

GABRIELA CAMARGO PACHER DE MOURA

**Diagnóstico de Leishmaniose Cutânea usando Espectroscopia FTIR e Aprendizagem de  
Máquina: um estudo em modelo animal**

**CAMPO GRANDE**

**2023**

GABRIELA CAMARGO PACHER DE MOURA

**Diagnóstico de Leishmaniose Cutânea usando Espectroscopia FTIR e Aprendizagem de Máquina: um estudo em modelo animal**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas da Universidade Federal de Mato Grosso do Sul como um dos requisitos para obtenção do título de Mestre em Ciências Farmacêuticas.

Orientadora: Prof<sup>ª</sup>. Dr<sup>ª</sup>. Carla Cardozo

Co-orientador: Prof. Dr. Cícero Cena

**CAMPO GRANDE**

**2023**

*Dedico ao meu maior apoiador: meu marido.*

## AGRADECIMENTOS

A Deus, que merece toda honra e louvor, para todo o sempre.

Todo erro e falha desta dissertação devem ser concedidos a mim. Os méritos e congratulações, todavia, deverão ser divididos com todos aqueles que me auxiliaram, apoiaram e seguraram as pontas quando eu me tornei mãe. Nas diversas vezes que pensei em desistir, pude sentir-me inspirada olhando para a caminhada destas pessoas.

Ao meu marido, todo agradecimento que existe. Aos olhos dele, fui perfeita em tudo, mesmo caminhando aos trancos e barrancos. Ao meu marido e à minha filha: **nunca foram peso, e sim combustível.**

À minha orientadora Carla, por ter sido paciente, parceira e carinhosa. Não mediu esforços para me ajudar. Carregarei, para sempre, todos os ensinamentos que ganhei. Se eu pudesse, a retribuiria com tudo e mais um pouco.

Ao meu coorientador Cícero, por ter me ensinado a mesma coisa um milhão de vezes. Mesmo eu, uma *loser* para física, consegui aprender o básico graças à paciência dele.

À Natália, por ter sido minhas mãos e pés quando não pude processar algumas amostras. Por ter segurado as pontas quando eu não pude, nem sequer, amarrá-las.

À Aline e Jucelei, por terem sido colegas de trabalho. Pessoas que me impulsionaram.

Aos meus pais que, mesmo na distância, torceram por mim.

À minha irmã que, mesmo na distância, me ajudou de várias formas.

Aos meus amigos mais próximos, aqueles que sempre me perguntaram como estava o mestrado e, mesmo sem entender sobre o projeto e/ou burocracias, me davam força para continuar.

A Raquel, mais que amiga, babá da Laura por alguns dias. Quantas vezes cuidou da minha filha para que eu pudesse cuidar do mestrado.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) e a Coordenação de Aperfeiçoamento de Pessoal de Ensino Superior (CAPES) pelo apoio financeiro.

Ao Programa de Pós-graduação em Farmácia e aos professores de cada disciplina que finalizei.

Ao Hospital Universitário pela parceria.

Ao Laboratório de Parasitologia Humana da Universidade Federal de Mato Grosso do Sul, por ter sido essencial na minha formação.

A todos que estiveram comigo: obrigada!

## RESUMO

A leishmaniose cutânea (LC) é uma doença polimórfica e espectral da pele, causada por protozoários parasitas do gênero *Leishmania*. O diagnóstico da LC é complexo, pois os métodos usuais são pouco sensíveis, morosos e onerosos. A espectroscopia de infravermelho por transformada de Fourier (FTIR), com algoritmos de aprendizado de máquina (ML), tem sido explorada como uma alternativa para obter resultados rápidos e precisos para o diagnóstico de doenças. Além da boa precisão exibida em muitos estudos, as diferenças espectrais entre os grupos infectados e não infectados são sutis para serem notadas. Uma vez que a heterogeneidade do conjunto de amostras (sexo, idade, dieta, etc) geralmente carrega grande variação de dados e dificulta uma compreensão mais detalhada das características espectrais e respostas imunes, neste estudo, exploramos esta metodologia inovadora para diagnóstico de LC em um estudo de modelo animal. Para isso, soro sanguíneo, lesões cutâneas e amostras dos linfonodos poplíteos drenantes foram obtidos de camundongos BALB/C infectados experimentalmente com *Leishmania (Leishmania) amazonensis*. Algoritmos FTIR e ML apresentaram precisão razoavelmente boa, podendo diferenciar entre amostras infectadas (grupo LC) e não infectadas (grupo controle). A melhor precisão geral (88,9%) foi obtida usando a Análise de Componentes Principais (PCA) e a Máquina de Vetor de Suporte (SVM) na faixa de 4000-700  $\text{cm}^{-1}$  para amostras de soro sanguíneo. A acurácia obtida para lesões cutâneas e amostras de linfonodos poplíteos foi em torno de 77,8%. Os resultados sugerem que um bom desempenho dos modelos de predição está mais associado à análise dos dados do que diferenças na composição molecular das amostras.

**Palavras-Chave:** Espectroscopia de infravermelho; leishmaniose cutânea experimental; *Leishmania amazonensis*; análise multivariada; fotodiagnóstico; biofluidos.

## ABSTRACT

Cutaneous Leishmaniasis (CL) is a polymorphic and spectral skin disease caused by *Leishmania* spp protozoan parasites. CL diagnosis is difficult, as usual methods are time-consuming, expensive, and not very sensitive. Fourier transform infrared spectroscopy (FTIR) with machine learning (ML) algorithms has been explored as an alternative to achieve fast and accurate results for many disease diagnoses. Besides the good accuracy exhibited in many studies, the spectral differences between infected and non-infected groups are too subtle to be noticed. Since the sample set inhomogeneity (sex, age, diet, etc) usually carries great data variance and hinders a more detailed understanding of the spectral characteristics and immune responses, in this study, we explore this innovative methodology for CL diagnosis in an animal model study. Blood serum, skin lesions, and draining popliteal lymph node samples were obtained from BALB/C mice experimentally infected with *Leishmania* (*Leishmania*) *amazonensis*. FTIR and ML algorithms with reasonably good accuracy could differentiate between infected (CL group) and non-infected (control group) samples. The best overall accuracy (88.9%) was obtained using Principal Component Analysis (PCA) and Support Vector Machine (SVM) in the 4000-700 cm<sup>-1</sup> range for blood serum samples. The accuracy obtained for skin lesions and popliteal lymph node samples was around 77.8%. The results suggest that a good performance of the prediction models is more associated with the data analysis than molecular composition differences in the samples.

**Keywords:** Infrared spectroscopy; experimental cutaneous leishmaniasis; *Leishmania amazonensis*; multivariate analysis; photodiagnosis; biofluids

## LISTA DE FIGURAS

<b>Figura 1.</b> Lesão típica de LC. ....	02
<b>Figura 2.</b> Lesão típica de LM. ....	02
<b>Figura 3.</b> Esplenomegalia e hepatomegalia características da LV. ....	03
<b>Figura 4.</b> Amastigota: forma intracelular – apontada pela seta – encontrada no interior da célula hospedeira (macrófago).....	04
<b>Figura 5.</b> Promastigota: forma com flagelo livre encontrada no inseto vetor.. ....	04
<b>Figura 6.</b> Distribuição das principais espécies de Leishmania, responsáveis pela transmissão da LC, no Brasil.....	06



## SUMÁRIO

### CAPÍTULO 1 – REVISÃO DE LITERATURA

<b>1. REVISÃO DE LITERATURA</b> .....	.01
1.1 Aspectos gerais das leishmanioses .....	01
1.2 Agente etiológico, vetor e transmissão.....	04
1.3 Epidemiologia das leishmanioses.....	05
1.4 Leishmaniose cutânea.....	06
1.5 Diagnóstico da leishmaniose cutânea.....	07
1.5.1 Diagnóstico clínico-epidemiológico.....	08
1.5.2 Métodos parasitológicos.....	08
1.5.3 Métodos imunológicos.....	09
1.5.4 Métodos moleculares.....	10
1.6 Espectroscopia de infravermelho por transformada de Fourier e aprendizagem de máquina.....	12
1.7 Leishmaniose cutânea experimental.....	13
<b>REFERÊNCIAS</b> .....	15

### CAPÍTULO 2 – ARTIGO

Artigo .....	29
--------------	----

<b>ANEXOS</b> .....	46
---------------------	----

# 1. REVISÃO DE LITERATURA

## 1.1 Aspectos gerais das leishmanioses

Predominantes na África, Ásia e Américas, as doenças tropicais negligenciadas (DTN) são um grupo de doenças que ocorrem em climas tropicais e subtropicais e afetam os países pobres, onde o acesso à informação, ao saneamento adequado e aos cuidados básicos com a saúde são limitados. Além disso, as pessoas vivem próximas a vetores de agentes infecciosos (Engels; Zhou, 2020).

Mesmo que as DTN não possuam prioridade na agenda global de saúde, sabe-se que, na última década, elas foram alvo de preocupação porque causam diversas morbidades e deficiências físicas nos pacientes, levando-os à exclusão do trabalho, da educação e do bem-estar em geral (Engels; Zhou, 2020).

Como exemplo de doenças tropicais negligenciadas, as leishmanioses, ocasionadas por protozoários parasitas do gênero *Leishmania*, são classificadas como um complexo de doenças infecciosas, não contagiosas e com caráter zoonótico que podem acometer a pele, as mucosas e /ou os órgãos viscerais do paciente (BRASIL, 2017; BRASIL, 2022).

As leishmanioses possuem ampla distribuição geográfica e afetam principalmente populações que vivem em regiões com condições sanitárias precárias, constituindo um grave problema de saúde pública. Estima-se que são registrados 1,3 milhão de novos casos a cada ano (OPAS, 2021), porém é válido afirmar que esses números não retratam a realidade devido à subnotificação e às infecções inaparentes, levando a considerar que as verdadeiras taxas de incidência são maiores.

Atualmente, são conhecidas 21 espécies de *Leishmania* causadoras de leishmanioses em seres humanos (Brindha, 2021) e, devido a essa ampla variedade de espécies e à resposta imunológica do hospedeiro, pode-se observar um largo espectro de manifestações clínicas, classificadas em: leishmaniose cutânea (LC), leishmaniose mucosa/mucocutânea (LM) e leishmaniose visceral (LV) (OPAS, 2021).

A LC, forma clínica mais frequente, é uma doença em que lesões se manifestam na pele do paciente (Figura 1). No Brasil, a LC é uma das afecções dermatológicas que merece mais atenção devido às consequências físicas, psicológicas e econômicas que ela acarreta na vida do paciente (BRASIL, 2017 e BRASIL, 2022), restringindo-os da vida profissional e social. Sobre as áreas endêmicas, dados mostram que a LC tem maior ocorrência no Afeganistão, Brasil, Irã, Iraque e Síria (OPS, 2021).

**Figura 1** - Lesão típica de LC.



**Fonte:** Portal Prefeitura de Florianópolis, 2014.

Já a leishmaniose mucocutânea ocorre quando os parasitas se disseminam para a mucosa oral, nasal, faríngea e/ou laríngea do indivíduo, provocando lesões nestas regiões. Na maioria dos casos, o parasita mais envolvido nessa forma clínica no Brasil é *Leishmania (Viannia) braziliensis*. A LM (Figura 2) raramente apresenta cura espontânea, então, se não tratada, pode progredir para perda extrema de peso, perda da funcionalidade do órgão acarretando, por exemplo, obstrução das vias aéreas, deformidades faciais, propiciando infecções oportunistas e, também, óbito (Bezemer et al., 2022).

**Figura 2** - Lesão típica de LM.



**Fonte:** Marie; Petri Jr., 2022.

Por mais que as leishmanioses cutânea e mucosa possuam baixa taxa de letalidade (BRASIL, 2017, BRASIL<sup>b</sup>, 2020; BRASIL, 2022), elas são, com base no indicador DALY (Disability Adjusted Life Years), responsáveis por 2,35 milhões de anos de vida perdidos por

incapacidade (Alvar, 2012; OPAS, 2019), uma vez que essas podem causar desfiguração física grave, isolando o paciente de atividades comuns do dia-a-dia.

Considerada a forma mais grave entre as leishmanioses, devido à sua alta incidência e alta letalidade (BRASIL, 2014, BRASIL<sup>c</sup>, 2020 e BRASIL, 2022), a LV, causada no Brasil pelo protozoário *Leishmania (Leishmania) infantum*, é a forma clínica em que os parasitas se espalham pela corrente sanguínea para órgãos viscerais como baço, fígado, linfonodos e medula óssea, podendo causar esplenomegalia e hepatomegalia (Figura 3).

**Figura 3** - Esplenomegalia e hepatomegalia características da LV.



**Fonte:** Jorge, 2018.

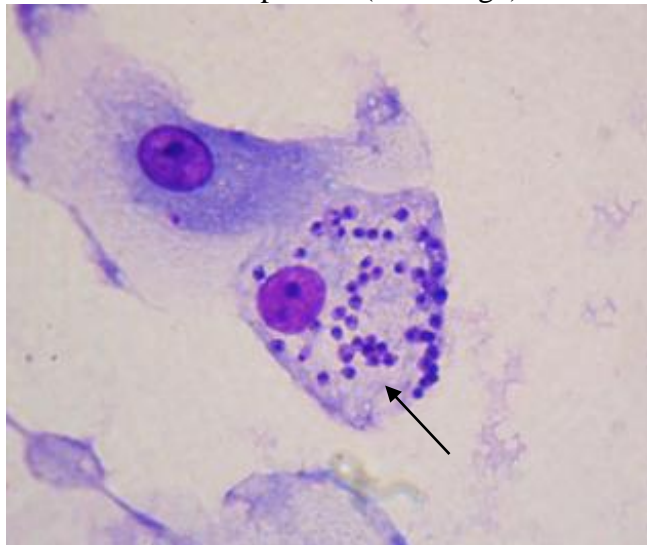
As últimas informações sobre as técnicas utilizadas para a diagnose das leishmanioses mostraram que, nas Américas, 81,1% dos casos de LC e LM foram diagnosticados por testes realizados em laboratórios, 12,7% por critérios clínico-epidemiológicos e 6% dos casos não possuem essa informação. Já sobre a LV, 89% das diagnoses ocorreram por testes laboratoriais e 11% por critério clínico-epidemiológico (OPAS, 2022).

