate the docking results according to energy, cluster members, and the number of hydrogen bonds. As a result, the best-docked structure between TLR2 (PDB ID: 2z7x) and the multi-epitope vaccine was selected based on minimum energy value, -50.76; exhibited 18 hydrogen bonds and 22 hydrophobic interactions. Thus, this docked complex was selected for molecular dynamics simulation (**Figure 4**).

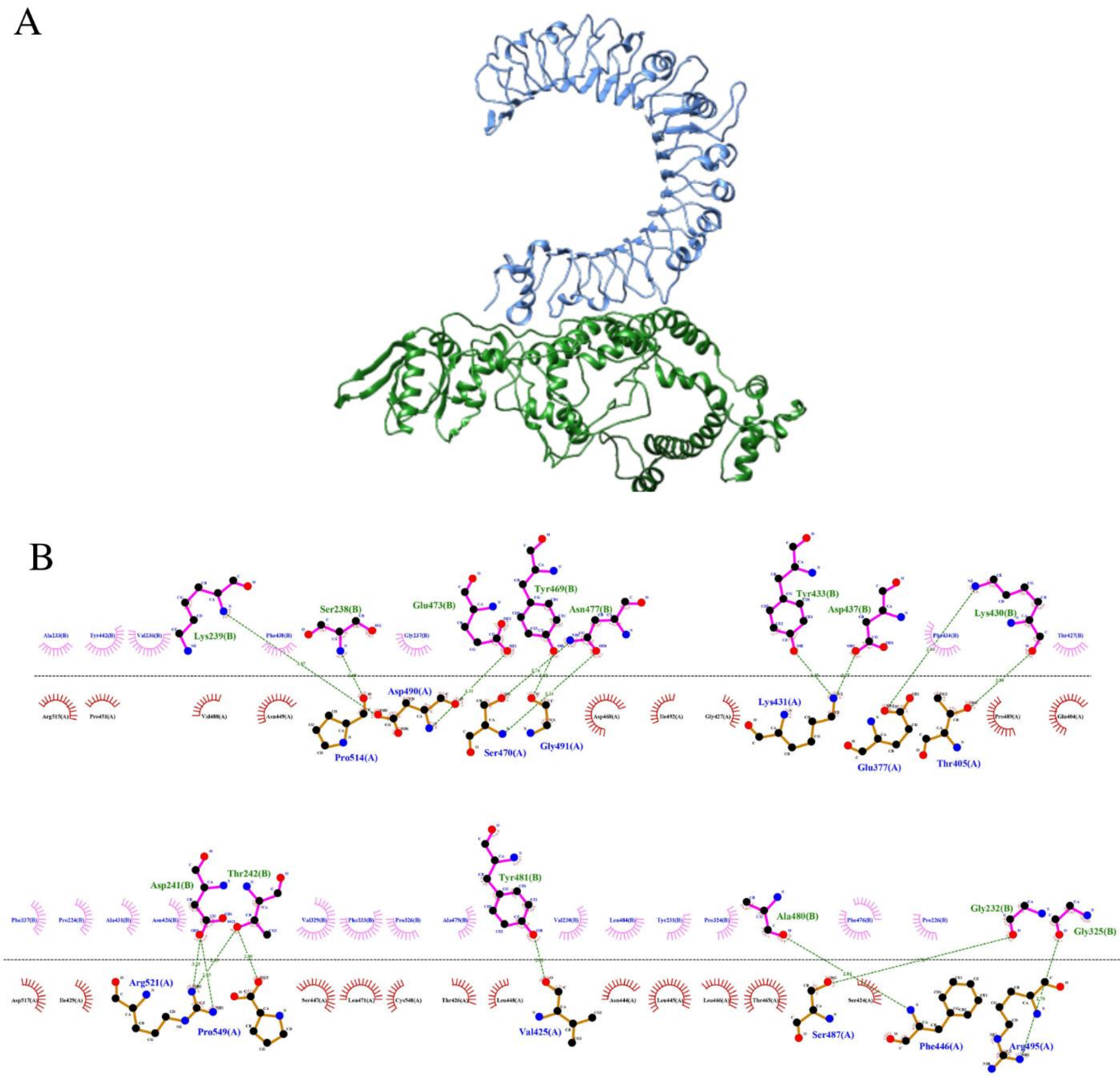


Figure 4: Molecular docking of the TLR2multi-epitope vaccine.

(A) The best dock complex was selected as the most favorable docking structure. (A) The 3D representation of the dock complex ligand (vaccine protein) is shown in green, and the receptor (TLR2) is shown in blue. (B) The 2D representation of the dock complex showing residues of chimeric protein (vaccine) - receptor (TLR) complex.

3.13 Molecular dynamics simulation of vaccine-TLR2 complex

The simulation was performed for the best-docked structure to check stable interactions between docked complexes. The minimization phase was performed to ensure that the system has no steric clashes or inappropriate geometry by 100 ps. The equilibration phases of

temperature and pressure were done to equilibrate the solvent and ions around the protein by 100 ps. The temperature and pressure evaluation plot indicate that the system maintains 300 K (**Figure 5A**) and 1 bar (**Figure 5B**). This phase was performed to check the stabilities and then run the MD production.

The system's equilibration phases are shown in figure 5A (temperature) and 5B (pressure). In both figures, we can see that the system quickly reaches the target value (300 K and 1 bar) and remains stable over the remainder of the equilibration. Two analyses were used to check the MD simulation result's stability: root-mean-square deviation (RMSD) and root-mean-square fluctuation (RMSF) Figure 5C & 5D, respectively.

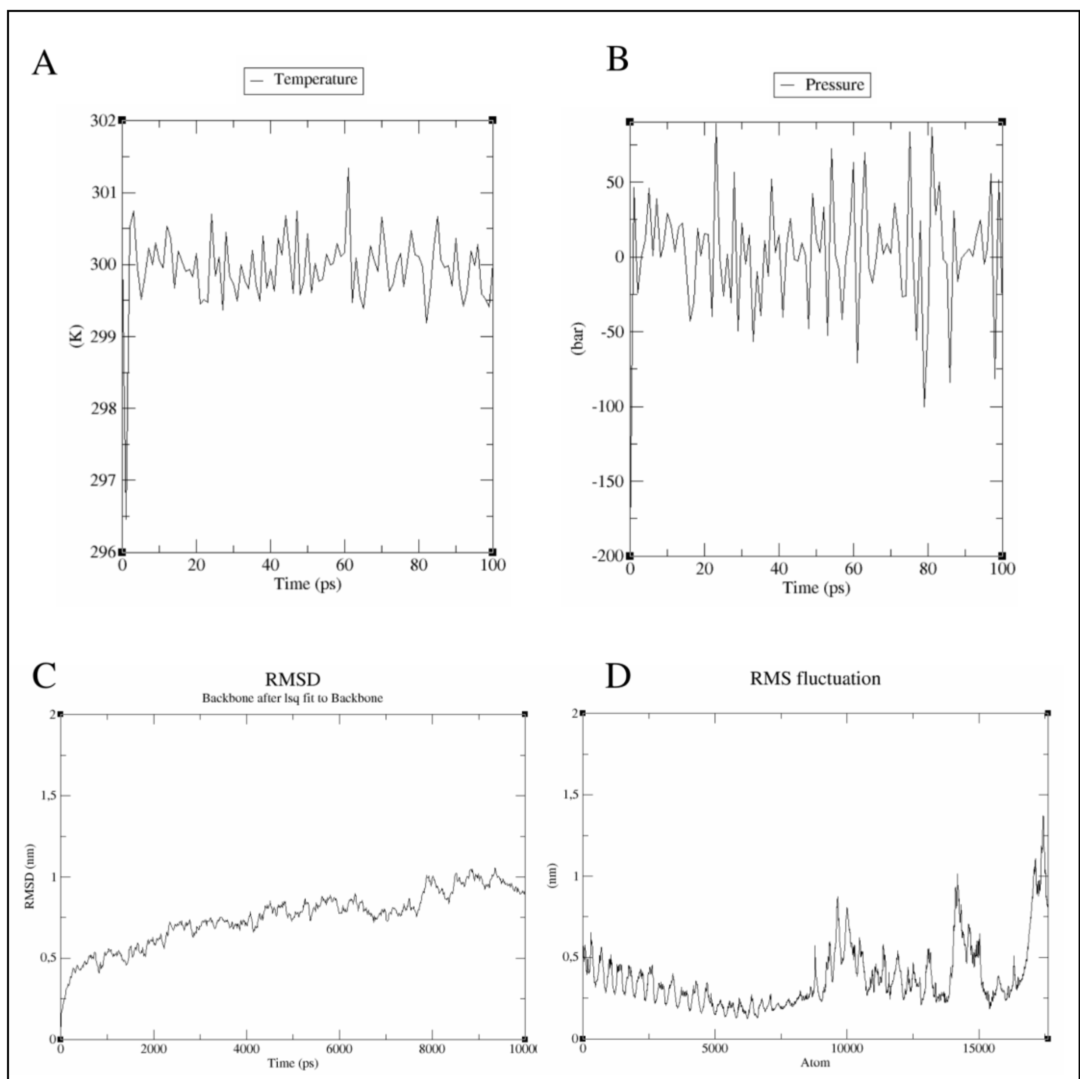


Figure 5: Molecular dynamics simulation plots of the ligand-receptor complex:

(A) Plots of the ligand-receptor (vaccine and human TLR-2) complex demonstrate that the temperature of the system reaches more than 300 K and remains around 300 k indicating the equilibration phase (100 ps) (B) These plots display the fluctuation of precision during the equilibration phase of 100 ps with pressure value on average of 0.5 bar. (C) RMSD – Root

Mean Square Deviation of the ligand-receptor complex indicates no expressive deviation, which means stable microscopic interaction. (D) RMSF- Root Mean Square Fluctuation plot, reflecting that the vaccine-TLR2 complex is flexible in the docking protein complex's side chain.

The root mean square deviation (RMSD) and the root mean square fluctuations (RMSF) were calculated to analyze the interaction between the ligand (multi-epitope vaccine) and receptor (TLR-2).

At the end of the production MD, the RMSD plot (**Figure 5C**) shows that error tends to remain constant during the rest of the simulation time, showing us the system's stability. RMSD analysis relative to the first frame in the MD trajectories showed an equilibrium tendency in the system. On the other hand, Figure 5D shows that the RMSF of all residues were computed. These fluctuations in some residues indicate the flexibility of complex with many loops area.

3.14 Codon adaptation and *in silico* cloning

Jcat software analysis for codon adaptation resulted in 51.21% of GC content in the optimum range is 30-70%. Also, the CAI index calculated was 1.0, also inside the allowed range. Using the SnapGene software, we created *BlpI* and *BamHI* restriction sites and inserted our vaccine sequence into the expression vector pET28a(+) vector. The complete length of the clone was 1721 bp (**Figure 6**).

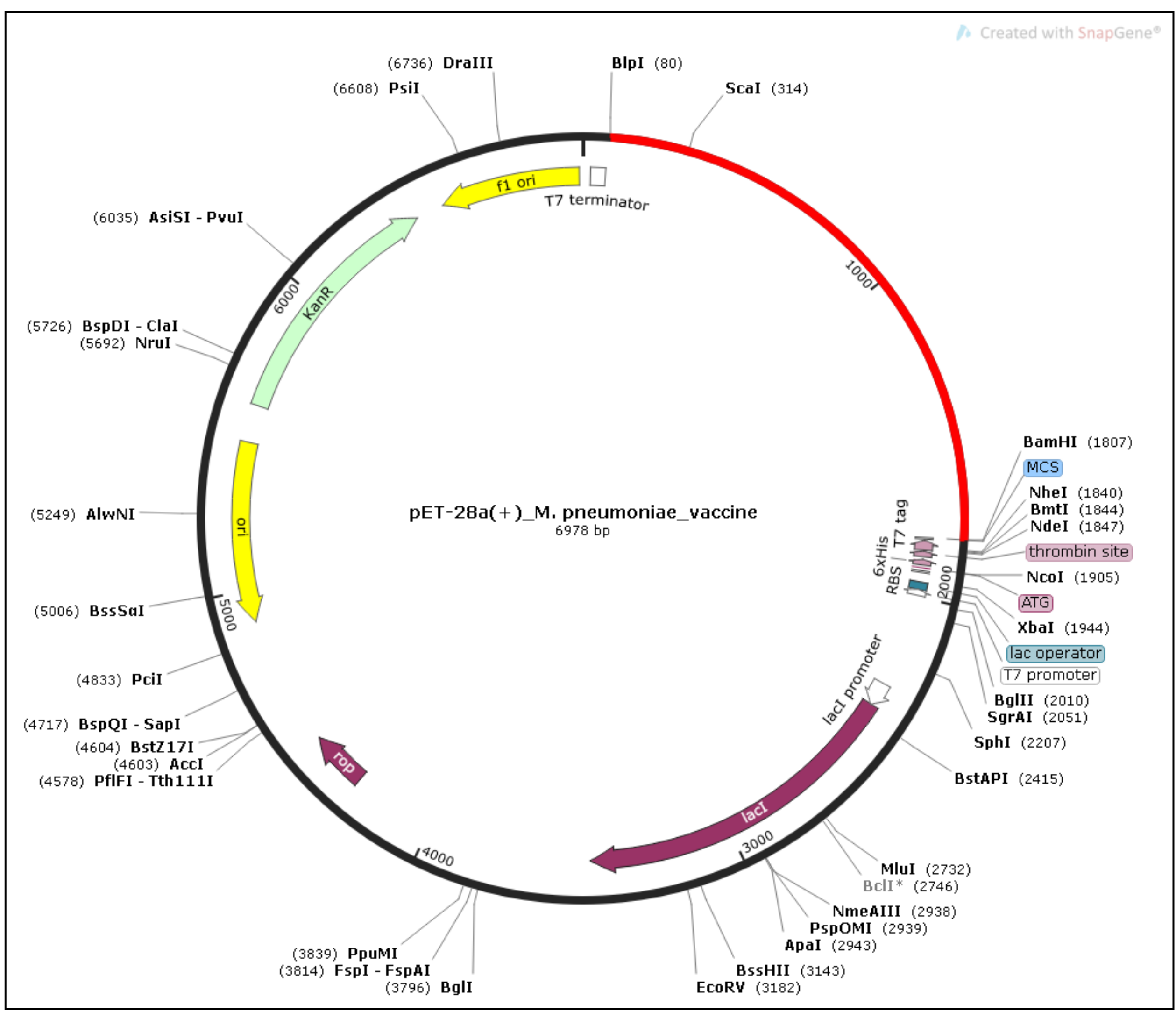


Figure 6: ***In silico*** **cloning through SnapGene (+) software**

(A) The optimized multi-epitope vaccine sequence (represented in red in vector) was inserted into the pET28a(+) vector (black circle) via the restriction enzymes BlpI and BamHI.

3.16 ***In silico*** **immune simulation**

The results generated by C-ImmSim were compatible with the real and robust development of immunity. Innate immunity plays a crucial role in initiating immune responses, antigen presentation and T cell activation. In turn, after activation, cells from adaptive immunity amplify the innate immune response. Thus, it is clear that strategies aiming at modulating both immune branches, represent the most promising candidates for a protective immunity, and therefore, vaccine development. Our vaccine was able to activate and stimulate natural killer cells (NK) and macrophages. In general, it was observed an increase, mainly in the secondary and tertiary response, on cell types and cytokines essential to sustain an effective immune response with a peak on injection days. As expected, it was observed a decrease in the level of

active cells a few days after the third injection and, together with the evidence of the putative induction of anti-inflammatory cytokines, IL-10 and TGF-β, it is plausible to assume an attempt of the immune system to control the inflammatory response, thus, avoiding undesirable exacerbated inflammation (**Figure 7, A-E**).

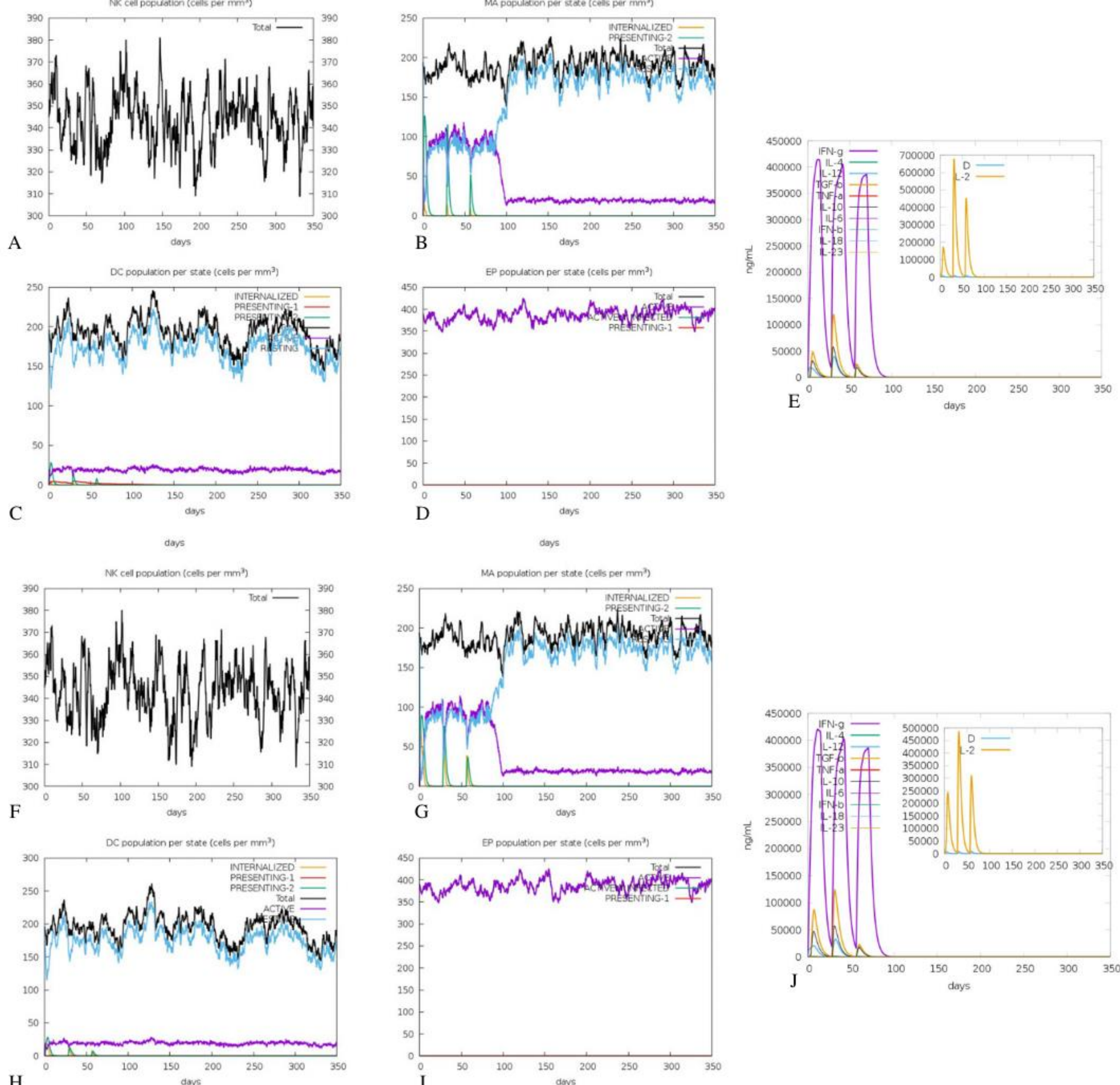


Figure 7: Immuno simulation results regarding innate immunity and cytokine production. A, B, C, D and E graphs are about the designed multi-epitope vaccine and the F-J graphs are the results regarding to the adjuvant simulation. Those graphs represents the immune simulation results about: (A and F) natural-killer cells population, (B and G) macrophages, (C and H) dendritic cells (D and I) Epithelial cellspopulation, (E and J) Cytokine levels induced during the injections and the D value, which is related to diversity of T-cells clones.

