



Ministério da Saúde

FIOCRUZ

Fundação Oswaldo Cruz



ESCOLA POLITÉCNICA DE SAÚDE
JOAQUIM VENÂNCIO

Larissa Emiliano de Araujo

**O USO DA BIOINFORMÁTICA PARA A DESCOBERTA DE NOVOS ALVOS PARA O
DIAGNÓSTICO DE INFECÇÕES ATIVAS POR *LEPTOSPIRA* PATOGÊNICAS**

Rio de Janeiro

2021

Larissa Emiliano de Araujo

**O USO DA BIOINFORMÁTICA PARA A DESCOBERTA DE NOVOS ALVOS PARA O
DIAGNÓSTICO DE INFECÇÕES ATIVAS POR *LEPTOSPIRA* PATOGÊNICAS**

**Monografia apresentada à Escola Politécnica de
Saúde Joaquim Venâncio – Fundação Oswaldo
Cruz (EPSJV-Fiocruz) como requisito parcial
para aprovação no Curso Técnico em
Biotecnologia.**

Orientador(a): Fernando de Paiva Conte

Coorientador(a): Rodrigo Nunes Rodrigues da Silva

Rio de Janeiro

2021

*Dedico esse trabalho a
minha mãe, um exemplo
de mulher forte e
generosa, e ao meu pai,
um homem trabalhador
e persistente. Dedico
juntamente a toda
comunidade acadêmica
que possa desfrutar
desta presente pesquisa*

AGRADECIMENTOS

Agradeço primeiramente à Deus, por me conceder perseverança, saúde e disposição para concluir meus projetos, assim como a resiliência necessária para ultrapassar todos os grandes obstáculos que encontrei no caminho deste.

Agradeço a Fundação Oswaldo Cruz (Fiocruz) pela honra e oportunidade de aprender e crescer dentro dessa instituição que visa acima de tudo o bem público. Agradeço a Escola Politécnica de Saúde Joaquim Venâncio pela chance de poder realizar esta monografia e por toda a rica jornada que pude experienciar nesses quatro anos e meio.

Agradeço a todos os professores do ensino médio e técnico, com quem pude trocar experiências e aprender de forma ampla e variada, em especial ao professor de história André Dantas, cujas aulas só comprovam a sua inteligência, aptidão e paixão pelo que ensina.

De nível técnico, agradeço em especial a Flávia Ribeiro, por lecionar com tanto empenho e carinho, e assim me despertar o interesse por imunologia, o que motivou diretamente na escolha do tema da monografia. Agradeço também por ela ter estado presente durante todo o tempo me dando suporte e sempre ter acreditado nas minhas capacidades.

Em seguida, agradeço aos meus orientadores Fernando de Paiva e Rodrigo Nunes primeiramente pela oportunidade de participar deste projeto de pesquisa que considero de extrema relevância social, e secundamente pela atenção plena que recebi durante o processo de pesquisa e escrita. Agradeço pelas correções e toda a aprendizagem técnica, pois foi através delas que pude extrair o melhor de mim e conseqüentemente concluir o projeto da melhor forma possível, seguindo na direção correta. Agradeço juntamente a Biomanguinhos e ao Laboratório de Tecnologia de Anticorpos Monoclonais.

Agradeço a minha família por me mandarem boas energias e por sempre acreditarem no meu potencial, em especial meu pai pelo grande incentivo ao meu progresso, minha irmã Hosana Emiliano, que se demonstra sempre disposta a me ajudar, e a Tatiana Neres por me ouvir e sempre

ter algo positivo para me dizer. Agradeço também ao Gabriel Emiliano, pela sua visão sensível e seus feedbacks sempre construtivos, que em momentos de exaustão me foram de grande valia para ampliar a minha visão.

Agradeço aos meu amigos: Caroline Sousa, Giulia Brum e Luan Avelino, que sempre estiveram dispostos a me ouvir e dar todo o apoio e ânimo que eu precisasse durante essa jornada, ao Guilherme Batalha, por sempre me fazer enxergar o lado cômico de tudo e me lembrar de que a vida não precisa ser levada tão a sério. Agradeço a Ana Beatriz, por sua capacidade de acolhimento e empatia, que a tornaram capaz de me oferecer o maior conforto em momentos de altos e baixos. Agradeço a Nicolly Martins, por sempre me lembrar dos meus pontos fortes e me incentivar, como também pela sua objetividade, que me fizeram ter um olhar mais resoluto diante dos obstáculos. Agradeço a Bianca Santana, por todo o suporte e palavras de incentivo que ela me dirigiu sempre que achou que eu precisasse ouvir. Agradeço aos meus amigos Gabriel Melo e Pedro Henrique por sempre me encorajarem a seguir a diante.

Agradeço a Ana Beatriz Barbosa, que se tornou para mim uma grande inspiração e foi fonte de muito aprendizado e crescimento como pessoa, com ela eu aprendi que o conhecimento precisa envolver generosidade.

Conhecimento sem generosidade é apenas vaidade

Ana Beatriz Barbosa Silva

RESUMO

A leptospirose é uma infecção sistêmica associada a bactérias patogênicas do gênero *Leptospira*, presentes na urina de roedores infectados que, em regiões periféricas nas áreas urbanas, estão associadas ao aparecimento cíclico de surtos de casos de infecção em decorrências de enchentes e inundações anuais. Dados do Ministério da Saúde demonstram que cerca de 3500 novos casos são confirmados ao longo do ano no Brasil, com uma estimativa de 60 mil mortes anuais no mundo. Contudo, o teste diagnóstico ouro, a prova por soroadglutinação microscópica (SAM) pode demorar até 15 dias para apontar resultados, o que pode resultar no agravamento do quadro clínico do paciente. Em contrapartida, os testes sorológicos rápidos existentes, capazes de detectar apenas anticorpos IgM, não são capazes de detectar reinfecções ativas da doença. Em decorrência disso, diversos casos de infecção por bactérias do gênero *Leptospira* podem estar associados a diagnósticos errôneos, levando à subnotificação de casos e inúmeros óbitos. A bioinformática surge nesse contexto como um campo promissor e interdisciplinar que possibilita, por exemplo, estudos de identificação de alvos de anticorpos específicos, explorados por diversas ferramentas computacionais e banco de dados de genoma e secretômica, para o desenvolvimento de novos testes sorológicos. Com isso, a imunoinformática se torna uma vantajosa alternativa para auxiliar nos problemas de diagnóstico da leptospirose, sendo possível a predição de proteínas conservadas e secretadas por diferentes espécies de leptospiros patogênicas no curso de uma infecção ativa. Dessa forma, o uso da bioinformática nos permite a identificação de novos alvos e, futuramente, o desenvolvimento de novos e melhores reativos de diagnóstico capazes de detectar infecção ativa no paciente, ainda que numa condição de reexposição. Sendo assim, essa pesquisa se consiste em duas etapas: a primeira baseia-se em um estudo de prospecção, com o intuito de identificar proteínas secretadas durante a infecção por leptospiros patogênicas, através de levantamento bibliográfico. Após a escolha das proteínas candidatas, na segunda etapa do projeto, estas foram analisadas *in silico*, utilizando uma combinação de diferentes ferramentas de bioinformática, a fim de avaliar o potencial de secreção, virulência e antigenicidade das proteínas candidatas. Foram selecionadas ao todo cinco proteínas candidatas e ao final do estudo foram consideradas como secretadas as proteínas: Papalisina 1, Termolisina e Lipoproteína P, que acreditamos ser uma proteína do tipo *moonlighting*, na qual pode

desempenhar diversas funções. As proteínas protease dependente de zinco tldD e Metionina aminopeptidase (Map) foram apontadas como não-secretadas. Quanto a análise de virulência e antigenicidade, as três proteínas consideradas secretadas foram todas associadas a virulência e antigenicidade. Diante dos resultados, espera-se que este presente estudo possa contribuir para futuras pesquisas que busquem gerar um diagnóstico mais eficaz para leptospirose.

