



**UNIVERSIDADE FEDERAL DA BAHIA
INSTITUTO DE QUÍMICA
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA**

ESTÉFANE DA CRUZ NUNES

**A METABOLÔMICA NA COMPREENSÃO DE DOENÇAS
NEUROLÓGICAS: ANÁLISE *UNTARGETED* DE AMOSTRAS DE
SORO DE BEBÊS COM MICROCEFALIA**

Salvador – BA
2021

ESTÉFANE DA CRUZ NUNES

**A METABOLÔMICA NA COMPREENSÃO DE DOENÇAS
NEUROLÓGICAS: ANÁLISE *UNTARGETED* DE AMOSTRAS DE
SORO DE BEBÊS COM MICROCEFALIA**

Dissertação apresentada ao Programa de Pós-Graduação em Química, Instituto de Química, Universidade Federal da Bahia, como requisito para obtenção do grau de Mestre em Química Analítica.

Orientadora: Prof^a. Dra. Gisele André Baptista Canuto

Salvador – BA
2021



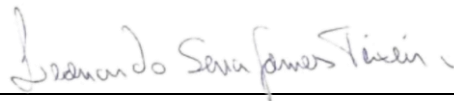
TERMO DE APROVAÇÃO

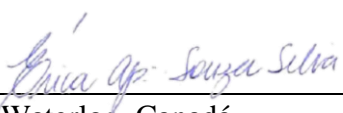
ESTÉFANE DA CRUZ NUNES

"A METABOLÔMICA NA COMPREENSÃO DE DOENÇAS NEUROLÓGICAS:
ANÁLISE UNTARGETED DE AMOSTRAS DE SORO DE BEBÊS COM
MICROCEFALIA"

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre(a) em Química,
Universidade Federal da Bahia, pela seguinte banca examinadora:

Prof^a. Dr^a. Gisele André Baptista Canuto 
Doutorado em Ciências, Universidade de São Paulo (USP)
Universidade Federal da Bahia

Prtof. Dr. Leonardo Sena Gomes Teixeira 
Doutorado em Química, Universidade Federal da Bahia (UFBA)
Universidade Federal da Bahia

Prof^a. Dr^a. Érica Aparecida Souza Silva 
Doutorado PhD em Química, University of Waterloo, Canadá
Universidade Federal de São Paulo

Salvador, 16 de dezembro de 2021.



Emitido em 16/12/202

TERMO DE ANUÊNCIA E AUTORIZAÇÃO Nº 613/2021 - IQUI (12.01.15)

(Nº do Protocolo: NÃO PROTOCOLADO)

(Assinado eletronicamente em 17/02/2022 11:52)

LUCIANA SANTOS DA SILVA

ASSISTENTE EM ADMINISTRACAO

PPGQ (12.01.15.07)

Matrícula: 2048239

Para verificar a autenticidade deste documento entre em <https://sipac.ufba.br/documentos/> informando seu número: **613**, ano: **2021**, tipo: **TERMO DE ANUÊNCIA E AUTORIZAÇÃO**, data de emissão: **17/02/2022** e o código de verificação: **3a8a5604e4**

Ficha catalográfica elaborada pela Biblioteca Universitária de Ciências e Tecnologias Prof. Omar Catunda, SIBI - UFBA.

N972 Nunes, Estéfane da Cruz

A metabolômica na compreensão de doenças neurológicas: análise *untargeted* de amostras de soro de bebês com microcefalia/ Estéfane da Cruz Nunes. – Salvador, 2021.

70 f.

Orientadora: Prof^a. Dr^a Gisele André Baptista Canuto

Dissertação (Mestrado) – Universidade Federal da Bahia. Instituto de Química, 2021.

1. Zika. 2. Aminoácidos. 3. GC-MS. 4. Microcefalia. I. Canuto, Gisele André Baptista. II. Universidade Federal da Bahia. III. Título.

CDU 574.192

À minha mãe Sonia, e minha irmã Isabel.

AGRADECIMENTOS

À professora Gisele Canuto, pela orientação, ensinamentos e oportunidades durante o mestrado.

