



UNIVERSIDADE FEDERAL DO MARANHÃO
CENTRO DE CIÊNCIAS BIOLÓGICAS E DA SAÚDE
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA SAÚDE –
PPGCS
MESTRADO



**APLICAÇÃO DE ESPECTROSCOPIA DE INFRAVERMELHO
COM TRANSFORMADA DE FOURIER NA BUSCA DE
MARCADORES DIAGNÓSTICOS EM PACIENTES COM
LEUCEMIA MIELOIDE AGUDA**

MARCOS GUIMARÃES FERREIRA

São Luís, MA

2021

MARCOS GUIMARÃES FERREIRA

**APLICAÇÃO DE ESPECTROSCOPIA DE INFRAVERMELHO
COM TRANSFORMADA DE FOURIER NA BUSCA DE
MARCADORES DIAGNÓSTICOS EM PACIENTES COM
LEUCEMIA MIELOIDE AGUDA**

Dissertação de Mestrado apresentado ao Programa de Pós-Graduação em Ciências da Saúde da Universidade Federal do Maranhão como requisito para obtenção do título de Mestre em Ciências da Saúde.

Orientadora: Profa. Dra. Silma Regina Ferreira Pereira

Co-orientador: Prof. Dr. Robinson Sabino da Silva

São Luís, MA

2021

MARCOS GUIMARÃES FERREIRA

**APLICAÇÃO DE ESPECTROSCOPIA DE INFRAVERMELHO
COM TRANSFORMADA DE FOURIER NA BUSCA DE
MARCADORES DIAGNÓSTICOS EM PACIENTES COM
LEUCEMIA MIELOIDE AGUDA**

Dissertação de Mestrado apresentado ao
Programa de Pós-Graduação em Ciências da
Saúde da Universidade Federal do Maranhão
como requisito para obtenção do título de
Mestre em Ciências da Saúde.

BANCA EXMINADORA

Profª. Dra. Silma Regina Ferreira Pereira
Universidade Federal do Maranhão

Prof. Dr. Lúcio Roberto Cançado Castelhana
Universidade Federal da Paraíba

Prof. Dr. Clenilton Costa dos Santos
Universidade Federal do Maranhão

Profª. Dra. Luciana Magalhães Rebelo Alencar
Universidade Federal do Maranhão

Ficha gerada por meio do SIGAA/Biblioteca com dados fornecidos pelo(a) autor(a).
Diretoria Integrada de Bibliotecas/UFMA

Guimarães Ferreira, Marcos.

Aplicação de espectroscopia de infravermelho por transformada de Fourier na busca de marcadores diagnósticos em pacientes com leucemia mieloide aguda / Marcos Guimarães Ferreira. - 2021.

69 p.

Coorientador(a): Robinson Sabino-Silva.

Orientador(a): Silma Regina Ferreira Pereira.

Dissertação (Mestrado) - Programa de Pós-graduação em Ciências da Saúde/ccbs, Universidade Federal do Maranhão, São Luís, 2021.

1. Algoritmos de aprendizado de máquinas. 2. Espectroscopia ATR-FTIR. 3. Leucemia. 4. Triagem. I. Regina Ferreira Pereira, Silma. II. Sabino-Silva, Robinson. III. Título.

“Seja quem você for, seja qual for a posição social que você tenha na vida, a mais alta ou a mais baixa, tenha sempre como meta muita força, muita determinação e sempre faça tudo com muito amor e com muita fé em Deus, que um dia você chega lá. De alguma maneira você chega lá.”

(Ayrton Senna)

Dedico essa dissertação a Deus, à minha família e a todos os que colaboraram comigo.

AGRADECIMENTOS

Ao grande arquiteto do universo por ter concedido saúde, força, ânimo e o encorajamento para a conclusão de mais uma etapa da minha vida.

À minha Orientadora, Profa. Dra. Silma Regina Ferreira Pereira pelos ensinamentos, pela dedicação, paciência e cuidado ao longo desses anos. Que fique registrado minha admiração pela sua pessoa.

Ao meu Co-orientador, Prof. Dr. Robinson Sabino-Silva pelo tempo, dedicação e cuidado. Seus ensinamentos foram, sem dúvidas, uma parte importante dessa etapa na minha jornada.

À minha família pelo apoio incondicional durante todos os momentos da minha vida, no qual sempre estiveram me incentivando a permanecer firme na caminhada. Jamais teria chegado tão longe sem esse apoio.

À minha namorada, Letícia Silva Teixeira, pelo incentivo, companheirismo e paciências ao longo desse tempo. Como diz Thales Grulha: “uma boa companhia encurta os cansativos percursos e prolonga a boa caminhada”.