Palavras-chave: leptospirose; diagnóstico; bioinformática

Sumário

1. INTRODUÇÃO	10
2. JUSTIFICATIVA	15
3. OBJETIVOS	19
3.1 Objetivos específicos	19
4. METODOLOGIA	20
4.1 Seleção de proteínas alvo	20
4.2 Predição do potencial de secreção	20
4.3 Predição do papel das proteínas quanto a virulência	25
4.4 Avaliação da capacidade de antigenicidade das proteínas	26
5. RESULTADOS	27
5.1 Proteínas selecionadas	27
5.2 Predição de secreção	28
5.2 Análise de virulência	31
5.3 Avaliação de antigenicidade	32
6. CONCLUSÃO	33
7. CRONOGRAMA	34
8. REFERÊNCIAS BIBLIOGRÁFICAS	35
8. ANEXO 1 – Sisgen – Cadastro N° A041B9C	38

LISTA DE FIGURAS

Figura 1. Morfologia da bactéria <i>Leptospira interrogans</i>	10
Figura 2. Ciclo de transmissão da leptospirose (adaptado)	11
Figura 3. Fotomicrografia da técnica SAM	13
Figura 4. Teste rápido (Imunocromatografia de fluxo lateral)	14
Figura 5. Teste por cultivo bacteriano	15
Figura 6. Reação em cadeia da polimerase (PCR)	15
Figura 7. Fluxograma de trabalho	21

LISTA DE QUADROS

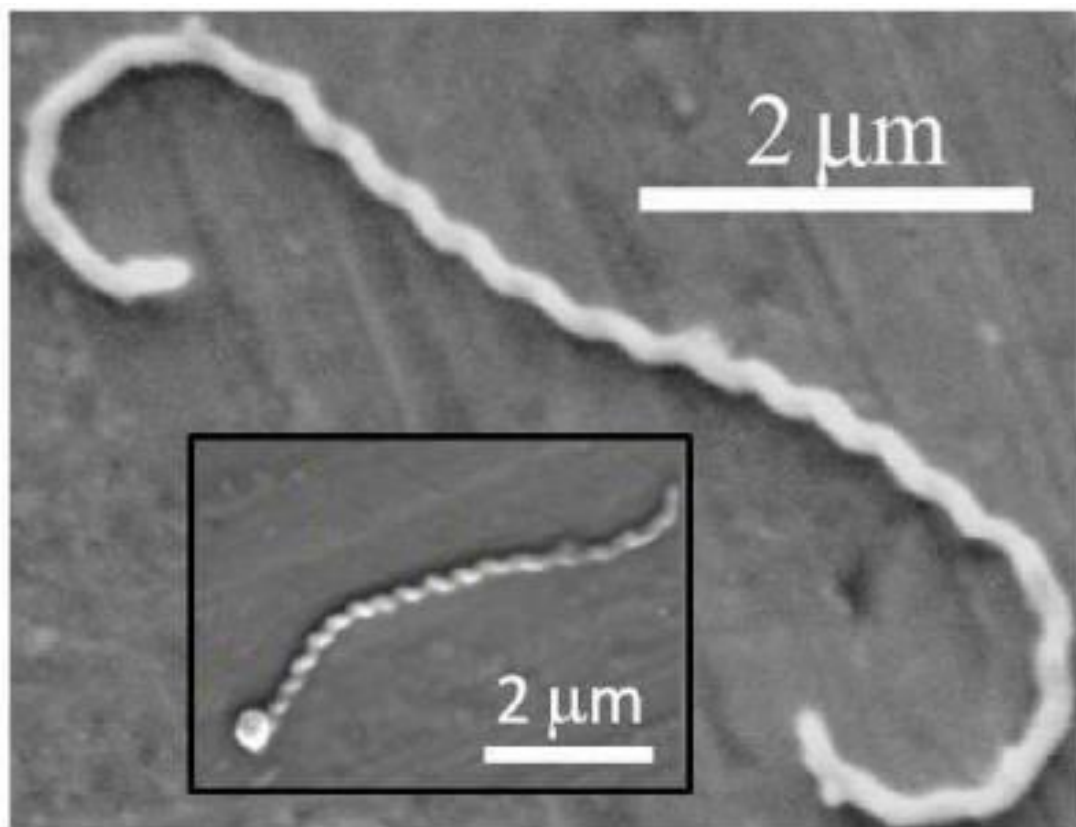
Quadro 1. Diagnósticos	16
Quadro 2. Proteínas selecionadas	27
Quadro 3. Resultados da predição de secreção	29
Quadro 4. Resultados da predição de secreção pela ferramenta VICMpred	30
Quadro 5. Resultados de virulência pela ferramenta Virulentpred	31
Quadro 6. Resultados de antigenicidade pela ferramenta Vaxijen V 2.0	32

1. INTRODUÇÃO

1.1 *Leptospira* sp.

A leptospirose tem como agente etiológico a bactéria aeróbica obrigatória, Gram-negativa e patogênica do gênero *Leptospira* (OLIVEIRA; GUIMARÃES; MEDEIROS, 2009). Vale ressaltar que atualmente já são conhecidas mais de 60 espécies bacterianas, que podem ser divididas em: 17 espécies patogênicas ao homem, 21 espécies de patogenicidade intermediária e 24 espécies saprófitas, contando com 300 sorovares que se subdividem em 24 sorogrupos de acordo com a sua complexidade antigênica (VINCENT, et al. 2019).