The multi-epitope vaccine increased T helper populations with a strong Th1 differentiation. Th1 lymphocytes are significant inducers of cytotoxic T lymphocyte proliferation and also increase their cytotoxic capacity, in addition, the vaccine stimulated the production of IFN-γ, a key cytokine in immune response against intracellular pathogens such as *M.pneumoniae* [76]. The results also indicate the growth of the active cytotoxic cell population with a decrease in the number of resting cells, which are essential to control intracellular bacterial infections **(Figure 8, A-E)**

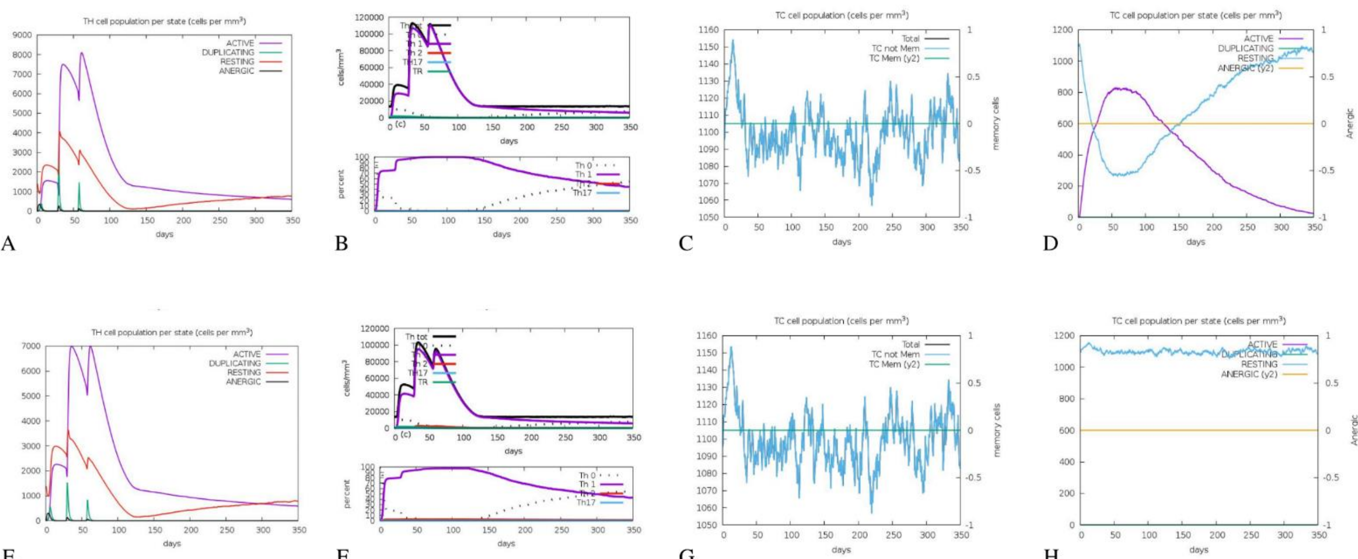


Figure 8: The Immuno simulation results regarding T lymphocytes.

The results about the multi-epitope vaccine simulation are in the graphs A, B, C and D. The simulations regarding the adjuvant are display in the E, F, G and H graphs. (A and E) Represents the helper T-cell population per state, (B and F) effector helper T-cell types fluctuation during the injections, (C and G) memory or not memory cytotoxic T-cells graph, (D and H) Cytotoxic T-cell populations per state after the injections, purple line represents the active cells and the blue line represents resting cells.

Regarding the B cell population, it is a clear growing of memory cells throughout vaccine injections with strong differentiation and production of IgM and IgG, with a decrease in non-memory B-cells. Further, there is a significant production of IgM+IgG, IgG1, and IgG1+IgG2, which are increased over the injections (**Figure 8 A, B and C).**

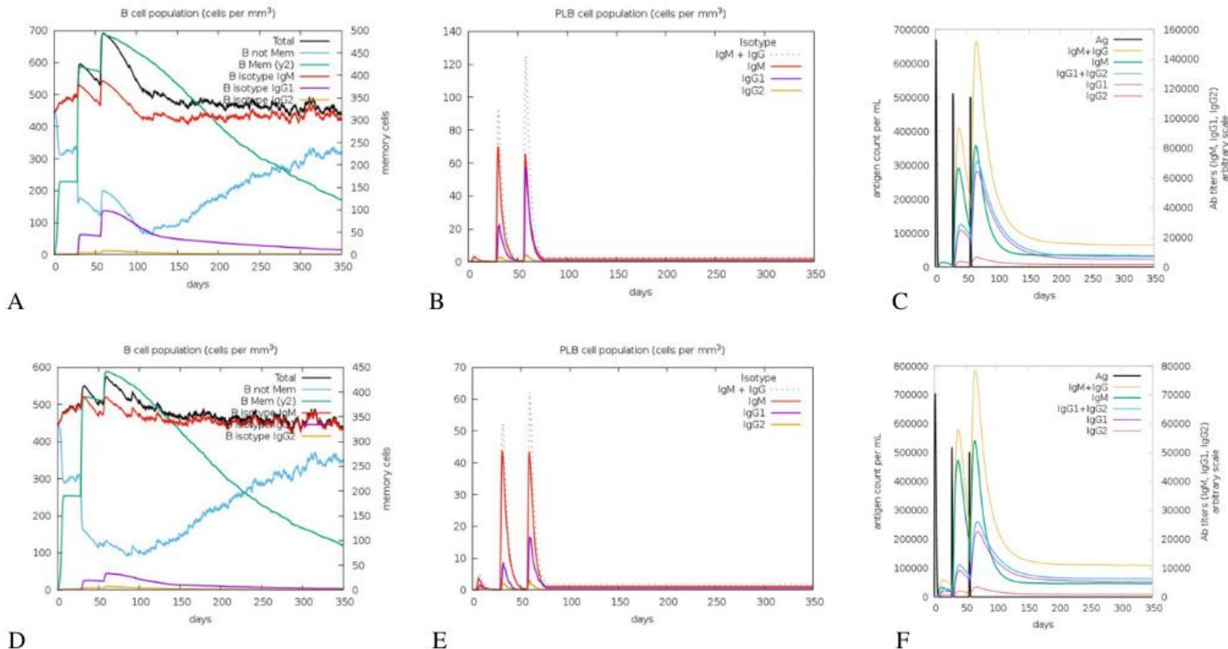


Figure 9: Immuno simulation results regarding B cell population and immunoglobulins production. The A-C are the graphs related to the multi-epitope vaccine's simulation and D-F are related to the adjuvant sequence simulation. (A, D) B-cell population per state, (B, E) changings in B-cell population in the three injections, (C, F) Differential production of immunoglobulins (colored peaks) during the antigen injections (black vertical lines).

Comparing the results of the simulation using the full vaccine (with adjuvant) to those observed with the adjuvant only and the simulation with the vaccine construct without adjuvant (**Supplementary Figure 3, 4 and 5**), the vaccine without the adjuvant induced a less pronounced proliferation of B cells however the levels of cytokine production shown an inverse result (Figure 9 A and D, C and F). Thus, suggesting the contribution of the selected epitopes to the induction of B lymphocytes *in silico*. On the other hand, the production of immunoglobulins seemed to be more dependent on the presence of the adjuvant (**Figure 9**). Further, the stimulation of CTL was very similar in both simulations (**Figure 8**).

4.0 DISCUSSION

Vaccines are the best strategy to prevent infectious diseases and to elicit protective immunity. Current immunoinformatics approaches have accelerated the production process and allowed the screening of genomes to identify targets with the greatest potential to be used as vaccine candidates, such as multi-epitope vaccines. This vaccine platform is designed from

selected epitopes to stimulate innate and adaptative immune responses, thus inducing protective memory. Also, this approach is associated with a reduced occurrence of side effects and spontaneous reversions that may occur in attenuated vaccines [77]. Recent studies, have demonstrated positive results with animal tests with multi-epitope vaccines, indicating this plataform as a promising and safe method [78,79].

Unni and collaborators proposed, through immunoinformatics tools [80], epitopes with good potential to induce immune memory responses and be used as a vaccine against *M. pneumoniae*. Also Marvat and associates designed a epitope based vaccine against *M. pneumoniae* utilizing the proteome of one pathogenic strain as data base [81]. Their methodology besides quite similar to ours, differs in some important points. The proteins submitted for epitope prediction used at the beginning of our work, belongs to the core genome of the 88 known *M.pneumoniae* strains and was chosen through reverse vaccinology approaches [33]. This aspect implies a important role to strains coverage, and improves the vaccine potential regarding the diversity of the specie. Further, to understand the immunological potential of the selected epitopes, we performed the immune simulation with the different pieces of the designed chimeric protein, with the aim to check if the result generated by the complete protein is not just an adjuvant effect.

Another important concern in vaccine development concerns the effectiveness among global populations, due to the great number of human leukocyte antigen (HLA) polymorphisms existing. In this regard, in the primary steps we predicted the epitopes that can be recognized by frequently HLA alleles in the global population; these epitopes overlapped among CTL, HTL, and B-cell epitopes to stimulate both cellular and humoral immunity. The assembly between epitopes to construct the vaccine was achieved by using linkers. In the perspective of multi-epitope vaccines, the main advantages of using linkers are the reduced probability of junctional antigens formation, ability to improve antigen processing and presentation, structural flexibility and rigidity [53].

Despite the ability of inducing immune response, classical vaccines platforms can be formed by regions able to induce allergic reactions as they are formulated with several proteins from the target organism[82]. In contrast, the epitope-based vaccines, like the one proposed in this study, was designed with carefully selected epitopes that drastically reduced the possibility of undesirable side effects problems [83], thus, our vaccine candidate can be considered safe and non-allergenic. The half-life results suggests that our vaccine is a promising candidate as one of the biggest challenges regarding the use of proteins for therapeutic purposes is its low half-life [84]. Besides, the antigenic potential evidenced by our vaccine suggest a proper

activation of the immune system. Bench tests should be performed to ensure the safety of the vaccine. In terms of antigenicity; it needs an enhancer due to the lack of antigenic determinants, contrasted with whole pathogen vaccines. The heat-labile enterotoxin from *E,coli* was chosen as adjuvant due its strong antigenic properties to enhance and regulate immunogenicity (85, 86). Those aspects, together with dynamics simulation results, are crucial as they demonstrate stability resulting from the analyzes, suggesting adequate perspectives for large-scale production. Further, the *in silico* cloning was performed properly, with parameters within the expected boundaries, strengthening this production potential.

Even with high performance algorithms of the *in silico* predictors, one of the biggest challenges in the field of vaccines is to stimulate the immune system properly [85]. *M. pneumoniae* is known to induce strong inflamatory responses during infections and its virulence factors as lipoproteins and adhesines are important mediators of this process [86]. The selected epitopes belong to the protein part of lipoproteins (WP_010874999 and WP_014325517), the P1 adhesine (belonging to both types, FH – WP_053344075 and M129 – WP_010874498) and a pro-liporpotein diacylglyceryl transferase (WP_010874581.1). The lipoproteins are key components of *M.pneumoniae* virulence and immunogenicity, being recognized by Toll-like receptors, which improves the pro-inflammatory cytokine liberation and the cellular recruitment [87]. The P1 adhesine is essential in cytoadherence process, and this interaction is related to TLR-4 recognition in inflammasome induction and authophagy mechanisms [88].

To build memory immunity against a specific pathogen and protect to possible infections, the vaccine must trigger innate immunity components which are indispensable in the primary recognition. The innate system also makes the link to the adaptative immunity where the CD4 and CD8 lymphocytes recognize MHC molecules to start specific responses targeting the vaccine components. The adaptative immune system is all connected to also induce the production of antibodies by specialized B-cells, which are crucial in pathogen neutralization during infections, this capacity of vaccines to induce the production of specific antibodies was one of the biggest challenges and that was explored in this work.

The *in silico* immune simulation was performed to understand our vaccine's immune capacity predicted against *M.pneumoniae*. The results demonstrated that our multi-epitope vaccine is a potent inducer of different innate and adaptive immunity, with wide differentiation of B-cell populations and production of antibodies vital in pulmonary clearance in *M. pneumoniae* infections [89]. Further, the immune simulation showed good ability to induce the production of key cytokines in immunity against *M.pneumoniae* (IFN- γ) and immune balance such as TGF-β and Il-10 which give us a reasonable expectation not to induce exaggerated

responses. The *in silico* immune simulation comparation allowed the differentiation of responses mediated by the adjuvant from those induced by our vaccine, with the selected epitopes using the methods proposed here. Although some aspects of the immunology were quite similar, some interesting points were found, such as the proposed vaccine's unique ability to induce cytotoxic T cells' proliferation, which is also crucial in *M.pneumoniae* immunity development [87].

5.0 CONCLUSION

In this study immunoinformatics was used to develop a potent multi-epitope vaccine candidate against *M.pneumoniae*. Several criteria were applied to select the best epitopes to induce both humoral and cellular responses. The analysis also related to allergenicity and toxicity showed that the vaccine is safe. Therefore, a promising vaccine candidate for *M. pneumoniae* was designed with less potential to induce exacerbated reaction, however, to ensure the immunologic efficiency and memory development, validation experiments need to be performed.

6.0 ACKNOWLEDGMENT

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Finance Code 001, and FAPEMIG (Fundação de Amparo à Pesquisa de Minas Gerais).

7.0 AUTHOR CONTRIBUTIONS

TCVR, ST, conceived, designed the protocol, collected and analyzed initial data, wrote the paper; TCVR, ST, HSC, SCS, VA, coordinated and led the entire project: AKJ, ST, SFOT, RBK, SCS, MVS, MRL, DB,VA, LCJA. Cross-checked all data, re-analysis: All authors read and approved the manuscript.

8.0 CONFLICTS OF INTEREST

The authors declare that they have no conflict of interest.

9.0 REFERENCES

[1] T.M. File, Community-acquired pneumonia, in: Lancet, Lancet Publishing Group, 2003: pp. 1991–2001. https://doi.org/10.1016/S0140-6736(03)15021-0.

[2] S. Jain, D.J. Williams, S.R. Arnold, K. Ampofo, A.M. Bramley, C. Reed, C. Stockmann, E.J. Anderson, C.G. Grijalva, W.H. Self, Y. Zhu, A. Patel, W. Hymas, J.D. Chappell, R.A. Kaufman, J.H. Kan, D. Dansie, N. Lenny, D.R. Hillyard, L.M. Haynes, M. Levine, S. Lindstrom, J.M. Winchell, J.M. Katz, D. Erdman, E. Schneider, L.A. Hicks, R.G. Wunderink, K.M. Edwards, A.T. Pavia, J.A. McCullers, L. Finelli, Community-Acquired Pneumonia Requiring Hospitalization among U.S. Children, N. Engl. J. Med. 372 (2015) 835–845. https://doi.org/10.1056/NEJMoa1405870.

[3] D.C. Angus, W.T. Linde-Zwirble, J. Lidicker, G. Clermont, J. Carcillo, M.R. Pinsky,

Epidemiology of severe sepsis in the United States: Analysis of incidence, outcome, and associated costs of care, Crit. Care Med. 29 (2001) 1303–1310. https://doi.org/10.1097/00003246-200107000-00002.

[4] A.K.C. Leung, A.H.C. Wong, K.L. Hon, Community-Acquired Pneumonia in Children, Recent Pat. Inflamm. Allergy Drug Discov. 12 (2018) 136–144. https://doi.org/10.2174/1872213X12666180621163821.

[5] M. Sultana, A.R. Sarker, N. Ali, R. Akram, L. Gold, Economic evaluation of community acquired pneumonia management strategies: A systematic review of literature, PLoS One. 14 (2019). https://doi.org/10.1371/journal.pone.0224170.

[6] K. McIntosh, Community-Acquired Pneumonia in Children, N. Engl. J. Med. 346 (2002) 429–437. https://doi.org/10.1056/NEJMra011994.

[7] R.G. Wunderink, G.W. Waterer, Community-acquired pneumonia, N. Engl. J. Med. 370 (2014) 543–551. https://doi.org/10.1056/NEJMcp1214869.

[8] Q.-S. Xing, G.-J. Li, Y.-H. Xing, T. Chen, W.-J. Li, W. Ni, K. Deng, R.-Q. Gao, C.-Z. Chen, Y. Gao, Q. Li, G.-L. Yu, J.-N. Tong, W. Li, G.-L. Hao, Y. Sun, A. Zhang, Q. Wu, Z.-P. Li, S.-L. Pan, Precautions are Needed for COVID-19 Patients with Coinfection of Common Respiratory Pathogens, MedRxiv. (2020) 2020.02.29.20027698. https://doi.org/10.1101/2020.02.29.20027698.

[9] G. 2016 C. of D. Collaborators, Global, regional, and national age-sex specifc mortality for 264 causes of death, 1980-2016: A systematic analysis for the Global Burden of Disease Study 2016, Lancet. 390 (2017) 1151–1210. https://doi.org/10.1016/S0140-6736(17)32152-9.

[10] J. Ottosen, H. Evans, Pneumonia, Surg. Clin. North Am. 94 (2014) 1305–1317. https://doi.org/10.1016/j.suc.2014.09.001.

[11] S. Razin, D. Yogev, Y. Naot, Molecular Biology and Pathogenicity of Mycoplasmas, Microbiol. Mol. Biol. Rev. 62 (1998) 1094–1156. https://doi.org/10.1128/mmbr.62.4.1094-1156.1998.

[12] N. Cimolai, Mycoplasmas, in: Lab. Diagnosis Bact. Infect., CRC Press, 2001: pp. 861–892. https://doi.org/10.1136/bmj.1.5688.99-c.

[13] C. Bébéar, S. Pereyre, O. Peuchant, *Mycoplasma pneumoniae* : susceptibility and resistance to antibiotics, Future Microbiol. 6 (2011) 423–431. https://doi.org/10.2217/fmb.11.18.

[14] S. Pereyre, J. Goret, C. Bébéar, Mycoplasma pneumoniae: Current knowledge on macrolide resistance and treatment, Front. Microbiol. 7 (2016) 974. https://doi.org/10.3389/fmicb.2016.00974.

[15] J.W. Dorigo-Zetsma, B. Wilbrink, J. Dankert, S.A.J. Zaat, Mycoplasma pneumoniae P1 type 1- and type 2-specific sequences within the P1 cytadhesin gene of individual strains, Infect. Immun. 69 (2001) 5612–5618. https://doi.org/10.1128/IAI.69.9.5612-5618.2001.

[16] J.B. Baseman, R.M. Cole, D.C. Krause, D.K. Leith, Molecular basis for cytadsorption of Mycoplasma pneumoniae, J. Bacteriol. 151 (1982) 1514–1522. https://doi.org/10.1128/jb.151.3.1514-1522.1982.

[17] Molecular Biology and Pathogenicity of Mycoplasmas, Springer US, 2002. https://doi.org/10.1007/b113360.

[18] Mycoplasma pneumoniae | Home | CDC, (n.d.). https://www.cdc.gov/pneumonia/atypical/mycoplasma/index.html (accessed March 25, 2021).

[19] I. Linchevski, E. Klmenet, R. Nir-Paz, Mycoplasma pneumoniae vaccine protective efficacy and adverse reactions-Systematic review and meta-analysis, Vaccine. 27 (2009) 2437–2446. https://doi.org/10.1016/j.vaccine.2009.01.135.