Em relação ao tratamento das leishmanioses, observa-se que os fármacos de primeira escolha utilizados - antimoniais pentavalentes, como antimoniato de N-metil glucamina - de forma geral, podem se relacionar à resistência ao parasito e a efeitos colaterais ao paciente, devido à alta toxicidade (Lamotte et al., 2017). Além disso, a resposta ao medicamento estará relacionada com a espécie de *Leishmania* causadora da infecção, o que influencia na eficácia do processo terapêutico (Handler et al., 2015).

## 1.2 Agente etiológico, vetor e transmissão

Transmitido através da picada de insetos flebotomíneos fêmeas pertencentes à Ordem Diptera, *Leishmania* spp. são protozoários parasitos com duas formas evolutivas principais: amastigota (Figura 4), célula intracelular encontrada nos tecidos dos hospedeiros vertebrados, esférica, imóvel e sem flagelo livre, e promastigota (Figura 5), célula encontrada no tubo digestivo do vetor, alongada, com flagelo longo e livre e cinetoplasto anterior (STERVEDING, 2017). Esse protozoário consegue alternar seu ciclo de vida entre o flebotomíneo e os hospedeiros mamíferos (Muxel, 2019).

**Figura 4.** Amastigota: forma intracelular – apontada pela seta – encontrada no interior da célula hospedeira (macrófago).



Fonte: Lima, 2010.

**Figura 5.** Promastigota: forma com flagelo livre encontrada no inseto vetor.



Fonte: Lima, 2010.

O ciclo de vida desses protozoários inicia-se quando os vetores – insetos flebotomíneos fêmeas – se infectam com amastigotas durante o repasto sanguíneo do hospedeiro vertebrado infectado. As amastigotas, por consequência, transformam-se em promastigotas e multiplicam-se no trato digestório do vetor, tornando-se em formas infectantes que posteriormente serão transmitidas para o hospedeiro através da picada do inseto. Já no hospedeiro, na pele, as formas infectantes são fagocitadas por macrófagos e, no interior dessas células, ocorre a diferenciação em amastigotas e a multiplicação das mesmas até a ruptura celular, causando a disseminação da infecção (Burza et al., 2018).

### **1.3 Epidemiologia das leishmanioses**

Endêmicas em 102 países, anualmente, são registrados aproximadamente 1,3 milhão de novos casos de leishmanioses nos cinco continentes. Especificamente sobre a LV, 90% dos casos ocorrem no Brasil, Índia, Sudão, Sudão do Sul, Etiópia e Quênia. Já sobre a LC, 95% ocorrem nas Américas, Mediterrâneo e Centro e Médio Oriente da Ásia (OPAS, 2019). Sendo assim, O Brasil, juntamente com Bolívia, Peru, Afeganistão, Irã, Iraque e Síria, possui os maiores registros da LC e LM.

Especialmente nas Américas, nos últimos 20 anos, foram notificados à Organização Pan-Americana da Saúde (OPAS) aproximadamente 1 milhão de novos casos de LC e LM, com média de 50 mil por ano, e aproximadamente 70 mil casos novos de LV, com uma média anual de 2 mil casos (OPAS, 2022). Para as três manifestações clínicas, duas informações são relevantes: quando comparado com outros países das Américas, o Brasil continua sendo um dos países com maior número de casos e os pacientes mais afetados são do sexo masculino, entre 20 e 50 anos (OPAS, 2022).

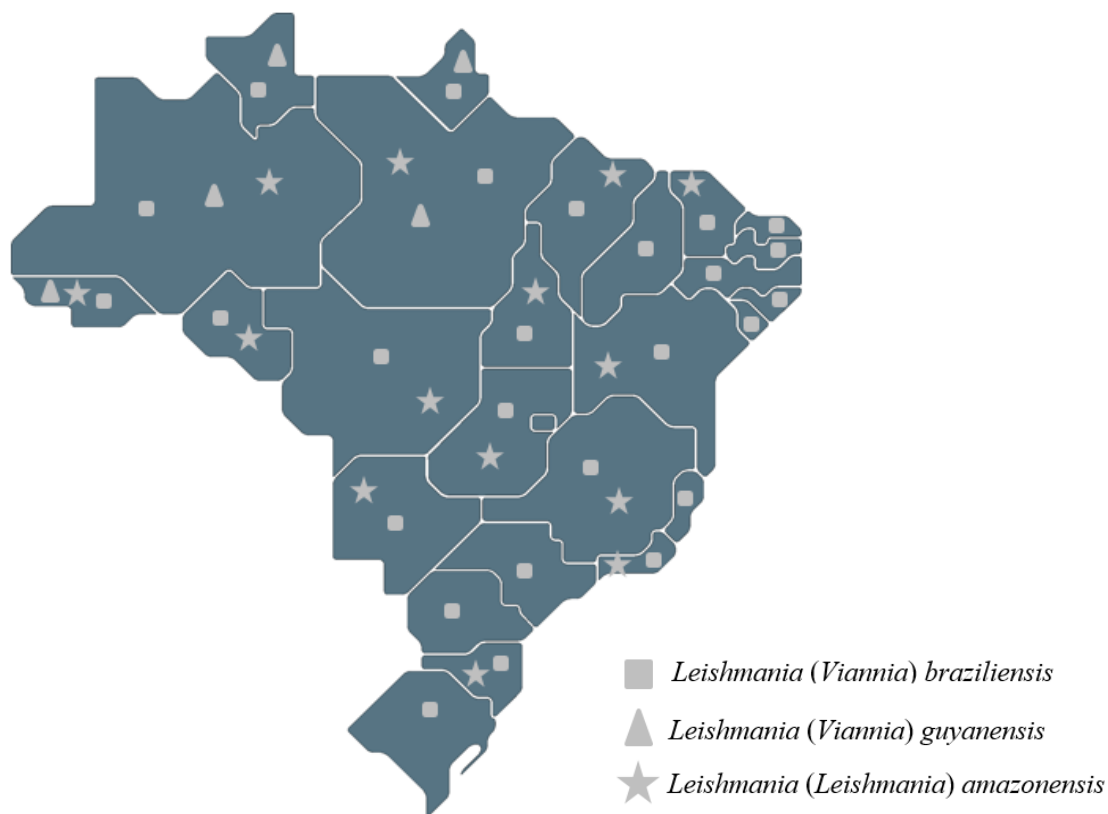
No território brasileiro, a LV vem sendo notificada, principalmente, nos municípios como Campo Grande (MS), Rio de Janeiro (RJ), Corumbá (MS), Belo Horizonte (MG), Palmas (TO), entre outros (BRASIL, 2014). Já a LC e LM, possuem os maiores registros nas regiões Norte, Nordeste e Centro-oeste (Mattos, 2023). Dados do site Informações de Saúde (TABNET) do Departamento do Sistema Único de Saúde (DATASUS), do Ministério da Saúde, indicam que no ano de 2022 foram confirmados 13.596 casos de LC no país, sendo que 112 casos foram notificados no estado de Mato Grosso do Sul (BRASIL<sup>a</sup>, 2023).

#### 1.4 Leishmaniose cutânea (LC)

A LC é a manifestação clínica mais comum entre as leishmanioses, e as características clínicas dependem da espécie de *Leishmania* em causa e da resposta imunológica do hospedeiro. Pacientes com leishmaniose cutânea apresentam lesões únicas ou múltiplas no local da picada do vetor, geralmente áreas expostas do corpo. Após a picada, o período de incubação varia de duas semanas a dois anos. Uma mácula e, posteriormente, pápula, precedem à lesão ulcerada (OPS, 2019, BRASIL, 2017; BRASIL, 2022). Esta, por sua vez, é caracterizada pela forma arredondada com base endurecida, bordas regulares e elevadas, fundo granuloso e cor rósea (BRASIL, 2017; BRASIL, 2022).

No Brasil, as principais espécies encontradas são *Leishmania (Viannia) braziliensis*, *Leishmania (Viannia) guyanensis* e *Leishmania (Leishmania) amazonensis* (Figura 6) (BRASIL, 2017 e BRASIL, 2022). Vale lembrar que não há transmissão de pessoa a pessoa e que o período de incubação da doença no ser humano é, em média, de dois a três meses, podendo variar de duas semanas a dois anos.

**Figura 6.** Distribuição das principais espécies de *Leishmania*, responsáveis pela transmissão da LC, no Brasil.



**Fonte:** Adaptado de Brasil, 2017.

Considerada uma das seis mais importantes doenças infecciosas do mundo, a LC registra, aproximadamente, 1,5 milhão de novos casos por ano. Nas Américas, de acordo com a OPAS, 17 países endêmicos registraram uma média anual de 55.317 novos casos por ano de LC entre 2001 e 2017. Em 2020, o Brasil registrou o maior número de casos de LC, sendo as principais áreas endêmicas as regiões Norte, Centro-Oeste e Nordeste (OPAS, 2021). Entre essas três regiões, a região Centro-oeste apresenta mais registros que a região Nordeste, uma vez que possui 15,4 casos a cada 100.000 habitantes (Mattos, 2023).

Em relação à resposta imunológica, sabe-se que a infecção pode desencadear um perfil de citocinas Th1 ou Th2. Também, pode causar um aumento de células T e de citocinas, como IFN- $\gamma$ , IL-4, IL-3, IL-10 e IL-12. Sendo assim, pacientes com LC desenvolvem uma resposta do tipo Th1 para que o controle da infecção seja feito e para que ocorra a eliminação dos parasitas a partir do aumento das atividades dos macrófagos parasitados (Reis et al., 2007; Cardoso et al., 2020).

Especificamente, para os casos de LC, o tratamento recomendado pela OMS se baseia na utilização de antimoniais pentavalentes (BRASIL, 2017a), sendo o antimoniato de meglumina a opção terapêutica mais administrada no Brasil. Além disso, em outros casos, pode-se utilizar a pentamidina, a anfotericina B, a paromomicina e a miltefosina (Machado-Silva et al., 2014).

### **1.5 Diagnóstico da leishmaniose cutânea**

Para o diagnóstico da LC, há a possibilidade de se utilizar os métodos clínicos, parasitológicos, imunológicos/sorológicos e/ou moleculares (BRASIL, 2017; BRASIL, 2022). Mesmo com essa variedade de técnicas, é difícil definir o melhor método, visto que eles ou são pouco sensíveis ou são dispendiosos em tempo/recursos. Além disso, os testes podem apresentar reações cruzadas com outros protozoários e sua sensibilidade pode mudar de acordo com a forma clínica da doença.

Os métodos atuais têm precisão limitada, o que exige o desenvolvimento de novos testes e estratégias. Recentemente, foram identificadas algumas prioridades, como o desenvolvimento de testes mais acessíveis e precisos, incluindo a detecção da infecção assintomática, o desenvolvimento de testes do tipo “*point-of-care*” e o encontro de biomarcadores que possam ser usados para prever o monitoramento eficaz de indivíduos submetidos a regimes de tratamento ou recaídas (Burza et al., 2018).



A carência de um método eficaz e rápido dificulta o tratamento adequado e precoce, diminuindo as chances no sucesso terapêutico. Por esse motivo, são necessários a pesquisa e o desenvolvimento de métodos diagnósticos não invasivos, rápidos, com baixo custo e alta sensibilidade para a LC, especialmente em áreas remotas onde a parasitose é endêmica.

Para avaliar a eficácia dos métodos diagnósticos, são observados dois aspectos: sensibilidade e especificidade. A sensibilidade corresponde ao percentual de resultados positivos dos indivíduos que possuem a doença. Já a especificidade diz respeito ao percentual de resultados negativos nos indivíduos que não têm a doença. No caso da LC, vale lembrar que a combinação de diferentes métodos diagnósticos aumentam a eficácia destes dois parâmetros citados acima (Vargas; Maria, 2023).

### **1.5.1 Diagnóstico clínico-epidemiológico**

O exame clínico é o primeiro método utilizado quando há a suspeita de LC. Essa primazia se deve ao fato de que o médico realiza a anamnese do paciente e avalia os sintomas apresentados, sendo as lesões cutâneas os mais característicos. Além disso, o levantamento de dados epidemiológicos, como a região onde o paciente vive e/ou trabalha, contribui para a formulação da hipótese diagnóstica (BRASIL, 2017).

Esta observação inicial pode auxiliar no diagnóstico da leishmaniose e, então, proporcionar um bom encaminhamento deste paciente para os cuidados e exames complementares necessários (BRASIL, 2022). A vantagem deste exame é seu caráter não invasivo; por outro lado, esta avaliação não é suficiente para um diagnóstico assertivo e definitivo, visto que a aparência das lesões varia de acordo com a carga parasitária, com a resposta imunológica do paciente, com a espécie causadora e, também, com o tempo de negligência da doença por parte do infectado. Ademais, os sintomas se assemelham com outras infecções dérmicas, como, por exemplo, sífilis, hanseníase e tuberculose (BRASIL, 2017; Masmoudi, 2013).