Figura 1 – Morfologia da bactéria *Leptospira interrogans*
(Imagem de microscopia eletrônica de varredura)



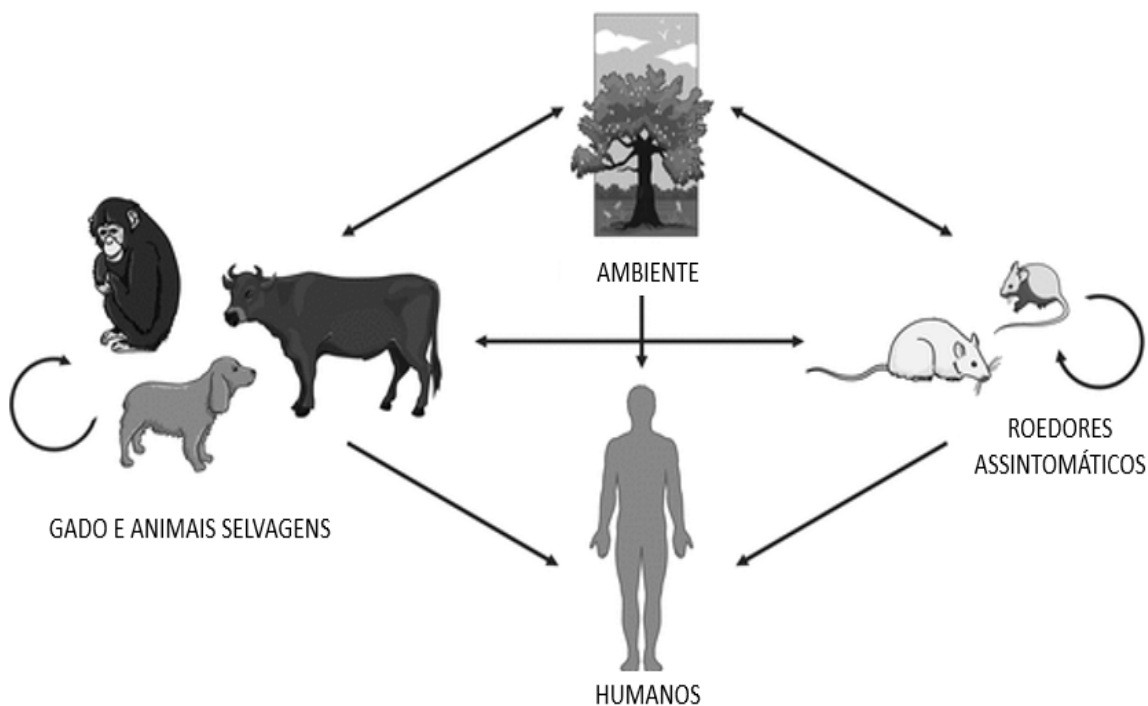
Fonte: (KOENS; LAUGA, 2014)

1.2 Transmissão

A leptospirose é uma doença zoonótica infecciosa, naturalmente transmitida ao ser humano através do contato com a urina e secreções de animais infectados pela doença.

Apesar dos roedores figurarem entre os principais vetores de transmissão da leptospirose ao homem, o envolvimento de uma variedade de hospedeiros, incluindo mamíferos e répteis, já foi demonstrada. Além disso, a transmissão também pode ocorrer pelo contato direto do ser humano com água ou lama contaminados, e sabe-se que, uma vez no ambiente, a bactéria pode permanecer viável por até 180 dias, o que aumentam as chances de infecção (PELISSARI, et al. 2011).

Figura 2 – Ciclo de transmissão da leptospirose (adaptado)



Fonte: https://media.springernature.com/original/springer-static/image/chp%3A10.1007%2F978-3-319-684932_8/MediaObjects/420908_1_En_8_Fig2_HTML.gif

1.3 Leptospirose – A doença

As manifestações clínicas associadas à leptospirose, podem incluir: febre alta, dor de cabeça, calafrios, sangramento, dores musculares, e vômitos, cujos sintomas e eventuais sinais da doença apresentam um período de incubação que pode variar entre o primeiro até o trigésimo dia, sendo mais comum o aparecimento dos sintomas entre uma semana a quinze dias.

Como uma infecção sistêmica, a leptospirose pode evoluir para a fase grave denominada Síndrome de Weil (SW), que se caracteriza por quadros de icterícia, disfunção renal e hemorragias. (PEREIRA, et al. 2020).

De acordo com o Ministério da Saúde, aproximadamente 15% dos pacientes progride para a fase tardia da doença, onde o óbito pode ocorrer nas primeiras vinte e quatro horas de internação. (MINISTÉRIO DA SAÚDE, 2009).

1.4 Leptospirose - Epidemiologia

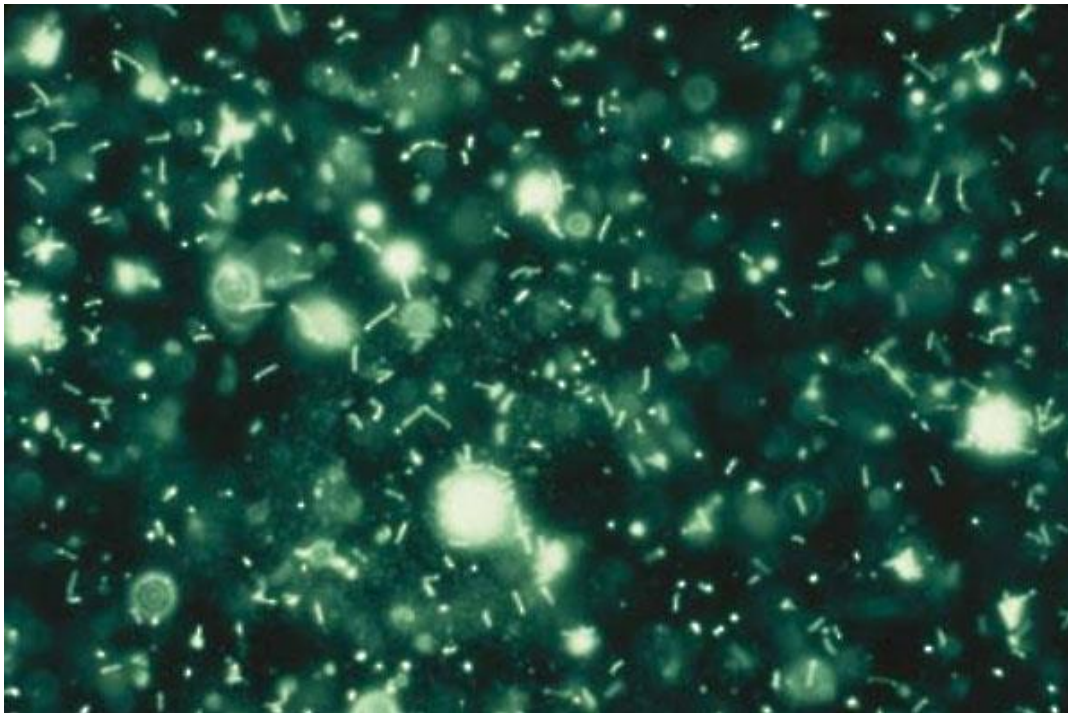
A doença apresenta uma estimativa de mais de um milhão de casos e cerca de 60 mil mortes por ano no mundo. (COSTA, et al. 2015). Enquanto no Brasil, de acordo com a estimativa do Ministério da Saúde, uma média de 13.000 casos são notificados por ano, sendo 3.500 confirmados, associando-se a um percentual de hospitalização de 75% dos casos (MINISTÉRIO DA SAÚDE, 2014).

A leptospirose é uma enfermidade de caráter endêmico que está presente especialmente em regiões tropicais e zonas urbanas, principalmente em áreas periféricas (MARINHO; 2012). Embora os casos da doença sejam registrados em todos os meses do ano no Brasil, o aumento significativo no número de casos coincide com períodos de alto índice pluviométrico e a locais onde infraestrutura e saneamento básico são precários. (GENOVEZ; 2009). Dessa forma, as áreas periféricas urbanas são as mais propensas a ocorrências de enchentes e inundações, que em conjunto com a precariedade dos serviços de saneamento básico contribuem para a proliferação de roedores, os principais vetores da leptospirose.