[20] T.P. Atkinson, M.F. Balish, K.B. Waites, Epidemiology, clinical manifestations,

pathogenesis and laboratory detection of Mycoplasma pneumoniae infections, FEMS Microbiol. Rev. 32 (2008) 956–973. https://doi.org/10.1111/j.1574-6976.2008.00129.x.
[21] V. Chauhan, T. Rungta, K. Goyal, M.P. Singh, Designing a multi-epitope based vaccine to combat Kaposi Sarcoma utilizing immunoinformatics approach, Sci. Rep. 9 (2019) 1–15. https://doi.org/10.1038/s41598-019-39299-8.
[22] A.B. Mara, T.D. Gavitt, E.R. Tulman, S.J. Geary, S.M. Szczepanek, Lipid moieties of Mycoplasma pneumoniae lipoproteins are the causative factor of vaccine-enhanced disease, Npj Vaccines. 5 (2020). https://doi.org/10.1038/s41541-020-0181-x.
[23] A. Sette, J. Fikes, Epitope-based vaccines: An update on epitope identification, vaccine design and delivery, Curr. Opin. Immunol. 15 (2003) 461–470. https://doi.org/10.1016/S0952-7915(03)00083-9.
[24] A. Sette, M. Newman, B. Livingston, D. McKinney, J. Sidney, G. Ishioka, S. Tangri, J. Alexander, J. Fikes, R. Chestnut, Optimizing vaccine design for cellular processing, MHC binding and TCR recognition, Tissue Antigens. 59 (2002) 443–451. https://doi.org/10.1034/j.1399-0039.2002.590601.x.
[25] L. Zhang, Multi-epitope vaccines: A promising strategy against tumors and viral infections, Cell. Mol. Immunol. 15 (2018) 182–184. https://doi.org/10.1038/cmi.2017.92.
[26] N. Hajighahramani, N. Nezafat, M. Eslami, M. Negahdaripour, S.S. Rahmatabadi, Y. Ghasemi, Immunoinformatics analysis and in silico designing of a novel multi-epitope peptide vaccine against Staphylococcus aureus, Infect. Genet. Evol. 48 (2017) 83–94. https://doi.org/10.1016/j.meegid.2016.12.010.
[27] W.Y. Zhou, Y. Shi, C. Wu, W.J. Zhang, X.H. Mao, G. Guo, H.X. Li, Q.M. Zou, Therapeutic efficacy of a multi-epitope vaccine against Helicobacter pylori infection in BALB/c mice model, Vaccine. 27 (2009) 5013–5019. https://doi.org/10.1016/j.vaccine.2009.05.009.
[28] Y. Lu, J. Ding, W. Liu, Y.H. Chen, A candidate vaccine against influenza virus intensively improved the immunogenicity of a neutralizing epitope., Int. Arch. Allergy Immunol. 127 (2002) 245–250. https://doi.org/10.1159/000053869.
[29] S.F. de O. Tosta, M.S. Passos, R. Kato, Á. Salgado, J. Xavier, A.K. Jaiswal, S.C. Soares, V. Azevedo, M. Giovanetti, S. Tiwari, L.C.J. Alcantara, Multi-epitope based vaccine against yellow fever virus applying immunoinformatics approaches, J. Biomol. Struct. Dyn. (2020). https://doi.org/10.1080/07391102.2019.1707120.
[30] R. He, X. Yang, C. Liu, X. Chen, L. Wang, M. Xiao, J. Ye, Y. Wu, L. Ye, Efficient control of chronic LCMV infection by a CD4 T cell epitope-based heterologous prime-boost vaccination in a murine model, Cell. Mol. Immunol. 15 (2018) 815–826. https://doi.org/10.1038/cmi.2017.3.
[31] M. Yayoshi, M. Araake, E. Hayatsu, T. Takezawa, M. Yoshioka, Immunogenicity and Protective Effect of Hemolysis Mutants of *Mycoplasma pneumoniae*, Microbiol. Immunol. 29 (1985) 1029–1037. https://doi.org/10.1111/j.1348-0421.1985.tb00893.x.
[32] I. Linchevski, E. Klmenet, R. Nir-Paz, Mycoplasma pneumoniae vaccine protective efficacy and adverse reactions-Systematic review and meta-analysis, Vaccine. 27 (2009) 2437–2446. https://doi.org/10.1016/j.vaccine.2009.01.135.
[33] T.C. Vilela Rodrigues, A.K. Jaiswal, A. de Sarom, L. de Castro Oliveira, C.J. Freire Oliveira, P. Ghosh, S. Tiwari, F.M. Miranda, L. de Jesus Benevides, V. Ariston de Carvalho Azevedo, S. de Castro Soares, Reverse vaccinology and subtractive genomics reveal new therapeutic targets against *Mycoplasma pneumoniae* : a causative agent of pneumonia, R. Soc. Open Sci. 6 (2019) 190907. https://doi.org/10.1098/rsos.190907.
[34] A.S. Juncker, M. V. Larsen, N. Weinhold, M. Nielsen, S. Brunak, O. Lund, Systematic characterisation of cellular localisation and expression profiles of proteins containing

MHC ligands, PLoS One. 4 (2009). https://doi.org/10.1371/journal.pone.0007448.
[35] S. Martini, M. Nielsen, B. Peters, A. Sette, The Immune Epitope Database and Analysis Resource Program 2003–2018: reflections and outlook, Immunogenetics. 72 (2020) 57–76. https://doi.org/10.1007/s00251-019-01137-6.
[36] R. Vita, S. Mahajan, J.A. Overton, S.K. Dhanda, S. Martini, J.R. Cantrell, D.K. Wheeler, A. Sette, B. Peters, The Immune Epitope Database (IEDB): 2018 update, Nucleic Acids Res. 47 (2019) D339–D343. https://doi.org/10.1093/nar/gky1006.
[37] J. Greenbaum, J. Sidney, J. Chung, C. Brander, B. Peters, A. Sette, Functional classification of class II human leukocyte antigen (HLA) molecules reveals seven different supertypes and a surprising degree of repertoire sharing across supertypes, Immunogenetics. 63 (2011) 325–335. https://doi.org/10.1007/s00251-011-0513-0.
[38] M.V. Larsen, C. Lundegaard, K. Lamberth, S. Buus, S. Brunak, O. Lund, M. Nielsen, An integrative approach to CTL epitope prediction: A combined algorithm integrating MHC class I binding, TAP transport efficiency, and proteasomal cleavage predictions, Eur. J. Immunol. 35 (2005) 2295–2303. https://doi.org/10.1002/eji.200425811.
[39] M. V. Larsen, C. Lundegaard, K. Lamberth, S. Buus, O. Lund, M. Nielsen, Large-scale validation of methods for cytotoxic T-lymphocyte epitope prediction, BMC Bioinformatics. 8 (2007) 424. https://doi.org/10.1186/1471-2105-8-424.
[40] H. Lin, G. Zhang, S. Tongchusak, E.L. Reinherz, V. Brusic, Evaluation of MHC-II peptide binding prediction servers: applications for vaccine research, BMC Bioinformatics. 9 (2008) S22. https://doi.org/10.1186/1471-2105-9-s12-s22.
[41] P. Wang, J. Sidney, C. Dow, B. Mothé, A. Sette, B. Peters, A systematic assessment of MHC class II peptide binding predictions and evaluation of a consensus approach, PLoS Comput. Biol. 4 (2008). https://doi.org/10.1371/journal.pcbi.1000048.
[42] K.K. Jensen, M. Andreatta, P. Marcatili, S. Buus, J.A. Greenbaum, Z. Yan, A. Sette, B. Peters, M. Nielsen, Improved methods for predicting peptide binding affinity to MHC class II molecules, Immunology. 154 (2018) 394–406. https://doi.org/10.1111/imm.12889.
[43] S. Southwood, J. Sidney, A. Kondo, M.F. del Guercio, E. Appella, S. Hoffman, R.T. Kubo, R.W. Chesnut, H.M. Grey, A. Sette, Several common HLA-DR types share largely overlapping peptide binding repertoires., J. Immunol. 160 (1998) 3363–73.
[44] M. Nielsen, O. Lund, NN-align. An artificial neural network-based alignment algorithm for MHC class II peptide binding prediction, BMC Bioinformatics. 10 (2009) 296. https://doi.org/10.1186/1471-2105-10-296.
[45] K.A. Galanis, K.C. Nastou, N.C. Papandreou, G.N. Petichakis, V.A. Iconomidou, Linear B-cell epitope prediction: a performance review of currently available methods, BioRxiv. (2019) 833418. https://doi.org/10.1101/833418.
[46] Prediction of B-cell Linear Epitopes with a Combination of Support Vector Machine Classification and Amino Acid Propensity Identification, (n.d.).
[47] Y. El-Manzalawy, V. Honavar, Recent advances in B-cell epitope prediction methods, Immunome Res. 6 (2010) S2. https://doi.org/10.1186/1745-7580-6-S2-S2.
[48] J. Chen, H. Liu, J. Yang, K.C. Chou, Prediction of linear B-cell epitopes using amino acid pair antigenicity scale, Amino Acids. 33 (2007) 423–428. https://doi.org/10.1007/s00726-006-0485-9.
[49] S. Saha, G.P.S. Raghava, Prediction of continuous B-cell epitopes in an antigen using recurrent neural network, Proteins Struct. Funct. Genet. 65 (2006) 40–48. https://doi.org/10.1002/prot.21078.
[50] Y. EL-Manzalawy, D. Dobbs, V. Honavar, PREDICTING FLEXIBLE LENGTH LINEAR B-CELL EPITOPES, Comput. Syst. Bioinforma. 7 (2008) 121.
[51] W. Zhao, X. Sher, Systematically benchmarking peptide-MHC binding predictors: From

synthetic to naturally processed epitopes, PLoS Comput. Biol. 14 (2018). https://doi.org/10.1371/journal.pcbi.1006457.
[52] J.J.A. Calis, M. Maybeno, J.A. Greenbaum, D. Weiskopf, A.D. De Silva, A. Sette, C. Keşmir, B. Peters, Properties of MHC Class I Presented Peptides That Enhance Immunogenicity, PLoS Comput. Biol. 9 (2013). https://doi.org/10.1371/journal.pcbi.1003266.
[53] X. Chen, J.L. Zaro, W.C. Shen, Fusion protein linkers: Property, design and functionality, Adv. Drug Deliv. Rev. 65 (2013) 1357–1369. https://doi.org/10.1016/j.addr.2012.09.039.
[54] C. Zhu, S. Wang, S. Hu, M. Yu, Y. Zeng, X. You, J. Xiao, Y. Wu, Protective efficacy of a Mycoplasma pneumoniae P1C DNA vaccine fused with the B subunit of Escherichia coli heat-labile enterotoxin, Can. J. Microbiol. 58 (2012) 802–810. https://doi.org/10.1139/W2012-051.
[55] Y. Ma, Recent advances in nontoxic Escherichia coli heat-labile toxin and its derivative adjuvants, Expert Rev. Vaccines. 15 (2016) 1361–1371. https://doi.org/10.1080/14760584.2016.1182868.
[56] I.A. Doytchinova, D.R. Flower, VaxiJen: A server for prediction of protective antigens, tumour antigens and subunit vaccines, BMC Bioinformatics. 8 (2007) 4. https://doi.org/10.1186/1471-2105-8-4.
[57] S.K. Dhanda, P. Vir, G.P.S. Raghava, Designing of interferon-gamma inducing MHC class-II binders, Biol. Direct. 8 (2013) 30. https://doi.org/10.1186/1745-6150-8-30.
[58] S. Gupta, P. Kapoor, K. Chaudhary, A. Gautam, R. Kumar, G.P.S. Raghava, In Silico Approach for Predicting Toxicity of Peptides and Proteins, PLoS One. 8 (2013) e73957. https://doi.org/10.1371/journal.pone.0073957.
[59] I. Dimitrov, I. Bangov, D.R. Flower, I. Doytchinova, AllerTOP v.2 - A server for in silico prediction of allergens, J. Mol. Model. 20 (2014) 1–6. https://doi.org/10.1007/s00894-014-2278-5.
[60] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, in: Proteomics Protoc. Handb., Humana Press, Totowa, NJ, 2005: pp. 571–607. https://doi.org/10.1385/1-59259-890-0:571.
[61] W. Pirovano, J. Heringa, Protein secondary structure prediction., Methods Mol. Biol. 609 (2010) 327–348. https://doi.org/10.1007/978-1-60327-241-4_19.
[62] D.T. Jones, Protein secondary structure prediction based on position-specific scoring matrices, J. Mol. Biol. 292 (1999) 195–202. https://doi.org/10.1006/jmbi.1999.3091.
[63] M. Källberg, H. Wang, S. Wang, J. Peng, Z. Wang, H. Lu, J. Xu, Template-based protein structure modeling using the RaptorX web server, Nat. Protoc. 7 (2012) 1511–1522. https://doi.org/10.1038/nprot.2012.085.
[64] R. Gosto, Assessment of Accuracies of Protein 3-Dimensional Prediction Software, Southeast Eur. J. Soft Comput. 7 (2018). https://doi.org/10.21533/scjournal.v7i2.161.
[65] L.A. Kelley, S. Mezulis, C.M. Yates, M.N. Wass, M.J.E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis, Nat. Protoc. 10 (2015) 845–858. https://doi.org/10.1038/nprot.2015.053.
[66] J. Yang, Y. Zhang, Protein Structure and Function Prediction Using I-TASSER, Curr. Protoc. Bioinforma. 52 (2015) 5.8.1-5.8.15. https://doi.org/10.1002/0471250953.bi0508s52.
[67] GalaxyWEB server for protein structure prediction and refinement, (n.d.). https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3394311/ (accessed August 4, 2020).
[68] J. Ponomarenko, H.H. Bui, W. Li, N. Fusseder, P.E. Bourne, A. Sette, B. Peters, ElliPro: A new structure-based tool for the prediction of antibody epitopes, BMC Bioinformatics.

9 (2008). https://doi.org/10.1186/1471-2105-9-514.
[69] W. Love, N. Dobbs, L. Tabor, J.W. Simecka, Toll-like receptor 2 (TLR2) plays a major role in innate resistance in the lung against murine mycoplasma, PLoS One. 5 (2010) 10739. https://doi.org/10.1371/journal.pone.0010739.
[70] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, T.E. Ferrin, UCSF Chimera?A visualization system for exploratory research and analysis, J. Comput. Chem. 25 (2004) 1605–1612. https://doi.org/10.1002/jcc.20084.
[71] SwarmDock: a server for flexible protein–protein docking | Bioinformatics | Oxford Academic, (n.d.). https://academic.oup.com/bioinformatics/article/29/6/807/183866 (accessed August 4, 2020).
[72] M. Torchala, P.A. Bates, Predicting the structure of protein–protein complexes using the Swarmdock web server, Methods Mol. Biol. 1137 (2014) 181–197. https://doi.org/10.1007/978-1-4939-0366-5_13.
[73] R.A. Laskowski, M.B. Swindells, LigPlot+: Multiple ligand-protein interaction diagrams for drug discovery, J. Chem. Inf. Model. 51 (2011) 2778–2786. https://doi.org/10.1021/ci200227u.
[74] N. Rapin, O. Lund, M. Bernaschi, F. Castiglione, Computational immunology meets bioinformatics: The use of prediction tools for molecular binding in the simulation of the immune system, PLoS One. 5 (2010) 9862. https://doi.org/10.1371/journal.pone.0009862.
[75] P. Smialowski, A.J. Martin-Galiano, A. Mikolajka, T. Girschick, T.A. Holak, D. Frishman, Protein solubility: sequence based prediction and experimental verification, Bioinformatics. 23 (2007) 2536–2542. https://doi.org/10.1093/bioinformatics/btl623.
[76] J. Yang, W.C. Hooper, D.J. Phillips, D.F. Talkington, Cytokines in Mycoplasma pneumoniae infections, Cytokine Growth Factor Rev. 15 (2004) 157–168. https://doi.org/10.1016/j.cytogfr.2004.01.001.
[77] P.M. Moyle, I. Toth, Modern Subunit Vaccines: Development, Components, and Research Opportunities, ChemMedChem. 8 (2013) 360–376. https://doi.org/10.1002/cmdc.201200487.
[78] Y. Cao, Z. Lu, Y. Li, P. Sun, D. Li, P. Li, X. Bai, Y. Fu, H. Bao, C. Zhou, B. Xie, Y. Chen, Z. Liu, Poly(I: C) combined with multi-epitope protein vaccine completely protects against virulent foot-and-mouth disease virus challenge in pigs, Antiviral Res. 97 (2013) 145–153. https://doi.org/10.1016/j.antiviral.2012.11.009.
[79] B. Vakili, N. Nezafat, B. Zare, N. Erfani, M. Akbari, Y. Ghasemi, M.R. Rahbar, G.R. Hatam, A new multi-epitope peptide vaccine induces immune responses and protection against Leishmania infantum in BALB/c mice, Med. Microbiol. Immunol. 209 (2020) 69–79. https://doi.org/10.1007/s00430-019-00640-7.
[80] P.A. Unni, A.M.M.T. Ali, M. Rout, A. Thabitha, S. Vino, S.S. Lulu, Designing of an epitope-based peptide vaccine against walking pneumonia: an immunoinformatics approach, Mol. Biol. Rep. 46 (2019) 511–527. https://doi.org/10.1007/s11033-018-4505-0.
[81] M. Mahmood, A. Javaid, F. Shahid, U.A. Ashfaq, Rational design of multimeric based subunit vaccine against Mycoplasma pneumonia: Subtractive proteomics with immunoinformatics framework., Infect. Genet. Evol. (2021) 104795. https://doi.org/10.1016/j.meegid.2021.104795.
[82] M.M. McNeil, F. DeStefano, Vaccine-associated hypersensitivity, J. Allergy Clin. Immunol. 141 (2018) 463–472. https://doi.org/10.1016/j.jaci.2017.12.971.
[83] A. Sette, J. Fikes, Epitope-based vaccines: An update on epitope identification, vaccine design and delivery, Curr. Opin. Immunol. 15 (2003) 461–470. https://doi.org/10.1016/S0952-7915(03)00083-9.

[84] D. Mathur, S. Prakash, P. Anand, H. Kaur, P. Agrawal, A. Mehta, R. Kumar, S. Singh, G.P.S. Raghava, PEPlife: A Repository of the Half-life of Peptides, Sci. Rep. 6 (2016). https://doi.org/10.1038/srep36617.

[85] A. Six, B. Bellier, V. Thomas-Vaslin, D. Klatzmann, Systems biology in vaccine design, Microb. Biotechnol. 5 (2012) 295–304. https://doi.org/10.1111/j.1751-7915.2011.00321.x.

[86] J. Yang, W.C. Hooper, D.J. Phillips, D.F. Talkington, Regulation of proinflammatory cytokines in human lung epithelial cells infected with Mycoplasma pneumoniae, Infect. Immun. 70 (2002) 3649–3655. https://doi.org/10.1128/IAI.70.7.3649-3655.2002.

[87] A. Christodoulides, N. Gupta, V. Yacoubian, N. Maithel, J. Parker, T. Kelesidis, The Role of Lipoproteins in Mycoplasma-Mediated Immunomodulation, Front. Microbiol. 9 (2018) 1682. https://doi.org/10.3389/fmicb.2018.01682.

[88] T. Shimizu, Inflammation-inducing factors of Mycoplasma pneumoniae, Front. Microbiol. 7 (2016) 414. https://doi.org/10.3389/fmicb.2016.00414.

[89] P.M. Meyer Sauteur, A.C.J.M. De Bruijn, C. Graça, A.P. Tio-Gillen, S.C. Estevão, T. Hoogenboezem, R.W. Hendriks, C. Berger, B.C. Jacobs, A.M.C. Van Rossum, R. Huizinga, W.W.J. Unger, Antibodies to protein but not glycolipid structures are important for host defense against mycoplasma pneumoniae, Infect. Immun. 87 (2019). https://doi.org/10.1128/IAI.00663-18.

7. CONCLUSÃO CAPÍTULO II

Existem grandes desafios quando se trata da construção de vacinas multi-epítopos, principalmente a respeito da capacidade de induzir respostas imunes adequadas. Nesta prospecção, através de abordagens de imunoinformática, selecionamos de forma criteriosa, epítopos de alta cobertura populacional e alta afinidade com receptores do sistema imunológico para construir um imunógeno multi-epítopo contra o *M. pneumoniae*. A proteína construída juntamente com um adjuvante, foi submetida à diversas análises, demonstrando ser segura, estável e com alto potencial de induzir respostas imunes protetoras.

Vacinas contra o *M. pneumoniae* ainda não estão disponíveis no mercado e muito se deve as respostas exacerbadas causadas por vacinas de proteínas inteiras ou inativadas. Portanto, a estratégia aqui abordada demonstra ser promissora, mas testes *in vitro* e *in vivo*, o que chamamos de bancada, são necessários para validar a estabilidade, segurança e imunogenicidade do imunógeno descrito.

8. DISCUSSÃO

O *M.pneumoniae* é um patógeno único, suas características morfológicas conferem interações com o sistema imune hospedeiro capazes de induzir respostas inflamatórias potentes e ao mesmo tempo, esses microorganismos conseguem variar seus antígenos para promover o escape imunológico e proceder com a replicação e colonização do epitélio respiratório (CHAMBAUD; WRÓBLEWSKI; BLANCHARD, 1999). Essas propriedades corroboram com a alta incidência da espécie nos casos de PAC, o que se torna ainda mais agravante devido à resistência aos antimicrobianos. O repertório gênico restrito do *M.pneumoniae*, apoia as suas capacidades metabólicas limitadas, de forma que o cultivo em laboratório desse tipo de patógeno é lento, podendo levar cerca de 3 semanas, e requer meios de cultura específicos enriquecidos (ATKINSON; BALISH; WAITES, 2008). Esses fatos dificultam significantemente o desenvolvimento de vacinas e drogas específicas para esse patógeno e suportam a ideia de que estratégias *in silico* podem ser alternativas chave, de grande valia nesse processo. Diante da situação apresentada, propusemos construir um imunógeno de alta cobertura contra essa espécie e predizer potenciais alvos de drogas.