### **1.5.2 Métodos parasitológicos**

Os métodos parasitológicos ainda podem ser considerados como padrão-ouro e baseiam-se na observação do parasita, com a pesquisa das formas amastigotas, no fragmento da borda da lesão cutânea, deixando a diagnose mais assertiva. Estes testes se dividem em duas vertentes: o método parasitológico direto e o método parasitológico indireto. A sensibilidade destas técnicas varia de 15% a 90% (Sato, 2017) pois dependem de fatores externos – serão

citados a partir do próximo parágrafo –, porém, quando se combinam ambas técnicas, ela pode aumentar para mais de 85% (Elmahallawy et al., 2014; Sato 2017).

O exame direto é realizado com raspagem, punção ou biópsia da lesão, corada por Giemsa ou Leishman (BRASIL, 2017). É realizado quando uma amostra de tecido, geralmente obtida através de uma raspagem da lesão cutânea, é examinada ao microscópio para a detecção direta do parasita. Por mais que a visualização do protozoário represente um diagnóstico assertivo, a sensibilidade do método depende, principalmente, da qualidade da amostra e da quantidade de parasitas presentes na mesma, podendo ocorrer resultados falso negativos (Sato, 2017).

Já o teste parasitológico indireto é realizado a partir da cultura do parasita *in vitro* no meio ágar sangue Novy e McNeal modificado por Nicolle (NNN). Para este método, é necessário que uma amostra de tecido seja macerada e o material cultivado em meio apropriado para o crescimento dos parasitas, podendo levar semanas para a obtenção dos resultados. A vantagem é que esta técnica permite o isolamento e a identificação do parasita, porém, requer um tempo de incubação prolongado e tem menor sensibilidade em comparação com outros métodos (BRASIL, 2017).

Embora esses métodos sejam os mais utilizados para o diagnóstico da leishmaniose, há desvantagens que diminuem a eficácia dessas técnicas, como a baixa sensibilidade, o processamento correto da amostra, o conhecimento técnico necessário para a realização do procedimento e o caráter invasivo, visto que o processo de raspagem, biópsia e/ou punção é doloroso para o paciente (Thakur, 2020).

### **1.5.3 Métodos imunológicos**

Os métodos imunológicos se baseiam na resposta imunológica do hospedeiro infectado, podendo compreender os métodos sorológicos e os que detectam a reação de hipersensibilidade tardia aos antígenos parasitários. Com especificidade em torno de 75% (Fagundes, 2012; Gomes, 2014), os métodos imunológicos apresentam resultados satisfatórios, mensurando a resposta imunológica do sujeito infectado a partir da presença de anticorpos específicos, por exemplo (Basano; Camargo, 2004). Os testes sorológicos podem apresentar limitações de sensibilidade, uma vez que os anticorpos estão, na maioria dos casos, indetectáveis ou presentes em títulos baixos devido à resposta humoral deficiente (Elmahallawy et al., 2014). Também, em áreas endêmicas, a exposição prévia ao parasita é comum, podendo confundir a diagnose.

Um método imunológico que pode ser utilizado é o Ensaio de Imunoabsorção Enzimática (ELISA), que tem como objetivo detectar anticorpos IgG e IgM em pacientes infectados (Ryan et al., 2002). Este método pode apresentar alta sensibilidade e especificidade – 95% e 100%, respectivamente –, no entanto, estes resultados dependem da preparação e da pureza dos antígenos (Sato, 2017) e da infecção ser localizada ou não. Ademais, outra desvantagem é a necessidade de uma equipe experiente e tecnologias especializadas, o que impede seu uso em campo (Elmahallawy et al., 2014).

A Intradermorreação de Montenegro (IDRM), que representa o exame complementar mais sensível e prático para o diagnóstico da LC (Santos, 2021), consiste na inoculação dos antígenos de *Leishmania* spp. por via intradérmica do paciente, sendo este avaliado após 48 horas. Caso seja encontrado, no local da aplicação, uma placa endurecida com diâmetro maior de 5 milímetros, considera-se positivo para leishmaniose (BRASIL, 2017). A vantagem deste teste está, principalmente, na facilidade e praticidade para a execução e no baixo custo (Cerutti et al., 2017; Guedes et al., 2017). Porém, é necessário reconhecer alguns fatores que diminuem a sua eficácia, como: não diferencia infecções atuais das passadas, costuma ser negativo em pacientes imunossuprimidos, não distingue isoladamente doença de infecção (Andrade et al., 2005). Infelizmente, este exame não está mais disponível no Brasil devido a problemas relacionados a limitações na produção e distribuição das preparações antigênicas (Guedes et al., 2017).

#### **1.5.4 Métodos moleculares**

Os métodos moleculares, apesar de não disponíveis na rotina da maioria dos serviços de saúde, baseiam-se na detecção de sequências de DNA específicas do parasita, sendo a reação em cadeia da polimerase (PCR) e a PCR quantitativa em tempo real (qPCR) as principais técnicas utilizadas (Goto e Lindoso 2010; Molina et al. 2013; Sudarshan et al. 2015; Galluzzi et al., 2018). Esses métodos possuem maior sensibilidade – em torno de 100% (Silva et al, 2012; Saab et al., 2015; Sato, 2017) –, especificidade e rapidez para o diagnóstico da LC se comparados com os métodos parasitológicos e imunológicos, especialmente para amostras que apresentam baixa carga parasitária (Sato, 2017).

Na Reação em Cadeia da Polimerase (PCR), pode-se utilizar qualquer material biológico que contenha fragmentos de DNA de *Leishmania* spp. No entanto, esta técnica requer profissionais habilitados e alto investimento em estrutura. Sendo assim, o seu uso é limitado

por ser caro, de execução complexa e com alto risco de contaminação, ficando limitado em geral aos grandes centros e laboratórios de pesquisa.

Várias sequências do genoma de *Leishmania* podem ser utilizadas como alvos para amplificação e incluem genes para RNA ribossômico,  $\beta$ -tubulina, locus gp63, locus hsp70, proteinases de cisteína e minicírculos de DNA de cinetoplasto (kDNA) (Singh 2006; Reithinger; Dujardin 2007). Algumas amostras clínicas podem conter uma pequena quantidade de DNA do parasita, e o uso de sequências de alto número de cópias como alvo (por exemplo, minicírculo de kDNA, rDNA) pode garantir a sensibilidade adequada (Goto; Lindoso 2010; Molina et al. 2013; Sudarshan et al. 2015). Diferentes protocolos podem ser utilizados para todas as etapas da reação, como variações nas amostras, no pré-tratamento das amostras, na extração do DNA e na escolha dos *primers* (Galluzzi et al. 2018; De Brito et al., 2020).

Outra possibilidade é a utilização da técnica chamada Amplificação Isotérmica Mediada por Loop (LAMP). Ao contrário da PCR, a LAMP possibilita a amplificação de DNA sem a necessidade do termociclador. Este método já apresentou 95% de sensibilidade e 86% de especificidade em um estudo conduzido com 105 pacientes com LC (Kothalawala; Katunaweera, 2016). Porém, a desvantagem é que a sua sensibilidade é menor se comparada com a PCR e o parasitológico direto (Kothalawala; Katunaweera, 2016).

A Tabela 1 resume as principais vantagens e desvantagens dos métodos utilizados atualmente para o diagnóstico da LC.

**Tabela 1.** Resumo das principais vantagens e desvantagens dos métodos de diagnóstico para a LC.

<b>Método</b>	<b>Principal vantagem</b>	<b>Principal desvantagem</b>
Clínico	Não invasivo.	Os sintomas são semelhantes a outras infecções, não proporcionando uma diagnose assertiva.
Parasitológico	Visualização do parasita.	Invasivo e dependente de profissionais instruídos.

Sorológico	Prático e com baixo custo.	Baixa sensibilidade, principalmente em pacientes imunossuprimidos.
Molecular	Alta sensibilidade.	Alto custo e execução complexa.

Fonte: Autora, 2023.

## 1.6 Espectroscopia de infravermelho por transformada de Fourier (FTIR) e aprendizagem de máquina

Como descrito acima, os métodos usuais de diagnóstico não possuem alta sensibilidade, necessitam de expertise profissional e/ou são caros e demorados, dificultando a confirmação da doença e impedindo o tratamento adequado e precoce.

As tendências em pesquisas para o diagnóstico de doenças convergem para a busca de inovações que facilitem a obtenção de resultados precisos aliados ao uso de metodologias não invasivas, além de possibilitar um diagnóstico em tempo inferior ao de métodos tradicionais (Pepe et al., 2001; Etzioni et al., 2003; Garnett et al., 2012;). Neste cenário, o emprego de técnicas de espectroscopia óptica vem se destacando pelo potencial demonstrado no diagnóstico preciso e altamente seletivo de diversas doenças.

Os principais alvos desta técnica têm sido amostras de biofluidos como urina, plasma e soro sanguíneos (Alum et al., 2008; Psychogios et al., 2011; Garnett, 2012). A urina é comumente utilizada devido à alta concentração de compostos, sendo possível a identificação de xenobióticos provenientes de medicamentos e contaminações. Já o soro e o plasma fornecem informações sobre as perturbações do controle homeostático, como alterações nas proteínas, lipídios e peptídeos (Alum et al., 2008). Estes biofluidos possuem características essenciais para servirem como meio diagnóstico de fácil acesso para identificação dos biomarcadores de doenças (Bellisola; Sorio, 2012), partindo-se do princípio de que a doença distribuirá “informações” por todo o organismo (PSYCHOGIOS et al., 2011).

Assim, o emprego desta técnica no estudo de biofluidos para diagnóstico de doenças pode permitir a detecção de analitos importantes para um diagnóstico clínico em tempo real (Kumar et al., 2013), constituindo um método não invasivo, de fácil aplicabilidade e alta reprodutividade, e atingindo índices consideráveis de precisão, especificidade, sensibilidade de

baixa taxa de erros falsos positivos e verdadeiros negativos (Bellisola; Sorio, 2012; Dunn; Jegalian; Greenwald, 2010).

Na FTIR, um feixe de radiação infravermelho atinge a amostra, provocando deslocamentos atômicos e vibrações moleculares. Essas vibrações produzem espectros contendo informações químicas e biológicas das amostras (Cameron et al., 2018) por meio da análise direta dos espectros obtidos e, em combinação com a análise multivariada de dados, é possível separar grupos infectados de grupos não infectados.

Ao recorrer à literatura, observa-se que essas técnicas foram capazes de diagnosticar várias doenças infecciosas com alta precisão: a análise do espectro da saliva forneceu 100% de precisão geral para o diagnóstico de COVID-19 (Martinez-Cuazitl et al., 2021) e, também, precisão, sensibilidade e especificidade de 91,6%, 92,5% e 88,8%, respectivamente (Desai, 2020). Ainda utilizando saliva, outro estudo alcançou até 91% de precisão para a detecção de SARS-CoV-2 (Karas, 2023).

Em amostras de soro sanguíneo de pacientes com dengue, os resultados mostraram 89% de precisão (Naseer et al., 2010). A identificação de outros microrganismos mostrou-se promissora, como *Salmonella* sp (Cordovana et al., 2021) e *Staphylococcus aureus* (Guliev et al., 2020), com acurácia acima de 87% e 100% por meio de amostras ambientais e de sangue, respectivamente. A infecção parasitária de *Trypanosoma* sp., *Leishmania infantum* (Aguiar et al., 2018) e *Plasmodium* sp. (Heraud et al. 2019) foi identificada com precisão acima de 92%, analisando a cultura *in vitro* e o sangue dos pacientes, respectivamente.

Nosso grupo demonstrou o potencial uso da espectroscopia FTIR com aprendizagem de máquina para o diagnóstico de leishmaniose visceral canina usando soro sanguíneo com precisão acima de 85%, atribuído a alterações nas bandas vibracionais relacionadas a moléculas de proteína (Larios et al., 2021). Os algoritmos de ML também foram capazes de fornecer uma precisão geral de 91,67% para o diagnóstico de paracoccidiodomicose usando espectros FTIR de soro sanguíneo humano, quando as principais alterações vibracionais foram atribuídas a moléculas de proteínas e lipídios (Brito et al., 2022).

### **1.7. Leishmaniose cutânea experimental**

Estudos indicam que as manifestações clínicas e as consequências das leishmanioses estão relacionadas à natureza da resposta imunológica desenvolvida pelo hospedeiro (Nadere e Mcconville, 2007, Taheri et al., 2013, Muxel et al., 2019, Santos-Pereira et al., 2019 e Salari et al., 2020). A infecção por *Leishmania* inicia-se a partir da capacidade dos macrófagos,

células hospedeiras de *Leishmania*, em eliminar os parasitos. Imunologicamente, ocorre a indução da produção de citocinas do perfil Th1, regulando mecanismos envolvidos nas atividades anti-inflamatórias e anti-*Leishmania* (Muxel, 2019). Nos tecidos linfoides, células T são ativadas em células Th0, e posteriormente em células Th1, ao entrarem em contato com células dendríticas e macrófagos em ambiente rico em IL-12 (Siewe et al. 2016). Células T CD4<sup>+</sup> reconhecem os antígenos ligados ao Complexo Principal de Histocompatibilidade (MHC) e, então, produzem IL-12, promovendo a ativação das células T CD8<sup>+</sup>, citotóxicas. Tanto as células T CD8<sup>+</sup>, como as células T CD4<sup>+</sup>, produzem IFN- $\gamma$  que, por sua vez, ativa os macrófagos para eliminar os parasitas (Siewe et al. 2016).

No entanto, o protozoário pode vencer a ativação de macrófagos para estabelecer o processo infeccioso. A progressão da doença está, então, relacionada com o padrão de resposta imunológica Th2, juntamente com um aumento dos níveis de IL-10, IL-5, IL-4 e TGF- $\beta$  (Torres-Guerrero et al. 2017).

