



Fundação Universidade Federal de Mato Grosso do Sul
Faculdade de Ciências Farmacêuticas, Alimentos e Nutrição
Programa de Pós-Graduação em Ciências Farmacêuticas



Vilma Almeida da S. Oliveira

**ANÁLISE ESPECTROSCÓPICA (FTIR) ASSOCIADA À ANÁLISE
MULTIVARIADA DE CEPAS DE *LEISHMANIA* sp. PROVENIENTES
DE CULTIVO AXÊNICO**

CAMPO GRANDE

2025



Fundação Universidade Federal de Mato Grosso do Sul
Faculdade de Ciências Farmacêuticas, Alimentos e Nutrição
Programa de Pós-Graduação em Ciências Farmacêuticas



Vilma Almeida da S. Oliveira

**ANÁLISE ESPECTROSCÓPICA (FTIR) ASSOCIADA À ANÁLISE
MULTIVARIADA DE CEPAS DE *LEISHMANIA* sp. PROVENIENTES
DE CULTIVO AXÊNICO**

Dissertação apresentada como requisito para a obtenção do título de mestre em ciências farmacêuticas, no Programa de Pós-Graduação em Farmácia. Na Faculdade de Ciências Farmacêuticas, Alimentos e Nutrição da Universidade Federal de Mato Grosso do Sul.

Orientadora: Dra. Carla Cardozo Pinto de Arruda.

Co-orientador: Dr. Cícero Cena.

CAMPO GRANDE

2025



Vilma Almeida da S. Oliveira

**ANÁLISE ESPECTROSCÓPICA (FTIR) ASSOCIADA À ANÁLISE
MULTIVARIADA DE CEPAS DE *LEISHMANIA* sp. PROVENIENTES
DE CULTIVO AXÊNICO**

Dissertação apresentada como requisito para a obtenção do título de Mestre em Ciências Farmacêuticas, no Programa de Pós-Graduação em Farmácia. Na Faculdade de Ciências Farmacêuticas, Alimentos e Nutrição da Universidade Federal de Mato Grosso do Sul.

Orientadora: Dra. Carla Cardozo Pinto de Arruda.

Co-orientador: Dr. Cícero Cena.

Banca Avaliadora

Presidente: Prof. Dr^a Carla Cardozo Pinto de Arruda - Orientadora, UFMS

Examinadora Externa: Prof. Dr^a Celeste da Silva Freitas de Souza, IOC/FIOCRUZ

Examinadora Interna: Prof. Dr^a Thalita Bachelli Riul, UFMS

Suplente

Examinadora Interna: Prof. Dr^a Eliane Mattos Piranda - UFMS



AGRADECIMENTOS

À Deus em primeiro lugar, pela vida e oportunidade de mais essa etapa de aprendizado.

À minha família que é minha maior motivação para querer crescer e me capacitar cada vez mais.

À minha orientadora, professora Carla Cardozo, pela excelência, dedicação, sabedoria e paciência em todas as etapas do mestrado.

Ao meu co-orientador, professor Cícero Cena, por me apresentar a esse novo mundo do fotodiagnóstico e a todas essas possibilidades e perspectivas que me trouxeram.

A todos os colaboradores e parceiros do Laboratório de Parasitologia Humana - IMBIO.

À minha amiga Fernanda, pelo apoio, e por não me deixar desistir nos momentos de crise.

À FUNDECT e à CAPES, pelo apoio financeiro.

Em geral a todos que de maneira direta ou indireta contribuíram para a elaboração deste trabalho.



Fundação Universidade Federal de Mato Grosso do Sul
Faculdade de Ciências Farmacêuticas, Alimentos e Nutrição
Programa de Pós-Graduação em Ciências Farmacêuticas



**“Feliz o Homem que encontrou a
sabedoria e alcançou o
entendimento, porque a sabedoria
vale mais do que prata, e dá mais
lucro que o ouro.”**

(Provérbios 3, 13-14)



RESUMO

As leishmanioses são doenças infectoparasitárias zoonóticas neotrópicas. Tais parasitoses acometem principalmente as regiões subdesenvolvidas, sendo consideradas negligenciadas e com alta incidência no Brasil. O isolamento de cepas de *Leishmania* spp. a partir do diagnóstico parasitológico indireto permite a identificação da espécie em causa no caso clínico, contribuindo para a vigilância e o controle epidemiológico da parasitose, além de direcionar a conduta clínica para uma abordagem terapêutica mais efetiva. A caracterização taxonômica do parasito é realizada usualmente pela eletroforese de isoenzimas ou por métodos moleculares, técnicas laboriosas realizadas em laboratórios de referência. A espectroscopia de absorção no infravermelho por transformada de Fourier (FTIR), associada à análise multivariada de dados por algoritmos, tem sido utilizada para identificação de microrganismos, uma vez que a análise de seus espectros infravermelhos fornece informações sobre o perfil vibracional dos seus componentes. Esse método inovador, de fácil aplicabilidade e alta reprodutibilidade atinge índices consideráveis de precisão, especificidade, sensibilidade, baixa taxa de erros falsos positivos e verdadeiros negativos. O objetivo deste trabalho foi padronizar uma nova metodologia para a identificação de espécies de *Leishmania* spp. mantidas em cultivo axênico a partir da análise multivariada de espectros de absorção no infravermelho das culturas dos parasitos, identificando as contribuições espectrais mais importantes para a identificação e associando-as aos componentes biológicos das amostras. As amostras submetidas à análise de FTIR foram cepas que já foram devidamente identificadas e caracterizadas por métodos convencionais: *Leishmania (Leishmania) amazonensis* (cepa IFLA/BR/1967/PH8 e MHOM/BR/2022/LT 013_22) e *Leishmania (Viannia) braziliensis* (cepa MHOM/BR/75/M2904). A absorção das amostras na região do infravermelho médio (MID, 4000-800cm⁻¹) foi registrada e os dados foram submetidos à análise estatística multivariada usando a análise de componentes principais (PCA) para diferenciar os grupos de amostras e destacar os principais fatores atribuídos aos modos vibracionais. A classificação automatizada das amostras foi obtida com o uso de algoritmos de aprendizado de máquina e o *Support Vector Machine* (SVM) foi usado para criar um modelo de previsão e classificação de amostras. Os espectros FTIR foram adquiridos a partir de 80 amostras de culturas e pré-processados utilizando a correção da variante normal padrão (SNV) e a filtragem da transformada rápida de Fourier (FFT). A PCA revelou um claro agrupamento espécie-específico, impulsionado por diferenças espectrais nas bandas vibracionais de proteínas, lipídios e ácidos nucleicos. Os modelos SVM foram treinados utilizando os resultados da PCA, atingindo uma precisão superior a 90% em todas as configurações testadas. O melhor modelo, utilizando um núcleo linear e os três primeiros componentes principais, atingiu 100% de exatidão, sensibilidade e especificidade na validação externa. Os resultados demonstram que a espectroscopia FTIR, em combinação com SVM, oferece uma estratégia rápida, de baixo custo e ampliável para o rastreamento e classificação de espécies de *Leishmania*, com aplicações promissoras para diagnósticos clínicos e de campo. Essa abordagem poderia facilitar um diagnóstico mais preciso e um tratamento mais eficaz para os pacientes.

Palavras-chave: Leishmanioses; espectroscopia de infravermelho; análise multivariada; fotodiagnóstico.



ABSTRACT

Leishmaniasis are neotropical vector-borne infectious-parasitic diseases. These diseases primarily affect underdeveloped regions and are considered neglected, with a high incidence in Brazil. The isolation of *Leishmania* spp. strains based on indirect parasitological diagnosis allows for the identification of the species involved in the clinical case. This contributes to the surveillance and epidemiological control of the parasitosis, as well as directing clinical management towards a more effective therapeutic approach. The taxonomic characterization of the parasite is typically performed through the use of isoenzyme electrophoresis or molecular methods. These techniques, however, are considered laborious and are typically executed in reference laboratories. The application of Fourier-transform infrared spectroscopy (FTIR), in conjunction with multivariate data analysis utilizing algorithms, has been instrumental in the identification of microorganisms. The analysis of their infrared spectra offers insights into the vibrational profile of their components. This innovative method is characterized by its ease of application and high reproducibility, resulting in substantial precision, specificity, sensitivity, and low false positive and true negative error rates. The objective of this study was to standardize a novel methodology for identifying *Leishmania* species maintained in axenic culture. This methodology was based on multivariate analysis of infrared absorption spectra of parasite cultures, with the aim of identifying the most significant spectral contributions to identification and associating them with the biological components of the samples. The samples submitted to FTIR analysis were strains that had previously undergone proper identification and characterization through conventional methods. The parasites in question are *Leishmania* (*Leishmania*) *amazonensis* (IFLA/BR/1967/PH8 and MHOM/BR/2022/LT 013_22) and *Leishmania* (*Viannia*) *braziliensis* (MHOM/BR/75/M2904). The absorption of the samples in the mid-infrared region (MIR, 4000-800 cm⁻¹) was recorded, and the resulting data were subjected to multivariate statistical analysis using principal component analysis (PCA). This analysis was performed to differentiate the groups of samples and highlight the main factors attributed to the vibrational modes. The automated classification of the samples was achieved through the implementation of machine learning algorithms, with the Support Vector Machine (SVM) being employed to construct a model for predicting and classifying samples. A total of 80 culture samples were subjected to FTIR, followed by preprocessing that entailed singular value decomposition (SVD) correction and fast Fourier transform (FFT) filtering. PCA revealed clear species-specific clustering, driven by spectral differences in the vibrational bands of proteins, lipids, and nucleic acids. The SVM models were trained using PCA results, achieving an accuracy of over 90% in all configurations tested. The optimal model, employing a linear kernel and the initial three principal components, attained 100% accuracy, sensitivity, and specificity in the external validation. The findings of this study demonstrate that FTIR spectroscopy, in conjunction with SVM, provides a rapid, cost-effective, and scalable approach for the screening and classification of *Leishmania* species. This strategy holds considerable promise for clinical and field diagnostics. This approach has the potential to facilitate more accurate diagnoses and more effective treatments for patients.

Keywords: leishmaniasis; photodiagnosis; *Leishmania amazonensis*; *Leishmania braziliensis*; FTIR spectra; machine learning.



LISTA DE ABREVIATURAS

DTN - Doença Tropical Negligenciada

OPAS - Organização Pan-Americana de Saúde

OMS - Organização Mundial de Saúde

WHO - World Health Organization

LV - Leishmaniose Visceral

LC - Leishmaniose Cutânea

LM - Leishmaniose Mucocutânea

FTIR - Espectroscopia de Infravermelho por Transformada de Fourier

IOC - Instituto Oswaldo Cruz

MLEE - *Multilocus Enzyme Electrophoresis*

PCR - *Polymerase Chain Reaction*

RFLP - *PCR-Restriction Fragment Length Polymorphism*



LISTA DE FIGURAS

Figura 1. Forma amastigota (A) e promastigota (B) do protozoário do gênero <i>Leishmania</i>	16
Figura 2. Lesões provocadas pela LC	26
Figura 3. Lesão causada por LM.....	28
Figura 4. Ciclo de vida dos parasitos do gênero <i>Leishmania</i>	31



SUMÁRIO

CAPÍTULO I

Revisão de Literatura:

1. Introdução	11
2. Doenças Tropicais Negligenciadas	12
3. Leishmanioses.	12
3.1 Leishmaniose Visceral	13
3.2 Leishmaniose cutânea e mucocutânea	14
3.3 Ciclo de vida de <i>Leishmania</i> sp	16
3.4 Diagnóstico e técnicas para identificação das espécies de <i>Leishmania</i>	17
4. Espectroscopia no infravermelho por transformada de fourier (FTIR)	20
5. OBJETIVOS	22
5.1 Objetivo geral	22
5.2 Objetivos específicos	22
6. REFERÊNCIAS	23

CAPÍTULO II:

ARTIGO: Easy identification of <i>L. amazonensis</i> and <i>L. braziliensis</i> species by using FTIR spectroscopy and machine learning algorithms	26
--	----



REVISÃO DE LITERATURA

1. Introdução

As leishmanioses são doenças infectoparasitárias zoonóticas neotrópicas que afetam principalmente as populações mais vulneráveis de quatro regiões ecoepidemiológicas do mundo: Américas, África Oriental, Norte de África e Ásia Ocidental e Sudeste Asiático (OPAS, 2022). Estas parasitoses são ocasionadas por protozoários parasitas do gênero *Leishmania*, transmitidos através da picada de insetos flebotomíneos. Devido à ampla variedade de espécies e à resposta imunológica do hospedeiro, pode-se observar um largo espectro de manifestações clínicas (Silveira; Lainson; Corbett, 2004), variando de lesões localizadas na pele (leishmaniose cutânea, LC), disseminação para a mucosa do trato respiratório superior (leishmaniose mucosa, LM), ou a forma sistêmica, que engloba os casos mais graves, associados à leishmaniose visceral (LV) (Altamirano-Enciso et al., 2003; WHO, 2019).