A vacinologia reversa, tem demonstrado ser uma ferramenta poderosa para identificar alvos antigênicos através do repertório genômico de um organismo (DEL TORDELLO; RAPPUOLI; DELANY, 2017). Por mais que as linhagens de *M.pneumoniae* apresentem um alto padrão de similaridade, variações existem e no contexto de desenvolver vacinas e buscar alvos drogáveis, a intenção é que ambas estratégias tenham capacidade de induzir respostas imunes ou proteger contra a maior cobertura de cepas possível. Dessa forma, foram identificados os genes do genoma central (core) oriundos de 88 genomas completos e incompletos de *M.pneumoniae* para iniciar nossa busca. A genômica subtrativa foi uma abordagem chave para o estudo como um todo, ao permitir identificar os genes centrais que não apresentam homologia com o hospedeiro humano. Esses genes apresentam características importantes no âmbito do desenvolvimento das vacinas e de drogas, são essenciais para a sobrevivência do patógeno, auxiliando nas capacidades adaptativas, sendo componentes fundamentais para o metabolismo e estabilidade físico-química do patógeno (BARH et al., 2011). Foram encontrados 5 candidatos à alvos de drogas com possíveis ligantes oriundos do banco de dados ZINC. Além disso, 8 proteínas de superfície e extracelulares foram identificadas como potenciais alvos vacinais. Dentre essas proteínas, 7 são lipoproteínas (ou putativas lipoproteínas), algumas ainda não caracterizadas e a última é a prolipopoproteína diacilgliceril transferase, com perfil enzimático. Essas proteínas foram selecionadas através da metodologia proposta, por estarem localizadas em regiões de alto contato com o sistema imune hospedeiro, por apresentarem propriedades imunogênicas. Ademais, 2 dessas proteínas foram encontradas em ilhas de patogenicidade putativas,

altamente conservadas entre os genomas estudados, demonstrando o potencial virulento dessas proteínas que está também relacionado com capacidade inflamatória.

As 8 proteínas encontras por meio da vacinologia reversa, mais as proteínas adesinas P1 e P30 foram então submetidas às análises de imunoinformática, afim de construir um imunógeno multi-epítopo contra o *M.pneumoniae.* Após a predição de epítopos, diversas filtragens foram realizadas para selecionar aqueles reconhecidos com alta afinidade por moléculas de MHC, codificadas por alelos frequentes na população mundial, e capacidade de induzir tanto respostas imunes humorais quanto celulares. Dessa forma, foram selecionados para compor à proteína quimérica, 29 epítopos pertencentes à 5 proteínas, 4 delas uma lipoproteína do tipo 3, uma lipoproteína da família P80, uma proteína da família DUF31, lipoproteína putativa que possivelmente tem função de peptidase, a prolipopoproteína diacilgliceril transferase e a proteína adesina P1. As lipoproteínas são componentes extremamente imunogênicos e já foram relacionadas com o potencial de induzir respostas imunes exacerbadas quando submetidas à testes como alvos vacinais, enquanto a remoção das porções lipídicas dessas proteínas evitavam essas reações (MARA et al., 2020). Dessa forma, a abordagem multi-epítopo, por utilizar somente determinantes imunogênicos específicos e em baixa quantidade, se comparado com a proteína inteira, tem menor chances de induzir respostas intensivas. Esse aspecto foi sustentado pelo resultado obtido pela simulação imune *in silico,* onde o potencial de indução de resposta imune avaliado demonstrou boas perspectivas e ao mesmo tempo citocinas de caráter anti-inflamatório como TGF-β e IL-10 sugeriram a capacidade de equilíbrio da resposta imunológica. De toda forma, para validar esses aspectos estudos *in vivo* precisão ser realizados. A enzima prolipopoproteína diacilgliceril transferase atua na acetilação de lipoproteínas, processo que interfere no reconhecimento por receptores do tipo Toll (ZUO; WU; YOU, 2009), os epítopos de MHC-II dela selecionados, apresentaram cobertura populacional variando entre 74.36% e 98.27%, com os melhores índices encontrados nestas análises. Já a proteína adesina P1, é uma das mais estudadas, já que é um potente fator de virulência de alta propriedade imunogênica e inclusive já foi alvo de estudos para construção de proteínas quiméricas para desenvolvimento de vacinas contra *M.pneumoniae* (UNNI et al., 2019). Esses achados, demonstram uma boa propensão do imunógeno construído em ser eficaz e induzir imunidade na população em geral e em proteger contra o *M.pneumoniae* sem induzir reações exacerbadas.

Dentre as vantagens da imunoinformática para desenvolvimento de vacinas baseadas em epítopos, está a diminuição nos riscos de efeitos adversos ou reversão para modo virulento que podem ocorrer em vacinas de organismo morto ou atenuado. Essas adversidades, condicionadas pela alta carga antigênica dessas vacinas, podem ser minimizadas ao usar vacinas de epítopos, que contém poucos motivos antigênicos, mas que

foram criteriosamente selecionados para induzir respostas imunes específicas contra o patógeno de interesse. No entanto, a quantidade reduzida de determinantes imunogênicos leva ao maior desafio acerca das vacinas multi-epítopos, a indução de respostas imunes eficazes e duradouras (LI et al., 2014). Para contornar esse viés, o uso de adjuvantes é indispensável nessa abordagem e uma ampla gama desses potencializadores imunológicos são utilizados em estudos similares e no nosso caso, o adjuvante oriundo de uma enterotoxina de *E. coli* foi utilizado visto sua ampla utilização e potencial. Vacinas multi-epítopo também se destacam quanto ao uso de recursos financeiros, com o desenvolvimento tecnológico, a síntese de proteínas tem se tornado mais acessível, além disso, por não necessitar de cultivo do patógeno, da vacina ser um produto proteico, o transporte e armazenamento pode se dar pela forma liofilizada, o que também reduz significativamente os custos gerais (PURCELL; MCCLUSKEY; ROSSJOHN, 2007). No geral, a proteína quimérica construída através dessa abordagem, demonstrou características físico-químicas e estruturais positivas, o que remete à uma boa perspectiva quanto a sua síntese.

A pneumonia é uma doença severa com impactos alarmantes em diversos aspectos. Além de contribuir enormemente nos índices de morbidade e mortalidade no mundo todo, muitas vezes a doença requer cuidados hospitalares de custos exorbitantes (ANDREWS, 2006). Apesar de ser um problema de saúde global, regiões com piores índices socioeconômicos são os mais afetados. Atualmente existem apenas duas vacinas disponíveis para prevenir a PAC, e mesmo assim, os programas de vacinação conseguiram até o momento disponibilizar a vacina a somente 50% da população mundial (“WHO | Pneumococcal conjugate 3rd dose (PCV3) immunization coverage”, 2020). Além disso, a ampla resistência à antimicrobianos de diversas classes é uma preocupação que mostra ainda mais a necessidade de medidas inovadoras para tentar diminuir as taxas epidemiológicas alarmantes acerca da doença (PEYRANI et al., 2019).

Dessa forma, a proposta apresentada demonstrou um desfecho favorável na construção in sílico de um imunógeno contra *M.pneumoniae*. Nossa estratégia foi buscar alvos presentes no core genoma, essenciais e presentes na superfície pois é considerado o método tradicional dentro dessa abordagem e corrobora com o nosso intuito de construir uma vacina de ampla cobertura. Entretanto, não descartamos a possibilidade de estudos futuros com os alvos excluídos através dessa metodologia para verificar a potencialidade deles como imunógenos já que existem estudos utilizando esses diferentes alvos.

9. CONCLUSÕES

Através de ferramentas de Vacinologia Reversa e Imunoinformática, construímos um imunógeno multi-epítopo de alta cobertura e com alto grau de afinidade por moléculas de reconhecimento imune frequentes na população, contra uma bactéria de distribuição mundial, o *M.pneumoniae*. Por meio das simulações *in silico*, interações com receptores do sistema imune e a indução de respostas foram avaliadas e revelaram achados promissores em relação a capacidade protetora sem sinais de exacerbação inflamatória. Além disso, as propriedades físico-químicas demonstraram estabilidade e qualidade estrutural da proteína quimérica formada. Potenciais alvos de drogas também foram preditos e a interação com os ligantes também gerou resultados satisfatórios.

Esses achados sugerem uma alternativa para produção de vacinas e drogas contra organismos de crescimento laborioso como *M.pneumoniae*, sem a necessidade do cultivo do patógeno, podendo ser aplicada contra outros patógenos emergentes. Contudo, os resultados encontrados por meio desta prospecção, necessitam de futuros ensaios *in vitro/in vivo* para validação experimental.

10. REFERÊNCIAS

AGARWALA, Richa et al. Database resources of the National Center for Biotechnology Information. Nucleic Acids Research, v. 46, n. D1, p. D8–D13, 1 jan. 2018. Disponível em: <www.ncbi.nlm.nih.gov/genbank/>.

ANDREWS, Kelly et al. TLR2 Regulates Neutrophil Recruitment and Cytokine Production with Minor Contributions from TLR9 during Hypersensitivity Pneumonitis. PLoS ONE, v. 8, n. 8, p. 73143, 30 ago. 2013. Disponível em: <www.plosone.org.

ANDREWS, R M. The National Hospital Bill: The Most Expensive Conditions by Payer, 2006: Statistical Brief #59. Healthcare Cost and Utilization Project (HCUP) Statistical Briefs, p. 1–12, 2006. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/21735581>.

ANGUS, Derek C. et al. Epidemiology of severe sepsis in the United States: Analysis of incidence, outcome, and associated costs of care. Critical Care Medicine, v. 29, n. 7, p. 1303–1310, 2001.

ARANCIBIA, Francisco et al. Antimicrobial treatment failures in patients with community-acquired pneumonia: Causes and prognostic implications. American Journal of Respiratory and Critical Care Medicine, v. 162, n. 1, p. 154–160, 2000. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/10903235/>.

ARAÚJO, Carlos Leonardo *et al.* Prediction of new vaccine targets in the core genome of Corynebacterium pseudotuberculosis through omics approaches and reverse vaccinology. *Gene*, v. 702, p. 36–45, 20 jun. 2019. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/30928361/>.

ATKINSON, Thomas Prescott; BALISH, Mitchell F.; WAITES, Ken B. Epidemiology, clinical manifestations, pathogenesis and laboratory detection of Mycoplasma pneumoniae infections. FEMS Microbiology Reviews. [S.l.]: Oxford Academic. Disponível em: <https://academic.oup.com/femsre/article/32/6/956/2683353>.

AZMI, Fazren et al. Recent progress in adjuvant discovery for peptide-based subunit vaccines. Human Vaccines and Immunotherapeutics. [S.l.]: Landes Bioscience. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/24300669/>, 2014.

BACKERT, Linus; KOHLBACHER, Oliver. Immunoinformatics and epitope prediction in the age of genomic medicine. Genome Medicine. [S.l.]: BioMed Central Ltd. Disponível em: http://genomemedicine.com/content/7/1/119>,20 nov. 2015.

BAH, Saikou Y. et al. Highlights on the Application of Genomics and Bioinformatics in the Fight Against Infectious Diseases: Challenges and Opportunities in Africa. Frontiers in Genetics, v. 9, p. 575, 27 nov. 2018. Disponível em: <www.frontiersin.org>.

BARH, Debmalya et al. In silico subtractive genomics for target identification in human bacterial pathogens. Drug Development Research, v. 72, n. 2, p. 162–177, 1 mar. 2011. Disponível em: <http://doi.wiley.com/10.1002/ddr.20413>.

BAJANTRI, Bharat; VENKATRAM, Sindhaghatta; DIAZ-FUENTES, Gilda. Mycoplasma pneumoniae : A Potentially Severe Infection . Journal of Clinical Medicine Research, v. 10, n. 7, p. 535–544, 2018. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29904437

BISCEVIC-TOKIC, Jasmina; TOKIC, Nedim; MUSANOVIC, Adnan. Pneumonia as the most common lower respiratory tract infection. Medical archives (Sarajevo, Bosnia and Herzegovina), v. 67, n. 6, p. 442–445, 1 dez. 2013. Disponível em: </pmc/articles/PMC4272454/>.Causas da pneumonia | CDC. Disponível em: <https://www.cdc.gov/pneumonia/causes.html>.

BLÖTZ, Cedric et al. Characterization of an Immunoglobulin Binding Protein (IbpM) From Mycoplasma pneumoniae. Frontiers in Microbiology, v. 11, p. 685, 16 abr. 2020. Disponível em: <www.frontiersin.org>.

CACCIOTTO, Carla et al. Mycoplasma lipoproteins are major determinants of neutrophil extracellular trap formation. Cellular Microbiology, v. 18, n. 12, 1 dez. 2016. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/27166588/>.

CAO, Bin *et al.* Overview of antimicrobial options for *M ycoplasma pneumoniae* pneumonia: focus on macrolide resistance. *The Clinical Respiratory Journal*, v. 11, n. 4, p. 419–429, 1 jul. 2017. Disponível em: <http://doi.wiley.com/10.1111/crj.12379>.

CHAMBAUD, Isabelle; WRÓBLEWSKI, Henri; BLANCHARD, Alain. Interactions between mycoplasma lipoproteins and the host immune system. Trends in Microbiology. [S.l.]: Trends Microbiol. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/10603485/>

CHASTRE, Jean; FAGON, Jean Yves. Ventilator-associated pneumonia. American Journal of Respiratory and Critical Care Medicine. [S.l.]: American Thoracic Society. Disponível em: http://www.atsjournals.org/doi/abs/10.1164/ajrccm.165.7.2105078,1 abr. 2002.

CHEN, Xiaoying; ZARO, Jennica L.; SHEN, Wei Chiang. Fusion protein linkers: Property, design and functionality. Advanced Drug Delivery Reviews. [S.l.]: NIH Public Access. Disponível em: </pmc/articles/PMC3726540/?report=abstract>.

CHOI, Sang Yong et al. Reactive oxygen species mediate Jak2/Stat3 activation and IL-8 expression in pulmonary epithelial cells stimulated with lipid-associated membrane proteins from Mycoplasma pneumonia. Inflammation Research, v. 61, n. 5, p. 493–501, 24 maio 2012. Disponível em: <https://link.springer.com/article/10.1007/s00011-012-0437-7>.

CLARK, Karen et al. GenBank. Nucleic Acids Research, v. 44, n. D1, p. D67–D72, 4 jan. 2016. Disponível em: <https://academic.oup.com/nar/article/44/D1/D67/2503088>.

DEL TORDELLO, E.; RAPPUOLI, R.; DELANY, I. Reverse Vaccinology: Exploiting Genomes for Vaccine Design. Human Vaccines: Emerging Technologies in Design and Development. [S.l.]: Elsevier Inc., 2017. p. 65–86.

DIBARDINO, David M.; WUNDERINK, Richard G. Aspiration pneumonia: A review of modern trends. Journal of Critical Care. [S.l.]: W.B. Saunders. , 1 fev. 2015

DORIGO-ZETSMA, J. W. et al. Mycoplasma pneumoniae P1 type 1- and type 2-specific sequences within the P1 cytadhesin gene of individual strains. Infection and Immunity, v. 69, n. 9, p. 5612–5618, 2001.

DRASBEK, M et al. Interaction between the P1 protein of Mycoplasma pneumoniae and receptors on HEp-2 cells. [S.d.]. Disponível em: <http://mic.sgmjournals.org>.

FILE, Thomas M. Community-acquired pneumonia. 13 dez. 2003, [S.l.]: Lancet Publishing Group, 13 dez. 2003. p. 1991–2001.

FOY, Hjordis M. et al. Viral and mycoplasmal pneumonia in a prepaid medical care group during an eight-year period. American Journal of Epidemiology, v. 97, n. 2, p. 93–102, 1973. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/4347511/>.

GARCIA, Josefina et al. A Mycoplasma fermentans-derived synthetic lipopeptide induces AP-1 and NF-κB activity and cytokine secretion in macrophages via the activation of mitogen-activated protein kinase pathways. Journal of Biological Chemistry, v. 273, n. 51, p. 34391–34398, 18 dez. 1998. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/9852105/>

General Recommendations on Immunization. Disponível em: <https://www.cdc.gov/mmwr/preview/mmwrhtml/rr6002a1.htm>.

GROSSHENNIG, Stephanie et al. Hydrogen sulfide is a novel potential virulence factor of Mycoplasma pneumoniae: Characterization of the unusual cysteine desulfurase/desulfhydrase HapE. Molecular Microbiology, v. 100, n. 1, p. 42–54, 1 abr. 2016. Disponível em: <https://onlinelibrary.wiley.com/doi/full/10.1111/mmi.13300>.

GROSSO, Alejandra; FAMIGLIETTI, Angela; LUNA, CarlosM. Community-acquired pneumonia due to gram-negative bacteria. Community Acquired Infection, v. 2, n. 4, p. 117, 2015. Disponível em: <http://www.caijournal.com/text.asp?2015/2/4/117/172651>.

HAMMERSCHLAG, M. R. Mycoplasma pneumoniae infections. Current Opinion in Infectious Diseases, v. 14, n. 2, p. 181–186, 2001. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/11979130/>.

HASSELBRING, Benjamin M.; KRAUSE, Duncan C. Proteins P24 and P41 function in the regulation of terminal-organelle development and gliding motility in Mycoplasma pneumoniae. Journal of Bacteriology, v. 189, n. 20, p. 7442–7449, out. 2007. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17693502/>.

HASSELBRING, Benjamin M.; SHEPPARD, Edward S.; KRAUSE, Duncan C. P65 truncation impacts P30 dynamics during mycoplasma pneumoniae gliding. Journal of Bacteriology, v.

194, n. 11, p. 3000–3007, jun. 2012. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/22544269/>.

HASTINGS, Deborah L et al. Mycoplasma pneumoniae outbreak in a long-term care facility--Nebraska, 2014. MMWR. Morbidity and mortality weekly report, v. 64, n. 11, p. 296–9, 27 mar. 2015. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/25811678>.

HE, Jun et al. Insights into the pathogenesis of Mycoplasma pneumoniae (Review). Molecular Medicine Reports. [S.l.]: Spandidos Publications. Disponível em: <http://www.spandidos-publications.com/10.3892/mmr.2016.5765/abstract>, 1 nov. 2016.

HIMMELREICH, Ralf et al. Complete sequence analysis of the genome of the bacterium Mycoplasma pneumoniae. Nucleic Acids Research, v. 24, n. 22, p. 4420–4449, 1996. Disponível em: <http://zmbh.uni-heidelberg.de/M_pneumoniae>.

HOEK, Kristen L. et al. A role for the Mycoplasma pneumoniae adhesin P1 in interleukin (IL)-4 synthesis and release from rodent mast cells. Microbial Pathogenesis, v. 39, n. 4, p. 149–158, out. 2005. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/16169702/>.

HOWE, Kevin et al. WormBase. Worm, v. 1, n. 1, p. 15–21, 20 jan. 2012. Disponível em: <http://www.tandfonline.com/doi/abs/10.4161/worm.19574>.

HSIEH, Yu-Chia et al. Life-threatening Pneumonia Caused by Macrolide-resistant Mycoplasma pneumoniae. Pediatric Infectious Disease Journal, v. 31, n. 2, p. 208–209, fev. 2012. Disponível em: <https://journals.lww.com/00006454-201202000-00028>.

HU, P. C. et al. Demonstration of antibodies to mycoplasma pneumoniae attachment protein in human sera and respiratory secretions. Infection and Immunity, v. 41, n. 1, p. 437–439, 1983. Disponível em: </pmc/articles/PMC264800/?report=abstract>.

