

**UNIVERSIDADE FEDERAL DO ESPÍRITO SANTO
CENTRO DE CIÊNCIAS EXATAS
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

**Biofluidos e espectrometria de massas para triagem de
pacientes para COVID-19**

Biofluids and Mass Spectrometry for COVID-19 Patient Screening

Camila Medeiros de Almeida

Tese de Doutorado em Química

**VITÓRIA
2024**

CAMILA MEDEIROS DE ALMEIDA

**Biofluidos e espectrometria de massas para triagem de
pacientes para COVID-19**

Tese apresentada ao Programa de Pós-Graduação em Química do Centro de Ciências Exatas da Universidade Federal do Espírito Santo como requisito parcial para obtenção do Título de Doutora em Química

Área de Concentração: Química

Linha de Pesquisa: Química Analítica

Orientador: Prof. Dr. Wanderson Romão

Coorientador: Prof. Dr. José Geraldo Mill

VITÓRIA
2024

Ficha catalográfica disponibilizada pelo Sistema Integrado de Bibliotecas - SIBI/UFES e elaborada pelo autor

A447b Almeida, Camila Medeiros de, 1991-
Biofluidos e espectrometria de massas para triagem de pacientes para COVID-19 / Camila Medeiros de Almeida. - 2024.

147 f. : il.

Orientador: Wanderson Romão.

Coorientador: José Geraldo Mill.

Tese (Doutorado em Química) - Universidade Federal do Espírito Santo, Centro de Ciências Exatas.

1. Química analítica. 2. Espectrometria de massa. 3. Quimiometria. 4. COVID-19. I. Romão, Wanderson. II. Mill, José Geraldo. III. Universidade Federal do Espírito Santo. Centro de Ciências Exatas. IV. Título.

CDU: 54

BIOFLUIDOS E ESPECTROMETRIA DE MASSAS PARA TRIAGEM DE PACIENTES PARA COVID-19

Camila Medeiros de Almeida

Tese submetida ao Programa de Pós-Graduação em Química do Centro de Ciências Exatas da Universidade Federal do Espírito Santo como requisito parcial para a obtenção do Grau de Doutor(a) em Química.

Aprovada em 28/02/2024 por:

Prof.(a) Dr.(a) Wanderson Romão
Orientador(a)
UFES

Prof.(a) Dr.(a) José Geraldo Mill
Coorientador(a)
UFES

Prof.(a) Dr.(a) Andrea Rodrigues Chaves
UFG

Prof.(a) Dr.(a) Luciene Cristina Gastalho Campos
UESC

Prof.(a) Dr.(a) Alvaro Cunha Neto
UFES

Prof.(a) Dr.(a) Paulo Roberto Filgueiras
UFES

Universidade Federal do Espírito Santo
Vitória, fevereiro de 2024





DocumentosBancaDefesaDoutorado_Camila Medeiros de Almeida

Data e Hora de Criação: 29/02/2024 às 10:15:47

Documentos que originaram esse envelope:

- FolhaAssinaturas_Camila Medeiros de Almeida.pdf (Arquivo PDF) - 1 página(s)
- FolhaRegistro_Camila Medeiros de Almeida.pdf (Arquivo PDF) - 1 página(s)
- AtaAprovação_Camila Medeiros de Almeida.pdf (Arquivo PDF) - 1 página(s)



Hashs únicas referente à esse envelope de documentos

[SHA256]: 5148c74103666898293bc3ecff8edc1662fb025b9d7710b091efe9fff8291e66

[SHA512]: 4d40e2bf3dd7446de2856ef43cc540fd0c482a6f48e7105555e9aadcf524be94c5c9f593edb59867cfa395875c726c71542fc99542700b06d15ffcb91bc7b223

Lista de assinaturas solicitadas e associadas à esse envelope



ASSINADO - Alvaro Cunha Neto (alvarocunhaneto@gmail.com)

Data/Hora: 01/03/2024 - 08:23:05, IP: 200.137.65.106

[SHA256]: 62bf7147435998685bba219474e760ac70b72c7ecf5a2930e1548638532b261a



ASSINADO - Andrea Rodrigues Chaves (andrea_chaves@ufg.br)

Data/Hora: 29/02/2024 - 19:09:05, IP: 177.158.31.222, Geolocalização: [-16.729204, -49.266585]

[SHA256]: b8f05058a65fcd17546cd6500b0487d564400e64df43682bcb9c385374025240



ASSINADO - José Geraldo Mill (jose.mill@ufes.br)

Data/Hora: 01/03/2024 - 20:30:56, IP: 45.236.240.226, Geolocalização: [-3.1028233, -60.023463]

[SHA256]: 56b1f58fccc608a48e95175d4b38abc18009756dbf73b9b4b024140d0ee1e27b



ASSINADO - Luciene Cristina Gastalho Campos (lcgcluiz@uesc.br)

Data/Hora: 11/03/2024 - 20:33:41, IP: 186.218.253.5

[SHA256]: 9427e34711dd29d1850b84bc2fcd0a36c76ebf834c7ad32f34edee32fc323098



ASSINADO - Paulo Roberto Filgueiras (paulo.filgueiras@ufes.br)

Data/Hora: 01/03/2024 - 16:50:35, IP: 200.137.65.107, Geolocalização: [-20.286241, -40.301194]

[SHA256]: a287803d5612e1c206df9bf8a4f5e448288ee384df48b5fc9582a822a48472b4

Paulo Roberto Filgueiras



ASSINADO - Wanderson Romão (wandersonromao@gmail.com)

Data/Hora: 01/03/2024 - 06:56:27, IP: 179.102.134.228, Geolocalização: [-20.275061, -40.284465]

[SHA256]: e48d8cf783d7fab76e896503fd0ead0350bfeb64e6c789bca7af768c5666c171

Wanderson Romão

Histórico de eventos registrados neste envelope

- 11/03/2024 20:33:42 - Envelope finalizado por lcgcluiz@uesc.br, IP 186.218.253.5
- 11/03/2024 20:33:42 - Assinatura realizada por lcgcluiz@uesc.br, IP 186.218.253.5
- 11/03/2024 20:33:22 - Envelope visualizado por lcgcluiz@uesc.br, IP 186.218.253.5
- 01/03/2024 20:30:56 - Assinatura realizada por jose.mill@ufes.br, IP 45.236.240.226
- 01/03/2024 20:30:42 - Envelope visualizado por jose.mill@ufes.br, IP 45.236.240.226
- 01/03/2024 16:50:35 - Assinatura realizada por paulo.filgueiras@ufes.br, IP 200.137.65.107
- 01/03/2024 16:49:55 - Envelope visualizado por paulo.filgueiras@ufes.br, IP 200.137.65.107
- 01/03/2024 08:23:05 - Assinatura realizada por alvarocunhaneto@gmail.com, IP 200.137.65.106
- 01/03/2024 06:56:27 - Assinatura realizada por wandersonromao@gmail.com, IP 179.102.134.228
- 29/02/2024 19:09:05 - Assinatura realizada por andrea_chaves@ufg.br, IP 177.158.31.222
- 29/02/2024 16:55:58 - Envelope registrado na Blockchain por michel.chaves@ufes.br, IP 200.137.65.102
- 29/02/2024 16:55:57 - Envelope encaminhado para assinaturas por michel.chaves@ufes.br, IP 200.137.65.102
- 29/02/2024 10:15:49 - Envelope criado por michel.chaves@ufes.br, IP 200.137.65.102

Dedico esse trabalho à minha amada família e às famílias de todos aqueles que perderam suas vidas por conta da doença COVID-19.

AGRADECIMENTOS

A Deus que me deu força para que eu não desistisse no meio do caminho, que foi meu consolo quando as coisas davam errado e minha gratidão quando eu encontrava o caminho certo. O meu maior companheiro durante toda essa nova e desafiante caminhada.

Aos meus pais, Lucas e Valéria, que me deram a vida e me ensinaram a desfrutá-la com dignidade e bondade, sempre apoiando todas as minhas escolhas. Vocês são minha base, com quem eu sei que posso contar em todos os momentos.

A minha irmã Isabela que me ajudou em vários momentos, juntamente com meu cunhado, Roney, sempre estiveram ao meu lado.

Aos meus amigos que tiravam o peso e a aflição quando a ansiedade batia e me mostraram que eu nunca estarei sozinha. Quem tem amigo tem tudo! Ao meu namorado, Bruno, que se tornou o meu melhor amigo. Você não sabe quão mais leve deixou minha caminhada. Te amo!

A minha madrinha, Ivana, que mesmo de longe se dedicou para me ajudar nessa etapa final, com todo seu cuidado, também deu um toque especial nesse lindo trabalho.

Ao meu orientador, Wanderson Romão, por me apresentar novos caminhos e pessoas incríveis que foram essenciais no desenvolvimento dessa pesquisa e por todos os ensinamentos, sendo o principal o amor pela ciência. E ao meu coorientador José Geraldo Mill por sempre estar disposto a colaborar para o meu desenvolvimento e a conclusão do trabalho.

Ao Laboratório Petrolômica e Forense por ser muito mais que um simples grupo de pesquisa, afinal, juntos, dividimos, o espaço, as cobranças, as tristezas, as experiências e principalmente as alegrias. Vocês tornam meus dias mais felizes.

Amanda, Luciara, Marcos e Natália, nessa reta final, vocês entraram na minha vida para alegrar, me acalmar e me tranquilizar, com risos, ensinamentos e muito companheirismo.

A Larissa obrigada por estar ao meu lado em quase todo doutorado, a você eu tenho muito mais que agradecer, choramos, rimos, brigamos, viajamos, caminhamos e até moramos juntas por um tempo. Você foi a irmã que a pesquisa me deu. Obrigada por acreditar e sempre fazer eu acreditar em mim. Você foi essencial.

Agradeço também a Gabriely pela paciência, gentileza e parceria desde o mestrado. Sem vocês duas com certeza esse trabalho não seria concluído.

Àqueles que eu conheci no laboratório e ganharam um lugar especial e eterno no meu coração, Bruna, Clebson, Eliane, Flávia, Fernandinha, Iago, Juliana, Lindamara, Natã, Nayara, Radigya, Rayana e Thieres obrigada pela amizade e apoio. E aos alunos novos, obrigada pela confiança e sucesso na caminhada.

Agradeço a Ricardo e Almir pela paciência e ensinamentos nas análises realizadas em Goiânia e a Rosiane e José nas análises em Brasília.

Aos professores PPGQUI, por todos os ensinamentos, em especial aos professores Álvaro, Valdemar e Paulo que me auxiliaram em diversos momentos.

Aos membros da banca, os professores Álvaro, Paulo, Andrea e Luciene, por aceitaram o convite e pelas considerações e pelo acréscimo neste trabalho.

A FAPES, CAPES CNPq pelo apoio financeiro. Ao Laboratório Tommasi (Vitória ES), Hospital Vila Velha, Hospital Universitário Cassiano Antônio Moraes, Pronto Atendimento da Glória e Hospital Roberto Arnizaut Silves pelo fornecimento das amostras e a todos os envolvidos na coleta das amostras.

Ao PPGQUI – UFES, especialmente ao coordenador prof Cleocir, por todo suporte durante esses anos.

Ao LabPetro e seus funcionários, Suzy, Carlão, Oscar, Cleide e Maria pela atenção e carinho.

A todos que em algum momento estiveram presentes, seja com um sorriso, um bom dia ou alguma ajuda para concretizar esse sonho, o meu MUITO OBRIGADA

"Os rios não bebem sua própria água; as árvores não comem seus próprios frutos. O sol não brilha para si mesmo; e as flores não espalham sua fragrância para si. Viver para os outros é uma regra da natureza. [...] A vida é boa quando você está feliz; mas a vida é muito melhor quando os outros estão felizes por sua causa".
(Papa Francisco)

"Coragem. Gentileza. Amizade. Essas são as qualidades que nos definem como seres humanos e nos impulsionam, ocasionalmente, à grandeza."
(Extraordinário)

"Devemos manter a nossa certeza de que depois dos dias ruins, os bons virão novamente."
(Marie Curie)

LISTA DE FIGURAS

Figura1. 1 Nuvem de palavras utilizando 727 artigos da busca de dados da Web of Science e PubMed usando as palavras-chave "mass spectrometry" AND biofluid* AND disease.	22
Figura1. 2. Ilustração da contribuição científica por países, demonstrada a partir da intensidade de cor, com a temática “biofluidos”, “MS” e “doenças”.	23
Figura1. 3. Diagrama de fluxo de trabalho para diagnósticos de doenças, detectando biomarcadores em biofluidos por MS.	24
Figura1. 4. Diagrama esquemático de um espectrômetro de massas.	43
Figura1. 5. Esquema do processo MALDI.	45
Figura1. 6. Processo de ionização da fonte Eletrospray.	46
Figura1. 7. Gráfico de scores e loads do modelo PCA de espectros MALDI(+) FT-ICR de amostras de saliva positivas ou negativas para SARS-CoV-2.	52
Figura2. 1 Espectros de massas MALDI(+) de amostras de saliva com e sem preparo de amostra: A) antes e B) após digestão trípica e processamento em coluna spin C18.	95
Figura2. 2 Espectros de massas MALDI(+) da amostra de saliva e A) Sem digestão trípica, (B- E) Após digestão trípica: B)10 µL tripsina e 2 h digestão trípica, C)10µL tripsina e 16h digestão trípica, D) 5 µL tripsina e 2 h digestão trípica, e E) 5 µL tripsina e 16 h digestão trípica	97
Figura2. 3 Gráfico de scores e loadings da PCA do conjunto de dados dos espectros de massas MALDI(+) de amostras de saliva positivas e negativas para SARS-CoV-2.	101
Figura2. 4 Sinal característico nos espectros MALDI(+) individuais entre amostras positivas para SARS-CoV-2 (verde) versus amostras negativas (cinza).	104
Figura3. 1. Espectros de ESI(±)-Orbitrap MS de extratos lipídicos do conjunto de amostras de soro humano após preparo de amostra para purificação de lipídios contendo (A) 239 amostras para ESI(+) e (B) 232 para ESI(-).	122
Figura3. 2. Gráficos biplot de PCA dos dados de ESI(±)-Orbitrap MS de extratos lipídicos de amostras de soro positivas ou negativas para SARS-CoV-2: (A) ESI(+) e (B) ESI; e gráficos de scores das componentes da PLS-DA	124

Figura3. 3. Parâmetros de validação cruzada para o modelo PLS-DA. R2, e Q2 representam as estimativas pela validação cruzada. A-B) ESI(+)-Orbitrap MS. C-D) ESI(-)-Orbitrap MS.	125
Figura3. 4. Sinais m/z discriminantes identificadas de acordo com a pontuação de Importância Variável na Projeção (VIP) (A) para dados de ESI(-) e (B) ESI(+)-Orbitrap MS.....	126
Figura3. 5. Classes lipídicas detectadas entre os 100 principais sinais VIP para dados de (A) ESI(+) MS, (B) ESI(-) MS e (C) ESI(±) MS. Estes são relevantes para distinguir entre classes positivas e negativas para a COVID-19.....	129
Figura3. 6. Espectros de MALDI(+)-TOF MS de amostra de soro humano (A) sem matriz (B); contendo matriz CHCA à 1 mg/mL (C) e à 10mg/mL.	133
Figura3. 7. 900 espectros de massas MALDI(+)-TOF MS de 300 amostras de soro humano após preparo de amostra utilizando Spin Column C18, sendo que 450 correspondem ao soro positivo pra COVID-19 (vermelho) e 450 ao soro negativo para COVID-19 (azul).	134
Figura3. 8. Análise estatística dados MALDI(+)-TOF MS. (A) Volcano plot indicando os sinais presentes no conjunto de amostras, com diferença estatística de acordo com as contagens do espectro, cada mancha representa um sinal. No eixo y, Log2 (valor p), no eixo x, Log2 (FC), os valores negativos indicam sinais mais relevantes em amostras negativas e valores positivos mostram sinais mais relevantes em amostras positivas. (B) Gráfico de PCA de espectros MALDI(+)-TOF MS de amostras de soro positivas e negativas para SARS-CoV-2. (C) Heatmap de espécies moleculares diferenciais em pacientes com COVID-19 em comparação com pacientes negativos.	135
Figura3. 9. Análise estatística de dados MALDI(+)-TOF MS. (A) O gráfico de scores de PLS-DA ilustra a segregação do perfil proteômico entre pacientes positivos para COVID-19 e controles. (B) Sinais m/z discriminantes baseados em VIP scores. Os 25 principais valores m/z com VIP scores >1,4. (C-D) Parâmetros de validação cruzada para PLS-DA, incluindo resultados R2 e Q2 estimados através de validação cruzada.	137

LISTA DE TABELAS

Tabela 1. 1. Artigos publicados nos últimos 5 anos sobre diagnósticos de doenças a partir do uso de MS e amostras de sangue.	27
Tabela 1. 2. Artigos publicados nos últimos 5 anos para diagnósticos de doenças a partir do uso de MS e amostras de saliva.	38
Tabela 1.3. Comparação entre os analisadores de massas.	48
Tabela 1. 4. Estudos utilizando as técnicas MALDI MS e LC-ESI MS e aprendizagem de máquina nas análises de biofluidos com SARS-CoV-2.	57
Tabela 2. 1 Dados dos participantes.	91
Tabela 2. 2 Peptídeos identificados em amostras de saliva após a digestão de proteínas e SPE.	96
Tabela 2. 3 Número de sinais detectados com S/N > 4%.	98
Tabela 2. 4 Sequência peptídica identificada em cada tipo de protocolo de proteína de digestão trípica.	99
Tabela 2. 5 Número de sinais detectados e tempo de aquisição dos espectros de massas do MALDI FT-ICR em função da otimização dos parâmetros.	100
Tabela 2. 6 . Desempenho dos parâmetros dos modelos SVM pelos espectros MALDI com Discriminante de Fisher (FD).	103
Tabela3. 1 Métricas do modelo SVM.	120
Tabela3. 2. Desempenho de Parâmetros de Modelos SVM com dados de ESI(±)-Orbitrap MS de amostra de soro humano, após o preparo de amostra para purificação de lipídios, com Discriminante de Fisher (FD).	127
Tabela3. 3. Desempenho de Parâmetros de Modelos SVM com espectros de massa MALDI(+)-TOF MS de amostra de soro humano após preparo e otimização das análises.	138

LISTA DE ABREVIATURAS

ACN – Acetonitrila

ADA, ADA tree boosting

AmBic - Bicarbonato de amônio

ANN - *Artificial Neural Network*

AO - osteoartrite primária

AR - artrite reumatoide

CHCA - *α -cyano-4-hydroxycinnamic acid*

DHB - *2,5-dihydroxybenzoic acid*

DIA - *Data Independent Acquisition*

DMTC - Doença mista do tecido conjuntivo

DNN - *deep neural network*

DPOC - Doença pulmonar obstrutiva crônica

DT - *Decision Tree*

DTT – Ditiotreitól

ESI - *Electrospray ionization*

ESPEC - Especificidade

FC - *Fold-Change*

FD - Discriminante de Fisher

FN - Falso negativo

FP - Falso positivo

FT-ICR - *Fourier-transform ion cyclotron resonance*

FWHM - *Full width at Half maximum*

GA - *Genetic Algorithm*

GBDT - *Gradient Boosted Decision Trees*

GBM - *Gradient Boost Machine*

GDB - *gradient tree boosting*

HPLC - *High-performance liquid chromatography*

IAA - Iodoacetamida

IgG - Imunoglobulina G

IgM - Imunoglobulina M

IOP - Insuficiência ovariana prematura

IPA- *Ingenious Path Analysis*

IT - *Ion trap*

KNNk - *Nearest Neighbors*

LC - *Liquid chromatography*

LDA - *Linear Discriminant Analysis*

LR - *Logistic Regression*

MALDI - *Matrix assisted laser desorption ionization*

MCC - Coeficiente de correlação de Matthews

ML - *Machine learning*

MLP - *Multilayer Perceptron*

MS - *Mass spectrometry*

NB - *Naïve Bayes*

nLVs - número ideal de variáveis latentes

OMS - Organização Mundial de Saúde

OPLS-DA - *Orthogonal Partial Least Squares Discriminant Analysis*

OPS - *Ordered predictor selection*

PC - *Principal Components*

PCA - *Principal component analysis*

PHS - Púrpura de Henoch-Schönlein

PLS-DA - *Partial least squares-discriminant analysis*

QC - *Quick Classifier*

Q-TOF – *Quadrupole time-of-flight*

RF - *Random Forest*

ROC - *Receiver Operating Characteristic*

RT-PCR – *Reverse Transcription Polymerase Chain Reaction*

SA - *Sinapinic acid*

SAFP - Síndrome antifosfolípide primária

SARS - Síndrome respiratória aguda grave

SENS - Sensibilidade

SNN - *Supervised Neural Network*

sPLS-DA - *Sparse partial least squares-discriminant analysis*

SVM - *support-vector machine*

TDAH - Transtorno de déficit de atenção e hiperatividade

TFA - Ácido trifluoroacético

TFN - Taxa de falso negativo

TFP - Taxa de falso positivo

TOF - *Time-of-flight*

TQ – triplo quadrupolo

UMAP - *uniform manifold approximation and projection*

UPLC - *Ultra Performance Liquid Chromatography*

UHPLC - *Ultra High Performance Liquid Chromatography*

Varsel - Variáveis selecionadas

VIP- *variable importance projection*

VN - Verdadeiro negativo

VNP - Valor negativo de prevalência

VP - Verdadeiro positivo

VPP - Valor positivo de prevalência

XGB - XGBoost

XRF- *extreme random forest*

LISTA DE SÍMBOLOS

% – Porcentagem

°C – grau Celsius

CHCl₃ - Clorofórmio

Da – Dalton

EtOH – etanol

G - força G

h - hora

Hz – Hertz

kV - kilovoltz

m – massa

m/Δm50% – poder de resolução de massa

MeOH – metanol

mg/mL – miligrama por mililitro

min - minutos

mL – mililitro

mM – miliMolar

ms - milisegundo

n – número de amostras

nm – Nanometro

ns - nanosegundos

rpm – rotação por minuto

S/N – sinal por ruido

scans – número de aquisições espectrais acumuladas

V – Voltz

v/v – volume por volume

z- carga

γ – gama

ε - margem

μg/μL – micrograma por microlitro

μL - microlitro

μM – micromolar

ns – nanosegundos

RESUMO

A doença COVID-19 foi e continua sendo uma preocupação na saúde mundial, a identificação de pacientes infectados em triagens rápidas e eficientes ainda são necessárias para conter a propagação. Os fluidos biológicos, como soro e saliva, oferecem facilidade de coleta e fornecem informações ricas sobre as alterações moleculares do corpo durante alguma doença. O uso da espectrometria de massas (do inglês *mass spectrometry*, MS) combinada com a aprendizagem de máquina (do inglês *machine learning*, ML) tem sido aplicado a biofluidos de pacientes portadores de doenças e controles, para a identificação de biomarcadores e realização de uma triagem rápida e eficaz. Assim, esta tese tem como objetivo apresentar os avanços na busca de biomarcadores de doenças, principalmente da COVID-19, utilizando tecnologias baseadas em ionização e dessorção a laser assistida por matriz (do inglês, *matrix assisted laser desorption ionization*, MALDI) MS e ionização por eletrospray (do inglês, *electrospray ionization*, ESI) MS e tratamentos de dados quimiométricos. Para isso, foi desenvolvida uma metodologia para triagens de pacientes com suspeita de COVID-19 a partir de amostras de saliva, utilizando MALDI MS mediante auxílio da aprendizagem por Máquina de Vetores de Suporte (do inglês, *support-vector machine*, SVM), otimizando o preparo de amostra e os parâmetros da análise. A maior eficiência em menor tempo de análise foi obtido com a digestão da saliva em 10 μ L de tripsina por 2 h e uso de 1M de resolução. Modelos SVM foram criados com os dados das análises de 149 amostras, sendo estas 97 positivas e 52 negativas para COVID-19 por RT-PCR. Dois modelos apresentaram os melhores resultados. O SVM1 selecionou 780 variáveis e possui taxa de falso negativo (TFN) de 0%, já o SVM2 selecionou somente 2 variáveis (525,4 Da e 1410,8 Da) com TFN de 3%. Outra aplicação da MS em biofluidos foi o desenvolvimento de um método multiômico para triagem de pacientes infectados com SARS-CoV-2 com base nos perfis lipídicos e proteômicos do soro. A ESI MS foi utilizada para investigar o perfil lipídico de 239 amostras de soro (119 positivas e 120 negativas para COVID-19 pelo teste ELISA). A MALDI MS foi utilizada para analisar o perfil proteômico de 300 amostras de soro (150 positivas e 150 negativas para COVID-19 pelo teste ELISA). Após o processamento dos dados de MS e a seleção de variáveis, análises estatísticas, como Volcano plot, o Heatmap, a análise de componentes principais (do inglês, *principal component analysis*, PCA), a análise discriminante de mínimos quadrados parciais (do inglês, *partial least squares-discriminant analysis*, PLS-DA) e a SVM, foram realizadas para distinguir as variáveis mais relevantes para classificar amostras positivas e negativas para COVID-19. Nas análises lipídicas usando ESI(\pm)-Orbitrap MS e modelos SVM, observou-se sensibilidades de 96,67% e 100%, especificidades de 82,14% e 96,88%, e acurácias de 89,66% e 98,44% para análises de modo de íon positivo e negativo, respectivamente. Já nas análises proteômicas usando MALDI(+)-TOF MS, o modelo linear PLS-DA demonstrou uma precisão de 99,10%. Sendo assim, a combinação das técnicas MS com tratamento de dados quimiométricos, demonstrou alternativas promissoras com alta sensibilidade e especificidade para discriminar amostras de biológicas infectadas e não infectadas pelo SARS-CoV-2.

Palavras-chave: Biofluidos. Espectrometria de massas. Aprendizagem de máquina. COVID-19.

ABSTRACT

The COVID-19 disease has been and continues to be a global health concern. The identification of infected patients through rapid and efficient screenings remains necessary to contain its spread. Biological fluids, such as serum and saliva, offer ease of collection and provide rich information about molecular changes in the body during illness. The use of mass spectrometry (MS) combined with machine learning (ML) has been applied to biofluids from patients with diseases and controls to identify biomarkers and conduct rapid and effective screenings. Therefore, this thesis aims to present advancements in the search for disease biomarkers, particularly for COVID-19, using technologies based on Matrix-Assisted Laser Desorption Ionization Mass Spectrometry (MALDI MS) and Electrospray Ionization Mass Spectrometry (ESI MS), along with chemometric data treatments. To achieve this, a methodology was developed for screening patients suspected of having COVID-19 based on saliva samples, using MALDI MS with the assistance of Support Vector Machine (SVM) learning. This involved optimizing sample preparation and analysis parameters. The most efficient results in a shorter analysis time were obtained by digesting saliva with 10 μ L of trypsin for 2 hours. Optimization of the parameters at 1M resolution was ideal for the analyses. SVM models were created using data from the analysis of 149 samples, 97 positive and 52 negative for COVID-19. Two models yielded the best results. SVM1 selected 780 variables with a false negative rate (FNR) of 0%, while SVM2 selected only 2 variables (525.4 Da and 1410.8 Da) with a 3% FNR. Another application of MS in biofluids was the development of a multiomic method for screening patients infected with SARS-CoV-2 based on serum lipid and proteomic profiles. ESI MS was used to investigate the lipid profile of 239 serum samples (119 positive and 120 negative for COVID-19). MALDI MS was used to analyze the proteomic profile of 300 serum samples (150 positive and 150 negative for COVID-19). After processing MS data and variable selection, statistical analyses such as Volcano plot, Heatmap, principal component analysis (PCA), partial least squares discriminant analysis (PLS-DA), and SVM were performed to distinguish the most relevant variables for classifying positive and negative samples for COVID-19. In lipidomic analyses using ESI(\pm)-Orbitrap MS and SVM models, sensitivities of 96.67% and 100%, specificities of 82.14% and 96.88%, and accuracies of 89.66% and 98.44% were observed for positive and negative ion mode analyses, respectively. In proteomic analyses using MALDI(+) MS, the linear PLS-DA model demonstrated an accuracy of 99.10%. Thus, the combination of MS techniques with chemometric data treatments has shown promising alternatives with high sensitivity and specificity to discriminate infected and non-infected biological samples by SARS-CoV-2.

Keywords: Biofluids. Mass spectrometry. Machine learning. COVID-19.

SUMÁRIO

CAPÍTULO I – Estado da arte	20
1.1 INTRODUÇÃO	20
1.2 BIOFLUIDOS.....	21
1.2.1 Sangue	25
1.2.2. Saliva	35
1.3 ESPECTROMETRIA DE MASSAS	42
1.3.1 MALDI MS.....	44
1.3.2 ESI MS.....	46
1.3.3 Analisadores de massas	48
1.3.4 Processamento de dados e análise estatística.....	50
1.4 COVID-19.....	53
1.4.1 COVID e MS	54
1.5 CONCLUSÃO	66
1.6 REFERÊNCIAS	67
CAPITULO II – MALDI(+) FT-ICR combinada com aprendizagem de máquina para triagem diagnóstica para COVID-19 baseada em amostras de saliva.....	87
2.1 INTRODUÇÃO	88
2.2 METODOLOGIA	90
2.2.1 Amostras e reagentes	90
2.2.2 Análise por MALDI-FT-ICR MS	92
2.2.3 Análise de dados.....	92
2.3 RESULTADOS E DISCUSSÃO	94
2.4 CONCLUSÃO	106
2.5 REFERÊNCIAS	107

CAPITULO III – Triagem para diagnóstico para COVID-19 baseado em dados multiômicos por espectrometria de massas de alta resolução (MALDI(+)-TOF MS e ESI(±)-Orbitrap MS)	112
3.1 INTRODUÇÃO	113
3.2 METODOLOGIA	116
3.2.1 Amostras e reagentes	116
3.2.2 Extração de lipídios	116
3.2.3 Preparo das amostras para análise MALDI(+)-TOF MS	117
3.2.4 Análise por ESI(±)-Orbitrap MS	117
3.2.5 Análise por MALDI(+)-TOF MS	118
3.2.6 Processamento de dados e análise estatística de dados	118
3.3 RESULTADOS E DISCUSSÃO	121
3.4 CONCLUSÃO	141
3.5 REFERÊNCIAS	142

CAPÍTULO I – Estado da arte

1.1 INTRODUÇÃO

A combinação de biofluidos e espectrometria de massas (do inglês, *mass spectrometry*, MS) surge como uma abordagem inovadora e promissora para a triagem eficaz de pacientes suspeitos de algumas doenças, incluindo a COVID-19, quando a rapidez do diagnóstico é fundamental para adoção de medidas que evitem a proliferação da doença. Diante da necessidade urgente de métodos diagnósticos rápidos e precisos, essa interseção entre a análise de biofluidos e a tecnologia de MS tem se destacado como uma ferramenta valiosa na identificação de biomarcadores associados à infecção pelo SARS-CoV-2.¹

Biofluidos, como sangue e saliva, oferecem uma fonte rica de informações sobre o estado fisiológico do organismo e, conseqüentemente, podem conter marcadores específicos da presença de alguma doença. A MS, por sua vez, é uma técnica analítica poderosa que permite a detecção de diversas moléculas em uma só análise, com base em sua massa molecular. Ao combinar essas duas abordagens, a avaliação dos padrões dos perfis moleculares distintos de biofluidos associados à COVID-19 proporcionam uma perspectiva única para o diagnóstico precoce e preciso da doença.²

Por fornecer uma enorme quantidade de dados, a combinação da MS com a quimiometria é crucial para obtenção de resultados mais rápidos e confiáveis. Essa combinação inovadora representa um passo essencial no desenvolvimento de estratégias diagnósticas mais eficazes para combater a propagação da COVID-19, oferecendo a possibilidade de intervenções rápidas e personalizadas para melhor gerenciamento clínico e controle da doença.³

Nesta interseção entre biofluidos e MS, é possível analisar uma variedade de biomarcadores, incluindo proteínas, lipídios e seus metabólitos, que podem refletir as alterações bioquímicas induzidas pela infecção viral.⁴ À medida que a pesquisa avança nesse campo, a integração de tecnologias de ponta na análise de biofluidos e MS promete aprimorar ainda mais a sensibilidade e a especificidade dos métodos de triagem não só da COVID-19, mas também de diferentes doenças.

1.2 BIOFLUIDOS

Os biofluidos, como soro, plasma, saliva, líquido cefalorraquidiano, urina e meios condicionados pela medula óssea, são originários do interior do corpo humano, sendo fontes importantes de informações sobre as condições fisiológicas dos indivíduos. Portanto, apresentam um grande número de moléculas como proteínas, lipídios, peptídeos e substâncias compostas por componentes excretados pelas células ao longo do corpo. Esses biofluidos ricos em informações químicas podem conter moléculas que, quando identificadas ou quantificadas, podem indicar a ocorrência de alterações moleculares que ocorrem no corpo em condições de doença. Essas alterações nas moléculas relacionadas à doença podem então ser usadas como biomarcadores de doenças.⁵

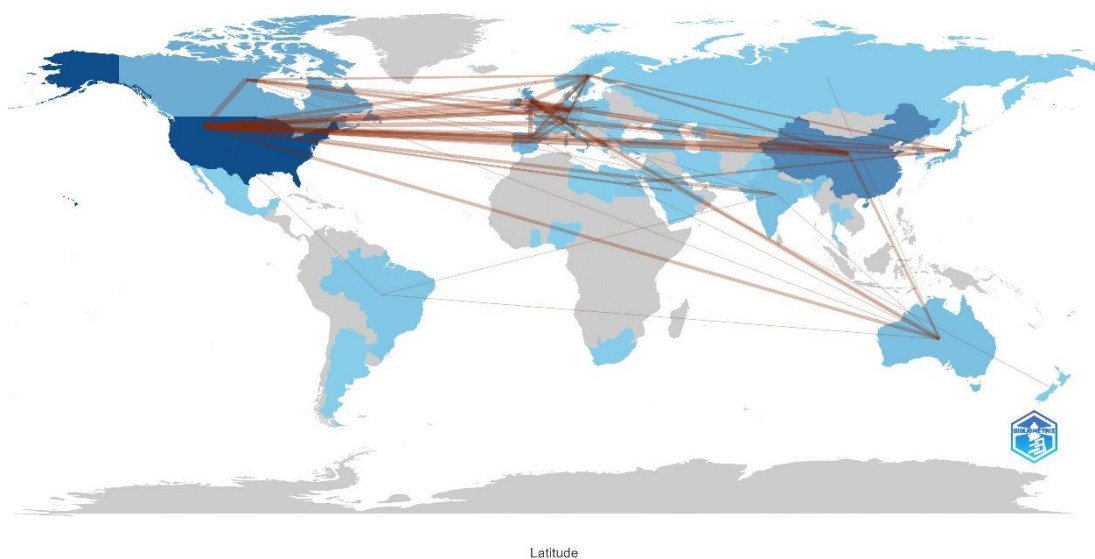
Os biomarcadores são características mensuráveis e objetivas que podem ser indicadores de processos biológicos normais, patológicos ou respostas a uma intervenção terapêutica. Podem ser usados principalmente para avaliar o estado de saúde, a presença de uma doença ou fazer previsões prognósticas. Muitas destas substâncias podem ser detectadas em estágios precoces de uma doença, quando as manifestações clínicas ainda não são tão evidentes. Dependendo do marcador, diferentes tipos de amostras biológicas podem ser utilizadas com esta finalidade.⁶

Alguns biofluidos estão entre as formas mais fáceis de obtenção das amostras clínicas. A coleta de saliva e urina são processos minimamente invasivos e de baixo custo, enquanto o sangue representa a espécime biológica mais utilizada na prática clínica, para realização de exame clínico de rotina.⁷

A necessidade de detectar biomarcadores de doenças em amostras de pacientes de forma rápida e precoce, é de extrema importância no combate a surtos e epidemias. A detecção da doença e diagnóstico de pacientes é uma ferramenta crítica para monitorar a disseminação da doença, orientar decisões terapêuticas e profiláticas e elaborar protocolos de distanciamento social, quando necessário.⁸

O desenvolvimento de técnicas rápidas e de resultados confiáveis, fez com que o uso da MS de rápida aquisição se tornasse uma ferramenta promissora para pesquisa e diagnósticos de doenças, devido à flexibilidade da técnica que permite a análise proteômica, metabolômica e lipidômica em uma ampla gama de matrizes biológicas, incluindo os biofluidos.^{4,5}

Figura1. 2. Ilustração da contribuição científica por países, demonstrada a partir da intensidade de cor, com a temática “biofluidos”, “MS” e “doenças”.



Fonte: *Bibliometrix*.⁹

As diversas aplicações da MS para análises de biofluidos abre um campo de oportunidades para identificações de possíveis biomarcadores de diferentes doenças. O avanço no estudo na área da proteômica, metabolômica e lipidômica auxilia no processo de diagnósticos rápidos e precoces.

Na Figura1. 3 observa-se um esquema para o processo de possíveis diagnósticos utilizando biofluidos e MS. Os diferentes biofluidos, por serem amostras complexas, necessitam de um preparo antes da obtenção dos espectros, para que não haja supressão iônica dos analitos de interesse. Após a aquisição dos dados espectrais o uso de ferramentas computacionais auxilia na rapidez e eficiência do tratamento de dados.¹⁰⁻¹²

Figura1. 3. Diagrama de fluxo de trabalho para diagnósticos de doenças, detectando biomarcadores em biofluidos por MS.



Fonte: Adaptado de Mahmud *et al.*¹⁰ e Dunphy *et al.*¹¹

A qualidade do preparo da amostra é um fator importante para o sucesso de todo procedimento analítico. A extração dos analitos de matrizes biológicas complexas, como soro, plasma, saliva, entre outras, removerá componentes da matriz que poderão interferir na análise e, no caso de analitos de baixa abundância, o procedimento pode incluir uma etapa de pré-concentração para atingir os limites de detecção da técnica analítica aplicada e obter melhores relações sinal-ruído.^{13–15}

Alguns dos preparos mais utilizados em análises de biofluidos por MS são: i) a extração em fase sólida, remove analitos interferentes de uma matriz líquida, muito utilizado na remoção de sais;¹⁶ ii) extração líquido-líquido, separa em diferentes fases os compostos de interesse, como na aplicação do protocolo *Bligh e Dyer* em estudo lipidômico;¹⁷ iii) precipitação de proteínas, usada tanto para remoção de proteínas em estudos metabolômicos, quanto para sua concentração em estudos proteômicos;¹⁸ iv) filtros de massa molar, utilizados em estudos de moléculas de massa molar mais baixa como peptídeos, aminoácidos e lipídios;¹⁹ v) digestão de proteínas, para análise do perfil proteômico;²⁰ entre outros. O método de preparo da amostra escolhido dependerá do analito de interesse e da matriz biológica utilizada.^{13–15}

1.2.1 Sangue

O sangue é um tecido circulante que transporta gases (oxigênio e gás carbônico) e outras moléculas para todos os órgãos e tecidos do organismo. Caracteriza-se por apresentar uma matriz líquida (plasma), onde se encontram suspensos os elementos celulares do sangue (hemácias, leucócitos e plaquetas).²¹

O plasma é a parte líquida do sangue não coagulado, sem as células sanguíneas, é composto por água e várias substâncias dissolvidas. No plasma são encontrados: eletrólitos (íons sódio, potássio, cálcio, magnésio, cloreto, bicarbonato, fosfato, entre outros) que desempenham papéis cruciais na regulação do equilíbrio hídrico e na função celular, proteínas, como a albumina (que contribui para a pressão osmótica do sangue, ajudando a manter o equilíbrio de fluidos entre o sangue e os tecidos), globulinas (que incluem as imunoglobulinas (anticorpos) e outras), fibrinogênio (que desempenha um papel crucial na coagulação sanguínea), dentre outros compostos químicos. Além disso, também estão presentes no plasma: nutrientes como, glicose, aminoácidos, lipídios (que fornecem energia e auxiliam nas funções celulares); produtos do metabolismo celular, hormônios, gases sanguíneos, vitaminas, enzimas e diversas substâncias químicas necessárias para a homeostase e funcionamento adequado do corpo.²¹

O soro é semelhante ao plasma em sua composição, porém não possui os fatores de coagulação sanguínea. Para a obtenção do soro, coleta-se o sangue na ausência de anticoagulantes, e nessas condições, forma-se um coágulo de fibrina que é então removido, restando o soro com uma concentração de proteínas um pouco inferior à do plasma.²²

Assim, mesmo com uma coleta um pouco mais invasiva comparada à coleta de saliva e urina, o soro é um fluido corporal frequentemente utilizado em rotinas de exames clínicos, por melhor refletir o estado fisiológico global de um organismo. Podendo ser uma excelente fonte para descoberta de novos biomarcadores de doenças devido ao rico repertório de biomoléculas provenientes de todo o corpo humano.^{11,23}

Mesmo sendo um biofluido rico em informação molecular, identificar biomarcadores de doenças no proteoma sérico pode ser desafiador, apesar de apresentar uma alta concentração de proteínas, 99% do seu conteúdo proteico correspondem a proteínas do próprio sangue (albumina, transferrinas,

imunoglobulinas). O outro 1% geralmente é composto por proteínas de maior interesse, as proteínas secretadas pelos tecidos.¹¹

Técnicas como a MS podem ser ferramentas promissoras na pesquisa de proteomas e, principalmente, de biomarcadores devido à sua alta seletividade, sensibilidade e versatilidade.²⁴ No entanto, essa abundância de proteínas é um fator limitante no resultado das análises.

Independente da técnica de ionização utilizada, a ionização das proteínas predominantes pode acarretar uma forte supressão de íons causada, por exemplo, pela abundância da albumina, esse fenômeno resulta em perda de informações sobre os possíveis biomarcadores, que geralmente são os objetos de interesse na pesquisa da doença.²⁵

Portanto, fluxos de trabalho dedicados à análise de biomoléculas no soro por MS incluem algumas etapas de pré-tratamento de amostras destinadas à redução da complexidade da amostra, reduzindo ou eliminando as proteínas altamente abundantes e favorecendo os analitos de maior interesse.²⁶ A sensibilidade da análise MS depende da natureza da amostra, conseqüentemente, a etapa de preparação para obter amostras purificadas, concentradas e ricas em moléculas de interesse é crucial.²⁷

Técnicas como a extração, precipitação e filtros de corte de massa molar, têm sido aplicadas para melhorar a qualidade da análise por MS, aumentando a sensibilidade e eficiência de identificação de biomarcadores em análises de plasma e soro.^{15,27} A Tabela 1. 1 mostra alguns dos artigos publicados nos últimos 5 anos que utilizam a MS para diferenciar amostras de sangue de indivíduos saudáveis de amostras de pacientes com diferentes tipos de doenças.

Tabela 1. 1. Artigos publicados nos últimos 5 anos sobre diagnósticos de doenças a partir do uso de MS e amostras de sangue.