As leishmanioses compõem o grupo de Doenças Tropicais Negligenciadas (DTNs) e estão entre as principais doenças infectoparasitárias ainda em expansão mundial. São presentes principalmente em países subdesenvolvidos de clima tropical e subtropical, que oferecem condições ideais para seu vetor, o inseto flebotomíneo (Diptera: Psychodidae: Phlebotominae), conhecido popularmente como “mosquito palha”. Estes insetos têm hábitos noturnos e são encontrados principalmente em áreas florestais e de Cerrado. Com a expansão das áreas urbanas e trabalhos de extrativismo florestal, as leishmanioses tornaram-se presentes em áreas urbanas, onde a população mais exposta está, em sua maioria, em áreas pobres e com condições de saneamento básico precários.

Estima-se que um bilhão de pessoas vivam expostas à infecção em áreas endêmicas para a leishmaniose. Dados epidemiológicos apontam que ocorram 30.000 novos casos de LV e mais de um milhão de novos casos de LC anualmente (WHO, 2024). Entre o período de 2001 e 2021, um total de 1.105.545 casos de LC e LM foram notificados à Organização Pan-Americana da Saúde (OPAS, 2022), correspondendo a uma média de 52.645 casos por ano. No mesmo período, foram registrados 69.665



novos casos de LV, com média anual de 2.488 casos e taxa de letalidade de cerca de 8% (OPAS, 2023).

Como doença tropical negligenciada, existe baixo investimento em novos tratamentos, métodos diagnósticos e de identificação de espécies. Os atuais métodos diagnósticos são morosos e/ou onerosos o Sistema Único de Saúde, o que demanda a necessidade de desenvolvimento de novas metodologias para a obtenção de resultados que sejam mais rápidos, precisos e de custo mais acessível.

As técnicas de espectroscopia ótica surgiram recentemente como abordagem promissora no campo do diagnóstico, o que é de importância significativa no contexto das doenças infecciosas. A espectroscopia no infravermelho com transformada de Fourier (FTIR) tem sido abordada em diversos estudos para identificação de microrganismos, possibilitando diagnósticos precisos e com baixo custo, dois fatores primordiais para a saúde pública e para o combate às doenças negligenciadas.

2. Doenças Tropicais Negligenciadas

As doenças tropicais negligenciadas (DTNs) abrangem diversas patologias causadas por uma variedade de patógenos, como vírus, bactérias, parasitos, fungos e toxinas. Atingem principalmente a população mais pobre de áreas tropicais, onde as condições sanitárias são precárias e os indivíduos têm pouca ou nenhuma informação de prevenção ou tratamento. Tais enfermidades podem ter reflexos bastante nocivos em diversos aspectos no âmbito social, econômico e da saúde.

A OMS traz dados alarmantes de mortalidade e morbidade: 200.000 mortes e 19 milhões de vidas vêm sendo incapacitadas por consequência das DTNs. Os custos diretos com saúde, perda de produtividade e redução do nível socioeconômico e educacional equivale a bilhões de dólares; em contrapartida, o investimento em prevenção e tratamento por meio de políticas públicas e privadas é ainda muito baixo perto do necessário (WHO, 2024).

3. Leishmanioses

As leishmanioses representam um grupo de doenças emergentes e altamente negligenciadas que estão presentes nos cinco continentes. Elas são causadas por



protozoários parasitos do gênero *Leishmania*, transmitidos pela picada de insetos flebotomíneos fêmeas que, durante o repasto sanguíneo, inoculam via regurgitação o parasito em animais e humanos. Estima-se que ocorram 1,3 milhão de novos casos e de 20.000 a 30.000 mortes por ano. As formas clínicas visceral, cutânea e mucocutânea representam uma preocupação significativa para a saúde pública nas Américas. A falta de investimento em novos tratamentos, prevenção, métodos de diagnóstico e identificação de espécies ou cepas representa um desafio significativo para o controle dessas doenças (OPAS, 2022).

A leishmaniose cutânea, enfermidade parasitária de relevância médico-sanitária, não é um fenômeno restrito à modernidade. Pelo contrário, sua presença entre populações humanas remonta a períodos históricos bastante remotos, conforme atestam evidências arqueológicas, iconográficas e paleogenéticas. Há indícios substanciais de que a doença já era conhecida – ainda que de forma empírica e não sistematizada – por civilizações da Antiguidade.

Entre as evidências mais emblemáticas estão aquelas provenientes de povos pré-incaicos estabelecidos nas regiões do atual Equador e Peru, onde cerâmicas datadas do século I d.C. retratam figuras humanas com deformidades faciais e lesões ulceradas, compatíveis com manifestações cutâneas típicas da leishmaniose. Tais representações sugerem não apenas a ocorrência da doença, mas também sua notoriedade no imaginário coletivo dessas populações (Altamirano-Enciso et al., 2003).

Registros datados do século VII a.C., provenientes de tábuas mesopotâmicas pertencentes à biblioteca do rei assírio Assurbanípal, descrevem lesões dérmicas compatíveis com a “oriental sore” (Lima et al., 2017). Entre 2050 e 1650 a.C., exames paleogenéticos realizados em múmias egípcias da era do Médio Reino confirmaram a presença de DNA mitocondrial de *Leishmania donovani*, atestando a ocorrência da forma visceral da doença naquela civilização (Zink et al., 2006). O *Papiro de Ebers* (c. 1500 a.C.) menciona uma afecção cutânea denominada “espinha do Nilo” (*Nile pimple*), possivelmente associada à leishmaniose cutânea (Zink et al., 2017).

No continente americano, há registros arqueológicos e iconográficos entre povos pré-incaicos da região andina. Em múmias peruanas datadas de 800 a.C., foi detectada, por análise imuno-histológica, a presença de *Leishmania*, com manifestações cutâneas visíveis (Guillen & Allison, 2005). Complementarmente, cerâmicas rituais datadas do século I d.C., encontradas no Equador e Peru, representam figuras humanas com lesões ulceradas faciais, compatíveis com a forma cutânea da doença (Altamirano-Enciso et al., 2003).

O reconhecimento científico da leishmaniose como uma entidade parasitária ocorreu apenas no final do século XIX. Em 1885, o patologista William Leishman, na Índia, descreveu pela primeira vez o agente etiológico da leishmaniose visceral ao observar amastigotas do parasito em esfregaços de baço de um soldado britânico infectado. No mesmo período, Cunningham já havia registrado observações semelhantes de protozoários em amostras humanas, ainda que sem a correta taxonomia do gênero *Leishmania* (Basano & Camargo, 2004). No Brasil, a primeira descrição clínica da leishmaniose cutânea ocorreu em 1855, sendo então denominada “botão de Biskra” - uma alusão à forma ulcerada da doença observada no Norte da África, especialmente na Argélia. Essa terminologia foi utilizada por médicos brasileiros à época, refletindo a influência da medicina europeia na caracterização de doenças tropicais (Basano & Camargo, 2004).

Mesmo se tratando de uma doença que possui relatos tão antigos, ainda na atualidade as leishmanioses representam um grande problema de saúde pública, principalmente nas Américas, África Oriental, Norte da África e Oeste e Sudeste Asiático (OPAS, 2024).

O protozoário parasito, pertencente à Ordem Kinetoplastida e à Família Trypanosomatidae, possui ciclo de vida heteroxênico, envolvendo as formas evolutivas promastigota e amastigota, presentes, respectivamente, nos hospedeiros invertebrados (vetores) e vertebrados (Figura 1).

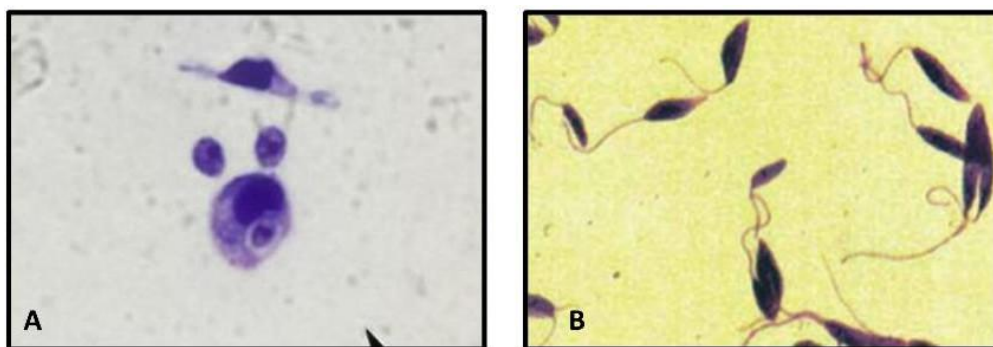


Figura 1. Formas amastigotas (A) e promastigotas (B) de protozoários do gênero *Leishmania*.
Fonte: A- autor e B- Brasil (2017).

3.1 Leishmaniose Visceral

A leishmaniose visceral é a forma clínica mais grave entre as leishmanioses. É uma doença sistêmica que causa febre, aumento do volume do baço e do fígado e,



quando não tratada, pode levar ao óbito em 90% dos casos. Acomete principalmente crianças abaixo de cinco anos e adultos com mais de 50 anos e tem alta incidência em indivíduos imunossuprimidos. Foi registrada em 13 países das Américas, mas 93% dos casos ocorreram no Brasil (OPAS, 2022). Nos últimos cinco anos, foram registrados cerca de 2.850 casos de LV, com letalidade de 8,2%.

A principal espécie causadora da leishmaniose visceral no Brasil é a *Leishmania (Leishmania) infantum*.

3.2 Leishmaniose cutânea e mucocutânea

As formas clínicas cutânea e mucocutânea são tradicionalmente classificadas como leishmaniose tegumentar (LT), que se apresenta como lesões ulcerativas na pele e nas membranas mucosas. A LT representa um importante problema em saúde pública em 89 países das Américas, Europa, África e Ásia, com uma estimativa de 1 a 1,5 milhão de casos documentados anualmente (WHO, 2024). Em 2021, 55 países notificaram à OMS 222.000 novos casos (OPAS, 2022).

A OMS considera a LT como uma das seis doenças infecciosas mais importantes devido à sua alta prevalência e capacidade de causar deformidades, juntamente com seu impacto psicológico e consequências sociais e econômicas. Dos nove países do mundo com maior número de casos, três estão nas Américas: Brasil, Colômbia e Peru. Dados epidemiológicos demonstram o registro de 52.645 casos de LC e LM nos últimos 20 anos. Em 2021, foram notificados 37.786 casos de LC, que está presente em 21 países e é considerada endêmica em 19 (OPAS, 2022).

A LC é uma das formas clínicas com maior incidência no Brasil, estando amplamente distribuída, com casos documentados em todas as regiões. Essa doença é caracterizada pelas lesões que provoca na pele (Fig. 1), deixando consequências físicas e psicológicas, além do impacto social e profissional na vida dos acometidos (BRASIL, 2017). No Brasil, os principais agentes etiológicos são *Leishmania (Viannia) braziliensis*, *L. (V.) guyanensis* e *L. (Leishmania) amazonensis* (Brasil, 2017).

Há alguns anos, a LT era uma enfermidade quase que exclusiva de pessoas que sobreviviam do extrativismo florestal e adentravam as florestas para realização dos trabalhos a eles empregados, porém hoje essa já não é uma doença que atinge com exclusividade essa parcela da população (Berton et al. 2023).



Figura 2. Lesões provocadas pela LC. Fonte: Brasil (2017).

A leishmaniose mucosa (LM) é a manifestação clínica que ocorre quando o parasito se dissemina pelas mucosas oral, nasal, laríngea e faríngea. A cura espontânea da LM é rara e sua progressão causa deformidades faciais graves (Fig. 2), podendo evoluir para óbito (Bezemer, et al., 2022).



Figura 3 - Lesão causada por LM. Fonte: Brasil (2017).

3.3 CICLO DE VIDA DE *Leishmania* sp

O ciclo desse parasito (Figura 3) é bastante complexo e envolve várias espécies de hospedeiros. A leishmaniose é uma zoonose complexa, causada por protozoários do gênero *Leishmania*, cuja transmissão ocorre por meio da picada de insetos flebotomíneos fêmeas infectadas. O ciclo de vida do parasita compreende duas fases

principais: a promastigota (no vetor) e a amastigota (no hospedeiro vertebrado). A transmissão entre hospedeiros ocorre de forma heteroxênica, envolvendo ao menos um vetor invertebrado e um reservatório vertebrado.

No Brasil, os vetores pertencem majoritariamente ao gênero *Lutzomyia*, enquanto nas regiões do Velho Mundo predominam espécies do gênero *Phlebotomus*. Esses insetos têm hábitos crepusculares e noturnos, e seus ciclos reprodutivos estão fortemente associados a ambientes quentes e úmidos, com matéria orgânica em decomposição (Brasil, 2022).

Além dos vetores, uma diversidade de vertebrados atua como reservatórios naturais da leishmaniose. Em áreas urbanas, os cães domésticos (*Canis lupus familiaris*) representam a principal fonte de infecção para a forma visceral, causada por *Leishmania infantum*. Em áreas silvestres, há uma ampla gama de mamíferos envolvidos na manutenção do ciclo enzoótico, como raposas, tatus, primatas e roedores (Basano & Camargo, 2004).