Aos colegas do BioMet, Emile, Raphaela, Thiara e Heiter pelo companheirismo, e em especial à Taiane pela parceria fundamental para realização deste trabalho.

À equipe da Fiocruz, Dr. Luiz Alcântara, Dr^a. Ana Filippis, Dr^a. Nieli Faria, Dr^a. Isadora Siqueira, pela parceria e fornecimento de amostras para o estudo.

À Álvaro Salgado, do Laboratório de Genética Celular e Molecular da UFMG, pela colaboração com as análises por *machine learning*.

À professora Fernanda Barreto, do Instituto Multidisciplinar em Saúde da UFBA, pela contribuição importantíssima na interpretação biológica.

Aos profissionais da Maternidade de Referência Professor José Maria de Magalhães Neto, pela coleta das amostras.

Aos responsáveis pelos recém-nascidos, por aceitarem participar do estudo.

Ao laboratório LEPETRO, pela colaboração com as análises por CG-MS.

Aos professores da Instituto de Química da UFBA por contribuírem significativamente para minha formação.

À minha mãe, Sonia e meus irmãos, Isabel, Diogo e Ricardo pelo incentivo e apoio em todos os momentos da minha vida.

Às minhas sobrinhas Sofia e Emily, por trazerem alegria e leveza aos meus dias.

Aos meus grandes amigos, Felipe, Islânia, Lucas, Clícia, Luísa e Vanessa, pela amizade, apoio e torcida.

À CAPES pela concessão da bolsa de mestrado.

RESUMO

A zika é uma doença provocada pelo Zika vírus (ZIKV), cuja principal forma de transmissão é pela picada dos mosquitos *Aedes aegypti* e *Aedes albopictus*. Casos da doença já foram registrados em mais de 80 países no mundo todo. A infecção por ZIKV tem sido associada a diversas desordens neurológicas e anormalidades auditivas e na visão. Uma preocupação na comunidade médica gira em torno da infecção durante o período gestacional, uma vez que o ZIKV afeta o desenvolvimento fetal e pode provocar complicações na gravidez. Uma das principais consequências é a microcefalia congênita, uma anomalia neurológica que provoca redução do perímetro cefálico do recém-nascido. O conhecimento do impacto do ZIKV no metabolismo é escasso e se mostra fundamental para o desenvolvimento de tratamentos e compreensão da patogênese. A metabolômica, análise abrangente e/ou quantitativa dos metabólitos presentes em um sistema biológico, é uma ferramenta interessante para obter informações acerca das alterações metabólicas de diversas doenças, incluindo a infecção por ZIKV. Neste trabalho, o perfil metabólico de amostras de soro sanguíneo de recém-nascidos com diagnósticos positivos para ZIKV, com e sem microcefalia, foi avaliado através da abordagem metabolômica global. O tratamento das amostras envolveu a precipitação de proteínas, seguido de derivatização, utilizando reações de oximação e siliilação para subseqüente análise por cromatografia gasosa acoplada a espectrometria de massas (GC-MS). Os dados foram processados no software XCMS e os metabólitos identificados no AMDIS. A análise estatística utilizando métodos univariados e multivariados foi aplicada para busca dos metabólitos discriminantes. Os resultados obtidos apontaram alterações significativas em vias metabólicas de aminoácidos, ácidos orgânicos, carboidratos e seus derivados. A interpretação biológica dessas alterações indica a possibilidade da microcefalia estar associada a uma redução nos níveis de carboidratos, provocada pela replicação do vírus no organismo. O aumento dos níveis de aminoácidos nos pacientes com microcefalia reforça essa relação devido ao envolvimento de intermediários da vida da glicólise na síntese de aminoácidos. Pouco se conhece sobre as alterações metabólicas provocadas pela infecção com ZIKV, sobretudo em recém-nascidos. Desse modo, os resultados obtidos neste estudo fornecem informações iniciais relevantes da associação entre a zika e casos de microcefalia.