Amostra	Preparo de amostra	Equipamento	Tratamento de dados	Resultados	Ano
228 amostras de plasma (55 controles, e 173 pacientes com doenças autoimunes sistêmicas)	Precipitação das proteínas: Adição de EtOH:MeOH e centrifugação. O sobrenadante coletado, evaporado e reconstituído com H ₂ O (ácido fórmico 0,1%) e MeOH.	HPLC-ESI-QTOF MS (Agilent Technologies)	<i>Software:</i> R Tratamento de dados: PCA, PLS Banco de dados: METLIN, LipidMaps, KEGG, massbank, e HMDB	Associação da maioria das patologias das doenças autoimunes sistêmicas investigadas com alterações metabólicas no plasma. Os metabólitos que distinguem entre controles e doenças autoimunes sistêmicas incluíam ácidos graxos insaturados, acilcarnitinas, acilglicinas e aminoácidos.	2020 ²⁸
25 amostras de soro (5 controles e 20 pacientes com diferentes estágios de doença pulmonar obstrutiva crônica DPOC).	A extração de lipídios: Adição de clorofórmio e MeOH contendo BHT, agitação e centrifugação. As camadas inferiores foram lavadas com um solução de KCl.	ESI TQ MS/MS (Biosystems)	<i>Software:</i> <i>Metaboanalyst</i> Tratamento de dados: OPLS e curva ROC	Um painel de 50 metabólitos lipídicos, incluindo fosfolipídios, esfingolipídios, glicerolipídios e ésteres de colesterol, pode ser usado para diferenciar a presença de DPOC. Análises multivariadas sugerem que quatro espécies lipídicas poderiam ser potenciais biomarcadores para DPOC.	2020 ²⁹
38 amostras de soro (10 controles e 28 pacientes com fibrose cística)	Extração de proteínas: Utilização do kit de depleção de albumina e IgG de esferas. A concentração proteica de cada amostra foi determinada utilizando o kit 2D-Quant.	MALDI- TOF MS/MS (Bruker Daltonics) UPLC- TQ MS/MS (Waters Corporation)	<i>Software:</i> <i>Metaboanalyst</i> Tratamento de dados: PLS-DA, curva ROC Banco de dados: MASCOT. IPA	Proteínas relacionadas à inflamação e reparos teciduais, como antiproteases, foram perturbadas nos soros de fibrose cística. As proteínas de transporte de vitamina A e D e as lipoproteínas foram reguladas negativamente, sugerindo possíveis explicações para as suas deficiências na fibrose cística.	2020 ³⁰
161 amostras de soro (64 controles, 67 pacientes com câncer da bexiga e 30 pacientes com presença de cálculos no trato urinário).	Purificação das amostras: Utilização de esferas magnéticas de troca catiônica fraca (WCX-MB, Bioyong). Extração de lipídios: método MTBE, agitação, centrifugação e uso de concentrador de amostra.	MALDI Clin-TOF II (Bioyong), LC- Orbitrap MS/MS (Thermo Fisher Scientific)	<i>Software:</i> BioExplorer Tratamento de dados: teste de Wilcoxon, KNN e curva ROC Banco de dados: proteoma humano (UniprotKB)	Foi contruido um modelo de diagnóstico contendo cinco peptídeos. Para o câncer de bexiga em estágio inicial, a sensibilidade e a especificidade foram de 92,31% e 93,75%, respectivamente; a sensibilidade do câncer de bexiga precoce de baixo grau foi de 90,00%; e o valor da AUC foi 0,944.	2020 ³¹

Amostra	Preparo de amostra	Equipamento	Tratamento de dados	Resultados	Ano
259 amostras de soro (100 controles, 60 pacientes com comprometimento cognitivo leve e 99 com doença de Alzheimer).	Separação de peptídeos e proteínas: eletroforese em gel de poliacrilamida BLOTCHIP® Extração de peptídeos: cartucho de extração em fase sólida Sep-Pak C18 e um concentrador	MALDI-TOF/TOF MS (Bruker Daltonics) e LC-MS/MS (Thermo Fisher Scientific).	<i>Software:</i> ClinProTools e R Tratamento de dados: teste de Wilcoxon, curva ROC Banco de dados: MASCOT e Swiss-Prot	Foi descoberto um novo biomarcador diagnóstico com conjunto de 4 peptídeos para discriminar controle, comprometimento cognitivo leve e Alzheimer com 87% de sensibilidade e 65% de especificidade entre controle e DA.	2020 ³²
105 amostras de soro (35 controles, 35 de pacientes com artrite reumatóide (AR), 35 de pacientes com osteoartrite primária (AO))	Separação de peptídeos de baixo massa molar: Utilização de esferas de cromatografia de interação hidrofóbica com esferas magnéticas (MB-HIC C8).	MALDI-TOF/TOF MS (Bruker Daltonics)	<i>Software:</i> ClinPro Tools 3.0. Tratamento de dados: SNN, GA e QC.	113 sinais discriminaram AR de AO, 101 sinais discriminaram AR de controles e 95 sinais discriminaram OA de controles. O perfil peptidômico é valioso para diferenciar indivíduos com AR de OA e controles saudáveis.	2020 ¹⁹
41 (16 controles e 25 pacientes com câncer colorretal)	Precipitação de proteínas: Adição de MeOH:ACN. Sobrenadante coletado e filtrado.	UPLC-QTOF-ESI MS (Thermo Fisher Scientific)	<i>Software:</i> <i>Metaboanalyst</i> 5.0 Tratamento de dados: fold change, Vulcano plot, PCA, PLS-DA e RF e curva ROC Banco de dados: LIPID MAPS Lipidomics Gateway e HMDB	Várias subclasses de lipídios, incluindo fosfatidilglicerol, ácidos graxos e ésteres de esteroide, bem como ceramidas, confirmaram o “fenótipo lipogênico” específico para o desenvolvimento do câncer colorretal, ou seja, a lipogênese regulada positivamente associada à progressão tumoral.	2021 ³³
22 amostras de plasma (11 controles e 11 pacientes com transtorno depressivo maior).	Extração de metabólitos: Adição de MeOH agitação e centrifugação. Ao sobrenadante adição de H ₂ O e clorofórmio, agitação e centrifugação. Fase superior filtrada e analisada.	UHPLC-QTOF-ESI MS (ThermoFisher Scientific) (Bruker Daltonics)	<i>Software:</i> DataAnalysis e <i>Metaboanalyst</i> Tratamento de dados: <i>Fold Change</i> , <i>Volcano Plot</i> , PCA, PLS-DA, RF e curva ROC Bases de dados: LIPID MAPS ® Lipidomics Gateway e HMDB Metabolome Humano	Perfil metabolômico não direcionado em amostras de plasma de pacientes com transtorno depressivo maior antes e depois do tratamento com escitalopram e controles saudáveis. Foi encontrado níveis significativamente aumentados de fosfatidilserina (PS (16:0/16:1)) e ácido fosfatídico (PA (18:1/ 18:0)) e diminuição dos níveis de PS (18:3/20:4) após tratamento antidepressivo.	2021 ³⁴
57 amostras de plasma (32 controles e 43 indivíduos com insuficiência ovariana)	Extração de metabólitos e lipídios: Adição de clorofórmio:MeOH:H ₂ O, agitação	LC-QTOF ESI MS (Agilent Zorbax)	<i>Software:</i> XCMS e <i>Isotopologue Parameter Optimization</i> Banco de dados: HMDB	Dois grupos foram separados com sucesso através da PCA. Entre os picos, fenil alanina, decanoil-L-carnitina, 1-palmitoil lisofosfatidilcolina e PC (O-16:0/2:0) foram identificados como biomarcador	2021 ³⁵

Amostra	Preparo de amostra	Equipamento	Tratamento de dados	Resultados	Ano
prematura)	e centrifugação. Na fase superior (metabólitos) e na fase inferior (lipídios).			plasmático para insuficiência ovariana prematura e monitoramento dos pacientes em período de tratamento.	
146 amostras de soro (146 COVID-19 e 152 controles (73 com sintomas clínicos semelhantes, 33 pacientes com tuberculose e 46 controles saudáveis)).	Diluição em tampão (kit PMFpre, Bioyong).	MALDI-TOF MS (Bioyong Technologies)	<i>Software: Metaboanalyst 4.0</i> Tratamento de dados: LR, SVM, NB, DT, RF, GBDT, KNN.	Um modelo de aprendizado de máquina de LR com 25 sinais de recursos alcançou precisão de 99%, com sensibilidade de 98% e especificidade de 100%, para detecção de COVID-19.	2021 ³⁶
173 amostras de soro (76 controles e 97 pacientes com artrite reumatóide (AR))	Extração de proteínas e pequenos peptídeos: Utilização de esferas magnéticas de troca catiônica.	MALDI-TOF MS (PBS IIc; Ciphergen Biosystems).	<i>Software: Ciphergen ProteinChip e SPSS.</i> Tratamento de dados: teste t de Student e o teste U de Mann-Whitney	Série de modelos diagnósticos envolvendo AR e AR com complicações. Além disso, a proteína C-C motif chemokine 24 (CCL24) pode ser um valioso indicador auxiliar de diagnóstico para AR.	2021 ³⁷
69 amostras de soro (32 controles e 37 pacientes cirróticos com carcinoma hepatocelular)	Precipitação de proteínas: Adição de MeOH, agitação e centrifugação. O sobrenadante foi coletado e filtrado e analisado.	UHPLC-QTOF-ESI MS (ThermoFisher Scientific) (Bruker Daltonics)	<i>Software: Metaboanalyst</i> Tratamento de dados: PCA, PLS-DA, ANOVA, curvaROC Banco de dados: <i>Lipid Maps</i> , HMDB e <i>PubChem</i> .	1,25-dihidroxicolesterol, palmitato de miristila, 25-hidroxitamina D2, ácido 12-cetodesoxicólico, lysoPC (21:4) e lysoPE (22:2) representam biomarcadores notáveis que diferenciam a cirrose do carcinoma hepatocelular precoce.	2022 ³⁸
86 amostras de plasma (30 controles e 26 com síndrome metabólica (SM) pré-diabéticos sem tratamento e 30 pacientes com SM normoglicêmica).	Extração de metabólitos: adição de MeOH, agitação e centrifugação. As amostras ressuspensas com de ácido fórmico a 0,1% em H ₂ O e filtradas.	UHPLC-QTOF MS (Bruker Daltonik).	<i>Software MetaboScape e Metaboanalyst</i> Tratamento de dados: sPLS-DA, ANOVA Banco de dados: HMDB	A análise ANOVA unidirecional revelou que 59 metabólitos diferiram significativamente entre os três grupos. Glutamina, 5-hidroxi-L-triptofano, L-sorbose e hipurato foram altamente associados à SM. No entanto, o ácido 9-metilúrico, a esfinganina e o ácido treônico foram altamente associados ao pré-diabetes/SM.	2022 ³⁹
129 amostras soro (51 controles e 78 pacientes com diferentes estágios de carcinoma de células renais)	Precipitação de proteínas: Adição de MeOH e H ₂ O resfriado. Incubação e centrifugação. O sobrenadante foi misturado com MeOH e de ácido fórmico 0,1%.	ESI-QTOF MS (Bruker Daltonics)	<i>Software: DataAnalysis, Metaboanalyst</i> Tratamento de dados: OPLS-DA Banco de dados: HMDB e METLIN.	A maioria dos metabólitos detectados foi regulada negativamente em pacientes com câncer. Uma alteração do nível de metabólitos plasmáticos exibiu uma correlação com a progressão do tumor.	2022 ⁴⁰

Amostra	Preparo de amostra	Equipamento	Tratamento de dados	Resultados	Ano
48 amostras de plasma (15 controles, 20 pacientes com mieloma múltiplo e 13 pacientes com leucemia de células plasmáticas)	Remover detritos celulares: centrifugação. Extração de proteínas indesejáveis: Mistura com ACN:H ₂ O, sonicação e centrifugação. O Sobrenadante foi analisado.	MALDI-TOF MS (Shimadzu)	<i>Software:</i> R pacote MALDIquant, MALDIrppa Tratamento de dados: PCA, PLS-DA, RF, DT e ANN	O protocolo de extração em duas etapas melhorou a análise de amostras de plasma de sangue periférico de gamopatia monoclonal por MS e forneceu uma ferramenta para abordar a complexa etiologia molecular das gamopatias monoclonais.	2023 ⁴¹
222 amostras de soro (83 controles e 139 pacientes com diferentes estágios de câncer cervical)	Purificação com coluna Ziptip C18 (Millipore, Burlington, MA, EUA).	MALDI-TOF/TOF MS (Bruker Daltonics)	<i>Software:</i> FlexAnalysis, ClinProTools Tratamento de dados: PCA, GA, ANOVA e SNN	A PCA mostrou que todos os grupos foram distribuídos separadamente, padrões de peptidoma sérico podem servir como ferramentas diagnósticas para o diagnóstico de câncer cervical em diferentes estágios.	2023 ⁴²
25 amostras de plasma (10 obesos (Ob), 8 pacientes com obesidade e diabetes (OD) e 7 pacientes obesos e diabéticos em tratamento com metformina)	Extração de Proteínas: uso de coluna Pierce™ Top 12. A concentração proteica determinada utilizando o kit 2D-Quant.	MALDI- TOF MS/MS (Bruker Daltonics)	<i>Software:</i> BioTools e Progenesis Same Spots Tratamento de dados: PCA IPA e ANOVA Banco de dados: Mascot e PANTHER	Possíveis biomarcadores alterados pelo tratamento com metformina em pacientes obesos com e sem diabetes mellitus tipo 2. Mudanças significativas em dezesseis proteínas plasmáticas foram observadas no grupo diabéticos em tratamento com metformina quando comparado aos grupos Ob e OD	2023 ⁴³
20 amostras de soro (10 controles e 10 crianças com diagnóstico de TDAH).	Remoção de múltiplas afinidades HSA / IgG: Utilização do cartucho giratório e coluna com resina de agarose Proteína A colocadas em uma coluna vazia do Bio-Spin®. Utilização de concentradores centrífugos	MALDI-TOF MS (Bruker Daltonics)	<i>Software:</i> flexAnalysis Excel e <i>GlycoWorkbench</i>	Nas amostras de pacientes com TDAH, foi observado um aumento da fucosilação antenal, diminuição dos N-glicanos di- / triantenários com bissecção da N-acetilglucosamina (GlcNAc) e diminuição da sialilação de α 2-3.	2023 ⁴⁴
60 amostras de soro (22 controles e 38 pacientes com púrpura de Henoch-Schönlein)	Extração de proteínas e pequenos peptídeos: Utilização de esferas MB-WCX com um separador magnético,	MALDI-TOF MS (Bruker Daltonics) nano UPLC –ESI	<i>Software:</i> ClinProTools, SPSS e GraphPad Prism. Tratamento de dados: Testes t, ANOVA	Identificação de potenciais biomarcadores peptídicos séricos para PHS em crianças chinesas Han; cinco proteínas foram identificadas como potenciais biomarcadores séricos.	2023 ⁴⁵

Amostra	Preparo de amostra	Equipamento	Tratamento de dados	Resultados	Ano
e pós-terapia)		MS/MS (Waters Corporation) (Thermo Fisher Scientific)	unidirecional, testes U de Mann-Whitney e testes H de Kruskal-Wallis e curva ROC. Banco de dados: UniProt		
144 amostras de soro (48 controles, 48 pacientes com câncer de fígado e 48 pacientes com hepatite B crônica)	Adição de (MeOH/diclorometano). Adição de padrão: TG 51:0 e centrifugação	MALDI-TOF/TOF MS (AB Sciex) com matriz de material de carbono grafitado	<i>Software:</i> mzMine, Origin e SIMCA, <i>Metaboanalyst</i> , Python Tratamento de dados: ANOVA, curva ROC, PLS-DA, LDA, LR, SVM, MLP, RF	Para a classificação de controles, câncer de fígado e hepatite B crônica, a LDA superou outras quatro técnicas de aprendizado de máquina com precisão, sensibilidade e especificidade de 92,3%, 92,3% e 96,2%, respectivamente. A análise ANOVA unidirecional revelou que numerosos TGs foram regulados negativamente no grupo com câncer de fígado.	2024 ⁴⁶

curva ROC (*Receiver Operating Characteristic*); DT (*Decision Tree*); ESI (*Electrospray ionization*); GA (*Genetic Algorithm*); GBDT (*Gradient Boosted Decision Trees*); HPLC (*High-performance liquid chromatography*); IPA (*Ingenious Path Analysis*); KNNk (*Nearest Neighbors*); LC (*Liquid chromatography*); LDA (*Linear Discriminant Analysis*); LR (*Logistic Regression*); MALDI (*Matrix assisted laser desorption ionization*); MLP (*Multilayer Perceptron*); MS (*Mass spectrometry*); OPLS-DA - *Orthogonal Partial Least Squares Discriminant Analysis* OPS (*Ordered predictor selection*); PCA (*Principal component analysis*); PLS-DA (*Partial least squares-discriminant analysis*); QC (*Quick Classifier*); Q-TOF (*Quadrupole time-of-flight*); RF (*Random Forest*); SNN (*Supervised Neural Network*); sPLS-DA (*Sparse partial least squares-discriminant analysis*); SVM (*support-vector machine*); TOF (*Time-of-flight*); TQ (triplo quadrupolo); UPLC (*Ultra Performance Liquid Chromatography*)

Além do preparo das amostras antes das análises, o uso de fontes de ionização capazes de identificar biomoléculas por MS, como a ionização e dessorção a laser assistida por matriz (do inglês, *matrix assisted laser desorption ionization* (MALDI)), abriu portas para a identificação de biomarcadores em amostras de soro, devido à fácil operação e excelente cobertura de uma ampla gama de moléculas.⁴⁷

Estudos mostraram que existem diferenças no perfil espectral de amostras de soro ou plasma analisadas por MALDI MS entre controles saudáveis e pacientes com certas doenças. O uso da fonte MALDI destaca-se pela excelente cobertura de uma ampla gama de moléculas, principalmente grandes biomoléculas como peptídeos e proteínas.⁴⁷

Uma das doenças muito estudada na busca por biomarcadores é o câncer. Biomarcadores específicos de câncer no soro como os autoanticorpos, que têm como alvo antígenos específicos associados a tumores, ou até mesmo os próprios antígenos específicos, têm sido utilizados como parâmetros de monitoramento clínico de rotina.⁴⁸ No entanto, o diagnóstico baseado em um único biomarcador pode apresentar resultados falsos positivos e baixa especificidade devido a outros eventos imunológicos.⁴⁸

Dessa forma, a busca por padrões proteômicos e peptidômicos séricos capazes de diferenciar indivíduos saudáveis de indivíduos com câncer, destacou o uso da técnica MALDI MS, que ganhou destaque apresentando resultados bastante satisfatórios nesses diagnósticos. Como no estudo de Rungkamoltip e colaboradores (2023),⁴² que a partir do uso da análise de componentes principais (do inglês, *principal component analysis*, PCA) observaram padrões peptídicos de discriminação entre o grupo saudável e todos os diferentes estágios do câncer cervical.⁴² Outro estudo que avaliou a discriminação do perfil espectral sérico, nesse caso perfil proteômico, foi realizado por Park *et al.* (2019)⁴⁹ que relataram uma diferença estatisticamente significativa no perfil de proteínas séricas totais baseadas em MALDI-TOF MS (do inglês, *time-of-flight*, TOF) entre pacientes com câncer de fígado e controles saudáveis pela análise discriminante de mínimos quadrados parciais (do inglês, *partial least squares-discriminant analysis*, PLS-DA).⁴⁹

Outros tipos de câncer, como câncer de pulmão,⁵⁰ bexiga,³¹ e ovário,⁵¹ também já foram estudados, a fim de desenvolver ferramentas diagnósticas a partir da análise por MALDI MS do perfil sérico.

Além do câncer, o uso da técnica MALDI-TOF MS para identificar novos biomarcadores de proteínas séricas relacionados a outras doenças como artrite reumatoide,¹⁹ mostrou-se eficaz. No estudo de Abdelati e colaboradores (2020),¹⁹ onde foram detectados 113 sinais que discriminavam pacientes com artrite reumatoide de pacientes com osteoartrite primária, outra doença que afeta as articulações, 101 sinais que discriminavam artrite reumatoide de pacientes controle e 95 sinais que discriminavam osteoartrite primária de pacientes controle.¹⁹

Assim, diagnósticos avaliando o perfil sérico por MALDI MS em diferentes tipos de doenças, como Alzheimer,³² transtorno de déficit de atenção e hiperatividade (TDAH),⁴⁴ fibrose cística,³⁰ a doença autoimune púrpura de Henoch-Schönlein,⁴⁵ vêm mostrando excelentes resultados nos últimos anos.

Além das proteínas, o plasma e o soro abrigam uma variedade de outras moléculas, que circulam pelo corpo interagindo com tecidos e células, como metabólitos e lipídios. Os lipídios são uma classe de moléculas orgânicas que desempenham várias funções essenciais nos seres vivos. Devido às suas importantes funções fisiológicas, os lipídios estão intimamente envolvidos em distúrbios orgânicos, incluindo síndrome metabólica, inflamação, distúrbios neurológicos e cânceres.¹²

O acoplamento da cromatografia a MS oferece uma excelente solução para análises de misturas complexas, como biofluidos, e é uma das técnicas mais utilizada em metabolômica e lipidômica. O preparo de amostras para análise lipidômica e metabolômica é similar, normalmente são utilizadas formas de precipitação de proteínas e extração de metabólitos ou lipídios.^{13,52}

Taşcı e colaboradores (2020)³⁵ realizaram análises de metabólitos e lipídios não direcionados por LC-ESI-Q-TOF MS (do inglês, *liquid chromatography electrospray ionization quadrupole time-of-flight*) para identificar potenciais biomarcadores em amostras de plasma de insuficiência ovariana prematura. Nesse estudo, 83 picos metabólicos e 213 picos lipídicos foram encontrados semi-quantitativamente e são estatisticamente diferentes entre as amostras de pacientes com insuficiência ovariana prematura e os indivíduos controle. Os dois grupos foram separados com sucesso através da PCA e os compostos, fenil alanina, decanoil-L-carnitina, 1-palmitoil lisofosfatidilcolina e PC (O-16:0/2:0) foram identificados e propostos como biomarcador plasmático para insuficiência ovariana prematura e para monitoramento dos pacientes em período de tratamento.³⁵

Em outro estudo, agora avaliando o perfil metabólico lipídico sérico de amostras de pacientes com doenças autoimunes sistêmicas, entre elas o lúpus eritematoso sistêmico, a esclerose sistêmica ou esclerodermia, a síndrome de Sjögren, a artrite reumatoide, a síndrome antifosfolípide primária, a doença mista do tecido conjuntivo e a doença indiferenciada do tecido conjuntivo. Fernández-Ochoa *et al* (2020),²⁸ relataram que os modelos PLS aplicados no tratamento de dados das análises por HPLC-ESI-QTOF-MS (do inglês, *high-performance liquid chromatography*, HPLC) foram capazes de classificar a maioria das doenças autoimunes sistêmicas em comparação com controles saudáveis, com exceção da síndrome antifosfolípide primária e da doença mista do tecido conjuntivo. Os metabólitos de diferenciação consistiam predominantemente de ácidos graxos insaturados, acilglicinas, acilcarnitinas e aminoácidos.²⁸

O uso da infusão direta no espectrômetro de massas, sem a separação cromatográfica, vem proporcionando análises mais rápidas e com maior reprodutibilidade de amostras biológicas para detecção de lipídios.⁵³ Para suprir a desvantagem de possíveis interferências isobáricas (interferência que pode ser resolvida com a cromatografia) em análises de amostras complexas, o uso de analisadores de massas de alta resolução e precisão, como TOF, ressonância ciclôtrica de íons com transformada de Fourier (do inglês, *fourier-transform ion cyclotron resonance*, FTICR) ou *Orbitrap*, são um facilitador na identificação das moléculas.^{12,54}

Na literatura é relatado que metabólitos lipídicos no sangue podem ser potenciais biomarcadores precoces de algumas doenças, como no estudo de Anand *et al* (2016), que detectou 23 potenciais biomarcadores lipídicos séricos de pré-eclâmpsia em gravidez entre 12 e 14 semanas, utilizando a análise por infusão direta em um espectrômetro de massas com analisador TOF (*Agilent Technologies*) por meio de uma fonte ESI operada no modo positivo.⁵³ Outro estudo de lipídios séricos forneceu informações sobre biomarcadores da doença de Alzheimer, ao encontrar diferenças nos lipídios e alguns metabólitos discriminando grupo de pacientes e de controles saudáveis, utilizando uma fonte de ESI e um analisador Q-TOF em um sistema QSTAR XL Hybrid (*Applied Biosystems*).⁵⁵

Um estudo recente (2020), também utilizou da infusão direta ESI MS para análise do perfil lipídico sérico de amostras de pacientes com diferentes estágios de doença pulmonar obstrutiva crônica e um painel de 50 metabólitos lipídicos,

incluindo fosfolipídios, esfingolipídios, glicerolipídios e ésteres de colesterol, pôde ser usado para diferenciar a presença de dessa doença.²⁹

O estudo do perfil lipídico de forma rápida, reprodutível e detalhada associado a conjuntos de dados, revela que os perfis lipídicos são influenciados por fatores fisiológicos, de estilo de vida e genéticos. Assim, a análise do perfil lipidômico MS é uma ferramenta valiosa para identificar lipídios que podem se tornar novos biomarcadores usados em aplicações clínicas, como em triagens de doenças cardíacas coronárias.⁵⁶

A técnica analítica MS pode auxiliar nos diagnósticos de diferentes doenças, a partir da análise do perfil proteômico, metabolômico e lipidômico sérico de indivíduos controles e pacientes com a doença. E apesar do sangue ser o biofluido mais utilizado e rico em diversas biomoléculas, outros biofluidos também podem ser interessantes matrizes na busca de diagnósticos de doenças.

1.2.2. Saliva

A saliva é uma secreção complexa produzida pelas glândulas salivares (parótida, submandibular e sublingual), juntamente com outras células secretoras situadas na cavidade bucal (lábios, língua, véu palatino, etc).⁵⁷ Aproximadamente 99% da saliva é água e o outro 1% é composto pelas moléculas orgânicas e inorgânicas.⁵⁸

Fisiologicamente, entre as funções da saliva estão a lubrificação de todos os componentes da cavidade oral, facilitando a fonação, a percepção do sabor dos alimentos (paladar), a ação bactericida, conferida pela enzima lisozima, e a digestão do amido através da ação da amilase salivar, além do tamponamento de variações do pH bucal.^{57,59}

Essas funções e propriedades são atribuídas à presença de algumas substâncias como: os eletrólitos, responsáveis pelo equilíbrio iônico na boca; as enzimas, como amilase-salivar e lisozima, que participam do processo de digestão; as proteínas, como as imunoglobulinas e as glicoproteínas; e lipídios.⁶⁰ A saliva também contém um fluido gengival com componentes séricos, bactérias e metabólitos bacterianos, células epiteliais esfoliadas e leucócitos.⁶¹ Além disso, o fluxo e a composição química da saliva podem ser afetados pelo sexo, idade, estado

emocional, hidratação, exercício físico, medicação, doenças sistêmicas, abuso de substâncias e nutrição.^{58,62}

Assim como o soro, a saliva contém diversas substâncias, e muitos desses constituintes entram na saliva a partir do sangue por difusão passiva, transporte ativo ou ultrafiltração extracelular.⁶³ Logo, o perfil molecular da saliva também pode refletir muitas vezes um estado fisiológico ou patológico que cursam com alterações no meio interno (sangue).⁵⁷

A saliva, como um biofluido de diagnóstico, oferece vantagens sobre o soro por ter uma amostragem fácil, não invasiva e indolor.⁶² No entanto, deve-se tomar cuidado ao substituir sangue por saliva, pois as relações entre esses biofluidos diferem dependendo da doença e do biomarcador.⁶⁴ O sangue continua sendo o melhor fluido corporal para avaliação de muitos biomarcadores e sua substituição deve ser usada com cautela.⁶⁴

Até agora, a saliva tem sido utilizada com sucesso para o diagnóstico de algumas doenças, com base na determinação qualitativa e/ou quantitativa de biomarcadores.⁶⁵ As patologias previstas foram: doenças bucais,^{66,67} cardíacas,⁶⁸ diabetes,^{69,70} síndrome de Sjörger,⁷¹ fibrose cística⁷² e HIV.^{73,74}

Algumas metodologias para confirmação dessas doenças utilizam técnicas caras, demoradas e específicas, com a identificação de somente um biomarcador. O uso da MS nas análises dos biofluidos, como a saliva, para diagnósticos de doenças, abrange a identificação de uma grande quantidade de compostos, possíveis biomarcadores, em uma única análise e tem sido cada vez mais explorado por pesquisadores nos últimos anos.⁷⁵

O maior foco dessa busca em saliva é por proteínas e, quando analisadas por MS, podem ser identificadas por métodos de ionização que minimizam a fragmentação da amostra, como ionização por ESI e MALDI.⁷⁶

Até o momento, o uso da técnica MALDI MS para auxiliar em diagnósticos de doenças vem se destacando com estudos que buscam moléculas ou perfis espectrais que possam diferenciar pacientes saudáveis de doentes, de maneira rápida e não invasiva.⁷⁷ Entretanto, outro problema associado ao uso da saliva é a sua composição variável e heterogênea. A concentração de biomarcadores pode ser várias ordens de grandeza menor do que no soro, as variações diurnas na composição podem causar problemas na interpretação dos resultados e sua composição complexa pode causar problemas no preparo das amostras.⁶⁵

Devido à variação na composição da saliva, dependendo da hora do dia, da hidratação, da posição corporal, da ingestão de medicamentos, da alimentação, entre outros, protocolos bem definidos devem ser realizados desde a coleta das amostras de saliva que pode ser por estimulação ou simplesmente em repouso.⁷⁸

O estudo de Zhu *et al* (2020)⁷⁹ investigou as variações do peptidoma salivar por MALDI-TOF MS, em resposta a diferentes condições de estimulação, saliva não estimulada e três tipos de amostras de saliva estimulada (saliva estimulada pelo olfato, pela gustação e pela mastigação). Por comparações do peptidoma salivar, os três tipos de saliva estimulada exibiram peptidoma significativamente diferentes. A saliva estimulada coletada em resposta à estimulação olfatória apresentou menores variações em comparação à estimulação gustativa/mastigatória. Sendo assim, sugere-se que apenas um tipo de método estimulante seja usado ao longo de uma pesquisa de peptidoma.⁷⁹

Embora existam algumas desvantagens em relação ao soro, a saliva tem sido estudada extensivamente como uma ferramenta de diagnóstico em potencial na última década devido à sua facilidade de coleta, juntamente com sua diversidade de biomarcadores, como material genético e proteínas.⁶³ A Tabela 1. 2 mostra alguns dos artigos publicados nos últimos 5 anos utilizando a MS para diagnósticos de doenças, tendo como a amostra saliva de indivíduos controle e pacientes com diferentes doenças.

Tabela 1. 2. Artigos publicados nos últimos 5 anos para diagnósticos de doenças a partir do uso de MS e amostras de saliva.

Amostra	Preparo de amostra	Equipamento	Tratamento de dados	Resultados	Ano
50 amostras de saliva (16 controles, 17 pacientes com periodontite crônica e 17 com gengivite)	Preparo: não utilizar quaisquer práticas de higiene bucal, coleta das 8h às 9h Remoção de materiais insolúveis: centrifugação e adição de um inibidor de proteinase. Purificação: esferas magnética de troca catiônica	MALDI-TOF MS (Bioyong Tech). nanoLC-ESI MS/MS (Thermo Fisher Scientific),	<i>Software:</i> BioExplorer e SPSS Statistics Tratamento de dados: testes t de Student ou o teste de Wilcoxon, Heatmap e curva ROC Base de dados: Mascot	A maioria dos peptídeos expressos diferencialmente exibiram um aumento em participantes com periodontite crônica e gengivite em comparação com controles saudáveis.	2019 ⁸⁰
141 pacientes (74 controles e 67 pacientes com periodontites),	Preparo: Não comer, beber ou escovar os dentes uma hora antes da coleta da amostra.	MALDI-TOF MS (Bruker Daltonics)	<i>Software:</i> BioTyper RTC™ programa R, com biblioteca MALDIQuant. Tratamento de dados: Análise Discriminante Binária	Foram detectados 217 picos na saliva, 114 foram significativamente diferentes entre os 2 grupos. Entre os principais picos de classificação, 8 apresentaram níveis mais elevados no grupo de periodontite, e 2 apresentaram níveis mais elevados no grupo controle.	2020 ⁸¹
45 amostras de saliva (15 controles, 15 pacientes com periodontite grave e 15 após tratamento periodontal ativo)	Precipitação de proteínas: Adição de acetona glacial: metanol: fosfato de tributil e centrifugação. Digestão de proteínas com tripsina.	MALDI-TOF MS (Bruker Daltonics)	<i>Software:</i> SPSS Tratamento de dados: ANOVA, teste de Friedman, teste de Tukey, teste de Dunn. Banco de dados: NCBI e Mascot	Um número maior de proteínas foi identificado no proteoma da saliva total de indivíduos saudáveis e tratada em comparação com o grupo com periodontite grave. A cistatina SN (CST1), correlacionou-se negativamente com a porcentagem de sítios patológicos e foi	2023 ⁸²

Amostra	Preparo de amostra	Equipamento	Tratamento de dados	Resultados	Ano
				eficaz na discriminação da periodontite ativa do estado periodontal saudável (seja saudável ou tratada).	
35 amostras de saliva (17 controles e 18 de pacientes com disfunções da articulação temporomandibular (DTM))	Preparo: saliva total não estimulada. Não realizar qualquer higiene bucal, comer ou beber por pelo menos uma hora antes da coleta. Centrifugação e filtração.	MALDI-TOF/TOF MS (Bruker Daltonic)	<i>Software:</i> ClinProTools Tratamento de dados: teste <i>t</i> de Student ou de um teste de Wilcoxon curva ROC, QC, GA e SNN. Banco de dados: Bruker MALDI Biotyper System	Dos 23 picos que exibiram o maior poder discriminatório, foi selecionado um painel de marcadores salivares que previram os pacientes com DTM. Houve a diferenciação entre amostras salivares de indivíduos com DTM e indivíduos saudáveis, confirmando a utilidade do perfil proteômico no monitoramento da doença.	2020 ⁸³
65 amostras de saliva (35 controles e 30 pacientes com doença renal crônica (DRC))	Preparo: Enxague da boca com água limpa após o café da manhã, e coleta após 10 minutos. Centrifugação e adição do inibidor de protease. Esferas magnéticas foram usadas para fracionar amostras de saliva.	MALDI-TOF MS nano-UPLC/ESI MS/MS (Waters) (Thermo Fisher Scientific)	<i>Software:</i> Bio Explorer (Bioyong Tech), Tratamento de dados: KNN, curva ROC Banco de dados: Mascot	Espera-se que Histatina-1, Mucina-7 e aPRP 1/2 sejam candidatos a biomarcadores salivares de pacientes com DRC em HD. Assim, a saliva pode ser um biofluido diagnóstico não invasivo promissor.	2019 ⁸⁴
53 amostras de saliva (30 controles e 23 pacientes com deficiências de anticorpos (DAP))	Preparo: Sem comer nem beber pelo menos 2 horas antes da coleta, realizada entre 10h e 12h. Adição de TFA 0.2% e	HPLC-ESI MS/MS, (ThermoFisher Scientific)	<i>Software:</i> MS-Product, GraphPad Prism Tratamento de dados: teste de Mann-Whitney e teste <i>t</i> , ANOVA.	Os pacientes exibiram menor abundância de α -defensinas 1-4, cistatinas S1 e S2, e maior abundância de cistatina B glutationilada e cistatina SN do que os controles. Os pacientes poderiam ser	2020 ⁸⁵

Amostra	Preparo de amostra	Equipamento	Tratamento de dados	Resultados	Ano
	centrifugação.		Banco de dados humanos UniProt-KB	agrupados em dois grupos com base em diferentes níveis de cistatina SN, S1 e S2, sugerindo que estas proteínas podem desempenhar papéis diferentes na doença.	
40 amostras de saliva (20 controles e 20 pacientes com diagnóstico de diabetes mellitus tipo 2)	Remoção de resíduos insolúveis. Centrifugação. Precipitação de proteínas: Mistura com ACN:H ₂ O, agitação e centrifugação.	novo chip NIMS, FEP@SiNWs MALDI-TOF/TOF e UPLC-MS/MS (Bruker Daltonics)	<i>Software:</i> FlexAnalysis, ClinproTool e MATLAB Tratamento de dados: PCA e OPLS-DA Banco de Dados: HMDB	Com o auxílio da análise multivariada, foram encontrados e identificados 22 candidatos metabólicos que podem discriminar diabetes mellitus tipo 2 e voluntários saudáveis.	2021 ⁸⁶
Amostras de saliva divididas em três grupos: 1) Forma leve, 2) Moderada, 3) Grave para COVID-19 .	Preparo: Uso do dispositivo de algodão – Salivette™. Não beber, comer ou escovar os dentes 2 horas antes da coleta da amostra. Coletadas das 7h às 11h.	MALDI-TOF Autoflex (Bruker Daltonics)	<i>Software:</i> R com pacotes MALDIquantForeign e MALDIquant Tratamento de dados: GBM, SVM, NNET, NB e RF.	Classificação entre condições leves/moderadas e graves. apresentou resultados de precisão, sensibilidade e especificidade de 67,18%, 52,17% e 75,60%, respectivamente.	2021 ⁸⁷
156 amostras de saliva (51 controles e 105 pacientes positivos para COVID-19 (coletadas no dia da inclusão do paciente	Preparo: Realização de bochechos antes da coleta de saliva com Neutral Salivette® e centrifugação.	MALDI-TOF MS (Bruker Daltonics)	<i>Software:</i> Flex Control e R com pacotes MALDIquant e MALDIrppa Tratamento de dados: PCA, SVM, RF e KNN e	9 picos distinguiram significativamente os dois grupos. A aplicação do modelo SVM em conjuntos de dados independentes confirmou seu desempenho com 88,9% e 80,8% de classificação correta para amostras coletadas em D0 e D30,	2022 ⁸⁸

Amostra	Preparo de amostra	Equipamento	Tratamento de dados	Resultados	Ano
(D0) e dez (D10) e trinta (D30) dias depois))			LDA.	respectivamente. Por outro lado, no D10, a proporção de classificação correta caiu para 64,3%.	
149 amostras de saliva (97 com RT-PCR positivo para SARS-CoV-2 e 52 negativas)	Preparo: Digestão com tripsina e purificação com spin C18.	MALDI MS FT-ICR MS (Bruker Daltonics)	<i>Software: DataAnalysis e</i> MATLAB 9.0 Tratamento de dados: PCA e SVM	Na otimização do método, o melhor preparo da amostra foi obtido com a digestão da saliva em 10 µL de tripsina por 2 h. Os modelos SVM1 e SVM2 apresentaram os melhores resultados. O SVM1 selecionou 780 variáveis e possui taxa de falso negativo (FNR) de 0%, enquanto o SVM2 selecionou apenas duas variáveis com FNR de 3%.	2022 ⁸⁹

curva ROC (*Receiver Operating Characteristic*); ESI (*Electrospray ionization*); GA (*Genetic Algorithm*); GBM (*Gradient Boost Machine*); HPLC (*High-performance liquid chromatography*); KNNk (*Nearest Neighbors*); LC (*Liquid chromatography*); LDA (*Linear Discriminant Analysis*); MALDI (*Matrix assisted laser desorption ionization*); MS (*Mass spectrometry*); NB (*Naïve Bayes*); NNET (*Neural Networks*); OPLS-DA - *Orthogonal Partial Least Squares Discriminant Analysis*; PCA (*Principal component analysis*); OPLS-DA (*orthogonal Partial least squares-discriminant analysis*); QC (*Quick Classifier*); RF (*Random Forest*); SNN (*Supervised Neural Network*); SVM (*support-vector machine*); TOF (*Time-of-flight*); UPLC (*Ultra Performance Liquid Chromatography*)

Devido à sua conexão anatômica direta com a cavidade oral, o aparelho digestório e o sistema endócrino, o uso clínico de testes de saliva para doenças relacionadas a esses campos, vem ganhando destaque.⁷⁸ Como em estudos analisando amostras de pacientes com periodontites, uma condição inflamatória crônica que afeta os tecidos de suporte dos dentes, conhecidos como tecidos periodontais, por MS em amostras de saliva, que já revelaram resultados satisfatórios na busca por alterações na saliva correlacionadas a doença periodontite.⁹⁰

Antezack e colaboradores (2020)⁸¹ desenvolveram, pela primeira vez em 2020, testes diagnósticos baseados em perfis proteicos de saliva, entre pacientes com periodontite e indivíduos saudáveis, detectando 114 sinais significativamente diferentes entre os dois grupos com sensibilidade igual a 70,3% e especificidade igual a 77,8% com análises utilizando MALDI-TOF MS.⁸¹

Outra doença em destaque no uso de amostras de saliva e técnicas utilizando MS nesses últimos anos foi a COVID-19. Por ser uma doença causada por um vírus respiratório a busca por uma diferença espectral salivar é de grande valia.⁸⁷⁻⁸⁹

Entretanto, anteriormente, diferentes estudos também já detectaram proteínas salivares por MS em uma variedade de doenças, sem necessariamente terem correlações com a saliva, como câncer de ovário,⁹¹ doença autoimune sistêmica, como a síndrome de Sjögren⁹² e diabetes⁹³.

O desenvolvimento de métodos para a identificação de biomarcadores por MS em amostras biológicas nos últimos tempos tem desempenhado papéis cada vez mais importantes na avaliação de risco de doenças, triagem, prognóstico e monitoramento de propagação, melhorando assim a saúde humana.⁴

1.3 ESPECTROMETRIA DE MASSAS

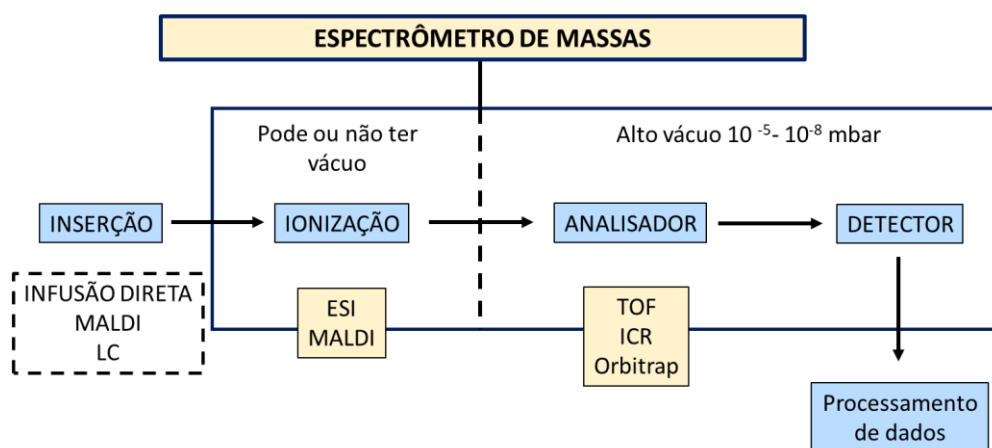
Avanços nas descobertas de biomarcadores de doenças baseados em MS estão continuamente expandindo o panorama do diagnóstico clínico.⁴

A MS é uma técnica que se destaca na química analítica, por fornecer informações qualitativas para uma ampla classe de compostos. Nessa técnica, amplamente utilizada para a identificação de diversos compostos, as moléculas são convertidas em íons em fase gasosa, subsequentemente, separados no espectrômetro de massas de acordo com sua razão massa (m) sobre carga (z), m/z .

Além disso, a partir dos experimentos de fragmentação de íons, é possível inferir a respeito da estrutura molecular do analito.⁹⁴

O funcionamento dos espectrômetros de massas constitui-se basicamente de cinco etapas: i) sistema de introdução de amostra podendo exigir ou não o preparo de amostra; ii) fonte de ionização (ionização das moléculas de interesse, analito); iii) analisador de massas (separação dos íons formados de acordo com a razão m/z); iv) detector, que realiza a “contagem” dos íons e transforma o sinal em corrente elétrica; v) processador que, posteriormente, converte a magnitude do sinal elétrico em razão m/z , ocorrendo a aquisição de um espectro de massas correspondente.⁹⁵ O processo descrito é ilustrado na Figura1. 4.

Figura1. 4. Diagrama esquemático de um espectrômetro de massas.