A seguir, a Tabela 1 resume as principais espécies envolvidas no ciclo de transmissão da leishmaniose no Brasil e em outras regiões endêmicas, evidenciando os agentes vetores, os principais reservatórios e a espécie de *Leishmania* comumente associada.

Tabela 1 – Espécies envolvidas no ciclo de vida da leishmaniose

Categoria	Espécie	Função no ciclo	Espécie de <i>Leishmania</i> associada
Vetores (flebotomíneos)	<i>Lutzomyia longipalpis</i>	Principal vetor urbano	<i>L. infantum</i> (visceral)
	<i>Lutzomyia whitmani</i>	Vetor silvestre e periurbano	<i>L. braziliensis</i> (cutânea)
	<i>Lutzomyia intermedia</i>	Vetor periurbano	<i>L. braziliensis</i>
	<i>Lutzomyia flaviscutellata</i>	Vetor amazônico	<i>L. amazonensis</i> (cutânea difusa)
	<i>Phlebotomus papatasi</i>	Vetor em regiões da África e Oriente Médio	<i>L. major</i> (cutânea)
Reservatórios domésticos	<i>Canis lupus familiaris</i> (cão)	Principal reservatório urbano	<i>L. infantum</i>
Reservatórios silvestres	<i>Cercopithecus thomasi</i> (raposa)	Reservatório silvestre	<i>L. infantum</i> , <i>L. braziliensis</i>
	<i>Dasypus novemcinctus</i> (tatu)	Hospedeiro natural	<i>L. braziliensis</i>
	<i>Callithrix</i> spp., <i>Cebus</i> spp.	Primatas com infecção natural	Diversas espécies
	Roedores (<i>Proechimys</i> , <i>Rattus</i>)	Reservatórios silvestres	<i>L. amazonensis</i>
Hospedeiro acidental	<i>Homo sapiens</i> (ser humano)	Acidental (ou terminal, em zonas urbanas)	Todas as formas (cutânea e visceral)

Fonte: adaptado de Brasil (2022); Basano e Camargo (2004); WHO (2023).

Para explicar esse ciclo, iniciamos pela picada do flebotomíneo fêmea que, durante o repasto sanguíneo, regurgita formas promastigotas metacíclicas sobre o local da picada, permitindo que o parasito penetre no organismo do hospedeiro. Estas formas possuem maior potencial de infecção devido às alterações bioquímicas ocorridas durante a metaciclogênese, que resultam na perda de adesão ao epitélio do flebotomíneo. Uma vez no organismo do hospedeiro, as formas promastigotas são fagocitadas pelos macrófagos, onde se diferenciam em formas amastigotas (BRASIL, 2017; Burza; Croft; Boelaert, 2018).

Após a diferenciação, as amastigotas se multiplicam por divisão binária até o rompimento da célula hospedeira, permitindo que as amastigotas infectem novas células e continuem o desenvolvimento da patogenia. Durante esse processo, um flebotomíneo pode se infectar ao ingerir sangue contendo as formas amastigotas. No intestino do inseto, as amastigotas se diferenciam em promastigotas, que também se multiplicam por divisão binária; migram, então, para a probóscide do flebotomíneo, e doravante terão o potencial de infectar novos hospedeiros (Burza; Croft; Boelaert, 2018).

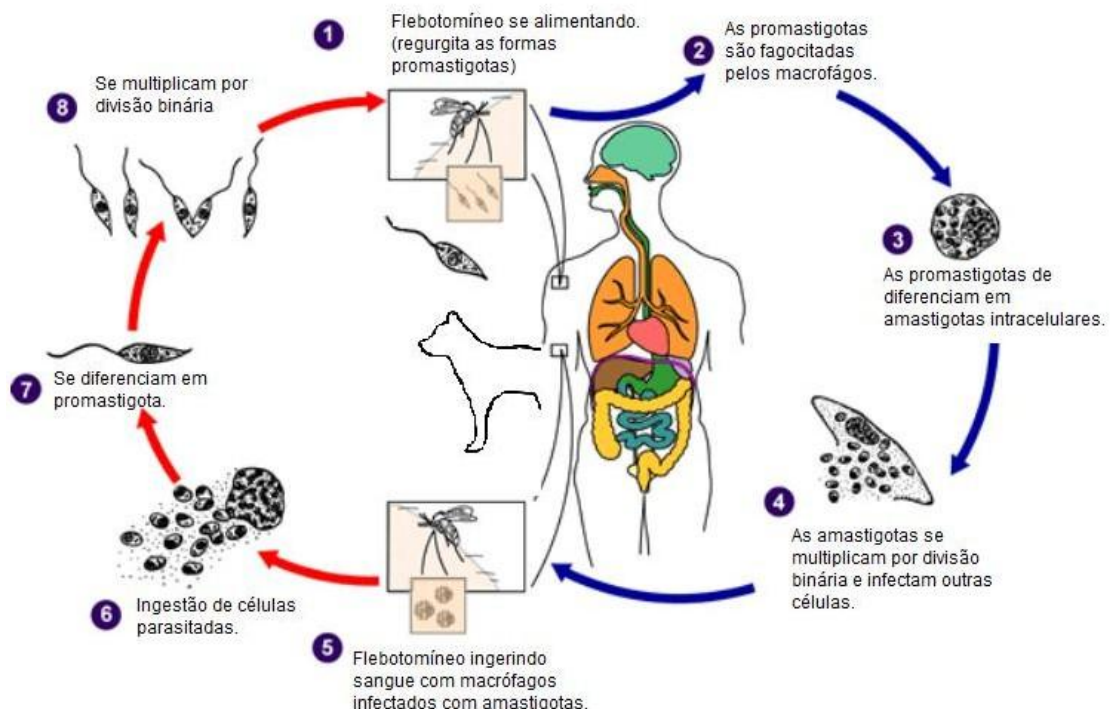


Figura 2. Ciclo de vida dos parasitos do gênero *Leishmania*.

Fonte: Adaptado de CDC, <https://www.cdc.gov/parasites/leishmaniasis/biology.html>.



3.4 Diagnóstico e técnicas para identificação das espécies de *Leishmania*

O diagnóstico das leishmanioses é um desafio devido ao amplo espectro de manifestações clínicas presentes nos pacientes, com sobreposição de sintomas com doenças como tuberculose, febre tifoide e malária, cujos tratamentos são muito diferentes. Além disso, as técnicas parasitológicas e sorológicas convencionais para o diagnóstico das leishmanioses possuem limitações (de Paiva-Cavalcanti et al. 2015), com desempenho variável dependendo da forma clínica e tempo de evolução da doença, espécie de *Leishmania* envolvida, carga parasitária, presença de infecções secundárias, forma de coleta e tipo de material biológico utilizado, além de estrutura laboratorial e corpo técnico disponível para análise. Além disso, todos os resultados dos exames deverão ser interpretados pelo médico, que indicará a conduta terapêutica correta para cada paciente.

Estas limitações levaram à investigação e aplicação de métodos moleculares, dada sua sensibilidade e especificidade (Azizi et al., 2012; Tlamcani et al., 2016). Porém, devido à operacionalização, estas funcionam de forma complementar e alternativa ao diagnóstico.

As técnicas usuais de diagnóstico das leishmanioses são os exames parasitológicos direto e indireto. Os exames parasitológicos baseiam-se em técnicas de exame direto e consistem na visualização microscópica das formas amastigotas no material obtido por escarificação, aspiração ou biópsia da lesão cutânea ou mucosa (para a LC e LCM) e no aspirado de medula óssea (para a LV), sendo o material biológico obtido corado por Giemsa ou Leishman após realização de esfregaço em lâmina (Andrade et al., 2005). As formas amastigotas de *Leishmania* também podem ser visualizadas em biópsias de lesões enviadas para avaliação histopatológica. O parasitológico indireto se dá partir do cultivo do parasito proveniente do material obtido, *in vitro*, em meio ágar-sangue de Novy e McNeal modificado por Nicolle (NNN) e permite a visualização das formas promastigotas após um período de crescimento no meio, possibilitando a identificação posterior da espécie (Brasil, 2017).

O isolamento de cepas de *Leishmania* ssp. a partir do diagnóstico parasitológico indireto permite a identificação da espécie em causa no caso clínico, contribuindo para a vigilância e o controle epidemiológico da parasitose e fortalecendo as linhas de pesquisa. A identificação de espécies e cepas circulantes do parasito



permite o mapeamento geográfico dos agentes etiológicos das leishmanioses nas Américas, contribuindo para o conhecimento de vários aspectos de sua epidemiologia. De acordo com a OPAS (IOC/FIOCRUZ, 2023), a identificação de espécies de *Leishmania* é importante em situações de surtos, novos focos de transmissão, focos endêmicos sem conhecimento prévio das espécies circulantes, focos endêmicos que apresentam uma situação epidemiológica não usual, e em situações clínicas especiais em que a identificação das espécies pode auxiliar no manejo clínico de, por exemplo, pacientes imunodeprimidos, em quadros clínicos atípicos (Alves et al., 2024).

Desta forma, uma vez que as leishmanioses constituem um grupo de doenças negligenciadas de grande incidência no Brasil, é crucial a identificação específica, dado o amplo espectro clínico relacionado às espécies causadoras. Além disso, a caracterização da espécie permite melhor abordagem terapêutica para tratamento mais adequado e controle da enfermidade.

Ao longo do tempo, foram desenvolvidos vários métodos de identificação, como a eletroforese enzimática multilocus (MLEE) e a eletroforese em gel de campo pulsado, além dos ensaios baseados na reação em cadeia da polimerase (PCR).

A análise isoenzimática de parasitos cultivados é a abordagem convencional para a caracterização taxonômica de *Leishmania* spp. O padrão ouro é a MLEE, que requer a cultura e o isolamento de parasitos e pode levar semanas ou meses para produzir resultados (Rioux et al., 1990). A MLEE permite a identificação de microrganismos por meio da mobilidade eletroforética de várias enzimas intracelulares (Ovalle-Bracho et al. 2018). As isoenzimas desempenham grande papel nos processos metabólicos em diferentes componentes celulares e são produzidas por vários genes. O tipo típico ou o padrão de mobilidade específico da espécie pode ser confirmado utilizando um conjunto definido de isoenzimas. A MLEE é considerada um método de referência para a identificação de espécies de *Leishmania* e pode também discriminar entre espécies do Velho Mundo e do Novo Mundo. No entanto, não permite a diferenciação de subpopulações devido a um menor número de marcadores, homoplasia ou heterozigossidade de genes destinados que causam mais de um esquema de mobilidade (Jamjoom et al. 2004).

Apesar de ser de grande importância para estudos epidemiológicos, a MLEE não é rápida o suficiente para orientar decisões terapêuticas de primeira linha. Assim, na maioria dos casos, os médicos tomam decisões iniciais de tratamento com base



apenas em dados clínicoepidemiológicos. Além disso, sua operacionalização é complexa devido à necessidade de equipamento especializado, à necessidade de cultivo dos parasitos e à falha na discriminação de algumas espécies e populações (Schönian et al. 2011).

A PCR e as suas modificações, como a PCR quantitativa (qPCR), a *semi-nested-PCR* e a *nested-PCR* vêm sendo utilizadas no diagnóstico das leishmanioses utilizando várias amostras e regiões-alvo (da Silva Solcà et al., 2012; Reis et al., 2013). Inicialmente, a técnica de qPCR havia sido utilizada para discriminar o DNA o agente causador em diferentes amostras humanas e para estimar quantitativamente a carga parasitária em indivíduos infectados (de Paiva-Cavalcanti et al., 2015).

A caracterização das espécies de *Leishmania* é também uma função da PCR (Mohammadiha et al., 2013), sendo possível identificar e diferenciar as espécies de *Leishmania* que causam a infecção usando *primers* específicos e genes-alvo, incluindo ITS (DNA espaçador não codificante), sequências repetitivas de DNA nuclear, genes do citocromo b, mini-exon, G6PD, cpb, gp63 e hsp70. Entretanto, o alto custo e a necessidade de laboratórios bem equipados e pessoal qualificado são desvantagens da técnica (Foulet et al., 2007).

O emparelhamento da PCR com diferentes metodologias, como a análise do polimorfismo de comprimento de fragmentos de restrição (RFLP) e o sequenciamento genético, tem ajudado na identificação de espécies em investigações epidemiológicas (Wang et al. 2011). Entretanto, as mesmas desvantagens se aplicam à identificação por sequenciamento genético, que inclui muitas abordagens de software para identificar genes codificadores de proteínas e RNA (Myler, Stuart, 2000).

Em vista disso, é importante estudar e aplicar novas técnicas para identificar espécies de *Leishmania*, pois elas estão diretamente relacionadas ao manejo terapêutico e no direcionamento de medidas de saúde pública com base em dados epidemiológicos.

4.ESPECTROSCOPIA NO INFRAVERMELHO POR TRANSFORMADA DE FOURRIER (FTIR)



As técnicas de espectroscopia óptica surgiram recentemente como uma abordagem promissora no campo do diagnóstico de doenças infecciosas. Essas técnicas permitem a análise de espectros obtidos de amostras biológicas, fornecendo resultados precisos de forma não invasiva. Essa abordagem oferece a possibilidade de reduzir o tempo necessário para o diagnóstico, o que é significativo no contexto das doenças infecciosas (Naseer et al., 2019; Martinez-Cuazitl et al., 2021).