Os modelos experimentais murinos têm sido extremamente úteis para o estudo dos mecanismos imunológicos que medeiam a patogênese da leishmaniose e a resistência do hospedeiro, bem como para o estudo de novos fármacos antileishmaniais e métodos de diagnóstico. Nestes modelos, a progressão ou cura da doença podem ser influenciadas pelo perfil de expressão de citocinas (Th1 ou Th2) ou pela genética do hospedeiro (Taheri et al., 2013 e Aoki, 2019).

A literatura mostra que camundongos BALB/c são claramente suscetíveis à infecção por *L. (L.) amazonensis*, desenvolvendo lesões graves e progressivas que levam ao óbito. Esse fenótipo de não cura está associado a uma resposta Th2 específica do parasita, caracterizada pelo aumento da expressão de citocinas de macrófagos, como IL-4, IL-10 e TGF- $\beta$  (Loría-Cervera et al., 2014).

Desta forma, este modelo foi utilizado para a padronização da técnica FTIR associada à análise multivariada, com o intuito de discriminar as amostras de materiais biológicos de animais infectados ou não por *L. (L.) amazonensis*. Objetiva-se, portanto, com este trabalho, o emprego da FTIR para a detecção da LC a partir do estudo de biofluidos dos animais infectados, que, aliada a métodos estatísticos adequados, pode vir a se tornar um método de diagnóstico mais sensível, seletivo e preciso para esta enfermidade.

## REFERÊNCIAS

- AGUIAR, J. C. et al. Using FT-IR spectroscopy for the identification of the *T. cruzi*, *T. rangeli*, and the *L. chagasi* species. **Experimental Parasitology**, v. 192, p. 46–51, set.
- 2018.ALUM, M. F. et al. 4,4-Dimethyl-4-silapentane-1-ammonium trifluoroacetate (DSA), a promising universal internal standard for NMR-based metabolic profiling studies of biofluids, including blood plasma and serum. **Metabolomics**, v. 4, n. 2, p. 122–127, 3 fev. 2008.
- ANDRADE, B. B. et al. Métodos diagnósticos da Leishmaniose Tegumentar: fatos, falácias e perspectivas. **Gazeta Médica da Bahia**, Bahia, v. 75, n. 1, p. 75-82, set. 2005.
- AOKI, J. et al. Differential immune response modulation in early *Leishmania amazonensis* infection of BALB/c and C57BL/6 macrophages based on transcriptome profiles. **Scientific Reports**, v. 9, n. 1, dez. 2019.
- BAKER, M. J. et al. Developing and understanding biofluid vibrational spectroscopy: a critical review. **Chemical Society Reviews**, v. 45, n. 7, p. 1803-1818, 2015.
- BASANO, S. DE A.; CAMARGO, L. M. A. Leishmaniose tegumentar americana: histórico, epidemiologia e perspectivas de controle. **Revista Brasileira de Epidemiologia**, v. 7, n. 3, p. 328–337, 1 set. 2004.
- BELLISOLA, G.; SORIO, C. Infrared Spectroscopy and Microscopy in Cancer Research and Diagnosis. **American Journal of Cancer Research**, v. 2, n. 1, p. 1-21, jan. 2012.
- BEZEMER, J. M. et al. Clinical criteria for Mucosal Leishmaniasis diagnosis in rural South America: A systematic literature review. **PLOS Neglected Tropical Diseases**, v. 16, n. 8, ago. 2022.
- BRASIL, Ministério da Saúde. Guia de Vigilância em Saúde. **Biblioteca Virtual em Saúde**, 2022; 5ª edição. Disponível em: < <https://dive.sc.gov.br/phocadownload/guia-vigilancia/2022/GV5-5edicao-04-05-2022.pdf>>. Acesso em: 02 ago.2023.
- BRASIL, Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis. **Manual de vigilância da leishmaniose tegumentar**. [recurso eletrônico] / Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância das Doenças Transmissíveis. – Brasília: Ministério da Saúde, 2017.
- BRASIL. Ministério da Saúde, Secretaria de Vigilância em Saúde. **Manual de Vigilância da Leishmaniose Tegumentar Americana**. 1. ed. atual. Brasília: Editora do Ministério da Saúde, 2017. Acesso em: 05 mai.2022.
- BRASIL, Ministério da Saúde, Secretaria de Vigilância em Saúde. **Manual de Vigilância da Leishmaniose Visceral**. 1. ed. atual. Brasília: Editora do Ministério da Saúde, 2014. Acesso em: 05 mai.2022.



BRASIL<sup>a</sup>. Ministério da Saúde. Sistema de Informação de Agravos de Notificação (SINAN). **Casos de Leishmaniose Tegumentar no Brasil**. Disponível em: <http://tabnet.datasus.gov.br/cgi/tabcgi.exe?sinanet/cnv/ltabr.def>. Acesso em: 30 jun.2023.

BRASIL<sup>b</sup>. Ministério da Saúde – MS. Secretária de Vigilância em Saúde. **Situação epidemiológica da Leishmaniose Tegumentar, 2020**. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/l/lt/arquivos/lt-graficos-e-mapas.pdf> Acesso em: 02 ago.2023.

BRASIL<sup>c</sup>. Ministério da Saúde – MS. Secretária de Vigilância em Saúde. **Situação epidemiológica da Leishmaniose Visceral, 2020**. Disponível em: <https://www.gov.br/saude/pt-br/assuntos/saude-de-a-a-z/l/leishmaniose-visceral/arquivos/lv-graficos-e-mapas.pdf> Acesso em: 02 ago.2023.

BRINDHA, J.; BALAMURALI, M.; CHANDA, K. An Overview on the Therapeutics of Neglected Infectious Diseases—Leishmaniasis and Chagas Diseases. **Frontiers in Chemistry**, v. 9, 12 mar. 2021.

BRITO, E. C. A. et al. Paracoccidioidomycosis screening diagnosis by FTIR spectroscopy and multivariate analysis. **Photodiagnosis and Photodynamic Therapy**, v. 39, p. 102921–102921, 1 set. 2022

BURZA, S.; CROFT, S. L.; BOELAERT, M. Leishmaniasis. **The Lancet**, v. 392, n. 10151, p. 951–970, set. 2018.

CAMERON, J. M. et al. Biofluid spectroscopic disease diagnostics: A review on the processes and spectral impact of drying. **Journal of Biophotonics**, v. 11, n. 4, 5 mar. 2018.

CARDOSO, F. O. et al. (2020) Modulation of Cytokines and Extracellular Matrix Proteins Expression by *Leishmania amazonensis* in Susceptible and Resistant Mice. **Frontiers in Microbiology**, v. 11, 31 ago. 2020.

CERUTTI, P. et al. Métodos diagnósticos da leishmaniose tegumentar americana: uma revisão de literatura. **Revista de Patologia do Tocantins**, v. 4, n. 4, p. 55, 28 nov. 2017.

CORDOVANA M. et al., Classification of *Salmonella enterica* of the (para-)typhoid fever group by Fourier-Transform Infrared (FTIR) Spectroscopy. **Microorganism**, v. 9, n. 853, 2021.

DA SILVA, J. G. L. et al. Comparison among three polymerase chain reaction assays on detection of DNA from *Leishmania* in biological samples from patients with american cutaneous leishmaniasis, **Revista da Sociedade Brasileira de Medicina Tropical**, v. 45, n. 2, 2012.

DE BRITO, R. C. F. et al. Recent advances and new strategies in Leishmaniasis diagnosis. **Applied Microbiology and Biotechnology**, v. 104, n. 19, p. 8105–8116, 26 ago. 2020.

DESAI, S. et al. Raman spectroscopy-based detection of RNA viruses in saliva: A preliminary report. **Journal of Biophotonics**, v. 13, n. 10, 10 ago. 2020.

DUNN, B. K.; JEGALIAN, K.; GREENWALD, P. Biomarkers for Early Detection and as Surrogate Endpoints in Cancer Prevention Trials: Issues and Opportunities. **Recent results in cancer research**, p. 21–47, 1 jan. 2010.

ELMAHALLAWY, E. K. et al. Diagnosis of leishmaniasis. **The Journal of Infection in Developing Countries**, v. 8, n. 08, p. 961–972, 13 ago. 2014.

ENGELS, D., ZHOU X. N. Neglected tropical diseases: an effective global response to local poverty-related disease priorities. **Infect Dis Poverty**, v. 9, n. 1, 2020.

ETZIONI, R. et al. The case for early detection. **Nature Reviews Cancer**, v. 3, n. 4, p. 243–252, abr. 2003.

FAGUNDES, A. et al. Comparison between in vivo measurement of the Montenegro skin test and paper recording. **International Journal of Dermatology**, v. 51, n. 5, p. 618–619, 19 abr. 2012.

GALLUZZI, L. et al. Real-time PCR applications for diagnosis of leishmaniasis. **Parasites & Vectors**, v. 11, n. 1, 2 maio 2018.

GARNETT, M. J. et al. Systematic identification of genomic markers of drug sensitivity in cancer cells. **Nature**, v. 483, n. 7391, p. 570–575, 28 mar. 2012.

GOMES, C. M. et al. Complementary exams in the diagnosis of american tegumentary leishmaniasis. **Anais Brasileiros de Dermatologia**, v. 89, n. 5, p. 701–709, set. 2014.

GOTO, H.; LINDOSO, J. A. L. Current diagnosis and treatment of cutaneous and mucocutaneous leishmaniasis. **Expert Review of Anti-infective Therapy**, v. 8, n. 4, p. 419–433, abr. 2010.

GUEDES, DC. et al., New strategy to improve quality control of Montenegro skin test at the production level. **Revista Da Sociedade Brasileira De Medicina Tropical**, v. 50, n. 6, p. 788–794, 1 dez. 2017..

GULIEV, R. R. et al. Optical methods for bacterial detection and characterization. **Analytical Chemistry**, v. 92, n. 4943, 2020.

HANDLER, M.Z. et al. Cutaneous and mucocutaneous leishmaniasis. **Journal of The American Academy of Dermatology**, v. 73, p. 911-926, dez. 2015.

HERAUD, P. et al. Infrared spectroscopy coupled to cloud-based data management as a tool to diagnose malaria: a pilot study in a malaria-endemic country, **Malaria Journal**, v. 18, n. 348, out. 2019.

JORGE, Stéfano Gonçalves. **Hepatomegalia**. (2018). Disponível em: <https://hepcentro.com/hepatomegalia/>. Acesso em: 02 agor.2023.

KARAS, BY. et al. ATR-FTIR spectrum analysis of plasma samples for rapid identification of recovered COVID-19 individuals. **Journal of Biophotonics**, jul. 2023.

KOTHALAWALA, H. S.; KARUNAWEERA, N. D. Loop-mediated isothermal amplification assay as a sensitive diagnostic tool for *Leishmania donovani* infections in Sri Lanka. **Ceylon Medical Journal**, v. 61, n. 2, p. 68, 23 jun. 2016.

LACHAUD, L. et al. Optimized PCR Using Patient Blood Samples For Diagnosis and Follow-Up of Visceral Leishmaniasis, with Special Reference to AIDS Patients. **Journal of Clinical Microbiology**, v. 38, n. 1, p. 236–240, 1 jan. 2000.

LAMOTTE, S. et al. The enemy within: targeting host–parasite interaction for antileishmanial drug discovery. **Plos Neglected Tropical Diseases**, v. 11, n. 6, p.1-12, 2017.

LARIOS, G. et al. A new strategy for canine visceral leishmaniasis diagnosis based on FTIR spectroscopy and machine learning. **Journal of Biophotonics**, v. 14, n. 11, 26 ago. 2021.

LIMA, A. C. V. M. DA R. **Estudo da variabilidade genética de *Leishmania (Viannia) braziliensis* Vianna, 1911 de diferentes regiões do Brasil**. Disponível em: <[https://bdtd.ibict.br/vufind/Record/UFMG\\_d6060c18df906ccf2e38ca12f3f42bc4](https://bdtd.ibict.br/vufind/Record/UFMG_d6060c18df906ccf2e38ca12f3f42bc4)>. Acesso em: 10 mar. 2022.

MACHADO-SILVA, A. et al. New perspectives for leishmaniasis chemotherapy over current anti-leishmanial drugs: a patent landscape. **Expert Opinion on Therapeutic Patents**, v. 25, n. 3, p. 247–260, 20 dez. 2014.

MARIE, C.; PETRI, W. A. **Leishmaniose**. Disponível em: <<https://www.msmanuals.com/pt/profissional/doen%C3%A7as-infecciosas/protozo%C3%A1rios-extraintestinais/leishmaniose>>..

MARTINEZ-CUAZITL, A. et al. ATR-FTIR spectrum analysis of saliva samples from COVID-19 positive patients. **Scientific Reports**, v. 11, n. 1, p. 19980, 7 out. 2021.

MASMOUDI, A. et al. Old World cutaneous leishmaniasis: diagnosis and treatment. **Journal of Dermatological Case Reports**, v. 7, n. 2, 30 jun. 2013.

MATTOS, A. B. N. DE; TUMELERO, J. L. Perfil epidemiológico da Leishmaniose tegumentar no Brasil de 2015-2020. **Research, Society and Development**, v. 12, n. 3, p. e17212340385, 6 mar. 2023.

MOLINA, I. et al. Ultrasensitive Real-Time PCR for the Clinical Management of Visceral Leishmaniasis in HIV-Infected Patients. **American Journal of Tropical Medicine and Hygiene**, v. 89, n. 1, p. 105–110, 10 jul. 2013.