Fonte: Adaptado de Romão.⁹⁶

Destacam-se como componentes mais importantes no espectrômetro de massas a fonte de ionização e o analisador. Cada método de ionização produz maior abundância de tipos específicos de compostos, as diferentes formas de ionização e os analisadores de massas determinam a aplicabilidade da MS.⁹⁵

No final da década de 1980, a introdução de novos métodos de dessorção-ionização, como ESI e MALDI, por John Fenn e Koichi Tanaka, revolucionou a MS de matrizes biológicas.⁹⁷ A descrição do genoma, a exploração do proteoma, bem como a identificação do metaboloma, metabólitos biossintetizados por um organismo, favoreceram a busca por possíveis biomarcadores de doenças como proteínas e metabólitos.⁹⁷ A ionização branda intrínseca dessas duas técnicas

permite a análise de moléculas intactas, proporcionando a geração de informações sobre as massas moleculares intactas com maior precisão.⁹⁸

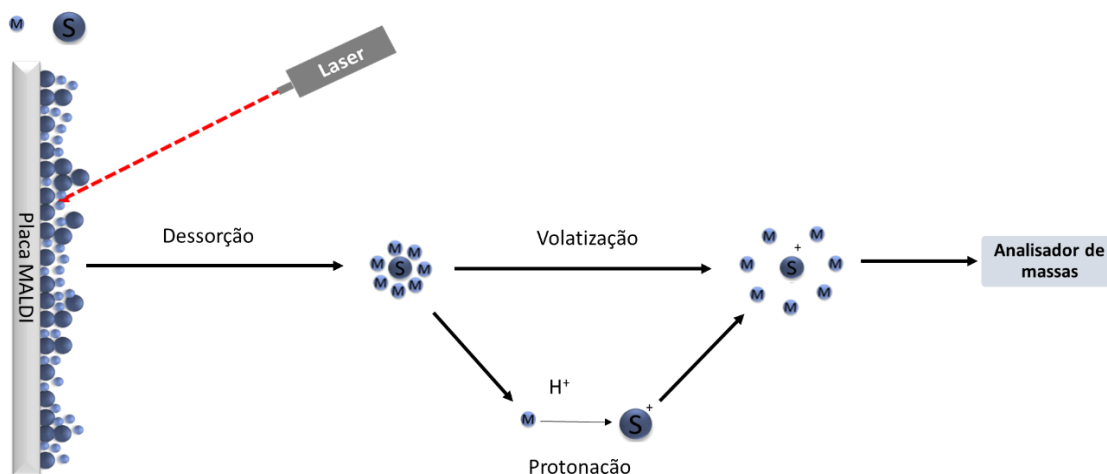
1.3.1 MALDI MS

Alguns tipos de ionização são altamente energéticos e podem causar a fragmentação das ligações químicas, já em outros métodos a energia utilizada é suficiente apenas para produzir os íons e não causar a fragmentação, como as fontes ESI e MALDI, que são técnicas de ionização suave.^{99,100}

No procedimento de ionização utilizando a técnica MALDI MS, a amostra é misturada por uma matriz e esse conjunto aplicado a uma placa metálica. A matriz consiste em substâncias, geralmente, compostos orgânicos aromáticos de baixa massa molecular, que absorvem energia no comprimento de onda do laser incidente, para a extração de moléculas de interesse e auxílio na dessorção e ionização destas.¹⁰¹

Cristais da matriz junto com o analito são formados durante a secagem sobre a placa. Na análise desse conjunto, a amostra e matriz depositada na placa metálica, é irradiado um laser pulsado (de 337 ou 355 nm), a matriz absorve a energia emitida e transfere-a para os compostos da amostra, ao mesmo tempo desencadeia-se um processo de dessorção, o que viabiliza a passagem da amostra do estado sólido para o gasoso, induzindo a dessorção e ionização desses cristais ricos em analitos. Após a ionização, os íons de analitos, em fase gasosa, são direcionados para o analisador de massas, que os organiza de acordo com o valor de m/z .^{102,103} Um esquema da ionização na fonte MALDI é mostrado na Figura 1. 5.

Figura1. 5. Esquema do processo MALDI.



Fonte: Adaptado de Hoffmann e Vincent.⁹⁹

Uma das vantagens da técnica MALDI está na facilidade do preparo das amostras, ampla faixa de m/z e na habilidade de produzir íons com pouca ou nenhuma fragmentação, obtendo a massa exata de uma molécula durante a análise de MS.¹⁰⁴

A eficiência de dessorção e ionização de diferentes tipos de moléculas depende da escolha do composto usado como matriz. O *α -cyano-4-hydroxycinnamic acid* (CHCA), por exemplo, é uma matriz muito utilizada para identificação de peptídeos, o *Sinapinic acid* (SA) para proteínas intactas e o *2,5-dihydroxybenzoic acid* (DHB) para glicanos.^{47,105}

Nos últimos anos, além das aplicações da tecnologia MALDI MS para análises qualitativas e quantitativas de proteínas, essa tecnologia também tem sido amplamente utilizada em pesquisas lipidômicas e metabolômicas em amostras biológicas.¹⁰⁶ Devido à sua ampla faixa de massa e capacidade de detectar rapidamente grandes biomoléculas em quantidades vestigiais, MALDI MS é uma ferramenta ideal para caracterizar biomarcadores de doenças em pesquisa de diagnóstico clínico.¹⁰⁷

Alguns diagnósticos já foram realizados na identificação de infecções bacterianas e fúngicas utilizando a técnica MALDI-TOF MS, podendo superar os métodos convencionais de diagnóstico em velocidade e precisão.¹⁰⁸ Além dos domínios do mundo microbiano, estudos demonstraram que a técnica MALDI MS também é promissora no diagnóstico rápido e precoce de câncer pulmonar,¹⁰⁹ de ovário¹¹⁰ e de trato aerodigestivo.¹¹¹ A técnica também se mostrou eficiente na

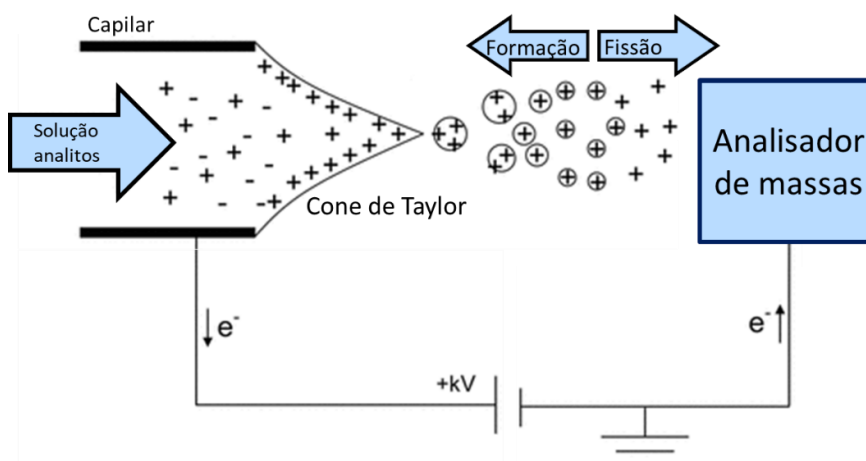
genotipagem de alto rendimento de vírus do papiloma humano oncogênico¹¹² e na identificação genérica e rápida de vírus respiratórios cultivados, além de diversas outras aplicações já relatadas no texto anteriormente.¹¹³

1.3.2 ESI MS

Ambas as fontes de ionização, ESI e MALDI, podem gerar íons $[M + zH]^{z+}$. A fonte MALDI normalmente forma íons monocarregados, ($z = 1$), enquanto a fonte ESI tem a facilidade de formar compostos com múltipla carga ($z \gg 1$). A facilidade da formação de íons multicarregados permite a detecção de moléculas com maior carga em uma faixa de m/z limitada, além de facilitar a dissociação nos experimentos de MS em *tandem*.¹¹⁴

Em uma fonte de formação de íons por *Eletrospray*, a solução contendo o(s) analito(s) de interesse passa por um capilar de metal, que é mantido sob um potencial elétrico (± 2 - ± 6 kV). O campo elétrico na ponta do capilar induz um acúmulo de cargas na superfície do líquido ao final do capilar, onde gotas altamente carregadas são formadas. Um cone de Taylor (estrutura líquida cônica) é criado pelo campo em um aerossol de gotículas altamente carregadas na solução na ponta do capilar. Este processo de pulverização é geralmente assistido por um fluxo de gás coaxial que evapora o solvente das gotículas carregadas. A alta densidade da gotícula carregada e as forças repulsivas entre os íons fazem com que a gota se desintegre em gotas menores, produzindo assim, íons carregados.^{98,100,114} Esse processo está ilustrado na Figura 1.6.

Figura 1. 6. Processo de ionização da fonte Eletrospray.



Fonte: Adaptado de Hoffmann e Vincent.⁹⁹

Uma das vantagens do uso dessa fonte é o acoplamento direto com a cromatografia líquida (do inglês, *liquid chromatography* (LC)) para a separação de misturas complexas antes da MS. A passagem da amostra por uma fase estacionária auxilia na separação e identificação dos analitos em uma amostra complexa. Com essas vantagens o LC-MS tornou-se o método padrão para análises biomédicas.¹⁰⁰

A análise por ESI MS e experimentos MS/MS tem sido usado no diagnóstico de doenças, ganhando grande importância nos últimos anos em metabolômica e particularmente em lipidômica, nas aplicações clínicas.¹¹⁵ O estudo da lipidômica abre portas para interpretar as alterações no metabolismo lipídico ou na regulação das vias metabólicas lipídicas ligadas a doenças numa perspectiva fisiológica e/ou patológica.¹¹⁶ Por esta razão, as investigações lipidômicas geralmente se concentram na medição de alterações de lipídios em níveis de sistemas indicativos de doença, perturbações ambientais ou resposta a dieta, drogas e toxinas.

Os perfis lipídicos em investigações clínicas relacionados a pessoas que estão em um estado de doença tornam-se a base para a detecção de potenciais biomarcadores relacionados à doença, quando comparados aos controles.¹¹⁶ Estudos apontaram a possibilidade de diagnosticar diversas doenças pelo perfil lipídico utilizando a MS, como câncer,¹¹⁷ doenças neurodegenerativas, como Huntington, Parkinson, Alzheimer,¹¹⁸ esquizofrenia e transtorno bipolar,¹¹⁹ doenças cardiovasculares, diabetes, obesidade e outras doenças crônicas como artrite reumatoide.^{119,120}

A infusão direta de amostras no espectrômetro proporciona uma análise mais rápida e com maior reprodutibilidade, podendo ser uma ferramenta de triagem de alto rendimento, centenas de amostras por dia com um tempo de análise de menos de 1 minuto geralmente e o espectro de massa pode ser usado para a classificação da amostra.⁵⁴

Por apresentar interferências isobáricas, a análise de amostras complexas requer o uso de instrumentos de alta resolução e precisão, ademais a elucidação estrutural pode ser feita por meio de experimentos MS/MS.^{12,121}

Estudos já demonstraram o uso da infusão direta ESI MS na caracterização lipidômica da doença de Alzheimer.¹²² González-Domínguez e colaboradores

(2014),⁵⁵ identificaram alguns compostos como possíveis marcadores da doença de Alzheimer em amostras de soro utilizando ESI MS. As principais diferenças foram encontradas nos lipídios, embora alguns metabólitos de baixa massa molar também tenham mostrado alterações significativas.⁵⁵

Além disso, Haijes e colaboradores (2019)¹²³ demonstraram que o estudo da metabolômica em amostras de plasma e manchas de sangue seco por infusão direta de alta resolução espectrometria de massas é um método muito consistente, de alto rendimento e não seletivo para investigar o metaboloma em doenças genéticas.¹²³

O conhecimento atual dos métodos de extração de lipídios, abordagens de MS, análise de dados bioestatísticos, incluindo fluxos de trabalho para a interpretação de dados lipidômicos de alto rendimento, vem se expandindo e, assim, novos mediadores lipídicos ou biomarcadores de doenças podem ser analisados por técnicas de MS apoiadas por bioinformática e bioestatística sofisticadas.¹²

1.3.3 Analisadores de massas

Após a ionização da amostra, os íons gerados são direcionados para um analisador de massas. Os íons então são separados de acordo com a razão existente entre suas massas e cargas, ou seja, razão m/z .¹⁰¹

Diferentes analisadores de massa separam os íons por uma variedade de metodologias, a escolha do mais apropriado deve ser efetuada considerando a aplicação, desempenho desejado e custo.^{100,101}

Os principais analisadores de massas são: setor magnético, quadrupolo, aprisionamento de íons (do inglês, *ion trap* (IT)), TOF, *orbitrap* e FT-ICR. A principal diferença entre os analisadores são os princípios físicos utilizados em cada um para discriminar as razões m/z , o que leva a espectros de massas com magnitudes de resolução e exatidão distintos.^{95,100} A Tabela 1.3 apresenta uma comparação entre alguns dos analisadores utilizados nos MS.⁹⁹

Tabela 1.3. Comparação entre os analisadores de massas.

Analisador	Limite de massas (Da)	Poder de Resolução (FWHM)	de Exatidão (ppm)	Princípio de Separação
Tempo de voo linear	1.000.000	5.000	200	Velocidade

Tempo de voo reflecton	10.000	30.000	10	Velocidade
Orbitrap	6.000	1000.000	< 1	Frequência ressonância
FT-ICR	4.000	10.000.000	< 1	Frequência ressonância

Fonte: Adaptado de Lanças.¹⁰¹

Os parâmetros de desempenho mais importantes são: sensibilidade, resolução de massa e faixa de massa. O desenvolvimento dos analisadores ICR e *orbitrap*, que geram dados de forma diferente dos demais, requerendo o uso da transformada de Fourier para transformar os dados obtidos pelo equipamento em espectros de massas, apresentam magnitudes maiores em relação ao poder de resolução e exatidão de massa, possibilitando a atribuição de fórmulas moleculares a partir das medidas de m/z .¹⁰¹

O poder de resolução de um analisador de massas pode ser definido como a habilidade em produzir sinais distintos no espectro de massas quando são analisados íons que têm uma pequena diferença na razão m/z . O poder de resolução é dado pela equação 1:

$$R = \frac{m}{\Delta m} \quad (\text{equação 1})$$

Em que m representa a massa e Δm a largura do sinal a meia altura (do inglês, *full width at Half maximum*, (FWHM)).⁹⁹ Um analisador com um alto poder de resolução (TOF, FT-ICR MS ou Orbitrap) é capaz de medir a massa exata de cada isótopo mais abundante, assim há a possibilidade da atribuição de fórmulas moleculares para cada sinal resolvido, de acordo com o seu defeito de massa (diferença entre a massa exata e a nominal).¹²⁴

Para a obtenção de boa acurácia na determinação das massas, geralmente é necessário um espectrômetro de massas de alta resolução, pois interferentes da amostra podem comprometer a acurácia das medidas efetuadas. A exatidão em massa indica o quão próximo o valor experimental está do valor verdadeiro e é geralmente expressada em ppm, pela equação 2.⁹⁹

$$\text{Erro(ppm)} = \frac{m_{\text{medida}} - m_{\text{teorico}}}{m_{\text{teorico}}} \times 10^6 \quad (\text{equação 2})$$

Ao se trabalhar com amostras complexas como fluidos biológicos, o uso de técnicas com alto poder de resolução e exatidão de massas torna-se um fator importante na determinação da identidade de compostos com maior segurança e obtenção de resultados mais confiáveis em estudos clínicos.

1.3.4 Processamento de dados e análise estatística

Nos experimentos utilizando MS de alta resolução é possível obter conjuntos de dados com muitas informações, que compreendem vários espectros de massas representando intensidades medidas em uma grande quantidade de sinais de m/z , descrevendo até centenas de moléculas diferentes. Quase sempre isso é uma grande vantagem, porém pode se tornar uma limitação, quando um grande conjunto de dados precisa ser avaliado sinal a sinal.⁶⁴

A grande quantidade de informações geradas pela MS de alta resolução permite a diferenciação de compostos presentes em uma faixa de massa estreita, possibilitando a identificação de uma maior quantidade de possíveis biomarcadores em pesquisas clínicas. Contudo, essa grande quantidade de dados obtidos em espectros de MS requer a combinação da saída desta técnica com estratégias estatísticas robustas para identificar padrões de diagnóstico.⁸⁹

A combinação de métodos estatísticos pode fornecer uma avaliação de conjuntos de dados complexos de forma mais rápida, permitindo acesso a um grande número de espectros individuais e reduzindo a complexidade no tratamento de dados. Entretanto, é possível ocorrer variações sistemáticas nas intensidades dos sinais, intrínsecas à aquisição de dados e não necessariamente devido a uma resposta biológica. Estas alterações oriundas dos efeitos de supressões iônicas e de interferências analíticas podem ocorrer mesmo após a otimização do método. Dessa forma, vários pré-tratamentos de dados geralmente são necessários antes que a análise estatística possa ser realizada para obter informações significativas sobre amostras de pacientes controle (saudáveis) e doentes.¹²⁵

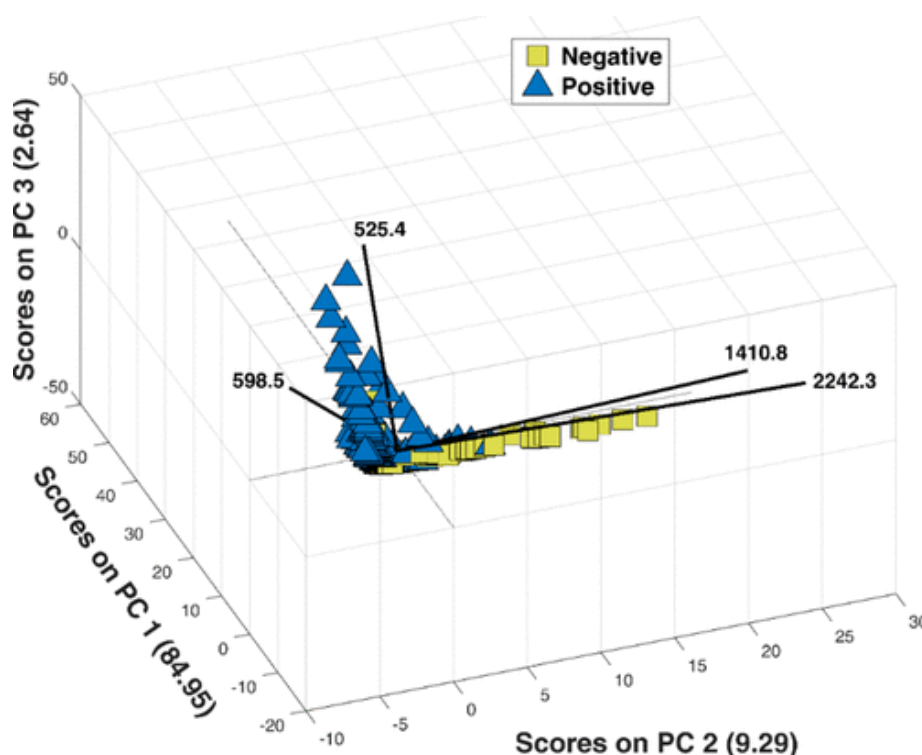
A fim de identificar as diferenças químicas e biológicas relevantes entre o conjunto de dados obtidos, pode ser empregada a quimiometria. Diversos métodos de reconhecimento de padrões são capazes de identificar as variâncias entre os sinais e informações que discriminem as amostras no conjunto de dados.¹²⁶

Métodos quimiométricos multivariados estão sendo amplamente utilizados para extrair informações da grande quantidade de dados gerados em análises de amostras biológicas. Essas abordagens estatísticas são amplamente classificadas em duas categorias: métodos 'não supervisionados' e métodos 'supervisionados'.¹²⁷

Uma técnica não supervisionada, a PCA, é uma ferramenta quimiométrica que, a partir de um determinado conjunto de dados, é capaz de condensar as informações mais relevantes em um número reduzido de novas variáveis.¹²⁸ O método não requer nenhuma entrada relacionada a informações de classe e é usado para projetar os dados para destacar as fontes máximas de variação usando um número mínimo de componentes principais (do inglês *Principal Components* (PC)). As PCs são combinações lineares das variáveis originais, entretanto, são não correlacionadas. Sendo assim, este modelo usa a variabilidade e correlação de cada variável (m/z) para criar um conjunto simplificado de novas variáveis (PCs) e obter o reconhecimento de padrões.¹⁰⁰

O uso do método quimiométrico PCA pode classificar os espectros sem o conhecimento da classe de espécimes biológicos (como doença ou controle) usando os valores de m/z e suas intensidades. Como demonstrado no estudo de Almeida e colaboradores,⁸⁹ em que analisaram amostras de saliva de pacientes positivos e negativos para SARS-CoV-2 por MALDI-FT-ICR e com o modelo de PCA construído, foi observado uma tendência de separação entre os dois grupos principais, com influência de m/z 525,4 e 598,5 nas amostras positivas e 1410,4 e 2242,3 nas amostras negativas, Figura 1. 7.⁸⁹

Figura1. 7. Gráfico de scores e loads do modelo PCA de espectros MALDI(+) FT-ICR de amostras de saliva positivas ou negativas para SARS-CoV-2.



Fonte: De Almeida *et al.*⁸⁹

Nos métodos supervisionados existe a necessidade de um conjunto de dados de treinamento com diagnóstico conhecido (como doente ou saudável), usado para construir um modelo preditivo. Após o treinamento, o modelo pode ser utilizado em um conjunto de teste para classificar amostras desconhecidas e medir a precisão preditiva do modelo. Métodos supervisionados são muito úteis para detectar diferenças sutis entre amostras semelhantes, a fim de identificar potenciais candidatos a biomarcadores.¹²⁷ Um modelo supervisionado muito utilizado é o PLS-DA.

Entretanto, matrizes complexas, como as biológicas, ou espectros mais discretos, como a MS, podem não ser modelados corretamente utilizando métodos lineares como, PLS-DA.¹²⁹ Como alternativa, é possível aplicar métodos de aprendizagem de máquina (do inglês, *machine learning*, ML) para tratamento de dados que possuem baixa correlação e/ou respostas mais complexas.

Para extrair a informação latente dos dados de alta dimensão, vários métodos de seleção de características têm sido aplicados. A máquina de vetores de suporte (do inglês, *Support Vector Machine*, SVM), por exemplo, é um método de

aprendizado de máquinas que tem demonstrado uma excelente capacidade de analisar os dados de sistemas biológicos complexos.¹³⁰

A tendência atual de associar algoritmos rotulados como inteligência artificial e técnicas “ômicas” rendeu plataformas envolvendo ML para análise de dados de MS, com o objetivo de identificar biomarcadores de doenças, incluindo doenças que viraram pandemias como a COVID-19.¹³¹

Estudos aplicaram a análise multivariada de dados, ML, em dados de análises utilizando a técnica MALDI MS para o estudo de amostras biológicas de pacientes com suspeita de contaminação pelo SARS-CoV-2.¹³²

1.4 COVID-19

A COVID-19 é uma doença causada por um coronavírus denominado de SARS-CoV-2 com transmissão por aerossóis decorrentes de tosse, espirro ou contato com secreções em suspensão no ar, com alto risco contaminação.^{8,133} A Organização Mundial de Saúde declarou, em 11 de março de 2020, a doença como pandemia e até novembro de 2023, mais de 772 milhões de casos confirmados e mais de seis milhões de mortes já tinham sido relatadas.¹³⁴ Os casos são classificados como leve, moderado, grave e crítico com base no nível de gravidade do dano celular nos pulmões, variando de assintomático a risco de vida.¹³⁵

Apesar das novas variantes apresentarem menor patogenicidade e dos avanços na vacinação pelo mundo, epidemiologistas especialistas acreditam que o SARS-CoV-2 continuará a se espalhar globalmente por muitos anos.¹³⁶ Assim, a pandemia de COVID-19 continua sendo uma preocupação global e estudos que investigam mecanismos de ação do SARS-CoV-2, os efeitos da infecção no metabolismo do hospedeiro, e principalmente a procura de biomarcadores para diagnósticos e prognósticos da infecção, continuam a ser extremamente necessários.¹³⁷

A definição de métodos de detecção precoce e rápida da COVID-19 foi e continua sendo crucial para interromper as vias de transmissão na comunidade, além de fornecer tratamento oportuno e rastreamento de indivíduos que mantiveram contatos com os pacientes contaminados.¹³⁸

A análise do genoma do SARS-CoV-2 mostrou que este novo vírus compartilha cerca de 80% da identidade de sequência com o SARS-CoV original.^{111,133,139} Os

métodos para identificação da doença podem servir tanto para detecção de SARS-CoV-2, quanto para diagnóstico de COVID-19. Os principais métodos podem detectar, em amostras de soro, plasma, sangue, swabs nasais, nasofaríngeos e orofaríngeos, o ácido nucleico, RNA, ou o antígeno do vírus SARS-CoV-2 ou anticorpos.¹⁴⁰

Uma reação em cadeia da polimerase com transcrição reversa em tempo real (do inglês, *reverse transcription polymerase chain reaction*, RT-PCR) é a técnica padrão-ouro para identificar os ácidos nucleicos virais e, portanto, encontra suas aplicações no diagnóstico de infecções virais por SARS-CoV-2. O RT-PCR oferece alta especificidade e sensibilidade, porém, amostras de swab contendo baixa carga viral e a qualidade do RNA viral extraído podem afetar significativamente a sensibilidade e a precisão da detecção. Além disso, a alta viscosidade e a alta concentração de proteínas e outros componentes interferentes representam um desafio adicional para a extração de RNA. Esses são alguns dos fatores responsáveis pela baixa sensibilidade e pelos resultados falso-negativos da detecção de COVID-19 baseada em RT-PCR.^{3,141}

Além disso, este ensaio é laborioso, de alto custo e mais lento, pois precisa respeitar os tempos de polimerização de acordo com os protocolos padronizados do método.^{133,142} Assim, são necessárias abordagens de diagnóstico rápidas e implantáveis para enfrentar os aspectos desafiadores dos surtos virais, em que existe alta procura por testes e diagnósticos.

1.4.1 COVID e MS

Os métodos disponíveis para detecção de doenças como COVID-19 empregam informações do genoma ou a presença de proteínas oriundas da resposta imune. A detecção do RNA viral em amostra de secreção nasal ou da garganta, coletada com swab pelo método RT-PCR, o padrão ouro, é muito precisa se for realizada no período adequado. A coleta ideal do material biológico deve ser feita na fase aguda da doença, do primeiro ao oitavo dia do aparecimento dos sintomas. A desvantagem do uso desse método está na baixa sensibilidade em casos assintomáticos, ou fora do período, fazendo com que um único teste PCR negativo seja inconclusivo. Além disso, o resultado do teste pode demorar até 48h.¹⁴³

Os testes rápidos não detectam o coronavírus, e sim os anticorpos, podendo ser detectados tanto os anticorpos IgM (imunoglobulina M), que indicam contato recente, infecção aguda, presente no momento da realização do exame, ou os anticorpos IgG (imunoglobulina G), indicadores de contato prévio já com desenvolvimento de resposta imune. Estes testes são fáceis de fazer e oferecem resultado em cerca de 15 minutos a partir da coleta de apenas uma gota de sangue. Eles devem ser realizados após sete dias do aparecimento dos sintomas, pois coletas anteriores apresentam sensibilidade de apenas 30%, assim, há grande risco de falsos negativos.¹⁴³

Ao analisar o sangue de pacientes com suspeita de COVID-19, o teste mais conhecido é o ELISA que busca anticorpos no plasma, bem mais preciso que o teste rápido, porém também há a necessidade de aguardar sete dias após o surgimento dos sintomas e os resultados demoram cerca de 24h, por ser necessário o uso de laboratórios especializados para a realização.¹⁴³

Sendo assim, uma alternativa a estes métodos, com alta sensibilidade e especificidade, seria a aplicação de ferramentas de diagnóstico baseadas em MS, identificando biomarcadores específicos para SARS-CoV-2, mesmo na completa ausência de informações do genoma.¹³⁸

A MS é uma técnica analítica poderosa que permite identificar e quantificar compostos em amostras complexas, incluindo amostras de COVID-19. Desde o início da pandemia, houve um grande interesse em aplicar a MS em análises de amostras de COVID-19, incluindo amostras de saliva, sangue, urina, entre outras.¹⁰ A diversidade de fontes de ionização e técnicas que podem ser hifenadas à MS permite análises multiômicas que auxiliam no rastreamento e diagnósticos de pacientes doentes e saudáveis.¹⁰ A capacidade de fornecer dados qualitativos e informações quantitativas de vários analitos (por exemplo, ácidos nucleicos, proteínas, peptídeos e lipídios) é uma das vantagens da MS sobre os métodos padrão ouro, como RT-PCR e ELISA.¹⁴¹

A maioria das pesquisas utilizando a MS e análises de amostras de COVID-19 está relacionado ao estudo de proteínas. As abordagens proteômicas são adequadas para estudar a COVID-19, pois podem fornecer quantificação e expressão diferencial de proteínas e suas modificações pós-traducionais, representando condutores de mecanismos celulares patológicos e status de doença individual.¹⁴⁴

Como já relatado, uma das principais técnicas MS para análises proteômicas é MALDI MS por sua tolerância relativamente alta de misturas e contaminantes biológicos. Além de conseguir detectar macromoléculas (de alto m/z), ela exige o mínimo de reagentes, tem alto rendimento e rápida aquisição de dados, e para amostras de COVID-19 não seria diferente.¹⁴⁵

Em 2003, quando surgiu a primeira síndrome respiratória aguda grave (SARS) produzida por coronavírus, Krokhin *et al* (2003)¹³⁸ identificaram uma proteína por MALDI-TOF MS, analisando soros provenientes de vários pacientes com SARS, e verificaram que essa correspondia a uma nova proteína do nucleocapsídeo de SARS, prevista anteriormente por Marra *et al* (2003).^{138,139}

A utilização da abordagem MALDI MS não direcionada, em que não há detecção do vírus SARS-CoV-2 em amostra clínica, vem se mostrando um método promissor. Embora as identidades químicas dos biomarcadores usados para a detecção de SARS-CoV-2 permaneçam não identificadas, o uso de várias estratégias de aprendizado de máquina para diferenciar perfis espectrais específicos da COVID-19 tem se mostrado muito favorável.¹⁴⁶

Além dos estudos proteômicos, pesquisas recentes relataram alterações no perfil lipídico e metabólico associadas a COVID-19.¹⁴⁷ Os lipídios desempenham diferentes papéis metabólicos essenciais, são componentes estruturais da membrana, moléculas de sinalização e fontes de energia. Devido à vasta gama de funções que os lipídios têm em humanos, a caracterização lipidômica de pacientes com infecção tem o potencial não apenas de identificar novos biomarcadores lipídicos, mas também de melhorar a compreensão do mecanismo da infecção e facilitar o monitoramento de sua evolução a longo prazo.^{148,149}

Vários estudos relataram alterações importantes no perfil lipídico de pacientes com infecção por COVID-19.^{147,150–153} Na MS, o uso de técnicas separativas como a cromatografia líquida, LC-ESI MS/MS, é o método mais comumente utilizado nas aplicações clínicas. Sendo assim, estudos utilizando a LC-MS na caracterização do perfil lipídico e metabólico de amostras de soro também já foram relatadas.¹⁵⁴

A Tabela 1. 4 apresenta artigos disponíveis nos bancos de dados PubMed, Scopus e Web of Science, utilizando as palavras-chaves “(COVID-19 or SARS-CoV-2 or coronavírus) AND mass spectrometry AND machine learning)” entre esses houve a seleção dos que utilizaram as fontes MALDI e ESI para a realização das análises dos biofluidos.

Tabela 1. 4.Estudos utilizando as técnicas MALDI MS e LC-ESI MS e aprendizagem de máquina nas análises de biofluidos com SARS-CoV-2.

Amostra	Tratamento de amostras	Técnica Equipamento MS	Composto identificado	Tratamento dos dados	Teste SARS-CoV-2	Ref
Swab nasofaríngeo	Sem preparo	MALDI-TOF (Autoflex, Bruker Daltonics)	Perfil espectral	FS, PCA e ML Linguagem R	RT-PCR	155
	Sem preparo	MALDI-TOF FlexControl (Bruker Daltonics)	Perfil espectral	GA	RT-PCR	132
	Inativação	MALDI-TOF Autoflex maX (Bruker Daltonics)	Perfil espectral	PCA, XBOOST e SVM	RT-PCR	156
	Sem preparo	MALDI-TOF-MS analisador Shimadzu 8020 (Shimadzu Scientific Instruments)	Perfil espectral	DNN, LR, NB, kNN, SVM, RF e XGBoost GBM	RT-PCR	131
	Sem preparo	MALDI-TOF-MS analisador Shimadzu 8020 (Shimadzu Scientific Instruments)	Perfil espectral	Teste A kNN, GBM, LR, MLP-NN, NB, RF, SVM Teste B AP-NN, BDT, LBFSGS, LGBM, SVM linear, LR, SDCA, SGD	RT-PCR	157
Saliva	Inativação viral,	MALDI-TOF Autoflex (Bruker Daltonics)	Perfil proteico	SVM, GBM, NNET, NB e RF Pacote R	RT-PCR.	87
	Sem preparo	Microflex LT MALDI-TOF (Bruker Daltonics),	Perfil espectral	SVM, RF, KNN e LDA Pacote R	NPSs	88
	4 Diferentes protocolos	Microflex LT MALDI-TOF (Bruker Daltonics),	Perfil proteico	ANOVA, diferença mínima significativa de Fisher (LSD)	PCR	158
	Digestão de proteínas	MALDI FT-ICR MS (Bruker Daltonics)	Perfil proteico	PCA, SVM	RT-PCR	89

Amostra	Tratamento de amostras	Técnica Equipamento MS	Composto identificado	Tratamento dos dados	Teste SARS-CoV-2	Ref
amostras de gargarejo	Extração de proteínas	MALDI-ToF MS (Shimadzu Kratos Analytical)	Perfil espectral	Area sob a curva (AUC)	RT-PCR	159
Plasma	Extração de proteínas	MALDI-TOF Autoflex (Bruker Daltonics)	Perfil proteico	PCA e SVM, KNN, rede neural NNET, NB e RF	PCR	160
Soro	Sem preparo	MALDI-TOF MS (Clin-TOF-II; Bioyong Technologies Inc.)	Perfil espectral	PLS-DA foi realizada usando o <i>Metaboanalyst</i> 4.0 LR, SVM, NB, DT, RF, GBDT, KNN	RT-PCR	36
	Extração de metabólitos	MALDI-TOF (Autoflex Max, Bruker Daltonics)	Perfil metabólico	PCA, UMAP SVM, KNN, RF, MLP e XGB	RT-PCR e ELISA	161
	Sem preparo	MALDI-TOF MS (Clin-TOF-II; Bioyong Technologies Inc.) nanoLC (Thermo-ESAY-nLC) Orbitrap Fusion Lumos (Thermo Fisher Scientific,)	Perfil proteômico Proteínas	PCA e PLS-DA <i>Metaboanalyst</i> UniProtKB	monitoramento de anticorpos	162
	Concentração e dessalinização de proteínas e peptídeos / Digestão de proteínas	Autoflex III MALDI-TOF (Bruker Daltonics) / nanoLC system (Agilent Technologies), Q-Exacte Plus Hybrid Quadrupole-Orbitrap MS (Thermo Fisher Scientific)	Perfil proteico / Identificação de peptídeos	PCA, LR, SVM, NB e RF Proteome Discoverer (Thermo Fisher Scientific)	RT-PCR	163
Swab nasofaringe	Uso de kits de perfis de metabólitos	UHPLC-MS/MS ExionLC AC Series QTRAP 5500 espectrômetro de	Perfil metabólico (fosfatidilcolinas (PC), lisofosfatidilcolinas (LysoPC),	PLS-DA Pacote R	(qRT-PCR)	164

Amostra	Tratamento de amostras	Técnica Equipamento MS	Composto identificado	Tratamento dos dados	Teste SARS-CoV-2	Ref
		massa (Sciex)	esfingomielinas (SM) e hidroxiesfingomielinas SM(OH))			
	corte de massa molar usando um filtro de membrana	RSLC (Thermo Scientific). tribrido Orbitrap Fusion Lumos (Thermo Fisher Scientific)	Identificação de proteínas/ peptídeos	UNIProt	RT-PCR	165
	Digestão de proteínas	LC-MS/MS usando 3000 RSLCnano (Thermo Scientific, San Jose, CA) Orbitrap Eclipse MS (Thermo Fisher Scientific).	Identificação de proteínas/ peptídeos	UniProt	RT-PCR	166
Saliva	Remoção de proteínas (extração metanoica)	RP/UPLC-MS/MS (Waters ACQUITY (UPLC) e um espectrômetro de massa Thermo Scientific Q-Exactive)	Perfil metabólico	PCA, LR e SVM	ELISA	167
	Extração lipídica	Infusão direta ESI Q-Exactive Orbitrap MS (Thermo Fisher Scientific)	Perfil lipídico (esfingolipídios, ceramidas, fosfolipídios eglicerolipídios)	PLS-DA e OPSDA	RT-PCR	152
fluido de lavagem broncoalveolar (BALF)	Extração e digestão de proteínas BALF	EASY-nLC 1200Q Exactive HF-X (Thermo Fisher Scientific)	Proteínas Tenascina-C (TNC), Mucina-1 (KL-6 ou MUC1), Lipocalina-2 (LCN2), periostina (POSTN), Chitinase 3-like 1 (CHI3L1 ou YKL40) e S100A12 , e 6D/E48 (LY6D), CD9	maxquant UniProt e no banco de dados de proteínas NCBI SARS-CoV-2	RT-PCR e ELISA	168

Amostra	Tratamento de amostras	Técnica Equipamento MS	Composto identificado e CD177	Tratamento dos dados	Teste SARS-CoV-2	Ref
Soro	Extração de proteínas	nanofluxo de alta pressão nanoElute timsTOF (Bruker Daltonics, Bremen, Alemanha)	Perfil proteico	SVM, NB, RF e PLS-DA. MATLAB	qPCR ELISA	169
	Extração lipídica	UHPLC-QTOF 6550 (Agilent Technologies, Santa Clara, CA, EUA).	Perfil lipídico	PCA e PLS-DA	PCR	170
	Digestão de proteínas	no sistema Eksigent NanoLC 400 (Eksigent, Dublin, CA, EUA) acoplado ao sistema TripleTOF 5600 e 6600 (SCIEX, CA, EUA)	Perfil proteico	RF	RT-PCR	171
Plasma	Inativação do vírus digestão de proteína	HPLC EASY-Spray NG Orbitrap Fusion Lumos Tribrid (Thermo Scientific)	Detecção do RNA do coronavírus 2	SVM	RT-qPCR e ELISA	172
	Extração de lipídios	UHPLC-TIMS foram realizadas UHPLC TimsTOF (Q-TOF) (Bruker Daltonics, Bremen, Alemanha)	perfil lipídico	<i>Metaboanalyst</i> 5.0 PCA, PLS-DA, RF	PCR	173
	Inativação do vírus digestão de proteína	nano LC 1200 Orbitrap Fusion Tribrid (Thermo Fischer Scientific)	Perfil proteico proteínas	PLS-DA, PCA, SVM Orange ferramentas Python e MATLAB	RT-PCR	174
	Digestão de proteínas	UHPLC-MS Agilent 6495C	Perfil proteico	PCA Pacote R	PCR	175
	Digestão de proteínas	LC-TripleTOF 6600 (SCIEX)	Perfil proteico	aprendizado de máquina baseado em árvores intensificadas por gradiente	PCR	176

Amostra	Tratamento de amostras	Técnica Equipamento MS	Composto identificado	Tratamento dos dados	Teste SARS-CoV-2	Ref
	Extração de metabólitos	Vanquish UHPLC acoplado on-line a um espectrômetro de massa Q Exactive (Thermo Fisher, Bremen, Alemanha).	Metabólitos e lipídios	<i>Metaboanalyst</i> 4.0 RF e SVM	RT-PCR, ELISA	177
		HESI-Q Exactive Orbitrap-MS (Thermo Scientific, Bremen, Alemanha)	Perfil metabólico	ADA, GDB, RF, XRF, PLS e SVM	RT-PCR	178

ADA, (ADA tree boosting); curva ROC (Receiver Operating Characteristic); DNN (deep neural network); DT (Decision Tree); ESI (Electrospray ionization); GA (Genetic Algorithm); GBDT (Gradient Boosted Decision Trees); GBM (Gradient Boost Machine); GDB (gradient tree boosting); HPLC (High-performance liquid chromatography); KNN (Nearest Neighbors); LC (Liquid chromatography); LDA (Linear Discriminant Analysis); LR (Logistic Regression); MALDI (Matrix assisted laser desorption ionization); ML (machine learning) MLP (Multilayer Perceptron); MS (Mass spectrometry); NB (Naïve Bayes); OPS (Ordered predictor selection); PCA (Principal component analysis); PLS-DA (Partial least squares-discriminant analysis); RF (Random Forest); SNN (Supervised Neural Network); SVM (support-vector machine); TOF (Time-of-flight); UPLC (Ultra Performance Liquid Chromatography); UHPLC (Ultra High Performance Liquid Chromatography); UMAP (uniform manifold approximation and projection); XGB (XGBoost); XRF (extreme random forest)

Logo no início da pandemia, começo de 2020, os estudos buscavam diversas formas de identificar os indivíduos contaminados, visando isolar pacientes com a doença, dada a alta contagiosidade do SARS-CoV-2. Um dos primeiros estudos utilizando a MS para triagem de pacientes foi em 2020. Nachtigall *et al* (2020)¹⁵⁵ relataram o desenvolvimento de um método MALDI MS para o diagnóstico de infecção por SARS-CoV-2. Neste estudo, amostras de swabs nasais foram analisadas diretamente por MALDI MS e um padrão espectral foi utilizado para classificar os pacientes suspeitos como infectados ou não. O modelo de SVM utilizado forneceu precisão de 93,9%, com 7% de falsos positivos e 5% de falsos negativos, uma aplicação interessante de MS para a detecção de infecções virais. Entretanto, seria interessante a inclusão de amostras com vírus ou bactérias, potenciais falsos positivos, além de apenas amostras caracterizadas como positivas ou negativas para SARS-CoV-2 por RT-PCR.¹⁵⁵

Outro estudo também utilizando amostras residuais de swab nasal, análise por MALDI-TOF MS e ML foi realizado por Tran e colaboradores.¹³¹ Dois modelos de ML otimizados foram identificados, exibindo acurácia de 98,3%, concordância percentual positiva (PPA) de 100%, concordância percentual negativa (NPA) de 96% e acurácia de 96,6%, PPA de 98,5% e NPA de 94%, respectivamente.¹³¹

O uso de amostra sem ou com mínimo preparo faz com que seja necessário a aplicação de ferramentas estatísticas mais sofisticadas e avançadas como a ML. Sendo assim, a observação de semelhanças e diferenças nas aplicações de diferentes métodos quimiométricos se torna interessante para encontrar um preditivo ideal para um determinado conjunto de dados.¹⁵⁷

Além de amostras de swab nasal, o biofluido padrão para o teste RT-PCR, diferentes biofluidos foram estudados e analisados por MALDI MS e ML como saliva e soro. Apesar da necessidade de um protocolo de coleta de saliva e preparo de amostras bem definido, a singularidade dos perfis de MS entre os indivíduos faz com que a identificação de perfis espectrais específicos associados ao status de infecção por COVID-19 exija análises bioestatísticas avançadas, como o uso da seleção de variáveis e o desenvolvimento de abordagens de ML¹⁵⁸

Sendo assim, foi observado na pesquisa de Almeida *et al* (2022)⁸⁹ o uso do discriminante de Fisher para a seleção de variáveis e do ML para classificação de amostras de saliva de pacientes com suspeita de COVID-19, utilizando MALDI-FT-ICR MS após a digestão de proteínas e purificação. Os modelos SVM apresentaram uma taxa de falso negativo de 0%, selecionando 780 variáveis e de 3%, selecionando apenas duas

variáveis. O melhor modelo SVM apresentou sensibilidade de 100%, especificidade de 87.5%, e acurácia de 95.6%.⁸⁹

Já no estudo de Bernado et al.(2024),¹⁵² avaliando o perfil lipídico de amostras de saliva, foi utilizado com sucesso o método seleção ordenada de preditores (do inglês, *ordered predictor selection*, OPS) integrado com PLS-DA para classificar amostras de saliva de pacientes com COVID-19. O modelo de classificação também demonstrou excelente previsibilidade com uma acurácia de 93,2%, especificidade de 90,5% no grupo positivo para COVID-19, e sensibilidade de 94,4% para o grupo saudável.¹⁵²

Em estudos utilizando o soro ou plasma como matriz de pesquisa também foi observada a necessidade da utilização da ML com seleção de sinais de interesse, por apresentar resultados mais significativos. Entretanto, o protocolo de coleta e preparo de amostra do soro é mais simples do que o da saliva.¹⁷⁹

Como no estudo de Yan *et al.* (2021)³⁶ com amostras de soro, em que foi necessário apenas a mistura de um tampão e da matriz auxiliar para ionização para análise por MALDI-TOF MS e para o tratamento de dados, um modelo de ML de regressão logística com 25 sinais selecionados de recursos alcançou acurácia de 99%, com sensibilidade de 98% e especificidade de 100%, para a detecção de COVID-19.³⁶

Além de estudos para triagem de pacientes, amostras de soro analisadas por MS podem fornecer informações sobre a resposta à imunização humana após a vacinação contra a COVID-19. Zhang e colaboradores (2022)¹⁶² observaram uma diferença significativa nos perfis peptidoma sérico por MALDI-TOF MS, junto com métodos de aprendizado de máquina, após a vacinação. Foi possível distinguir indivíduos positivos para anticorpos neutralizantes de indivíduos negativos para esses anticorpos após a vacinação completa.¹⁶²

Esses estudos demonstram a simplicidade e o menor custo da aplicação do MALDI MS combinada com a aprendizagem de máquina, avaliando o perfil proteômicos em diferentes biofluidos. Os resultados demonstraram um grande potencial do método para triagem, diagnóstico de COVID-19 e estudos sobre a eficiência da vacinação.