A espectroscopia de infravermelho com transformada de Fourier (*Fourier-transform infrared spectroscopy*, FTIR) tem sido empregada em vários estudos para identificar microrganismos, permitindo diagnósticos precisos a um baixo custo, dois fatores cruciais para a saúde pública e o combate a doenças negligenciadas (Larios et al., 2021; Pacher et al., 2024). A FTIR é um método analítico com amplas possibilidades de aplicação, sendo muito utilizado nas indústrias e meios acadêmicos, pois possibilita a compreensão das estruturas moleculares do analito, que pode ser desde produtos sintéticos a amostras biológicas (Cameron et al., 2018).

A FTIR usa energia de infravermelho médio para analisar a estrutura molecular e a composição de substâncias, medindo a absorção de frequências específicas de radiação infravermelha e fornecendo uma impressão digital exclusiva para identificação e análise das amostras. A radiação infravermelha interage quando atinge frequências de energia vibracionais específicas das ligações de grupos funcionais presentes na molécula, revelando assim características específicas na composição da amostra analisada. Esse método oferece uma abordagem rápida, precisa, minimamente invasiva e econômica para identificar as espécies de micro-organismos (Baker et al., 2016). Essa tecnologia tem se mostrado muito promissora, pois é necessário um manuseio mínimo de amostras para obtenção de resultados (Maquelin et al., 2002). Tem o potencial de beneficiar os pacientes, pois permite a seleção de um tratamento adequado, além de contribuir significativamente para a coleta de dados epidemiológicos (Baker et al., 2016).

A FTIR tem sido usada em estudos para identificar espécies de *Leishmania*. Aguiar et al. (2015) diferenciaram as espécies *L. (L.) amazonensis*, *L. (L.) chagasi* (sin. *L. (L.) infantum*) e *L. major* por meio da deconvolução das bandas por ajuste de curva nessas regiões, diferenciando quantitativamente polissacarídeos, amida III, fosfolipídios, proteínas e ácidos nucleicos. As espécies apresentaram variações significativas, que contribuíram para sua diferenciação. Hornemann et al. (2017)



diferenciaram espécies de *Leishmania* do Velho Mundo (*L. major*, *L. tropica*, *L. infantum* e *L. donovani*) e híbridos de espécies diferentes usando impressões digitais espectroscópicas vibracionais.

Estudo realizado em 2018 para tentar separar e identificar espécies de *Trypanosoma cruzi*, *Trypanosoma rangeli* e *Leishmania infantum* apresentou resultados bastante promissores, onde a análise multivariada e da curva de ajuste gaussiana foram suficientes para distinguir as espécies (Aguiar et al., 2018).

A FTIR vem sendo combinada com algoritmos de aprendizagem de máquina (*machine learning*, ML) para fornecer resultados rápidos e precisos para muitos diagnósticos (Marangoni-Ghoreyshi et al., 2023; Pacher et al., 2024). A análise de componentes principais (PCA) é uma técnica multivariada usada para gerar índices e agrupar indivíduos de acordo com a variação em suas características, facilitando a análise de padrões entre eles para melhor visualizar as variáveis que contribuem para a separação entre grupos (Hongyu et al., 2016). O *Support Vector Machine* (SVM), por outro lado, é um algoritmo matemático que consiste em reconhecer padrões, separar os dados em regiões e determinar os riscos juntamente com os intervalos de confiança para sua separação. O SVM cria um plano usando diferentes funções para separar dois grupos; o SVM linear calcula a média dos dois grupos para a tomada de decisão (Ferrão et al., 2007).

O objetivo do presente estudo é discriminar espécies de *Leishmania* causadoras de leishmaniose cutânea e mucocutânea, provenientes do cultivo axênico de cepas de referência, a partir da análise multivariada de espectros de absorção no infravermelho, identificando as contribuições espectrais mais importantes para a identificação e associando-as aos componentes biológicos das amostras.

5. OBJETIVOS

5.1. OBJETIVO GERAL: desenvolver uma nova metodologia/protocolo para a diferenciação e identificação de espécies de *Leishmania* spp. cultivadas *in vitro*, a partir da análise multivariada de espectros de absorção no infravermelho.

5.2. OBJETIVOS ESPECÍFICOS:



- discriminar espécies de *Leishmania* spp. isoladas a partir das culturas de cepas de referência de *Leishmania* sp. mantidas no Laboratório de Parasitologia Humana/INBIO/UFMS.
- identificar as contribuições espectrais mais importantes para a identificação das espécies/cepas de *Leishmania* spp. e associá-las aos componentes biológicos das amostras.



6. 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, 2018. doi: 10.1016/j.exppara.2018.07.008.
- AGUIAR, J. C. et al. Differentiation of *Leishmania* species by FT-IR spectroscopy. **Spectrochim Acta A Mol Biomol Spectrosc.**, vol.5, n. 142, p. 80-5, 2015. doi: 10.1016/j.saa.2015.01.008.
- ALTAMIRANO-ENCISO, A. J. et al. Sobre a origem e dispersão das leishmanioses cutânea e mucosa com base em fontes históricas pré e pós colombianas. **História, Ciências, Saúde Manguinhos**, v. 10, n. 3, p. 853-88, 2003.
- ALVES, N. O. et al. VL-HIV co-infection with *Leishmania* containing skin lesions resembling para-kala-azar dermal leishmaniasis. **PLoS Neglected Tropical Diseases**, v. 18, n. 8, e-0012438, 2024.
- ANDRADE, B. B. et al. Métodos diagnósticos da Leishmaniose Tegumentar: fatos, falácias e perspectivas. **Gazeta Médica da Bahia**, v. 75, n. 1, p. 75-82, 2005.
- AZIZI, K.; SOLTZNI, A.; ALIPOUR, H. Molecular detection of *Leishmania* isolated from cutaneous leishmaniasis patients in Jask County, Hormozgan Province, Southern Iran, 2008. **Asian Pac J Trop Med.**, v. 5, n.7, p. 514-517, 2012.
- BAKER, M. J. et al. Using Fourier transform IR spectroscopy to analyze biological materials. **Nature protocols**, v. 9, n. 8, p. 1771-1791, 2014.
- BAKER, M. J., et al. Developing and understanding biofluid vibrational spectroscopy: a critical review. **Chem. Soc. Rev.**, v. 45, p. 1803–1818, 2016.
- BASANO, S. 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, 2004.
- BERTON, N. C. et al. Prevalence of American Tegumentary Leishmaniasis in Brazil - A systematic review and meta-analysis. **Revista Contexto & Saúde**, v. 23, n.47, e13946, 2023.
- 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, 2022.
- BRASIL. Ministério da Saúde. Secretaria de Vigilância em Saúde. Departamento de Vigilância das Doenças Transmissíveis Ministério da Saúde. **Manual de vigilância da leishmaniose tegumentar**. Brasília: Ministério da Saúde, 2017.
- BURZA, S.; CROFT, S. L.; BOELAERT, M. Leishmaniasis. **Lancet**, v. 392, p.951-970, 2018. Disponível em:10.1016/S0140-6736(18)31204-2.



CAMERON, J. M. et al. Biofluid spectroscopic disease diagnostics: A review on the processes and spectral impact of drying. **J. Biophoton.**, v. 11, e201700299, 2018. <https://doi.org/10.1002/jbio.201700299>.

SOLCÀ, M. S. et al. Qualitative and quantitative polymerase chain reaction (PCR) for detection of *Leishmania* in spleen samples from naturally infected dogs. **Vet Parasitol.**, v. 184, n. 2-4, p.133-140, 2012.

de PAIVA-CAVALCANTI M., et al. Leishmaniasis diagnosis: an update on the use of immunological and molecular tools. **Cell Biosci.**, v. 5, n. 1, p. 31, 2015.

FERRÃO, M. F. et al. LS-SVM: uma nova ferramenta quimiométrica para regressão multivariada. Comparação de modelos de regressão LS-SVM e PLS na quantificação de adulterantes em leite em pó empregando NIR. **Quim. Nova**, v. n. 30, p. 852-859, 2007.

FOULET, F. et al. Detection and identification of *Leishmania* species from clinical specimens by using a real-time PCR assay and sequencing of the cytochrome B gene. **J. Clin. Microbiol.**, v. 45, p. 2110-2115, 2007.

GUILLEN, S.; ALLISON, M. An early case of South American *Leishmaniasis* in Peru. In: *1st Paleopathology Association Meeting in South America*, 2005. p. 61.

HORNEMANN, A. et al. A pilot study on fingerprinting *Leishmania* species from the Old World using Fourier transform infrared spectroscopy. **Anal. Bioanal. Chem.**, v. 409, p. 6907-6923, 2017.

HONGYU, K.; SANDANIELO, V. L. M.; De OLIVEIRA JR, G. J. Análise de componentes principais: resumo teórico, aplicação e interpretação. **E&S Engin. & Sci.**, v. 5, p. 83-90. 2016.

INSTITUTO OSWALDO CRUZ, COLEÇÃO DE LEISHMANIA (CLIOC). Programa Regional de Leishmanioses - **OPAS/OMS**. Disponível em: <http://clioc.fiocruz.br/index?opas>. Acesso em: 28/03/2023.

JAMJOOM, M. B. et al. *Leishmania donovani* is the only cause of visceral leishmaniasis in East Africa; previous descriptions of *L. infantum* and "*L. archibaldi*" from this region are a consequence of convergent evolution in the isoenzyme data. **Parasitology**, v. 129, n. 4, p. 399-409, 2004.

LIMA, J. A. et al. História da leishmaniose: uma análise histórica e científica. *Revista Interdisciplinar*, v. 10, n. 1, p. 73-84, 2017.

LARIOS G. et al. A new strategy for canine visceral leishmaniasis diagnosis based on FTIR spectroscopy and machine learning. **J Biophotonics**, v. 14, n. 11, e202100141, 2021. doi: 10.1002/jbio.202100141.

MARANGONI-GHOREYSHI, Y. G. et al. Multi-resistant diarrheagenic *Escherichia coli* identified by FTIR and machine learning: a feasible strategy to improve the group classification. **RSC Adv.**, v. 13. 2023, Doi: 10.1039/d3ra03518b.

MARTINEZ-CUAZITL, A. et al. ATR-FTIR spectrum analysis of saliva samples from COVID-19 positive patients. **Sci. Rep.**, v. 11, 19980, 2021.

MYLER, P. J.; STUART, K. D. Recent developments from the *Leishmania* genome



Fundação Universidade Federal de Mato Grosso do Sul
Faculdade de Ciências Farmacêuticas, Alimentos e Nutrição
Programa de Pós-Graduação em Ciências Farmacêuticas



project. **Cur. Opin. Microbiol.**, v. 3, p. 412-416, 2000.

ZINK, A. R. et al. Leishmaniasis in Ancient Egypt and Upper Nubia. *Emerging Infectious Diseases*, v. 12, n. 10, p. 1616-1617, 2006.

ZINK, A. R. et al. Ancient *Leishmania donovani* DNA in Egyptian Middle Kingdom mummies (2050-1650 BC). *Parasites & Vectors*, v. 10, n. 1, p. 1-9, 2017.



- MOHAMMADIHA, A. et al. Comparison of real-time PCR and conventional PCR with two DNA targets for detection of *Leishmania (Leishmania) infantum* infection in human and dog blood samples. **Exp Parasitol.**, v. 133, n. 1, p. 89-94, 2013.
- NASEER, K. et al. FTIR spectroscopy of freeze-dried human sera as a novel approach for dengue diagnosis. **Infrared Phys. Technol.**, v. 102, 102998, 2019.
- OVALLE-BRACHO, C. et al. Molecular typing of *Leishmania (Leishmania) amazonensis* and species of the subgenus Viannia associated with cutaneous and mucosal leishmaniasis in Colombia: a concordance study. **Biomédica**, v. 38, n. 1, p. 86-95, 2018.
- Organização Pan-Americana de Saúde (OPAS). **Diretrizes para o tratamento das leishmanioses na Região das Américas**. 2ª edição, 2022. Disponível em: <https://doi.org/10.37774/9789275725030>. Acesso em: 12 jan. 2024.
- Organização Pan-Americana da Saúde (OPAS). **Leishmanioses: Informe epidemiológico das Américas**. No 11 (Dezembro de 2022) - OPAS/OMS. Disponível em: <https://www.paho.org/pt/documentos/leishmanioses-informe-epidemiologico-das-americas-no-11-dezembro-2022>. Acesso em: 20 mar. 2024.
- Organização Pan-Americana da Saúde (OPAS). **Leishmanioses: Informe epidemiológico das Américas**. N.12 (dezembro de 2023), 2024. Disponível em: <https://iris.paho.org/handle/10665.2/59170>. Acesso em: 12 jan. 2024.
- Organização Pan-Americana da Saúde (OPAS). Manual of procedures for leishmaniasis surveillance and control in the Region of the Americas. 2ª edition, 2024. Disponível em: <https://iris.paho.org/handle/10665.2/51838>. Acesso em: 10 mar. 2024.
- PACHER G. et al. Diagnosis of Cutaneous Leishmaniasis Using FTIR Spectroscopy and Machine Learning: An Animal Model Study. **ACS Infect Dis.** v. 9;10, n. 2, p. 467-474, 2024. doi: 10.1021/acsinfectdis.3c00430.
- RIOUX, J. A. et al. Taxonomy of *Leishmania*. Use of isoenzymes. Suggestions for a new classification. **Annales de parasitologie humaine et comparee**, v. 65, p. 111-125, 1990.
- SCHÖNIAN, G.; KUHLS K.; MAURICIO I. L. Molecular approaches for a better understanding of the epidemiology and population genetics of *Leishmania*. **Parasitology**, v.138, n. 4, p. 405-425, 2011.
- SILVEIRA, F. T.; LAINSON, R.; CORBETT, C. E. P. Clinical and immunopathological spectrum of American cutaneous leishmaniasis with special reference to the disease in Amazonian Brazil: a review. **Mem. Inst. Oswaldo Cruz**, v. 99, p. 239-251, 2004.
- TLAMCANI, Z. Visceral leishmaniasis: an update of laboratory diagnosis. **Asian Pac J Trop Dis.**, v. 6, n. 7, p. 505-508, 2016.