HU, Weiguo et al. A flexible peptide linker enhances the immunoreactivity of two copies HBsAg preS1 (21-47) fusion protein. Journal of Biotechnology, v. 107, n. 1, p. 83–90, 8 jan. 2004

IKEBE, Mitsuo et al. A hinge at the central helix of the regulatory light chain of myosin is critical for phosphorylation-dependent regulation of smooth muscle myosin motor activity. Journal of Biological Chemistry, v. 273, n. 28, p. 17702–17707, 10 jul. 1998. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/9651368/>.

JAIN, Seema; SELF, Wesley H.; et al. Community-Acquired Pneumonia Requiring Hospitalization among U.S. Adults. New England Journal of Medicine, v. 373, n. 5, p. 415–427, 30 jul. 2015. Disponível em: <http://www.nejm.org/doi/10.1056/NEJMoa1500245>.

JAIN, Seema; WILLIAMS, Derek J.; et al. Community-Acquired Pneumonia Requiring Hospitalization among U.S. Children. New England Journal of Medicine, v. 372, n. 9, p. 835–845, fev. 2015.

JAISWA, Arun Kumar Jaiswa *et al.* Reverse vaccinology and subtractive genomics approaches for identifying common therapeutics against Mycobacterium leprae and Mycobacterium lepromatosis. *Journal of Venomous Animals and Toxins including Tropical*

Diseases, v. 27, 2021. Disponível em: <https://www.scielo.br/scielo.php?script=sci_arttext&pid=S1678-91992021000100308&lng=en&nrm=iso&tlng=en>.

JOHANSSON, Nielas et al. Etiology of Community-Acquired pneumonia: Increased microbiological yield with new diagnostic methods. Clinical Infectious Diseases, v. 50, n. 2, p. 202–209, 15 jan. 2010. Disponível em: https://academic.oup.com/cid/article-lookup/doi/10.1086/648678, 2021.

JONES, Simon A. Directing Transition from Innate to Acquired Immunity: Defining a Role for IL-6. The Journal of Immunology, v. 175, n. 6, p. 3463–3468, 15 set. 2005. Disponível em: <http://www.jimmunol.org/content/175/6/3463>.

KANAMPALLIWAR, Amol M. Reverse vaccinology and its applications. Methods in Molecular Biology. [S.l.]: Humana Press Inc., 2020. v. 2131. p. 1–16. Disponível em: <https://link.springer.com/protocol/10.1007/978-1-0716-0389-5_1>.

KANNAN, Thirumalai R. et al. Functional mapping of community-acquired respiratory distress syndrome (CARDS) toxin of Mycoplasma pneumoniae defines regions with ADP-ribosyltransferase, vacuolating and receptor-binding activities. Molecular Microbiology, v. 93, n. 3, p. 568–581, 2014. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/24948331/>.

KELLY, Dominic F.; RAPPUOLI, Rino. Reverse vaccinology and vaccines for serogroup B Neisseria meningitidis. Advances in Experimental Medicine and Biology. [S.l.]: Springer Science and Business Media Deutschland GmbH. Disponível em: <www.genomesonline.org>, 2005.

KOGA, Shunsuke et al. A case of meningoencephalitis associated with macrolide-resistant Mycoplasma pneumoniae infection. Pediatrics International, v. 54, n. 5, p. 724–726, out. 2012. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/23005909/>.

KOH, Y. Y. et al. Levels of interleukin-2, interferon-gamma, and interleukin-4 in bronchoalveolar lavage fluid from patients with Mycoplasma pneumonia: implication of tendency toward increased immunoglobulin E production. Pediatrics, v. 107, n. 3, 2001. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/11230620/>.

KOTSIAS, Fiorella; CEBRIAN, Ignacio; ALLOATTI, Andrés. Antigen processing and presentation. International Review of Cell and Molecular Biology. [S.l.]: Elsevier Inc., 2019. v. 348. p. 69–121.

KUMAR JAISWAL, Arun *et al.* An In Silico Identification of Common Putative Vaccine Candidates against Treponema pallidum: A Reverse Vaccinology and Subtractive Genomics Based Approach. *International journal of molecular sciences*, v. 18, n. 2, 14 fev. 2017. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/28216574>.

LANKS, Charles W.; HSIA, David W. Community-acquired Pneumonia and Hospital- acquired Pneumonia. Medical Clinics of North America, v. 103, n. 3, p. 487–501, 1 maio

2019. Disponível em:
<https://www.sciencedirect.com/science/article/abs/pii/S0025712518301731?via%3Dihub>.

LANKS, Charles W.; MUSANI, Ali I.; HSIA, David W. Community-acquired Pneumonia and Hospital-acquired Pneumonia. Medical Clinics of North America. [S.l.]: W.B. Saunders. Disponível em: <http://www.medical.theclinics.com/article/S0025712518301731/fulltext>, 1 maio 2019.

LEE, Joon Kee et al. Comparative genomics of Mycoplasma pneumoniae isolated from children with pneumonia: South Korea, 2010-2016. BMC Genomics, v. 20, n. 1, 29 nov. 2019. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/31783732/>.

LEI, Yao et al. Application of built-in adjuvants for epitope-based vaccines. PeerJ, v. 2019, n. 1, 2019. Disponível em: </pmc/articles/PMC6336016/>.

LI, Weidang et al. Peptide vaccine: Progress and challenges. Vaccines. [S.l.]: MDPI AG. Disponível em: </pmc/articles/PMC4494216/>, 2 jul. 2014

LIEBERMAN, David et al. Serological evidence of Mycoplasma pneumoniae infection in acute exacerbation of COPD. Diagnostic Microbiology and Infectious Disease, v. 44, n. 1, p. 1–6, set. 2002. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/12376023/>.

LYNCH, J. P. Hospital-acquired pneumonia: Risk factors, microbiology, and treatment. Chest. [S.l.]: American College of Chest Physicians. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/11171773/>, 2001.

M. WINCHELL, Jonas. Mycoplasma pneumoniae; A National Public Health Perspective. 2013.

MAHMOOD, Marvah et al. Rational design of multimeric based subunit vaccine against Mycoplasma pneumonia: Subtractive proteomics with immunoinformatics framework. Infection, genetics and evolution : journal of molecular epidemiology and evolutionary genetics in infectious diseases, p. 104795, 2 mar. 2021. Disponível em:
<http://www.ncbi.nlm.nih.gov/pubmed/33667723>.

MANDELL, Lionel A. et al. Infectious Diseases Society of America/American Thoracic Society Consensus Guidelines on the management of community-acquired pneumonia in adults. Clinical Infectious Diseases. [S.l.]: Oxford University Press. Disponível em:
</pmc/articles/PMC7107997/>, 1 mar. 2007.

MARA, Arlind B. et al. Lipid moieties of Mycoplasma pneumoniae lipoproteins are the causative factor of vaccine-enhanced disease. npj Vaccines, v. 5, n. 1, p. 1–5, 1 dez. 2020. Disponível em: <https://doi.org/10.1038/s41541-020-0181-x>.

MARTIN-LOECHES, Ignacio; RODRIGUEZ, Alejandro H.; TORRES, Antoni. New guidelines for hospital-acquired pneumonia/ventilator-associated pneumonia. Current Opinion in Critical Care, v. 24, n. 5, p. 347–352, out. 2018. Disponível em:
<http://journals.lww.com/00075198-201810000-00006>.

MENÉNDEZ, Rosario; PERPIÑÁ, Miguel; TORRES, Antoni. Evaluation of nonresolving and progressive pneumonia. Seminars in Respiratory Infections, v. 18, n. 2, p. 103–111, jun. 2003. Disponível em: <https://link.springer.com/chapter/10.1007/978-1-4757-5548-0_17>.

METLAY, Joshua P. et al. Diagnosis and treatment of adults with community-acquired pneumonia. American Journal of Respiratory and Critical Care Medicine, v. 200, n. 7, p. E45–E67, 1 out. 2019. Disponível em: <https://www.atsjournals.org/doi/10.1164/rccm.201908-1581ST>.

MICHAEL SCHELD, W.; MANDELL, Gerald L. Nosocomial pneumonia: Pathogenesis and recent advances in diagnosis and therapy. Reviews of Infectious Diseases, v. 13, p. S717–S732, 1991.

MILLIGAN, Rachael et al. Vaccines for preventing typhoid fever. Cochrane Database of Systematic Reviews. [S.l.]: John Wiley and Sons Ltd. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29851031/>.

MOROZUMI, Miyuki; TAKAHASHI, Takashi; UBUKATA, Kimiko. Macrolide-resistant Mycoplasma pneumoniae: Characteristics of isolates and clinical aspects of community-acquired pneumonia. Journal of Infection and Chemotherapy. [S.l.]: Japanese Society for Cancer Chemotherapy. Disponível em: <http://www.rna.ccbb.>,22 jan. 2010.

MOYLE, Peter; TOTH, Istvan. Self-Adjuvanting Lipopeptide Vaccines. Current Medicinal Chemistry, v. 15, n. 5, p. 506–516, 31 jan. 2008. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/18289006/>.

Mycoplasma pneumoniae | Home | CDC. Disponível em: <https://www.cdc.gov/pneumonia/atypical/mycoplasma/index.html>.

NEGAHDARIPOUR, Manica et al. Structural vaccinology considerations for in silico designing of a multi-epitope vaccine. Infection, Genetics and Evolution, v. 58, p. 96–109, 1 mar. 2018.

NONIKOV, V. E. Atypical Pneumonia. Antibiotiki i Khimioterapiya, v. 46, n. 6, p. 32–37, 2001. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/00325481.1993.11701702>.

OLI, Angus Nnamdi et al. <p>Immunoinformatics and Vaccine Development: An Overview</p>. ImmunoTargets and Therapy, v. Volume 9, p. 13–30, 26 fev. 2020. Disponível em: <https://www.dovepress.com/immunoinformatics-and-vaccine-development- an-overview-peer-reviewed-article-ITT>.

OPITZ, Oliver et al. Cytokine gene expression in immune mice reinfected with Mycoplasma pneumoniae: The role of T cell subsets in aggravating the inflammatory response. Immunobiology, v. 196, n. 5, p. 575–587, 1 mar. 1997. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/9145334/>.

PAHAL, Parul; SHARMA, Sandeep. Pneumonia, Typical Bacterial. [S.l.]: StatPearls Publishing, 2018. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/30485000>.

PARVIZPOUR, Sepideh et al. Epitope-based vaccine design: a comprehensive overview of bioinformatics approaches. Drug Discovery Today. [S.l.]: Elsevier Ltd. , 1 jun. 2020

PEYRANI, Paula et al. The burden of community-acquired bacterial pneumonia in the era of antibiotic resistance. Expert Review of Respiratory Medicine. [S.l.]: Taylor and Francis Ltd. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/17476348.2019.1562339>, 1 fev. 2019.

PICKETT, Brett E. et al. ViPR: An open bioinformatics database and analysis resource for virology research. Nucleic Acids Research, v. 40, n. D1, p. D593–D598, 1 jan. 2012. Disponível em: <www.pdb.org>.

POLLARD, Andrew J.; BIJKER, Else M. A guide to vaccinology: from basic principles to new developments. Nature Reviews Immunology. [S.l.]: Nature Research. Disponível em:<https://doi.org/10.1038/>, 1 fev. 2021.

PURCELL, Anthony W.; MCCLUSKEY, James; ROSSJOHN, Jamie. More than one reason to rethink the use of peptides in vaccine design. Nature Reviews Drug Discovery. [S.l.]: Nat Rev Drug Discov. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/17473845/>, maio 2007.

RALSTON, David|Cochran, Burt. A College Epidemic of Mycoplasma Pneumoniae. Journal of the American College Health Association, v. 27, n. 5, abr. 1979.

RAMIREZ, Julio A. et al. Adults Hospitalized with Pneumonia in the United States: Incidence, Epidemiology, and Mortality. Clinical Infectious Diseases, v. 65, n. 11, p. 1806–1812, 1 dez. 2017. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/29020164/>.

RAMYA, Chellammal M. Bacterial pneumonia. Research Journal of Pharmacy and Technology, v. 7, n. 8, p. 942–945, 2 jan. 2014. Disponível em:<https://www.ncbi.nlm.nih.gov/books/NBK513321/>.

RAPPUOLI, Rino. Reverse vaccinology, a genome-based approach to vaccine development. Vaccine, v. 19, n. 17–19, p. 2688–2691, 21 mar. 2001. Disponível em: <https://linkinghub.elsevier.com/retrieve/pii/S0264410X00005545>.

RAPPUOLI, Rino. Reverse vaccinology. Current Opinion in Microbiology, v. 3, n. 5, p. 445–450, 1 out. 2000. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1369527400001193>

RAUTANEN, Anna et al. Genome-wide association study of survival from sepsis due to pneumonia: An observational cohort study. The Lancet Respiratory Medicine, v. 3, n. 1, p. 53–60, 1 jan. 2015. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/25533491/>.

RAZIN, Shmuel. Mycoplasmas. Medical Microbiology. Baron S, editor 4th edition. Chapter 37. 1996. Available from: https://www.ncbi.nlm.nih.gov/books/NBK7637/

RAZIN, S; YOGEV, D; NAOT, Y. Molecular biology and pathogenicity of mycoplasmas. Microbiology and molecular biology reviews : MMBR, v. 62, n. 4, p. 1094–156, dez. 1998. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/9841667>.

ROGERS, M. J. et al. Construction of the mycoplasma evolutionary tree from 5S rRNA sequence data. Proceedings of the National Academy of Sciences of the United States of America, v. 82, n. 4, p. 1160–1164, 1 fev. 1985. Disponível em: <https://www.pnas.org/content/82/4/1160>.

RINAUDO, C. Daniela et al. Vaccinology in the genome era. Journal of Clinical Investigation. [S.l.]: American Society for Clinical Investigation. Disponível em: </pmc/articles/PMC2735939/>.

RUBIN, Lorry G. et al. 2013 IDSA clinical practice guideline for vaccination of the immunocompromised host. Clinical Infectious Diseases, v. 58, n. 3, fev. 2014.

SAID, Maria A. et al. Estimating the Burden of Pneumococcal Pneumonia among Adults: A Systematic Review and Meta-Analysis of Diagnostic Techniques. PLoS ONE, v. 8, n. 4, 2 abr. 2013. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/23565216/>.

SANCHES, Rodrigo C.O. *et al.* Immunoinformatics Design of Multi-Epitope Peptide-Based Vaccine Against Schistosoma mansoni Using Transmembrane Proteins as a Target. *Frontiers in Immunology*, v. 12, 2 mar. 2021. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/33737928/>.

SANTOS, Anderson R. *et al.* Mature Epitope Density - A strategy for target selection based on immunoinformatics and exported prokaryotic proteins. *BMC Genomics*, v. 14, n. Suppl 6, p. S4, 25 out. 2013. Disponível em: <http://bmcgenomics.biomedcentral.com/articles/10.1186/1471-2164-14-S6-S4>.

SAROM, Alissa De et al. Putative vaccine candidates and drug targets identified by reverse vaccinology and subtractive genomics approaches to control Haemophilus ducreyi, the causative agent of chancroid. Journal of the Royal Society Interface, v. 15, n. 142, 1 maio 2018. Disponível em: <http://dx.doi.org/10.1098/rsif.2018.0032>.

SEIB, K. L.; ZHAO, X.; RAPPUOLI, R. Developing vaccines in the era of genomics: A decade of reverse vaccinology. Clinical Microbiology and Infection. [S.l.]: Blackwell Publishing Ltd. , 2012

SELF, Wesley H. et al. Procalcitonin as a Marker of Etiology in Adults Hospitalized With Community-Acquired Pneumonia. Clinical Infectious Diseases, v. 65, n. 2, p. 183–190, 15 jul. 2017. Disponível em: <https://academic.oup.com/cid/article/65/2/183/3605415>.

SHIMIZU, Takashi et al. Cytadherence of mycoplasma pneumoniae induces inflammatory responses through autophagy and toll-like receptor 4. Infection and Immunity, v. 82, n. 7, p. 3076–3086, 2014. Disponível em: </pmc/articles/PMC4097642/>.

SHIMIZU, Takashi. Inflammation-inducing factors of Mycoplasma pneumoniae. Frontiers in Microbiology. [S.l.]: Frontiers Media S.A. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/27065977/>, 31 mar. 2016.

STELMACH, Iwona et al. Humoral and cellular immunity in children with Mycoplasma pneumoniae infection: A 1-year prospective study. Clinical and Diagnostic Laboratory Immunology, v. 12, n. 10, p. 1246–1250, out. 2005. Disponível em: </pmc/articles/PMC1247828/>.

SULTANA, Marufa et al. Economic evaluation of community acquired pneumonia management strategies: A systematic review of literature. PLoS ONE, v. 14, n. 10, 2019.

The top 10 causes of death. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/the-top-10-causes-of-death>.

SUGIYAMA, M. et al. Activation of inflammasomes in dendritic cells and macrophages by Mycoplasma salivarium. Molecular Oral Microbiology, v. 31, n. 3, p. 259–269, 2016. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/26177301/>.

TAKEUCHI, Osamu et al. Cutting Edge: Role of Toll-Like Receptor 1 in Mediating Immune Response to Microbial Lipoproteins. The Journal of Immunology, v. 169, n. 1, p. 10–14, 1 jul. 2002. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/12077222/>.

TOMAR, Namrata; DE, Rajat K. Immunoinformatics: An integrated scenario. Immunology. [S.l.]: Wiley-Blackwell. Disponível em: </pmc/articles/PMC2967261/>, out. 2010.

TURVEY, Stuart E.; BROIDE, David H. Innate immunity. Journal of Allergy and Clinical Immunology, v. 125, n. 2 SUPPL. 2, p. S24, fev. 2010. Disponível em:</pmc/articles/PMC2832725/>.

TOSTA, Stephane Fraga de Oliveira *et al.* Multi-epitope based vaccine against yellow fever virus applying immunoinformatics approaches. *Journal of Biomolecular Structure and Dynamics*, v. 39, n. 1, p. 219–235, 2021. Disponível em: <https://www.tandfonline.com/doi/abs/10.1080/07391102.2019.1707120>.

UNNI, P. Ambili et al. Designing of an epitope-based peptide vaccine against walking pneumonia: an immunoinformatics approach. Molecular Biology Reports, v. 46, n. 1, p. 511–527, 1 fev. 2019. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/30465133/>.

VAN WERKHOVEN, Cornelis H.; HUIJTS, Susanne M. Vaccines to Prevent Pneumococcal Community-Acquired Pneumonia. Clinics in Chest Medicine. [S.l.]: W.B. Saunders. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/30390745/>, 1 dez. 2018.

VILELA RODRIGUES, Thaís Cristina 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, p. 190907, 31 jul. 2019. Disponível em: <https://royalsocietypublishing.org/doi/10.1098/rsos.190907>.

WAITES, Ken B. et al. Mycoplasma pneumoniae from the respiratory tract and beyond. Clinical Microbiology Reviews. [S.l.]: American Society for Microbiology. Disponível em: </pmc/articles/PMC5475226/>, 1 jul. 2017.

WAITES, Ken B.; TALKINGTON, Deborah F. Mycoplasma pneumoniae and its role as a human pathogen. Clinical Microbiology Reviews. [S.l.]: Clin Microbiol Rev. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/15489344/>, out. 2004.

WALLER, Jessica L. et al. Detection and characterization of Mycoplasma pneumoniae during an outbreak of respiratory illness at a university. Journal of Clinical Microbiology, v. 52, n. 3, p. 849–853, 1 mar. 2014. Disponível em: <http://jcm.asm.org/>.

WARRENFELTZ, Susanne et al. EuPathDB: The eukaryotic pathogen genomics database resource. Methods in Molecular Biology. [S.l.]: Humana Press Inc., 2018. v. 1757. p. 69–113.

WIEMKEN, Timothy L. *et al.* Socioeconomic Position and the Incidence, Severity, and Clinical Outcomes of Hospitalized Patients With Community-Acquired Pneumonia. *Public Health Reports*, v. 135, n. 3, p. 364–371, 1 maio 2020. Disponível em: <http://journals.sagepub.com/doi/10.1177/0033354920912717>.

WHO | Pneumococcal conjugate 3rd dose (PCV3) immunization coverage. WHO, 2020. Disponível em: <http://www.who.int/gho/immunization/pneumococcal/en/>.