Palavras-chave: Zika; Metabolômica; GC-MS; Microcefalia

ABSTRACT

Zika is a disease caused by the Zika virus (ZIKV), whose main form of transmission is through the bite of *Aedes aegypti* and *Aedes albopictus* mosquitoes. The disease has been registered in more than 80 countries worldwide. ZIKV infection has been associated with several neurological disorders and auditory and vision abnormalities. A concern in the medical community revolves around infection during the gestation period, as ZIKV affects fetal development and can lead to pregnancy complications. One of the main consequences is congenital microcephaly, a neurological anomaly that causes a reduction in the newborn's head circumference. The knowledge of the impact of ZIKV on metabolism is scarce and is essential for the development of treatments and understanding of pathogenesis. Metabolomics, a comprehensive and quantitative analysis of the metabolites present in a biological system, is an interesting tool to obtain information about the metabolic alterations of several diseases, including ZIKV infection. In this work, the metabolic profile of serum samples from newborns with positive diagnoses for ZIKV, with and without microcephaly was evaluated using untargeted metabolomics. Sample treatment involved protein precipitation, followed by derivatization, using oximation and silylation reactions for subsequent analysis by gas chromatography hyphenated to mass spectrometry (GC-MS). Data were processed in XCMS software and metabolites identified in AMDIS. Statistical analysis using univariate and multivariate methods was applied to find discriminating metabolites. The results showed significant changes in metabolic pathways of amino acids, organic acids, carbohydrates and derivatives. The biological interpretation of these alterations indicates that microcephaly is associated with a reduction in carbohydrate levels caused by virus replication. The increase of amino acid in patients with microcephaly reinforces this relationship due to the involvement of glycolysis intermediaries in amino acid synthesis. Little is known about the metabolic changes caused by infection with ZIKV, especially in newborns. Thus, the results obtained in this study provide relevant information on the association between Zika and cases of microcephaly.

Keywords: Zika, Metabolomics; GC-MS; Microcephaly.

LISTA DE ILUSTRAÇÕES

Figura 1 – Ciclo de transmissão vetorial do Zika vírus.....	15
Figura 2 – Perímetro cefálico de recém-nascidos com e sem microcefalia.....	16
Figura 3 – Fluxo de trabalho da metabolômica.....	19
Figura 4 – Sobreposição dos cromatogramas obtidos das análises das amostras QCs.....	44
Figura 5 – Modelos de análise multivariada de todos os grupos de estudo. (a) Gráfico de escores do modelo de PCA e (b) Gráfico de escores do modelo de PLS-DA. Dados normalizados pela mediana e escalonados por Pareto.....	48
Figura 6 – Gráficos de escores para os modelos de PLS-DA. (a) ZPMN vs ZNMN ($R^2= 0,9593$; $Q^2= 0,00055$), (b) ZPMP vs ZNMN ($R^2= 0,9568$; $Q^2= 0,04034$) e (c) ZPMP vs ZPMN ($R^2= 0,9377$; $Q^2= 0,1362$). Dados normalizados pela mediana e escalonados por Pareto.....	49
Figura 7 – Análise de enriquecimento dos resultados obtidos na análise metabolômica.....	52

LISTA DE TABELAS

Tabela 1 – Compilado das informações de destaque dos trabalhos de metabolômica aplicada à zika.....	32
Tabela 2 – Dados clínicos dos pacientes em estudo.....	40
Tabela 3 – Lista de calibração para análise por RI.....	45
Tabela 4 – Parâmetros do XCMS avaliados para processamento dos dados.....	47
Tabela 5 – Resultados estatísticos dos metabólitos (ou classes químicas) alterados.....	50