1.5 Leptospirose – Métodos diagnósticos

Quanto aos métodos laboratoriais de diagnóstico disponíveis, destacam-se: o método padrão-ouro que consiste na prova de soroaglutinação microscópica (SAM); os testes sorológicos para detecção de anticorpos IgG e IgM; o teste rápido (imunocromatografia de fluxo lateral); o teste por cultivo bacteriano e os testes moleculares utilizando a reação em cadeia da polimerase (PCR) (SIMÕES, et al. 2015).

Figura 3: Fotomicrografia da técnica SAM
(Utilizando microscopia de campo escuro)



Fonte: CDC - Centers for Disease Control and Prevention, Mrs. M. Gatton, figura 2888 Site:
<https://www.zoetis.com.br/prevencao/caesegatos/posts/c%C3%A3es/como-a-soroaglutina%C3%A7%C3%A3o-microsc%C3%B3pica-atua-no-diagn%C3%B3stico-da-leptospirose-canina.aspx>

A prova de soroaglutinação microscópica (SAM) é realizada através da utilização do soro do paciente, para assim testar a capacidade dos anticorpos aglutinarem em meio líquido espécimes de leptospiras e assim identificar a presença da infecção. A problemática desse método diagnóstico se dá pelo longo período (15 dias) para obtenção de uma resposta conclusiva nos ensaios por SAM, o que pode acarretar o agravamento do quadro clínico do paciente.

Figura 4: Teste rápido (Imunocromatografia de fluxo lateral)



Fonte: <https://portalarquivos2.saude.gov.br/images/jpg/2020/fevereiro/13/capa-13022020.jpg>

O diagnóstico realizado pelos testes rápidos atua visando a detecção da presença de anticorpos, quando o organismo teve contato com a bactéria. Porém os testes rápidos disponíveis, limitam-se apenas na detecção do anticorpo IgM, marcador sorológico que indica uma exposição recente ao patógeno, não sendo possível apontar um caso de reinfecção, que só poderia ser detectada precocemente pela sinalização do anticorpo IgG nos testes rápidos. Por isso, eles são incapazes de detectar reinfecções ativas da doença, se tornando então, uma problemática quando consideramos o curto prazo do período de incubação da infecção (PEREIRA, et al. 2020).

Figura 5: Teste por cultivo bacteriano



Fonte: <https://factdr.com/health-conditions/leptospirosis/>

O teste por cultivo bacteriano se dá pelo isolamento e identificação do agente patogênico através do cultivo bacteriano, que é cultivado em um meio específico para o seu crescimento, e é realizado com materiais específicos recolhidos no sítio da infecção (SIMÕES, et al. 2015). Como é preciso aguardar a fase de crescimento da bactéria, não é capaz de fazer uma detecção precoce da doença.

Figura 6: Reação em cadeia da polimerase (PCR)



Fonte: <https://www.blogs.unicamp.br/synbiobrasil/2011/03/09/oque-e-um-pcr-como-montar-um-programa-de-pcr/>

A reação em cadeia pela polimerase, mais conhecido como PCR, é capaz de identificar possíveis infecções bacterianas, através da ampliação de sequência de ácidos nucleicos dos patógenos, identificando assim a presença da bactéria (SIMÕES, et al. 2015). Portanto, embora seja capaz de realizar uma detecção precoce da infecção e reconhecer infecções, a técnica requer um procedimento de alto custo, o que o torna inviável em larga escala.

Em suma, os desafios associados ao diagnóstico da leptospirose se dão pelo longo período de tempo (15 dias) para obtenção de uma resposta conclusiva nos ensaios por SAM, e também pois em casos de reinfecção do paciente, não é possível fazer a diferenciação de uma exposição prévia para uma ativa através de outras técnicas de diagnóstico, o que mascara o diagnóstico precoce e impede o tratamento correto, podendo levar o paciente á óbito (Quadro 1).

Quadro 1: Diagnósticos

Diagnóstico	SAM	Teste Molecular	Teste por Cultivo	Teste rápido
Detecção precoce	X	V	X	X
Reconhecimento de reinfecções	V	V	V	X
Resultado rápido	X	V	X	V
Diagnóstico tardio	V	X	V	V

Com isso, torna-se extremamente necessário um teste diagnóstico que possa, simultaneamente, detectar a doença de forma precoce bem como reconhecer reinfecções, além de ser economicamente viável.

1.6 Bioinformática

Frente a esses entraves no diagnóstico da leptospirose, estudos *in silico* surgem como uma estratégia promissora para o avanço da pesquisa em leptospirose, uma vez que possibilita a identificação de alvos que podem ser direcionados ao desenvolvimento de reativos diagnósticos, com base na exploração de dados do genoma, transcriptoma, proteômica ou secretômica, por diferentes ferramentas de bioinformática e biológica computacional, que sistematicamente analisam e interpretam informações biológicas. (RAPPUOLI, 2000; RAPPUOLI, BOTTOMLEY et al. 2016; BIDMOS, SIRIS et al. 2018). Esta abordagem vem sendo aplicada com sucesso contra diferentes patógenos, propiciando a descoberta de novos alvos vacinais e melhores diagnósticos. (RAPPUOLI, 2014; RAPPUOLI PIZZA et al. 2014; DE LA FUENTE, KOPACEK et al. 2016; MICHALIK, DJAHANSHIRI et al. 2016; BITENCOURT CHAVES, PERCE-DA-SILVA et al. 2017; RODRIGUES DA-SILVA, SOARES et al. 2017; CHAVES, PERCE-DA-SILVA et al. 2019; CONTE, TINOCO et al. 2019, contudo, ainda há poucos relatos do uso desta tecnologia aplicados à leptospirose (DELLAGOSTIN, GRASSMANN et al. 2017). Desse modo, a bioinformática pode permitir a maximização da identificação de alvos vacinais e diagnósticos, contribuindo para o desenvolvimento de métodos diagnósticos e formulações vacinais mais eficazes.

2. JUSTIFICATIVA

A leptospirose é uma doença zoonótica emergente causada por bactérias patogênicas do gênero *Leptospira*, que representa um problema de extrema relevância para a sociedade pois está associada à falta de saneamento básico e investimento público em infraestrutura.

Em um contexto global, estima-se o surgimento de mais de 1 milhão de novos casos por ano, como consequência do desenfreado aumento da população moradora de periferias urbanas, locais mais vulneráveis e propícios para o alastramento da doença.

Nesta perspectiva, deve-se ressaltar a importância de se investir em testes diagnósticos mais eficientes que sejam capazes de apontar uma infecção ativa e de discriminar entre outras doenças hemorrágicas com sintomatologia clínica semelhante (ex. dengue, febre amarela, hantavirose). O principal problema no diagnóstico atual para leptospirose baseia-se, dentre outros, na demora da entrega dos resultados pela técnica padrão-ouro (SAM), podendo chegar a 15 dias, aliado à deficiência de detecção dos casos de reinfecção pelos testes sorológicos atuais (pesquisa de anticorpos IgM), o que pode estar associado à subnotificação dos casos e ao aumento do número de óbitos.