WINCHELL, Jonas. Mycoplasma pneumoniae – A National Public Health Perspective. Current Pediatric Reviews, v. 9, n. 4, p. 324–333, 31 dez. 2013. Disponível em: <http://www.eurekaselect.com/openurl/content.php?genre=article&issn=1573-3963&volume=9&issue=4&spage=324>.

XIAO, Li et al. Comparative genome analysis of Mycoplasma pneumoniae. BMC Genomics, v. 16, n. 1, ago. 2015a.

XIAO, Li et al. Comparative genome analysis of Mycoplasma pneumoniae. BMC Genomics, v. 16, n. 1, 16 ago. 2015b. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/26275904/>.

XIROUCHAKI, Nektaria et al. Lung ultrasound in critically ill patients: Comparison with bedside chest radiography. Intensive Care Medicine, v. 37, n. 9, p. 1488–1493, set. 2011. Disponível em: <https://pubmed.ncbi.nlm.nih.gov/21809107/>.

YAMAMOTO, Takeshi et al. Mpn491, a secreted nuclease of Mycoplasma pneumoniae, plays a critical role in evading killing by neutrophil extracellular traps. Cellular Microbiology, v. 19, n. 3, p. e12666, 1 mar. 2017. Disponível em: <https://doi.org/10.1111/cmi.12666>.

ZUO, Ling Ling; WU, Yi Mou; YOU, Xiao Xing. Mycoplasma lipoproteins and Toll-like receptors. Journal of Zhejiang University: Science B. [S.l.]: Zhejiang University Press. Disponível em: </pmc/articles/PMC2613965/>, 2009

11. ANEXOS

11. 1 MATERIAL SUPLEMENTAR REFERENTE AO CAPÍTULO I:

MHOLline Supplementary Table 1: Template ID used by MHOLline for Identified targets for modeling with identity.

Protein Name	Template ID	Idenity (%)
WP_010874513.1	1pa4	>=75%
WP_010874670.1	1n0e	>=75%
WP_010874705.1	1t71	>=75%
WP_010874779.1	2i15	>=75%
WP_014325598.1	1zxj	>=75%

MHOLline Supplementary Table 2: Classification of each sequence in accordance with criteria in table 4.

Sequence ID	Sequence name	MHOLline group name (only G2)	Filters Quality
6	WP_010874513.1	G2	Very High
12	WP_010874670.1	G2	Very High
13	WP_010874705.1	G2	Very High
15	WP_010874779.1	G2	Very High
40	WP_014325598.1	G2	Very High
36	WP_014325566.1	G2	High
1	WP_010874387.1	G2	Good
2	WP_010874410.1	G2	Medium to Good
8	WP_010874577.1	G2	Medium to Good
25	WP_010875025.1	G2	Medium to Good
29	WP_014325292.1	G2	Medium to Good
37	WP_014325567.1	G2	Medium to Good
42	WP_014325643.1	G2	Medium to Good
43	WP_014325646.1	G2	Medium to Good
44	WP_014325666.1	G2	Medium to Good
46	WP_019830475.1	G2	Medium to Low
47	WP_019830488.1	G2	Medium to Low
5	WP_010874483.1	G2	Medium to Low
20	WP_010874917.1	G2	Medium to Low
27	WP_010875045.1	G2	Medium to Low
32	WP_014325427.1	G2	Medium to Low
35	WP_014325526.1	G2	Medium to Low
39	WP_014325596.1	G2	Medium to Low
17	WP_010874873.1	G2	Low
48	WP_019830439.1	G2	Low
45	WP_014325669.1	G2	Very Low

MHOLline Supplementary Table 3: MHOLline criteria from G2Group

Quality	Identity	Lenght Variation Index
Very High	>=75%	<= 0.1
High	>=50% and <75%	<= 0.1
Good	>=50%	> 0.1 and <= 0.3
Medium to Good	>=35% and <50%	<= 0.3
Medium to Low	>=25% and 35%	<= 0.3
Low	>=25%	> 0.3 and <= 0.5
Very Low	>=25%	> 0.5 and <= 0.7

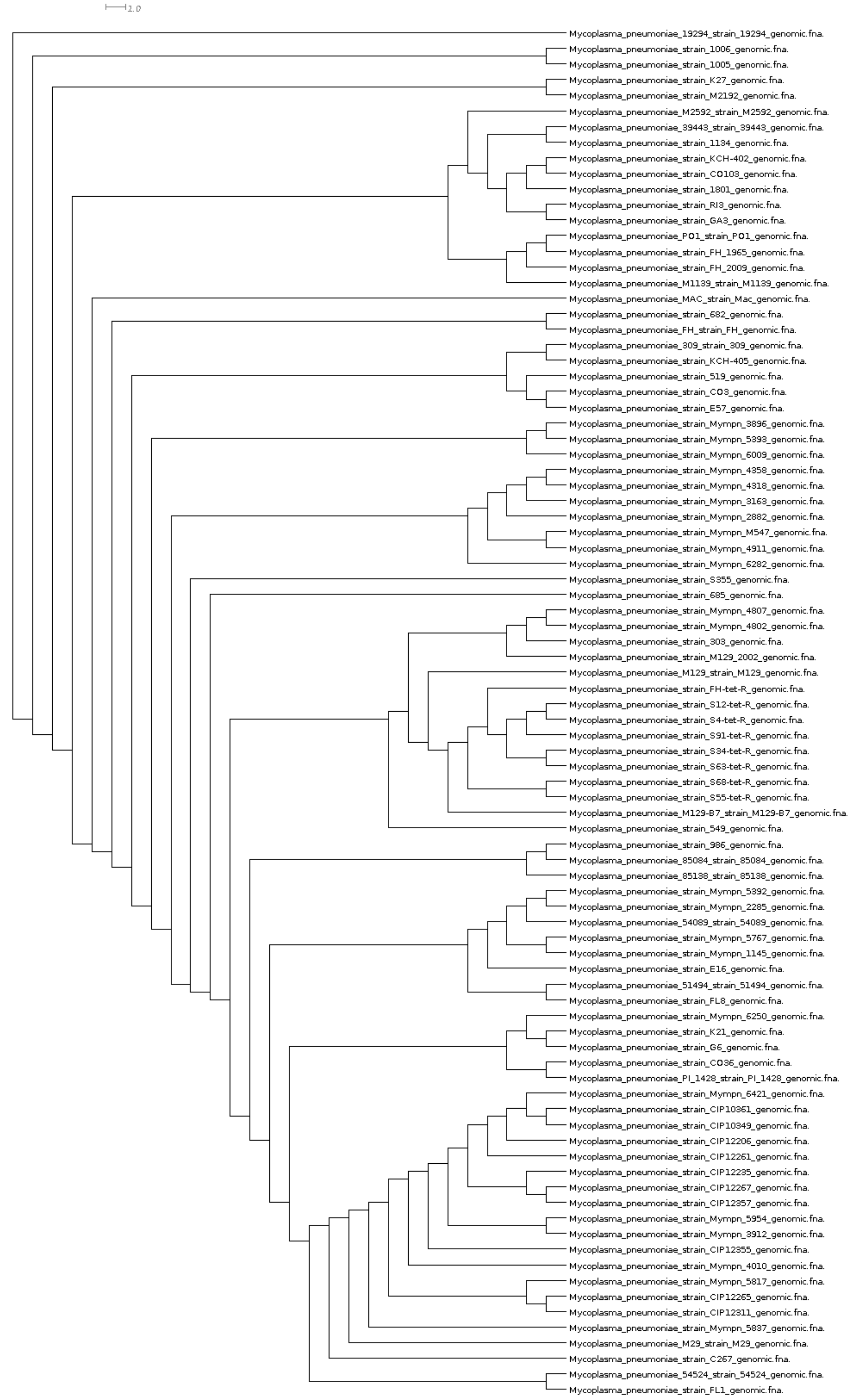
Supplementary table 4: Top 10 Molecules from Virtual Screening for each Identified Targets.

ZINC Compound ID		Autodock vina Binding Affinity	No of H-bond/Residues
	30S ribosome-binding factor (WP_010874513.1)		
ZINC04259381		-10.5	3/ASN18, ARG15
ZINC04259578		-8.5	3/ALA78
ZINC04259719		-6.0	2/ARG75
ZINC04235426		-9.7	2/HIS37, LEU40
ZINC05396219		-10.5	1/ASN11
ZINC20503175		-7.2	1/ASN27
ZINC04259499		-9.1	1/THR32
ZINC08300280		-6.6	1/THR29
ZINC31154666		-8.5	2/ARG75
	division/cell wall cluster transcriptional repressor MraZ (WP_010874670.1)		
ZINC04237100		-9.4	1/GLN54
ZINC08635277		-8.5	1/ASP64
ZINC04237105		-10.0	1/ARG43

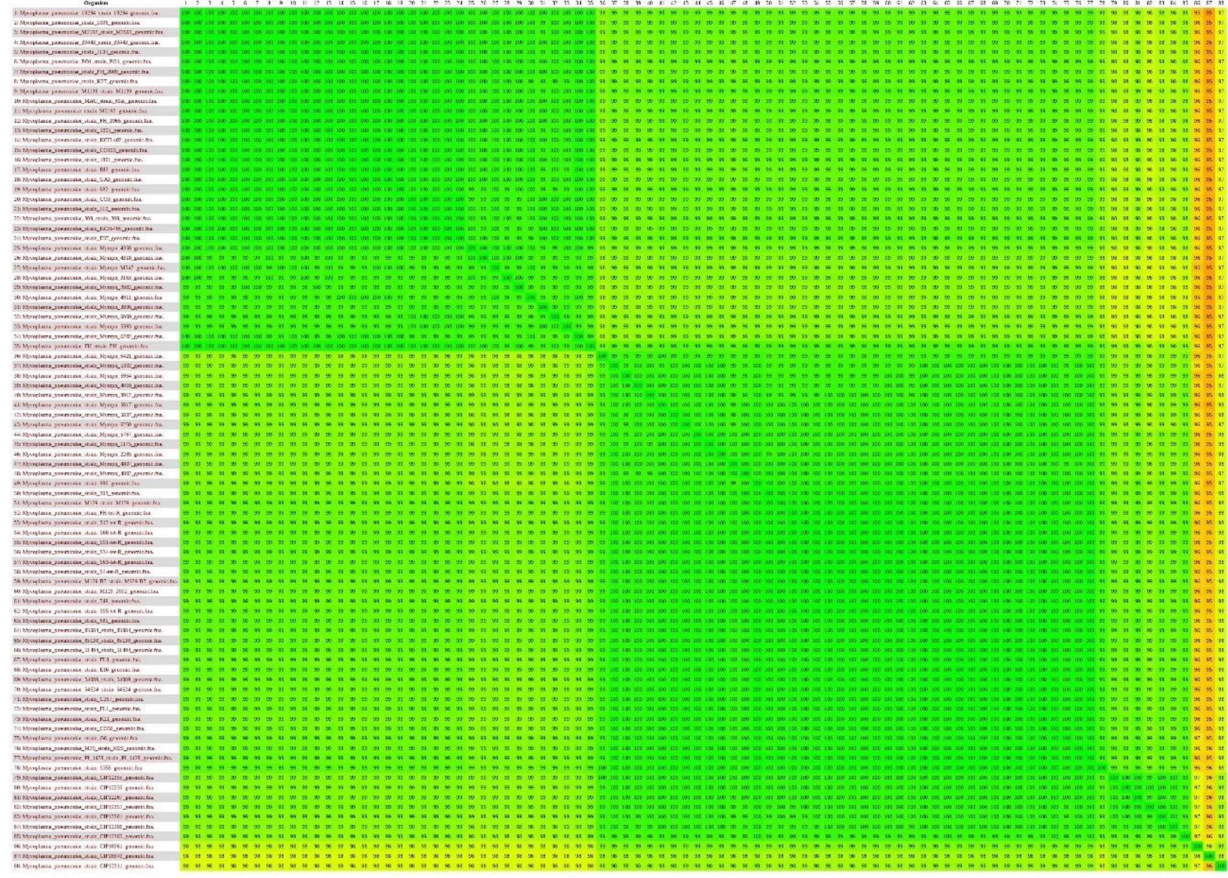
ZINC04258871		-10.4	2/ARG43, GLN54
ZINC04235924		-10.2	1/ARG34
ZINC04237101		-10.7	
ZINC04237091		-9.0	1/ASN38
ZINC04260408		-9.5	1/ARG43
1ZINC04222214		-10.6	1/CYS39
ZINC04270981		-10.7	2/ARG34, GLU41
	dTIGR00282 family metallophosphoesterase (WP_010874705.1)		
ZINC05396856		-8.5	1/LYS49
ZINC04259703		-8.9	3/LYS49, ASN71
ZINC04235924		-9.5	1/ASN71
ZINC08300419		-8.7	1/LYS139
ZINC04237082		-8.5	1/ASN71
ZINC04259578		-9.7	1/ASN71
ZINC20503625		-8.2	1/ASN71
ZINC04270628		-9.5	3/LYS196, GLY174, THR177
ZINC04259588		-9.6	1/ASN71
ZINC04277685		-9.6	1/ARG264
	Hypothetical protein (WP_010874779.1)		
ZINC08300419		-11.9	1/CYS90
ZINC05415832		-11.1	1/PHE93
ZINC05415069		-12.3	1/LYS108
ZINC04236421		-10.5	1/SER51
ZINC04235924		-11.0	1/CYS89
ZINC15709489		-11.1	1/CYS89
ZINC04237100		-11.5	1/TYR115
ZINC08635277		-11.4	1/TYR115
ZINC20503551		-11.1	1/TYR115
ZINC20503308		-9.9	1/CYS89
	Hypothetical protein (WP_014325598.1)		
ZINC04235924		-11.8	1/TYR143

ZINC04235884		-11.6	1/SER43
ZINC15672005		-10.2	2/GLU49, LYS45
ZINC04235972		-11.5	1/GLU145
ZINC04258871		-9.5	1/HIS36
ZINC04235928		-10.6	1/SER43
ZINC04236030		-10.3	2/LYS45, TYR154
ZINC04235880		-10.1	1/ARG39
ZINC04235930		-10.8	2/SER43, ASN148
ZINC03839937		-9.4	1/LYS62

Supplementary figure 1: Phylogenetic tree correlating the evolutive distance among the strains of *M. pneumonia*



Supplementary figure 2: Heatmap of similarity between the genomes of *M. pneumoniae*



11.2 MATERIAL SUPLEMENTAR REFERENTE AO CAPÍTULO II:

Table 1: **List of selected *Mycoplasma Pneumoniae* proteins**

List of selected proteins from the literature or reverse vaccinology, used in this work for the immunoinformatics analysis, and corresponding ID of the proteins.

ID	Name
WP_010874999.1	Specific mycoplasma lipoprotein type 3
WP_014574866.1	Hypothetical lipoprotein
WP_010874581.1	Prolipoprotein diacylglyceryl transferase
WP_010874862.1	P80 family lipoprotein
WP_014325486.1	P80 family lipoprotein
WP_014325517.1	P80 family lipoprotein
WP_014325659.1	Lipoprotein
WP_014325660.1	MPN647 family lipoprotein
WP_010874498.1	Adhesin P1 (M129 type)
WP_010874809.1	Adhesin P30
WP_053344075.1	Adhesin P1 (FH type)

Table 2: Predicted IFN-gamma inducing epitopes.

57 MHC-II binding epitopes capable to induce IFN-gamma (positive) were predicted using IFNepitope server hybrid method (MERCI and SVM) default parameters.

Epitope sequence	Method	Result	Score
GPGPGGYFWTYFFLY	MERCI	POSITIVE	2
PGPGGYFWTYFFLYG	MERCI	POSITIVE	2
GPGGYFWTYFFLYGW	MERCI	POSITIVE	2
PGGYFWTYFFLYGWY	MERCI	POSITIVE	2
GGYFWTYFFLYGWYK	MERCI	POSITIVE	2
GYFWTYFFLYGWYKV	MERCI	POSITIVE	2
YFWTYFFLYGWYKVG	MERCI	POSITIVE	2
TYFFLAAYWTYFFLY	MERCI	POSITIVE	2
YFFLAAYWTYFFLYG	MERCI	POSITIVE	2

FFLAAYWTYFFLYGW	MERCI	POSITIVE	2
FLAAYWTYFFLYGWA	MERCI	POSITIVE	2
LAAYWTYFFLYGWAA	MERCI	POSITIVE	2
AAYWTYFFLYGWAAY	MERCI	POSITIVE	2
AYWTYFFLYGWAAYV	MERCI	POSITIVE	2
AGVSKTFKDKCASTT	MERCI	POSITIVE	1
IPGGQYYPENYLSNE	MERCI	POSITIVE	1
PGGQYYPENYLSNEM	MERCI	POSITIVE	1
GGQYYPENYLSNEMR	MERCI	POSITIVE	1
GQYYPENYLSNEMRK	MERCI	POSITIVE	1
QYYPENYLSNEMRKI	MERCI	POSITIVE	1
YYPENYLSNEMRKIA	MERCI	POSITIVE	1
YPENYLSNEMRKIAM	MERCI	POSITIVE	1
PENYLSNEMRKIAMA	MERCI	POSITIVE	1
ENYLSNEMRKIAMAA	MERCI	POSITIVE	1
NYLSNEMRKIAMAAV	MERCI	POSITIVE	1
IVRGPGPGGYFWTYF	MERCI	POSITIVE	1
VRGPGPGGYFWTYFF	MERCI	POSITIVE	1
RGPGPGGYFWTYFFL	MERCI	POSITIVE	1
FWTYFFLYGWYKVGP	MERCI	POSITIVE	1
WTYFFLYGWYKVGPG	MERCI	POSITIVE	1
TYFFLYGWYKVGPGP	MERCI	POSITIVE	1
YFFLYGWYKVGPGPG	MERCI	POSITIVE	1
TDFTINRIAAYVFEN	MERCI	POSITIVE	1
DFTINRIAAYVFENW	MERCI	POSITIVE	1
FTINRIAAYVFENWN	MERCI	POSITIVE	1
TINRIAAYVFENWNE	MERCI	POSITIVE	1
INRIAAYVFENWNEL	MERCI	POSITIVE	1
NRIAAYVFENWNELL	MERCI	POSITIVE	1
RIAAYVFENWNELLA	MERCI	POSITIVE	1
IAAYVFENWNELLAA	MERCI	POSITIVE	1

AAYVFENWNELLAAY	MERCI	POSITIVE	1
FSYFLAAYGYFWTYF	MERCI	POSITIVE	1
SYFLAAYGYFWTYFF	MERCI	POSITIVE	1
YFLAAYGYFWTYFFL	MERCI	POSITIVE	1
FLAAYGYFWTYFFLA	MERCI	POSITIVE	1
LAAYGYFWTYFFLAA	MERCI	POSITIVE	1
AAYGYFWTYFFLAAY	MERCI	POSITIVE	1
AYGYFWTYFFLAAYW	MERCI	POSITIVE	1
YGYFWTYFFLAAYWT	MERCI	POSITIVE	1
GYFWTYFFLAAYWTY	MERCI	POSITIVE	1
YFWTYFFLAAYWTYF	MERCI	POSITIVE	1
FWTYFFLAAYWTYFF	MERCI	POSITIVE	1
WTYFFLAAYWTYFFL	MERCI	POSITIVE	1
YWTYFFLYGWAAYVA	MERCI	POSITIVE	1
WTYFFLYGWAAYVAA	MERCI	POSITIVE	1
TYFFLYGWAAYVAAF	MERCI	POSITIVE	1
YFFLYGWAAYVAAFF	MERCI	POSITIVE	1

Table 3: Predicted B cell conformational epitopes.

The EliPro server was used to predict the conformational B cell binding epitopes via the 3D structure of the *Mycoplasma pneumoniae* multi-epitope vaccine.