Pesquisas recentes relataram alterações no perfil metabólico e lipídico associadas a COVID-19.^{147,161} Wan e colaboradores (2021)¹⁶¹ promoveram um estudo a fim de detectar a infecção assintomática por SARS-CoV-2. Padrões de metabólitos e lipídios foram extraídos de amostras de soro e analisados por MALDI MS. Um novo modelo de aprendizado de conjunto foi desenvolvido usando a estratégia de empilhamento com um novo algoritmo de votação. Essa abordagem foi validada em uma grande coorte de 274

amostras (92 COVID-19 assintomáticas e 182 controles saudáveis) e forneceu alta precisão de 93,4%, com apenas 5% de falsos negativos e 7% de falsos positivos. Também foi identificado um painel de biomarcadores de dez metabólitos e lipídios, além de vias metabólicas alteradas durante a infecção assintomática por SARS-CoV-2.¹⁶¹

Além de MALDI MS, outra técnica muito utilizada no estudo de amostras de COVID-19 foi a LC MS. No estudo de Gray e colaboradores (2021)¹⁸⁰ foi desenvolvido um método eficaz usando perfis lipídicos plasmáticos para discriminar infecção por SARS-CoV-2, controles saudáveis e outras infecções respiratórias. O perfil lipídico direcionado por UHPLC-MS/MS foi realizado e um perfil direcionado de biomarcadores lipídicos desenvolvido. Usando um painel de lipídios, a modelagem de regressão logística resultou em um forte modelo classificador que pode determinar a infecção por SARS-CoV-2 em comparação com controles saudáveis. Em um subconjunto de amostras, o painel lipídico também foi capaz de distinguir indivíduos positivos para SARS-CoV-2 daqueles que relataram sintomas, mas testaram negativo para o vírus, indicando que o painel pode ser capaz de diferenciar entre SARS-CoV-2 e outras infecções virais presentes na comunidade.¹⁸⁰

Em outro estudo utilizando UHPLC-MS para identificar perfis proteômicos e metabolômicos de soros de indivíduos com COVID-19 e controles, foram identificadas alterações moleculares no soro de pacientes com COVID-19 em comparação com outros grupos, implicando desregulação de macrófagos, plaquetas degranulação, vias do sistema complemento e supressão metabólica maciça. Este estudo revelou alterações características de proteínas e metabólitos no soro de pacientes graves com COVID-19, que podem ser usadas na seleção de potenciais biomarcadores sanguíneos para avaliação da gravidade.¹⁵⁴

Na metabolômica direcionada, substâncias químicas pré-definidas são medidas com base em procedimentos padronizados, e é a mais utilizada quando se quer detectar o vírus já existente. No entanto, uma abordagem não direcionada com análise mais abrangente, pode nos dar informações de forma mais rápida, medindo todos os metabólitos detectáveis em uma determinada amostra, incluindo metabólitos desconhecidos, e fornecendo um perfil espectral daquela amostra. Apesar de mais rápida, existe a necessidade de um extenso processamento de dados e validação estatística grande para gerar resultados clinicamente significativos, devido a quantidade de dados brutos gerados.¹⁴⁶

Outra opção para análise de lipídios e metabólitos de forma rápida e eficiente é a ESI MS por infusão direta. Estudos já mostraram sua eficiência em análise direta de swabs clínicas para triagem de COVID-19, permitindo a detecção rápida de perfis lipídicos ricos. Dois classificadores estatísticos foram gerados com base nas informações lipídicas adquiridas, produzindo uma exatidão do conjunto de validação de 89,6%, sensibilidade de 100%, e especificidade de 85,3%.¹⁵¹

No estudo de Bernardo et al, o uso da ESI MS por infusão direta na análise de extratos lipídicos de amostras de saliva de pacientes com COVID-19 combinada com os PLS-DA, observou que os lipídios mais importantes para a classificação da COVID-19 foram os esfingolipídios, ceramidas, fosfolipídios e glicerolipídios.¹⁵²

Ademais, ao comparar o perfil metabólico sérico e salivar, capaz de diferenciar pacientes positivos para COVID-19 e controles por LC-MS, Spick e colaboradores(2022)¹⁷⁹ relataram que o soro teve melhor desempenho por análise multivariada (sensibilidade e especificidade de 0,97), com alterações dominantes nos níveis de triglicerídeos e ácidos biliares, observando uma dislipidemia da infecção por COVID-19.¹⁷⁹

Sendo assim, durante esses 4 anos após o início da pandemia houve grande publicação de artigos relacionando as palavras-chave “(COVID-19 or SARS-CoV-2 or coronavírus) AND mass spectrometry AND machine learning)” com ápice em 2022, com decréscimo posterior. Mas ainda existem estudos que buscam aperfeiçoar os diagnósticos e auxiliar na tomada de decisão, empregando estudos multiômicos e inteligência artificial (IA) em amostras de COVID-19 analisadas por MS.^{181,182} Então é observado que testes não direcionados a espécies moleculares específicas, usando a MS para detecção de COVID-19, estão sendo avaliados como métodos de triagem rápida para atenuar surtos virais e têm o potencial de melhorar o diagnóstico clínico.

1.5 CONCLUSÃO

O presente estado da arte buscou apresentar estudos atuais sobre o uso da MS combinada a tratamentos quimiométricos e bioestatísticos em análises de biofluidos, a fim de obter informações sobre as condições fisiológicas do indivíduo, observando alterações moleculares e identificando possíveis biomarcadores de doenças. As aplicações clínicas baseadas em MS utilizando as fontes de ionização suave ESI e MALDI mostraram-se ferramentas promissoras para detecção de perfis espectrais, sendo em análises proteômicas, metabolômicas ou lipidômicas. O preparo simplificado das amostras e o uso de análises diretas aumenta a velocidade das análises e conseqüentemente dos resultados dos testes diagnósticos, entretanto necessitam de tratamentos de dados mais sofisticados e complexos. Em estudos de amostras de pacientes com COVID-19 a necessidade de avaliar o desenvolvimento e sequelas dessa doença ainda despertam interesse. Além disso, o aperfeiçoamento dos resultados, otimizando os tratamentos de dados com auxílio da IA, pode ser um campo promissor na pesquisa clínica.

1.6 REFERÊNCIAS

- (1) Lima, N. M.; Fernandes, B. L. M.; Alves, G. F.; de Souza, J. C. Q.; Siqueira, M. M.; Patrícia do Nascimento, M.; Moreira, O. B. O.; Sussulini, A.; de Oliveira, M. A. L. Mass Spectrometry Applied to Diagnosis, Prognosis, and Therapeutic Targets Identification for the Novel Coronavirus SARS-CoV-2: A Review. *Anal. Chim. Acta* **2022**, *1195*, 339385.
- (2) Griffin, J. H.; Downard, K. M. Mass Spectrometry Analytical Responses to the SARS-CoV2 Coronavirus in Review. *TrAC Trends Anal. Chem.* **2021**, *142*, 116328.
- (3) Mohanty, A.; Fatrekar, A. P.; Krishnan, S.; Vernekar, A. A. A Concise Discussion on the Potential Spectral Tools for the Rapid COVID-19 Detection. *Results Chem.* **2021**, *3*, 100138.
- (4) Banerjee, S. Empowering Clinical Diagnostics with Mass Spectrometry. *ACS Omega* **2020**, *5* (5), 2041–2048.
- (5) Chen, Y. T.; Liao, W. R.; Wang, H. T.; Chen, H. W.; Chen, S. F. Targeted Protein Quantitation in Human Body Fluids by Mass Spectrometry. *Mass Spectrom. Rev.* **2022**, No. September 2021, 1–25.
- (6) Zamora-Obando, H. R.; Godoy, A. T.; Amaral, A. G.; de Mesquita, A. S.; Simões, B. E. S.; Reis, H. O.; Rocha, I.; Dallaqua, M.; Baptista, M.; Fernandes, M. C. V.; Lima, M. F.; Simionato, A. V. C. BIOMARCADORES MOLECULARES DE DOENÇAS HUMANAS: CONCEITOS FUNDAMENTAIS, MODELOS DE ESTUDO E APLICAÇÕES CLÍNICAS. *Quim. Nova* **2022**, *45* (9), 1098–1113.
- (7) Greco, V.; Piras, C.; Pieroni, L.; Ronci, M.; Putignani, L.; Roncada, P.; Urbani, A. Applications of MALDI-TOF Mass Spectrometry in Clinical Proteomics. *Expert Rev. Proteomics* **2018**, *15* (8), 683–696.
- (8) Kucharski, A. J.; Russell, T. W.; Diamond, C.; Liu, Y.; Edmunds, J.; Funk, S.; Eggo, R. M.; Sun, F.; Jit, M.; Munday, J. D.; Davies, N.; Gimma, A.; van Zandvoort, K.; Gibbs, H.; Hellewell, J.; Jarvis, C. I.; Clifford, S.; Quilty, B. J.; Bosse, N. I.; Abbott, S.; Klepac, P.; Flasche, S. Early Dynamics of Transmission and Control of COVID-19: A Mathematical Modelling Study. *Lancet. Infect. Dis.* **2020**, *20* (5), 553–558.
- (9) Aria, M.; Cuccurullo, C. Bibliometrix: An R-Tool for Comprehensive Science Mapping Analysis. *J. Informetr.* **2017**, *11* (4), 959–975.
- (10) Mahmud, I.; Garrett, T. J. Mass Spectrometry Techniques in Emerging Pathogens Studies: COVID-19 Perspectives. *J. Am. Soc. Mass Spectrom.* **2020**, *31* (10), 2013–

2024.

- (11) Dunphy, K.; O'mahoney, K.; Dowling, P.; O'gorman, P.; Bazou, D. Clinical Proteomics of Biofluids in Haematological Malignancies. *Int. J. Mol. Sci.* **2021**, *22* (15), 1–21.
- (12) Wu, Z.; Bagarolo, G. I.; Thoröe-Boveleth, S.; Jankowski, J. "Lipidomics": Mass Spectrometric and Chemometric Analyses of Lipids. *Adv. Drug Deliv. Rev.* **2020**, *159*, 294–307.
- (13) Züllig, T.; Trötz Müller, M.; Köfeler, H. C. Lipidomics from Sample Preparation to Data Analysis: A Primer. *Anal. Bioanal. Chem.* **2019**, *412* (10), 2191–2209.
- (14) Dettmer, K.; Aronov, P. A.; Hammock, B. D. Mass Spectrometry-Based Metabolomics. *Mass Spectrom. Rev.* **2007**, *26* (1), 51–78.
- (15) Thomas, S. L.; Thacker, J. B.; Schug, K. A.; Maráková, K. Sample Preparation and Fractionation Techniques for Intact Proteins for Mass Spectrometric Analysis. *J. Sep. Sci.* **2021**, *44* (1), 211–246.
- (16) Wang, W.; Becker, C. H.; Zhou, H.; Lin, H.; Roy, S.; Shaler, T. A.; Hill, L. R.; Norton, S.; Kumar, P.; Anderle, M. Quantification of Proteins and Metabolites by Mass Spectrometry without Isotopic Labeling or Spiked Standards. *Anal. Chem.* **2003**, *75* (18), 4818–4826.
- (17) Bligh, E. G.; Dyer, W. J. A RAPID METHOD OF TOTAL LIPID EXTRACTION AND PURIFICATION. *Can. J. Biochem. Physiol.* **1959**, *37*, 911–917.
- (18) Nickerson, J. L.; Doucette, A. A. Rapid and Quantitative Protein Precipitation for Proteome Analysis by Mass Spectrometry. *J. Proteome Res.* **2020**, *19* (5), 2035–2042.
- (19) Abdelati, A. A.; Elnemr, R. A.; Kandil, N. S.; Dwedar, F. I.; Ghazala, R. A. Serum Peptidomic Profile as a Novel Biomarker for Rheumatoid Arthritis. *Int. J. Rheumatol.* **2020**, *2020*.
- (20) *Towards defining the urinary proteome using liquid chromatography-tandem mass spectrometry I. Profiling an unfractionated tryptic digest - Spahr - 2001 - PROTEOMICS - Wiley Online Library.*
[https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/1615-9861\(200101\)1:1%3C93::AID-PROT93%3E3.0.CO;2-3](https://analyticalsciencejournals.onlinelibrary.wiley.com/doi/10.1002/1615-9861(200101)1:1%3C93::AID-PROT93%3E3.0.CO;2-3) (accessed 2024-01-11).
- (21) TORTORA, Gerard J.; DERRICKSON, B. *Corpo Humano*, 10ª edição.; 2016.
- (22) Issaq, H. J.; Xiao, Z.; Veenstra, T. D. Serum and Plasma Proteomics. *Chem. Rev.*

2007, 107 (8), 3601–3620.

- (23) Jelonek, K.; Ros, M.; Pietrowska, M.; Widlak, P. Cancer Biomarkers and Mass Spectrometry-Based Analyses of Phospholipids in Body Fluids. *Clin. Lipidol.* **2013**, 8 (1), 137–150.
- (24) Zhou, X.; Zhang, W.; Ouyang, Z. Recent Advances in On-Site Mass Spectrometry Analysis for Clinical Applications. *TrAC Trends Anal. Chem.* **2022**, 149, 116548.
- (25) Donnelly, D. P.; Rawlins, C. M.; DeHart, C. J.; Fornelli, L.; Schachner, L. F.; Lin, Z.; Lippens, J. L.; Aluri, K. C.; Sarin, R.; Chen, B.; Lantz, C.; Jung, W.; Johnson, K. R.; Koller, A.; Wolff, J. J.; Campuzano, I. D. G.; Auclair, J. R.; Ivanov, A. R.; Whitelegge, J. P.; Paša-Tolić, L.; Chamot-Rooke, J.; Danis, P. O.; Smith, L. M.; Tsybin, Y. O.; Loo, J. A.; Ge, Y.; Kelleher, N. L.; Agar, J. N. Best Practices and Benchmarks for Intact Protein Analysis for Top-down Mass Spectrometry. *Nat. Methods* **2019**, 16 (7), 587–594.
- (26) Finoulst, I.; Pinkse, M.; Van Dongen, W.; Verhaert, P. Sample Preparation Techniques for the Untargeted LC-MS-Based Discovery of Peptides in Complex Biological Matrices. *J. Biomed. Biotechnol.* **2011**, 2011, 14.
- (27) Callesen, A. K.; Madsen, J. S.; Vach, W.; Kruse, T. A.; Mogensen, O.; Jensen, O. N. Serum Protein Profiling by Solid Phase Extraction and Mass Spectrometry: A Future Diagnostics Tool? *Proteomics* **2009**, 9 (6), 1428–1441.
- (28) Fernández-Ochoa, Á.; Brunius, C.; Borrás-Linares, I.; Quirantes-Piné, R.; Cádiz-Gurrea, M. D. L. L.; Alarcón Riquelme, M. E.; Segura-Carretero, A. Metabolic Disturbances in Urinary and Plasma Samples from Seven Different Systemic Autoimmune Diseases Detected by HPLC-ESI-QTOF-MS. *J. Proteome Res.* **2020**, 19 (8), 3220–3229.
- (29) Liu, D.; Meister, M.; Zhang, S.; Vong, C. I.; Wang, S.; Fang, R.; Li, L.; Wang, P. G.; Massion, P.; Ji, X. Identification of Lipid Biomarker from Serum in Patients with Chronic Obstructive Pulmonary Disease. *Respir. Res.* **2020**, 21 (1), 1–13.
- (30) Benabdelkamel, H.; Alamri, H.; Okla, M.; Masood, A.; Jabar, M. A.; Alanazi, I. O.; Alfadda, A. A.; Nizami, I.; Dasouki, M.; Rahman, A. M. A. Serum-Based Proteomics Profiling in Adult Patients with Cystic Fibrosis. *Int. J. Mol. Sci.* **2020**, Vol. 21, Page 7415 **2020**, 21 (19), 7415.
- (31) Ding, D.; Chen, M.; Xiao, X.; Cao, P.; Li, S. Novel Serum Peptide Model Revealed by MALDI-TOF-MS and Its Diagnostic Value in Early Bladder Cancer. *Int. J. Biol. Markers* **2020**, 35 (3), 59–66.

- (32) Abe, K.; Shang, J.; Shi, X.; Yamashita, T.; Hishikawa, N.; Takemoto, M.; Morihara, R.; Nakano, Y.; Ohta, Y.; Deguchi, K.; Ikeda, M.; Ikeda, Y.; Okamoto, K.; Shoji, M.; Takatama, M.; Kojo, M.; Kuroda, T.; Ono, K.; Kimura, N.; Matsubara, E.; Osakada, Y.; Wakutani, Y.; Takao, Y.; Higashi, Y.; Asada, K.; Senga, T.; Lee, L. J.; Tanaka, K. A New Serum Biomarker Set to Detect Mild Cognitive Impairment and Alzheimer's Disease by Peptidome Technology. *J. Alzheimer's Dis.* **2020**, *73* (1), 217–227.
- (33) Răchieriu, C.; Eniu, D. T.; Moiş, E.; Graur, F.; Socaciu, C.; Socaciu, M. A.; Hajjar, N. Al. Lipidomic Signatures for Colorectal Cancer Diagnosis and Progression Using UPLC-QTOF-ESI+MS. *Biomol.* **2021**, *Vol. 11, Page 417* **2021**, *11* (3), 417.
- (34) Homorogan, C.; Nitusca, D.; Enatescu, V.; Schubart, P.; Moraru, C.; Socaciu, C.; Marian, C. Untargeted Plasma Metabolomic Profiling in Patients with Major Depressive Disorder Using Ultra-high Performance Liquid Chromatography Coupled with Mass Spectrometry. *Metabolites* **2021**, *11* (7), 466.
- (35) Taşçı, Y.; Fındık, R. B.; Pekcan, M. K.; Kaplan, O.; Celebier, M. UPLC-Q-TOF/MS Based Untargeted Metabolite and Lipid Analysis on Premature Ovarian Insufficiency Plasma Samples. *Curr. Pharm. Anal.* **2020**, *17* (4), 474–483.
- (36) Yan, L.; Yi, J.; Huang, C.; Zhang, J.; Fu, S.; Li, Z.; Lyu, Q.; Xu, Y.; Wang, K.; Yang, H.; Ma, Q.; Cui, X.; Qiao, L.; Sun, W.; Liao, P. Rapid Detection of COVID-19 Using MALDI-TOF-Based Serum Peptidome Profiling. *Anal. Chem.* **2021**, [acs.analchem.0c04590](https://doi.org/10.1021/acs.analchem.0c04590).
- (37) Ma, D.; Liang, N.; Zhang, L. Establishing Classification Tree Models in Rheumatoid Arthritis Using Combination of Matrix-Assisted Laser Desorption/Ionization Time-of-Flight Mass Spectrometry and Magnetic Beads. *Front. Med.* **2021**, *8*, 609773.
- (38) Nenu, I.; Stefanescu, H.; Procopet, B.; Sparchez, Z.; Minciuna, I.; Mocan, T.; Leucuta, D.; Morar, C.; Grigorescu, M.; Filip, G. A.; Socaciu, C. Navigating through the Lipid Metabolism Maze: Diagnosis and Prognosis Metabolites of Hepatocellular Carcinoma versus Compensated Cirrhosis. *J. Clin. Med.* **2022**, *11* (5), 1292.
- (39) Alsoud, L. O.; Soares, N. C.; Al-Hroub, H. M.; Mousa, M.; Kasabri, V.; Bulatova, N.; Suyagh, M.; Alzoubi, K. H.; El-Huneidi, W.; Abu-Irmaileh, B.; Bustanji, Y.; Semreen, M. H. Identification of Insulin Resistance Biomarkers in Metabolic Syndrome Detected by UHPLC-ESI-QTOF-MS. *Metabolites* **2022**, *12* (6), 508.
- (40) Maslov, D. L.; Trifonova, O. P.; Lichtenberg, S.; Balashova, E. E.; Mamedli, Z. Z.; Alferov, A. A.; Stilidi, I. S.; Lokhov, P. G.; Kushlinskii, N. E.; Archakov, A. I. Blood Plasma Metabolome Profiling at Different Stages of Renal Cell Carcinoma. *Cancers*

(Basel). **2023**, 15 (1), 140.

- (41) Pečinka, L.; Vlachová, M.; Moráň, L.; Gregorová, J.; Porokh, V.; Kovačovicová, P.; Almáši, M.; Pour, L.; Štork, M.; Havel, J.; Ševčíková, S.; Vaňhara, P. Improved Screening of Monoclonal Gammopathy Patients by MALDI-TOF Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2023**, 34 (12), 2646–2653.
- (42) Rungkamoltip, P.; Roytrakul, S.; Navakanitworakul, R. MALDI-TOF MS Analysis of Serum Peptidome Patterns in Cervical Cancer. *Biomed. 2023, Vol. 11, Page 2327* **2023**, 11 (8), 2327.
- (43) Alshahrani, A.; Aljada, A.; Masood, A.; Mujammami, M.; Alfadda, A. A.; Musambil, M.; Alanazi, I. O.; Al Dubayee, M.; Abdel Rahman, A. M.; Benabdelkamel, H. Proteomic Profiling Identifies Distinct Regulation of Proteins in Obese Diabetic Patients Treated with Metformin. *Pharmaceuticals* **2023**, 16 (10), 1345.
- (44) Kianičková, K.; Pažitná, L.; Kundalia, P. H.; Pakanová, Z.; Nemčovič, M.; Baráth, P.; Katrlíková, E.; Šuba, J.; Trebatická, J.; Katrlík, J. Alterations in the Glycan Composition of Serum Glycoproteins in Attention-Deficit Hyperactivity Disorder. *Int. J. Mol. Sci.* **2023**, Vol. 24, Page 8745 **2023**, 24 (10), 8745.
- (45) Liu, L.; Liu, H.; Zhu, K.; Zhang, L.; Yin, X.; Han, L.; Wang, M.; Gao, S.; Xiao, X.; Yang, J.; Huang, C.; Huang, Y. Proteome Analysis Reveals Novel Serum Biomarkers for Henoch-Schönlein Purpura in Chinese Children. *J. Proteomics* **2023**, 276, 104841.
- (46) Wu, Q.; Yu, J.; Zhang, M.; Xiong, Y.; Zhu, L.; Wei, B.; Wu, T.; Du, Y. Serum Lipidomic Profiling for Liver Cancer Screening Using Surface-Assisted Laser Desorption Ionization MS and Machine Learning. *Talanta* **2024**, 268, 125371.
- (47) Ng, E. W. Y. E. E. W.; Wong, M. Y. M. Y.; Poon, T. C. W. C. T.; Ng, E. W. Y. E. E. W.; Poon, T. C. W. C. T.; Ka Shing, L.; Wong Li Ka, M. *Advances in MALDI Mass Spectrometry in Clinical Diagnostic Applications*; 2013; Vol. 336, pp 139–175.
- (48) Tan, H. T.; Low, J.; Lim, S. G.; Chung, M. C. M. Serum Autoantibodies as Biomarkers for Early Cancer Detection. *FEBS J.* **2009**, 276 (23), 6880–6904.
- (49) Park, H. G.; Jang, K. S.; Park, H. M.; Song, W. S.; Jeong, Y. Y.; Ahn, D. H.; Kim, S. M.; Yang, Y. H.; Kim, Y. G. MALDI-TOF MS-Based Total Serum Protein Fingerprinting for Liver Cancer Diagnosis. *Analyst* **2019**, 144 (7), 2231–2238.
- (50) Yildiz, P. B.; Shyr, Y.; Rahman, J. S. M.; Wardwell, N. R.; Zimmerman, L. J.; Shakhtour, B.; Gray, W. H.; Chen, S.; Li, M.; Roder, H.; Liebler, D. C.; Bigbee, W. L.; Siegfried, J. M.; Weissfeld, J. L.; Gonzalez, A. L.; Ninan, M.; Johnson, D. H.;

Carbone, D. P.; Caprioli, R. M.; Massion, P. P. Diagnostic Accuracy of MALDI Mass Spectrometric Analysis of Unfractionated Serum in Lung Cancer. *J. Thorac. Oncol.* **2007**, *2* (10), 893–901.

- (51) Swiatly, A.; Horal, A.; Hajduk, J.; Matysiak, J.; Nowak-Markwitz, E.; Kokot, Z. J. MALDI-TOF-MS Analysis in Discovery and Identification of Serum Proteomic Patterns of Ovarian Cancer. *BMC Cancer* **2017**, *17* (1), 1–9.
- (52) Lei, Z.; Huhman, D. V.; Sumner, L. W. Mass Spectrometry Strategies in Metabolomics. *J. Biol. Chem.* **2011**, *286* (29), 25435–25442.
- (53) Anand, S.; Young, S. A.; Esplin, M. S.; Peadar, B.; Tolley, H. D.; Porter, T. F.; Varner, M. W.; D'Alton, M. E.; Jackson, B. J.; Graves, S. W. Detection and Confirmation of Serum Lipid Biomarkers for Preeclampsia Using Direct Infusion Mass Spectrometry. *J. Lipid Res.* **2016**, *57* (4), 687–696.
- (54) Dunn, W. B.; Ellis, D. I. Metabolomics: Current Analytical Platforms and Methodologies. *TrAC - Trends Anal. Chem.* **2005**, *24* (4), 285–294.
- (55) González-Domínguez, R.; García-Barrera, T.; Gómez-Ariza, J. L. Metabolomic Study of Lipids in Serum for Biomarker Discovery in Alzheimer's Disease Using Direct Infusion Mass Spectrometry. *J. Pharm. Biomed. Anal.* **2014**, *98*, 321–326.
- (56) Harshfield, E. L.; Koulman, A.; Ziemek, D.; Marney, L.; Fauman, E. B.; Paul, D. S.; Stacey, D.; Rasheed, A.; Lee, J. J.; Shah, N.; Jabeen, S.; Imran, A.; Abbas, S.; Hina, Z.; Qamar, N.; Mallick, N. H.; Yaqoob, Z.; Saghir, T.; Rizvi, S. N. H.; Memon, A.; Rasheed, S. Z.; Memon, F. U. R.; Qureshi, I. H.; Ishaq, M.; Frossard, P.; Danesh, J.; Saleheen, D.; Butterworth, A. S.; Wood, A. M.; Griffin, J. L. An Unbiased Lipid Phenotyping Approach to Study the Genetic Determinants of Lipids and Their Association with Coronary Heart Disease Risk Factors. *J. Proteome Res.* **2019**, *18* (6), 2397–2410.
- (57) Griffin Timothy Griffin, T. J.; Griffin, T. J. Human Saliva Proteome: An Overview. *Sens. Technol. Glob. Heal. Mil. Med. Environ. Monit. IV* **2014**, *9112* (5), 29–38.
- (58) Maddu, N.; Maddu, N. Functions of Saliva. *Saliva and Salivary Diagnostics* **2019**.
- (59) Wang, A.; Wang, C. P.; Tu, M.; Wong, D. T. W. Oral Biofluid Biomarker Research: Current Status and Emerging Frontiers. *Diagnostics* **2016**, *6* (4), 45.
- (60) Tiwari, M. Science behind Human Saliva. *J. Nat. Sci. Biol. Med.* **2011**, *2* (1), 53.
- (61) Chojnowska, S.; Baran, T.; Wilińska, I.; Sienicka, P.; Cabaj-Wiater, I.; Knaś, M. Human Saliva as a Diagnostic Material. *Adv. Med. Sci.* **2018**, *63* (1), 185–191.
- (62) Agatonovic-Kustrin, S.; Morton, D. W.; Smirnov, V.; Petukhov, A.; Gegechkori, V.;

- Kuzina, V.; Gorpinchenko, N.; Ramenskaya, G. Analytical Strategies in Lipidomics for Discovery of Functional Biomarkers from Human Saliva. *Dis. Markers* **2019**, *2019*.
- (63) Javaid, M. A.; Ahmed, A. S.; Durand, R.; Tran, S. D. Saliva as a Diagnostic Tool for Oral and Systemic Diseases. *J. Oral Biol. Craniofacial Res.* **2016**, *6* (1), 67–76.
- (64) Williamson, S.; Munro, C.; Pickler, R.; Grap, M. J.; Elswick, R. K. Comparison of Biomarkers in Blood and Saliva in Healthy Adults. *Nurs. Res. Pract.* **2012**, *2012*, 1–4.
- (65) Gug, I. T.; Tertis, M.; Hosu, O.; Cristea, C. Salivary Biomarkers Detection: Analytical and Immunological Methods Overview. *TrAC Trends Anal. Chem.* **2019**, *113*, 301–316.
- (66) Prakasam, S.; Srinivasan, M. Evaluation of Salivary Biomarker Profiles Following Non-Surgical Management of Chronic Periodontitis. *Oral Dis.* **2014**, *20* (2), 171–177.
- (67) Teanpaisan, R.; Thitasomakul, S.; Piwat, S.; Thearmontree, A.; Pithpornchaiyakul, W.; Chankanka, O. Longitudinal Study of the Presence of Mutans Streptococci and Lactobacilli in Relation to Dental Caries Development in 3-24 Month Old Thai Children. *Int. Dent. J.* **2007**, *57* (6), 445–451.
- (68) Out, D.; Hall, R. J.; Granger, D. A.; Page, G. G.; Woods, S. J. Assessing Salivary C-Reactive Protein: Longitudinal Associations with Systemic Inflammation and Cardiovascular Disease Risk in Women Exposed to Intimate Partner Violence. *Brain. Behav. Immun.* **2012**, *26* (4), 543–551.
- (69) Border, M. B.; Schwartz, S.; Carlson, J.; Dibble, C. F.; Kohltfarber, H.; Offenbacher, S.; Buse, J. B.; Bencharit, S. Exploring Salivary Proteomes in Edentulous Patients with Type 2 Diabetes. *Mol. Biosyst.* **2012**, *8* (4), 1304–1310.
- (70) Rao, P. V.; Reddy, A. P.; Lu, X.; Dasari, S.; Krishnaprasad, A.; Biggs, E.; Roberts, C. T.; Nagalla, S. R. Proteomic Identification of Salivary Biomarkers of Type-2 Diabetes. *J. Proteome Res.* **2009**, *8* (1), 239–245.
- (71) Tzioufas, A. G.; Kapsogeorgou, E. K. Saliva Proteomics Is a Promising Tool to Study Sjögren Syndrome. *Nat. Rev. Rheumatol.* **2015**, *11* (4), 202–203.
- (72) Livnat, G.; Bentur, L.; Kuzmishsky, E.; Nagler, R. M. Salivary Profile and Oxidative Stress in Children and Adolescents with Cystic Fibrosis. *J. Oral Pathol. Med.* **2010**, *39* (1), 16–21.
- (73) Pai, N. P.; Balram, B.; Shivkumar, S.; Martinez-Cajas, J. L.; Claessens, C.; Lambert, G.; Peeling, R. W.; Joseph, L. Head-to-Head Comparison of Accuracy of a Rapid

Point-of-Care HIV Test with Oral versus Whole-Blood Specimens: A Systematic Review and Meta-Analysis. *Lancet Infect. Dis.* **2012**, *12* (5), 373–380.

- (74) Krishnamurthy, S.; Vasudeva, S. B.; Vijayasathy, S. Salivary Gland Disorders: A Comprehensive Review. <http://www.wjgnet.com/> **2015**, *4* (2), 56–71.
- (75) Pappa, E.; Vougas, K.; Zoidakis, J.; Vastardis, H. Proteomic Advances in Salivary Diagnostics. *Biochim. Biophys. Acta - Proteins Proteomics* **2020**, *1868* (11), 140494.
- (76) Lee, Y.-H. H.; Wong, D. T.; Felix, D.; Yip Endowed Professor, M.; Dean, A. Saliva: An Emerging Biofluid for Early Detection of Diseases. *Am. J. Dent.* **2009**, *22* (4), 241–248.
- (77) Patel, R. MALDI-TOF MS for the Diagnosis of Infectious Diseases. *Clin. Chem.* **2015**, *61* (1), 100–111.
- (78) Martina, E.; Campanati, A.; Diotallevi, F.; Offidani, A. Saliva and Oral Diseases. *J. Clin. Med.* **2020**, *Vol. 9*, Page 466 **2020**, *9* (2), 466.
- (79) Zhu, C.; Yuan, C.; Wei, F.-Q.; Sun, X.-Y.; Zheng, S.-G. Comparative Evaluation of Peptidome and Microbiota in Different Types of Saliva Samples. *Ann. Transl. Med.* **2020**, *8* (11), 686–686.
- (80) Tang, H.; Yuan, C.; Ma, Z.; Zhu, C.; Tong, P.; Gallagher, J. E.; Zheng, S.; Sun, X. The Potentiality of Salivary Peptide Biomarkers for Screening Patients with Periodontal Diseases by Mass Spectrometry. *Clin. Chim. Acta* **2019**, *495*, 278–286.
- (81) Antezack, A.; Chaudet, H.; Tissot-Dupont, H.; Brouqui, P.; Monnet-Corti, V. Rapid Diagnosis of Periodontitis, a Feasibility Study Using MALDI-TOF Mass Spectrometry. *PLoS One* **2020**, *15* (3), e0230334.
- (82) Romano, F.; Franco, F.; Corana, M.; Abbadessa, G.; Di Scipio, F.; Pergolizzi, B.; Castrignano, C.; Aimetti, M.; Berta, G. N. Cystatin SN (CST1) as a Novel Salivary Biomarker of Periodontitis. *Int. J. Mol. Sci.* **2023**, *Vol. 24*, Page 13834 **2023**, *24* (18), 13834.
- (83) LaputkovÃai, G.; Talian, I.; SchwartzovÃai, V.; SchwartzovÃai, Z. MALDI-TOF MS Profiling in the Discovery and Identification of Salivary Proteomic Patterns of Temporomandibular Joint Disorders. *Open Chem.* **2020**, *18* (1), 1173–1180.
- (84) Tong, P.; Yuan, C.; Sun, X.; Yue, Q.; Wang, X.; Zheng, S. Identification of Salivary Peptidomic Biomarkers in Chronic Kidney Disease Patients Undergoing Haemodialysis. *Clin. Chim. Acta* **2019**, *489*, 154–161.
- (85) Contini, C.; Firinu, D.; Serrao, S.; Manconi, B.; Olianias, A.; Cinetto, F.; Cossu, F.; Castagnola, M.; Messina, I.; Del Giacco, S.; Cabras, T. RP-HPLC-ESI-IT Mass

Spectrometry Reveals Significant Variations of the Human Salivary Protein Profile Associated with Predominantly Antibody Deficiencies. *J. Clin. Immunol.* **2020**, *40* (2), 329–339.

- (86) Jiang, X.; Chen, X.; Wang, T.; Li, Y.; Pan, A.; Wu, J. Perfluorinated Polymer Modified Vertical Silicon Nanowires as Ultra Low Noise Laser Desorption Ionization Substrate for Salivary Metabolites Profiling. *Talanta* **2021**, *225*, 122022.
- (87) Lazari, L. C.; Zerbinati, R. M.; Rosa-Fernandes, L.; Santiago, V. F.; Rosa, K. F.; Angeli, C. B.; Schwab, G.; Palmieri, M.; Sarmiento, D. J. S.; Marinho, C. R. F.; Almeida, J. D.; To, K.; Giannecchini, S.; Wrenger, C.; Sabino, E. C.; Martinho, H.; Lindoso, J. A. L.; Durigon, E. L.; Braz-Silva, P. H.; Palmisano, G. MALDI-TOF Mass Spectrometry of Saliva Samples as a Prognostic Tool for COVID-19. *J. Oral Microbiol.* **2022**, *14* (1).
- (88) Costa, M. M.; Martin, H.; Estellon, B.; Dupé, F. X.; Saby, F.; Benoit, N.; Tissot-dupont, H.; Million, M.; Pradines, B.; Granjeaud, S.; Almeras, L. Exploratory Study on Application of MALDI-TOF-MS to Detect SARS-CoV-2 Infection in Human Saliva. *J. Clin. Med.* **2022**, *11* (2), 295.
- (89) De Almeida, C. M.; Motta, L. C.; Folli, G. S.; Marcarini, W. D.; Costa, C. A.; Vilela, A. C. S.; Barauna, V. G.; Martin, F. L.; Singh, M. N.; Campos, L. C. G.; Costa, N. L.; Vassallo, P. F.; Chaves, A. R.; Endringer, D. C.; Mill, J. G.; Filgueiras, P. R.; Romão, W. MALDI(+) FT-ICR Mass Spectrometry (MS) Combined with Machine Learning toward Saliva-Based Diagnostic Screening for COVID-19. *J. Proteome Res.* **2022**, *21* (8), 1868–1875.
- (90) Hu, H.; Leung, W. K. Mass Spectrometry-Based Proteomics for Discovering Salivary Biomarkers in Periodontitis: A Systematic Review. *Int. J. Mol. Sci.* **2023**, *24* (19), 14599.
- (91) Tajmul, M.; Parween, F.; Singh, L.; Mathur, S. R.; Sharma, J. B.; Kumar, S.; Sharma, D. N.; Yadav, S. Identification and Validation of Salivary Proteomic Signatures for Non-Invasive Detection of Ovarian Cancer. *Int. J. Biol. Macromol.* **2018**, *108*, 503–514.
- (92) Baldini, C.; Giusti, L.; Ciregia, F.; Da Valle, Y.; Giacomelli, C.; Donadio, E.; Ferro, F.; Galimberti, S.; Donati, V.; Bazzichi, L.; Bombardieri, S.; Lucacchini, A. Correspondence between Salivary Proteomic Pattern and Clinical Course in Primary Sjögren Syndrome and Non-Hodgkin's Lymphoma: A Case Report. *J. Transl. Med.* **2011**, *9* (1), 1–8.

- (93) Caseiro, A.; Vitorino, R.; Barros, A. S.; Ferreira, R.; Calheiros-Lobo, M. J.; Carvalho, D.; Duarte, J. A.; Amado, F. Salivary Peptidome in Type 1 Diabetes Mellitus. *Biomed. Chromatogr.* **2012**, *26* (5), 571–582.
- (94) Robert M. Silverstein; Francis X. Webster; David J. Kiemle. *Spectrometric Identification of Organic Compounds*; 2005.
- (95) Gross, J. H. *Mass Spectrometry A Textbook*, 1st ed.; Springer, Ed.; Berlin, 2004.
- (96) ROMÃO, W. *Novas Aplicações Da Espectrometria de Massas Em Química Forense*. 2010.
- (97) Tabet, J. C.; Rebuffat, S. [Nobel Prize 2002 for Chemistry: Mass Spectrometry and Nuclear Magnetic Resonance]. *Med. Sci. (Paris)*. **2003**, *19* (8–9), 865–872.
- (98) Nadler, W. M.; Waidelich, D.; Kerner, A.; Hanke, S.; Berg, R.; Trumpp, A.; Rösli, C. MALDI versus ESI: The Impact of the Ion Source on Peptide Identification. *J. Proteome Res.* **2017**, *16* (3), 1207–1215.
- (99) Hoffmann, E. de; Stroobant Vincent. *Mass Spectrometry: Principles and Applications*, 3rd ed.; Wiley: Inglaterra, 2007.
- (100) Alsaleh, M.; Barbera, T.; ... R. A.-J. of clinical and; 2019, undefined. *Mass Spectrometry: A Guide for the Clinician*. Elsevier.
- (101) Lanças, F. M. *Espectrometria de Massas: Fundamentos, Instrumentação e Aplicações*; Átomo: Campinas, SP, 2019.
- (102) La Scola, B.; Raoult, D. Direct Identification of Bacteria in Positive Blood Culture Bottles by Matrix-Assisted Laser Desorption Ionisation Time-of-Flight Mass Spectrometry. *PLoS One* **2009**, *4* (11), e8041.
- (103) Brioude, G.; Brégeon, F.; Trousse, D.; Flaudrops, C.; Secq, V.; De Dominicis, F.; Chabrières, E.; D'journo, X. B.; Raoult, D.; Thomas, P. A. Rapid Diagnosis of Lung Tumors, a Feasibility Study Using Maldi-Tof Mass Spectrometry. *PLoS One* **2016**, *11* (5), e0155449.
- (104) Diamandis, E. P. Mass Spectrometry as a Diagnostic and a Cancer Biomarker Discovery Tool: Opportunities and Potential Limitations. *Mol Cell Proteomics* **2004**, *3* (4), 367–378.
- (105) Knochenmuss, R.; Zenobi, R. MALDI Ionization: The Role of In-Plume Processes. *Chem. Rev.* **2003**, *103* (2), 441–452.
- (106) Sidransky, D.; Irizarry, R.; Califano, J. A.; Li, X.; Ren, H.; Benoit, N.; Mao, L. Serum Protein MALDI Profiling to Distinguish Upper Aerodigestive Tract Cancer Patients From Control Subjects. *JNCI J. Natl. Cancer Inst.* **2003**, *95* (22), 1711–1717.

- (107) Darie-Ion, L.; Whitham, D.; Jayathirtha, M.; Rai, Y.; Neagu, A. N.; Darie, C. C.; Petre, B. A. Applications of MALDI-MS/MS-Based Proteomics in Biomedical Research. *Mol.* **2022**, *Vol. 27, Page 6196* **2022**, *27* (19), 6196.
- (108) Söderlund-Strand, A.; Dillner, J.; Carlson, J. High-Throughput Genotyping of Oncogenic Human Papilloma Viruses with MALDI-TOF Mass Spectrometry. *Clin. Chem.* **2008**, *54* (1), 86–92.
- (109) Majchrzykiewicz-Koehorst, J. A.; Heikens, E.; Trip, H.; Hulst, A. G.; de Jong, A. L.; Viveen, M. C.; Sedee, N. J. A.; van der Plas, J.; Coenjaerts, F. E. J.; Paauw, A. Rapid and Generic Identification of Influenza A and Other Respiratory Viruses with Mass Spectrometry. *J. Virol. Methods* **2015**, *213*, 75–83.
- (110) Wu, Y.; Wang, J.; Luo, C.; Hu, S.; Lin, X.; Anderson, A. E.; Bruera, E.; Yang, X.; Wei, S.; Qian, Y. A Comparison of Burnout Frequency Among Oncology Physicians and Nurses Working on the Frontline and Usual Wards During the COVID-19 Epidemic in Wuhan, China. *J. Pain Symptom Manage.* **2020**, *60* (1), e60–e65.
- (111) Li, G.; He, X.; Zhang, L.; Ran, Q.; Wang, J.; Xiong, A.; Wu, D.; Chen, F.; Sun, J.; Chang, C. Assessing ACE2 Expression Patterns in Lung Tissues in the Pathogenesis of COVID-19. *J. Autoimmun.* **2020**, *112*, 102463.
- (112) Jebiril, N. World Health Organization Declared a Pandemic Public Health Menace: A Systematic Review of the Coronavirus Disease 2019 “COVID-19.” *SSRN Electron. J.* **2020**, *24* (9), 2784–2795.
- (113) Fessell, D.; Cherniss, C. Coronavirus Disease 2019 (COVID-19) and Beyond: Micropractices for Burnout Prevention and Emotional Wellness. *J. Am. Coll. Radiol.* **2020**, *17* (6), 746–748.
- (114) Konermann, L.; Ahadi, E.; Rodriguez, A. D.; Vahidi, S. Unraveling the Mechanism of Electrospray Ionization. *Anal. Chem.* **2013**, *85* (1), 2–9.
- (115) Wang, M.; Wang, C.; Han, R. H.; Han, X. Novel Advances in Shotgun Lipidomics for Biology and Medicine. *Prog. Lipid Res.* **2016**, *61*, 83–108.
- (116) Hu, C.; van der Heijden, R.; Wang, M.; van der Greef, J.; Hankemeier, T.; Xu, G. Analytical Strategies in Lipidomics and Applications in Disease Biomarker Discovery. *J. Chromatogr. B* **2009**, *877* (26), 2836–2846.
- (117) Wang, Y. Applications of Lipidomics in Tumor Diagnosis and Therapy. *Adv. Exp. Med. Biol.* **2021**, *1316*, 25–39.
- (118) Peña-bautista, C.; Álvarez, L.; Durand, T.; Vigor, C.; Cuevas, A.; Baquero, M.; Vento, M.; Hervás, D.; Cháfer-pericás, C. Clinical Utility of Plasma Lipid Peroxidation

Biomarkers in Alzheimer's Disease Differential Diagnosis. *Antioxidants* 2020, Vol. 9, Page 649 **2020**, 9 (8), 649.