Fundação Universidade Federal de Mato Grosso do Sul
Faculdade de Ciências Farmacêuticas, Alimentos e Nutrição
Programa de Pós-Graduação em Ciências Farmacêuticas



WANG J. Y. et al. The prevalence of canine *Leishmania infantum* infection in western China detected by PCR and serological tests. **Parasit Vector.**, v. 4, n. 1, p. 69, 2011.

World Health Organization (WHO). **Leishmaniasis**. Disponível em: <https://www.who.int/news-room/fact-sheets/detail/leishmaniasis>. Acesso em 20/09/2024.



1. MANUSCRITO SUBMETIDO AO PERIÓDICO ACS INFECTIOUS DISEASES

ACS Infectious Diseases

This document is confidential and is proprietary to the American Chemical Society and its authors. Do not copy or disclose without written permission. If you have received this item in error, notify the sender and delete all copies.

Easy identification of *L. amazonensis* and *L. braziliensis* species by using FTIR spectroscopy and machine learning algorithms

Journal:	<i>ACS Infectious Diseases</i>
Manuscript ID	id-2025-00435h
Manuscript Type:	Article
Date Submitted by the Author:	20-May-2025
Complete List of Authors:	Oliveira, Vilma; UFMS Fernandes, Vitoria; UFMS Silva, Fernanda; Universidade Federal de Mato Grosso do Sul, Faculdade de Ciências Farmacêuticas, Alimentos e Nutrição Franca, Thiago; UFMS, Physics Institute Calvani, Camila; UFMS, Physics Institute Marangoni, Bruno; UFMS, Physics Institute Pinto de Arruda, Carla Cardozo; Universidade Federal de Mato Grosso do Sul Cena, Cicero; UFMS, Physics Institute

SCHOLARONE™
Manuscripts

Easy identification of *L. amazonensis* and *L. braziliensis* species by using FTIR spectroscopy and machine learning algorithms

Vilma A.S. Oliveira¹, Vitoria S. Fernandes¹, Fernanda Silva¹, Thiago Franca², Camila
Calvani², Bruno Marangoni², Carla Arruda^{1*}, Cicero Cena^{2*}
*carla.arruda@ufms.br; *cicero.cena@ufms.br

¹UFMS – Universidade Federal de Mato Grosso do Sul, Laboratório de Parasitologia Humana,
Instituto de Biociências, Av. Costa e Silva s/n, 79070-900, Campo Grande-MS, Brazil.
²UFMS – Universidade Federal de Mato Grosso do Sul, Laboratório de Fotodiagnóstico, Instituto
de Física, Av. Costa e Silva s/n, 79070-900, Campo Grande-MS, Brazil.

ABSTRACT

Leishmaniasis is a neglected tropical disease requiring accurate species identification to ensure proper clinical management and epidemiological surveillance. Accurate species and strain identification depends on molecular and biochemical tools. While these conventional techniques are effective, they are often costly, time-consuming, and inaccessible in low-resource settings. In this study, we evaluated the potential of Fourier-transform infrared (FTIR) spectroscopy, combined with machine learning algorithms, for the discrimination of *Leishmania amazonensis* and *Leishmania braziliensis* species in liquid cultures. FTIR spectra were acquired from 80 culture samples and preprocessed using standard normal variate (SNV) correction and Fast Fourier Transform (FFT) filtering. Principal Component Analysis (PCA) revealed clear species-specific clustering driven by spectral differences in protein, lipid, and nucleic acid vibrational bands. Support Vector Machine (SVM) models were trained using PCA scores, achieving over 90% accuracy in all tested configurations. The best model, using a linear kernel and the first three principal components, reached 100% accuracy, sensitivity, and specificity in external validation. Our findings demonstrate that FTIR spectroscopy, in combination with SVM, offers a rapid, low-cost, and scalable strategy for the

screening and classification of *Leishmania* species, with promising applications for field and clinical diagnostics.

Keywords: Leishmaniasis, FTIR spectroscopy, Machine Learning, Parasite Identification.

Graphical Abstract



1. INTRODUCTION

Leishmaniasis encompasses a group of neglected tropical diseases (NTDs) that affect populations across all five continents. Each year, an estimated 1.3 million new cases and 20,000 to 30,000 deaths occur globally. In the Americas, the visceral, cutaneous, and mucocutaneous clinical forms pose a significant public health threat.

In Brazil alone, health authorities reported 1,648 cases in 2022, including 105 in the Centro-Oeste region [1].

Treatment, management, and control of leishmaniasis are constant challenges, mainly due to insufficient investment in novel therapies, preventive strategies, diagnostic tools, and even accurate identification of species and strains. These limitations hinder disease control efforts and disproportionately affect vulnerable populations worldwide [2]. Precise identification of *Leishmania* species remains critical for optimizing clinical interventions; however, research in this area remains underdeveloped [3].

Accurate species and strain identification depends on molecular and biochemical tools such as high-resolution melting (HRM) real-time polymerase chain reaction (PCR), DNA sequencing, PCR-restriction fragment length polymorphism (PCR-RFLP), and isoenzyme analysis. PCR-HRM demonstrates 96.4% sensitivity in biopsy samples [4] and offers rapid, adaptable analysis. DNA sequencing provides high-resolution identification ($\geq 99\%$ accuracy) [5] but requires specialized infrastructure [6]. PCR-RFLP and isoenzyme analysis remain valuable for species differentiation [6,7], although isoenzyme-based techniques are labor-intensive and vulnerable to contamination due to parasite culturing [7]. These methods enable the differentiation of key pathogenic species, such as *L. braziliensis* (mucocutaneous) and *L. donovani* (visceral), yet vary widely in cost, accuracy, and operational feasibility.

Despite their efficacy, the global application of these techniques encounters several barriers. These include the need for advanced infrastructure—limiting their use in resource-limited endemic regions, genetic variability among *Leishmania* species (e.g., kDNA variants), and sample degradation. Furthermore, the absence of standardized protocols for molecular markers (such as hsp70) and the presence of co-infections (e.g., *L. amazonensis* with *L. infantum*) complicate diagnosis. Traditional methods often rely on parasite culturing, which delays treatment and introduces additional logistical challenges [4–7].

Accurate species identification plays a critical role in guiding targeted therapies, as each clinical form requires distinct treatment approaches (e.g., antimonials for visceral leishmaniasis vs. amphotericin B for mucocutaneous leishmaniasis). Effective species-level identification also enhances epidemiological

surveillance by enabling precise mapping of reservoirs (e.g., dogs) and vectors (e.g., sand flies). Misdiagnosis not only delays treatment but also increases the risk of drug resistance. Integrating diagnostic data from human, animal, and environmental sources through a *One Health* framework is essential for breaking zoonotic transmission cycles and mitigating the global burden of leishmaniasis [8].

Challenges in *Leishmania* species identification reflect broader limitations in the diagnosis of other infectious diseases. In response, researchers have explored optical spectroscopy techniques as promising, low-cost alternatives. These techniques analyze spectral signatures from biological samples with minimal processing, offering a practical solution for low-resource settings.

Among these alternatives, Fourier-transform infrared (FTIR) spectroscopy has shown strong potential for microbial identification, combining low cost with easy implementation—two critical factors for public health and neglected disease control [9–11]. FTIR probes the vibrational modes of molecular bonds in the sample matrix. Subtle changes in concentration, molecular composition, conformation, and environmental conditions (e.g., pH) influence spectral signatures and allow group classification through proper data analysis. Integrating FTIR with machine learning enables enhanced classification, model validation, and interpretability of spectral data.

However, FTIR-based classification also faces limitations. Spectral features often reflect dominant molecular concentrations, which may obscure signals from diagnostically relevant but low-abundance molecules. In cultured samples, high promastigote concentrations do not guarantee informative spectra, as the surrounding matrix includes numerous non-specific biomolecules. Differences between species primarily arise at the molecular level within the promastigotes themselves. Standard FTIR spectrometers often lack the sensitivity to resolve these subtle differences.

To mitigate these issues, many studies analyze dried samples [9,10,12,13], which offer improved spectral resolution but introduce practical challenges, including longer preparation times, strict control of temperature and humidity, sample heterogeneity, and the need for spatial measurements across multiple regions.

This study aims to discriminate *Leishmania* species obtained from axenic cultivation of reference strains using FTIR spectroscopy in liquid sample form. We seek to identify key spectral features contributing to group classification and to develop a robust, reproducible protocol for simple and scalable sample analysis.

2. MATERIALS AND METHODS

2.1 Axenic cultivation of *Leishmania* sp. promastigotes

The samples that were subjected to FTIR analysis were reference strains that had already been properly identified and characterized using conventional methods: *Leishmania* (*Leishmania*) *amazonensis* (IFLA/BR/1967/PH8 and MHOM/BR/2022/LT 013_22 strains) and *Leishmania* (*Viannia*) *braziliensis* (MHOM/BR/75/M2904 strain). Parasites were maintained as promastigote forms at the Human Parasitology Laboratory/INBIO/UFMS in Schneider's Insect Medium (Sigma-Aldrich®, SP/Brazil), supplemented with 20% fetal bovine serum (Sigma-Aldrich®, SP/Brazil), 10,000 U/mL of penicillin, and 10 mg/mL of streptomycin (Sigma-Aldrich®, SP/Brazil) [14].

Cultures were expanded in the stationary growth phase and analyzed on the fifth day of cultivation. To prepare the samples, both cultures were centrifuged at 3,500 rpm for 10 minutes, after which the supernatant (culture medium) was discarded. Subsequently, the parasite concentrate was diluted in 500 µL of medium and characterized by infrared spectroscopy [14].

We divided the samples into two groups—*L. amazonensis* and *L. braziliensis* MHOM/BR/75/M2904)—with 40 samples in each group. The *L. amazonensis* group included two distinct strains: 20 samples from strain MHOM/BR/2022/LT 013_22 and 20 from strain IFLA/BR/1967/PH8. To increase intra-group variability and avoid bias associated with batch-specific effects, we performed independent culture replications across different weeks. Consequently, the samples were not derived from a single culture or measured under identical conditions on the same day. This approach aimed to capture natural variability within the group and improve the robustness and generalizability of the spectral classification.

We acquired the infrared spectra of each sample using an Agilent Cary 630 Fourier Transform Infrared (FTIR) spectrometer. For each measurement, we deposited 20 μL of the culture sample (liquid state) onto a solid attenuated total reflectance (ATR) crystal accessory at room temperature. Spectral acquisition focused on the mid-infrared (MIR) region, ranging from 1800 to 900 cm^{-1} , with a resolution of 8 cm^{-1} and 16 scans per sample.

2.3 Data analysis and prediction model validation

2.3.1 – data pre-processing

We performed all data analysis, predictive modeling, and figure generation using the Scikit-Learn library (version 1.3.0) in the Python programming language (version 3.11.5) [15]. Initially, the raw FTIR spectra underwent a pre-processing step. We applied a modified standard normal variate (SNV) normalization, using the absorbance at 1800 cm^{-1} as the reference baseline. This normalization technique minimizes baseline variability and compensates for multiplicative scatter effects across samples, improving spectral comparability [16].

However, we observed that the raw spectra exhibited low signal intensity and the presence of interference fringes, high-frequency oscillations resulting from optical path differences in the ATR setup. These fringes introduce frequency-dependent artifacts into the spectra, potentially biasing group classification. To address this, we applied a Fast Fourier Transform (FFT)-based filter to smooth the spectra and suppress high-frequency noise. The FFT filter effectively removes periodic noise components while preserving the relevant spectral features, thereby enhancing signal quality and the accuracy of subsequent multivariate analyses [17,18].

2.3.2 – Principal Component Analysis – Group Clustering.