Reiterando, a ausência de um diagnóstico sorológico rápido que possa detectar casos de reinfecções e infecções ativas se faz necessário e pode auxiliar na vigilância epidemiológica desta doença. Com isso, esse estudo baseia-se, em um primeiro momento, na pesquisa bibliográfica de proteínas ativamente secretadas por leptospirosas patogênicas durante o curso de uma infecção natural, utilizando para isso dados de secretoma recentemente publicados.

Em uma segunda etapa, de posse desses dados, avaliamos o potencial de secreção e de antigenicidade das proteínas selecionadas de bactérias patogênicas do gênero *Leptospira* através do uso de diferentes ferramentas de bioinformática.

Em conjunto, estes algoritmos nos permitiram selecionar alvos que sejam conservados entre diferentes espécies de leptospirosas patogênicas. Assim, acreditamos que ao final deste estudo possamos prever alvos potenciais para novos marcadores de diagnósticos, que poderão no futuro compor novos reativos diagnóstico para detecção sorológica de casos ativos de leptospirose.

3. OBJETIVO GERAL

Identificar potenciais proteínas bacterianas secretadas durante a manifestação clínica da leptospirose que apresentem potencial diagnóstico.

3.1 OBJETIVOS ESPECÍFICOS

- Identificar através de levantamento bibliográfico, proteínas bacterianas secretadas durante o curso da infecção por bactérias patogênicas do gênero *Leptospira*
- Predizer o potencial de secreção das proteínas-alvo
- Analisar o papel dessas proteínas quanto à virulência da bactéria
- Avaliar o potencial antigênico das proteínas preditas como secretadas

4. METODOLOGIA

4.1 Seleção de proteínas alvo

Com base em uma revisão de literatura, ocorreu por meio de busca em base de dados de artigos científicos de revistas indexadas, tais como Pubmed e Scielo, tendo como referência os seguintes descritores: secretoma; leptospirose; proteínas, onde foram selecionadas proteínas-candidatas ao diagnóstico da leptospirose descritas como alvos de anticorpos.

Esta etapa da seleção de proteínas alvo foi realizada com base na leitura de três artigos que apontaram as principais proteínas detectadas no sobrenadante de cultura de leptospiras patogênicas. Sendo eles: (AMAMURA; FRAGA, et al. 2017); (ESHGHI; PAPPALARDO, et al. 2015); (ZENG; ZHANG, et al. 2013)

4.2 Predição do potencial de secreção das proteínas-alvo

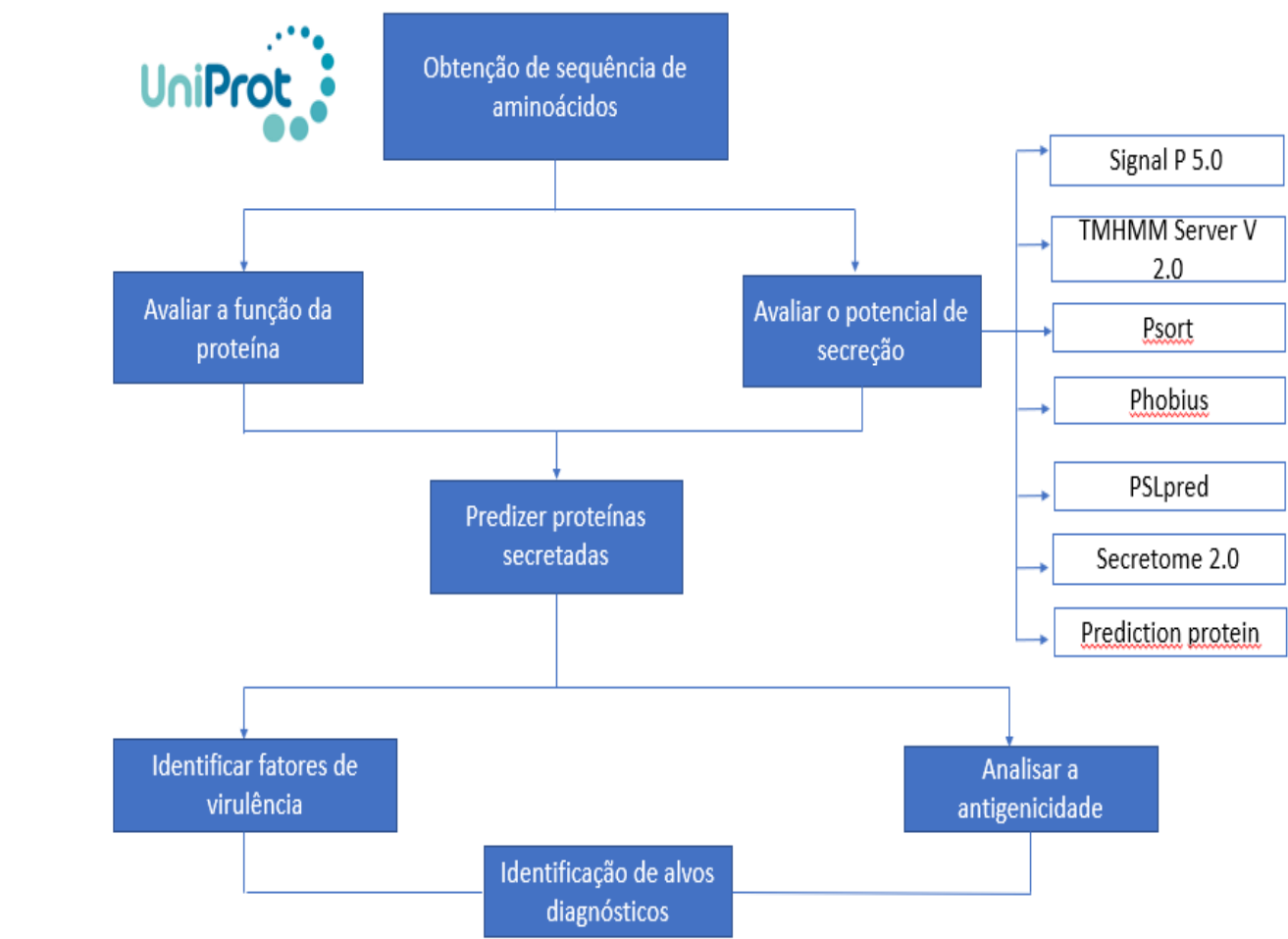
Para uma proteína ser dada como secretada, isto é, liberada para o exterior da célula, deve-se considerar um conjunto de fatores. Um deles é a existência do chamado peptídeo sinal, sequência de aminoácidos localizados na região N terminal da proteína e que tem como objetivo conduzi-la para as principais vias secretórias, por isso, é de extrema importância saber se uma proteína possui peptídeo sinal. Outro fator a ser considerado na análise de secreção é saber se a proteína em questão possui hélice transmembranar, visto que proteínas secretadas não possuem essas regiões pois do contrário, permaneceriam aderidas à membrana, de modo a não ser secretadas. Também é possível reconhecer uma proteína como secretada, através do conhecimento da sua localização celular, uma vez que esta se encontra no ambiente extracelular, enquanto uma proteína não secretada apresenta sublocalização citoplasmática.