- (119) B, A.; Waseem, M.; Kashif, M.; Srinivasan, H. Lipidomics: An Excellent Tool for Chronic Disease Detection. *Curr. Res. Transl. Med.* **2022**, 70 (4), 103346.
- (120) Zhou, G.; Lu, J.; Xu, T.; Lu, Y.; Chen, W.; Wang, J.; Ke, M.; Shen, Q.; Zhu, Y.; Shan, J.; Liu, S. Clinical Lipidomics Analysis Reveals Biomarkers of Lipid Peroxidation in Serum from Patients with Rheumatoid Arthritis. *Microchem. J.* **2021**, 169, 106607.
- (121) Want, E. J.; Nordstro, A.; Morita, H.; Siuzdak, G. From Exogenous to Endogenous: The Inevitable Imprint of Mass Spectrometry in Metabolomics. **2007**.
- (122) Zhao, N. N.; Sun, Y. F.; Zong, L.; Liu, S.; Song, F. R.; Liu, Z. Q.; Liu, S. Y. Serum Lipidomics Study of Ding-Zhi-Xiao-Wan Effect on Alzheimer's Disease Using Online Liquid Extraction Surface Analysis Coupled to Direct Infusion Mass Spectrometry. *Int. J. Mass Spectrom.* **2018**, 434, 29–36.
- (123) Haijes, H. A.; Willemsen, M.; van der Ham, M.; Gerrits, J.; Pras-Raves, M. L.; Prinsen, H. C. M. T.; van Hasselt, P. M.; de Sain-Van der Velden, M. G. M.; Verhoeven-Duif, N. M.; Jans, J. J. M. Direct Infusion Based Metabolomics Identifies Metabolic Disease in Patients' Dried Blood Spots and Plasma. *Metabolites* **2019**, 9 (1).
- (124) Beatriz, A.; Lacerda Jr, V. *Fundamentos de Espectrometria e Aplicações*; Atheneu, 2017.
- (125) Sousa, C.; Botelho, J.; Silva, L.; Grosso, F.; Nemeč, A.; Lopes, J.; Peixe, L. MALDI-TOF MS and Chemometric Based Identification of the Acinetobacter Calcoaceticus-Acinetobacter Baumannii Complex Species. *Int. J. Med. Microbiol.* **2014**, 304 (5–6), 669–677.
- (126) Raftery, D. Mass Spectrometry in Metabolomics: Methods and Protocols. *Methods Mol. Biol.* **2014**, 1198, 333–353.
- (127) Gowda, G. A. N.; Zhang, S.; Gu, H.; Asiago, V.; Shanaiah, N.; Raftery, D. Metabolomics-Based Methods for Early Disease Diagnostics: A. *Expert Rev. Mol. Diagn.* **2008**, 8 (5), 617.
- (128) De Francisco, T. M. G.; Zaramella, I. F.; Ao, J.; Gasparetto, C.; Letícia, L.; Cerqueira, B.; Piantavini, M. S.; Pontarolo, R.; Campos, F. R. Rapid Detection of Aspergillosis in Immunocompromised Patients Using DIMS and Chemometric Analysis. **2015**.
- (129) Galli, M.; Zoppis, I.; Smith, A.; Magni, F.; Mauri, G. Machine Learning Approaches in

- MALDI-MSI: Clinical Applications. *Expert Rev. Proteomics* **2016**, *13* (7), 685–696.
- (130) Luts, J.; Ojeda, F.; Van de Plas Raf, R.; De Moor, B.; Van Huffel, S.; Suykens, J. A. K. A Tutorial on Support Vector Machine-Based Methods for Classification Problems in Chemometrics. *Anal. Chim. Acta* **2010**, *665* (2), 129–145.
- (131) Tran, N. K.; Howard, T.; Walsh, R.; Pepper, J.; Loegering, J.; Phinney, B.; Salemi, M. R.; Rashidi, H. H. Novel Application of Automated Machine Learning with MALDI-TOF-MS for Rapid High-Throughput Screening of COVID-19: A Proof of Concept. *Sci. Reports* **2021**, *11* (1), 1–10.
- (132) Rocca, M. F.; Zintgraff, J. C.; Dattero, M. E.; Santos, L. S.; Ledesma, M.; Vay, C.; Prieto, M.; Benedetti, E.; Avaro, M.; Russo, M.; Nachtigall, F. M.; Baumeister, E. A Combined Approach of MALDI-TOF Mass Spectrometry and Multivariate Analysis as a Potential Tool for the Detection of SARS-CoV-2 Virus in Nasopharyngeal Swabs. *J. Virol. Methods* **2020**, *286*, 113991.
- (133) Corman, V. M.; Landt, O.; Kaiser, M.; Molenkamp, R.; Meijer, A.; Chu, D. K. W.; Bleicker, T.; Brünink, S.; Schneider, J.; Schmidt, M. L.; Mulders, D. G. J. C.; Haagmans, B. L.; Van Der Veer, B.; Van Den Brink, S.; Wijsman, L.; Goderski, G.; Romette, J. L.; Ellis, J.; Zambon, M.; Peiris, M.; Goossens, H.; Reusken, C.; Koopmans, M. P. G.; Drosten, C. Detection of 2019 Novel Coronavirus (2019-NCoV) by Real-Time RT-PCR. *Eurosurveillance* **2020**, *25* (3), 23–30.
- (134) Laboratory Testing for Coronavirus Disease (COVID-19) in Suspected Human Cases.
- (135) Yuki, K.; Fujiogi, M.; Koutsogiannaki, S. COVID-19 Pathophysiology: A Review. *Clin. Immunol.* **2020**, *215*, 108427.
- (136) Phillips, N. The Coronavirus Is Here to Stay - Here's What That Means. *Nature* **2021**, *590* (7846), 382–384.
- (137) Baiges-Gaya, G.; Iftimie, S.; Castañé, H.; Rodríguez-Tomás, E.; Jiménez-Franco, A.; López-Azcona, A. F.; Castro, A.; Camps, J.; Joven, J. Combining Semi-Targeted Metabolomics and Machine Learning to Identify Metabolic Alterations in the Serum and Urine of Hospitalized Patients with COVID-19. *Biomolecules* **2023**, *13* (1), 163.
- (138) Krokhin, O.; Li, Y.; Andonov, A.; Feldmann, H.; Flick, R.; Jones, S.; Stroehrer, U.; Bastien, N.; Dasuri, K. V. N.; Cheng, K.; Simonsen, J. N.; Perreault, H.; Wilkins, J.; Ens, W.; Plummer, F.; Standing, K. G. Mass Spectrometric Characterization of Proteins from the SARS Virus: A Preliminary Report. *Mol. Cell. Proteomics* **2003**, *2* (5), 346–356.

- (139) Marra, M. A.; Jones, S. J. M.; Astell, C. R.; Holt, R. A.; Brooks-Wilson, A.; Butterfield, Y. S. N.; Khattra, J.; Asano, J. K.; Barber, S. A.; Chan, S. Y.; Cloutier, A.; Coughlin, S. M.; Freeman, D.; Girn, N.; Griffith, O. L.; Leach, S. R.; Mayo, M.; McDonald, H.; Montgomery, S. B.; Pandoh, P. K.; Petrescu, A. S.; Robertson, A. G.; Schein, J. E.; Siddiqui, A.; Smailus, D. E.; Stott, J. M.; Yang, G. S.; Plummer, F.; Andonov, A.; Artsob, H.; Bastien, N.; Bernard, K.; Booth, T. F.; Bowness, D.; Czub, M.; Drebot, M.; Fernando, L.; Flick, R.; Garbutt, M.; Gray, M.; Grolla, A.; Jones, S. J. M.; Feldmann, H.; Meyers, A.; Kabani, A.; Li, Y.; Normand, S.; Stroher, U.; Tipples, G. A.; Tyler, S.; Vogrig, R.; Ward, D.; Watson, B.; Brunham, R. C.; Kraiden, M.; Petric, M.; Skowronski, D. M.; Upton, C.; Roper, R. L. The Genome Sequence of the SARS-Associated Coronavirus. *Science (80-)*. **2003**, *300* (5624), 1399–1404.
- (140) Zhang, L.; Guo, H. Biomarkers of COVID-19 and Technologies to Combat SARS-CoV-2. *Adv. Biomark. Sci. Technol.* **2020**, *2*, 1–23.
- (141) Wong, T. F.; So, P. K.; Yao, Z. P. Advances in Rapid Detection of SARS-CoV-2 by Mass Spectrometry. *TrAC Trends Anal. Chem.* **2022**, *157*, 116759.
- (142) Spielmann, N.; Wong, D. Saliva: Diagnostics and Therapeutic Perspectives. *Oral Dis.* **2011**, *17* (4), 345–354.
- (143) de Oliveira, M. A. L.; Watanabe, A. S. A.; Cesar, D. E.; Candido, J. M. B.; Lima, N. M.; Moreira, O. B. O.; Chellini, P. R. TESTES DIAGNÓSTICOS PARA O SARS-COV-2: UMA REFLEXÃO CRÍTICA. *Quim. Nova* **2022**, *45* (6), 760–766.
- (144) McArdle, A.; Washington, K. E.; Chazarin Orgel, B.; Binek, A.; Manalo, D. M.; Rivas, A.; Ayres, M.; Pandey, R.; Phebus, C.; Raedschelders, K.; Fert-Bober, J.; Van Eyk, J. E. Discovery Proteomics for COVID-19: Where We Are Now. *J. Proteome Res.* **2021**, *20* (10), 4627–4639.
- (145) Preianò, M.; Correnti, S.; Pelaia, C.; Savino, R.; Terracciano, R. MALDI MS-Based Investigations for SARS-CoV-2 Detection. *Biochem 2021, Vol. 1, Pages 250-278* **2021**, *1* (3), 250–278.
- (146) Hasan, M. R.; Suleiman, M.; Pérez-López, A. Metabolomics in the Diagnosis and Prognosis of COVID-19. *Front. Genet.* **2021**, *12*, 1358.
- (147) Rezaei, A.; Neshat, S.; Heshmat-Ghahdarjani, K. Alterations of Lipid Profile in COVID-19: A Narrative Review. *Curr. Probl. Cardiol.* **2022**, *47* (3), 100907.
- (148) Mahat, R. K.; Rathore, V.; Singh, N.; Singh, N.; Singh, S. K.; Shah, R. K.; Garg, C. Lipid Profile as an Indicator of COVID-19 Severity: A Systematic Review and Meta-Analysis. *Clin. Nutr. ESPEN* **2021**, *45*, 91–101.

- (149) Barman, H. A.; Pala, A. S.; Dogan, O.; Atıcı, A.; Yumuk, M. T.; Alici, G.; Sit, O.; Gungor, B.; Dogan, S. M. Prognostic Significance of Temporal Changes of Lipid Profile in COVID-19 Patients. *Obes. Med.* **2021**, *28*, 100373.
- (150) Bai, Y.; Huang, W.; Li, Y.; Lai, C.; Huang, S.; Wang, G.; He, Y.; Hu, L.; Chen, C. Lipidomic Alteration of Plasma in Cured COVID-19 Patients Using Ultra High-Performance Liquid Chromatography with High-Resolution Mass Spectrometry. *Biosci. Rep.* **2021**, *41* (3), 20204305.
- (151) Garza, K. Y.; Silva, A. A. R.; Rosa, J. R.; Keating, M. F.; Povilaitis, S. C.; Spradlin, M.; Sanches, P. H. G.; Varão Moura, A.; Marrero Gutierrez, J.; Lin, J. Q.; Zhang, J.; DeHoog, R. J.; Bensussan, A.; Badal, S.; Cardoso de Oliveira, D.; Dias Garcia, P. H.; Dias de Oliveira Negrini, L.; Antonio, M. A.; Canevari, T. C.; Eberlin, M. N.; Tibshirani, R.; Eberlin, L. S.; Porcari, A. M. Rapid Screening of COVID-19 Directly from Clinical Nasopharyngeal Swabs Using the MasSpec Pen. *Anal. Chem.* **2021**, *93* (37), 12582–12593.
- (152) Bernardo, R. A.; Roque, J. V.; de Oliveira Júnior, C. I.; Lima, N. M.; Machado, L. S.; Duarte, G. R. M.; Costa, N. L.; Sorgi, C. A.; Soares, F. F. L.; Vaz, B. G.; Chaves, A. R. Exploring Salivary Lipid Profile Changes in COVID-19 Patients: Insights from Mass Spectrometry Analysis. *Talanta* **2024**, *269*, 125522.
- (153) De Silva, I. W.; Nayek, S.; Singh, V.; Reddy, J.; Granger, J. K.; Verbeck, G. F. Paper Spray Mass Spectrometry Utilizing Teslin® Substrate for Rapid Detection of Lipid Metabolite Changes during COVID-19 Infection. *Analyst* **2020**, *145* (17), 5725–5732.
- (154) Shen, B.; Yi, X.; Sun, Y.; Bi, X.; Du, J.; Zhang, C.; Quan, S.; Zhang, F.; Sun, R.; Qian, L.; Ge, W.; Liu, W.; Liang, S.; Chen, H.; Zhang, Y.; Li, J.; Xu, J.; He, Z.; Chen, B.; Wang, J.; Yan, H.; Zheng, Y.; Wang, D.; Zhu, J.; Kong, Z.; Kang, Z.; Liang, X.; Ding, X.; Ruan, G.; Xiang, N.; Cai, X.; Gao, H.; Li, L.; Li, S.; Xiao, Q.; Lu, T.; Zhu, Y.; Liu, H.; Chen, H.; Guo, T. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **2020**, *182* (1), 59-72.e15.
- (155) Nachtigall, F. M.; Pereira, A.; Trofymchuk, O. S.; Santos, L. S. Detection of SARS-CoV-2 in Nasal Swabs Using MALDI-MS. *Nat. Biotechnol.* **2020**, *38* (10), 1168–1173.
- (156) Deulofeu, M.; García-Cuesta, E.; Peña-Méndez, E. M.; Conde, J. E.; Jiménez-Romero, O.; Verdú, E.; Serrando, M. T.; Salvadó, V.; Boadas-Vaello, P. Detection of SARS-CoV-2 Infection in Human Nasopharyngeal Samples by Combining MALDI-TOF MS and Artificial Intelligence. *Front. Med.* **2021**, *8*, 398–410.

- (157) Rashidi, H. H.; Pepper, J.; Howard, T.; Klein, K.; May, L.; Albahra, S.; Phinney, B.; Salemi, M. R.; Tran, N. K. Comparative Performance of Two Automated Machine Learning Platforms for COVID-19 Detection by MALDI-TOF-MS. *PLoS One* **2022**, *17* (7).
- (158) Costa, M. M.; Benoit, N.; Saby, F.; Pradines, B.; Granjeaud, S.; Almeras, L. Optimization and Standardization of Human Saliva Collection for MALDI-TOF MS. *Diagnostics* **2021**, *11* (8), 1304–1319.
- (159) Chivte, P.; LaCasse, Z.; Seethi, V. D. R.; Bharti, P.; Bland, J.; Kadkol, S. S.; Gaillard, E. R. MALDI-ToF Protein Profiling as a Potential Rapid Diagnostic Platform for COVID-19. *J. Mass Spectrom. Adv. Clin. Lab* **2021**, *21*, 31–41.
- (160) Lazari, L. C.; de Rose Ghilardi, F.; Rosa-Fernandes, L.; Assis, D. M.; Nicolau, J. C.; Santiago, V. F.; Dalçóquio, T. F.; Angeli, C. B.; Bertolin, A. J.; Marinho, C. R. F.; Wrenger, C.; Durigon, E. L.; Siciliano, R. F.; Palmisano, G. Prognostic Accuracy of MALDI-TOF Mass Spectrometric Analysis of Plasma in COVID-19. *Life Sci. Alliance* **2021**, *4* (8), 1–12.
- (161) Wan, Q.; Chen, M.; Zhang, Z.; Yuan, Y.; Wang, H.; Hao, Y.; Nie, W.; Wu, L.; Chen, S. Machine Learning of Serum Metabolic Patterns Encodes Asymptomatic SARS-CoV-2 Infection. *Front. Chem.* **2021**, *9*, 746134.
- (162) Zhang, W.; Li, D.; Xu, B.; Xu, L.; Lyu, Q.; Liu, X.; Li, Z.; Zhang, J.; Sun, W.; Ma, Q.; Qiao, L.; Liao, P. Serum Peptidome Profiles Immune Response of COVID-19 Vaccine Administration. *Front. Immunol.* **2022**, *13*, 956369.
- (163) Gomila, R. M.; Martorell, G.; Fraile-Ribot, P. A.; Doménech-Sánchez, A.; Albertí, M.; Oliver, A.; García-Gasalla, M.; Albertí, S. Use of Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry Analysis of Serum Peptidome to Classify and Predict Coronavirus Disease 2019 Severity. *Open Forum Infect. Dis.* **2021**, *8* (6), 1–8.
- (164) Bennet, S.; Kaufmann, M.; Takami, K.; Sjaarda, C.; Douchant, K.; Moslinger, E.; Wong, H.; Reed, D. E.; Ellis, A. K.; Vanner, S.; Colautti, R. I.; Prameet, & Sheth, M.; Sheth, P. M. Small-Molecule Metabolome Identifies Potential Therapeutic Targets against COVID-19. *Sci. Reports | 123AD*, *12*, 10029.
- (165) Tsai, H.; Phinney, B. S.; Grigorean, G.; Salemi, M. R.; Rashidi, H. H.; Pepper, J.; Tran, N. K. Identification of Endogenous Peptides in Nasal Swab Transport Media Used in MALDI-TOF-MS Based COVID-19 Screening. *ACS Omega* **2022**.
- (166) Renuse, S.; Vanderboom, P. M.; Maus, A. D.; Kemp, J. V.; Gurtner, K. M.;

- Madugundu, A. K.; Chavan, S.; Peterson, J. A.; Madden, B. J.; Mangalaparthy, K. K.; Mun, D. G.; Singh, S.; Kipp, B. R.; Dasari, S.; Singh, R. J.; Grebe, S. K.; Pandey, A. A Mass Spectrometry-Based Targeted Assay for Detection of SARS-CoV-2 Antigen from Clinical Specimens. *EBioMedicine* **2021**, *69*.
- (167) Pozzi, C.; Levi, R.; Braga, D.; Carli, F.; Darwich, A.; Spadoni, I.; Oresta, B.; Dioguardi, C. C.; Peano, C.; Ubaldi, L.; Angelotti, G.; Bottazzi, B.; Garlanda, C.; Desai, A.; Voza, A.; Azzolini, E.; Cecconi, M.; Mantovani, A.; Penna, G.; Barbieri, R.; Politi, L. S.; Rescigno, M.; Alessio, A.; Clement, A.; Salvatore, B.; Cristina, B.; Alice, B.; Sara, B.; Paola, B.; Francesca, C.; Michela, C.; Assunta, C.; Arianna, C.; Claudia, C.; Sara, C.; Silvia, C.; Valentina, C.; Maurizio, C.; Michele, C.; Nicolò, C.; Abbass, D.; de Nalda Ana, L.; Federica, D. P.; Rachele, D. D.; Elisabeth, D.; Barbara, D.; Maria, F. F.; Valentina, F.; Giulia, F.; Sara, F.; Antonio, G. G.; Silvia, G.; Rita, G. A.; Michela, L.; Antonino, L. C.; Alessia, M.; Alessandro, M.; Ilaria, M.; Bianca, O.; Fabio, P.; Anna, P.; Erica, P.; Chiara, P.; Chiara, P.; Valeria, R.; Monica, R.; Alice, S.; Carlo, S.; Alessandra, S.; Marina, S.; Ilaria, S.; Salvatore, S.; Gianmarco, S.; Domenico, S.; Paolo, T.; Aldo, U.; Sonia, V.; Antonio, V.; Elisa, Z.; Veronica, Z. A 'Multiomic' Approach of Saliva Metabolomics, Microbiota, and Serum Biomarkers to Assess the Need of Hospitalization in Coronavirus Disease 2019. *Gastro Hep Adv.* **2022**, *1* (2), 194–209.
- (168) Zeng, H. L.; Chen, D.; Yan, J.; Yang, Q.; Han, Q. Q.; Li, S. S.; Cheng, L. Proteomic Characteristics of Bronchoalveolar Lavage Fluid in Critical COVID-19 Patients. *FEBS J.* **2021**, *288* (17), 5190–5200.
- (169) Beltrán-Camacho, L.; Eslava-Alcón, S.; Rojas-Torres, M.; Sánchez-Morillo, D.; Martínez-Nicolás, M. ^aP; Martín-Bermejo, V.; de la Torre, I. G.; Berrocoso, E.; Moreno, J. A.; Moreno-Luna, R.; Durán-Ruiz, M. ^aC. The Serum of COVID-19 Asymptomatic Patients up-Regulates Proteins Related to Endothelial Dysfunction and Viral Response in Circulating Angiogenic Cells Ex-Vivo. *Mol. Med.* **2022**, *28* (1), 1–16.
- (170) Castañé, H.; Iftimie, S.; Baiges-Gaya, G.; Rodríguez-Tomás, E.; Jiménez-Franco, A.; López-Azcona, A. F.; Garrido, P.; Castro, A.; Camps, J.; Joven, J. Machine Learning and Semi-Targeted Lipidomics Identify Distinct Serum Lipid Signatures in Hospitalized COVID-19-Positive and COVID-19-Negative Patients. *Metabolism* **2022**, *131*, 155197.
- (171) Zhang, Y.; Cai, X.; Ge, W.; Wang, D.; Zhu, G.; Qian, L.; Xiang, N.; Yue, L.; Liang, S.;

- Zhang, F.; Wang, J.; Zhou, K.; Zheng, Y.; Lin, M.; Sun, T.; Lu, R.; Zhang, C.; Xu, L.; Sun, Y.; Zhou, X.; Yu, J.; Lyu, M.; Shen, B.; Zhu, H.; Xu, J.; Zhu, Y.; Guo, T. Potential Use of Serum Proteomics for Monitoring COVID-19 Progression to Complement RT-PCR Detection. *J. Proteome Res.* **2022**, *21* (1), 90–100.
- (172) Gutmann, C.; Takov, K.; Burnap, S. A.; Singh, B.; Ali, H.; Theofilatos, K.; Reed, E.; Hasman, M.; Nabeebaccus, A.; Fish, M.; McPhail, M. J.; O’Gallagher, K.; Schmidt, L. E.; Cassel, C.; Rienks, M.; Yin, X.; Auzinger, G.; Napoli, S.; Mujib, S. F.; Trovato, F.; Sanderson, B.; Merrick, B.; Niazi, U.; Saqi, M.; Dimitrakopoulou, K.; Fernández-Leiro, R.; Braun, S.; Kronstein-Wiedemann, R.; Doores, K. J.; Edgeworth, J. D.; Shah, A. M.; Bornstein, S. R.; Tonn, T.; Hayday, A. C.; Giacca, M.; Shankar-Hari, M.; Mayr, M. SARS-CoV-2 RNAemia and Proteomic Trajectories Inform Prognostication in COVID-19 Patients Admitted to Intensive Care. *Nat. Commun.* **2021**, *12* (1), 1–17.
- (173) Ciccarelli, M.; Merciai, F.; Carrizzo, A.; Sommella, E.; Di Pietro, P.; Caponigro, V.; Salviati, E.; Musella, S.; Sarno, V. di; Rusciano, M.; Toni, A. L.; Iesu, P.; Izzo, C.; Schettino, G.; Conti, V.; Venturini, E.; Vitale, C.; Scarpati, G.; Bonadies, D.; Rispoli, A.; Polverino, B.; Poto, S.; Pagliano, P.; Piazza, O.; Licastro, D.; Vecchione, C.; Campiglia, P. Untargeted Lipidomics Reveals Specific Lipid Profiles in COVID-19 Patients with Different Severity from Campania Region (Italy). *J. Pharm. Biomed. Anal.* **2022**, *217*, 114827.
- (174) Suvarna, K.; Biswas, D.; Pai, M. G. J.; Acharjee, A.; Bankar, R.; Palanivel, V.; Salkar, A.; Verma, A.; Mukherjee, A.; Choudhury, M.; Ghantasala, S.; Ghosh, S.; Singh, A.; Banerjee, A.; Badaya, A.; Bihani, S.; Loya, G.; Mantri, K.; Burli, A.; Roy, J.; Srivastava, A.; Agrawal, S.; Shrivastav, O.; Shastri, J.; Srivastava, S. Proteomics and Machine Learning Approaches Reveal a Set of Prognostic Markers for COVID-19 Severity With Drug Repurposing Potential. *Front. Physiol.* **2021**, *12*, 432.
- (175) Wang, Z.; Cryar, A.; Lemke, O.; Tober-Lau, P.; Ludwig, D.; Helbig, E. T.; Hippenstiel, S.; Sander, L. E.; Blake, D.; Lane, C. S.; Sayers, R. L.; Mueller, C.; Zeiser, J.; Townsend, S. J.; Demichev, V.; Mülleder, M.; Kurth, F.; Sirka, E.; Hartl, J.; Ralser, M. A Multiplex Protein Panel Assay for Severity Prediction and Outcome Prognosis in Patients with COVID-19: An Observational Multi-Cohort Study. *eClinicalMedicine* **2022**, *49*, 101495.
- (176) Demichev, V.; Tober-Lau, P.; Lemke, O.; Nazarenko, T.; Thibeault, C.; Whitwell, H.; Röhl, A.; Freiwald, A.; Szyrwił, L.; Ludwig, D.; Correia-Melo, C.; Aulakh, S. K.;

Helbig, E. T.; Stubbemann, P.; Lippert, L. J.; Grüning, N. M.; Blyuss, O.; Vernardis, S.; White, M.; Messner, C. B.; Joannidis, M.; Sonnweber, T.; Klein, S. J.; Pizzini, A.; Wohlfarter, Y.; Sahanic, S.; Hilbe, R.; Schaefer, B.; Wagner, S.; Mittermaier, M.; Machleidt, F.; Garcia, C.; Ruwwe-Glösenkamp, C.; Lingscheid, T.; Bosquillon de Jarcy, L.; Stegemann, M. S.; Pfeiffer, M.; Jürgens, L.; Denker, S.; Zickler, D.; Enghard, P.; Zelezniak, A.; Campbell, A.; Hayward, C.; Porteous, D. J.; Marioni, R. E.; Uhrig, A.; Müller-Redetzky, H.; Zoller, H.; Löffler-Ragg, J.; Keller, M. A.; Tancevski, I.; Timms, J. F.; Zaikin, A.; Hippenstiel, S.; Ramharter, M.; Witzentrath, M.; Suttorp, N.; Lilley, K.; Mülleder, M.; Sander, L. E.; Kleinschmidt, M.; Heim, K. M.; Millet, B.; Meyer-Arndt, L.; Hübner, R. H.; Andermann, T.; Doehn, J. M.; Opitz, B.; Sawitzki, B.; Grund, D.; Radünzel, P.; Schürmann, M.; Zoller, T.; Alius, F.; Knape, P.; Breitbart, A.; Li, Y.; Bremer, F.; Pergantis, P.; Schürmann, D.; Temmesfeld-Wollbrück, B.; Wendisch, D.; Brumhard, S.; Haenel, S. S.; Conrad, C.; Georg, P.; Eckardt, K. U.; Lehner, L.; Kruse, J. M.; Ferse, C.; Körner, R.; Spies, C.; Edel, A.; Weber-Carstens, S.; Krannich, A.; Zvorc, S.; Li, L.; Behrens, U.; Schmidt, S.; Rönnefarth, M.; Dang-Heine, C.; Röhle, R.; Lieker, E.; Kretzler, L.; Wirsching, I.; Wollboldt, C.; Wu, Y.; Schwanitz, G.; Hillus, D.; Kasper, S.; Olk, N.; Horn, A.; Briesemeister, D.; Treue, D.; Hummel, M.; Corman, V. M.; Drost, C.; von Kalle, C.; Ralser, M.; Kurth, F. A Time-Resolved Proteomic and Prognostic Map of COVID-19. *Cell Syst.* **2021**, *12* (8), 780-794.e7.

- (177) D'alessandro, A.; Thomas, T.; Akpan, I. J.; Reisz, J. A.; Cendali, F. I.; Gamboni, F.; Nemkov, T.; Thangaraju, K.; Katneni, U.; Tanaka, K.; Kahn, S.; Wei, A. Z.; Valk, J. E.; Hudson, K. E.; Roh, D.; Moriconi, C.; Zimring, J. C.; Hod, E. A.; Spitalnik, S. L.; Buehler, P. W.; Francis, R. O. Biological and Clinical Factors Contributing to the Metabolic Heterogeneity of Hospitalized Patients with and without Covid-19. *Cells* **2021**, *10* (9), 2293.
- (178) Delafiori, J.; Navarro, L. C.; Siciliano, R. F.; De Melo, G. C.; Busanello, E. N. B.; Nicolau, J. C.; Sales, G. M.; De Oliveira, A. N.; Val, F. F. A.; De Oliveira, D. N.; Eguti, A.; Dos Santos, L. A.; Dalçóquio, T. F.; Bertolin, A. J.; Abreu-Netto, R. L.; Salsoso, R.; Baía-Da-Silva, D.; Marcondes-Braga, F. G.; Sampaio, V. S.; Judice, C. C.; Costa, F. T. M.; Durán, N.; Perroud, M. W.; Sabino, E. C.; Lacerda, M. V. G.; Reis, L. O.; Fávaro, W. J.; Monteiro, W. M.; Rocha, A. R.; Catharino, R. R. Covid-19 Automated Diagnosis and Risk Assessment through Metabolomics and Machine Learning. *Anal. Chem.* **2021**, *93* (4), 2471–2479.

- (179) Spick, M.; Lewis, H. M.; Frampas, C. F.; Longman, K.; Costa, C.; Stewart, A.; Dunn-Walters, D.; Greener, D.; Evetts, G.; Wilde, M. J.; Sinclair, E.; Barran, P. E.; Skene, D. J.; Bailey, M. J. An Integrated Analysis and Comparison of Serum, Saliva and Sebum for COVID-19 Metabolomics. *Sci. Reports 2022* **12**, 12 (1), 1–12.
- (180) Gray, N.; Lawler, N. G.; Zeng, A. X.; Ryan, M.; Bong, S. H.; Boughton, B. A.; Bizkarguenaga, M.; Bruzzone, C.; Embade, N.; Wist, J.; Holmes, E.; Millet, O.; Nicholson, J. K.; Whiley, L. Diagnostic Potential of the Plasma Lipidome in Infectious Disease: Application to Acute SARS-CoV-2 Infection. *Metab. 2021*, Vol. 11, Page 467 **2021**, 11 (7), 467.
- (181) Seethi, V. D. R.; LaCasse, Z.; Chivte, P.; Bland, J.; Kadkol, S. S.; Gaillard, E. R.; Bharti, P.; Alhoori, H. An Explainable AI Approach for Diagnosis of COVID-19 Using MALDI-ToF Mass Spectrometry. *Expert Syst. Appl.* **2024**, 236, 121226.
- (182) Kugler, S.; Hahnefeld, L.; Kloka, J. A.; Ginzel, S.; Nürenberg-Goloub, E.; Zinn, S.; Vehreschild, M. J.; Zacharowski, K.; Lindau, S.; Ullrich, E.; Burmeister, J.; Kohlhammer, J.; Schwäble, J.; Gurke, R.; Dorochow, E.; Bennett, A.; Dauth, S.; Campe, J.; Knape, T.; Laux, V.; Kannt, A.; Köhm, M.; Geisslinger, G.; Resch, E.; Behrens, F. Short-Term Predictor for COVID-19 Severity from a Longitudinal Multi-Omics Study for Practical Application in Intensive Care Units. *Talanta* **2024**, 268, 125295.

CAPITULO II – MALDI(+) FT-ICR combinada com aprendizagem de máquina para triagem diagnóstica para COVID-19 baseada em amostras de saliva

“MALDI(+) FT-ICR Mass Spectrometry (MS) Combined with Machine Learning toward Saliva-Based Diagnostic Screening for COVID-19”

RESUMO

A identificação rápida de vírus respiratórios existentes e emergentes em amostras biológicas, como SARS-CoV-2, é de extrema importância em estratégias de combate e controle a surtos e pandemias. A técnica MALDI-FT-ICR MS (*matrix assisted laser desorption ionization Fourier-transform ion cyclotron resonance mass spectrometry*) possui alta sensibilidade e especificidade que combinada com aprendizagem de máquina, pode ser promissora na classificação de amostras positivas para a SARS-CoV-2, mesmo que em matrizes complexas e em baixas concentrações. Neste contexto, o trabalho teve como objetivo desenvolver uma metodologia rápida e eficaz para realizar triagens de pacientes com suspeita de COVID-19 a partir de amostras de saliva, utilizando a técnica de MALDI- FT-ICR MS mediante auxílio da aprendizagem por máquina de vetores de suporte (SVM). Para eficiência nas análises MALDI-FT-ICR MS, foi necessário otimizar o preparo de amostras e os parâmetros da análise. A maior eficiência em menor tempo de análise foi obtido com a digestão da saliva em 10 µL de tripsina por de 2 h. Já a otimização dos parâmetros da análise por MALDI, 1M de resolução foi satisfatório para as análises, apresentando um ótimo número de sinais em rápido tempo de análise. Modelos SVM foram criados com os dados das análises de 149 amostras sendo estas 97 positivas e 52 negativas, comprovadas por testes de reação em cadeia da polimerase com transcrição reversa (RT-PCR). Dois modelos apresentaram os melhores resultados, SVM1 e SVM2. O grupo de calibração obteve acurácia de 100% e o grupo de teste de 95,6% (SVM1) e 86,7% (SVM2). O SVM1 selecionou 780 variáveis e possui Taxa de Falso Negativo (TFN) de 0% já o SVM2 selecionou somente 2 variáveis (525,4 Da e 1410,8 Da) com TFN de 3%. Desta forma, a metodologia proposta apresenta-se como promissora ferramenta no auxílio de triagem prévia para diagnóstico de COVID-19.

Palavras-chave: MALDI FT-ICR MS; Machine learning; Saliva; COVID-19.

2.1 INTRODUÇÃO

A doença COVID-19 é uma infecção viral causada pelo coronavírus 2 da síndrome respiratória aguda grave (SARS-CoV-2). A disseminação pode ser decorrente do transporte aéreo de partículas aerossol da respiração normal ou de gotículas, em suspensão no ar oriundas de tosse ou de espirro.¹ A COVID-19 foi declarada como pandemia em 11 de março de 2020, pela Organização Mundial de Saúde (OMS) com mais de 420 000 000 contaminados e 5 800 000 óbitos em todo mundo até março de 2022.²

A definição de métodos de detecção rápida e precoce da COVID-19 é crucial para fornecer um tratamento oportuno, um rastreamento de indivíduos que mantiveram contatos com os pacientes contaminados e interromper as vias de transmissão.³

A reação em cadeia da polimerase de transcrição reversa (RT-PCR), usando material coletado de swab nasofaríngeo é rotineiramente usada e considerada padrão-ouro na detecção da infecção por SARS-CoV-2.^{4,5} A coleta de amostras de swab requer profissionais de saúde qualificados e grande preocupação com manuseio, transporte e armazenamento das amostras. Além do swab nasofaríngeo causar desconforto aos pacientes, também pode causar sangramento e por isso existem várias contraindicações relacionadas as condições clínicas, tais como a coagulopatia ou uso de terapia anticoagulante, e desvio significativo do septo nasal.⁶ Além disso, o teste RT-PCR possui limitações associadas às despesas e ao tempo que precisam ser superadas.^{7,8}

Para melhorar questões como transporte de materiais altamente contaminados, disponibilidade de equipamentos e a necessidade de testes mais rápidos, o uso de testes rápidos de antígenos para COVID-19 foi uma solução para detectar e rastrear rapidamente comunidades suspeitas de contrair COVID-19. Um teste rápido de antígeno é um teste de diagnóstico que detecta a presença do vírus diretamente no momento da análise.⁹ Apesar de atender as necessidades de uma resposta rápida em um programa de triagem em massa, ainda assim é necessário testes com maior sensibilidade e acurácia para o diagnóstico COVID-19.¹⁰

O diagnóstico para COVID-19 ainda encontra limitações para ampla utilização especialmente em países em desenvolvimento, em razão da técnica ouro utilizada e da necessidade de insumos para a aplicação de inúmeros testes na população. Uma alternativa a estes métodos, com alta sensibilidade e especificidade seria a aplicação de ferramentas de diagnóstico baseadas em espectrometria de massas (MS), utilizando a técnica de ionização e dessorção a laser assistida por matriz (MALDI) na identificação de

proteínas específicas de doenças. Recentemente, técnicas de MS foram exploradas para detectar COVID-19 com base na identificação de biomarcadores salivares.¹¹⁻¹³

A saliva é uma secreção produzida pelas glândulas salivares, possui diversas funções como proteção e limpeza da cavidade oral e auxílio na digestão. Além disso, devido à facilidade e segurança de coletar amostras de saliva por meio de técnicas não invasivas, este biofluido pode ser uma fonte alternativa de amostras para obtenção de biomarcadores de doenças.^{14,15} Assim, a COVID-19 poderia ser detectada por meio de testes de diagnóstico salivar, apresentando vantagens para profissionais de saúde e pacientes.¹⁶

A grande quantidade de dados obtidos usando espectros de MS requer a combinação de poderosas estratégias estatísticas para a identificação de padrões para diagnóstico.¹⁷ A tendência atual em associar algoritmos explicados por inteligência artificial e técnicas “ômicas” tem rendido plataformas envolvendo aprendizagem de máquina (ML) para análise de dados de MS, visando a identificação de biomarcadores de doenças, incluindo a COVID-19.¹⁸

Recentemente, Nachtigall *et al.* (2020)¹⁹ relataram o desenvolvimento de um método MALDI MS para o diagnóstico de infecção por SARS-CoV-2. Neste estudo, amostras de swabs nasais foram analisadas diretamente por MALDI-TOF MS e um padrão espectral foi usado para classificar os pacientes infectados ou não, uma aplicação interessante de MS para a detecção de infecções virais.¹⁹

Outro estudo, realizado por Chivte *et al.* (2021),¹² desenvolveu a pesquisa utilizando amostras de saliva e estabelecendo cinco sinais de proteínas virais, potenciais biomarcadores para COVID-19. As análises por MALDI-TOF apresentaram $\geq 90\%$ de concordância com os resultados de RT-PCR, concluindo que o MALDI-TOF pode ser utilizado para desenvolver ensaios para COVID-19 relativamente rápido e barato.¹²

Diante do exposto, o desenvolvimento de uma metodologia utilizando MALDI FT-ICR MS combinado com ML pode fornecer uma resposta rápida e de elevada acurácia para classificar amostras positivas para doenças, favorecendo diagnósticos valiosos que complementam as informações genômicas, permitindo aumentar o conhecimento atual da doença COVID-19.

2.2 METODOLOGIA

2.2.1 Amostras e reagentes

2.2.1.1 Reagentes

RapiGest SF 0,2% (*Waters*, Milford, MA, EUA), Tripsina, 0,05 µg/µL, (*Promega*, Madison, WI, EUA), matriz CHCA (*α-cyano-4-hydroxycinnamic acid*, *Sigma-Aldrich Chemicals*, St. Louis, MO, EUA) 1 mg/mL, 50% ACN (acetonitrila, *Sigma-Aldrich Chemicals*, EUA) : 0,1% TFA, AmBic 50 mM (bicarbonato de amônio, *Sigma-Aldrich Chemicals*, St. Louis, MO, EUA), DTT 100 mM (Ditiotreitól, *Sigma-Aldrich Chemicals*, St. Louis, MO, EUA) em AmBic 50 mM, IAA 300 mM (Iodoacetamida, *Sigma-Aldrich Chemicals*, St. Louis, MO, EUA) em AmBic 50 mM, TFA 5% (ácido trifluoroacético, *Sigma-Aldrich Chemicals*, St. Louis, MO, EUA)

2.2.1.2 Amostras de saliva

149 amostras de saliva (97 com RT-PCR positivo para SARS-CoV-2 e 52 negativas para COVID-19) foram incluídas no estudo. Dentre elas, 86 amostras (34 com RT-PCR positivo para SARS-CoV-2 e 52 negativo) foram disponibilizadas pelo grupo de pesquisa do Programa de Pós-graduação em Ciências Fisiológicas do Centro de Ciências da Saúde (comitê de ética CAAE: 30993920.1.0000.5071) e 63 amostras com RT-PCR positivo para SARS-CoV-2 foram disponibilizadas pelo grupo de pesquisa da faculdade de odontologia da Universidade Federal de Goiás (comitê de ética CAAE: 38088920.9.0000.5083, 38088920.9.3002.5082, 38088920.9.3001.5078).

A coleta de saliva ocorreu nos hospitais das de cidades, Goiânia, Vila Velha e São Mateus, em pacientes que tiveram pelo menos um sintoma característico de COVID-19 e espontaneamente compareceu aos hospitais para realizar o teste RT-PCR para COVID-19. Imediatamente após a coleta do swab orofaríngeo, cada paciente recebeu um tubo estéril para facilitar a autocoleta de saliva (cerca de 1 mL) sob supervisão direta de uma enfermeira.

A Tabela 2. 1 descreve os perfis demográficos dos participantes com gênero e sintomas.

Tabela 2. 1 Dados dos participantes

Sexo	Quantidade		
	Total (n= 149)	Positivo (65%)	Negativo (35%)
Feminino	74	41	33
Masculino	58	44	14
Sem informação	17	12	5
Sintomas	Quantidade		
Tosse	82	55	27
Febre	44	31	13
Coriza	27	25	22
Dor de garganta	34	30	14
Anosmia	32	30	2
Dispneia	37	27	10
Náusea	25	20	5
Vomito	14	14	0
Diarreia	22	18	4
Adinamia	52	44	8
Falta de apetite	30	30	0
Dor de cabeça	64	42	22
Mialgia	34	22	12

2.2.1.3 Preparo de amostra

Para adaptação do protocolo de preparo de amostra utilizou-se em uma amostra de saliva saudável, que teve sua concentração de proteína determinada pelo método Bradford.²⁰ Foi adicionado 10 µL de AmBic 50 mM, 25 µL da solução RapiGest SF 0,2%v/v em 50µL de amostra de saliva e foi agitado em vórtex; o tubo foi transferido para um bloco aquecido (37°C) e incubado por 60 min. Depois foi adicionado 2,5 µL de DTT 100 mM a solução, agitado em vórtex novamente e aquecido (37°C) e incubado por 40 min. Posteriormente, foi adicionado 2,5 µL de IAA 300 mM, agitado em vórtex e o tubo transferido para o escuro e incubado à temperatura ambiente por 30 min. Uma solução de tripsina 0,05 µg/µL foi adicionada e agitada com movimentos “*up and down*”, a digestão ocorreu a 37°C. O volume de tripsina e o tempo de digestão foram otimizados, variando o volume em 5 e 10 µL e o tempo em 2h e 16h (“*overnight*”).

Para término da digestão, foi adicionado 10 µL de TFA 5%v/v, agitado em vórtex e as amostras foram incubados a 37°C por 90 min. Posteriormente, a amostra foi

centrifugada a 14 000 rpm a 6°C, por 30 min. Para a purificação dos peptídeos foi utilizado Pierce™ C18 *Spin Columns* (Thermo Scientific, San Jose, CA, EUA).