We subjected the pre-processed FTIR-SNV-FFT data to multivariate statistical analysis using Principal Component Analysis (PCA) to enhance and highlight sample similarities and differences based on the main spectral features associated with molecular vibrational modes [19]. PCA reduces the dimensionality of the dataset while preserving the variance, enabling visualization of intrinsic

1
2
3 clustering patterns and discrimination among sample groups based on their
4 chemical composition.
5
6
7

8
9 Prior to interpretation, a one-time Hotelling's T^2 test was applied to the PCA
10 score matrix to identify and remove multivariate outliers [20]. This test quantifies
11 the distance of each observation from the multivariate center within the PCA-
12 reduced space, taking into account the covariance structure of the data.
13 Observations with T^2 values exceeding the confidence threshold were flagged as
14 outliers and excluded to ensure robust pattern recognition and prevent undue
15 influence on group separation.
16
17
18
19
20
21
22

23 In this dataset, PCA effectively captured the dominant sources of spectral
24 variation and separated samples according to strain- or species-specific biochemical
25 signatures. The PCA score plots revealed group clustering trends, while the
26 corresponding loading plots identified the wavenumbers (i.e., vibrational bands)
27 that most contributed to intergroup separation. These loadings provided insight into
28 the molecular components underlying the classification, supporting the
29 interpretation of biochemical differences between *Leishmania* species or strains.
30
31
32
33
34
35
36

37 2.3.3 – Prediction model – Support Vector Machine

38 Finally, we employed a Support Vector Machine (SVM) algorithm to construct
39 a predictive classification model using the PCA scores as input features [21]. SVM is
40 particularly well-suited for this type of analysis, as it constructs an optimal
41 hyperplane that maximizes the margin between different sample groups in the
42 multivariate space defined by the principal components. By utilizing a defined
43 number of PCs, SVM can capture complex nonlinear relationships when appropriate
44 kernel functions (e.g., linear, polynomial, radial basis function) are applied.
45
46
47
48
49
50

51 We optimized the model by systematically adjusting key hyperparameters—
52 such as the regularization parameter (C), kernel type, and kernel-specific
53 parameters—as well as by varying the number of principal components included in
54 the model. To identify the optimal number of PCs, we evaluated classification
55 accuracy in both training and validation datasets. We selected the configuration that
56 achieved the highest overall accuracy while minimizing the discrepancy between
57
58
59
60

training and validation performance, thereby reducing the risk of overfitting and improving model generalizability.

Once trained and validated, the final SVM model can be exported and applied to external datasets for automated classification, eliminating the need for re-accessing the original training data and enabling practical deployment in diagnostic workflows.

2.3.4 – Validation tests

The predictive accuracy of the model was estimated using the Leave-One-Out Cross-Validation (LOOCV) approach during the training phase [22]. Model training was performed using 70% of the total sample set. In LOOCV, each sample is sequentially excluded from the training set and used as an independent test case, while the remaining samples are used to train the model. This process is repeated iteratively until every sample has been used once as the validation set. The overall accuracy is calculated as the average classification performance across all iterations.

After identifying the optimal model conditions, the remaining 30% of the dataset—comprising samples not previously used during training—was employed as an independent test set. A confusion matrix was then constructed to assess the model's robustness by quantifying correct and incorrect classifications, as well as deriving key performance metrics such as sensitivity and specificity.

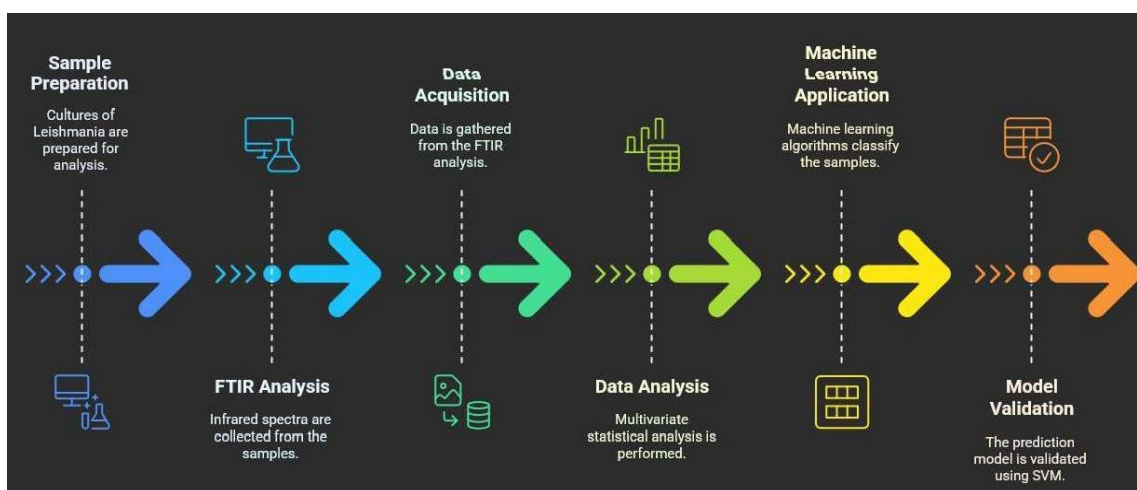


Figure 1: Flow chart of the experimental procedure summarizing the steps and activities until the achievement of a model for *Leishmania* species classification.

1
2
3
4
5 3. RESULTS and DISCUSSION
6

7 Figure 2(a,b) displays the average FTIR-SNV spectra of *Leishmania*
8 *amazonensis* and *Leishmania braziliensis*, both before and after application of a Fast
9 Fourier Transform (FFT) filter. The raw spectra exhibit characteristic interference
10 patterns—periodic oscillations superimposed along the spectral range—which are
11 likely associated with the nature of the liquid-drop samples and the single-reflection
12 ATR (Attenuated Total Reflectance) configuration used during data acquisition. An
13 overall reduction in signal intensity – which highlights the interference pattern in
14 the result – accompanies these spectral artifacts and can be attributed to a
15 combination of experimental and optical factors.
16
17

18 One contributing factor to the low signal intensity (below 10%) is the limited
19 concentration of biological material in the sample. When the concentration of cells
20 is insufficient, the interaction between infrared radiation and biomolecular
21 constituents is weakened, leading to faint or indistinct absorption bands. To address
22 this, increasing the cell density and minimizing the presence of residual culture
23 medium, such as BHI broth, are important steps to enhance the quality of the spectra,
24 such as improving the number of scans. However, additional procedures can
25 complicate the methodology and become time-consuming. To maintain the goal of a
26 fast and simple approach, we chose to use the samples as obtained to evaluate the
27 robustness of the data treatment.
28
29

30 The interference fringes observed in the spectra are indicative of thin-film
31 interference effects, which are common in ATR measurements involving non-
32 uniform or thin liquid films. These fringes originate from multiple internal
33 reflections occurring at the interfaces between the ATR crystal, the sample, and the
34 surrounding air. When the sample behaves like a semi-transparent thin film,
35 constructive and destructive interference between the reflected IR beams generates
36 periodic modulations in the spectrum. Furthermore, irregular sample thickness or
37 inhomogeneous distribution, such as that caused by the droplet curvature or
38 biofilm-like structures, leads to variations in optical path length, reinforcing the
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

formation of such interference patterns. Figure 2(c) illustrates a didactic scheme describing such phenomena.

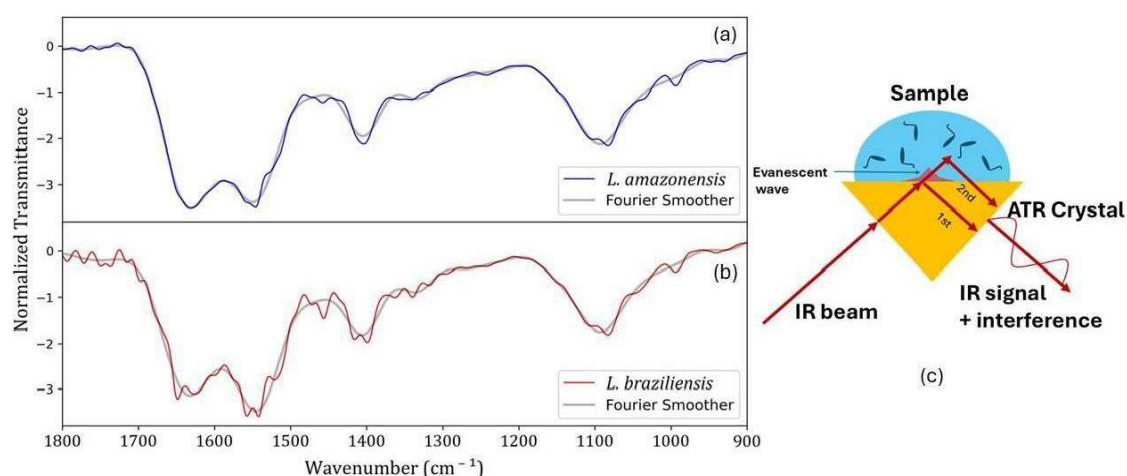


Figure 2: FTIR-SNV average spectra for: (a) *L. amazonensis* (blue trace), and (b) *L. braziliensis* (red trace) culture. (c) A diagram scheme to explain the origin of the interference pattern in the experiment.

To minimize the impact of interference fringes and improve spectral quality, post-acquisition data processing, such as baseline correction or the application of FFT filters, can be employed to suppress residual fringe artifacts and enhance the visibility of relevant spectral features. In this study, we applied an FFT filter to remove the periodic oscillations from the dataset. Figure 2(b), in particular, exhibits a more prominent effect of interference fringes, which could introduce bias in group classification during data analysis. The result after FFT filtering can also be identified in the figure (gray trace). This highlights the importance of a careful inspection of raw data before conducting multivariate analysis.

Figure 3 displays the FTIR-SNV-FFT average spectra for the *L. braziliensis* and *L. amazonensis* groups. In Figure 3(a), the shaded regions represent the standard deviation within each group. The $1440\text{--}900\text{ cm}^{-1}$ range shows more spectral variability for both groups, indicating greater signal dispersion. In particular, the *L. braziliensis* group shows higher deviation in the entire spectrum when compared to the *L. amazonensis* group.

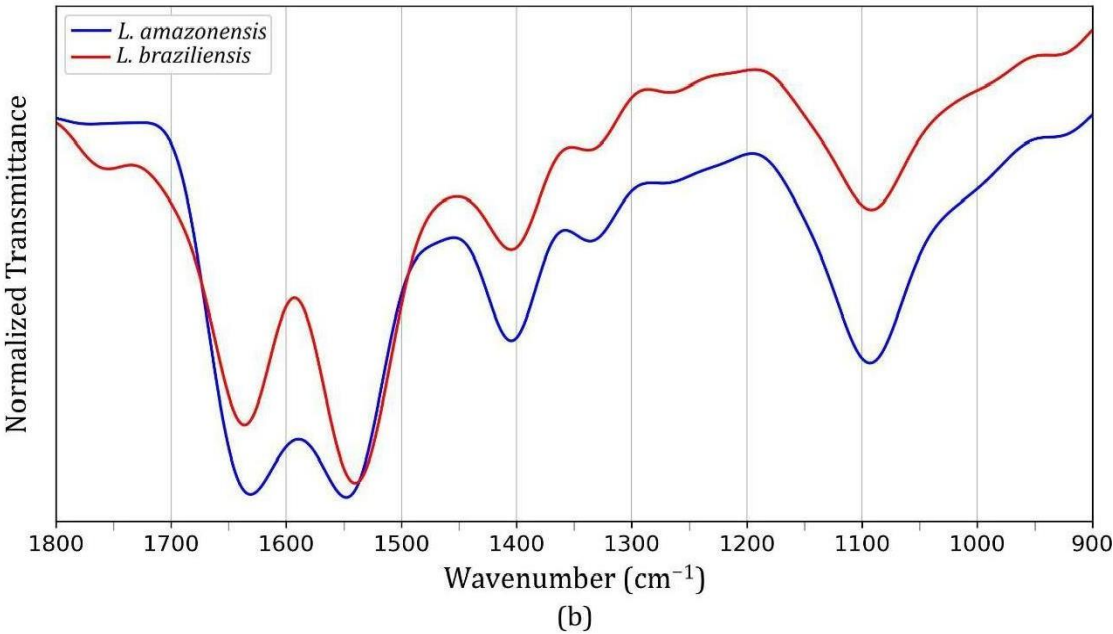
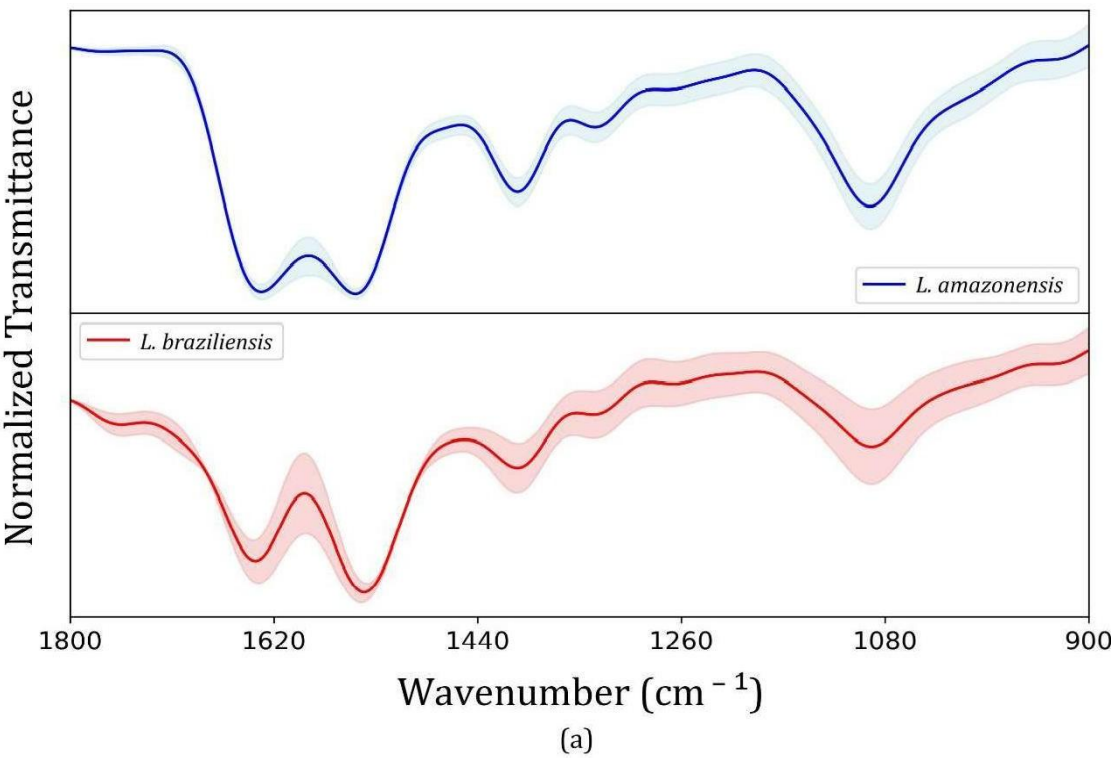


Figure 3: FTIR-SNV-FFT average spectra for *L. amazonensis* (blue trace) and *L. braziliensis* (red trace) cultures in Schneider's Insect medium. (a) shows the average spectra and the respective standard deviation for each group, (b) exhibits the superimposed average spectra for each group for comparison.

