

Determinação rápida e sustentável de doença de Chagas via *Machine Learning*

Gabriely S. Folli^{1*} (PG), Mathews S. Martins¹ (PG), Ângelo A. O. Silva² (PQ), Felipe S. S. de Jesus² (PQ), Emily F. Santos² (PQ), Daniel D. Sampaio² (PQ), Fred L. N. Santos (PQ)², Wanderson Romão³ (PQ), Valerio G. Barauna¹ (PQ), Paulo R. Filgueiras¹ (PQ).

**gabriely.folli@ufes.br*

¹ Universidade Federal do Espírito Santo, Vitória, ES, Brasil.

² Fundação Oswaldo Cruz, Salvador, BA, Brasil

³ Instituto Federal do Espírito Santo, Vila-Velha, ES, Brasil.

Palavras-Chave: Diagnóstico precoce, espectroscopia FT-IR, Reagent-free, suporte diagnóstico, triagem.

Introdução

A doença de Chagas, causada pelo protozoário *Trypanosoma cruzi*, apresenta duas fases clínicas (aguda – encontrado na corrente sanguínea – ou crônica – não encontrado na corrente sanguínea)¹. Com isso, não existe um exame único para identificação assertiva^{2,3}. No entanto, os exames utilizados carecem de reagentes específicos e tempo considerável de análise². Ademais, o diagnóstico é frequentemente dificultado pela natureza silenciosa da infecção (provavelmente pela baixa carga parasitária), que pode permanecer assintomática por anos e levar a complicações graves e irreversíveis com o diagnóstico tardio². Diante deste cenário, este trabalho visou desenvolver uma alternativa para determinação precoce e sustentável da doença de chagas com aplicação da espectroscopia de infravermelho médio com transformada de Fourier (FT-IR) associada à aprendizagem de máquina para auxílio da tomada de decisão médica.

Resultados e Discussão

Cem participantes (49 positivas e 51 negativas) foram obtidas e com diagnóstico confirmado via Laboratório Central de Saúde Pública (LACEN) e Fundação de Hematologia e Hemoterapia da Bahia (HEMOBA). Este estudo foi conduzido seguindo os princípios descritos na Declaração de Helsinque e recebeu aprovação do Comitê de Ética em Pesquisa (CEP) do Instituto Gonçalo Moniz, Fundação Oswaldo Cruz (IGM-FIOCRUZ) em Salvador, Bahia, Brasil (protocolo nº 67809417.0.0000.0040). Todas as amostras foram adquiridas em triplicata via ALPHA II (Bruker, Alemanha) com acessório de reflectância total atenuada (ATR-FTIR), no intervalo espectral de 4000 a 400 cm⁻¹, resolução de 4 cm⁻¹ e 32 varreduras. Os modelos de Análise por Componentes Principais (PCA) e Análise Discriminante por Mínimos Quadrados Parciais (PLS-DA) foram construídos no *software* Matlab R2023a. Inicialmente, houve a divisão do conjunto inicial em 70% para treinar o modelo e 30% para validação externa e diferentes pré-tratamentos foram testados. O melhor modelo foi avaliado a partir de diferentes parâmetros de desempenho para ambos os subgrupos amostrais. A capacidade preditiva do modelo desenvolvido foi averiguada a partir 200 ciclos de teste de permutação⁴. O melhor modelo de PLS-DA foi construído utilizando o filtro Savitzky-Golay. Os parâmetros de desempenho para as amostras de validação da classe negativa apresentaram valores de 0,89, 0,80, 0,69, 0,84 e 0,84 para sensibilidade, especificidade, coeficiente de correlação de Mathews, F1score e exatidão, respectivamente. Ademais, o teste de permutação evidenciou a não aleatoriedade e correspondência verdadeira entre as informações químicas obtidas no espectro de FT-IR. As variáveis 1500 – 1700 cm⁻¹ foram as mais importantes para classificação da doença de Chagas. Essa região está muito correlacionada a agrupamentos de amida II, amida I, nitrila, aminas primárias e carbonilas que podem estar associadas a proteínas, lipídios e ácidos graxos⁵.

Conclusões

Os resultados obtidos são promissores para metodologia de triagem precoce e assintomática da doença de chagas de forma rápida, sustentável e eficiente.

Agradecimentos

FAPES, CAPES, CNPq, UFES, LACEN, HEMOBA, IGM-FIOCRUZ.

¹ Lima, F.; Oliveira, P.; Mortara, R.; Silveira, J.; Bahia, D. *New Biotechnol.* 27, 837–843, 2010.

² Rassi Jr., A.; Rassi, A.; Rezende, J. *Infectious Disease Clinics of North America*, 26(2), 275-291, 2012.

³ Zhang, L.; Tarleton, R. L. *The Journal of Infectious Diseases*, 180, 480–486, 1999.

⁴ Folli, G.S.; Nascimento, M.H.C.; Lovatti, B.P.O.; Romão, W.; Filgueiras, P.R. *Chemom. Intell. Lab. Syst.* 251, 105154, 2024.

⁵ Najib, M.; El-Sayed; *et al. Science*, 309,409-415, 2005.