2.2.2 Análise por MALDI-FT-ICR MS

As análises MALDI MS foram realizadas com um espectrômetro de massa FT-ICR MS (modelo 9.4 T Solarix, Bruker Daltonics, Bremen, Alemanha), equipado com um laser Smartbeam II Nd:YAG (355 nm) e fonte MALDI.

Dois microlitros de cada amostra foram colocados em triplicata na placa alvo MALDI (Bruker Daltonics) seguido pela adição de 1 µL da solução matriz CHCA e secagem ao ar livre.²¹

Os espectros de massa foram adquiridos no modo positivo de ionização na faixa de massa m/z 200–4000. Os principais parâmetros experimentais utilizados foram: *laser frequency*: 200 Hz; *plate offset*, -100.0 V; *deflector plate*, -210.0 V; *number of laser shots*: 200; *laser power* 35%; *laser focus*, “small”; 16 *scans*, e modo *random walk*.

O tempo gasto na análise e a quantidade de sinais foram avaliados na análise MALDI FT-ICR, utilizando diferentes resolução (512k, 1M, 2M, 4M e 8M).

2.2.3 Análise de dados

O conjunto de dados obtido com o software *DataAnalysis* (Bruker Daltonics, EUA) foi exportado no formato ASC e importado no software MATLAB 9.0 (R2013a). Foram realizados pré-tratamentos como centralização na média, normalização, autoescalar e combinações entre eles.

2.2.3.1 Análise de componentes principais (PCA)

O modelo de reconhecimento de padrão por análise de componentes principais (PCA) foi desenvolvido para identificar possíveis amostras anômalas, visualização de similaridades e possíveis grupamentos naturais entre as amostras, e analisar o comportamento e dispersão das variáveis espectrais. Essa modelagem é dada pela decomposição da matriz **X** de estudo em duas matrizes: *scores* contendo informações da amostra e *loadings* contendo informações das variáveis.

2.2.3.2 *Máquina de vetores de suporte (SVM)*

Os modelos de classificação de SVM foram desenvolvidos para classificar amostras positivas e negativas para SARS-CoV-2. Para cada conjunto de técnicas e propriedades a serem estimadas, o conjunto amostral foi previamente dividido em calibração (70%) e predição (30%) por metodologia Kennard Stone.²² Os parâmetros custo (C), margem (ϵ) e gama (γ) da função kernel foram otimizados.

2.3 RESULTADOS E DISCUSSÃO

A técnica MALDI FT-ICR MS é uma ferramenta muito utilizada para o estudo de compostos biológicos devido à alta sensibilidade e precisão de massa. Sua combinação com ML pode favorecer diagnósticos moleculares, utilizando amostras de coleta econômica, segura, fácil e não invasiva, como a saliva.²³

A saliva é um fluido corporal produzido pelas glândulas salivares, composta de uma variedade de substâncias e um rico reservatório de proteínas e peptídeos; na verdade, reúne mais de 3.652 proteínas e 12.562 peptídeos. Em geral, os componentes proteicos da saliva apresentam moléculas de alto massa molar como, mucinas, lisozima, imunoglobulinas, proteína rica em prolina (PRP) e esterinas.^{14,24} Neste sentido, a digestão de proteínas é uma etapa essencial para as análises por MALDI FT-ICR MS, uma vez que a faixa de m/z trabalhada (200-4000) não identifica proteínas de alta massa molecular.

A identificação de proteínas é baseada em medições de m/z para um conjunto de peptídeos, produtos de digestão da proteína original.²⁵ No entanto, a análise de peptídeos pode apresentar uma série de problemas que afetam a resposta do detector, devido à presença de interferências indesejadas da matriz ou de reagentes e outros aditivos usados para facilitar a digestão de proteínas.²⁶ Dessa forma, uma alternativa para a eliminação desses sinais interferentes, é o uso de colunas spin C18 após a digestão trípica, que removem contaminantes interferentes e liberam peptídeos em soluções compatíveis com MS, resultando em maior sensibilidade e espectro de alta qualidade.²⁷

Após protocolo de digestão da saliva, foi possível observar a diferença espectral entre uma amostra de saliva sem preparo (Figura2. 1A) e outra após a etapa de digestão das proteínas (Figura2. 1B), e fazer uma comparação dos íons detectados pelas análises de MALDI FT-ICR MS, com sinais de alguns peptídeos já identificados em análises de saliva relatados na literatura (Tabela2. 2).

Figura2. 1 Espectros de massas MALDI(+) de amostras de saliva com e sem preparo de amostra: A) antes e B) após digestão trípica e processamento em coluna spin C18.

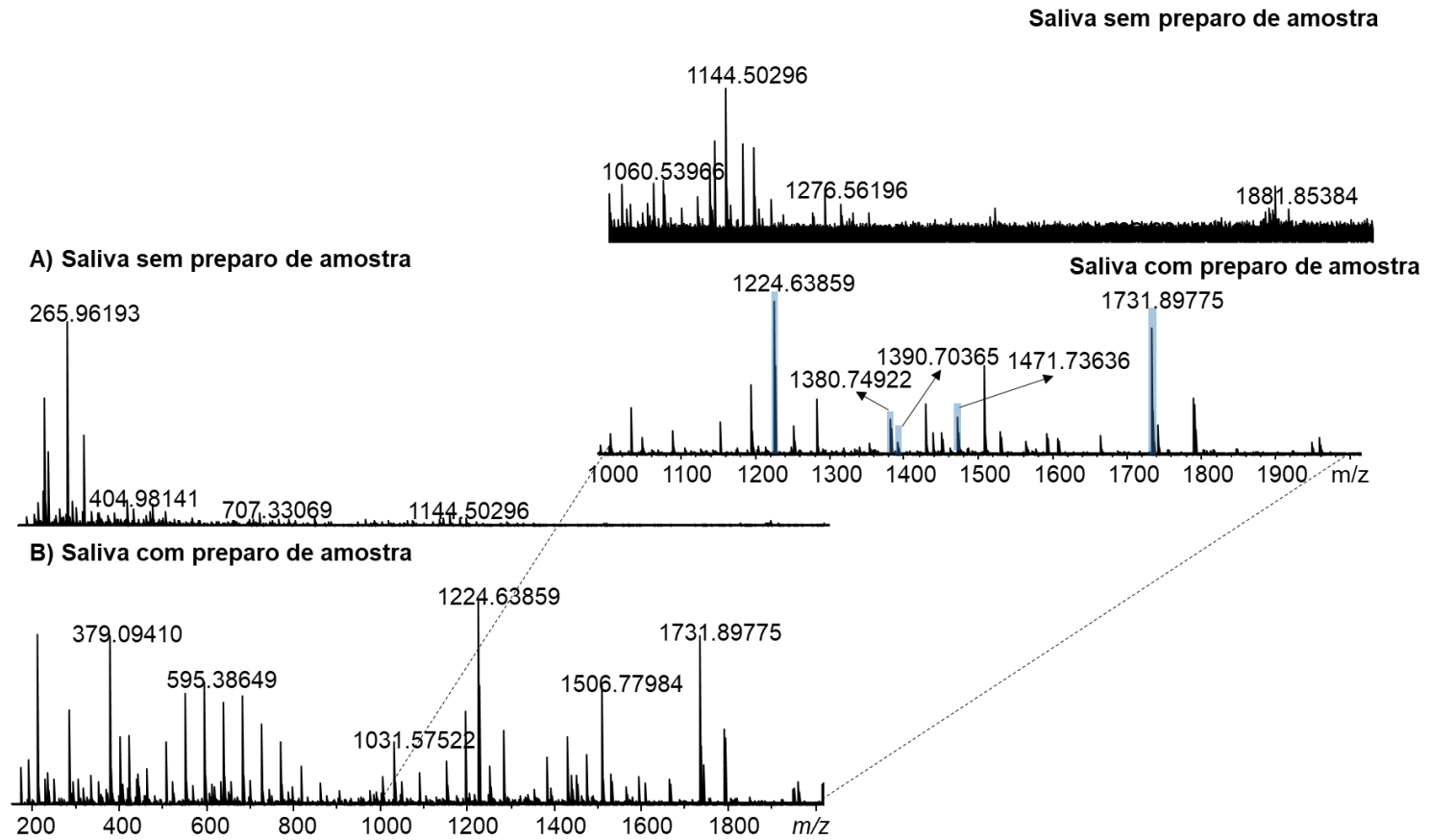


Tabela 2. 2 Peptídeos identificados em amostras de saliva após a digestão de proteínas e SPE.

<i>m/z</i> (M+H ⁺)	TIC (total ion current)	Sequência de peptídeos	Proteína	Ref
1224.63859	4.5 × 10 ⁷	GPPQQGGHPRPP	Fosfoproteína	28
1380.74922	1.0 × 10 ⁷	GPPQQGGHPRPPR	salivar ácida rica em	28
1390.70365	3.6 × 10 ⁶	GGRPQGGPPQGQSPQ	prolina ou	21,29
1471.73636	1.1 × 10 ⁷	GRPQGGPPQQGGHQQ	PRH1Protein	28
1731.89775	3.7 × 10 ⁷	GPPQQGGHPPPPQGRPQ	(P02810)	21

Para a digestão das proteínas na saliva foi utilizado 50 µL da amostra. Na adaptação do protocolo de digestão foram avaliados o volume de tripsina (0,05 µg/µL) e o tempo de digestão, a fim de otimizar e agilizar o processo de preparo de amostra. O volume utilizado foi de 5 e 10µL de tripsina e o tempo de digestão de 2h e 16h (“*overnight*”), totalizando 4 variações no processo de digestão da amostra de saliva.

Os resultados das análises por MALDI(+)-FT-ICR MS após a digestão, não mostraram diferenças espectrais significativas entre as quatro variações no protocolo de digestão (Figura2. 2) na faixa de *m/z* de 1000 a 2200. Comparando os espectros full scan das amostras que utilizaram 10 uL de tripsina para a digestão, esses apresentaram uma maior relação de números de sinais detectados com relação sinal ruído (S/N) > 4, comparados às amostras que utilizaram o volume de 5uL de tripsina para a digestão. As amostras que tiveram sua digestão *overnight* também apresentaram um maior número de sinais, porém o menor tempo (2h), foi satisfatório e agiliza o processo de diagnóstico em 14h no mínimo (Tabela2. 3). Além disso, a amostra com 10uL de tripsina e 2h de digestão apresentou um maior TIC para todas as sequências de peptídeos (Tabela2. 4). Todas as 4 amostras após a digestão apresentaram um espectro diferente do espectro da amostra de saliva sem a digestão (Figura2. 2A), onde foi possível identificar a sequência de peptídeos, representada pelos íons de *m/z* 1224; 1380; 1390; 1471 e 1731 que são derivados de fosfoproteína, rica em prolina ácida salivar (1/2), P02810.^{21,28}

Figura 2. 2 Espectros de massas MALDI(+) da amostra de saliva e A) Sem digestão triptica, (B- E) Após digestão triptica: B) 10 μ L tripsina e 2 h digestão triptica, C) 10 μ L tripsina e 16h digestão triptica, D) 5 μ L tripsina e 2 h digestão triptica, e E) 5 μ L tripsina e 16 h digestão triptica

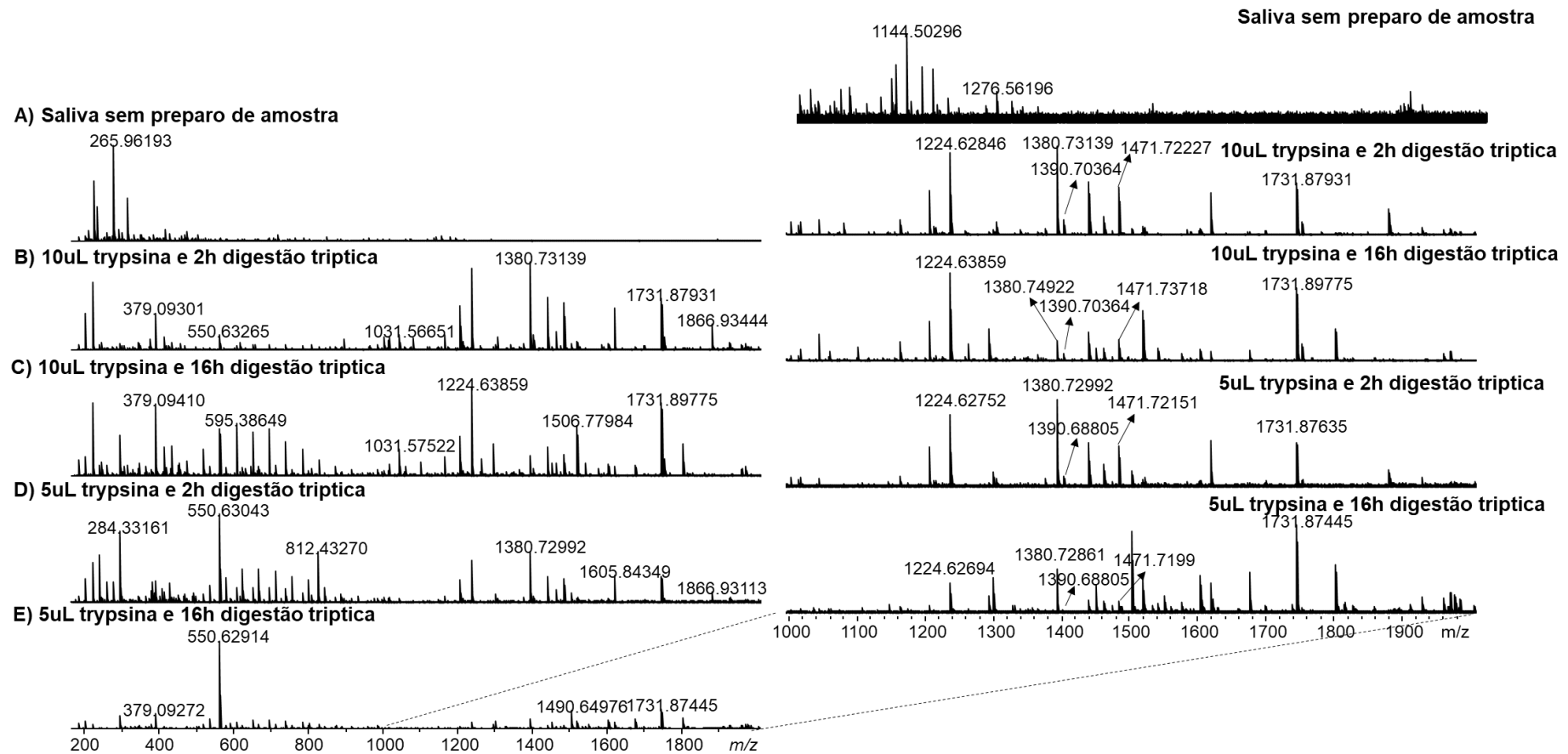


Tabela 2. 3 Número de sinais detectados com S/N > 4%.

Número de sinais detectados com sinal/ruído S/N > 4%.	
10 uL tripsina + 2 h para digestão	22786
10 uL tripsina + 16 h para digestão	22970
5 uL tripsina + 2 h para digestão	22538
5 uL tripsina + 16 h para digestão	22602

Tabela 2. 4 Sequência peptídica identificada em cada tipo de protocolo de proteína de digestão triptica.

10 uL tripsina + 2 h para digestão (m/z)	10 uL tripsina + 16 h para digestão (m/z)	5 uL tripsina + 2 h para digestão (m/z)	5 uL tripsina + 16 h para digestão (m/z)	Sequência de peptídeos	Proteína
1224.62846 TIC (5.2×10^7)	1224.63859 TIC (4.5×10^7)	1224.62752 TIC (9.4×10^6)	1224.62694 TIC (7.1×10^6)	GPPQQGGHPRPP	Fosfoproteína salivar ácida rica em prolina ou <i>PRH1</i> <i>Protein</i> (P02810)
1380.73139 TIC (5.6×10^7)	1380.74922 TIC (1.0×10^7)	1380.72992 TIC (1.1×10^7)	1380.72861 TIC (1.0×10^7)	GPPQQGGHPRPPR	
1390.68853 TIC (9.9×10^6)	1390.70365 TIC (3.7×10^6)	1390.68805 TIC (1.4×10^6)	1390.68878 TIC (6.2×10^5)	GGRPQGPPQGQSPQ	
1471.72227 TIC (3.0×10^7)	1471.73718 TIC (1.1×10^7)	1471.72151 TIC (5.2×10^6)	1471.71999 TIC (2.9×10^6)	GRPQGPPQQGGHQ	
1731.87931 TIC (3.2×10^7)	1731.89775 TIC (3.7×10^7)	1731.87635 TIC (5.7×10^6)	1731.87445 TIC (2.1×10^7)	GPPQQGGHPPPPQGRPQ	

Secreções salivares glandular humanas contêm várias fosfoproteínas ricas em prolina ácida. Essas proteínas têm importantes funções biológicas relacionadas ao fornecimento de um ambiente protetor para os dentes, e possuem outras atividades associadas à modulação da adesão de bactérias às superfícies orais.³⁰

Na otimização dos parâmetros de análises por MALDI MS foram avaliados o poder de resolução e o tempo de aquisição das amostras. O tamanho do conjunto de dados gerado está relacionado ao número de sinais no espectro de massas e este aumento de acordo com o poder de resolução, entretanto, a velocidade de aquisição diminui. Dessa forma, é necessário utilizar um poder de resolução ótimo que não afete a quantidade de informação adquirida no espectro de FT-ICR MS, de uma forma mais rápida.³¹

Nas análises dos espectros brutos, observamos na Tabela 2. 5, que o tempo de análise para uma aquisição com maior resolução (8M) é de 1,89 min. Ela gera um maior número de sinais com S/N > 4 (43.228), porém o tempo total para as 447 análises é de aproximadamente 14 h.

A filtragem de ruído é projetada para separar sinais de componente do fundo originado da matriz química ou interferência instrumental. Assim, os espectros foram corrigidos eliminando sinais com intensidade relativa abaixo do ruído.³²

Após essa correção, a quantidade de sinais diminuiu significativamente, tornando a diferença entre o tamanho do conjunto de dados mínima e a diferença em relação ao tempo significativa. Logo o poder de resolução escolhido para realizar as análises foi o 1M que apresentou 221 sinais com intensidade relativa acima do ruído e um tempo de análise de 3,8 h, apenas 1,1 h a mais que a análise mais rápida de resolução de 512k, no qual apresentou somente 169 sinais acima do ruído.

Tabela 2. 5 Número de sinais detectados e tempo de aquisição dos espectros de massas do MALDI FT-ICR em função da otimização dos parâmetros.

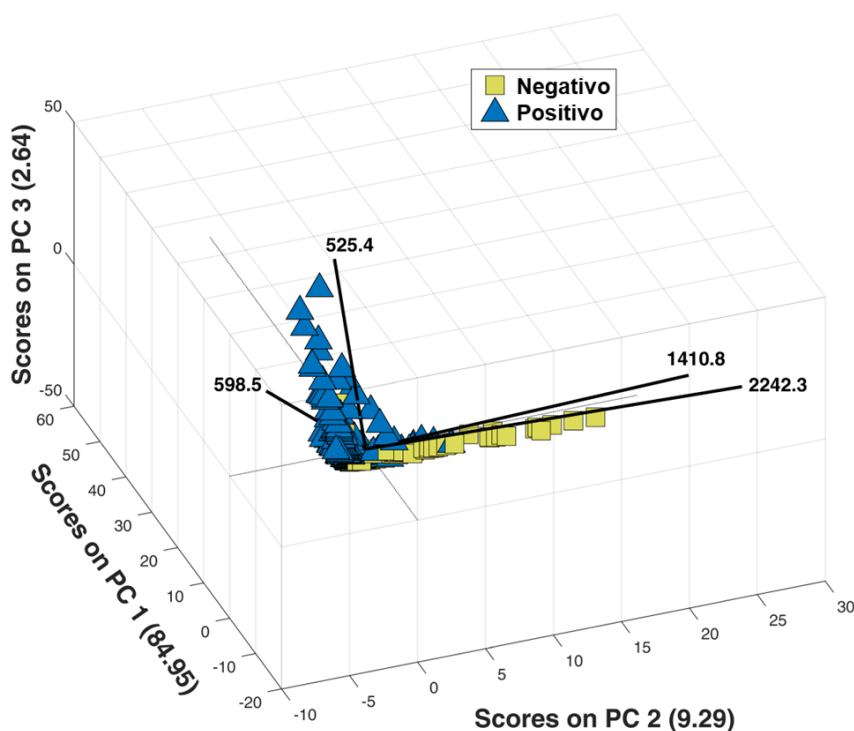
Resolução	Número de sinais com Sinal/ruído >4%	Número de sinais com intensidade relativa > % ruído	Tempo de aquisição (min)	Tempo total de aquisição (h)	TOTAL ION CURRENT
8M	43228	169	1,89	14,1	$1,1 \times 10^8$
4M	22699	234	1,06	7,9	$2,1 \times 10^8$
2M	11729	223	0,69	5,2	$1,9 \times 10^8$
1M	6493	221	0,51	3,8	$1,2 \times 10^8$
512k	3345	169	0,36	2,7	$3,9 \times 10^7$

Após a otimização dos parâmetros, 149 amostras (97 positivas para SARS-CoV-2 e 52 negativas) foram analisadas por MALDI FT-ICR MS. Uma análise de componente principal independente foi realizada como análise exploratória dos dados. O modelo PCA construído com as amostras de saliva positivas e negativas para SARS-CoV-2 apresentou 96,88% de variância total e o gráfico de *scores* e *loadings* das PC1 × PC2 × PC3 está contido na Figura 3. A partir dos *scores*, foi possível notar uma tendência de separação entre dois grupos principais, tendo a influência dos *m/z* 525,4 e 598,5 nas amostras positivas e *m/z* 1410,4 e 2242,3 nas negativas.

Nota-se que os ângulos entre os vetores das variáveis *m/z* 1410,4 e 2242,3 é quase zero, o que indica forte correlação entre essas variáveis e o mesmo é válido para as variáveis *m/z* 525,4 e 598,5. Já, o ângulo entre os vetores de *m/z* 1410,4 e 598,5 é quase perpendicular, indicando que a correlação é quase inexistente entre essas duas variáveis e o mesmo é válido para os de *m/z* 2242,3 e 525,4.

Posteriormente, a análise de ML foi empregada para identificar o modelo de melhor desempenho para distinguir os casos positivos e negativos para COVID-19.

Figura2. 3 Gráfico de *scores* e *loadings* da PCA do conjunto de dados dos espectros de massas MALDI(+) de amostras de saliva positivas e negativas para SARS-CoV-2.



Apesar das diferenças exibidas, nenhum biomarcador específico para qualquer um dos grupos de amostra positiva ou negativa foi encontrado pelas análises realizadas por MALDI FT-ICR MS. A abordagem ML foi então aplicada para classificar as amostras positivas para COVID, foi aplicado o FD (Discriminante de Fisher) para a seleção de variáveis. Dos modelos construídos, dois modelos apresentaram melhores resultados, um com FD de 30% e outro com 85%. Ao treinar os diferentes modelos, o modelo 1, com a seleção de variáveis com FD acima de 30%, apresentou um número maior de variáveis selecionadas (780) e melhor desempenho, com sensibilidade de 100%, especificidade de 87,5%, acurácia de 95,6%. Todavia, o modelo 2, que utilizou 85% do FD e selecionou apenas 2 variáveis, m/z 525,4 e m/z 1410,8 apresentou uma sensibilidade de 96,6%, especificidade de 68,7% e acurácia de 86,7%. As variáveis de maior importância foram as mesmas e apresentaram o mesmo perfil observado no modelo de PCA. No qual, variáveis com maiores m/z apresentaram maior correlação com amostras negativas, enquanto variáveis de menores m/z apresentaram forte correlação com as amostras positivas. Esse perfil também pode ser visto nos espectros brutos e após processamentos para as amostras positivas e negativas para SARS-CoV-2. Há maior número de sinais e em maior intensidade relativa em m/z mais altos para espectros de amostras negativas. Enquanto as amostras positivas possuem maior intensidade e números de sinais para m/z mais baixos.

Os modelos apresentaram um excelente desempenho e foi possível discriminar a classe positiva e negativa, já que há pouca variância espectral entre amostras de mesma classe. Os resultados do modelo 1, Tabela 2. 6, apresentaram uma TNR (Taxa de falso negativo) de 0,0% fazendo com que a metodologia utilizada seja eficiente em distinguir amostras positivas para SARS-CoV-2 das negativas em amostras de saliva uma matriz não invasiva, evitando diagnósticos de falso negativo.

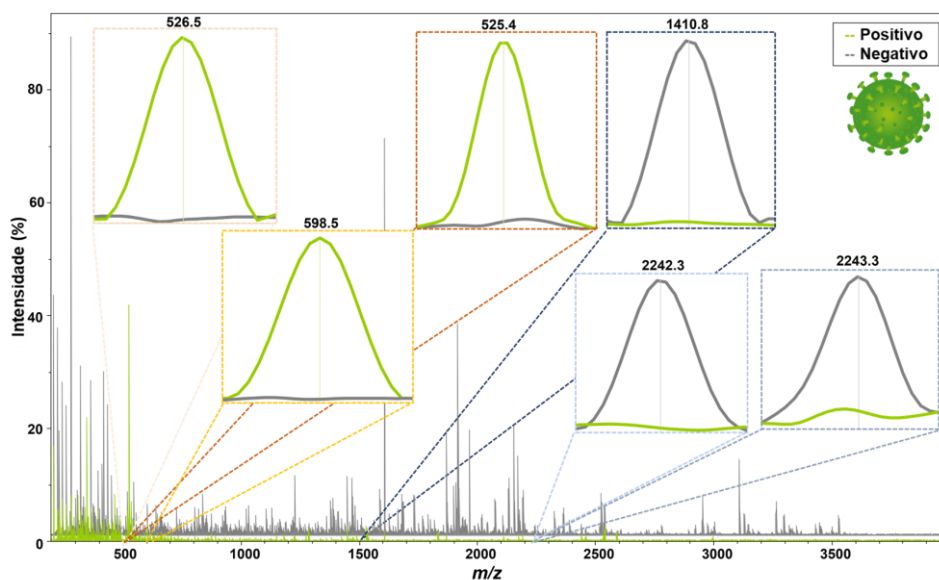
Tabela 2. 6 . Desempenho dos parâmetros dos modelos SVM pelos espectros MALDI com Discriminante de Fisher (FD).

Modelo	FD	Varsel	Grupo	VP	VN	FP	FN	SENS	ESPEC	ACC	VPP	VNP	TFP	TFN	MCC
1	30%	780	cal	204	107	1	0	100.0%	99.1%	99.7%	99.5%	100.0%	0.1%	0.0%	99.3%
			test	87	42	6	0	100.0%	87.5%	95.6%	93.6%	100.0%	12.5%	0.0%	90.5%
2	85%	2	cal	204	108	0	0	100.0%	100.0%	100.0%	100.0%	100.0%	0.0%	0.0%	100.0%
			test	84	33	15	3	96.6%	68.8%	86.7%	84.9%	91.7%	31.3%	3.5%	70.7%

FD (Discriminante de Fisher); Varsel (variáveis selecionadas); VP (verdadeiro positivo); VN (verdadeiro negativo); FP (falso positivo); FN (falso negativo); SENS (Sensibilidade); ESPEC (Especificidade); ACC (Acurácia); VPP (valor positivo de prevalência); VNP (valor negativo de prevalência); TFP (taxa de falsos positivos); TFN (taxa de falso negativo); MCC (coeficiente de correlação de Matthews).

Após obtenção dos modelos de classificação supervisionado e não supervisionado, observamos que sinais de m/z abaixo de 1000 possuem maior intensidade espectral em amostras positivas e de m/z acima de 1000 em amostras negativas, como demonstrado na Figura 2.4. O mesmo perfil é identificado na Figura 2.3, resultados da PCA, e nos resultados dos modelos SVM.

Figura 2.4 Sinal característico nos espectros MALDI(+) individuais entre amostras positivas para SARS-CoV-2 (verde) versus amostras negativas (cinza).



A partir dos resultados expostos, foi comparado com a literatura outros estudos que também aplicaram a análise multivariada de dados, ML, em dados obtidos pela técnica MALDI MS para o estudo de amostras biológicas de pacientes com suspeita de contaminação pelo SARS-CoV-2.

Rocca et.al. (2020)³³ obtiveram resultados preliminares para a detecção de perfis de proteínas diretamente de amostras de swabs nasofaríngeos positivas e negativas para SARS-CoV-2, onde a acurácia, sensibilidade e especificidade foram 68%, 62% e 72%, respectivamente.³³ No estudo de Nachtigall *et al.* (2020),¹⁹ um total de 362 amostras de swab nasal sem preparo prévio foram analisadas o MALDI MS com ML, o modelo apresentou maior acurácia (94%) com 7% de falsos positivos e 5% de falsos negativos.¹⁹ Já

Deulofeu *et al.* (2021)¹⁷ desenvolveram uma metodologia que combina MALDI-TOF MS com ML e o melhor modelo apresentou alto desempenho em termos de acurácia, sensibilidade e especificidade, atingindo em todos os casos, valores superiores a 90%.¹⁷

Diferente dos trabalhos citados acima, outros estudos, utilizam amostras de sangue (plasma ou soro). Lazari *et al.* (2021)³⁴ obtiveram resultados otimizados com acurácia de 92%, sensibilidade de 93% e especificidade de 92%, no conjunto de dados analisados.³⁴ Gomila *et.al.* (2021)³⁵ realizaram análise por MALDI-TOF MS de peptídeo sérico em pacientes clinicamente classificados como leve, grave, críticos e controle saudável. A abordagem ML classificou com 100% de acurácia em pacientes clínicos estáveis de acordo com a gravidade e previu corretamente a evolução dos pacientes instáveis em todos os casos.³⁵

Apesar de ótimos os resultados encontrados na literatura, relacionando a técnica MALDI MS com ML em amostras biológicas de pacientes suspeitos de COVID-19, os protocolos de coletas dessas amostras ainda são muito invasivos, incômodos e há um risco não desprezível de transmissão viral para o operador que realiza o procedimento.

Dessa forma, o uso da saliva como amostra biológica já despertou o interesse nos estudos de diagnósticos para COVID-19 como no estudo de Chivte *et al.* (2021)¹² e Costa *et al.* (2021)²³ que identificaram potenciais biomarcadores de infecção por SARS-CoV-2.^{12,23}

A alta precisão de medição de massa do FT-ICR MS reduz a probabilidade de uma atribuição incorreta de uma proteína no perfil de massa de peptídeos, além disso, a altíssima resolução do equipamento fornece uma maior quantidade de variáveis que torna a separação mais exata.³⁶

A ausência de estudos utilizando MALDI FT-ICR MS em diagnósticos rápidos em amostras biológicas torna o trabalho de grande valia para o desenvolvimento de novas técnicas com resultados de sensibilidade, especificidade e acurácia superiores a outras técnicas já estudadas anteriormente. Contudo, a singularidade dos perfis de MS entre indivíduos exige o desenvolvimento de análises bioestatísticas sofisticadas, como abordagens ML, tornando nosso estudo promissor no avanço das pesquisas utilizando a técnica MALDI MS combinada com ML para diagnósticos.

2.4 CONCLUSÃO

O SVM foi eficiente para diagnóstico de COVID-19 com TFN de 0.0%. Com isso, a metodologia criada, além de não invasiva, evita diagnósticos falso negativo. Além disso, o modelo SVM apresentou sensibilidade de 100%, especificidade de 87,5% e acurácia de 95,6%.

Na otimização da metodologia, o uso de 10uL de tripsina e 2h de digestão apresentou um maior TIC para todas as sequências de peptídeos, agilizando o processo de preparo das amostras. Já para as análises de MALDI FT-ICR MS o uso do poder de resolução de 1M apresentou uma ótima relação de sinais detectados (221 sinais) com intensidade relativa acima do ruído e um tempo de análise individual de meio minuto.

Os sinais com m/z abaixo de 1000 possuem maior intensidade em amostras positivas e sinais com m/z acima de 1000 em amostras negativas para os modelos de PCA e SVM.

A metodologia rápida de triagem de pacientes com suspeita de COVID-19 desenvolvida neste trabalho apresentou coleta simples, segura e não invasiva. A técnica MALDI FT-ICR MS possuiu alta sensibilidade, especificidade para detecção de COVID-19 combinada métodos de ML.

O artigo encontra-se publicado na Journal of Proteome Research, em Julho de 2022 Link para acesso: <https://pubs.acs.org/doi/10.1021/acs.jproteom.e.2c00148>

2.5 REFERÊNCIAS

- (1) Baloch, S.; Baloch, M. A.; Zheng, T.; Pei, X. The Coronavirus Disease 2019 (COVID-19) Pandemic. *Tohoku J. Exp. Med.* **2020**, *250* (4), 271–278.
- (2) Jebiril, N. World Health Organization Declared a Pandemic Public Health Menace: A Systematic Review of the Coronavirus Disease 2019 “COVID-19.” *SSRN Electron. J.* **2020**, *24* (9), 2784–2795.
- (3) Kucharski, A. J.; Russell, T. W.; Diamond, C.; Liu, Y.; Edmunds, J.; Funk, S.; Eggo, R. M.; Sun, F.; Jit, M.; Munday, J. D.; Davies, N.; Gimma, A.; van Zandvoort, K.; Gibbs, H.; Hellewell, J.; Jarvis, C. I.; Clifford, S.; Quilty, B. J.; Bosse, N. I.; Abbott, S.; Klepac, P.; Flasche, S. Early Dynamics of Transmission and Control of COVID-19: A Mathematical Modelling Study. *Lancet. Infect. Dis.* **2020**, *20* (5), 553–558.
- (4) Corman, V. M.; Landt, O.; Kaiser, M.; Molenkamp, R.; Meijer, A.; Chu, D. K. W.; Bleicker, T.; Brünink, S.; Schneider, J.; Schmidt, M. L.; Mulders, D. G. J. C.; Haagmans, B. L.; Van Der Veer, B.; Van Den Brink, S.; Wijsman, L.; Goderski, G.; Romette, J. L.; Ellis, J.; Zambon, M.; Peiris, M.; Goossens, H.; Reusken, C.; Koopmans, M. P. G.; Drosten, C. Detection of 2019 Novel Coronavirus (2019-nCoV) by Real-Time RT-PCR. *Eurosurveillance* **2020**, *25* (3), 23–30.
- (5) Laboratory Testing for Coronavirus Disease (COVID-19) in Suspected Human Cases.
- (6) Lippi, G.; Simundic, A. M.; Plebani, M. Potential Preanalytical and Analytical Vulnerabilities in the Laboratory Diagnosis of Coronavirus Disease 2019 (COVID-19). *Clin. Chem. Lab. Med.* **2020**, *58* (7), 1070–1076.
- (7) Kim, Y. G.; Yun, S. G.; Kim, M. Y.; Park, K.; Cho, C. H.; Yoon, S. Y.; Nam, M. H.; Lee, C. K.; Cho, Y. J.; Lim, C. S. Comparison between Saliva and Nasopharyngeal Swab Specimens for Detection of Respiratory Viruses by Multiplex Reverse Transcription-PCR. *J. Clin. Microbiol.* **2016**, *55* (1), 226–233.
- (8) Sasaki, T.; Inoue, O.; Ogihara, S.; Kubokawa, K.; Oishi, S.; Shirai, T.; Iwabuchi, K.; Suzuki-Inoue, K. Detection of SARS-CoV-2 RNA Using RT-

- QPCR in Nasopharyngeal Swab, Saliva, Lingual, and Buccal Mucosal Swab. *Jpn. J. Infect. Dis.* **2022**, *75*, 102–104.
- (9) Yamayoshi, S.; Sakai-Tagawa, Y.; Koga, M.; Akasaka, O.; Nakachi, I.; Koh, H.; Maeda, K.; Adachi, E.; Saito, M.; Nagai, H.; Ikeuchi, K.; Ogura, T.; Baba, R.; Fujita, K.; Fukui, T.; Ito, F.; Hattori, S. I.; Yamamoto, K.; Nakamoto, T.; Furusawa, Y.; Yasuhara, A.; Ujie, M.; Yamada, S.; Ito, M.; Mitsuya, H.; Omagari, N.; Yotsuyanagi, H.; Iwatsuki-Horimoto, K.; Imai, M.; Kawaoka, Y. Comparison of Rapid Antigen Tests for COVID-19. *Viruses* **2020**, *Vol. 12*, Page 1420 **2020**, *12* (12), 1420.
- (10) Scohy, A.; Anantharajah, A.; Bodéus, M.; Kabamba-Mukadi, B.; Verroken, A.; Rodriguez-Villalobos, H. Low Performance of Rapid Antigen Detection Test as Frontline Testing for COVID-19 Diagnosis. *J. Clin. Virol.* **2020**, *129*, 104455.
- (11) Ihling, C.; Tänzler, D.; Hagemann, S.; Kehlen, A.; Hüttelmaier, S.; Arlt, C.; Sinz, A. Mass Spectrometric Identification of SARS-CoV-2 Proteins from Gargle Solution Samples of COVID-19 Patients. *J. Proteome Res.* **2020**, *19* (11), 4389–4392.
- (12) Chivte, P.; LaCasse, Z.; Seethi, V. D. R.; Bharti, P.; Bland, J.; Kadkol, S. S.; Gaillard, E. R. MALDI-ToF Protein Profiling as a Potential Rapid Diagnostic Platform for COVID-19. *J. Mass Spectrom. Adv. Clin. Lab* **2021**, *21*, 31–41.
- (13) Frampas, C. F.; Longman, K.; Spick, M. P.; Lewis, H. M.; Costa, C. D. S.; Stewart, A.; Dunn-Walters, D.; Greener, D.; Evetts, G. E.; Skene, D.; Trivedi, D.; Pitt, A. R.; Hollywood, K.; Barran, P.; Bailey, M. J. Untargeted Saliva Metabolomics Reveals COVID-19 Severity. *medRxiv* **2021**, 6;21260080.
- (14) Spielmann, N.; Wong, D. Saliva: Diagnostics and Therapeutic Perspectives. *Oral Dis.* **2011**, *17* (4), 345–354.
- (15) Kawas, S. Al; Rahim, Z. H. A.; Ferguson, D. B. Potential Uses of Human Salivary Protein and Peptide Analysis in the Diagnosis of Disease. *Arch. Oral Biol.* **2012**, *57* (1), 1–9.
- (16) Baghizadeh Fini, M. Oral Saliva and COVID-19. *Oral Oncol.* **2020**, *108*, 104821.
- (17) Deulofeu, M.; García-Cuesta, E.; Peña-Méndez, E. M.; Conde, J. E.;

- Jiménez-Romero, O.; Verdú, E.; Serrando, M. T.; Salvadó, V.; Boadas-Vaello, P. Detection of SARS-CoV-2 Infection in Human Nasopharyngeal Samples by Combining MALDI-TOF MS and Artificial Intelligence. *Front. Med.* **2021**, *8*, 398–410.
- (18) Tran, N. K.; Howard, T.; Walsh, R.; Pepper, J.; Loegering, J.; Phinney, B.; Salemi, M. R.; Rashidi, H. H. Novel Application of Automated Machine Learning with MALDI-TOF-MS for Rapid High-Throughput Screening of COVID-19: A Proof of Concept. *Sci. Reports 2021 111* **2021**, *11* (1), 1–10.
- (19) Nachtigall, F. M.; Pereira, A.; Trofymchuk, O. S.; Santos, L. S. Detection of SARS-CoV-2 in Nasal Swabs Using MALDI-MS. *Nat. Biotechnol.* **2020**, *38* (10), 1168–1173.
- (20) Bradford, M. M. A Rapid and Sensitive Method for the Quantitation of Microgram Quantities of Protein Utilizing the Principle of Protein-Dye Binding. *Anal. Biochem.* **1976**, *72* (1–2), 248–254.
- (21) Huang, C.-M.; Zhu, W. Profiling Human Saliva Endogenous Peptidome via a High Throughput MALDI-TOF-TOF Mass Spectrometry. *Comb. Chem. High Throughput Screen.* **2009**, *12* (5), 521–531.
- (22) Kennard, R. W.; Stone, L. A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11* (1), 137–148.
- (23) Costa, M. M.; Benoit, N.; Saby, F.; Pradines, B.; Granjeaud, S.; Almeras, L. Optimization and Standardization of Human Saliva Collection for MALDI-TOF MS. *Diagnostics* **2021**, *11* (8), 1304–1319.
- (24) Lorenzo-Pouso, A. I.; Pérez-Sayáns, M.; Bravo, S. B.; López-Jornet, P.; García-Vence, M.; Alonso-Sampedro, M.; Carballo, J.; García-García, A.; Zalewska, A. Protein-Based Salivary Profiles as Novel Biomarkers for Oral Diseases. *Dis. Markers* **2018**, *2018*.
- (25) Bogdanov, B.; Smith, R. D. Proteomics by FTICR Mass Spectrometry: Top down and Bottom Up. *Mass Spectrom. Rev.* **2005**, *24* (2), 168–200.
- (26) Cantú, M. D.; Carrilho, E.; Wulff, N. A.; Palma, M. S. Seqüenciamento de Peptídeos Usando Espectrometria de Massas: Um Guia Prático. *Quim. Nova* **2008**, *31* (3), 669–675.
- (27) Chen, W.; Hwang, Y. Y.; Gleaton, J. W.; Titus, J. K.; Hamlin, N. J. Optimization of a Peptide Extraction and LC-MS Protocol for Quantitative

- Analysis of Antimicrobial Peptides. *Futur. Sci. OA* **2019**, *5* (1).
- (28) Lucchi, G.; Chambon, C.; Truntzer, C.; Pecqueur, D.; Ducoroy, P.; Schwartz, C.; Nicklaus, S.; Morzel, M. Mass-Spectrometry Based Characterisation of Infant Whole Saliva Peptidome. *Int. J. Pept. Res. Ther.* **2009**, *15* (3), 177–185.
- (29) Hardt, M.; Witkowska, H. E.; Webb, S.; Thomas, L. R.; Dixon, S. E.; Hall, S. C.; Fisher, S. J. Assessing the Effects of Diurnal Variation on the Composition of Human Parotid Saliva: Quantitative Analysis of Native Peptides Using ITRAQ Reagents. *Anal. Chem.* **2005**, *77* (15), 4947–4954.
- (30) Hay, D. I.; Bennick, A.; Schlesinger, D. H.; Minaguchi, K.; Madapallimattam, G.; Schluckebier, S. K. The Primary Structures of Six Human Salivary Acidic Proline-Rich Proteins (PRP-1, PRP-2, PRP-3, PRP-4, PIF-s and PIF-F). *Biochem. J.* **1988**, *255* (1), 15–21.
- (31) Ferey, J.; Marguet, F.; Laquerrière, A.; Marret, S.; Schmitz-Afonso, I.; Bekri, S.; Afonso, C.; Tebani, A. A New Optimization Strategy for MALDI FTICR MS Tissue Analysis for Untargeted Metabolomics Using Experimental Design and Data Modeling. *Anal. Bioanal. Chem.* **2019**, *411* (17), 3891–3903.
- (32) Yi, L.; Dong, N.; Yun, Y.; Deng, B.; Ren, D.; Liu, S.; Liang, Y. Chemometric Methods in Data Processing of Mass Spectrometry-Based Metabolomics: A Review. *Anal. Chim. Acta* **2016**, *914*, 17–34.
- (33) Rocca, M. F.; Zintgraff, J. C.; Dattero, M. E.; Santos, L. S.; Ledesma, M.; Vay, C.; Prieto, M.; Benedetti, E.; Avaro, M.; Russo, M.; Nachtigall, F. M.; Baumeister, E. A Combined Approach of MALDI-TOF Mass Spectrometry and Multivariate Analysis as a Potential Tool for the Detection of SARS-CoV-2 Virus in Nasopharyngeal Swabs. *J. Virol. Methods* **2020**, *286*, 113991.
- (34) Lazari, L. C.; de Rose Ghilardi, F.; Rosa-Fernandes, L.; Assis, D. M.; Nicolau, J. C.; Santiago, V. F.; Dalçóquio, T. F.; Angeli, C. B.; Bertolin, A. J.; Marinho, C. R. F.; Wrenger, C.; Durigon, E. L.; Siciliano, R. F.; Palmisano, G. Prognostic Accuracy of MALDI-TOF Mass Spectrometric Analysis of Plasma in COVID-19. *Life Sci. Alliance* **2021**, *4* (8), 1–12.
- (35) Gomila, R. M.; Martorell, G.; Fraile-Ribot, P. A.; Doménech-Sánchez, A.; Albertí, M.; Oliver, A.; García-Gasalla, M.; Albertí, S. Use of Matrix-

- Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry Analysis of Serum Peptidome to Classify and Predict Coronavirus Disease 2019 Severity. *Open Forum Infect. Dis.* **2021**, 8 (6), 1–8.
- (36) Yang, H. J.; Park, K. H.; Shin, S.; Lee, J. H.; Park, S.; Kim, H. S.; Kim, J. Characterization of Heme Ions Using MALDI-TOF MS and MALDI FT-ICR MS. *Int. J. Mass Spectrom.* **2013**, 343–344, 37–44.