Figure 3(b) shows the superimposed FTIR-SNV-FFT average spectra for the *L. braziliensis* (red trace) and *L. amazonensis* (blue trace) groups. From 1500 to 900 cm^{-1} , both groups exhibit a similar number of vibrational modes, with comparable spectral patterns and relative band intensities. However, notable differences emerge in the 1800 to 1500 cm^{-1} region. A minor band around 1750 cm^{-1} is observed exclusively in the *L. braziliensis* group. Additionally, the major bands centered at approximately 1650 and 1550 cm^{-1} display markedly different relative intensities between the two groups, suggesting distinct contributions for the vibrational modes emerging from the biomolecular matrix, which can be related to compositional and/or conformation changes.

The FTIR spectral analysis of *Leishmania* promastigotes cultured in BHI medium revealed distinct vibrational bands corresponding to molecular components derived from both the parasites and the culture medium. The prominent bands at $\sim 1750 \text{ cm}^{-1}$ and $\sim 1650 \text{ cm}^{-1}$ are attributed to the C=O stretching vibrations of amide I, primarily originating from protein constituents present in the BHI medium. The band at $\sim 1550 \text{ cm}^{-1}$ corresponds to the N-H bending vibration of amide II, characteristic of the protein content of *Leishmania* cells. The absorption at $\sim 1400 \text{ cm}^{-1}$ is consistent with symmetric stretching of CH_3 groups from lipids, as well as contributions from the C=O stretching of carboxylate groups (COO^-) and CH_2 bending of lipid chains, reflecting the mixed biochemical origin from both the parasites and residual medium components. The band observed at $\sim 1350 \text{ cm}^{-1}$ is indicative of in-plane bending vibrations of CH_3 groups, likely arising from methyl-bearing biomolecules such as phospholipids or branched-chain amino acids. A band at $\sim 1280 \text{ cm}^{-1}$ is assigned to overlapping amide III vibrations from protein backbones and P=O stretching modes from nucleic acids (DNA and RNA) of the parasite cells. Additionally, the symmetric stretching of PO_2^- groups, also observed in this region, can be attributed to the phosphodiester backbone of nucleic acids from *Leishmania*, although potential contributions from phosphate-containing polysaccharides in Schneider's Insect medium cannot be excluded [23,24].

Figure 4 presents principal component analysis (PCA) results derived from FTIR-SNV-FFT spectra of *L. amazonensis* (blue circle symbols) and *L. braziliensis*

(red diamond symbols) groups. On the left, the scores plot (PC1 vs. PC2) exhibits a clear separation between the two species, with PC1 accounting for 98.1% of the total variance and PC2 contributing 1.1%. This separation suggests strong biochemical differences in the spectral profiles of the two *Leishmania* species, most of which are captured by PC1, as expected, once we observed in the FTIR average spectra (Figure 3(b)) visible differences in the 1700 – 1500 cm^{-1} range.

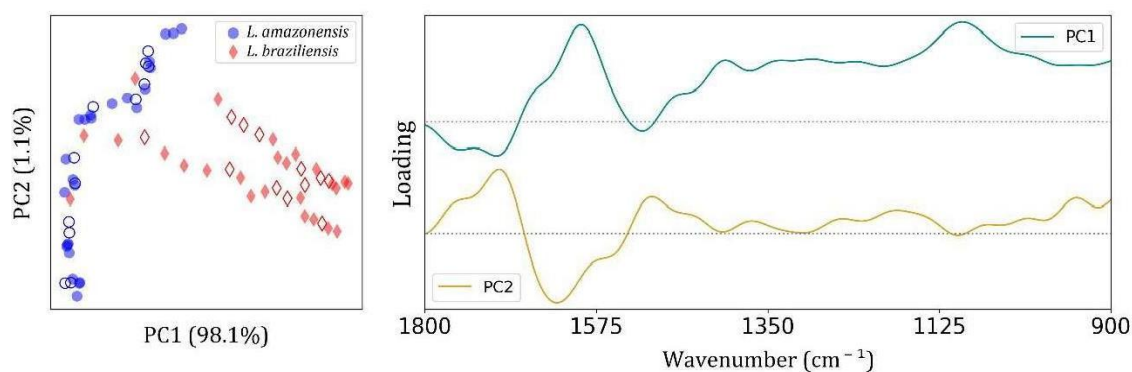


Figure 4: Principal Component Analysis (PCA) of FTIR spectra from *Leishmania amazonensis* and *Leishmania braziliensis* promastigotes culture in Schneider's Insect medium. (Left) PCA score plot showing clear separation between *L. amazonensis* (blue circles) and *L. braziliensis* (red diamonds). Filled symbols represent samples used for model training, while hollow symbols correspond to test samples reserved for model validation. (Right) PCA loading plots for PC1 and PC2, indicating the spectral regions contributing most to the variance. The dashed horizontal line denotes the zero-loading baseline for reference.

The corresponding loading plots (Figure 4 on the right) provide insight into the spectral regions responsible for the observed variance. For PC1, major contributions are observed around 1650 cm^{-1} , corresponding to C=O stretching vibrations of amide I from proteins, and at $\sim 1550 \text{ cm}^{-1}$, related to N-H bending vibrations of amide II, likely reflecting species-specific protein composition. Additional contributions appear near 1400 cm^{-1} (symmetric CH_3 stretching and COO^- stretching), and around 1240–1280 cm^{-1} , where overlapping amide III vibrations and P=O stretching from nucleic acids are found. The region between 1100 and 1000 cm^{-1} , where symmetric stretching of PO_2^- groups from nucleic acids and BHI polysaccharides occurs, also contributes significantly to PC1.

1
2
3
4
5 PC2 captures smaller variance, with notable features at ~ 1550 and ~ 1400
6 cm^{-1} , suggesting it may reflect subtler variations in protein or lipid composition
7 rather than species-level differentiation. Overall, the PCA analysis indicates that
8 spectral differences primarily related to proteins (amide I and II), lipids (CH_2 , CH_3),
9 and nucleic acid components ($\text{P}=\text{O}$ and PO_2^-) drive the discrimination between *L.*
10 *amazonensis* and *L. braziliensis* in the FTIR dataset.
11
12
13
14
15
16

17 Figure 5 presents the overall accuracy of SVM-based classification models
18 (left column) using different kernel functions (Linear, Quadratic, and Cubic) and
19 varying the regularization hyperparameter C (1, 10, and 100). For each
20 configuration, the number of principal components (PCs) was incrementally
21 increased, and overall accuracy was monitored. To mitigate the risk of overfitting,
22 model evaluation was halted upon reaching the first local maximum in accuracy. All
23 tested configurations yielded accuracies above 90%. Given that PC1 alone accounts
24 for 98% of the explained variance, the analysis was limited to a maximum of three
25 PCs to optimize performance while maintaining model generalizability.
26
27
28
29
30
31
32
33

34 The right column of Figure 5 displays the corresponding loading plots and
35 clustering of the data in the PCA space (swarm plot). The loadings indicate that the
36 first three principal components capture vibrational features relevant to class
37 discrimination. While PC1 contributes most significantly to class separation,
38 projections onto PC2 and PC3 also provide complementary information that
39 enhances the robustness of SVM classification.
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60

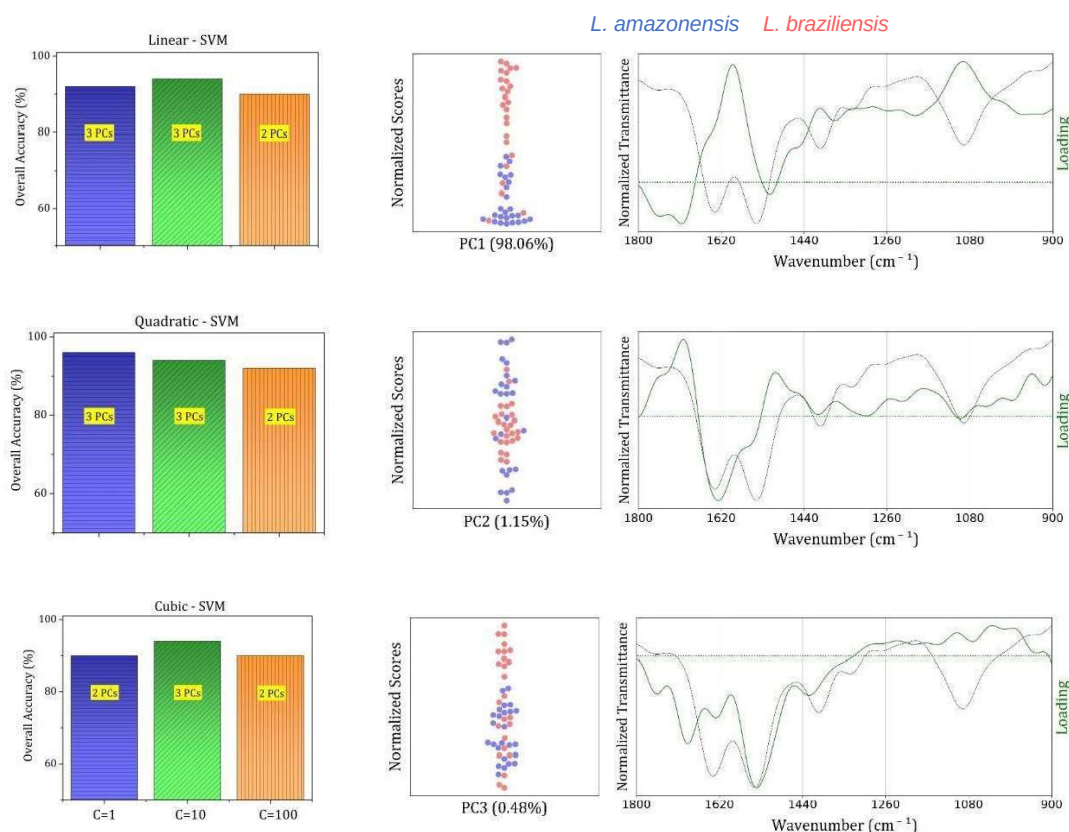


Figure 5: Performance of SVM classification models using line, quadratic, and cubic (top to bottom, left column) kernels for the discrimination of *Leishmania* species based on PCA data from FTIR-SNV-FFT spectral data. Left column: Overall classification accuracy for different kernel types and numbers of principal components (PCs), with regularization parameter $C=1, 10, 100$. For each kernel, the best performance was achieved using 2 or 3 PCs, as indicated on each bar. Middle column: PCA score plots showing separation along PC1 (98.06% of explained variance, top), PC2 (1.15%, middle), and PC3 (0.48%, bottom) corresponding to the components with the highest contribution to class separation. The species are represented by *L. amazonensis* (blue points) and *L. braziliensis* (red points). Right column: PCA loading plots showing the spectral regions (in cm^{-1}) that most contribute to PC1, PC2, and PC3. Green curves represent the loading values, the green dashed horizontal line indicates the zero-reference baseline, and the gray dashed line represents the FTIR-SNV-FFT average spectra.

Finally, the predictive models were validated through a blind test using 30% of the samples that were not included in the training phase. Figure 6 summarizes the best-performing model, which corresponds to the Linear SVM using the first

three principal components (PCs). This model demonstrated superior generalization capacity, as evidenced by its performance on the blind test. The small deviation observed between the leave-one-out cross-validation (LOOCV) and external validation results is not considered significant, likely due to sample size imbalance. In the blind test, the model achieved 100% accuracy, with both sensitivity and specificity also reaching 100%, indicating its robustness.