Ao todo, foram utilizadas sete ferramentas de bioinformática para realizar a predição de secreção das proteínas selecionadas e para isso as sequências de aminoácidos de cada proteína foram encontradas utilizando o servidor Uniprot (<https://www.uniprot.org/>).

Sendo assim, as proteínas selecionadas foram avaliadas através de uma combinação de algoritmos de bioinformática, e as proteínas identificadas como alvos de resposta imune de anticorpos e proteínas secretadas serão consideradas alvos promissores para o estudo. Para que cumpram esse requisito, os indicadores são: a presença do peptídeo sinal da proteína, ausência de região transmembranar e localização no ambiente extracelular.

É importante ressaltar a existência de proteínas que porventura possam atuar em diferentes funções, ou até mesmo em funções múltiplas, sendo assim, ocupando mais de uma localização subcelular. Estas são preditas como proteínas *moonlighting* e podem ser encontradas em diversas localizações subcelulares.

Figura 7: Fluxograma de trabalho



4.2.1 Signal P 5.0

Esta ferramenta de predição (<http://www.cbs.dtu.dk/services/SignalP>) é capaz de identificar a presença do denominado sinal peptídeo da proteína através de três vias principais de secreção: Signal peptide (Sec/SPI); TAT signal peptide (Tat/SPI) e Lipoprotein signal peptide (Sec/SPII).

Caso um peptídeo sinal seja previsto, a posição de clivagem também é revelada. As proteínas que possuem peptídeo sinal, serão avaliadas como potencialmente secretadas pela análise realizada por este algoritmo.

4.2.2 TMHMM Server v 2.0

Esta ferramenta (<http://www.cbs.dtu.dk/services/TMHMM/>) é capaz de identificar, através da sequência de aminoácidos, se a proteína obtém hélice transmembranar.

As proteínas que não possuem região transmembranar serão avaliadas como potencialmente secretadas pela análise realizada por este algoritmo.

4.2.3 Psort

Este algoritmo de predição (<http://www.psort.org/psortb/>), avalia simultaneamente a existência de peptídeo sinal e de região transmembranar.

As proteínas avaliadas como “Extracelullar”, serão consideradas como potencialmente secretadas pela análise realizada por este algoritmo.

4.2.4 Phobius

Essa ferramenta (<https://phobius.sbc.su.se/>), assim como Psort, também é capaz de identificar simultaneamente a existência de peptídeo sinal e região transmembranar, de acordo com a localização celular da proteína em questão.

As proteínas avaliadas como “NON CYTOPLASMIC”, serão consideradas como potencialmente secretadas pela análise realizada por este algoritmo.

4.2.5 PSLpred

Esta ferramenta (<https://webs.iitd.edu.in/raghava/pslpred/submit.html>) é capaz de comparar a sequência da proteína a um banco de dados, relacionando a localização celular de estruturas similares. Como resultado, apresenta a localização celular predita e acurácia estimada.

As proteínas avaliadas como “Extracelullar”, serão consideradas como potencialmente secretadas pela análise realizada por este algoritmo.

4.2.6 Secretome 2.0

Esta ferramenta (<http://www.cbs.dtu.dk/services/SecretomeP/>), pode reconhecer proteínas que são secretadas por vias não clássicas de secreção. Ela é útil para analisar proteínas que não tem um peptídeo sinal, porém, podem ser secretadas. Os resultados são mostrados de modo que as proteínas secretadas tem o Secp score maior que o threshold (0.5).

As proteínas preditas com score maior que 0,5 ($>0,5$) serão consideradas como potencialmente secretadas pela análise realizada por este algoritmo.

4.2.7 Prediction Protein

Esta ferramenta de predição (<https://predictprotein.org/>) é bastante versátil e avalia um conjunto de funções como: análise de região transmembranar, sítio de ligação, estrutura secundária, localização subcelular.

As proteínas avaliadas como “Extracelular”, serão consideradas como potencialmente secretadas pela análise realizada por este algoritmo.

4.2.8 VICMpred

Esta é uma ferramenta bastante completa (<https://webs.iiitd.edu.in/raghava/vicmpred/index.html>) que tem como objetivo principal classificar as proteínas microbianas patogênicas em quatro classes funcionais: **a.** Fatores de virulência: Adesões; Toxinas; e molécula hemolítica;

b. Armazenamento e processamento de informação: Translacional, ribossomal e biogênese; Transcrição; Replicação e reparo de DNA.;

c. Processo celular: Divisão celular e partição cromossômica; Biogênese do envelope celular; Motilidade celular; Modificação pós-tradução; Transporte de íons inorgânicos;

d. Metabolismo: Produção e conversão de energia; Transporte e metabolismo de aminoácidos; Transporte e metabolismo de nucleotídeos; Transporte e metabolismo de carboidratos; Metabolismo lipídico.

4.3 Análise do papel das proteínas quanto a virulência da bactéria

Analisar o papel das proteínas quanto a virulência é de extrema importância, pois implica-se com a descoberta de novos fatores associados à virulência, o quão potencial patogênico a proteína pode oferecer ao organismo quando secretada. Com isso, podem ser a base para o desenvolvimento de novos medicamentos contra patógenos, e também ajudar no entendimento dos complexos mecanismos relacionados a virulência no mesmo.

Virulentpred

A partir do uso do algoritmo Virulentpred (<http://203.92.44.117/virulent/submit.html>), será possível realizar a predição do papel das proteínas em relação a virulência da bactéria, na qual a ferramenta aponta um índice numeral de virulência de acordo com a proteína processada.

Resultados que apontassem um threshold maior que zero (>0) poderia avaliar a proteína como associada a fatores de virulência.

4.4 Avaliação do potencial antigênico das proteínas

A metodologia aplicada para a avaliação da capacidade de antigenicidade das proteínas selecionadas, se baseia na exploração da ferramenta Vaxijen V 2.0, um algoritmo de bioinformática que realiza a previsão de antígenos protetores e vacinas de subunidades.

Vaxijen v 2.0

Esta ferramenta de predição (<http://www.ddg-pharmfac.net/vaxijen/VaxiJen/VaxiJen.html>), através da sequência de aminoácidos da proteína, realiza a previsão de antígenos protetores e vacinas de subunidades, dispensando a necessidade de cultivo bacteriano e também apresentando vantagem na especificação de antigenicidade pelo método da classificação dos antígenos com base em suas propriedades físico-químicas, visto que comparando com a abordagem de alinhamento de sequência para a identificação de antígenos se torna uma grande melhoria, partindo do pressuposto que nem todas as proteínas apresentam similaridade de sequência notória.

Resultados que apontassem um limite maior ou igual que 0,4 ($\geq 0,4$) poderia avaliar a proteína como antigênica.

5. RESULTADOS

A realização da exploração das diferentes ferramentas de bioinformática possibilitou os seguintes resultados:

5.1 Proteínas selecionadas

Através de revisões bibliográficas dos artigos já citados foram selecionadas para o estudo as proteínas: Papalisina 1, Termolisina, Lipoproteína P, Protease dependente de zinco -tldD e Metionina aminopeptidase -Map. Maiores detalhes sobre estas proteínas encontram-se descritos no Quadro 2.