CAPITULO III – Triagem para diagnóstico para COVID-19 baseado em dados multiômicos por espectrometria de massas de alta resolução (MALDI(+)-TOF MS e ESI(±)-Orbitrap MS)

“DIAGNOSTIC SCREENING OF COVID-19 BASED ON MULTIOMICS DATA BY HIGH-RESOLUTION MASS SPECTROMETRY (MALDI (+)-TOF MS and ESI(±)-Orbitrap MS)”

RESUMO

A urgência de um diagnóstico rápido da COVID-19, doença que resultou na pandemia de 2020, ressaltou a importância do desenvolvimento de novos testes analíticos com respostas rápidas e eficazes, levando em conta a sua natureza altamente contagiosa. Dessa forma, este estudo teve como objetivo desenvolver um método eficaz para triagem de pacientes infectados com SARS-CoV-2 com base nos perfis lipídicos e proteômicos do soro. A espectrometria de massas por ionização por *electrospray* (ESI MS) foi utilizada para investigar o perfil lipídico de 239 amostras de soro (119 positivas e 120 negativas para COVID-19). Já para analisar o perfil proteômico de 300 amostras de soro (150 positivas e 150 negativas para infecção por SARS-CoV-2), foi utilizada a espectrometria de massas com ionização e dessorção a laser assistida por matriz (MALDI MS). Após o processamento dos dados de MS e a seleção de variáveis, análises estatísticas, como *Volcano plot*, o *Heatmap*, a análise de componentes principais (PCA), a análise discriminante de mínimos quadrados parciais (PLS-DA) e máquina de vetores de suporte (SVM), foram realizadas para distinguir as variáveis mais relevantes para classificar amostras positivas e negativas para COVID-19. Por exemplo, a análise lipidômica usando ESI(±)-Orbitrap MS e modelos SVM mostrou sensibilidades de 96,67% e 100%, especificidades de 82,14% e 96,88%, e acurácias de 89,66% e 98,44% para análises no modo positivo de ionização e modo negativo, respectivamente. Para análises proteômicas usando MALDI(+)-TOF MS, o modelo linear PLS-DA demonstrou uma precisão de 99,10%. No geral, a combinação das técnicas ESI-Orbitrap MS e MALDI-TOF MS com quimiometria empregada neste estudo demonstrou alternativas promissoras com alta sensibilidade e especificidade para discriminar amostras de soro infectadas e não infectadas pelo SARS-CoV-2. Porém, a investigação do perfil lipídico pelo DI-ESI MS representa uma abordagem inovadora que amplia a aplicação da espectrometria de massa em diagnósticos clínicos.

Palavras-chave: COVID-19, triagem, espectrometria de massas, MALDI-TOF MS, ESI-Orbitrap MS, quimiometria.

3.1 INTRODUÇÃO

A pandemia da COVID-19 apresentou um desafio significativo para a saúde pública e para a comunidade científica global. A necessidade de isolamento dos pacientes diagnosticados com esta doença impulsionou o desenvolvimento de testes rápidos e eficazes.¹

Para o diagnóstico de COVID-19, existem dois tipos amplamente utilizados de testes: o teste molecular de reação em cadeia da polimerase com transcrição reversa em tempo real (RT-PCR), utilizado para identificar a presença de RNA viral em fluidos corporais durante a fase aguda da infecção, e o teste sorológico, que detecta os anticorpos imunoglobulina G (IgG) e imunoglobulina M (IgM) produzidos como resposta imune ao vírus. No entanto, o momento preciso da administração do teste após o início dos sintomas é crucial para ambos os testes.² O teste RT-PCR possui uma taxa de falso positivo que varia dependendo do dia da coleta da secreção e pode chegar a 15 a 20%.^{3,4} Além disso, os resultados dos testes moleculares geralmente levam aproximadamente 72 h para serem liberados, dificultando as estratégias de isolamento e tratamento para indivíduos potencialmente infectados. Em contraste, os testes sorológicos para IgM e IgG devem ser realizados apenas após 14 dias do início dos sintomas.⁵⁻⁷ O RT-PCR é o principal método para o diagnóstico de COVID-19, pois a utilização de swabs nasofaríngeos para testes apresenta maior probabilidade de contaminação com o coronavírus da síndrome respiratória aguda grave 2 (SARS-CoV-2) em comparação com amostras de sangue.⁸ Entretanto, existe uma probabilidade de contaminação a profissionais e cuidado no transporte. Além disso, resultados falso negativos também podem contribuir para o aumento da disseminação da doença.⁵

A infecção por vírus ou bactéria resulta em alterações substanciais no metabolismo da célula hospedeira, incluindo mudanças na concentração e composição de várias espécies lipídicas, bem como modificações nos níveis de proteínas no plasma. Dessa forma, o soro do paciente pode refletir sua condição fisiológica ou patológica.^{9,10} Testes sanguíneos com múltiplos índices bioquímicos são rotineiramente realizados. A análise do soro de pacientes infectados com COVID-19 tem sido uma abordagem significativa para entender

a patologia da doença e identificar biomarcadores que podem contribuir para o diagnóstico e tratamento.¹¹

Entre as técnicas analíticas aplicadas em diagnósticos clínicos hoje, a espectrometria de massa (MS) tem ganhado destaque como uma técnica analítica poderosa. Suas vantagens sobre métodos tradicionais (RT-PCR e sorológicos) incluem alto rendimento, sensibilidade e seletividade aprimoradas, tempos de análise curtos, um pequeno volume inicial de amostra e a ausência da necessidade de identificar um único biomarcador.¹²

Além disso, o uso de técnicas de ionização suave, como ionização e dessorção a laser assistida por matriz (MALDI) e ionização por *eletrospray* (ESI), na detecção e caracterização de biomoléculas está se tornando cada vez mais comum no campo da medicina, especialmente na análise associada a infecções virais, como a COVID-19.¹²

A técnica MALDI MS foi adotada em inúmeros laboratórios de diagnóstico em todo o mundo, e sua análise pode ser caracterizada como de alto rendimento e ultrarrápida (<1 s/amostra) para amostras clínicas. Os espectros de massas obtidos fornecem informações abundantes sobre o perfil metabólico, que pode ser utilizado para a discriminação de espécies. Essa abordagem ganhou destaque no diagnóstico de várias doenças significativas, incluindo a análise de amostras de COVID-19.¹³⁻¹⁸ Embora as identidades químicas dos biomarcadores usados para detectar o SARS-CoV-2 permaneçam não identificadas, a aplicação de várias estratégias de aprendizado de máquina para identificar perfis espectrais específicos de COVID-19, tem se mostrado altamente promissora.^{9,18}

Em aplicações clínicas, além da MALDI MS, técnicas hífenadas são os sistemas MS mais explorados, especialmente a cromatografia líquida de alto ou ultra eficiência (HPLC ou UHPLC), devido à complexidade de fluidos biológicos. No entanto, a infusão direta de amostras em espectrômetros de alta resolução pode fornecer uma análise mais rápida e reproduzível, servindo potencialmente como uma ferramenta de triagem de alto rendimento, capaz de analisar centenas de amostras por dia com um tempo de análise geralmente inferior a 1 minuto.¹⁹

Estudos já demonstraram a infusão direta em ESI MS para caracterização lipídica em doenças como o Alzheimer²⁰ e o risco de pré-eclâmpsia.²¹ Além

disso, o perfil lipídico do soro de amostras de pacientes infectados com SARS-CoV-2 foi investigado. No entanto, até o momento, nenhum desses resultados encontrados na literatura empregou a infusão direta (DI) de amostras de soro contendo metabólitos no espectrômetro de massas, melhorando assim a velocidade e a reprodutibilidade da análise, a cobertura de metabólitos não direcionados e, conseqüentemente, a capacidade de triagem de alto rendimento.²²

A enorme quantidade de dados obtidos a partir dos espectros de MS exige técnicas estatísticas robustas para identificar padrões diagnósticos.^{17,23} Embora as identidades químicas dos biomarcadores usados para detectar SARS-CoV-2 permaneçam não identificadas, a aplicação de várias estratégias de aprendizado de máquina (ML) para identificar perfis espectrais específicos da COVID-19 provou ser altamente promissor.

A utilização de ML para analisar dados de MS e classificar amostras positivas para doenças emergiu como uma estratégia potente para a triagem de pacientes com COVID-19. Além disso, a utilização da DI-ESI MS para análises rápidas amplia o emprego de diferentes metodologias analíticas para diagnóstico rápido e eficaz em várias doenças. Assim, este estudo pode oferecer um método de triagem eficiente, utilizando a informação espectral de amostras de soro de pacientes infectados por SARS-CoV-2, empregando espectrometria de massa (ESI-Orbitrap MS e MALDI-TOF MS) combinada com quimiometria.

3.2 METODOLOGIA

3.2.1 Amostras e reagentes

3.2.1.1 Reagentes

Matriz CHCA (α -cyano-4-hydroxycinnamic acid, Sigma-Aldrich Chemicals, St. Louis, MO, USA), Clorofórmio (CHCl_3 , Sigma-Aldrich Chemicals, St. Louis, MO, USA), Metanol (MeOH, Sigma-Aldrich Chemicals, St. Louis, MO, USA), de Acetonitrila (Sigma-Aldrich Chemicals, St. Louis, MO, USA), ácido trifluoroacético (TFA, Sigma-Aldrich Chemicals, EUA).

3.2.1.2 Amostras de soro

As amostras de soro obtidas pelo Laboratório de análises clínicas (Vitória-ES, Brasil) e cedidas sem nenhuma identificação nominal e sem dados sobre pacientes, apenas os resultados do teste ELISA para COVID-19. As amostras foram coletadas seguindo os princípios éticos (Comitê de Ética em Pesquisa nº CAAE: 51803621.1.0000.5060) e armazenadas em freezer -80°C .

3.2.2 Extração de lipídios

Foram utilizadas um total de 239 amostras, compreendendo 119 positivas e 120 negativas para COVID-19, de acordo com um teste ELISA, para extração de lipídios. Os lipídios foram extraídos utilizando um protocolo adaptado de Bligh & Dyer.²⁴ Neste procedimento, foram adicionados 500 μL de CHCl_3 , 500 μL de MeOH e 260 μL de água ultrapura a 190 μL de soro. A solução foi então agitada e centrifugada a 15.000 G por 10 min, resultando em um sistema de três fases (fase aquosa, pellet de proteína e fase orgânica). A fase orgânica foi coletada em um frasco de vidro, e 500 μL de CHCl_3 foram adicionados à fase aquosa restante + pellet de proteína para aumentar a eficiência da extração. Esta solução foi então agitada, centrifugada a 15.000 G por 10 minutos, e a fase orgânica resultante foi coletada e misturada com a primeira fase orgânica. A mistura de fases orgânicas foi agitada e, em seguida, evaporada em um concentrador SpeedVac (Savant SPD131DDA, Thermo Scientific, Massachusetts, EUA). Finalmente, a amostra foi ressuspensa em 200 μL de

MeOH, centrifugada a 15.000 G por 10 min, e o sobrenadante foi analisado por ESI(±)-Orbitrap MS.

3.2.3 Preparo das amostras para análise MALDI(+)-TOF MS

Trezentas amostras de soro (150 positivas para COVID-19 e 150 negativas para COVID-19, com diagnóstico confirmado por teste ELISA) foram purificadas com uma coluna de spin C18 (Pierce C18 Spin Column, Thermo Scientific, San Jose, CA, EUA). Essas colunas foram ativadas por 2 × 200 µL de ACN a 50%v/v e equilibradas com 2 × 200 µL de TFA a 0,5%v/v. As amostras foram adsorvidas ao meio usando dois ciclos repetidos de 60 µL (composto por 15 µL de TFA a 2%v/v em ACN a 20%v/v, mais 45 µL de amostra), e a coluna foi lavada com 2 × 200 µL de TFA a 0,5%v/v. Finalmente, os peptídeos foram eluídos em 2 × 20 µL de ACN a 70%v/v. Todas as etapas, exceto a eluição, foram realizadas em uma centrífuga de bancada a 1.000 G por 2 min.

3.2.4 Análise por ESI(±)-Orbitrap MS

Para a aquisição de dados das amostras extraídas, foi utilizado um espectrômetro de massas Q-Exactive Orbitrap (*Thermo Scientific*, Bremen, Alemanha) acoplado a um cromatógrafo líquido de alta performance (*Agilent* 1260 *infinity* I, Santa Clara, EUA) para infusão direta automática. Uma alíquota de 40 µL da amostra foi injetada em um fluxo de 300 µL/min de MeOH (0,1%v/v de ácido fórmico), no modo positivo e MeOH (0,1%v/v de hidróxido de amônio) no modo negativo. Espectros de massa de alta resolução foram adquiridos em ambos os modos de ionização positiva e negativa, com uma voltagem de *electrospray* aplicada a +/- 3,5 kV e temperatura capilar a 256°C.

O modo de Aquisição Independente de Dados (do inglês, *Data Independent Acquisition*, DIA) foi utilizado para adquirir espectros de MS/MS para os íons detectados. Um espectro de MS completo (m/z 150–1200) foi adquirido no Orbitrap com uma resolução de 140.000, um alvo de controle de ganho automático de 10^6 e tempo de injeção máximo de 200 ms. O MS/MS foi realizado com um alvo de controle de ganho automático de 10^5 , tempo de

injeção máximo de 50 ms, resolução de massa de 70.000 e largura de isolamento do precursor de 0,4 m/z .

O padrão utilizado no modo positivo de ionização foi cafeína-(3-metil- ^{13}C) (700 ng/mL), e no modo negativo, Bisfenol-S e Prednisolona (500 ng/mL).

.3.2.5 Análise por MALDI(+)-TOF MS

Um microlitro de cada amostra foi aplicado em triplicata na placa MALDI (*Bruker Daltonics*), seguido da adição de 1 μL da solução de matriz CHCA preparada por dissolução em ACN/água 50:50 v/v contendo TFA a 10 mg/mL e seco em temperatura ambiente. A concentração da matriz e do ácido foram avaliadas, 0,1%v/v ou 2,5%v/v de TFA, e 1 mg/mL e 10 mg/mL de concentração de CHCA.¹⁴

As amostras foram analisadas em um espectrômetro de massa MALDI-TOF Autoflex speed smartbeam (*Bruker Daltonics*, Bremen, Alemanha) usando o *software* FlexControl (versão 3.3; *Bruker Daltonics*, Bremen, Alemanha). Os espectros foram registrados no modo linear positivo (*laser frequency*, 2000 Hz; *Pulsed ion Extraction time*, 170 ns; *Ion Source 1 high-voltage*, 19,59 kV; *Ion Source 2 high-voltage*, 18,09 kV; *lens high-voltage*, 9,48 kV; faixa de massa, 800-20.000 Da). Os espectros foram adquiridos usando o modo de execução automática para evitar interferência subjetiva na aquisição de dados. Para cada amostra, foram somados 20.000 *shots*, em passos de 2.000 *shots*. Todos os espectros foram calibrados usando o padrão de calibração de Proteína I (Insulina $[\text{M}+\text{H}]^+ = 5.734,52$, Citocromo C $[\text{M}+ 2\text{H}]^{2+} = 6.181,05$, Mioglobina $[\text{M}+ 2\text{H}]^{2+} = 8.476,66$, Ubiquitina I $[\text{M}+\text{H}]^+ = 8.565,76$, Citocromo C $[\text{M}+\text{H}]^+ = 12.360,97$, Mioglobina $[\text{M}+\text{H}]^+ = 16.952,31$) (*Bruker Daltonics*).

3.2.6 Processamento de dados e análise estatística de dados

3.2.6.1 Abertura de amostras

O conjunto de dados de ESI(\pm)-Orbitrap MS obtido com o *software* XCalibur 4.2 (*Thermo Scientific*, Bremen, Alemanha) e o conjunto de dados de MALDI(+)-TOF MS obtido com o *software* FlexControl v3.4 (*Bruker Daltonics*, Bremen, Alemanha) foram exportados em formato .txt e importados para o

software MATLAB 9.0 (R2013a) e MetaboAnalyst 5.0 (<https://www.metaboanalyst.ca>).²⁵

3.2.6.2 Discriminante de Fisher (FD)

A análise FD foi utilizada para reduzir variáveis com diferenças significativas entre as duas classes. A razão de Fisher (Equações (1), (2) e (3), Tabela 3. 1) é definida como a variação classe por classe da variável independente dividida pela soma das variações dentro da classe.²⁶

Foram utilizadas onze seções, variando de 10% a 95% de corte variável. Além da seleção de variáveis, um limiar de intensidade relativa de 1% foi aplicado para redução de dados/subtração de ruído.

3.2.6.3 Análise estatística dos dados

Após o corte das variáveis e pré-processamento dos espectros, uma investigação preliminar foi realizada utilizando análise de componentes principais (PCA) em cada um dos 3 conjuntos de dados, ESI(\pm) MS e MALDI(+) MS. Em seguida, foi construído o modelo de análise discriminante de mínimos quadrados parciais (PLS-DA) para discriminar pacientes indivíduos infectados (positivos) e não infectados (negativos) com SARS-CoV-2. A validação cruzada 5-fold foi realizada e o número ideal de variáveis latentes (nLVs) foi selecionado. Os nLVs foram selecionados maximizando a precisão na validação cruzada. Pontuações de importância variável na projeção (VIP) foram usadas para avaliar a significância do sinal. As pontuações dos recursos VIP são baseadas na quantidade de variação entre os pacientes positivos e negativos que é explicada por cada recurso nos componentes. Para todos os efeitos, as pontuações VIP refletem a importância relativa de um recurso na classificação dos participantes como doentes ou controles.

Além dos três modelos foram feitas as análises *Fold-Change* (FC), construindo um *Volcano plot* e *Heatmap* para os dados MALDI(+) MS, utilizando o MetaboAnalyst 5.0 (<https://www.metaboanalyst.ca>)²⁵ e o software MATLAB 9.0 (R2013a).

Os modelos de classificação de aprendizagem de Máquina por Vetores de Suporte (SVM), realizados com o software MATLAB 9.0 (R2013a), foi desenvolvido para classificar amostras positivas e negativas para COVID-19.

Os conjuntos de dados tanto do ESI(\pm)-Orbitrap MS quanto do MALDI(+)-TOF MS foram divididos em 70% do grupo amostral para o conjunto de treinamento e 30% para o conjunto de validação, por metodologia *Kennard-Stone*.²⁷ Os parâmetros custo (C), margem (ϵ) e gama (γ) da função *kernel* foram otimizados.

O desempenho do modelo foi avaliado considerando as métricas de acurácia, sensibilidade e especificidade. Essas métricas foram calculadas usando as seguintes figuras de mérito: falso positivo (FP), falso negativo (FN), verdadeiro positivo (VP) e verdadeiro negativo (VN), de acordo com as equações (Tabela3. 1).

Tabela3. 1 Métricas do modelo SVM

Métricas	Equação
Peso de Fisher	$FD = \frac{D}{d}$ (1)
Intraclasse (D)	$D = \frac{\sum(\bar{x}_i - \bar{x})^2 n_i}{(L-1)}$ (2)
Varição intraclasse (d)	$d = \frac{\sum(\sum(x_{ij} - \bar{x})^2) - (\sum((\bar{x}_i - \bar{x})^2 n_i))}{(N-L)}$ (3)
Acurácia	$Acc = \frac{VP + VN}{VP + VN + FP + FN}$
Sensibilidade	$Sens = \frac{VP}{VP + FN}$
Especificidade	$Espec = \frac{VN}{VN + FP}$
Taxa de falso positivo	$TFP = \frac{FP}{FP + VN}$
Taxa de falso negativo	$TFN = \frac{FN}{FN + VP}$
Valor preditivo positivo	$VPP = \frac{VP}{VP + FP}$
Valor preditivo negativo	$VPN = \frac{VN}{VN + FN}$

3.3 RESULTADOS E DISCUSSÃO

Lipidômica

Para extração de lipídios séricos totais, foi utilizado o método adaptado Bligh & Dyer.²⁴ Esse método envolve extração líquido-líquido em uma mistura de CHCl_3 , água e MeOH, que separa o homogeneizado em duas camadas: a camada de CHCl_3 contendo principalmente lipídios e a camada de MeOH contendo outros compostos endógenos. Um extrato lipídico purificado é obtido simplesmente isolando a camada de CHCl_3 . Após a secagem e ressuspensão em MeOH, as amostras foram analisadas por infusão direta usando o injetor automático de cromatografia líquida de alta performance acoplado ao espectrômetro de massas Orbitrap nos modos de ionização por *electrospray* positivo e negativo. Os espectros estão mostrados na Figura3. 1.

Alguns estudos já destacaram a importância de analisar lipídios em ambos os modos de ionização, positivo e negativo. Devido à sua diversidade estrutural, várias íons precursores e íons produtos derivados de espécies moleculares de lipídios exibem sensibilidades distintas que dependem do modo de ionização.^{28,29}

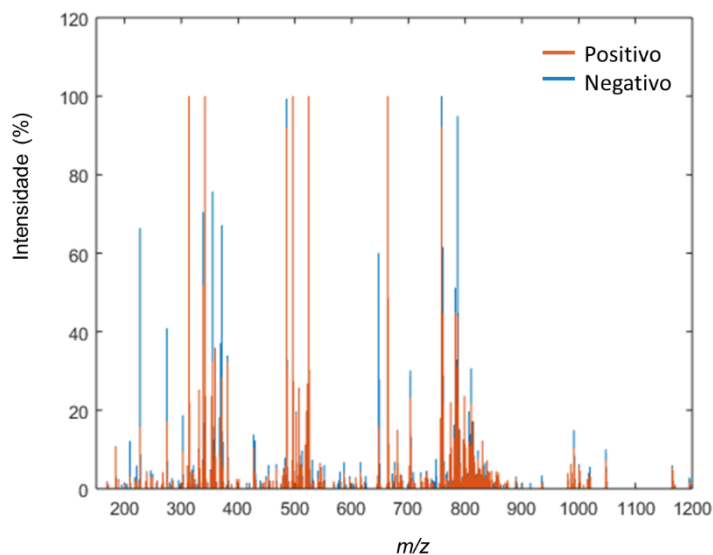
Nas análises de ESI(+)-Orbitrap MS, um perfil espectral semelhante foi observado para ambas as classes de amostras (positivas e negativas para SARS-CoV-2) (Figura3. 1A). De acordo com Schumann e colegas (2012), as classes de lipídios mais comumente encontradas em análises no modo de ionização positiva são colesterol (Chol), ésteres de colesterol (CholE) e triacilgliceróis (TAG). Eles examinaram extratos lipídicos de plasma sanguíneo humano usando nanoESI(+)-FT MS e detectaram 62 lipídios, incluindo 44 TAG, 17 Chol e 1 CholE.²⁸

Nas análises de ESI(-)-Orbitrap MS (Figura3. 1B), um maior número de sinais foi observado nos espectros das amostras negativas para COVID-19, com maior intensidade de sinal em valores de m/z maiores que 500, em comparação com as amostras positivas para COVID-19. Além disso, um maior número de sinais foi observado nas análises de ESI(-)-Orbitrap MS (2216 variáveis) (Figura3. 1B) em comparação com as análises de ESI(+)-Orbitrap MS (874 variáveis) (Figura3. 1A). Estudos anteriores relataram que no modo de ionização negativa, é possível identificar uma maior variedade de espécies

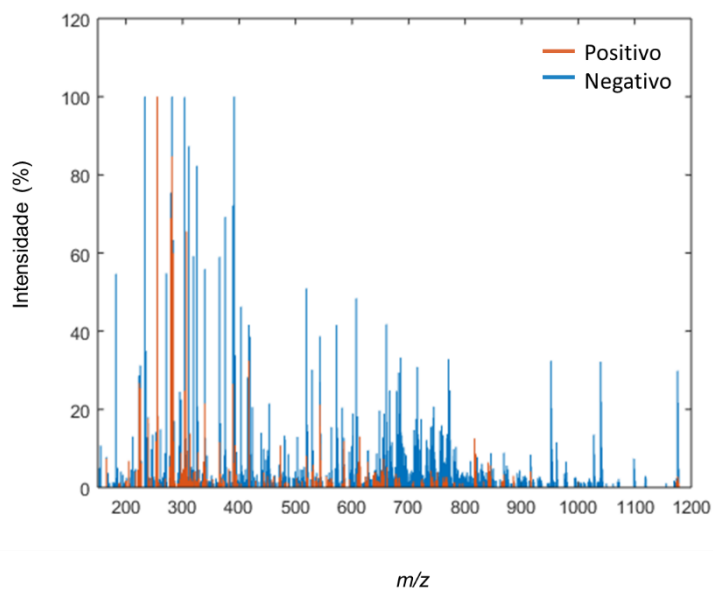
lipídicas, incluindo fosfatidilserina, fosfatidilglicerol, lisofosfatidilglicerol, ácido fosfático, fosfatidilcolinas, fosfatidiletanolaminas, diacilglicerol e ceramida.^{28,30}

Figura3. 1. Espectros de ESI(±)-Orbitrap MS de extratos lipídicos do conjunto de amostras de soro humano após preparo de amostra para purificação de lipídios contendo (A) 239 amostras para ESI(+) e (B) 232 para ESI(-).

A) ESI(+) – Orbitrap MS



B) ESI(-) – Orbitrap MS



Após analisar as amostras por ESI(±)-Orbitrap MS, que incluíram 239 amostras para ESI(+) e 232 amostras para ESI(-), uma análise multivariada foi

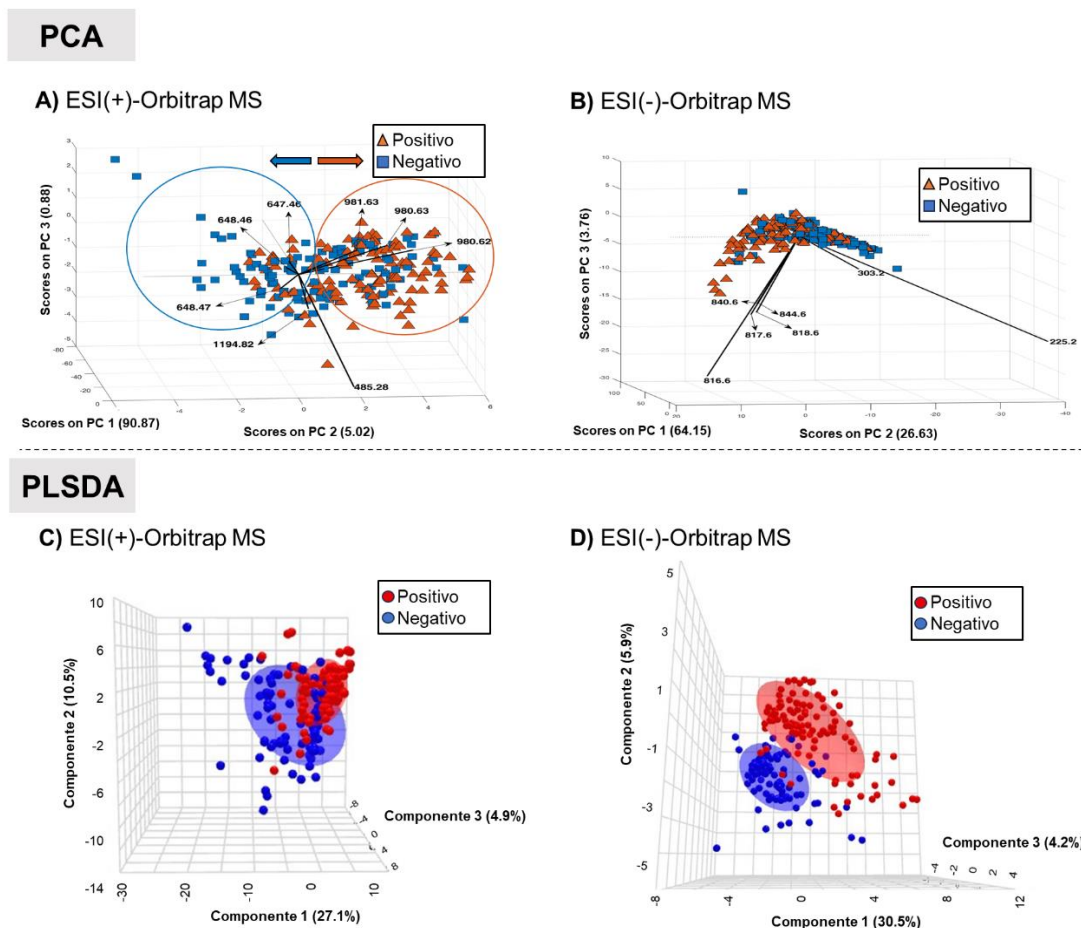
realizada para diferenciar entre pacientes com resultados positivos e negativos para a COVID-19 com base em seus perfis lipídicos. Inicialmente, uma investigação preliminar foi conduzida usando PCA.

Para os dados de ESI(+)-Orbitrap MS, não foi observada uma separação clara usando o modelo PCA, indicando uma mistura de amostras positivas e negativas para COVID-19 em todo o gráfico (Figura3. 2A). Em contraste, para os dados de ESI(-)-Orbitrap MS, o modelo PCA apresentou um bom desempenho na separação de amostras positivas e negativas para SARS-CoV-2, com 94,54% da variância total explicada por PC1 x PC2 x PC3. Os gráficos de *scores* e *loadings* de PC1 x PC2 x PC3 são mostrados na Figura3. 2B. A partir dos escores, é possível observar uma tendência de separação entre os dois principais grupos, com a influência de m/z 816,576 e 840,576 em amostras positivas e m/z 303,233 e 225,161 em amostras negativas.

Portanto, para visualizar melhor a separação de amostras de soro com resultados positivos e negativos para COVID-19, foi empregada a PLS-DA. Procedimentos para identificação e remoção de outliers (antes e após o pré-processamento dos dados) resultaram em um conjunto de dados composto por 220 amostras no modo positivo de ionização e 213 amostras no modo negativo.

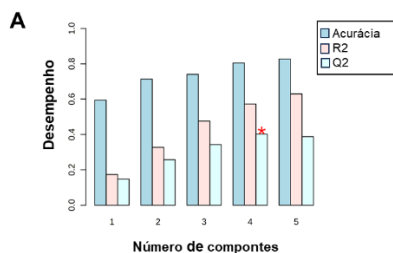
Além disso, a aplicação do discriminante de Fisher para seleção de variáveis atribuiu 35% FD ao conjunto de dados no modo positivo de ionização (138 variáveis) e 25% FD ao conjunto de dados no modo de negativo (416 variáveis). Como mostrado na Figura3. 2, o gráfico 3D de PLS-DA diferenciou claramente amostras de soro de indivíduos infectados e não infectados com SARS-CoV-2 no modo negativo, com uma variância total explicada por três componentes de 40,6%. No modo positivo de ionização, houve uma tendência de separação, como mostrado na Figura3. 2C, com 42,5% da variância total explicada por três componentes.

Figura3. 2. Gráficos biplot de PCA dos dados de ESI(±)-Orbitrap MS de extratos lipídicos de amostras de soro positivas ou negativas para SARS-CoV-2: (A) ESI(+) e (B) ESI; e gráficos de scores das componentes da PLS-DA



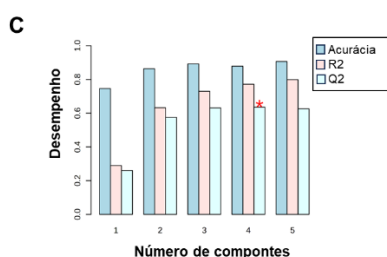
O desempenho desses modelos foi avaliado por meio de validação cruzada e teste de permutação 100 vezes. A verificação dupla dos modelos foi utilizada como recurso de validação estatística do PLS-DA. Foi observada uma tendência de agrupamento aparente com uma precisão de 80,42% e Q2 de 0,40 para ESI(+)-Orbitrap MS e uma precisão de 88,81% e Q2 de 0,65 para ESI(-)-Orbitrap MS (Figura3. 3 A-B). Como resultado, o conjunto de dados de ESI(-)-Orbitrap MS apresentou uma melhor resposta na classificação de amostras como positivas ou negativas para COVID-19 usando o modelo PLS-DA, atingindo uma precisão de 88,81% e mostrando uma tendência na separação de amostras pelo método não supervisionado PCA.

Figura3. 3. Parâmetros de validação cruzada para o modelo PLS-DA. R2, e Q2 representam as estimativas pela validação cruzada. A-B) ESI(+)-Orbitrap MS. C-D) ESI(-)-Orbitrap MS.



B

	1 comps	2 comps	3 comps	4 comps	5 comps
Acurácia	0.59494	0.71335	0.74043	0.80419	0.82713
R2	0.17307	0.32688	0.47543	0.57235	0.62934
Q2	0.1475	0.25726	0.3424	0.402	0.38777

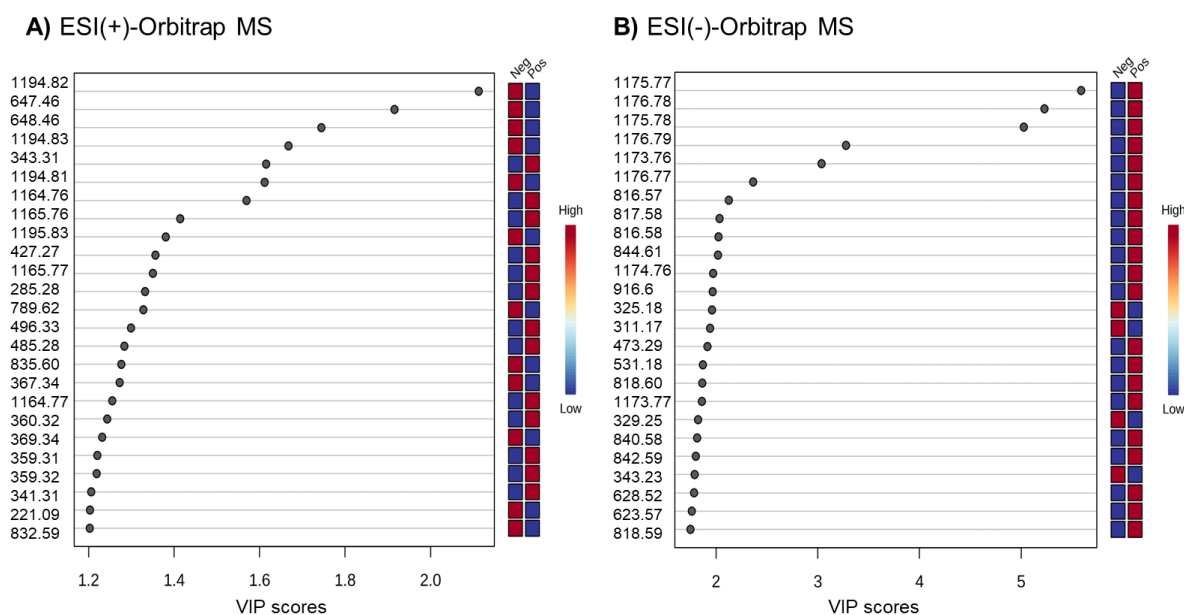


D

	1 comps	2 comps	3 comps	4 comps	5 comps
Acurácia	0.74667	0.88399	0.88822	0.88811	0.89298
R2	0.2895	0.6326	0.73004	0.77151	0.7985
Q2	0.2667	0.57843	0.63866	0.64762	0.62585

Os sinais que contribuíram para a separação de classes e discriminação entre os dois grupos foram relatados com base nos VIP scores extraídos pelo PLS-DA. Os 25 principais sinais com os maiores VIP score (score VIP $\geq 1,2$ no modo positivo de ionização e $\geq 1,8$ no modo negativo) são apresentados na Figura3. 4A para o modo de ionização positiva e na Figura3. 4B para o modo de ionização negativa. No modo de ionização negativa, os sinais com valores de m/z mais altos foram mais relevantes nas amostras positivas para COVID-19, corroborando as descobertas no modelo PCA.

Figura3. 4. Sinais m/z discriminantes identificadas de acordo com a pontuação de Importância Variável na Projeção (VIP) (A) para dados de ESI(-) e (B) ESI(+)-Orbitrap MS.



Posteriormente, a análise *machine learning* (ML) foi empregada e foram construídos distintos modelos de aprendizado de máquina SVM para classificação de amostras positivas e negativas para COVID-19.

Os melhores modelos SVM propostos demonstraram uma precisão de 89,66%, sensibilidade de 82,14% e especificidade de 96,67% para os dados de ESI(+)-Orbitrap MS, com a seleção de 138 variáveis (FD de 35%). Em contraste, para os dados de ESI(-)-Orbitrap MS, os resultados foram os seguintes: precisão de 98,44%, sensibilidade de 100% e especificidade de 96,88%, utilizando 416 variáveis selecionadas (FD de 25%) (Tabela3. 2). A taxa de falso negativo (TFN) foi de 0,00% para ESI(+)-Orbitrap MS e 3,33% para ESI(-)-Orbitrap MS.

Tabela3. 2. Desempenho de Parâmetros de Modelos SVM com dados de ESI(±)-Orbitrap MS de amostra de soro humano, após o preparo de amostra para purificação de lipídios, com Discriminante de Fisher (FD).

Modelo	FD	Pré-trat	grupo	VP	VN	FP	FN	SENS (%)	ESPEC (%)	ACC (%)	VPP (%)	VPN (%)	TFP (%)	TFN (%)
ESI(-)	25% (416 variáveis)	autoescalar e normalização na média	cal	74	75	0	0	100.00	100.00	100.00	100.00	100.00	0.00	0.00
			test	32	31	1	0	100.00	96.88	98.44	96.97	100.00	3.13	0.00
ESI(+)	35% (138 variáveis)	Nenhum	cal	84	78	0	0	100.00	100.00	100.00	100.00	100.00	0.00	0.00
			test	29	23	5	1	96.67	82.14	89.66	85.29	95.83	17.86	3.33

FD (Discriminante de Fisher); VP (verdadeiro positivo); VN (verdadeiro negativo); FP (falso positivo); FN (falso negativo); SENS (sensibilidade); ACC (acurácia); SPEC (especificidade); VPP (valor preditivo positivo); VPN (valor preditivo negativo); TFP (taxa de falsos positivos); TFN (taxa de falso negativo).

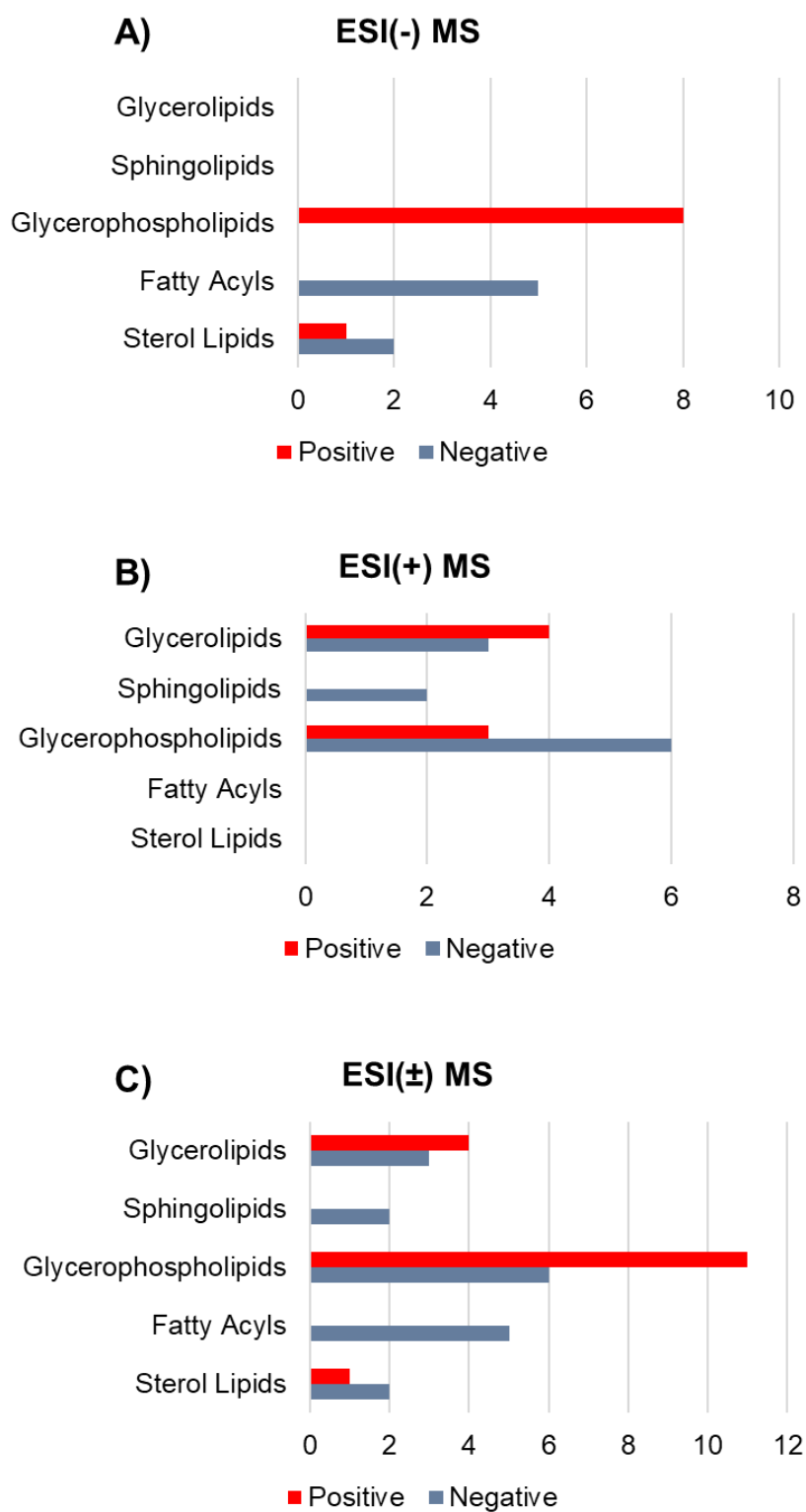
Os modelos PLS-DA mais relevantes foram obtidos com quatro componentes, valor Q2 de 0,65, e 0,40, acurácia de 88,81% e 80,42% no modo ESI(-) e ESI(+), respectivamente (Figura 3. 3).

Por outro lado, os modelos SVM não lineares apresentaram excelente desempenho em ambos os modos de ionização, possibilitando a discriminação das classes positivas e negativas para a COVID-19. Os resultados mostraram uma TFN de 3,33% e 0,00% para ESI(+)-Orbitrap MS e ESI(-)-Orbitrap MS, respectivamente (Tabela 3. 2). Isso destaca a eficiência da metodologia utilizada na distinção de amostras positivas para o SARS-CoV-2 de amostras negativas pela análise do perfil lipídico de amostras de soro.

Para interpretar os 50 sinais mais relevantes, considerando os maiores valores de VIP, foram utilizadas as plataformas LIPIDMAPS,³¹ Swisslipid³² e HMDB.³³ As classes lipídicas identificadas são apresentadas na Figura 3.5. Foi possível identificar aproximadamente 30% dos 50 sinais mais relevantes para discriminar amostras positivas e negativas para COVID-19 em ambos os modos de ionização. Houve 16 sinais nas análises ESI(-)-Orbitrap MS (Tabela 1 Anexo 1) e 18 sinais nas análises ESI(+)-Orbitrap MS (Tabela 2 Anexo 2 S3).