MUXEL, S. M. et al. Metabolomic Profile of BALB/c Macrophages Infected with *Leishmania amazonensis*: Deciphering L-Arginine Metabolism. **International Journal of Molecular Sciences**, v. 20, n. 24, p. 6248, 11 dez. 2019..

NADERER, T.; MCCONVILLE, M. J. The *Leishmania*-macrophage interaction: a metabolic perspective. **Cellular Microbiology**, v. 10, n. 2, p. 301–308, 9 dez. 2007.

NASEER, K. et al. FTIR spectroscopy of freeze-dried human sera as a novel approach for dengue diagnosis. **Infrared Physics & Technology**, v. 102, p. 102998, nov. 2019.

NUZUM, E. et al. Diagnosis Of Symptomatic Visceral Leishmaniasis By Use Of The Polymerase Chain Reaction On Patient Blood. **Journal of Infectious Diseases**, v. 171, n. 3, p. 751–754, 1 mar. 1995.

ORGANIZAÇÃO PAN AMERICANA DE SAÚDE. Manual de procedimentos para vigilância e controle de Leishmanioses nas américas. Washington, D.C. 2019. Disponível em: [https://iris.paho.org/bitstream/handle/10665.2/50524/9789275320631\\_spa.pdf](https://iris.paho.org/bitstream/handle/10665.2/50524/9789275320631_spa.pdf). Acesso em: 03 ago.2023.

ORGANIZAÇÃO PAN AMERICANA DE SAÚDE. Leishmanioses: informe epidemiológico das Américas, n. 10, 2021. Washington, D.C.: OPS; 2021. Disponível em: <https://iris.paho.org/handle/10665.2/51742>

ORGANIZAÇÃO PAN AMERICANA DE SAÚDE. Leishmanioses: informe epidemiológico das Américas, n. 11, 2022. Washington, D.C.: OPS; 2022. Disponível em: <https://iris.paho.org/handle/10665.2/56832>

PEPE, M. S. et al. Phases of Biomarker Development for Early Detection of Cancer. **Journal Of The National Cancer Institute**, v. 93, n. 14, p. 1054-1061, jul. 2001.

PSYCHOGIOS, N. et al. The Human Serum Metabolome. **PLoS ONE**, v. 6, n. 2, p. e16957, 16 fev. 2011.

**Secretaria Municipal de Saúde**. Disponível em: <<https://www.pmf.sc.gov.br/entidades/saude/index.php?cms=zoonoses+++leishmaniose+tegmentar+americana>>. Acesso em: 02 ago. 2023.

REIS, L. de C. et al. Mecanismos imunológicos na resposta celular e humoral na leishmaniose tegumentar americana. **Revista de Patologia Tropical / Journal of Tropical Pathology**, v. 35, n. 2, p. 103–116, 2007.

REITHINGER, R.; DUJARDIN, J.-C. Molecular Diagnosis of Leishmaniasis: Current Status and Future Applications. **Journal of Clinical Microbiology**, v. 45, n. 1, p. 21–25, 8 nov. 2006.

RYAN J. R. et al. Enzyme-linked immunosorbent assay based on soluble promastigote antigen detects immunoglobulin M (IgM) and IgG antibodies in sera from cases of visceral and cutaneous leishmaniasis. **J Clin Microbiol**, v. 40, n. 3, p. 1037-1043, mar. 2002.

SAAB, M. et al. Diagnosis of Cutaneous Leishmaniasis: why punch when you can scrape? **Am J Trop Med Hyg**, v. 92, n. 3, p. 518-522, mar. 2015.

SALARI, S. et al. Evaluation of a new live recombinant vaccine against cutaneous leishmaniasis in BALB/c mice. **Parasites & Vectors**, v. 13, n. 1, 12 ago. 2020.

SANTOS, Y. et al. **Ciências da saúde: ensino, formação e pesquisa**. Editora Amplla, 2021.

SANTOS-PEREIRA, S. et al. Leishmania amazonensis resistance in murine macrophages: Analysis of possible mechanisms. **PLOS ONE**, v. 14, n. 12, p. e0226837, dez. 2019.

SIEWE, N. et al. Immune response to infection by Leishmania: A mathematical model. **Mathematical Biosciences**, v. 276, p. 28–43, jun. 2016.

SINGH, S. New developments in diagnosis of leishmaniasis. **PubMed**, v. 123, n. 3, p. 311–30, 1 mar. 2006.

STERVEDING, D. The history of leishmaniasis. **Parasites & Vectors**, v. 10, n. 82, fev. 2017.

SUDARSHAN, M. et al. A Correlative Study of Splenic Parasite Score and Peripheral Blood Parasite Load Estimation by Quantitative PCR in Visceral Leishmaniasis. **Journal of Clinical Microbiology**, v. 53, n. 12, p. 3905–3907, dez. 2015.

TAHERI, A.R. et al. Plasma levels of interleukin-4 and Interferon- $\gamma$  in patients with chronic or healed cutaneous leishmaniasis. **Iran J Basic Med Sci**, v. 17, n. 3, 2013.

THAKUR, J. J; KAUR, S. Leishmaniasis diagnosis: an update on the use of parasitological, immunological and molecular methods. **Journal of Parasitic Diseases**, v. 44, p 253, mar. 2020.

TORRES-GUERRERO, E. et al. Leishmaniasis: a review. **F1000Research**, v. 6, n. 1, p. 750, maio 2017.

VAN DER MEIDE, W. F. et al. Quantitative nucleic acid sequence-based assay as a new molecular tool for detection and quantification of *Leishmania* parasites in skin biopsy samples. **J Clin Microbiol**, v. 43, n. 11, p. 5560-5566, nov. 2005.

VARGAS, M.; MARIA, A. Aspectos epidemiológicos da Leishmaniose Tegumentar Americana no Brasil em 2022. **Peer Review**, v. 5, n. 11, p. 294–305, jun. 2023.

## 2. Artigo

*Manuscrito submetido ao periódico ACS Infectious Diseases (Anexo)*

### **Diagnosis of cutaneous leishmaniasis using FTIR spectroscopy and machine learning: an animal model study**

Gabriela Pacher<sup>1</sup> | Thiago Franca<sup>2</sup> | Miller Lacerda<sup>2</sup> | Natália O. Alves<sup>1</sup> | Eliane M. Piranda<sup>1</sup> | Carla Arruda<sup>1\*</sup> | Cícero Cena<sup>2\*</sup>

\*carla.arruda@ufms.br, \*cicero.cena@ufms.br

<sup>1</sup>Laboratório de Parasitologia Humana, Instituto de Biociências, Universidade Federal de Mato Grosso do Sul, Campo Grande, Brazil

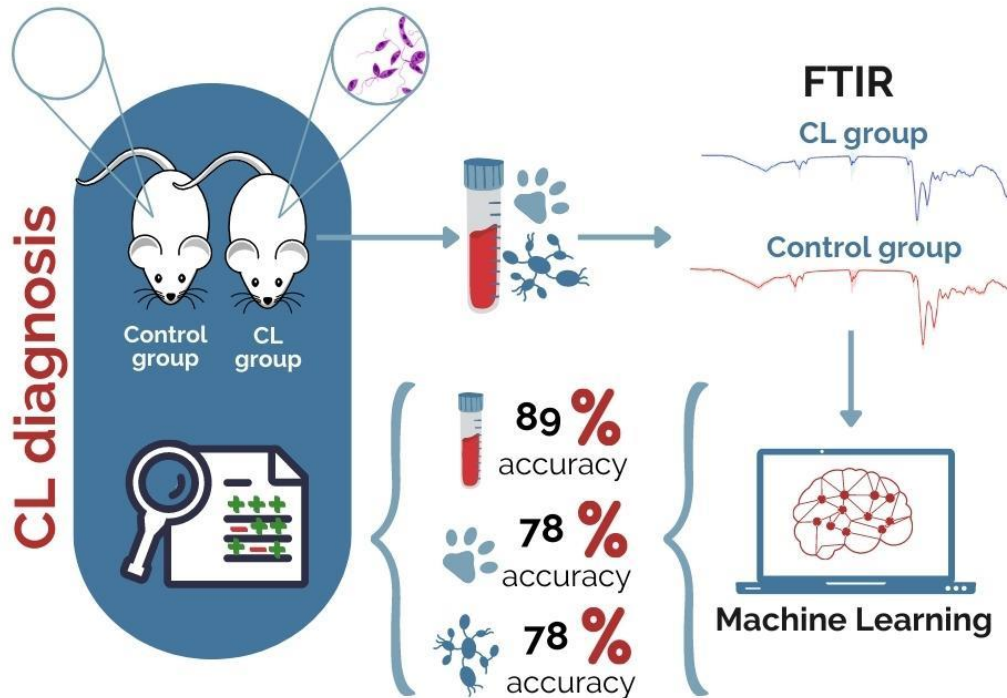
<sup>2</sup>Laboratório de Óptica e Fotônica (SISFOTON-UFMS), Instituto de Física, Universidade Federal de Mato Grosso do Sul, Campo Grande, Brazil

#### **ABSTRACT**

Cutaneous Leishmaniasis (CL) is a polymorphic and spectral skin disease caused by *Leishmania* spp protozoan parasites. CL is difficult to diagnose because conventional methods are time-consuming, expensive, and not very sensitive. Fourier transform infrared spectroscopy (FTIR) with machine learning (ML) algorithms has been explored as an alternative to achieve fast and accurate results for many disease diagnoses. Besides the good accuracy exhibited in many studies, the spectral differences between infected and non-infected groups are too subtle to be noticed. Since the sample set inhomogeneity (sex, age, diet, etc.) usually carries great data variance and hinders a more detailed understanding of the spectral characteristics and immune responses, in this study we explore this innovative methodology for CL diagnosis in an animal model study. Blood serum, skin lesions, and draining popliteal lymph node samples were obtained from BALB/C mice experimentally infected with *Leishmania (Leishmania) amazonensis*. The FTIR method and ML algorithms demonstrated reasonable accuracy in differentiating between infected (CL group) and non-infected (control group) samples. The best overall accuracy (88.9%) was obtained using Principal Component Analysis (PCA) and Support Vector Machine (SVM) in the 4000-700 cm<sup>-1</sup> range for blood serum samples. The accuracy obtained for skin lesions and popliteal lymph node samples was around 77.8%. According to the findings, successful functioning of prediction models is mostly related to data analysis rather than the differences in the molecular composition of the samples.

**Keywords:** Infrared spectroscopy; experimental cutaneous leishmaniasis; *Leishmania amazonensis*; multivariate analysis; photodiagnosis; biofluids.

## Graphical Abstract



## INTRODUCTION

Cutaneous Leishmaniasis (CL) is a polymorphic disease with wide clinical manifestations, where single or multiple ulcers affect skin, possibly able to spread mucous membranes. CL is caused by protozoan parasites of the genus *Leishmania*, transmitted by female phlebotomine sand fly bites. Classified by World Health Organization (WHO) as one of the most relevant neglected diseases [1], CL is distributed in subtropical and tropical regions, affecting 98 countries in America, Africa, Asia, and Europe [2]. In the Americas, in the last 20 years, approximately one million new cases of CL and mucous leishmaniasis were reported to the Pan American Health Organization (PAHO), with an average of 50,000 per year [3].

One of the main challenges for the control of CL lies in early diagnosis. Clinical diagnosis supported by epidemiological data with parasitological methods is the most common approach. Usually, the clinical method is based on the pathophysiology and localization of cutaneous lesions [4]. However, the differential diagnosis depends on reliable patient clinical data [5], then the identification of the parasite is mandatory for assertive diagnosis.

Parasitological methods are based on the observation of the parasite in the cutaneous lesion fragment and can be divided into direct and indirect techniques. The direct parasitological

test is performed with scraping, puncture, or biopsy of the lesion, stained by Giemsa or Leishman [4]. The indirect parasitological test is performed from the in vitro culture of the parasite in the bifasic medium Novy and McNeal blood agar modified by Nicolle (NNN), which may take a few weeks to provide results. Moreover, these methods are invasive and depend on lesion evolution, the nature of biological material collected, the sample processing, and mainly professional expertise [6]. To improve the accuracy of CL, molecular methods, such as real-time polymerase chain reaction, may be used for nuclear or kinetoplast DNA identification [7]; however, it demands technical training and high-cost resources. On the other hand, serological tests demonstrate low antibody titers, being less sensitive for the detection of CL.

To overcome these main drawbacks, new alternatives have been investigated by optical spectroscopy techniques. Fourier Transform Infrared Spectroscopy (FTIR) associated with multivariate analysis demonstrates great potential for a faster, non-invasive, lower-cost, and high-sensitivity diagnostic test development [8-10].

FTIR spectroscopy with machine learning (ML) was able to diagnose various infectious diseases with high accuracy. Saliva spectrum analysis provided 100% overall accuracy for COVID-19 diagnosis [11], and dengue-infected blood serum samples provided 89% accuracy [12]. The identification of other microorganisms was promising, such as *Salmonella* [13] and *Staphylococcus aureus* [14], with accuracy above 87% and 100% by using environmental samples and blood, respectively. Parasite infection from *Trypanosoma* sp, *Leishmania (Leishmania) infantum* [15], and *Plasmodium* sp [16] was identified with accuracy above 92% by analyzing in vitro culture and patients' blood, respectively.