Quadro 2: Resultado das proteínas selecionadas

PROTEÍNA	CÓDIGO / GENE	TAMANHO Nº de aminoácidos	MASSA (kDa)
Papalisina 1	EKO25673	480	51,83
Termolisina	LIC10715	800	90,33
Lipoproteína P	LIC10508	203	22,99
Protease dependente de zinco (tldD)	AAN49383.1	461	49,92
Metionina aminopeptidase (MAp)	AAN49656.2	240	26,59

5.2 Predição de secreção das proteínas-alvo

A análise realizada pela ferramenta Signal P 5.0 apontou positivo para a existência de peptídeo sinal para as proteínas Papalisina 1, Termolisina e Lipoproteína P, já as demais proteínas obtiveram resultado negativo para a existência do mesmo.

A ferramenta TMHMM result apontou a existência de uma região transmembranar somente na proteína Termolisina.

O algoritmo Psort avaliou a localização celular da proteína Papalisina 1 como desconhecida, a Termolisina como extracelular e as demais proteínas apresentaram localização citoplasmática por essa ferramenta.

O Phobius avaliou as proteínas Papalisina 1, Termolisina e Lipoproteína P como secretadas, porém as demais proteínas foram dadas como não-secretadas.

A ferramenta PSLpred apontou as proteínas Papalisina 1 e Termolisina como extracelulares, já as outras proteínas foram apontadas como citoplasmáticas.

O algoritmo Secretome 2.0 apontou as proteínas Papalisina 1 e Termolisina como secretadas, e as outras como não-secretadas.

Já a ferramenta Prediction Protein apontou que as proteínas Papalisina 1, Termolisina e Lipoproteína P são secretadas, e as demais foram apontadas como citoplasmáticas.

Ao final das análises, consideramos como secretadas: a proteína Papalisina 1, predita como secretada por 6 algoritmos, a proteína Termolisina, apontada como secretada por também 6 ferramentas, e a proteína Lipoproteína P, predita como secretada por 4 ferramentas. As proteínas protease dependente de zinco tldD e MAp foram consideradas não-secretadas por 6 ferramentas (Quadro 3).

Quadro 3: Resultados da predição de secreção analisados por sete ferramentas de bioinformática para verificação de requisitos que confirme ou não a secreção das proteínas.

	Signal P 5.0	TMHMM result	Psort	Phobius	PSLpred	Secretome 2.0	Prediction protein	Resultado
Papalisina 1 EKO25673	Possui peptídeo sinal	Não possui região transmembranar	Desconhecida	Secretada	Extracelular 53,10%	Secretada 0.9350	Secretada 87	Secretada
Termolisina LIC10715	Possui peptídeo sinal	Possui região transmembranar	Extracelular 9,72	Secretada	Extracelular 90,20%	Secretada 0.9082	Secretada 14	Secretada
Lipoproteína P LIC10508	Possui peptídeo sinal	Não possui região transmembranar	Citoplasmática 8,96	Secretada	Citoplasmática 68,10%	Não secretada 0.0860	Secretada 40	Secretada
Protease dependente de zinco (tldD) AAN49383.1	Não Possui peptídeo sinal	Não possui região transmembranar	Citoplasmática 9,97	Não secretada	Citoplasmática 71,10%	Não secretada 0.1446	Citoplasma 7	Não secretada
Metionina aminopeptidase (MAp) AAN49656.2	Não possui peptídeo sinal	Não possui região transmembranar	Citoplasmática 9,97	Não secretada	Citoplasmática 68,10%	Não secretada 0.1685	Citoplasma 33	Não secretada

Predição de secreção das proteínas quanto a função

A ferramenta VICMpred indicou que as proteínas Papalisina 1, Termolisina, Lipoproteína P e Protease dependente de zinco tldD exercem função de auxiliadoras no processo celular. Na análise de fatores virulentos, a proteína Termolisina foi associada à virulência, enquanto a proteína MAP apresentou função ligada ao metabolismo (Quadro 4).

Quadro 4: Resultados da análise de predição de secreção quanto a função das proteínas

Proteína	Função	Score
Papalisina 1 EKO25673	Processo Celular	2,6756765
Termolisina LIC10715	Processo Celular	1,3259607
	Fatores Virulentos	1,8027951
Lipoproteína P LIC10508	Processo Celular	1,0034036
Protease dependente de zinco tldD AAN49383.1	Processo Celular	1,1282831
Metionina aminopeptidase (Map) AAN49656.2	Metabolismo	2,450067

5.3 Análise do papel das proteínas quanto à virulência da bactéria

A ferramenta Virulentpred identificou possíveis fatores de virulência das proteínas Papalisina 1, Termolisina e Lipoproteína P, que são consideradas por esta ferramenta, portanto, proteínas associadas à virulência (Quadro 5).

Quadro 5: Resultados da análise de virulência das proteínas

Proteína	Virulentpred
Papalisina 1 EKO25673	1,1042 Associado à virulência
Termolisina LIC10715	1,0225 Associado à virulência
Lipoproteína P LIC10508	0,7709 Associado à virulência

5.4 Avaliação do potencial antigênico das proteínas

Em sua análise de antigenicidade, a ferramenta Vaxijen V 2.0 apontou as proteínas Papalisina 1, Termolisina e Lipoproteína P como portadoras de potencial antigênico (Quadro 6).

Quadro 6: Resultados da análise de antigenicidade das proteínas

Proteína	Vaxijen V 2.0
Papalisina 1 EKO25673	0,5531 Antigênico
Termolisina LIC10715	0,5668 Antigênico
Lipoproteína P LIC10508	0,6307 Antigênico

6. CONCLUSÃO

Concluimos que ao final do processo de execução das ferramentas que realizam a predição do potencial de secreção das proteínas alvo, foram consideradas como secretadas pela análise, as proteínas: Papalisina 1 e Termolisina. Apesar da proteína Lipoproteína P ser localizada no citoplasma, como indica as ferramentas Psort e PSLpred, ela possui peptídeo sinal e é considerada secretada pela ferramenta Prediction protein, portanto acredita-se que se trata de uma proteína do tipo *moonlighting*, na qual a mesma pode desempenhar funções diversas funções. A maioria das ferramentas indicou as proteínas protease dependente de zinco tldD e MAp como não-secretadas.

As proteínas Papalisina 1, Termolisina e Lipoproteína P, foram todas associadas a virulência e antigenicidade.

Diante dos resultados do presente estudo, podemos concluir que o fato das proteínas que foram selecionadas terem sido consideradas como secretadas e antigênicas, pode contribuir diretamente para futuras pesquisas com o objetivo de gerar um melhor diagnóstico para leptospirose, e assim solucionar os entraves presentes nos diagnósticos atuais.

8. REFERÊNCIAS BIBLIOGRÁFICAS

AMAMURA, Thais A. et al. Pathogenic *Leptospira* secreted proteases target the membrane attack complex: a potential role for thermolysin in complement inhibition. **Frontiers in microbiology**, v. 8, p. 958, 2017.

BIDMOS, Fadil A. et al. Bacterial Vaccine Antigen Discovery in the Reverse Vaccinology 2.0 era: Progress and Challenges. **Frontiers in immunology**, v. 9, p. 2315, 2018.
<https://doi.org/10.3389/fimmu.2018.02315>.