Aos nossos colaboradores da Universidade Federal de Uberlândia, na pessoa do Prof. Dr. Luís Goulart, do Prof. Dr. Murillo Guimarães, da Doutoranda Léia Sousa-Cardoso e de todos os envolvidos.

Aos amigos do laboratório de Genética e biologia Molecular da Universidade Federal do Maranhão, especialmente ao Doutorando Rossy-Eric Pereira Soares.

À Doutoranda Perla Lopes de Freitas, pelo incentivo e contribuição durante o início da minha vida acadêmica. Que fique registrado que você foi o elo inicial entre mim e a pesquisa científica. Gratidão!

A Fundação de Amparo à Pesquisa e ao Desenvolvimento Científico e Tecnológico do Maranhão (*FAPEMA*) pelo apoio ao projeto no qual está vinculada essa dissertação e a Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), por ter concedido a bolsa de mestrado durante uma parte da execução deste trabalho.

Ao Programa de Pós-Graduação em Ciências da Saúde da Universidade Federal do Maranhão (PPGCS-UFMA) por ter concedido a mim esta sublime oportunidade de alcançar mais uma etapa da minha vida acadêmica.

A Universidade Federal do Maranhão pela infraestrutura disponibilizada ao PPGCS e ao Laboratório de Genética e biologia Molecular, a qual faço parte.

RESUMO

A leucemia mieloide aguda (LMA) faz parte de um grupo de neoplasias hematológicas que se caracterizam pelo bloqueio maturativo e proliferação clonal exacerbada de progenitores mieloides na medula óssea ou no sangue periférico. Atualmente, os principais métodos diagnósticos dessa neoplasia, dependem da presença de pelo menos 20% de células malignas na medula óssea ou no sangue periférico, o que geralmente coincide com a fase avançada da doença. Além disso, essas metodologias requerem equipamentos e insumos caros, bem como mão de obra especializada, o que torna o diagnóstico da LMA restrito apenas aos grandes centros de diagnóstico ou pesquisa. Portanto, neste trabalho, implementamos o uso de Espectroscopia de Infravermelho com Transformada de Fourier e reflexão total atenuada (ATR-FTIR) univariada e acoplada com algoritmos de inteligência artificial, como uma ferramenta rápida, barata e com alta sensibilidade e especificidade para discriminação de pacientes com diagnóstico clínico e laboratorial de leucemia mieloide aguda. Para isso, realizamos uma análise espectral quantitativa do plasma (1µl) de 35 pacientes e 35 indivíduos controles, com uma resolução de 4 cm⁻¹ e 32 varreduras espectrais em triplicata, utilizando o espectro do ar como *background* das análises. Como resultado, foram identificados dez modos vibracionais (3095 cm⁻¹, 1648 cm⁻¹, 1635 cm⁻¹, 1540 cm⁻¹, 1284 cm⁻¹, 1170 cm⁻¹, 1120 cm⁻¹, 1078 cm⁻¹, 1031 cm⁻¹ e 850 cm⁻¹) com sensibilidade e especificidade ≥80% (p<0,0001) em discriminar as amostras patológicas em relação ao grupo controle. Além disso, duas regiões do espectro médio (1333 cm⁻¹ -1272 cm⁻¹ e 1096 cm⁻¹ - 996 cm⁻¹) também apresentaram valores de sensibilidade e especificidade igual ou superior a 80% (p<0,0001). Com relação a classificação dos espectros por inteligência artificial, os algoritmos de regressão logística, análise discriminante linear e floresta aleatória, apresentaram acurácias igual ou superior a 90% em pelo menos um conjunto de pré-processamento selecionado, com destaque para o algoritmo máquina de vetores de suporte, que discriminou o controle e LMA com acurácia de 100%. Em conclusão, a espectroscopia ATR-FTIR univariada ou acoplada com algoritmos de inteligência artificial foi capaz de discriminar o plasma de pacientes com LMA e controles, sendo promissor seu uso como teste rápido, de baixo custo e com alta sensibilidade e especificidade para triagem da LMA.

Palavras chaves: Leucemia; triagem; modos vibracionais; algoritmos de aprendizado de máquinas; Espectroscopia ATR-FTIR.