Our group have demonstrated the potential use of FTIR spectroscopy with machine learning for the diagnosis of canine visceral leishmaniasis (VL) by using blood serum with accuracy above 85%, assigned to changes in the vibrational bands related to protein molecules [17]. ML algorithms also provided an overall accuracy of 91.67% for paracoccidioidomycosis diagnosis using human blood serum FTIR spectra, where the main vibrational changes were assigned to proteins and lipids molecules [18]. In this work, we investigated the potential use of FTIR spectroscopy with ML for CL diagnosis in a murine experimental model using BALB/c mice blood serum, skin lesion samples, and the draining popliteal lymph node. We analyze strategies to improve the overall accuracy obtained by different ML algorithms to classify *Leishmania (L.) amazonensis* infected and non-infected samples once CL does not provide a specific biomarker in the bloodstream to be easily identified, as occurs in VL.



## **EXPERIMENTAL METHODS**

### **Murine model for experimental cutaneous leishmaniasis**

Sixty female BALB/c mice, aged 4-6 weeks, obtained from the central animal facility of the Institute of Biosciences of Federal University of Mato Grosso do Sul (UFMS) were kept in mini-isolators coupled to a ventilated rack, fed with balanced food (Nuvital®, Nuvilab®) and hydrated with filtered water *ad libitum*. The Ethics Committee on the Use of Animals (CEUA/UFMS), under Protocol Number 1.174 /2021, approved procedures.

The animals were distributed into two groups (infected mice, called *CL group*, and non-infected mice, called *control group*). In CL group, 30 animals were subcutaneously infected with *Leishmania (Leishmania) amazonensis* promastigote forms ( $1 \times 10^6/0.1$  mL/PBS) into the right hind footpad (RHFP). In control group, 30 animals received the same volume (0,05 mL) of PBS subcutaneously into RHFP.

The kinetics of the cutaneous lesions were evaluated weekly, up to 30 days post-infection. Footpad thickness was measured using a caliper with an accuracy of 0.01 mm (Worker®) and was expressed as the difference between the infected and the mean of five non-infected footpads. The animals were euthanized 30 days after infection, using inhalational anesthesia with isoflurane followed by intraperitoneal injection of thiopental.

Parasite load was evaluated in RHFP (n=8) and in draining lymph node (n=11) soon after euthanasia. The organs were removed, weighed and homogenized in 1 mL of Schneider's Insect Medium (Sigma-Aldrich®, SP/Brazil) supplemented with 20% foetal bovine serum (Sigma-Aldrich®, SP/Brazil), 10,000 U/mL penicillin and 10 mg/mL streptomycin (Sigma-Aldrich®, SP/Brazil). The limiting dilution assay was performed in duplicate [19]. Parasite load was calculated using the geometric mean reciprocal of positive titers obtained for the homogenate of each organ divided by the respective weight, and the number of parasites per nanogram of tissue was then calculated.

### **Blood serum and tissue sampling**

Blood was collected by cardiac puncture and serum was obtained by centrifugation at 5000 rpm for 10 minutes. A fragment of previously induced cutaneous lesion (RHFP) and its draining popliteal lymph node were removed and homogenized in 1 mL of complete Schneider's Insect Medium (Sigma-Aldrich®, SP/Brazil), as described above. Samples were stored at -20 °C.

### **Sample preparation for FTIR**

Serum samples and the macerates of cutaneous lesion and the draining popliteal lymph node from BALB/C mice were thawed at room temperature (27 °C) for 10 min. After homogenization, 90 µL of samples were deposited on silicon substrate in three deposition steps (30 µL in each step), followed by drying at 40 °C for 60 min. Each sample was analyzed in triplicate and 60 samples were measured.

### **FTIR spectra acquisition and data analysis**

Samples were analyzed by vibrational optical spectroscopy (FTIR, Perkin-Elmer, model Spectrum 100) with an attenuated total reflectance accessory. For analysis, we recorded infrared absorption spectra in the range from 4000 to 600  $\text{cm}^{-1}$ , 4  $\text{cm}^{-1}$  resolution, and 10 scans. Each sample was measured at two different points, from the center to the edge of the sample, to avoid any heterogeneity in the sample's composition due to the drying process [20].

The averaged FTIR spectra were pre-treated using the Standard Normal Variate (SNV) method to reduce possible experimental data random fluctuations [18]. Then, the data were processed by the principal component analysis (PCA), an unsupervised analysis used for dimensionality reduction, which aims to separate the data according to the data variance with the aid of Python (version 3.9.12) software using the Scikit-learn package (version 1.1.2).

Sample classification was conducted using predictive models generated by machine learning (ML) algorithms, harnessing the output data from PCA applied to 70% of the sample set. Our study employed the Support Vector Machine (SVM) method as a prediction algorithm. SVM is a technique that organizes sample classes by optimizing a hyperplane - this hyperplane can exhibit either linear or nonlinear characteristics, and its optimization aims to achieve peak performance in separating the classes [21].

Within the SVM framework, we explored diverse functions to construct the hyperplanes, encompassing Linear, Quadratic, Cubic, and various degrees of granularity such as Fine, Medium, and Coarse. Additionally, our investigation extended to different spectral ranges: 4000-600  $\text{cm}^{-1}$ , 3100-2800  $\text{cm}^{-1}$ , and 1800-700  $\text{cm}^{-1}$ . These spectral ranges were chosen to capture pertinent sample data patterns and facilitate classification.

First, we performed the sample classification using the first three PCs - which significantly contributed to data variance - avoiding overfitting. The prediction model performance was evaluated by its classification accuracy registered in a Leave-one out cross-validation test. The LOOCV method involves selecting a sample from the dataset for testing while employing the remaining samples to construct the prediction model (training).

Subsequently, the accuracy of the prediction model is assessed by testing it against the withdrawn sample data, reiterating this process until all samples have undergone testing [22].

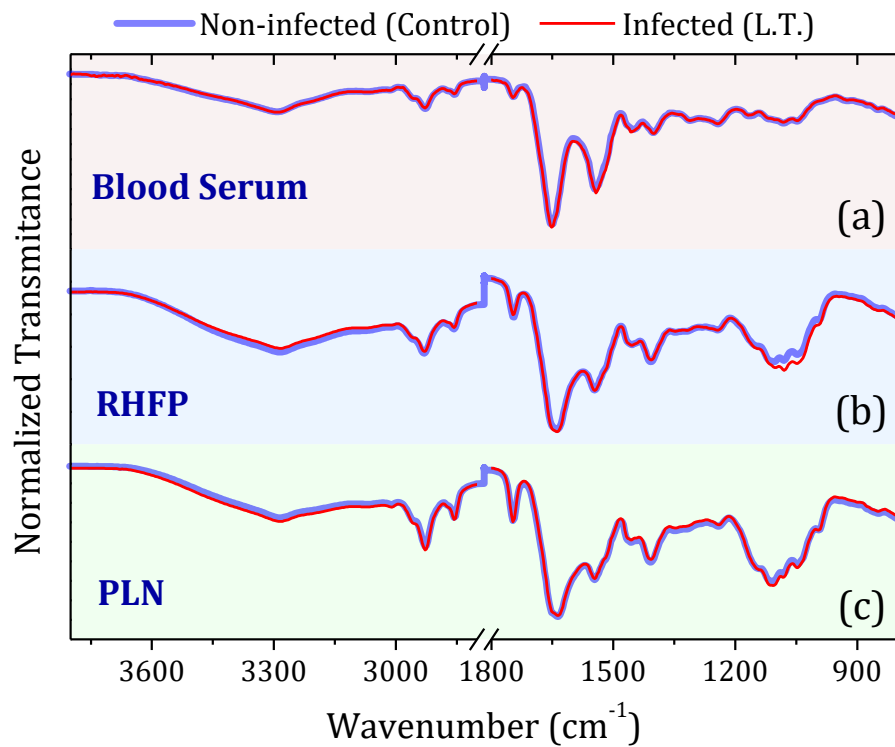
To improve the overall accuracy and still avoid overfitting, we adopted the Feature Selection Recursive Feature Elimination (RFE) method, which identifies the key PCs contributing significantly to achieving high accuracy while discarding PCs with minimal impact on accurate sample classification within the ML tests [23]. The inclusion or exclusion of specific PCs was determined by evaluating the accuracy of employing Linear Discriminant Analysis (LDA) in classifying samples within a Leave One Out Cross-Validation (LOOCV) framework. In a concise explanation, Discriminant Analysis (DA) categorizes samples by gauging the separation between sample data and the contour constructed through linear functions, which demarcate distinct classes (groups) [24].

Upon establishing the optimal number of PCs and identifying the most influential ones for sample classification in each analyzed spectral range, we executed a LOOCV test using the corresponding RFE-PCs. This test was carried out using SVM and different functions. Ultimately, we pinpointed the most effective spectral range, ML algorithm, and PCs to construct a predictive model. Its generalization capacity was evaluated through an external validation test utilizing 30% of the sample set.

## RESULTS AND DISCUSSION

*L. (L.) amazonensis* infected BALB/c mice (n=30) showed progressive footpad lesions until 4-weeks after infection. At the time of serum analysis (30 days p.i.), an average of  $7,18 \times 10^{26}$  and  $1,18 \times 10^{24}$  parasites per g of tissue were detected in the infection site (right hind footpad, RHFP) and in the popliteal draining lymph node, respectively. BALB/c mice are highly susceptible to *L. (L.) amazonensis* infection. In this murine model, a Th1/Th2 immune response co-exist, and the infection progresses as parasites disseminate from the cutaneous lesion to lymph nodes, spleen, and liver around 90 days post-infection [24,25].

The average FTIR-SNV spectra for murine blood serum, RHFP, and popliteal lymph nodes (PLN) from the control and CL groups, Figure 1, exhibits remarkable similarity for both groups for any kind of sample analyzed. The isolated spectra for each sample and group with respective standard deviation can be visualized in the supplementary material (SM). The main vibrational bands were assigned to proteins, carbohydrates, fatty acids, and lipids, and their vibrational assignments and molecular groups are displayed in Table 1.



**Figure 1:** Superimposed average FTIR-SNV spectra for murine (a) blood serum; (b) RHFP – right hind footpad (infection site); (c) PLN – popliteal lymph nodes. The control group is exhibited with a blue line, and the *L. (L.) amazonensis* (Cutaneous Leishmaniasis) Infected group with a red line.

The FTIR spectra obtained from different sample sources exhibited the same vibrational modes. A prominent vibrational band around  $1650\text{ cm}^{-1}$  assigned to Amide I is remarkable. Amides I and II usually play an important role in group classification in photodiagnosis studies. A great difference between blood serum samples and RHFP|PLN is evidenced by  $1546\text{ cm}^{-1}$  and the wide band centered around  $1103\text{ cm}^{-1}$ . The first band can be assigned with Amide II (C-N stretching and N-H bending), while the second is the result of several vibrational modes superimposed, which can be assigned with carbohydrates, phospholipids, and protein molecules. This set of overlapped bands shows a small difference between CL group and the infected RHFP samples, which may contribute to the sample classification.

**TABLE 1** Vibrational molecular mode assignments and related molecules groups from murine samples FTIR spectra [8, 17].