Conformational epitope	Residues and position	Number of residues	Score
1	_:Y524, _:G525, _:W526, _:A527, _:A528, _:Y529, _:V530, _:A531, _:A532, _:F533, _:F534, _:T535, _:T536, _:L537, _:F538, _:A539, _:A540, _:Y541, _:D542, _:Q543, _:I544, _:D545, _:F546, _:N547, _:R548, _:L549, _:F550, _:A551, _:A552, _:Y553, _:F554, _:N555, _:R556, _:L557, _:T559, _:H560, _:P561, _:V562, _:A563, _:A564, _:Y565, _:Q566, _:Y567	43	0.947

2	_:I568, _:P569, _:L570, _:F571, _:I572, _:D573, _:I574	7	0.931
3	_:P344, _:G345, _:P346, _:G347, _:D348, _:Q349, _:I350, _:D351, _:F352, _:N353, _:R354, _:L355, _:F356, _:T357, _:H358, _:P359, _:V360, _:T361, _:D362, _:G363, _:P364, _:G365, _:P366, _:G367, _:T368, _:M369, _:L370, _:V371, _:Y372, _:D373, _:Q374, _:Y375, _:I376, _:P377, _:L378, _:F379, _:I380, _:D381, _:I382, _:A383, _:A384, _:Y385, _:Q386, _:T387, _:D388, _:F389, _:T390, _:I391, _:N392, _:R393, _:I394, _:A395, _:A396, _:Y397, _:V398, _:F399, _:E400, _:N401, _:W402, _:N403, _:E404, _:L405, _:L406, _:A407, _:A408, _:Y409, _:S410, _:S411, _:N412, _:G413, _:L414, _:Y415, _:I416, _:N417, _:K418, _:A419, _:Y421, _:I422, _:A423, _:L425, _:N426, _:T427, _:Y429, _:F513, _:L514, _:A515, _:A516, _:Y517, _:W518, _:T519, _:Y520, _:F522, _:L523	93	0.718
4	_:M1, _:N2, _:F3, _:F29, _:K32, _:C33, _:A34, _:S35, _:T36, _:T37, _:A38, _:K39, _:V44, _:Q45, _:L46, _:V47, _:K48, _:L49, _:A50, _:S51, _:D52, _:T53, _:N54, _:K55, _:D56, _:S57, _:K58, _:G59, _:I60, _:Y61, _:I62, _:T63, _:D64, _:S65, _:T66, _:G67, _:K68, _:T69, _:R70, _:F71, _:I72, _:P73, _:G74, _:G75, _:Q76, _:Y77, _:E80, _:N81, _:Y82, _:L83, _:S84, _:N85, _:E86, _:M87, _:R88, _:K89, _:I90, _:A91, _:M92, _:A93, _:A94, _:V95, _:L96, _:S97, _:N98, _:V99, _:R100, _:V101, _:I103, _:C104, _:A105, _:S106, _:E107, _:A108, _:Y109, _:T110, _:P111, _:N112, _:H113, _:V114, _:W115, _:A116, _:I117, _:L119, _:A120, _:A121, _:E122, _:K249, _:N251, _:T252, _:S253, _:I254	92	0.707

Figure 1: The representation of secondary structure of the multi-epitope vaccine.

The result shows arrangement of alpha helices (46.0%), β-strand (9.0%), and coil formation (43.0%).

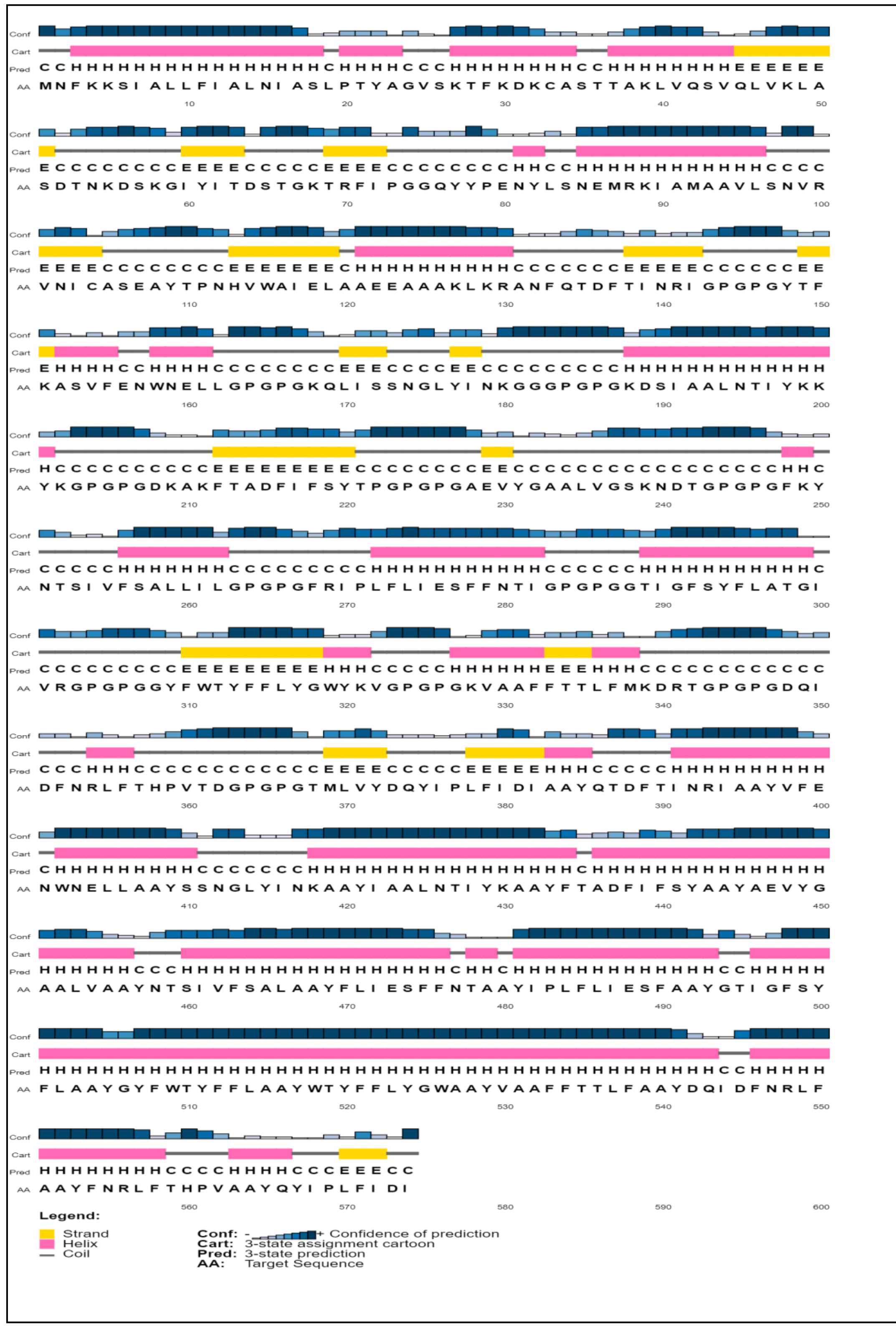


Figure 2. Scheme of using the script to search for overlap between epitopes.

The figure represents with lines and boxes the method for seeking overlap between epitopes. Below each circle there is a box with the method used to predict epitopes. Leaving the CTL and HTL boxes, the lines indicate a new box containing the confidence epitopes, which were predicted by the two methods indicated. The next lines indicate the search for overlap between CTL and B and HLT and B epitopes.

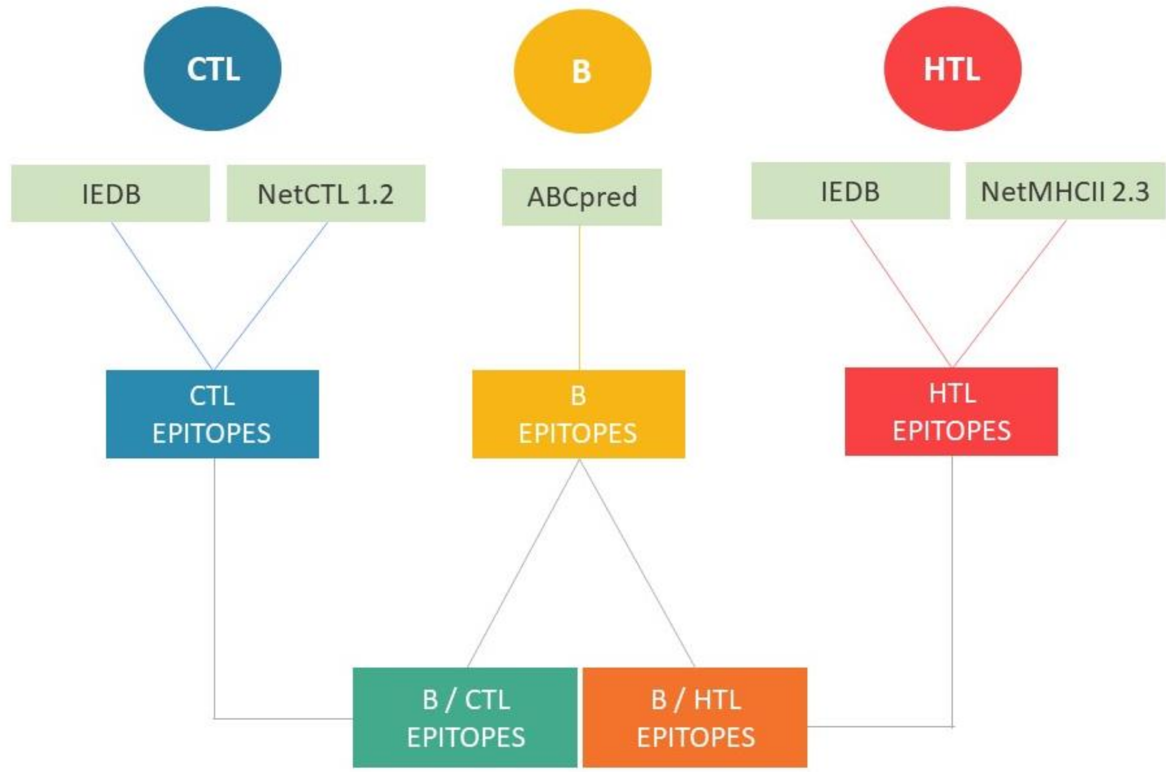


Figure 3: Immuno simulation results regarding innate immunity and cytokine production.

Those graphs represent the immune simulation with only the epitope sequence as injections. (A) natural-killer cells population fluctuation during the injections, (B) macrophages population, (C) dendritic cells population, (D) epithelial cells population, (E) cytokine levels induced during the injections and the D value, which is related to diversity of T-cells clones.

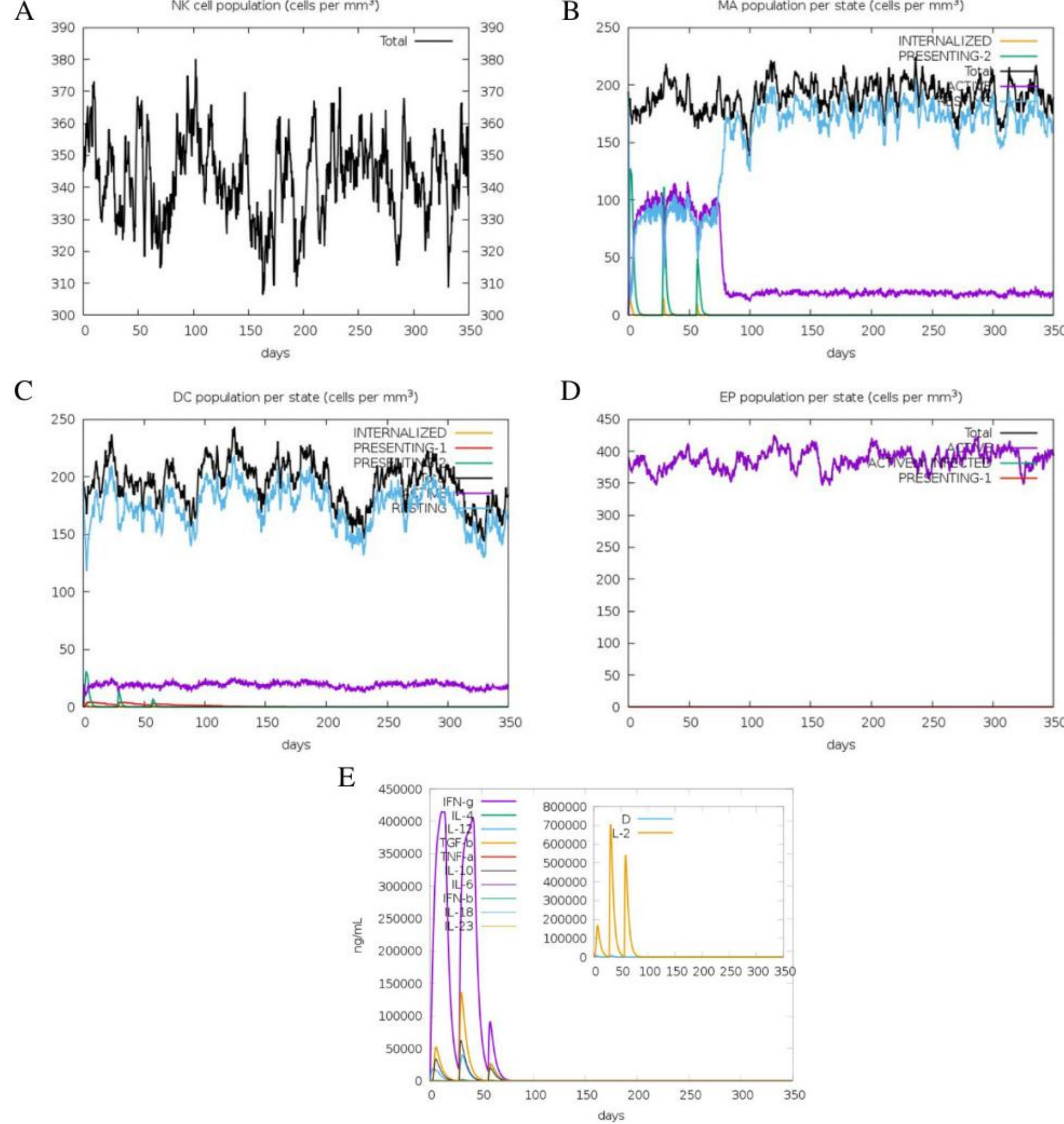


Figure 4: The Immuno simulation results regarding T lymphocytes.

(A) Memory and not memory cytotoxic T-cells graph, (B) Cytotoxic T-cell populations per state after the injections, purple line represents the active cells and the blue line represents resting cells, (C) Represents the helper T-cell population per state, (D) effector helper T-cell types fluctuation during the injections.

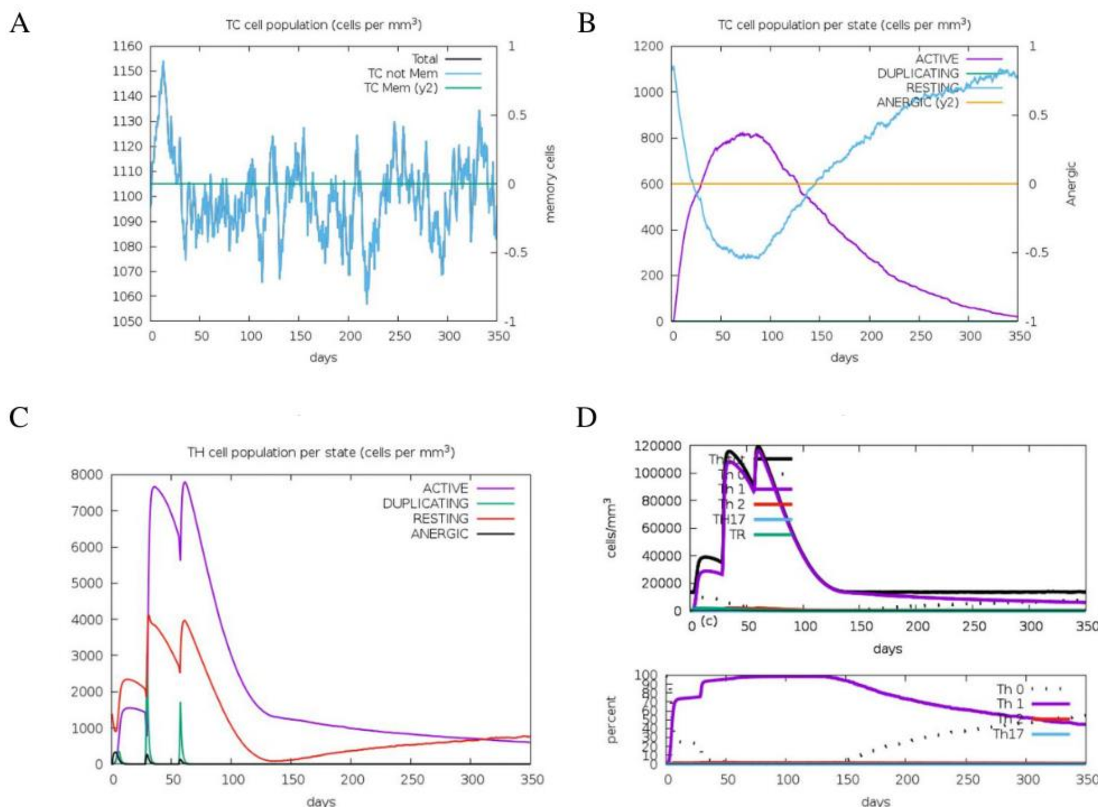
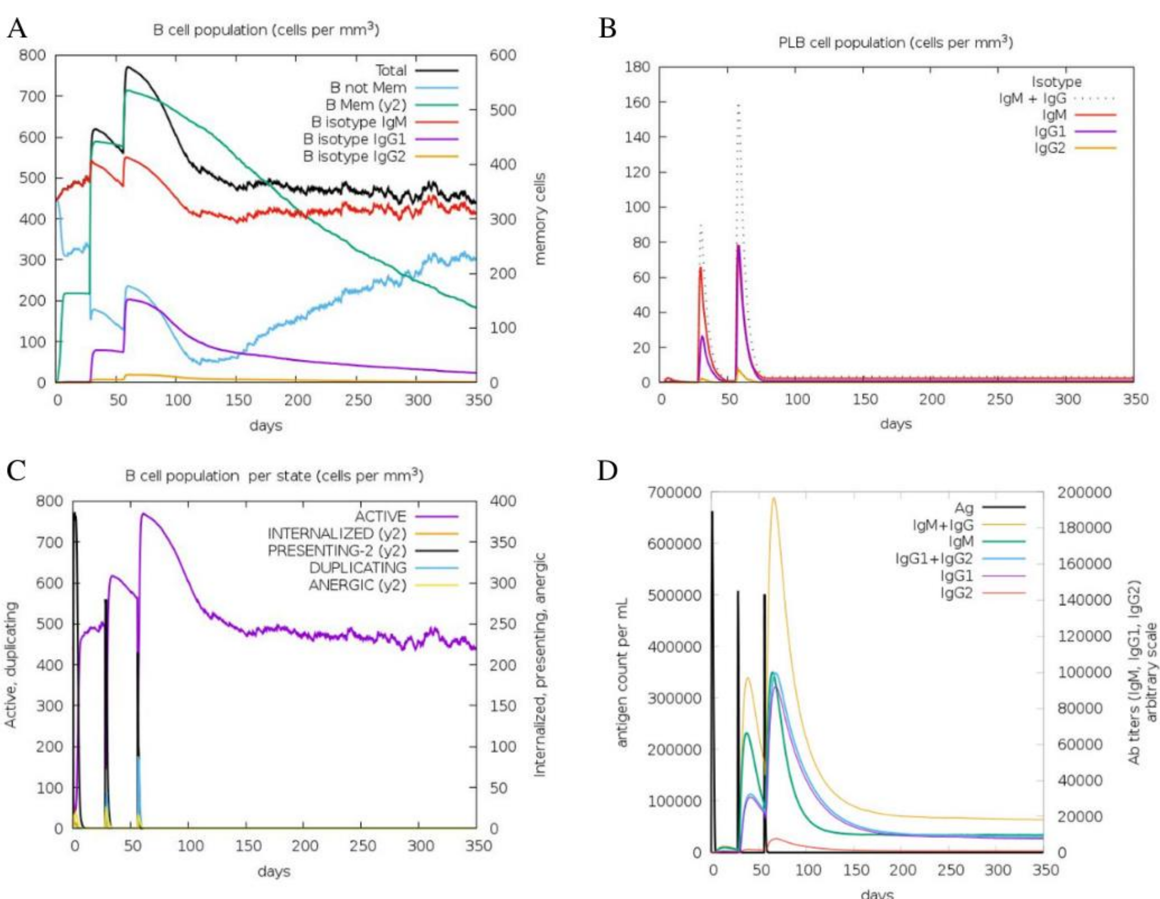


Figure 5: Immuno simulation results regarding B cell population and immunoglobulins production.