Notably, all tested kernel functions yielded models with overall accuracy above 90%, underscoring the potential of SVM-based approaches as reliable screening tools in microbiological laboratories. These models can serve as rapid and cost-effective alternatives to conventional gold standard methods. The primary advantages of the proposed approach include ease of implementation, minimal sample preparation requirements, and low operational cost, making it a promising strategy to accelerate diagnostic workflows across a variety of applications.

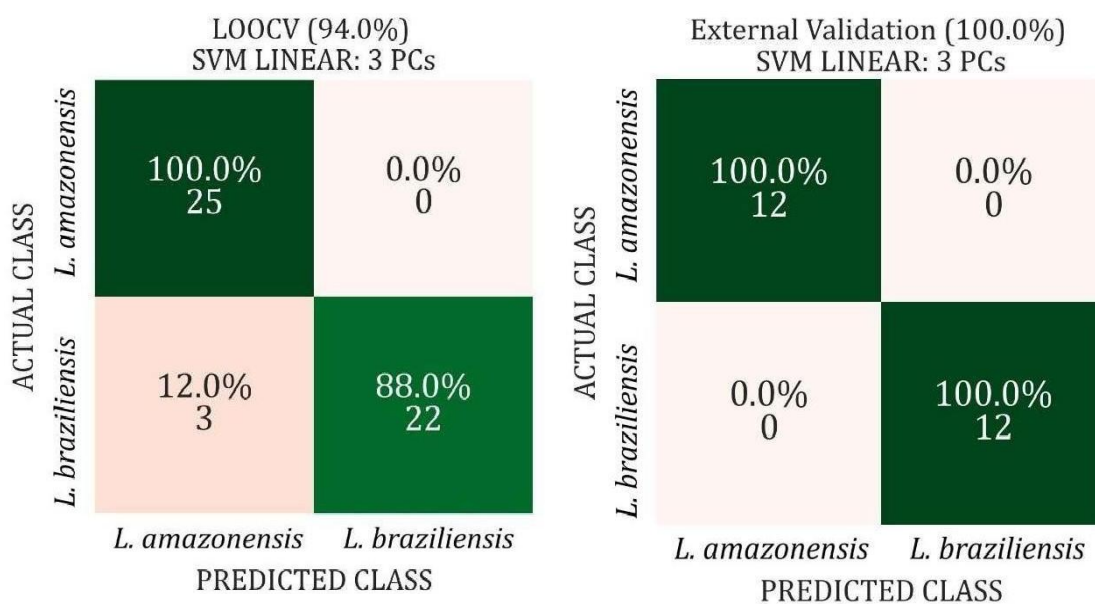


Figure 6: Confusion matrices illustrating the performance of the SVM classification model using three principal components (PCs) for cultured *Leishmania* promastigote samples. The left matrix shows the results of external validation using 30% of the dataset (blind test), while the right matrix presents the outcomes of leave-one-out cross-validation (LOOCV) on the remaining 70% of the data. The input features were derived from PCA scores obtained from FTIR spectra preprocessed

with Standard Normal Variate (SNV) and Fast Fourier Transform (FFT), within the spectral range of 1800–900 cm⁻¹.

Despite the promising results, this study should be interpreted as both robust and exploratory. While the spectral and computational approaches demonstrated high accuracy and generalization capacity, the dataset was limited to two *Leishmania* species and a restricted number of cultured promastigote samples. To further validate and expand the applicability of this methodology, future studies should incorporate a broader diversity of *Leishmania* strains and clinical isolates from different geographic and epidemiological backgrounds. The approach can also be extended to include additional *Leishmania* species, potentially enabling the development of a comprehensive, FTIR-based classification system for the genus.

4. CONCLUSION

This study demonstrates that FTIR spectroscopy, coupled with supervised machine learning algorithms, enables accurate discrimination between *Leishmania amazonensis* and *Leishmania braziliensis* in axenic culture. The methodology proved to be fast, low-cost, and compatible with standard laboratory infrastructure, requiring minimal sample preparation—making it especially advantageous for resource-limited settings.

Beyond its immediate diagnostic potential, the approach represents a breakthrough in facilitating the work of clinical and research laboratories by offering a practical and scalable alternative to conventional molecular methods. The integration of spectral data analysis with AI provides a promising foundation for high-throughput parasite screening.

Despite the promising performance, the study should be viewed as a proof of concept. The current validation is limited to two species under controlled conditions. Therefore, future studies are essential to expand the dataset, including a broader diversity of *Leishmania* species, strains, and clinical isolates from different geographical regions and host origins (human and animal). This will be critical to validate the method’s robustness and support its translation into field diagnostics.

Moreover, the proposed strategy aligns with One Health principles by providing a potential tool for integrated surveillance across human, animal, and environmental reservoirs. Its application could significantly enhance the monitoring and control of zoonotic transmission routes, contributing to more effective disease surveillance and management frameworks in endemic regions.

Consent to Publish

All participants were informed and consent for publication was obtained.

Author Contribution Statement

All authors have equal contributions for the study

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; 444388/2023-0. Fundação de Apoio ao Desenvolvimento do Ensino, Ciência e Tecnologia do Estado de Mato Grosso do Sul (FUNDECT), code 91/2023; 360/2022.

Competing interest statement

The authors declare that there is no conflict of interest.

Data availability statement

Data will be shared upon request.

REFERENCES

- [1] BRASIL, Epidemiological Bulletin Secretariat of Health and Environment Surveillance Ministry of Health. Neglected Tropical Diseases in Brazil Morbidity, Mortality and National Response in the Context of the Sustainable Development Goals 2016-2020. ISSN 9352-7864, 2024.
- [2] Wamai RG, Kahn J, McGloin J, Ziaggi G. Visceral leishmaniasis: a global overview. J Glob Health Sci. 2020 Apr;2(1):e3. DOI: 10.35500/jghs.2020.2.e3

- Foulet F, Botterel F, Buffet P, Morizot G, Rivollet D, Deniau M, Pratlong F, Costa JM, Bretagne S. Detection and identification of *Leishmania* species from clinical specimens by using a real-time PCR assay and sequencing of the cytochrome B gene. *J Clin Microbiol*. 2007 Jul;45(7):2110-5. DOI: 10.1128/JCM.02555-06.
- [3] Veasey JV, Zampieri RA, Lellis RF, Freitas THP, Winter LMF. Identification of *Leishmania* species by high-resolution DNA dissociation in cases of American cutaneous leishmaniasis. *An Bras Dermatol*. 2020;95:459–68. DOI: 10.1016/j.abd.2020.02.003
- [4] Foulet F, Botterel F, Buffet P, Morizot G, Rivollet D, Deniau M, Pratlong F, Costa J, Bretagne S. 2007. Detection and Identification of *Leishmania* Species from Clinical Specimens by Using a Real-Time PCR Assay and Sequencing of the Cytochrome b Gene. *J Clin Microbiol* 45. DOI:10.1128/jcm.02555-06.
- [5] Reithinger R, Dujardin JC. Molecular diagnosis of leishmaniasis: current status and future applications. *J Clin Microbiol*. 2007 Jan;45(1):21-5. DOI: 10.1128/JCM.02029-06.
- [6] Marfurt J, Nasereddin A, Niederwieser I, Jaffe CL, Beck H, Felger I. 2003. Identification and Differentiation of *Leishmania* Species in Clinical Samples by PCR Amplification of the Miniexon Sequence and Subsequent Restriction Fragment Length Polymorphism Analysis. *J Clin Microbiol* 41. DOI:10.1128/jcm.41.7.3147-3153.2003
- [7] Hong, A.; Zampieri, R.A.; Shaw, J.J.; Floeter-Winter, L.M.; Laranjeira-Silva, M.F. One Health Approach to Leishmaniasis: Understanding the Disease Dynamics through Diagnostic Tools. *Pathogens* 2020, 9, 809. DOI: 10.3390/pathogens9100809.
- [8] Y.G. Marangoni-Ghoreyshi, T. Franca, J. Esteves, A. Maranni, K.D.P. Portes, C. Cena, C.R.B. Leal. Multi-resistant diarrheagenic *Escherichia coli* identified by FTIR and machine learning: a feasible strategy to improve the group classification. *RSC Advances*, v. 13, n. 36, p. 24909–24917, 2023.
- [9] Gabriela Pacher, Thiago Franca, Miller Lacerda, Natália O. Alves, Eliane M. Piranda, Carla Arruda, and Cícero Cena
ACS Infectious Diseases 2024 10 (2), 467-474. DOI: 10.1021/acsinfecdis.3c00430
- [10] Casaril, A.E.; Santos, C.G.; Marangoni, B.S.; Lima, S.M.; Andrade, L.H.C.; Fernandes, W.S.; Infran, J.O.M.; Alves, N.O.; Borges, M.D.G.L.; Cena, C.; Oliveira, A.G.

Intraspecific differentiation of sandflies specimens by optical spectroscopy and multivariate analysis. J. of biophotonics 2021, DOI: 10.1002/jbio.202000412

[11] Brito, E.C.A.; Franca, T.; Canassa, T.; Weber, S.S.; Paniago, A.M.M.; Cena, C. Paracoccidioidomycosis screening diagnosis by FTIR spectroscopy and multivariate analysis. Photodiagnosis Photodyn. Ther 2022, DOI: 10.1016/j.pdpdt.2022.102921

[12] Cameron, J.M; Butler, H.J.; Palmer, D.S.; Baker, M.J. Biofluid spectroscopic disease diagnostics: A review on the processes and spectral impact of drying. J. of Biophotonics, 2018, DOI: 10.1002/jbio.201700299

[13] ALVES, NATÁLIA O. ; OSHIRO, JÉSSICA A. ; SILVA, YUNNA C. ; PACHER, GABRIELA C. ; CASARIL, ALINE E. ; RIZK, YASMIN S. ; UEHARA, SILVIA N. O. ; PANIAGO, ANAMARIA M. M. ; ANDRADE, ISADORA L. X. ; ARRUDA, CARLA C. P. ;

OLIVEIRA, ALESSANDRA G. . VL-HIV co-infection with Leishmania containing skin lesions resembling para-kala-azar dermal leishmaniasis. PLoS Neglected Tropical Diseases, v. 18, p. e0012438, 2024.

[14] Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D.; Brucher, M.; Perrot, M.; Duchesnay, E. Scikit-learn: Machine learning in Python, J. Mach. Learn. Res. 2011, 12(85), 2825-2830.

[15] Engel, J.; Gerretzen, J.; Szymanska, E.; Jansen, J.J.; Downey, G.; Blanchet, L.; Buydens, L.M.C. Breaking with trends in pre-processing, TrAC Trends Anal. Chem. 2013, DOI: 10.1016/j.trac.2013.04.015

[16] Franca, Thiago; Lacerda, Miller; Calvani, Camila; Arruda, Kelvyn; Maranni, Ana; Nicolodelli, Gustavo; KARTHIKEYAN, SIVAKUMARAN; Marangoni, Bruno; Nascimento, Carlos; Cena, Cicero. Improving Bovine Brucellosis Diagnostics: Rapid, Accurate Detection via blood serum infrared Spectroscopy and Machine Learning. ACS OMEGA, 2025.

[17] Kosarev, E.L.; Pantos, E. Optimal smoothing of 'noisy' data by fast Fourier transform. J. of Physics E: Sci. Inst. 1983, DOI: 10.1088/0022-3735/16/6/020

[18] Jolliffe, I.T.; Cadima, J. Principal Component Analysis: a review and recent developments, Phil. Trans. R. Soc. A 2016, DOI: 10.1098/rsta.2015.0202

[19] Jensen, D.R.; Ramirez, D.E. Use of Hotelling's T^2 : Outlier Diagnostics in Mixtures. Int. J. of Statistics and Prob. 2017, DOI: 10.5539/ijsp.v6n6p24

[20] Chauhan, V.K.; Dahiya, K.; Sharma, A. Problem formulations and solvers in linear



SVM: a review, *Artif. Intell. Rev.* 2018, DOI: 10.1007/s10462-018-9614-6

[21] Yates, L.A.; Aandahl, Z.; Richards, S.A.; Brook, B.W. Cross-validation for model selection: A review with examples from ecology. *Ecological Monographs*, 2023, DOI: 10.1002/ecm.1557

[22] Aguiar, J.C.; Mittmann, J.; Ferreira, I.; Ferreira-Strixino, J.; Raniero, L. Differentiation of *Leishmania* species by FT-IR spectroscopy. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 2015. Available in: <http://dx.doi.org/10.1016/j.saa.2015.01.008>.

[23] Thulya Chakkumpulakkal Puthan Veettil, Rebekah N. Duffin, Supti Roy, Philip C. Andrews, Bayden R. Wood, Biochemical characterization and discrimination of *Leishmania* major parasites and infected macrophages with Raman spectroscopy and chemometrics, *Clinical Spectroscopy*, V 5, 2023. DOI: 10.1016/j.clispe.2023.100024.