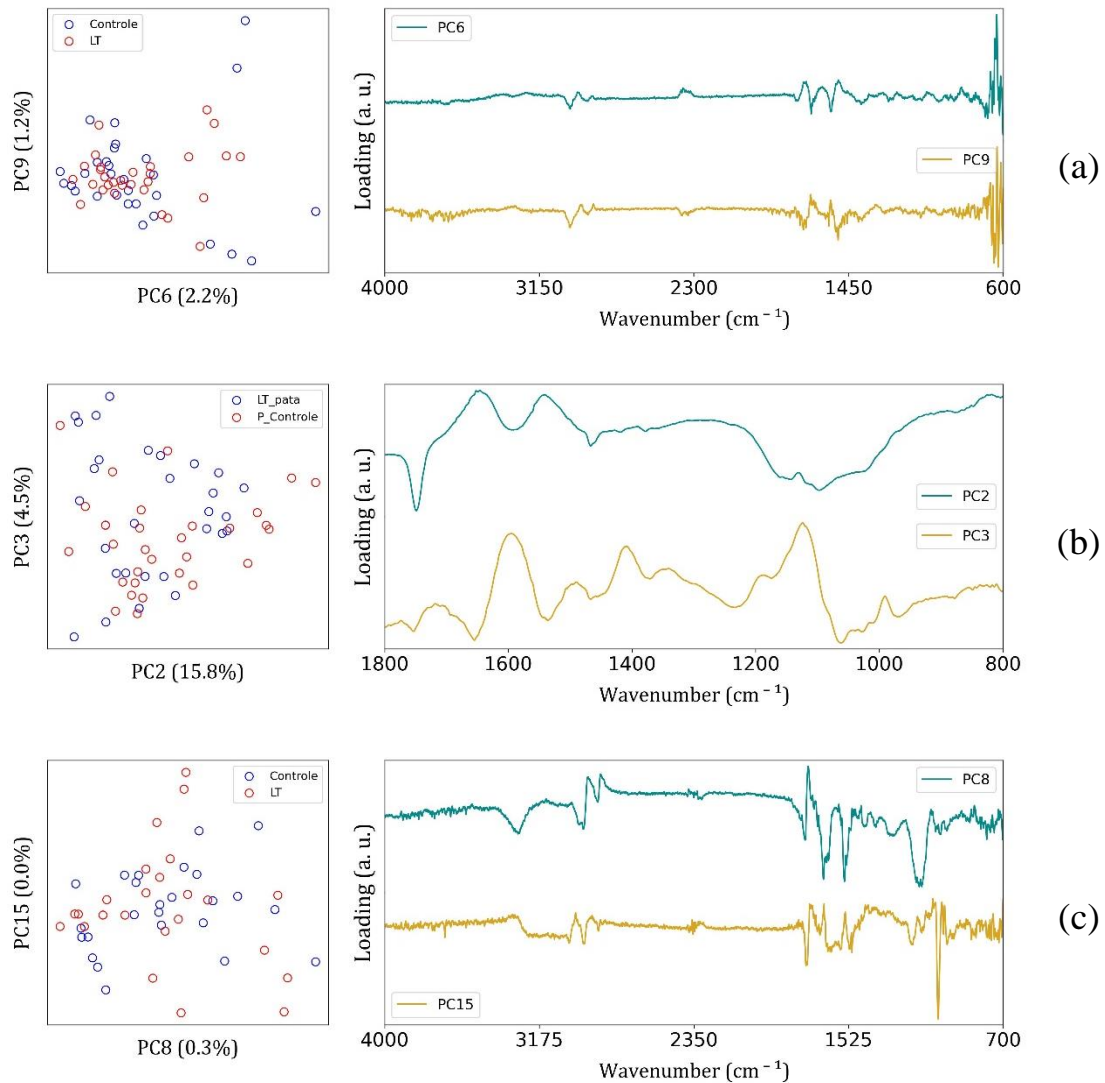
Band (cm <sup>-1</sup> )	Vibrational assignments	Organic group	Biomolecule	Samples
1048	C-O stretching	Sugar		Serum, RHFP and PLN
1085	PO <sub>2</sub> <sup>-</sup>	Sugar		Serum, RHFP and PLN
1110	C-O stretching	Carbohydrates; proteins		RHFP and PLN
1166	C-O stretching; C=O stretching; C-O-H; C-O-C	Carbohydrates; proteins	Serine; Apolipoprotein C; Threonine; Tyrosine	Serum, RHFP and PLN
1247	P-O stretching	Lipid phosphate; amide III		Serum, RHFP and PLN
1307		Amide III; collagen	Transferrin; glycoprotein; α <sub>1</sub> -acid	Serum, RHFP and PLN
1403	O-H bending; S=O stretching; COO <sup>-</sup> symmetric stretching	Amino acids; proteins	Haptoglobin; IgG1; IgA; IgM; Fibrinogen;	Serum, RHFP and PLN
1407	S=O stretching	Amino acids; proteins		RHFP and PLN
1455	CH <sub>2</sub> deformation; C-H scissoring bending; COO <sup>-</sup> symmetric stretching; C-HAmide II; lipids bending	Amide II; lipids	Apolipoprotein A	Serum, RHFP and PLN
1543	N-O stretching; N-H bending	Amide II	Albumin; IgG4	Serum, RHFP and PLN
1647	C=C stretching; C=O stretching	Amide I; lipids	IgG2, IgG3	Serum, RHFP and PLN
1744	C=O stretching	Fatty acids		RHFP and PLN
1751	C-H bending; C=O stretching	Amide I; fatty acids; phospholipids		Serum, RHFP and PLN
2858	N-H stretching; CH <sub>2</sub> symmetric stretching	Lipids; proteins	Apolipoprotein-B	Serum, RHFP and PLN
2925	C-H stretching; CH <sub>2</sub> asymmetric stretching	Lipids; proteins		Serum, RHFP and PLN (more RHFP and PLN)
2937	C-H	Lipids		RHFP
2962	C-H stretching; CH <sub>3</sub> asymmetric and symmetric stretching	Lipids; proteins; fatty acids		Serum, RHFP and PLN
3277	N-H	Proteins	Amide A	RHFP and PLN
3295	O-H and N-H stretching	Water; amide A		Serum, RHFP and PLN

The main prominent band identified for both groups was assigned for Amide I, around 1647-1751 cm<sup>-1</sup>, from the C=O stretching vibration mode. This was followed by 1247-1307 cm<sup>-1</sup>, assigned to Amide III, and lipid phosphate band (PO<sub>2</sub>) at 1085 and 1247 cm<sup>-1</sup> due to the symmetric P-O stretching vibration mode. Vibrational bands from N-H and C-H stretching assigned to lipids, protein and fatty acid molecules were evidenced in the 2858-2962 cm<sup>-1</sup> range. The wide vibrational band at around 3295 cm<sup>-1</sup> is attributed to the superposition of N-H and O-H molecular modes from proteins and water molecules.

The FTIR-SNV potential for group separation was analyzed by PCA. First, we used the PC1xPC2 score plot to investigate the clustering tendency and group separation for the three sample sets – blood serum, RHFP, PLN –to improve the cluster formation we also analyzed the spectral range contribution for clustering. The score plot (PC1xPC2) and loadings results obtained in the 4000-800, 3100-2800, and 1800-800 cm<sup>-1</sup> range for blood serum, RHFP, and PLN can be visualized in the SM. In summary, both groups form the same cluster, which hinders

the group classification with PC1 and PC2, and the loading shows that the data variance is concentrated in the main bands observed in each spectral range – as expected.

Previous studies have shown the potential use of high-order PCs to improve clustering and group separation [26,27]. Here, we selected these high-order PCs using the RFE (Feature Selection Recursive Feature Elimination) algorithm. Figure 2 shows the score plot and loading for the three sample sets in selected spectral ranges in which we could achieve the best results for sample classification.



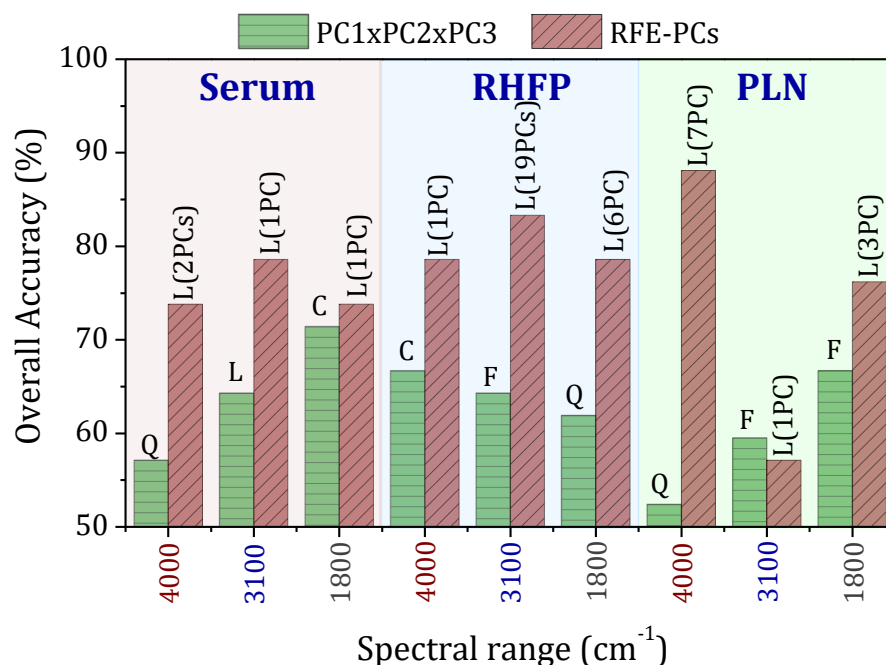
**Figure 2:** Principal Component Analysis for murine (a) blood serum; (b) RHFP – right hind footpad (infection site); (c) PLN – popliteal lymph nodes. The control group exhibited a red circle, and the *L. (L.) amazonensis* (Cutaneous Leishmaniasis) Infected group with a blue circle. The score plot and respective loading for the three different ranges were analyzed: (a) 4000 to 800  $\text{cm}^{-1}$ ; (b) 1800 to 800  $\text{cm}^{-1}$ ; and (c) 4000 to 800  $\text{cm}^{-1}$ .

Figure 2 shows that even the RFE-PCs could not provide a great clusterization with potential group separation – as observed in some cases [26,27]. Figure 2(a) represents only 3.4% of data variance for blood serum samples in the 4000 to 800  $\text{cm}^{-1}$ .

The loadings, Figure 2, exhibit the data variance for the two first PCs. The main data variance was observed for around 1746, 1652 and 1560  $\text{cm}^{-1}$ , assigned to Amide I and II vibrational bands, probably related to the presence of IgG1 and IgG2a. It's known that serum levels of anti-*L. (L.) amazonensis* IgG1 and IgG2 antibodies increase during infection in BALB/c mice, and the serum recognized amastigote antigens 5 weeks after infection, with the presence of IgG1 and IgG2 associated with a Th2 profile and the progression of the disease [28]. Similar bands were found relevant for visceral leishmaniasis diagnosis from canine blood serum [17].

Machine learning algorithms were tested using PCA data from murine blood serum, RHFP, and PLN FTIR-SNV spectra. Figure 3 shows the best overall accuracy results achieved for SVM algorithms in the 4000-800  $\text{cm}^{-1}$ , 1800-700  $\text{cm}^{-1}$ , and 3100-2700  $\text{cm}^{-1}$  range in the LOOCV test. LOOCV tests were performed for performance comparison using the first three PCs (PC1xPC2xPC3) and the RFE-PCs. Using only the first 3 PCs, we avoid overfitting, but also, we may fall into an underfitting situation – here, we found the best accuracy by changing the function (Fine, Quadratic, Linear, and Coarse) associated with the SVM algorithm to build the separation hyperplane.

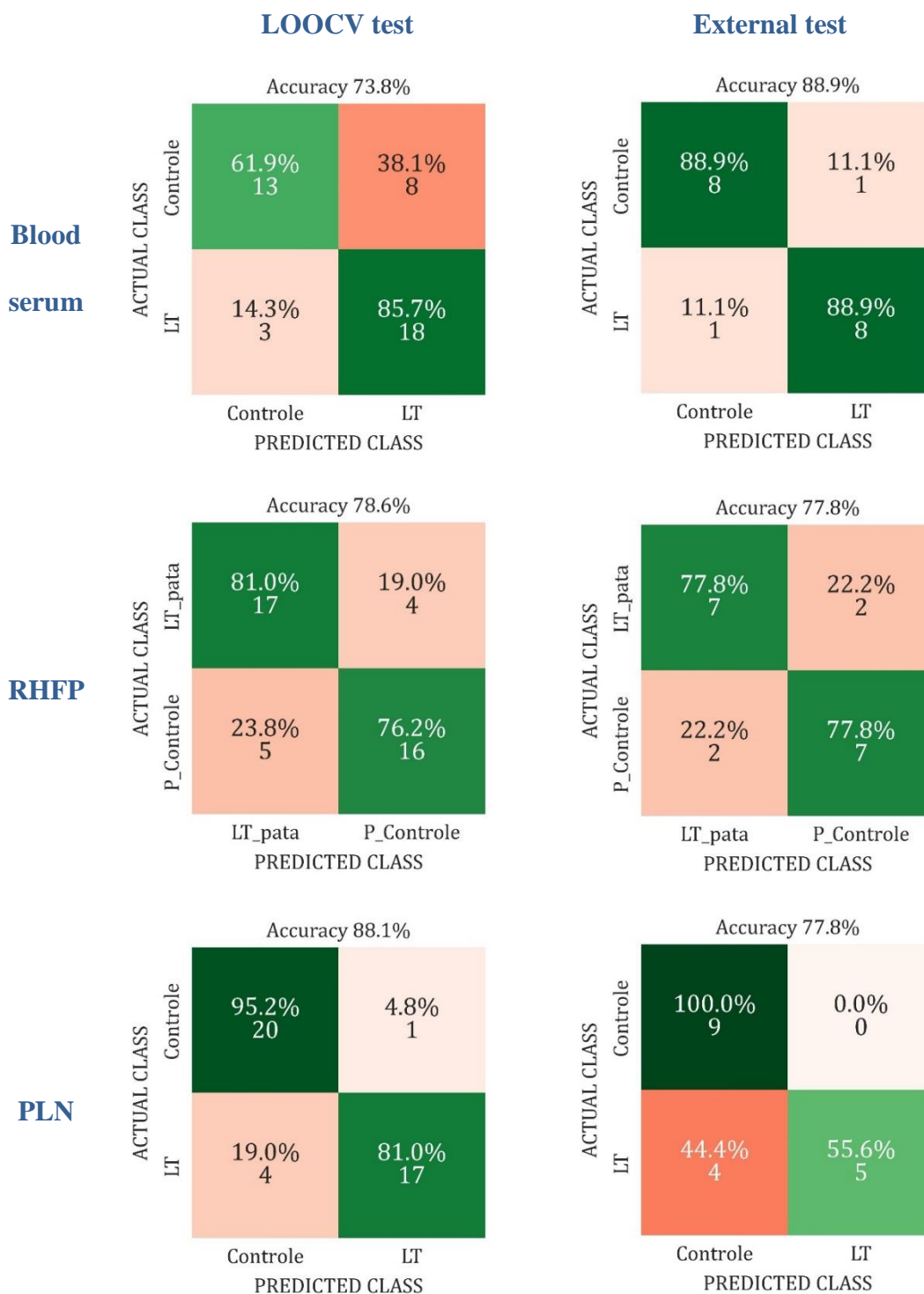
However, using RFE-PCs, the Linear SVM provided the best accuracy for any sample and spectral range analyzed. Note that the number of RFE-PCs varies; in some situations, only 1 PC is required to obtain a better performance in comparison with PC1xPC2xPC3, but on the other side for RHFP in the 3100-2800  $\text{cm}^{-1}$  range, 19 PCs were suggested by the algorithm – which may drive us for an overfitting situation.



**Figure 3:** Overall accuracy obtained in the LOOCV tests for Support vector machine (SVM) algorithms. Different ranges were analyzed by using PC1xPC2xPC3 (green bars with horizontal line pattern), and RFE-PCs (brown bars with left inclined line pattern): (i) 4000 to 800 cm<sup>-1</sup> (denoted as 4000 in the x-axis); (ii) 3100 to 2800 cm<sup>-1</sup> (denoted as 3100 in the x-axis); (iii) 1800 to 800 cm<sup>-1</sup> (denoted as 1800 in the x-axis). The highest overall accuracy was 88.1% for Linear SVM at 4000 to 800 cm<sup>-1</sup> range and PLN samples.