ABSTRACT

Acute myeloid leukemia (AML) is part of a group of hematologic neoplasms that are characterized by maturation blockade and exacerbated clonal proliferation of myeloid progenitors in the bone marrow or peripheral blood. Currently, the main diagnostic methods for this neoplasm depend on the presence of at least 20% of malignant cells in the bone marrow or peripheral blood, which generally coincides with the advanced stage of the disease. In addition, these methodologies require expensive equipment and supplies, as well as specialized labor, which makes the diagnosis of AML restricted only to large diagnostic or research centers. Therefore, in this work, we implement the use of Fourier Transform Infrared Spectroscopy and attenuated total reflection (ATR-FTIR) univariate and coupled with artificial intelligence algorithms, as a fast, inexpensive tool with high sensitivity and specificity for patient discrimination with clinical and laboratory diagnosis of acute myeloid leukemia. For this, we performed a quantitative spectral analysis of plasma (1 μ l) of 35 patients and 35 control subjects, with a resolution of 4 cm^{-1} and 32 spectral scans in triplicate, using the air spectrum as the *background* of the analysis. As a result, ten vibrational modes were identified (3095 cm^{-1} , 1648 cm^{-1} , 1635 cm^{-1} , 1540 cm^{-1} , 1284 cm^{-1} , 1170 cm^{-1} , 1120 cm^{-1} , 1078 cm^{-1} , 1031 cm^{-1} and 850 cm^{-1}) with sensitivity and specificity $\geq 80\%$ ($p < 0.0001$) in discriminating the pathological samples in relation to the control group. Furthermore, two regions of the medium spectrum (1333 cm^{-1} -1272 cm^{-1} and 1096 cm^{-1} -996 cm^{-1}) also presented sensitivity and specificity values equal to or greater than 80% ($p < 0.0001$). Regarding the classification of spectra by artificial intelligence, logistic regression, linear discriminant analysis and random forest algorithms presented accuracy equal to or greater than 90% in at least one selected pre-processing set, with emphasis on the vector machine algorithm of support, which discriminated the control and LMA with 100% accuracy. In conclusion, univariate ATR-FTIR spectroscopy or coupled with artificial intelligence algorithms was able to discriminate the plasma of AML patients and controls, promising its use as a fast, low-cost test with high sensitivity and specificity for AML screening.

Key words: Leukemia; screening; vibrational modes; machine learning algorithms; ATR-FTIR spectroscopy.

LISTA DE FIGURAS

	Páginas
Figura 1. Representação esquemática da hematopoese (adaptado de Metcalf, 2007)	19
Figura 2. Taxa bruta estimada da incidência de leucemia no estado do Maranhão para o ano de 2020 (INCA 2020)	22
Figura 3. Esquema de funcionamento do interferômetro de Michelson (BENETI, 2014)	24
Figura 4. Modos de análises por espectroscopia no infravermelho (BAKER. 2014)	25

LISTA DE ABREVIATURAS

ATR	Attenuated Total Reflexao (reflexão total atenuada)
CTH's	Células tronco hematopoiética
EDTA	Ethylenediamine tetraacetic acid (ácido etilenodiamino tetra-acético)
FAB	French-American-British (Franco-americano-britânico)
FLT-3	FMS -fator estimulador de colônia de macrófagos- like tyrosine kinase
FTIR	Fourier Transform in Infrared (Transformada de Fourier no Infravermelho)
INCA	Instituto Nacional do Câncer
K-RAS	Kirsten rat sarcoma (sarcoma de rato Kirsten)
LA's	Leucemias agudas
LC's	Leucemias crônicas
LLA	Leucemia Linfoide Aguda
LMA	Leucemia Mieloide Aguda
LMA-CN	Leucemia Mieloide Aguda –Cariótipo Normal
N-RAS	Neuroblastoma- rat sarcoma (neuroblastoma- sarcoma de rato)
miRNA's	micro- Ribonucleic acid (micro- Ácido Ribonucleicos)
OMS	Organização Mundial da saúde

SUMÁRIO

1 Introdução.....	15
2 Fundamentação teórica.....	17
2.1 Sistema hematopoiético: hematopoese normal x Leucemogênese.....	17
2.2 Leucemia mieloide aguda.....	20
2.2.1 Classificação.....	21
2.2.2 Epidemiologia e diagnóstico.....	22
2.3 Espectroscopia de Infravermelho com Transformada de Fourier.....	23
2.4 Plasma: um biofluido por excelência.....	27
3 Objetivos.....	29
3.1 Objetivo geral.....	29
3.2 Objetivos específicos.....	29
4. Metodologia.....	30
4.1 Sujeitos da pesquisa e aspectos legais.....	30
4.2. Critérios de inclusão.....	30
4.3 Critérios de exclusão.....	30
4.4 Imunofenotipagem.....	31
4.5 Mensuração dos espectros por espectroscopia ATR-FTIR.....	31
4.6 Processamento dos dados.....	31
4.7 Algoritmos de aprendizado de máquina para análise multivariada.....	32
4.8 Análise estatística.....	32
5 Resultados (Artigo).....	33
6 Considerações finais.....	56
7 Conclusões.....	57
8 Referências bibliográficas.....	58
Anexos A.....	68
Anexo B.....	69