BRASIL. MINISTÉRIO DA SAÚDE. Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis. **Leptospirose: diagnóstico e manejo clínico**. 2014.

BRASIL. MINISTÉRIO DA SAÚDE. Secretaria de Vigilância em Saúde. **Guia Leptospirose: Diagnóstico e Manejo Clínico**. 2009.

CHAVES, Lana Bitencourt et al. Plasmodium vivax ookinete surface protein (Pvs25) is highly conserved among field isolates from five different regions of the Brazilian Amazon. **Infection, Genetics and Evolution**, v. 73, p. 287-294, 2019.

CONTE, Fernando de Paiva et al. Identification and validation of specific B-cell epitopes of hantaviruses associated to hemorrhagic fever and renal syndrome. **PLoS neglected tropical diseases**, v. 13, n. 12, e0007915, 2019.
<https://doi.org/10.1371/journal.pntd.0007915>.

COSTA, Federico et al. Global Morbidity and Mortality of Leptospirosis: A Systematic Review. **PLOS Neglected Tropical**, v. 9, n. 9, p. e0003898, 2015.
<https://doi.org/10.1371/journal.pntd.0003898>.

DE LA FUENTE, José et al. Strategies for new and improved vaccines against ticks and tick-borne diseases. **Parasite immunology**, v. 38, n. 12, p. 754-769, 2016.

<https://doi.org/10.1111/pim.12339>

DELLAGOSTIN, Odir A. et al. Reverse vaccinology: an approach for identifying leptospiral vaccine candidates. **International Journal of Molecular Sciences**, v. 18, n. 1, p. 158, 2017.

<https://doi.org/10.3390/ijms18010158>.

DE OLIVEIRA, Denise Santos Correia, et al. Modelo produtivo para a leptospirose. **Revista de Patologia Tropical/Journal of Tropical Pathology**, v. 38, n. 1, p. 17-26, 2009.

<https://doi.org/10.5216/rpt.v38i1.6205>.

ESHGHI, Azad et al. Pathogenic *Leptospira interrogans* exoproteins are primarily involved in heterotrophic processes. **Infection and immunity**, v. 83, n. 8, p. 3061-3073, 2015.

GENOVEZ, Margareth Elide. Leptospirose: uma doença de ocorrência além da época das chuvas. **Biologico**, v. 71, n. 1, p. 1-3, 2009.

KOENS, Lyndon; LAUGA, Eric. The passive diffusion of *Leptospira interrogans*. **Physical biology**, v. 11, n. 6, p. 066008, 2014.

MARINHO, Marcia. Leptospirose: fatores epidemiológicos, fisiopatológicos e imunopatogênicos. **Veterinária e Zootecnia**, v. 15, n. 3, p. 428-434, 2012.

MICHALIK, Marcin et al. Reverse vaccinology: the pathway from genomes and epitope predictions to tailored recombinant vaccines. In: **Vaccine Design**. Humana Press, New York, v. 1, p. 87-106, 2016.

PELLISSARI, Daniele Maria et al. Revisão sistemática dos fatores associados à leptospirose no Brasil, 2000-2009. **Epidemiologia e Serviços da Saúde**, v. 20, n. 4, p. 565-574, 2011.

<https://doi.org/10.5123/S1679-49742011000400016>

PEREIRA, Igor Alcântara et al. Leptospirose em fase aguda evoluindo com síndrome de weil e seu frágil diagnóstico sorológico: Relato de um Caso. **Rev Med Minas Gerais**, v. 2019, n. 29, 2020.

RAPPUOLI, Rino et al. Reverse vaccinology 2.0: Human immunology instructs vaccine antigen design. **Journal of Experimental Medicine**, v. 213, n. 4, p. 469-481, 2016.

<https://doi.org/10.1084/jem.20151960>.

RAPPUOLI, Rino et al. Vaccines, new opportunities for a new society. **Proceedings of the National Academy of Sciences**, v. 111, n. 34, p. 12288-12293, 2014.

RAPPUOLI, Rino. Reverse vaccinology. **Current opinion in microbiology**, v. 3, n. 5, p. 445-450, 2000.

RAPPUOLI, Rino. Vaccines, emerging viruses, and how to avoid disaster. **BMC biology**, v. 12, n. 1, p. 1-2, 2014.

RODRIGUES-DA-SILVA, Rodrigo Nunes et al. Plasmodium vivax cell-traversal protein for ookinetes and sporozoites: naturally acquired humoral immune response and B-cell epitope mapping in brazilian amazon inhabitants. **Frontiers in immunology**, v. 8, p. 77, 2017.

<https://doi.org/10.3389/fimmu.2017.00077>.

SIMÕES, Luciana Senna et al. Leptospirose–Revisão. **PubVet**, v. 10, p. 111-189, 2015.

VINCENT, Antony T. et al. Revisiting the taxonomy and evolution of pathogenicity of the genus *Leptospira* through the prism of genomics. **PLoS neglected tropical diseases**, v. 13, n. 5, p. e0007270, 2019.

ZENG, Lingbing et al. Extracellular proteome analysis of *Leptospira interrogans* serovar Lai. **Omics: a journal of integrative biology**, v. 17, n. 10, p. 527-535, 2013.

9. ANEXO 1 – Sisgen – Cadastro Nº A041B9C



Ministério do Meio Ambiente
CONSELHO DE GESTÃO DO PATRIMÔNIO GENÉTICO
SISTEMA NACIONAL DE GESTÃO DO PATRIMÔNIO GENÉTICO E DO CONHECIMENTO TRADICIONAL ASSOCIADO

Comprovante de Cadastro de Acesso

Cadastro nº A041B9C

A atividade de acesso ao Patrimônio Genético, nos termos abaixo resumida, foi cadastrada no SisGen, em atendimento ao previsto na Lei nº 13.123/2015 e seus regulamentos.

Número do cadastro: **A041B9C**
Usuário: **Fiocruz**
CPF/CNPJ: **33.781.055/0001-35**
Objeto do Acesso: **Patrimônio Genético**
Finalidade do Acesso: **Pesquisa e Desenvolvimento Tecnológico**

Espécie

Leptospira interrogans

Título da Atividade: **Estudos in silico de proteínas secretadas por Leptospira sp. patogênicas:
Uma nova abordagem para identificação de alvos vacinais e diagnósticos.**

Equipe

Rodrigo Nunes Rodrigues da Silva	Fiocruz
Fernanda de Moraes Mais	Fiocruz
Silvia da Silva Fontes	Fiocruz
Fernando de Paiva Conte	Fiocruz
Larissa Emiliano de Araujo	Escola Politécnica de Saúde Joaquim Venâncio

Data do Cadastro: **27/02/2020 11:03:54**
Situação do Cadastro: **Concluído**

Conselho de Gestão do Patrimônio Genético
Situação cadastral conforme consulta ao SisGen em **10:39** de **06/05/2020**.



SISTEMA NACIONAL DE GESTÃO
DO PATRIMÔNIO GENÉTICO
E DO CONHECIMENTO TRADICIONAL
ASSOCIADO - **SISGEN**