(A) B-cell population per state, (B) changings in B-cell population in the three injections, (C) B-cell population per state in the stages of injections (black vertical lines) (D) Differential production of immunoglobulins (colored peaks) during the antigen injections (black vertical lines).



11. 3 PRODUÇÃO ACADÊMICA, *CURRICULUM VITAE*

Nome Thaís Cristina Vilela Rodrigues

Formação acadêmica/titulação

2019 Mestrado em Genética. Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil Título: Predição de vacina multiepitopos para *Mycoplasma pneumoniae* através de análises in silico Orientador: Vasco Ariston de Carvalho Azevedo Co-orientador: Siomar de Castro Soares Bolsista do(a): Coordenação de Aperfeiçoamento de Pessoal de Nível Superior

2015 - 2019 Graduação em Biomedicina. Universidade Federal do Triângulo Mineiro, UFTM, Uberaba, Brasil Título: Vacinologia reversa e genômica subtrativa revelam novos alvos

terapêuticos contra *Mycoplasma pneumoniae*, um agente causador da pneumonia Orientador: Siomar de Castro Soares

2012 - 2014 Ensino Médio (2o grau). Escola Estadual Vigário José Alexandre, EEVJA, Brasil, Ano de obtenção: 2014

Formação complementar

2020 - 2020 Curso de curta duração em COVID-19 Atualização e evidências para profissionais da saúde.

Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paul, HCFMUSP, Brasil

2020 - 2020 Curso de curta duração em WORKSHOP DE PYTHON PARA DADOS BIOLÓGICOS 2020. (Carga horária: 33h).

Universidade de São Paulo, USP, Sao Paulo, Brasil

2020 - 2020 Curso de curta duração em Diagnóstico Molecular. (Carga horária: 2h).

Sinapse Biotecnologia Ltda, SINAPSE, Sao Paulo, Brasil

2020 - 2020 Curso de curta duração em I Workshop Online de Bioinformática. (Carga horária: 30h).

Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil

2019 - 2019 Curso de curta duração em Introdução à Programação para Bioinformática. (Carga horária: 7h).

Udemy, UDEMY, Brasil

2019 - 2019 Curso de curta duração em Introdução à Linguagem R. (Carga horária: 4h).

Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil

2019 - 2019 Curso de curta duração em III Curso de Verão em Bioinformática da UFMG. (Carga horária: 30h).

Universidade Federal de Minas Gerais, UFMG, Belo Horizonte, Brasil

2018 - 2018 Curso de curta duração em IV Curso de Inverno em Oncologia Molecular. (Carga horária: 40h).

Instituto Social Hospital de Câncer de Barretos, ISCB, Barretos, Brasil

2018 - 2018 Curso de curta duração em Cultura de células e ensaios funcionais. (Carga horária: 4h).

Instituto Social Hospital de Câncer de Barretos, ISCB, Barretos, Brasil

2018 - 2018 Curso de curta duração em Análise proteica: western blotting e imuno-histoquímica. (Carga horária: 4h).

Instituto Social Hospital de Câncer de Barretos, ISCB, Barretos, Brasil

2018 - 2018 Curso de curta duração em Bioinformática. (Carga horária: 4h).

Instituto Social Hospital de Câncer de Barretos, ISCB, Barretos, Brasil

2018 - 2018 Curso de curta duração em PCR: convencional, tempo real e digital. (Carga horária: 4h).
Instituto Social Hospital de Câncer de Barretos, ISCB, Barretos, Brasil
2018 - 2018 Produção Oral: interações acadêmicas nível B1,. . (Carga horária: 32h).
Universidade Federal do Triângulo Mineiro, UFTM, Uberaba, Brasil
2017 - 2017 II Workshop de Doenças Infecciosas. . (Carga horária: 12h).
Universidade Federal de Uberlândia, UFU, Uberlandia, Brasil
2017 - 2017 Curso de curta duração em Utilização de Técnicas de Bilogia Molecular Aplicadas ao Estudo da Imunolog. (Carga horária: 8h).
Universidade Federal do Triângulo Mineiro, UFTM, Uberaba, Brasil
2017 - 2017 Curso de curta duração em Avaliação da viabilidade celular através da colorimetria e fluorescência. (Carga horária: 5h).
Universidade Federal de Uberlândia, UFU, Uberlandia, Brasil
2016 - 2016 Curso de curta duração em Perícia em crimes sexuais. (Carga horária: 6h).
Universidade Federal do Triângulo Mineiro, UFTM, Uberaba, Brasil
2015 - 2015 Curso de curta duração em Extração e quantificação de DNA de células sanguíneas. (Carga horária: 6h).
Universidade Federal do Triângulo Mineiro, UFTM, Uberaba, Brasil

Atuação profissional

1. Universidade Federal de Minas Gerais - UFMG

Vínculo institucional
2020 - 2021 Vínculo: Professor , Enquadramento funcional: Professor de estágio em docência, Regime: Parcial
Outras informações:
Participou das 15 semanas de aulas na disciplina Genética e Evolução (BIG 601), e ministrou 04 (quatro) aulas para a graduação da Medicina Veterinária
2020 - 2020 Vínculo: Professor estágio em docência , Enquadramento funcional: Professor de estágio em docência, Regime: Parcial

2. hospital de câncer de Barretos – Fundação Pio XII - HCB

Vínculo institucional
2019 - 2019 Vínculo: Visita observacional , Enquadramento funcional: Estágio observacional dep. Biomedicina , Carga horária: 56, Regime: Integral

3. Universidade Federal do Triângulo Mineiro - UFTM

Vínculo institucional

2018 - 2018 Vínculo: Colaborador , Enquadramento funcional: Monitora da disciplina de Imunologia , Carga horária: 10, Regime: Parcial
2018 - 2018 Vínculo: Voluntátio , Enquadramento funcional: Iniciação científica , Carga horária: 12, Regime: Parcial
Outras informações:
Estágio Supervisionado em Pesquisa I e II, da Universidade Federal do Triângulo Mineiro, UFTM, Uberaba/MG. O trabalho desenvolvido foi intitulado "Vacinologia reversa e genômica subtrativa revelam novos alvos terapêuticos contra Mycoplasma pneumonia, um agente causador da pneumonia", sob orientação do Prof. Dr. Siomar de Castro Soares.
2017 - 2018 Vínculo: Bolsista , Enquadramento funcional: Iniciação científica , Carga horária: 12, Regime: Parcial
Outras informações:
Projeto: ESTUDO MORFOMÉTRICO E MORFOLÓGICO DO INTESTINO DE FÊMEAS MIF NOCAUTE INFECTADAS POR TOXOPLASMA GONDII DURANTE A GESTAÇÃO Realizado pela bolsista, Thaís Cristina Vilela Rodrigues sob orientação de Angelica de Oliveira Gomes, sendo parte do Programa de Iniciação Científica BIC/FAPEMIG.
2017 - 2017 Vínculo: Voluntário , Enquadramento funcional: Monitora de Imunologia , Carga horária: 30, Regime: Parcial
2017 - 2017 Vínculo: Voluntário , Enquadramento funcional: Monitora de Imunologia , Carga horária: 30, Regime: Parcial

Prêmios e títulos

2021 Menção Honrosa na Categoria Melhor Trabalho do SIMBRAVAC - Análise Pan-Genômica de Moraxella catarrhalis com Predição de Novos Alvos Vacinais e Alvos de Drogas, SIMBRAVAC
2021 Menção Honrosa na Categoria Melhor Trabalho do SIMBRAVAC - Prospecção in silico de epítopos antigênicos para diagnóstico da Hanseníase, SIMBRAVAC
2021 Menção Honrosa na Categoria Melhor Trabalho do SIMBRAVAC - trabalho intitulado: Uso de Abordagens de Imunoinformática para Construção de uma Vacina Multi-Epítopo Contra Mycoplasma pneumoniae, SIMBRAVAC
2020 Prêmio de melhor poster para o trabalho intitulado: IMMUNOINFORMATICS APPROACHS TO DESIGN A CHIMERIC MULTIEPITOPE VACCINE AGAINST MYCOPLASMA PNEUMONIAE, na sessão de poster: Proteins and Proteomics no, X-meeting

Produção bibliográfica

Artigos completos publicados em periódicos

1. Clique para acessar o artigo FELICE, ANDREI G.; ALVES, LEANDRO G.; FREITAS, ALISSA S. F.; RODRIGUES, THAÍS C. V.; JAISWAL, ARUN K.; TIWARI, SANDEEP; GOMES, LUCAS G. R.; MIRANDA, FÁBIO M.; RAMOS, ROMMEL T. J.; AZEVEDO, VASCO; OLIVEIRA, LETÍCIA C.; OLIVEIRA, CARLO J.; SOARES, SIOMAR D. C.; BENEVIDES, LEANDRO J.
Pan-genomic analyses of 47 complete genomes of the genus and prediction of new vaccine targets and virulence factors of the species. JOURNAL OF BIOMOLECULAR STRUCTURE & DYNAMICS. Fator de Impacto(2018 JCR): 3,3100, v.39, p.1 - 15, 2021.
2. Clique para acessar o artigo VILELA RODRIGUES, THAÍS CRISTINA; JAISWAL, ARUN KUMAR; DE SAROM, ALISSA; DE CASTRO OLIVEIRA, LETÍCIA; FREIRE OLIVEIRA, CARLO JOSÉ; GHOSH, PREETAM; TIWARI, SANDEEP; MIRANDA, FÁBIO MALCHER; DE JESUS BENEVIDES, LEANDRO; ARISTON DE CARVALHO AZEVEDO, VASCO; DE CASTRO SOARES, SIOMAR
Reverse vaccinology and subtractive genomics reveal new therapeutic targets against Mycoplasma pneumoniae : a causative agent of pneumonia. Royal Society Open Science. Fator de Impacto(2019 JCR): 2,6470, v.6, p.190907 - , 2019.

Trabalhos publicados em anais de eventos (resumo)
1. MARCON, C. F.; RODRIGUES, T. C. V.; DE OLIVEIRA GOMES; FERREIRA, P.T.M.; CHICA, J.E.L.
ESTUDO MORFOLÓGICO E MORFOMÉTRICO DO INTESTINO DELGADO DE FÊMEAS MIF NOCAUTE INFECTADAS POR Toxoplasma gondii In: 4ª JORNADA INTEGRADA DE ENSINO, PESQUISA E EXTENSÃO da UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO, 2018, Uberaba.
Anais da IV Jornada Integrada de Ensino, Pesquisa e Extensão da UFTM, 2018. , 2018.
2. RODRIGUES, T. C. V.; DE OLIVEIRA GOMES; MARCON, C. F.; CHICA, J.E.L.; FERREIRA, P.T.M.
Estudo morfométrico e morfológico do intestino de fêmeas MIF nocaute infectadas por Toxoplasma gondii durante a gestação In: 4a JORNADA INTEGRADA DE ENSINO, PESQUISA E EXTENSÃO da UNIVERSIDADE FEDERAL DO TRIÂNGULO MINEIRO,, 2018, Uberaba.
Anais da IV Jornada Integrada de Ensino, Pesquisa e Extensão da UFTM, 2018.. , 2018.
3. OLIVEIRA , A. R.; LELIS, G. L.; RODRIGUES, T. C. V.; CIPRIANI, E. G.; GONCALVES , A. L. D.; MARCON, C. F.; LIMA, V. M.; CASTRO, J. A.
Liga de Biologia Molecular In: JORNADA INTEGRADA DE ENSINO, PESQUISA E EXTENSÃO - JIEPE, 2018, Uberaba.
4ª Jornada Integrada de Ensino, Pesquisa e Extensão - JIEPE - Uberaba. , 2018. v.4.

Trabalhos publicados em anais de eventos (resumo expandido)

1. RODRIGUES, T. C. V.; ARISTON DE CARVALHO AZEVEDO, VASCO; SALES-CAMPOS, H.; JAISWAL, ARUN KUMAR; KATO, R. B.; TOSTA, S. F. O.; GIOVANETTI, M.; SOARES, S. C.; TIWARI, S.
PREDIÇÃO DE VACINA MULTI-EPÍTOPO CONTRA Mycoplasma pneumoniae APLICANDO ABORDAGENS DE IMUNOINFORMÁTICA In: VII Simpósio de Microbiologia da UFMG - CONECTA SIM 2020, 2020
ANAIS do CONECTA SIM 2020. , 2020.

Artigos em jornal de notícias

1. RODRIGUES, T. C. V.
A Liga de Biologia Molecular. Biomed Informa. , p.1 - 1, 2018.

Apresentação de trabalho e palestra

1. WU, M. M. S.; CEBALLOS, V. A. S.; FELICE, A. G.; RODRIGUES, THAÍS C. V.; BENEVIDES, L. J.; OLIVEIRA, L.C.; SOARES, S. C.
Análise Pan-Genômica de Moraxella catarrhalis com Predição de Novos Alvos Vacinais e Alvos de Drogas, 2021. (Simpósio,Apresentação de Trabalho)
2. LEMES, M. R.; VILELA RODRIGUES, THAÍS CRISTINA; TIWARI, S.; SOARES, S. C.; SILVA, M. V.
Prospecção in Silico de Epítopos Antigênicos para Diagnóstico da Hanseníase, 2021. (Simpósio,Apresentação de Trabalho)
3. VILELA RODRIGUES, THAÍS CRISTINA; TIWARI, S.; AZEVEDO, V. A. C.; JAISWAL, A. K.; SALES-CAMPOS, H.; KATO, R. B.; SOARES, S. C.
Uso de Abordagens de Imunoinformática para Construção de uma Vacina Multi-Epítopo Contra Mycoplasma pneumoniae, 2021. (Simpósio,Apresentação de Trabalho)
4. VILELA RODRIGUES, THAÍS CRISTINA
Uso de abordagens de vacinologia reversa e imunoinformática para o design de vacina multi-epitopo contra Mycoplasma pneumoniae, 2021. (Conferência ou palestra,Apresentação de Trabalho)
5. VILELA RODRIGUES, THAÍS CRISTINA; DE CASTRO SOARES, SIOMAR; KATO, R. B.; ARISTON DE CARVALHO AZEVEDO, VASCO; TIWARI, S.
DESING OF MULTI-EPITOPE VACCINE AGAINST Mycoplasma pneumoniae APPLYING IMMUNOINFORMATICS APPROACHES, 2020. (Simpósio,Apresentação de Trabalho)
6. VILELA RODRIGUES, THAÍS CRISTINA
Docking Molecular: Bioinformática na predição de alvos de drogas, 2020. (Conferência ou palestra,Apresentação de Trabalho)

7. VILELA RODRIGUES, THAÍS CRISTINA; JAISWAL, A. K.; SALES-CAMPOS, H.; GIOVANETTI, M.; KATO, R. B.; TOSTA, S. F. O.; TIWARI, S.; AZEVEDO, V. A. C.; SOARES, S. C.
IMMUNOINFORMATICS APPROACHS TO DESIGN ACHIMERIC MULTI-EPITOPE VACCINE AGAINST MYCOPLASMA PNEUMONIAE, 2020. (Congresso,Apresentação de Trabalho)
8. RODRIGUES, THAÍS C. V.; TIWARI, S.; AZEVEDO, V. A. C.; KATO, R. B.; TOSTA, S. F. O.; SOARES, S. C.
In silico approaches for Mycoplasma pneumoniae multiepitope vaccine construction, 2020. (Congresso,Apresentação de Trabalho)
9. VILELA RODRIGUES, THAÍS CRISTINA; JAISWAL, A. K.; ANCHIETA, A. S. F. F.; OLIVEIRA, L.C.; OLIVEIRA, C. J. F.; GHOSH, P.; TIWARI, S.; MIRANDA, F. M.; BENEVIDES, L. J.; AZEVEDO, V. A. C.; SOARES, S. C.
New vaccine and drug targets of Mycoplasma pneumoniae revealed by reverse vaccinology and subtractive genomics, 2019. (Congresso,Apresentação de Trabalho)
10. RODRIGUES, THAÍS C. V.; JAISWAL, A. K.; ANCHIETA, A. S. F. F.; OLIVEIRA, L.C.; OLIVEIRA, C. J. F.; GHOSH, P.; TIWARI, S.; MIRANDA, F. M.; BENEVIDES, LEANDRO J.; AZEVEDO, V. A. C.; SOARES, S. C.
Reverse vaccinology and subtractive genomics reveal new therapeutic targets against Mycoplasma pneumoniae: a causative agent of pneumonia,, 2019. (Congresso,Apresentação de Trabalho)
11. RODRIGUES, T. C. V.; DE OLIVEIRA GOMES; MARCON, C. F.; CHICA, J.E.L.; FERREIRA, P.T.M.
Estudo morfométrico e morfológico do intestino de fêmeas MIF nocaute infectadas por Toxoplasma gondii durante a gestação, 2018. (Outra,Apresentação de Trabalho)
12. VILELA RODRIGUES, THAÍS CRISTINA; ANCHIETA, A. S. F. F.; SOARES, S. C.; OLIVEIRA, C. J. F.
Identificação de potenciais alvos vacinais de Mycoplasma pneumoniae via Genômica Comparativa, 2018. (Outra,Apresentação de Trabalho)
13. RODRIGUES, THAÍS C. V.; OLIVEIRA, C. J. F.; SOARES, S. C.; ANCHIETA, A. S. F. F.
Identificação de potenciais alvos vacinais de Mycoplasma pneumoniae via Genômica Comparativa, 2018. (Congresso,Apresentação de Trabalho)
14. OLIVEIRA , A. R.; GONCALVES , A. L. D.; MARCON, C. F.; CIPRIANI, E. G.; LELIS, G. L.; CASTRO, J. A.; RODRIGUES, THAÍS C. V.; LIMA, V. M.; PEDROSA, A. L.
Liga de Biologia Molecular, 2018. (Outra,Apresentação de Trabalho)

15. MARCON, C. F.; CHICA, J.E.L.; FERREIRA, P.T.M.; RODRIGUES, THAÍS C. V.; DE OLIVEIRA GOMES
MORFOLÓGICO E MORFOMÉTRICO DO INTESTINO DELGADO DE FÊMEAS MIF NOCAUTE INFECTADAS POR Toxoplasma gondii, 2018. (Outra,Apresentação de Trabalho)

Produção técnica

Trabalhos técnicos

1. RODRIGUES, THAÍS C. V.; OLIVEIRA , A. R.
Liga de Biologia Molecular, 2018

2. CIPRIANI, E. G.; RODRIGUES, THAÍS C. V.; GONCALVES , A. L. D.; CASTRO, J. A.; LIMA, V. M.; MARCON, C. F.; LELIS, G. L.; OLIVEIRA , A. R.; PEDROSA, A. L.
Liga de Biologia Molecular, 2017

Demais produções técnicas

1. RODRIGUES, THAÍS C. V.; SOARES, S. C.; FELICE, A. G.; CEBALLOS, V. A. S.; ZEN, F. L.
Bioinformática e Ciências Ômicas, 2020. (Extensão, Curso de curta duração ministrado)

2. VILELA RODRIGUES, THAÍS CRISTINA
Perspectivas do Mestrado na Genética da UFMG, 2020. (Extensão, Curso de curta duração ministrado)

3. RODRIGUES, T. C. V.
Aula teórica sobre Vacinologia Reversa e Genômica Subtrativa na Predição de Alvos Vacinais e de Drogas utilizando Bioinformática, 2019. (Extensão, Curso de curta duração ministrado)

Citações

Web of Science

Total de trabalhos: 1Total de citações: 6 Fator H: 1

Totais de produção

Produção bibliográfica

Artigos completos publicados em periódico	2
Jornais de Notícias	1
Trabalhos publicados em anais de eventos	4
Apresentações de trabalhos (Conferência ou palestra)	2
Apresentações de trabalhos (Congresso)	5
Apresentações de trabalhos (Simpósio)	4
Apresentações de trabalhos (Outra)	4

Produção técnica

Curso de curta duração ministrado (extensão)	3
Extensão tecnológica	2

Eventos

Participações em eventos (congresso)	12
Participações em eventos (seminário)	1
Participações em eventos (simpósio)	
Participações em eventos (encontro)	3
Participações em eventos (outra)	1
Organização de evento (congresso)	4
Organização de evento (outro)	3