Nos dados do ESI(-)-Orbitrap MS, dos 16 lipídios identificados, 9 foram mais relevantes nas amostras positivas para COVID-19, pertencentes aos grupos Glicerofosfolipídios e Esterol Lipídios. Por outro lado, nas amostras analisadas pelo ESI(+)-Orbitrap MS, dos 18 lipídios identificados, 7 foram mais relevantes nas amostras positivas para COVID-19, pertencentes aos grupos Glicerofosfolipídeos e Glicerolipídeos (Figura 3. 5).

Figura3. 5. Classes lipídicas detectadas entre os 100 principais sinais VIP para dados de (A) ESI(+) MS, (B) ESI(-) MS e (C) ESI(±) MS. Estes são relevantes para distinguir entre classes positivas e negativas para a COVID-19.



Avaliando ambos os modos de ionização, observamos a detecção de Glicerofosfolipídios, Glicerolipídios e Esterol Lipídios em amostras positivas para COVID-19, e Glicerofosfolipídios, Glicerolipídios, Esterol Lipídios, Esfingolipídios e Acilos Graxos em amostras negativas para COVID-19. Apesar dos glicerofosfolipídios serem discriminativos em ambas as classes de amostras, as classes lipídicas diferiram. Os glicerofosfolipídios observados nas amostras negativas foram fosfatidilcolinas (PC), enquanto nas amostras positivas foram fosfatidilserinas (PS).

Wu et al. (2020) observaram que as fosfatidilcolinas (PCs) foram gradualmente reduzidas em casos fatais de COVID-19. A diminuição na abundância de PC no plasma de pacientes com COVID-19 revelou α -citidiltransferase colina aberrante em macrófagos, que utilizaram PC para gerar CDP-colina, a molécula essencial para diversas funções celulares. Um padrão semelhante foi observado em pacientes com doença por vírus Ebola (DVE), onde a disfunção hepática e a deterioração do metabolismo da colina afetaram a síntese de LPCs e PCs e foram associadas à gravidade da doença.

Li et al. (2021), além de observarem diminuição do PC, também observaram diminuição do PG e aumento dos lipídios IP, PS ao comparar grupos controle com grupos positivos para COVID-19. PS é um componente importante das micropartículas plaquetárias pró-coagulantes secretadas por plaquetas ativadas. Níveis aumentados de PS no plasma durante a infecção por SARS-CoV-2 indicaram maior ativação plaquetária nesses pacientes. Os lipídios desempenham um papel importante no bom funcionamento do corpo. Na homeostase, um desequilíbrio de lipídios pode levar a vários tipos de condições de doença no hospedeiro. Além disso, os lipídios estão relacionados ao metabolismo anormal de várias doenças, incluindo câncer, doenças neurodegenerativas, doenças cardiovasculares, diabetes, obesidade e outras doenças crônicas.³⁶ Resultados encontrados na literatura já relataram mudanças no perfil lipídico do soro, em diferentes doenças e com a COVID-19 não seria diferente.

Shen et al. (2020)³⁵ relataram a diminuição de esfingolipídios em pacientes com COVID-19, graves e não graves, e uma diminuição contínua de glicerofosfolipídios (GPL) após infecção por SARS-CoV-2.³⁵ Esfingolipídios e glicerofosfolipídios são componentes importantes das biomembranas, regulam

diversos processos, incluindo regulação do crescimento, migração celular, adesão, apoptose, senescência e respostas inflamatórias.³⁷ Janneh e colaboradores (2021)³⁸ também observaram uma redução de quase 15 vezes em níveis séricos de esfingosina em pacientes sintomáticos com COVID-19, em comparação aos níveis em pacientes assintomáticos.³⁸

Outro estudo que também observou as alterações lipidômicas séricas de pacientes com COVID-19 foi de Wu e colaboradores (2020),³⁴ que identificaram uma maior abundância dos grupos diglicerídeos (DGs), ácidos graxos livres (FAAs) e triglicerídeos (TGs) no grupo de fatalidade, bem como aumento das abundâncias relativas desses lipídios com a deterioração da doença. Em contraste, as fosfatidilcolinas (PC) foram encontradas como lipídios gradualmente reduzidos ao longo dos casos fatais por COVID-19.³⁴

A dislipidemia, uma doença caracterizada por anomalias nos níveis de lipídios no sangue, principalmente do colesterol total e dos TGs, foi observada em pacientes curados de COVID-19 devido à infecção viral e ao tratamento médico no estudo de Bai e colaboradores (2021).³⁹ A lipidômica não direcionada para caracterização lipídica, usando UPLC com o Orbitrap MS e ferramentas de bioinformática indicaram que a alteração lipidômica foi drasticamente alterada, mesmo nos pacientes que atingiram os padrões de alta.³⁹

Além da identificação das alterações dos perfis lipídicos, estudos estão sendo realizados para classificar e prever amostras positivas para a COVID-19. Cobre *et al.* (2022)⁴⁰ relataram que o modelo PLS-DA desenvolvido com amostras de soro (n = 96) e plasma (n = 96) de pacientes com COVID-19 (doença aguda, grave e fatal), pode ajudar a prever o diagnóstico, a gravidade e a fatalidade de SARS-CoV-2 na prática diária, com taxas de precisão de 93%, 94% e 97%, respectivamente. Além disso, alguns biomarcadores normalmente com altos níveis no soro ou plasma de pacientes com COVID-19 (ou seja, ribotimidina, N-acetil-glucosamina-1-fosfato, l-ornitina, 5,6-di-hidro-5-metiluracil) devem ser avaliados como indicadores prognósticos da doença.⁴⁰

Uma metodologia para detecção de infecção assintomática por SARS-CoV-2 utilizando padrões metabólicos séricos analisados por MALDI-TOF MS combinado com aprendizado de máquina foi proposta. O modelo proporcionou uma alta precisão de 93,4%, com apenas 5% de falsos negativos e 7% de

falsos positivos. Além disso, identificaram 10 metabólitos e lipídios como potenciais biomarcadores para infecção assintomática por SARS-CoV-2.¹³

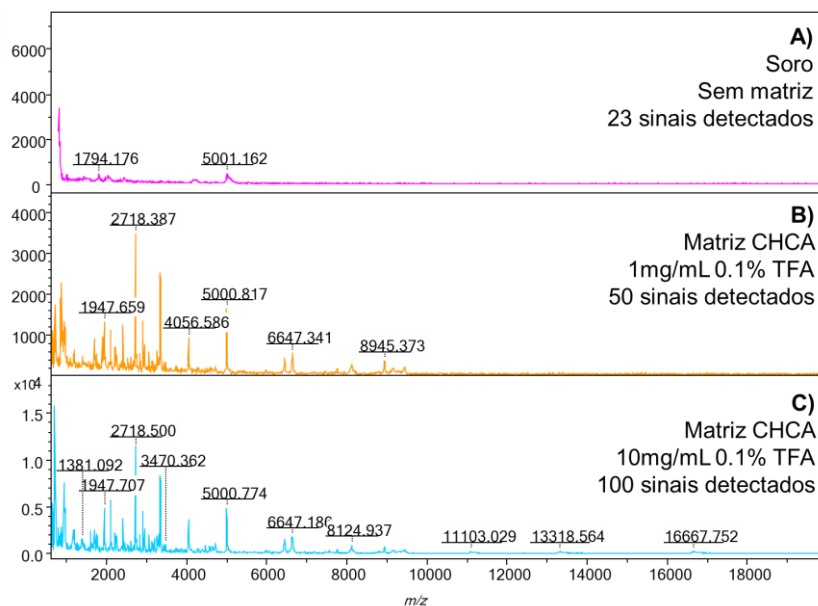
Portanto, a estratégia de estudar a lipidômica sérica por técnicas de aquisição rápida e de alta resolução, obtendo uma TFR de 0,00% pode não apenas auxiliar a classificar amostras positivas para COVID-19, mas também na identificação de biomarcadores potenciais da doença em estudos futuros.

Proteômica

Para análise de soro por MALDI(+)-TOF MS é necessário o preparo de amostras, utilizando uma *Spin Column* C18. A combinação da tecnologia *Spin Column* C18 e análise MALDI(+)-TOF MS nos permite uma limpeza de interferentes (dessalinização) e concentração dos analitos de interesse, permitindo gerar maior número de componentes espectrais (sinais de m/z) de interesse nos espectros de massas das amostras de soro.^{14,41}

A eficiência de dessorção e ionização das moléculas depende da seleção do composto utilizado como matriz e de sua concentração. Adicionalmente, a inclusão de um ácido na solução da matriz pode facilitar a ionização.⁴² A avaliação da concentração da matriz pode ser observada na Figura 3. 6, onde o uso de 10 mg/mL de CHCA permite a detecção de um maior número de sinais e proporciona maior sensibilidade. Isso é demonstrado pela intensidade absoluta do íon mais intenso, m/z 2718,500, que apresentou um valor mais alto de corrente total de íons (TIC) ($1,5 \times 10^4$), em comparação com o valor de TIC ($3,4 \times 10^3$) do íon mais intenso, m/z 2718,387, no espectro quando 1 mg/mL de CHCA foi utilizado.

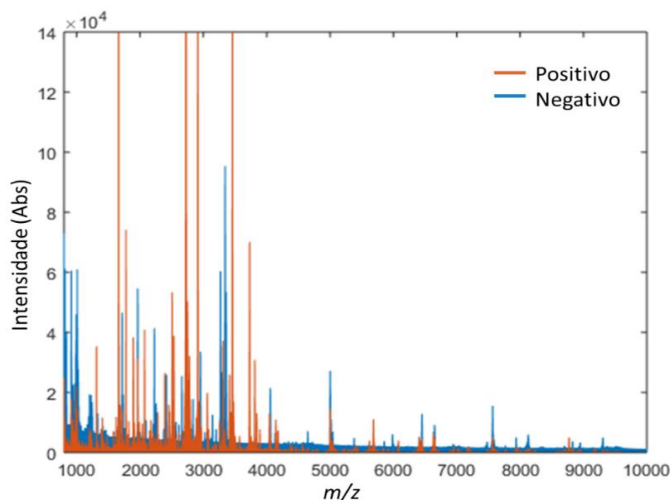
Figura3. 6. Espectros de MALDI(+)-TOF MS de amostra de soro humano (A) sem matriz (B); contendo matriz CHCA à 1 mg/mL (C) e à 10mg/mL.



Uma comparação entre a porcentagem de TFA utilizada no preparo da matriz foi observada após 31 amostras em triplicata, de um total de 100 amostras, não ionizarem no modo automático das análises, utilizando a matriz CHCA na concentração de 10 mg/mL 0,1% TFA. Dessa forma, o aumento da porcentagem de TFA na matriz para 2,5%, já descrito na literatura por Lazari *et al.*(2021),¹⁴ foi necessário e suficiente para ionizar e analisar todas as amostras em triplicata.

Após preparo e otimização das análises, foram analisadas 300 amostras (150 positivas para COVID-19 e 150 negativas para COVID-19) por MALDI(+)-TOF MS. Os espectros obtidos estão mostrados na Figura3. 7.

Figura3. 7. 900 espectros de massas MALDI(+)-TOF MS de 300 amostras de soro humano após preparo de amostra utilizando Spin Column C18, sendo que 450 correspondem ao soro positivo pra COVID-19 (vermelho) e 450 ao soro negativo para COVID-19 (azul).



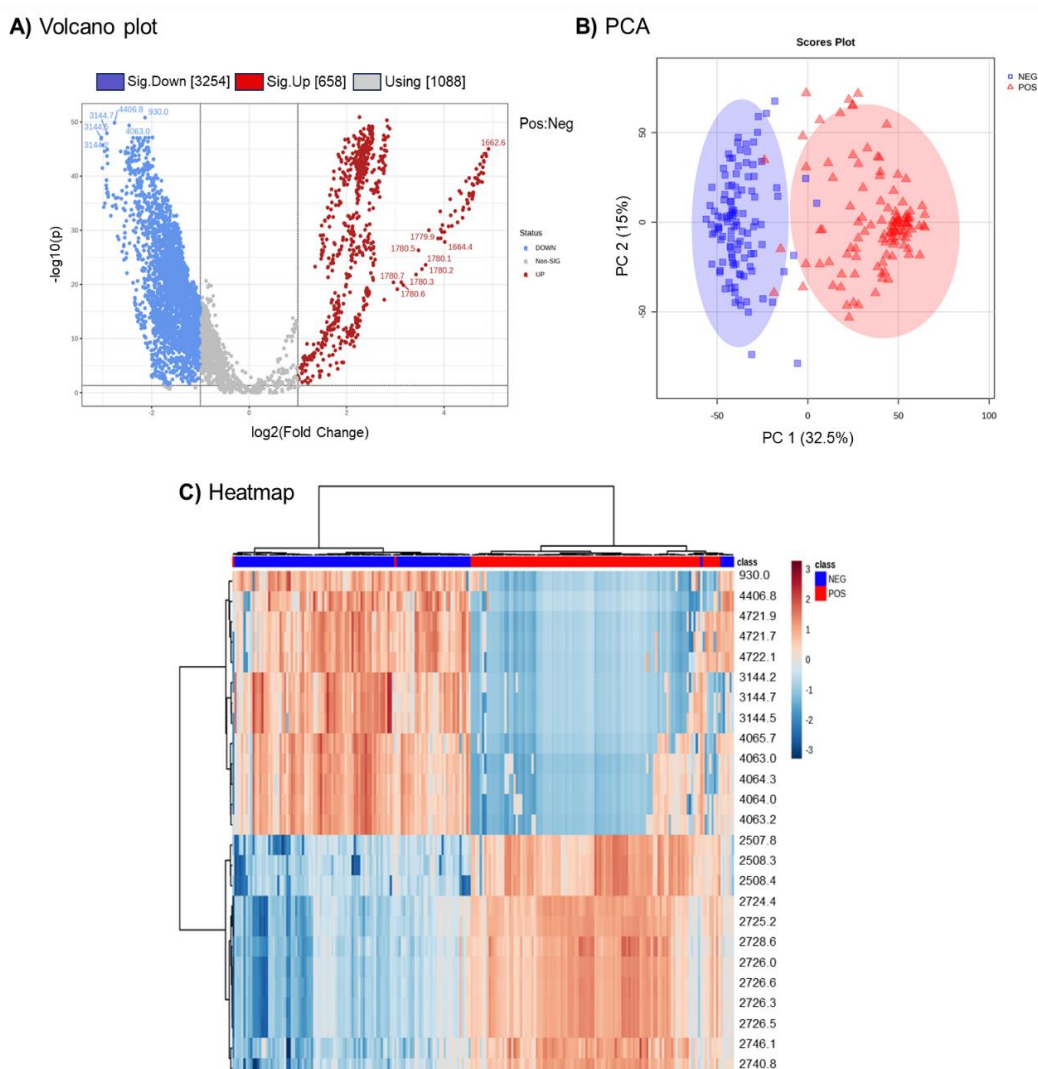
Os espectros de MALDI(+)-TOF MS obtidos foram então submetidos à alinhamento de sinais, seguido pela construção de modelos univariados e multivariados. A maioria dos sinais estavam localizados entre m/z 900 e 5.000. No total, 37018 sinais foram detectados após o processamento de dados de todos os arquivos brutos de MS.

A Figura3. 8 retrata a alta diversidade de sinais entre amostras controle e amostras positivas para COVID-19. A diferença significativa na abundância relativa entre os dois conjuntos de amostras por meio de contagens de espectro pode ser observada no gráfico *volcano plot*, em que o eixo x representa o \log_2 (*Fold change* (FC)) e o eixo y $-\log_{10}$ (p). Entre todos os sinais detectados, 3912 (78,24%) apresentaram diferença estatisticamente significativa em suas intensidades relativas entre as amostras, $p < 0.05$ e $FC \geq 2$, com 658 (16,82%) apresentando expressão elevada em pacientes positivos para SARS-CoV-2 e 3254 (83,18%) apresentando expressão elevada em pacientes negativo para COVID-19 (Figura3. 8A).

A fim de observar a discriminação dos dois grupos de amostras (pacientes positivos e negativos para a COVID-19, a PCA foi realizada nos perfis proteômicos sérico coletados. Tal como mostrado na Figura3. 8B, o conjunto de dados pode ser claramente discriminado, separando amostras de

pacientes controle de pacientes com COVID-19, com variância das componentes principais (PC): PC1 = 32,5%, PC2 = 15%.

Figura3. 8. Análise estatística dados MALDI(+)-TOF MS. (A) Volcano plot indicando os sinais presentes no conjunto de amostras, com diferença estatística de acordo com as contagens do espectro, cada mancha representa um sinal. No eixo y, Log2 (valor p), no eixo x, Log2 (FC), os valores negativos indicam sinais mais relevantes em amostras negativas e valores positivos mostram sinais mais relevantes em amostras positivas. (B) Gráfico de PCA de espectros MALDI(+)-TOF MS de amostras de soro positivas e negativas para SARS-CoV-2. (C) Heatmap de espécies moleculares diferenciais em pacientes com COVID-19 em comparação com pacientes negativos.

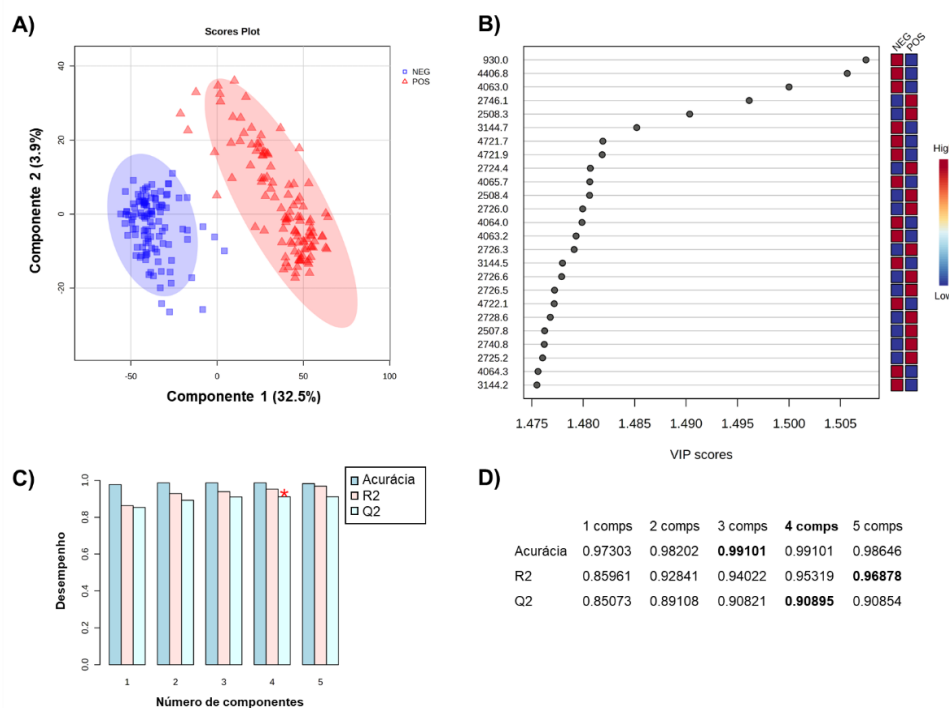


As diferenças proteômicas plasmáticas entre os dois grupos foram comparadas usando gráficos de *scores* de PLS-DA supervisionado para investigar os dados do agrupamento entre pacientes positivos para a COVID-19 e grupo controle (Figura3. 9). Os modelos PLS-DA foram estabelecidos após a normalização na média, transformação logarítmica e o método de escalonamento de Pareto. A validação cruzada e teste de permutação de 100 vezes foram realizados no PLS-DA para avaliar a qualidade dos modelos. Uma aparente tendência de agrupamento foi observada com acurácia de 99,10% e Q2 de 0,91, que indicaram perfis proteômicos plasmáticos distintos em resposta à infecção pelo vírus COVID-19 (Figura3. 9).

A medida de VIP foi utilizada para identificar os sinais mais discriminantes caracterizando os dois grupos. Os 25 sinais *m/z* com maior importância variável na projeção, $VIP > 1.4$, são destacados na Figura3. 9.

Para demonstrar ainda mais a classificação, foi realizado o *Heatmap* de análise de agrupamento nos perfis proteômicos sérico coletados, que mostrou a separação das amostras positivas e negativas para SARS-CoV-2 (Figura3. 8C). Sendo os sinais na faixa de *m/z* 3000-4000 relevante para as amostras negativas e os de *m/z* 2500-2800 nas amostras positivas para COVID-19 analisadas por MALDI(+)TOF MS.

Figura3. 9. Análise estatística de dados MALDI(+)-TOF MS. (A) O gráfico de scores de PLS-DA ilustra a segregação do perfil proteômico entre pacientes positivos para COVID-19 e controles. (B) Sinais m/z discriminantes baseados em VIP scores. Os 25 principais valores m/z com VIP scores $>1,4$. (C-D) Parâmetros de validação cruzada para PLS-DA, incluindo resultados R2 e Q2 estimados através de validação cruzada.



Além de modelos lineares, modelos não lineares de aprendizagem de máquina como SVM, foram desenvolvidos para analisar os dados na faixa de m/z trabalhada.

Dos modelos construídos, dois modelos apresentaram os melhores resultados. O modelo 1 com todas as variáveis (37018 variáveis) e o modelo 2 selecionando as variáveis com um 70 % do FD (1139 variáveis). Ao treinar os diferentes modelos e selecionar variáveis com FD acima de 70%, o modelo 2 apresentou desempenho similar ao modelo sem seleção de variáveis, ou seja, utilizando toda a faixa de m/z , com sensibilidade de 100%, especificidade de 100% e acurácia de 100% para as amostras de calibração e sensibilidade de 100%, especificidade de 97,78% e acurácia de 98,88% para as amostras de teste (Tabela3. 3).

Tabela3. 3. Desempenho de Parâmetros de Modelos SVM com espectros de massa MALDI(+)-TOF MS de amostra de soro humano após preparo e otimização das análises.

Modelo	FD	grupo	VP	VN	FP	FN	SENS (%)	ESPEC (%)	ACC (%)	VPP (%)	VNP (%)	TFP (%)	TFN (%)
1	Full (37018 variáveis)	Cal	315	314	0	0	100.00	100.00	100.00	100.00	100.00	0.00	0.00
		Test	134	133	2	0	100.00	97.78	98.88	97.81	100.00	2.22	0.00
2	70% (1139 variáveis)	Cal	315	314	0	0	100.00	100.00	100.00	100.00	100.00	0.00	0.00
		Test	134	132	3	0	100.00	97.78	98.88	97.81	100.00	2.22	0.00

FD (Discriminante de Fisher); VP (verdadeiro positivo); VN (verdadeiro negativo); FP (falso positivo); FN (falso negativo); SENS (sensibilidade); ACC (acurácia); ESPEC (especificidade); VPP (valor positivo da prevalência); VNP (valor negativo da prevalência); TFP (taxa de falsos positivos); TFN (taxa de falso negativo).

Os modelos SVM apresentaram excelente desempenho, sendo possível discriminar as classes positivas e negativas com uma TFN de 0,0% (Tabela 3. 2), tornando a metodologia utilizada eficiente em distinguir amostras de soro positivas para SARS-CoV-2 de negativas e evitando diagnósticos falso-negativos. Entretanto, os dados das análises de soro por MALDI(+)-TOF MS apresentam uma melhor acurácia nos modelos lineares, PLS-DA, tendo uma acurácia de 99,01%.

Estudos já publicados na literatura vêm mostrando a eficiência em classificar amostras positivas para COVID-19 utilizando dados de análises de biofluidos por MALDI MS combinados com técnicas ML. O primeiro estudo foi de Nachtigall e colaboradores (2020),¹⁵ que obtiveram espectros de massas de um total de 362 amostras de swab nasal sem preparação prévia. O modelo de SVM forneceu acurácia de 93,9%, com 7% de falsos positivos e 5% de falsos negativos.¹⁵

Um estudo preliminar semelhante de 311 amostras de swab nasal de pacientes foi conduzido por Rocca *et al.* (2020).¹⁶ Uma acurácia de aproximadamente 68%, sensibilidade de 62% e especificidade de 72% foram relatadas, além disso observou-se que nenhum dos sinais encontrados poderia ser molecularmente atribuído a proteínas específicas do vírus e provavelmente a algum tipo de componente viral do hospedeiro.¹⁶

Outro estudo utilizando 226 amostras de swab nasal e análises por MALDI MS realizado por Tran e colaboradores (2021),¹⁷ também apresentou bons resultados em seus modelos. Dois modelos de ML otimizados exibiram acurácia de 98,3%, concordância percentual positiva (PPA) de 100%, concordância percentual negativa (NPA) de 96% e acurácia de 96,6%, PPA de 98,5% e NPA de 94%.¹⁷

Apesar da coleta de swab ser mais rápida que a de soro, a periculosidade da presença do vírus é maior. A vantagem do uso de soro para diagnósticos e triagens se dá, por este já está presente nas rotinas de exames clínicos e apresentar ricas informações sobre o estado fisiológico ou patológico do paciente.

Lazari *et al.* (2021)¹⁴ obtiveram resultados otimizados com acurácia de 92%, sensibilidade de 93% e especificidade de 92% no conjunto de dados analisados de 117 amostras de soro de pacientes com COVID-19.¹⁴ O estudo de Wan *et al.* (2021)¹³ desenvolveu uma nova abordagem para detectar infecção assintomática por SARS-CoV-2 usando padrões metabólicos combinados com ML em 362 amostras de soro. Essa abordagem forneceu 93,4% de precisão com apenas 5% de falsos negativos e 7% de falsos positivos.¹³

Apesar dos excelentes resultados encontrados na literatura relacionando a técnica MALDI MS com ML em amostras biológicas de pacientes com suspeita de COVID-19, estes apresentaram uma taxa de falso negativo maior do que as encontradas por este estudo tanto pelas análises de MALDI(+)-TOF MS das amostras de soro, quanto pelas análises de lipídios por ESI(\pm)-Orbitrap MS.

Assim, estudos baseados em lipídios, peptídeos e proteínas usando soro de pacientes e técnicas espectrométricas de alto rendimento, prometem ser valiosos para a classificação de amostras positivas para COVID-19 e possível identificação de biomarcadores associados a doença.

3.4 CONCLUSÃO

Os modelos SVM, tanto para a análise de lipídios por ESI(-)-Orbitrap MS quanto para a análise de proteínas e peptídeos por MALDI(+)-TOF MS, são eficientes no diagnóstico da COVID-19 com uma taxa de Falso Negativo de 0,0% ao utilizar modelos ML não lineares. Embora a análise de lipídios por ESI(+)-Orbitrap MS apresente uma TFN de 3,3%, ainda está entre as taxas mais baixas relatadas na literatura.

Modelos lineares, como PLS-DA, forneceram resultados inferiores no conjunto de dados ESI(±)-Orbitrap MS; no entanto, tiveram um desempenho satisfatório nos resultados de MALDI(+)-TOF MS, alcançando uma precisão de 99,1%. Apesar dos resultados estatísticos superiores no conjunto de dados de análise proteômica, a análise de lipídios se destaca devido à sua preparação de amostras rápida, fácil e economicamente viável, juntamente com seus resultados promissores. Além disso, os lipídios desempenham um papel significativo no funcionamento do corpo.

As técnicas analíticas de MS empregadas neste estudo apresentaram alta sensibilidade e especificidade na detecção de valores de m/z associados a lipídios e proteínas em pacientes com COVID-19 quando combinadas com métodos estatísticos. A identificação potencial de biomarcadores de lipídios no futuro torna o uso de infusão direta ESI MS para triagem de pacientes e observação das diferenças nos perfis espectrais um estudo inovador e promissor na análise clínica.

3.5 REFERÊNCIAS

- (1) Spick, M.; Lewis, H. M.; Frampas, C. F.; Longman, K.; Costa, C.; Stewart, A.; Dunn-Walters, D.; Greener, D.; Evetts, G.; Wilde, M. J.; Sinclair, E.; Barran, P. E.; Skene, D. J.; Bailey, M. J. An Integrated Analysis and Comparison of Serum, Saliva and Sebum for COVID-19 Metabolomics. *Sci. Reports* **2022**, *12* (1), 1–12.
- (2) de Oliveira, M. A. L.; Watanabe, A. S. A.; Cesar, D. E.; Candido, J. M. B.; Lima, N. M.; Moreira, O. B. O.; Chellini, P. R. TESTES DIAGNÓSTICOS PARA O SARS-COV-2: UMA REFLEXÃO CRÍTICA. *Quim. Nova* **2022**, *45* (6), 760–766.
- (3) Ferrari, D.; Sabetta, E.; Ceriotti, D.; Motta, A.; Strollo, M.; Banfi, G.; Locatelli, M. Routine Blood Analysis Greatly Reduces the False-Negative Rate of RT-PCR Testing for COVID-19. *Acta Bio Medica Atenei Parm.* **2020**, *91* (3), e2020003.
- (4) Kim, J. Y.; Ko, J. H.; Kim, Y.; Kim, Y. J.; Kim, J. M.; Chung, Y. S.; Kim, H. M.; Han, M. G.; Kim, S. Y.; Chin, B. S. Viral Load Kinetics of SARS-CoV-2 Infection in First Two Patients in Korea. *J. Korean Med. Sci.* **2020**, *35* (7).
- (5) Tahamtan, A.; Ardebili, A. Real-Time RT-PCR in COVID-19 Detection: Issues Affecting the Results. *Expert Rev. Mol. Diagn.* **2020**, *20* (5), 453–454.
- (6) Chen, M.; Qin, R.; Jiang, M.; Yang, Z.; Wen, W.; Li, J. Clinical Applications of Detecting IgG, IgM or IgA Antibody for the Diagnosis of COVID-19: A Meta-Analysis and Systematic Review. *Int. J. Infect. Dis.* **2021**, *104*, 415–422.
- (7) Peeling, R. W.; Wedderburn, C. J.; Garcia, P. J.; Boeras, D.; Fongwen, N.; Nkengasong, J.; Sall, A.; Tanuri, A.; Heymann, D. L. Serology Testing in the COVID-19 Pandemic Response. *Lancet. Infect. Dis.* **2020**, *20* (9), e245.
- (8) Wang, W.; Xu, Y.; Gao, R.; Lu, R.; Han, K.; Wu, G.; Tan, W. Detection of SARS-CoV-2 in Different Types of Clinical Specimens. *JAMA* **2020**, *323* (18), 1843–1844.
- (9) Hasan, M. R.; Suleiman, M.; Pérez-López, A. Metabolomics in the Diagnosis and Prognosis of COVID-19. *Front. Genet.* **2021**, *12*, 1358.
- (10) Gomila, R. M.; Martorell, G.; Fraile-Ribot, P. A.; Doménech-Sánchez, A.; Albertí, M.; Oliver, A.; García-Gasalla, M.; Albertí, S. Use of Matrix-Assisted Laser Desorption Ionization Time-of-Flight Mass Spectrometry Analysis of Serum Peptidome to Classify and Predict Coronavirus Disease 2019 Severity.

- Open Forum Infect. Dis.* **2021**, *8* (6), 1–8.
- (11) Zhou, K.; Sun, Y.; Li, L.; Zang, Z.; Wang, J.; Li, J.; Liang, J.; Zhang, F.; Zhang, Q.; Ge, W.; Chen, H.; Sun, X.; Yue, L.; Wu, X.; Shen, B.; Xu, J.; Zhu, H.; Chen, S.; Yang, H.; Huang, S.; Peng, M.; Lv, D.; Zhang, C.; Zhao, H.; Hong, L.; Zhou, Z.; Chen, H.; Dong, X.; Tu, C.; Li, M.; Zhu, Y.; Chen, B.; Li, S. Z.; Guo, T. Eleven Routine Clinical Features Predict COVID-19 Severity Uncovered by Machine Learning of Longitudinal Measurements. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 3640–3649.
- (12) Lima, N. M.; Fernandes, B. L. M.; Alves, G. F.; de Souza, J. C. Q.; Siqueira, M. M.; Patrícia do Nascimento, M.; Moreira, O. B. O.; Sussulini, A.; de Oliveira, M. A. L. Mass Spectrometry Applied to Diagnosis, Prognosis, and Therapeutic Targets Identification for the Novel Coronavirus SARS-CoV-2: A Review. *Anal. Chim. Acta* **2022**, *1195*, 339385.
- (13) Wan, Q.; Chen, M.; Zhang, Z.; Yuan, Y.; Wang, H.; Hao, Y.; Nie, W.; Wu, L.; Chen, S. Machine Learning of Serum Metabolic Patterns Encodes Asymptomatic SARS-CoV-2 Infection. *Front. Chem.* **2021**, *9*, 746134.
- (14) Lazari, L. C.; de Rose Ghilardi, F.; Rosa-Fernandes, L.; Assis, D. M.; Nicolau, J. C.; Santiago, V. F.; Dalçóquio, T. F.; Angeli, C. B.; Bertolin, A. J.; Marinho, C. R. F.; Wrenger, C.; Durigon, E. L.; Siciliano, R. F.; Palmisano, G. Prognostic Accuracy of MALDI-TOF Mass Spectrometric Analysis of Plasma in COVID-19. *Life Sci. Alliance* **2021**, *4* (8), 1–12.
- (15) Nachtigall, F. M.; Pereira, A.; Trofymchuk, O. S.; Santos, L. S. Detection of SARS-CoV-2 in Nasal Swabs Using MALDI-MS. *Nat. Biotechnol.* **2020**, *38* (10), 1168–1173.
- (16) Rocca, M. F.; Zintgraff, J. C.; Dattero, M. E.; Santos, L. S.; Ledesma, M.; Vay, C.; Prieto, M.; Benedetti, E.; Avaro, M.; Russo, M.; Nachtigall, F. M.; Baumeister, E. A Combined Approach of MALDI-TOF Mass Spectrometry and Multivariate Analysis as a Potential Tool for the Detection of SARS-CoV-2 Virus in Nasopharyngeal Swabs. *J. Virol. Methods* **2020**, *286*, 113991.
- (17) Tran, N. K.; Howard, T.; Walsh, R.; Pepper, J.; Loegering, J.; Phinney, B.; Salemi, M. R.; Rashidi, H. H. Novel Application of Automated Machine Learning with MALDI-TOF-MS for Rapid High-Throughput Screening of COVID-19: A Proof of Concept. *Sci. Reports 2021 111* **2021**, *11* (1), 1–10.
- (18) De Almeida, C. M.; Motta, L. C.; Folli, G. S.; Marcarini, W. D.; Costa, C. A.;

- Vilela, A. C. S.; Barauna, V. G.; Martin, F. L.; Singh, M. N.; Campos, L. C. G.; Costa, N. L.; Vassallo, P. F.; Chaves, A. R.; Endringer, D. C.; Mill, J. G.; Filgueiras, P. R.; Romão, W. MALDI(+) FT-ICR Mass Spectrometry (MS) Combined with Machine Learning toward Saliva-Based Diagnostic Screening for COVID-19. *J. Proteome Res.* **2022**, *21* (8), 1868–1875.
- (19) Dunn, W. B.; Ellis, D. I. Metabolomics: Current Analytical Platforms and Methodologies. *TrAC - Trends Anal. Chem.* **2005**, *24* (4), 285–294.
- (20) González-Domínguez, R.; Sayago, A.; Fernández-Recamales, Á. Direct Infusion Mass Spectrometry for Metabolomic Phenotyping of Diseases. *Bioanalysis* **2017**, *9* (1), 131–148.
- (21) Anand, S.; Young, S. A.; Esplin, M. S.; Peadar, B.; Tolley, H. D.; Porter, T. F.; Varner, M. W.; D'Alton, M. E.; Jackson, B. J.; Graves, S. W. Detection and Confirmation of Serum Lipid Biomarkers for Preeclampsia Using Direct Infusion Mass Spectrometry. *J. Lipid Res.* **2016**, *57* (4), 687–696.
- (22) Draper, J.; Lloyd, A. J.; Goodacre, R.; Beckmann, M.; Draper, J.; Lloyd, Á. A. J.; Beckmann, Á. M.; Goodacre, R. Flow Infusion Electrospray Ionisation Mass Spectrometry for High Throughput, Non-Targeted Metabolite Fingerprinting: A Review. *Metabolomics* **2012**, *9* (1), 4–29.
- (23) Deulofeu, M.; García-Cuesta, E.; Peña-Méndez, E. M.; Conde, J. E.; Jiménez-Romero, O.; Verdú, E.; Serrando, M. T.; Salvadó, V.; Boadas-Vaello, P. Detection of SARS-CoV-2 Infection in Human Nasopharyngeal Samples by Combining MALDI-TOF MS and Artificial Intelligence. *Front. Med.* **2021**, *8*, 398–410.
- (24) Bligh, E. G.; Dyer, W. J. A RAPID METHOD OF TOTAL LIPID EXTRACTION AND PURIFICATION. *Can. J. Biochem. Physiol.* **1959**, *37*, 911–917.
- (25) Pang, Z.; Zhou, G.; Ewald, J.; Chang, L.; Hacariz, O.; Basu, N.; Xia, J. Using MetaboAnalyst 5.0 for LC–HRMS Spectra Processing, Multi-Omics Integration and Covariate Adjustment of Global Metabolomics Data. *Nat. Protoc.* **2022**, *17* (8), 1735–1761.
- (26) Lovatti, B. P. O.; Nascimento, M. H. C.; Rainha, K. P.; Oliveira, E. C. S.; Neto, Á. C.; Castro, E. V. R.; Filgueiras, P. R. Different Strategies for the Use of Random Forest in NMR Spectra. *J. Chemom.* **2020**, *34* (12), e3231.
- (27) Kennard, R. W.; Stone, L. A. Computer Aided Design of Experiments. *Technometrics* **1969**, *11* (1), 137–148.

- (28) Schuhmann, K.; Almeida, R.; Baumert, M.; Herzog, R.; Bornstein, S. R.; Shevchenko, A. Shotgun Lipidomics on a LTQ Orbitrap Mass Spectrometer by Successive Switching between Acquisition Polarity Modes. *J. Mass Spectrom.* **2012**, *47* (1), 96–104.
- (29) Gallart-Ayala, H.; Courant, F.; Severe, S.; Antignac, J. P.; Morio, F.; Abadie, J.; Le Bizec, B. Versatile Lipid Profiling by Liquid Chromatography–High Resolution Mass Spectrometry Using All Ion Fragmentation and Polarity Switching. Preliminary Application for Serum Samples Phenotyping Related to Canine Mammary Cancer. *Anal. Chim. Acta* **2013**, *796*, 75–83.
- (30) Yamada, T.; Uchikata, T.; Sakamoto, S.; Yokoi, Y.; Fukusaki, E.; Bamba, T. Development of a Lipid Profiling System Using Reverse-Phase Liquid Chromatography Coupled to High-Resolution Mass Spectrometry with Rapid Polarity Switching and an Automated Lipid Identification Software. *J. Chromatogr. A* **2013**, *1292*, 211–218.
- (31) Sud, M.; Fahy, E.; Cotter, D.; Brown, A.; Dennis, E. A.; Glass, C. K.; Merrill, A. H.; Murphy, R. C.; Raetz, C. R. H.; Russell, D. W.; Subramaniam, S. LMSD: LIPID MAPS Structure Database.
- (32) Aimo, L.; Liechti, R.; Hyka-Nouspikel, N.; Niknejad, A.; Gleizes, A.; Götz, L.; Kuznetsov, D.; David, F. P. A.; Van Der Goot, F. G.; Riezman, H.; Bougueleret, L.; Xenarios, I.; Bridge, A. The SwissLipids Knowledgebase for Lipid Biology. *Bioinformatics* **2015**, *31* (17), 2860–2866.
- (33) Wishart, D. S.; Tzur, D.; Knox, C.; Eisner, R.; Guo, A. C.; Young, N.; Cheng, D.; Jewell, K.; Arndt, D.; Sawhney, S.; Fung, C.; Nikolai, L.; Lewis, M.; Coutouly, M. A.; Forsythe, I.; Tang, P.; Shrivastava, S.; Jeroncic, K.; Stothard, P.; Amegbey, G.; Block, D.; Hau, D. D.; Wagner, J.; Miniaci, J.; Clements, M.; Gebremedhin, M.; Guo, N.; Zhang, Y.; Duggan, G. E.; MacInnis, G. D.; Weljie, A. M.; Dowlatabadi, R.; Bamforth, F.; Clive, D.; Greiner, R.; Li, L.; Marrie, T.; Sykes, B. D.; Vogel, H. J.; Querengesser, L. HMDB: The Human Metabolome Database. *Nucleic Acids Res.* **2007**, *35* (suppl_1), D521–D526.
- (34) Wu, D.; Shu, T.; Yang, X.; Song, J. X.; Zhang, M.; Yao, C.; Liu, W.; Huang, M.; Yu, Y.; Yang, Q.; Zhu, T.; Xu, J.; Mu, J.; Wang, Y.; Wang, H.; Tang, T.; Ren, Y.; Wu, Y.; Lin, S. H.; Qiu, Y.; Zhang, D. Y.; Shang, Y.; Zhou, X. Plasma Metabolomic and Lipidomic Alterations Associated with COVID-19. *Natl. Sci. Rev.* **2020**, *7* (7), 1157–1168.

- (35) Shen, B.; Yi, X.; Sun, Y.; Bi, X.; Du, J.; Zhang, C.; Quan, S.; Zhang, F.; Sun, R.; Qian, L.; Ge, W.; Liu, W.; Liang, S.; Chen, H.; Zhang, Y.; Li, J.; Xu, J.; He, Z.; Chen, B.; Wang, J.; Yan, H.; Zheng, Y.; Wang, D.; Zhu, J.; Kong, Z.; Kang, Z.; Liang, X.; Ding, X.; Ruan, G.; Xiang, N.; Cai, X.; Gao, H.; Li, L.; Li, S.; Xiao, Q.; Lu, T.; Zhu, Y.; Liu, H.; Chen, H.; Guo, T. Proteomic and Metabolomic Characterization of COVID-19 Patient Sera. *Cell* **2020**, *182* (1), 59-72.e15.
- (36) B, A.; Waseem, M.; Kashif, M.; Srinivasan, H. Lipidomics: An Excellent Tool for Chronic Disease Detection. *Curr. Res. Transl. Med.* **2022**, *70* (4), 103346.
- (37) Hannun, Y. A.; Obeid, L. M. Sphingolipids and Their Metabolism in Physiology and Disease. *Nat. Rev. Mol. Cell Biol.* **2017**, *19* (3), 175–191.
- (38) Janneh, A. H.; Kassir, M. F.; Dwyer, C. J.; Chakraborty, P.; Pierce, J. S.; Flume, P. A.; Li, H.; Nadig, S. N.; Mehrotra, S.; Ogretmen, B. Alterations of Lipid Metabolism Provide Serologic Biomarkers for the Detection of Asymptomatic versus Symptomatic COVID-19 Patients. *Sci. Reports* **2021**, *11* (1), 1–10.
- (39) Bai, Y.; Huang, W.; Li, Y.; Lai, C.; Huang, S.; Wang, G.; He, Y.; Hu, L.; Chen, C. Lipidomic Alteration of Plasma in Cured COVID-19 Patients Using Ultra High-Performance Liquid Chromatography with High-Resolution Mass Spectrometry. *Biosci. Rep.* **2021**, *41* (3), 20204305.
- (40) de Fátima Cobre, A.; Surek, M.; Stremel, D. P.; Fachi, M. M.; Lobo Borba, H. H.; Tonin, F. S.; Pontarolo, R. Diagnosis and Prognosis of COVID-19 Employing Analysis of Patients' Plasma and Serum via LC-MS and Machine Learning. *Comput. Biol. Med.* **2022**, *146*, 105659.
- (41) Naldrett, M. J.; Zeidler, R.; Wilson, K. E.; Kocourek, A. Concentration and Desalting of Peptide and Protein Samples with a Newly Developed C18 Membrane in a Microspin Column Format. *J. Biomol. Tech.* **2005**, *16* (4), 423.
- (42) Ng, E. W. Y. E. E. W.; Wong, M. Y. M. Y.; Poon, T. C. W. C. T.; Ng, E. W. Y. E. E. W.; Poon, T. C. W. C. T.; Ka Shing, L.; Wong Li Ka, M. *Advances in MALDI Mass Spectrometry in Clinical Diagnostic Applications*; 2013; Vol. 336, pp 139–175.