The algorithms' prediction performance was superior by using the RFE-PCs. The best overall accuracy (88.1%) was obtained using the Linear SVM in the 4000-800 cm<sup>-1</sup> range, with 7 PCs for PLN samples. Then to test the reliability of the prediction model, an external validation test was performed by using 30% of the sample set (not used during the training). The results obtained for RFE-PCs are summarized in Figure 4 by the confusion matrices - the complete set of results can be found in the SM.





**Figure 4:** Confusion matrix for the performance of Linear SVM algorithm using RFE-PCs. Classification results were obtained in the LOOCV test (right side) with a 70% data set and the External validation (left side) test with a 30% data set. Blood serum samples in the 4000 to 800  $\text{cm}^{-1}$  range using 2 PCs with 3.4% of data variance. RHFP samples in the 1800 to 800  $\text{cm}^{-1}$  range using 6 PCs with 20.3% of data variance. PLN samples in the 4000 to 800  $\text{cm}^{-1}$  range using 7 PCs with 32.1% of data variance.

Besides the good performance achieved in the LOOCV test for PLN samples, the prediction model did not show a good generalization power during the external validation test. The overall accuracy of 88.1% in LOOCV decreased to 77.8% at the external validation test in the 4000 to 800  $\text{cm}^{-1}$  range using 7 PCs (PC1xPC5xPC8xPC11xPC15xPC17xPC24) with 32.1% of data variance. Here, the best performance was observed for blood serum samples in the 4000 to 800  $\text{cm}^{-1}$  range using 2 PCs (PC6xPC9) with 3.4% of data variance – only one sample from each group was misclassified. For RHFP samples in the 1800 to 800  $\text{cm}^{-1}$  range using 6 PCs (PC2xPC3xPC6xPC9xPC15xPC18) with 20.3% of data variance, the overall accuracy achieved 77.8%.

We were expecting the best performance for RHFP and PLN samples, but our analysis must consider some aspects before making conclusions. First, there is a difference between how the overall accuracy is achieved in the LOOCV and external validation tests. In LOOCV, overall accuracy is calculated as the average value from each test made for each sample in the set, while in the external test, the prediction model is built from the training sample set, then the classification of the external sample is made, providing only one value of accuracy. Then, statistical fluctuation is expected in such cases since the number of samples is different.

From the spectral data analysis, we must consider two characteristics: RHFP and PLN must provide small deviations in the spectra not identified by the naked eye but important enough to be highlighted by PCA analysis since the control and CL groups must have different molecular composition due the presence of *L. (L.) amazonensis*. Besides that, we couldn't obtain a great group separation in the score plot even using RFE-PCs; probably, the difference between the groups is far below the detection capacity of the FTIR spectrometer – more studies are required. Additionally, the blood samples data analyzed in the 4000-800  $\text{cm}^{-1}$  range involves an O=C=O band, assigned to carbon dioxide in the bloodstream produced due the euthanasia, which can have an influence on this range but is not associated with the infection.

Finally, the immune responses in CL and, thereafter, the disease progression or cure are influenced by the cytokine expression profile (Th1 or Th2) induced by parasite/host interactions [29]. Although protective immunity against *Leishmania* spp. is dependent on a Th1 response, the Th1/Th2 dichotomy does not occur when BALB/c mice are infected with *L. (L.) amazonensis*, where the parasite promotes the impairment of multiple immune functions at early stages of infection, and a mixed Th1/Th2 response can be maintained [30]. In this way, numerous cytokines can be produced and even systemically detected, such as those with a pro-inflammatory phenotype (IFN- $\gamma$ , TNF- $\alpha$ , GM-CSF, etc.) and especially those with an anti-

inflammatory phenotype (IL-4, IL -3, IL-10, TGF- $\beta$ , M-CSF), linked to parasite replication and disease progression [29]. These cytokines may be included in the band ranges showing data variance, which could explain the best performance of blood serum samples for group classification.

## CONCLUSION

FTIR combined with ML analysis can be a prominent method for screening diagnosis tests for Cutaneous Leishmaniasis. Blood serum sample analyses could differentiate between infected (CL group) and non-infected animals (control group) with an overall accuracy of around 80% using the Linear SVM in the 4000-800  $\text{cm}^{-1}$  range, with only 2PCs (PC6 and PC9) and 3.4% data variance. RHFP and PLN samples also showed good overall accuracy - around 80% - for sample classification but required 6 and 7 PCs, respectively. Besides, the RHFP and PLN samples have different molecular compositions between the control and infected group - due to the presence of *L. (L.) amazonensis* protozoan – no clear evidence of such difference was found in the FTIR spectra and PCA analyses.

## ACKNOWLEDGMENTS

Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), code 001. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPQ), code 403651/2020-5; 302525/2022-0; 440214/2021-1. Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT), code 91/2023; 360/2022.

## DATA AVAILABILITY STATEMENT

Research data are not shared.

## REFERENCES

- [1] Pan American Health Organization. Manual of procedures for leishmaniasis surveillance and control in the Americas. Washington, D.C., 2019.
- [2] D. Sterveding, Parasites & Vectors. 2017, 10.
- [3] ORGANIZAÇÃO PAN AMERICANA DE SAÚDE. Leishmanioses: informe epidemiológico das Américas, n. 11, 2022. Washington, D.C.: OPS; 2022.
- [4] Manual de vigilância da leishmaniose tegumentar (Eds: Ministério da Saúde, Secretaria de Vigilância em Saúde). Brasília, Brasil, 2017, p. 66.

- [5] P. H. P. Cerutti, C. G. Lopes, F. G. L. Filho, and V. R. Guedes, *Revista de Patologia do Tocantins*. 2017, 4, 55.
- [6] S. Thakur, J. Joshi, and S. Kaur, *Journal of Parasitic Diseases*. 2020, 44, 253.
- [7] L. Galluzzi, M. Ceccarelli, A. Diotallevi, M. Menotta, and M. Magnani, *Parasites & Vectors*. 2018, 11, 1.
- [8] M. J. Baker, S. R. Hussain, L. Lovergne, V. Untereiner, C. Hughes, R. A. Lukaszewski, G. Thiéfinbg, and G. D. Sockalingum, *Chemical Society Reviews*. 2015, 45, 1803.
- [9] R. Etzioni, N. Urban, S. Ramsey, M. McIntosh, S. Schwartz, B. Reid, J. Radich, G. Anderson, and L. Hartwell, *Nature Reviews Cancer*. 2003, 3, 243.
- [10] M. S. Pepe, R. Etzioni, Z. Feng, J. D. Potter, M. L. Thompson, M. Thornquist, M. Winget, and Y. Yasui, *Journal of the National Cancer Institute*. 2021, 93, 1054.
- [11] A. Martinez-Cuazitl, G. J. Vazquez-Zapien, M. Sanchez-Brito, J. H. Limon-Pacheco, M. Guerrero-Ruiz, F. Garibay-Gonzalez, R. J. Delgado-Macuil, M. G. Garcia de Jesus, M. A. Corona-Perezgrovas, A. Pereyra-Talamantes, and M. M. Mata-Miranda, *Scientific Reports*. 2021, 11, 1.
- [12] K. Naseer, S. Ali, S. Mubarak, I. Hussain, B. Mirza, and J. Qazi, *Infrared Physics & Technology*. 2019, 102, 1.
- [13] M. Cordovana, N. Mauder, M. Kostrzewa, A. Wille, S. Rojak, R. M. Hagen, S. Ambretti, S. Pongolini, L. Soliani, U.S. Justesen, H. M. Holt, O. Join-Lambert, S. L. Hello, M. Auzou, A. C. Veloo, J. May, H. Frickmann, and D. Dekker, *Microorganisms*. 2021, 9, 853.
- [14] R. R. Guliev, A. Y. Suntsova, T. Y. Vostrikova, A. N. Shchegolikhin, D. A. Popov, M. A. Guseva, A. B. Shevelev, and I. N. Kurochkin, *Analytical Chemistry*. 2020, 92, 4943.
- [15] J.C. Aguiar, J. Mittmann, P.C. Caetano Júnior, and L. Raniero, *Experimental Parasitology*. 2018, 192, 46.
- [16] P. Heraud, P. Chatchawal, M. Wongwattanakul, P. Tippayawat, C. Doerig, P. Jearanaikoon, D. Perez-Guaita, and B. R. Wood, *Malaria Journal*. 2019, 18, 1.
- [17] G. Larios, M. Ribeiro, C. Arruda, S. L. Oliveira, T. Canassa, M. J. Baker, B. Marangoni, C. Ramos, and C. Cena, *Journal of Biophotonics*. 2021, 14, 1.
- [18] E. C.A. Brito, T. Franca, T. Canassa, S. S. Weber, A. M.M. Paniago, and C. Cena, *Photodiagnosis and Photodynamic Therapy*. 2022, 39.
- [19] NOBLE, William S. What is a support vector machine? *Nature biotechnology*, v. 24, n. 12, p. 1565-1567, 2006.
- [19] R. G. Titus, M. Marchand, T. Boon, and J. A. Louis, *Parasite Immunology*. 1985, 7, 545.

- [20] J. Cameron, H. Butler, D. S. Palmer, and M. J. Baker, *Journal of Biophotonics*. 2018, 11, 1.
- [21] WONG, T. Performance evaluation of classification algorithms by k-fold and leave-one-out cross-validation. *Pattern Recognition*, v. 48, n. 9, p. 2839-2846, 2015.
- [22] THEERTHAGIRI, P. Predictive analysis of cardiovascular disease using gradient boosting-based learning and recursive feature elimination technique. *Intelligent Systems with Applications*, v. 16, p. 200121, 2022.
- [23] WU, W. et al. Comparison of regularized discriminant analysis linear discriminant analysis and quadratic discriminant analysis applied to NIR data. *Analytica Chimica Acta*, v. 329, n. 3, p. 257-265, 1996.
- [24] R. P. Almeida, M. Barral-Netto, A. M. De Jesus, L. A. De Freitas, E. M. Carvalho, A. Barral, *American Journal of Tropical Medicine and Hygiene*. 1996, 54, 178.
- [25] S. M. N. Cupolilo, C. S. F. Souza, A. L. Abreu-Silva, K. S. Calabrese, and S. C. G. Costa, *Histology and Histopathology*. 2003, 18, 1059.
- [26] B. Resende, T. Franca, M. Paula, H. Cleveland, C. Cena, C. Nascimento. *Photochemistry and Photobiology*. 2023.
- [27] Y. Marangoni-Ghoreyshi, T. Franca, J. Esteves, A. Maranni, K. Portes, C. Cena, C. Leal. *RSC Advances*. 2023.
- [28] C. Hernández-Chinea, *Experimental Parasitology*. 2007, 116, 492.
- [29] F. O. Cardoso, T. Zaverucha-do-Valle, F. Almeida-Souza, A. L. Abreu-Silva, and K. S. Calabrese, *Frontiers in Microbiology*. 2020, 11, 1.
- [30] J. Ji, J. Sun, H. Qi, and L. Soong, *The American Journal of Tropical Medicine and Hygiene*. 2002, 66, 338.

# Submission Confirmation



Thank you for your submission

**Submitted to**  
ACS Infectious Diseases

**Manuscript ID**  
id-2023-00430b

**Title**  
Diagnosis of cutaneous leishmaniasis using FTIR spectroscopy and machine learning: an animal model study

**Authors**  
Pacher, Gabriela  
Franca, Thiago  
Lacerda, Miller  
Alves, Natalia  
Piranda, Eliane  
Pinto de Arruda, Carla Cardozo  
Cena, Cicero

**Date Submitted**  
21-Aug-2023

---

---

[Author Dashboard](#)



Serviço Público Federal  
Ministério da Educação

Fundação Universidade Federal de Mato Grosso do Sul



### CERTIFICADO

Certificamos que a proposta intitulada “Diagnóstico da Leishmaniose tegumentar utilizando espectroscopia ótica e aprendizagem de máquina”, registrada com o nº 1.174/2021, sob a responsabilidade de CARLA CARDOZO PINTO DE ARRUDA - que envolve a utilização de animais pertencentes ao filo Chordata, subfilo Vertebrata, para fins de pesquisa científica – encontra-se de acordo com os preceitos da Lei nº 11.794, de 8 de outubro de 2008, do Decreto nº 6.899, de 15 de julho de 2009, e com as normas editadas pelo Conselho Nacional de Controle de Experimentação Animal (CONCEA), e foi aprovada pela COMISSÃO DE ÉTICA NO USO DE ANIMAIS/CEUA DA UNIVERSIDADE FEDERAL DE MATO GROSSO DO SUL/UFMS, na 4ª reunião ordinária do dia 23/06/2021.

FINALIDADE	( ) Ensino ( x ) Pesquisa Científica
Vigência da autorização	02/08/2021 a 26/04/2023
Espécie/Linhagem/Raça	Mus musculus / BALB/c
Nº de animais	60
Peso/Idade	20-30 g / 4 semanas
Sexo	Fêmeas
Origem	Biotério Central da UFMS

Fábio José Carvalho Faria

Coordenador da CEUA/UFMS

Campo Grande, 25 de junho de 2021.



Documento assinado eletronicamente por Fabio Jose Carvalho Faria, Presidente de Comissão, em 25/06/2021, às 15:10, conforme horário oficial de Mato Grosso do Sul, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [https://sei.ufms.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://sei.ufms.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador 2641173 e o código CRC 94EA8E1D.