LISTA DE ABREVIATURAS E SIGLAS

AMDIS - *Automated Mass Spectral Deconvolution and Identification System*

ANOVA - Análise de Variância

ATP - Adenosina Trifosfato

BSTFA - Bis(trimetilsilil)trifluoroacetamida

bw - Largura de banda

CE - Eletroforese Capilar

CE-MS - Eletroforese Capilar acoplada à Espectrometria de Massas

CZS - Síndrome Congênita do Zika Vírus

EDTA - Ácido Etilenodiamino Tetra-acético

EI - Impacto de Elétrons

ELISA - Ensaio de Imunoabsorção Enzimática

FC - Fold Change

fwhm - Largura do pico

GC – Cromatografia Gasosa

GC-MS - Cromatografia Gasosa acoplada à Espectrometria de Massas

HMDB - *Human Metabolome Database*

IgM - Imunoglobulina M

LC - Cromatografia Líquida

LC-MS - Cromatografia Líquida acoplada à Espectrometria de Massas

LPC - Lisofosfatilcolina

m/z – Razão massa/carga

MBDST - N-terc-butildimetilsilil-N-metiltrifluoroacetamida

ML - *Machine Learning*

MS - Espectrometria de Massas

MSTFA - N-metil-N-trimetilsililtrifluoroacetamida

NAD - Nicotinamida Adenina Dinucleotídeo

NMR - Ressonância Magnética Nuclear

NPLC - Cromatografia Líquida de Fase Normal

OMS - Organização Mundial da Saúde

OPAS - Organização Pan-Americana da Saúde

PCA - Análise de Componentes Principais

PIP - Fosfatidilinositol Fosfato

PLS-DA - Análise Discriminante de Mínimos Quadrados Parciais

Q² - Previsibilidade do modelo

QC - Amostra de Controle de Qualidade

R² - Correlação linear

RI - Índice de Retenção

RPLC - Cromatografia Líquida de Fase Reversa

RT - Tempo de Retenção

RT-PCR - Reação em Cadeia da Polimerase por Transcriptase Reversa

snthresh - relação sinal/ruído

TCA - Ácido Tricarboxílico

TMCS - Trimetilclorosilano

VIP - *Variable Importance Projection*

ZIKV - Zika Vírus

SUMÁRIO

1	INTRODUÇÃO	14
1.1	ZIKA	14
1.1.1	Síndrome Congênita do Zika Vírus	16
1.2	METABOLÔMICA	17
1.2.1	Amostras Biológicas	21
1.2.2	Preparo de amostra	23
1.2.3	Técnicas analíticas	24
1.2.4	Tratamento de dados	27
1.3	ZIKA E METABOLÔMICA	29
2	OBJETIVOS	33
2.1	OBJETIVO GERAL	33
2.2	OBJETIVOS ESPECÍFICOS	33
3	PARTE EXPERIMENTAL	34
3.1	MATERIAIS E REAGENTES	34
3.2	DELINEAMENTO EXPERIMENTAL E COLETA DAS AMOSTRAS	34
3.3	PREPARO DA AMOSTRA	35
3.3.1	Precipitação das proteínas	35
3.3.2	Derivatização	35
3.4	ANÁLISE POR GC-MS	35
3.5	PROCESSAMENTOS DOS DADOS	36
3.6	IDENTIFICAÇÃO DOS METABÓLITOS	36
3.7	ANÁLISE ESTATÍSTICA	37
3.8	CORRELAÇÃO COM EFEITOS BIOLÓGICOS	38
4	RESULTADOS E DISCUSSÕES	39
5	CONCLUSÃO	55
	REFERÊNCIAS	56
	ANEXO A – Espectros de massas dos isômeros	66

1 INTRODUÇÃO

1.1 ZIKA

A zika é uma doença causada pela infecção por Zika vírus (ZIKV). O patógeno foi isolado pela primeira vez no ano de 1947, em sangue de macaco *rhesus*, na floresta de Zika na Uganda. As primeiras infecções em humanos foram identificadas em 1954, na Nigéria (PATTERSON; SAMMON; GARG, 2016; PLOURDE; BLOCH, 2016). O primeiro surto importante ocorreu somente 53 anos depois, em 2007, na ilha Yap onde cerca de 70% da população foi infectada (PLOURDE; BLOCH, 2016; ROMBI et al., 2020). Entre 2013 e 2014, o vírus se espalhou por diversas ilhas do oceano Pacífico, causando grandes surtos. No ano de 2015, o Brasil relatou os primeiros casos de zika nas Américas (MAYER; TESH; VASILAKIS, 2017; PATTERSON; SAMMON; GARG, 2016; PLOURDE; BLOCH, 2016).

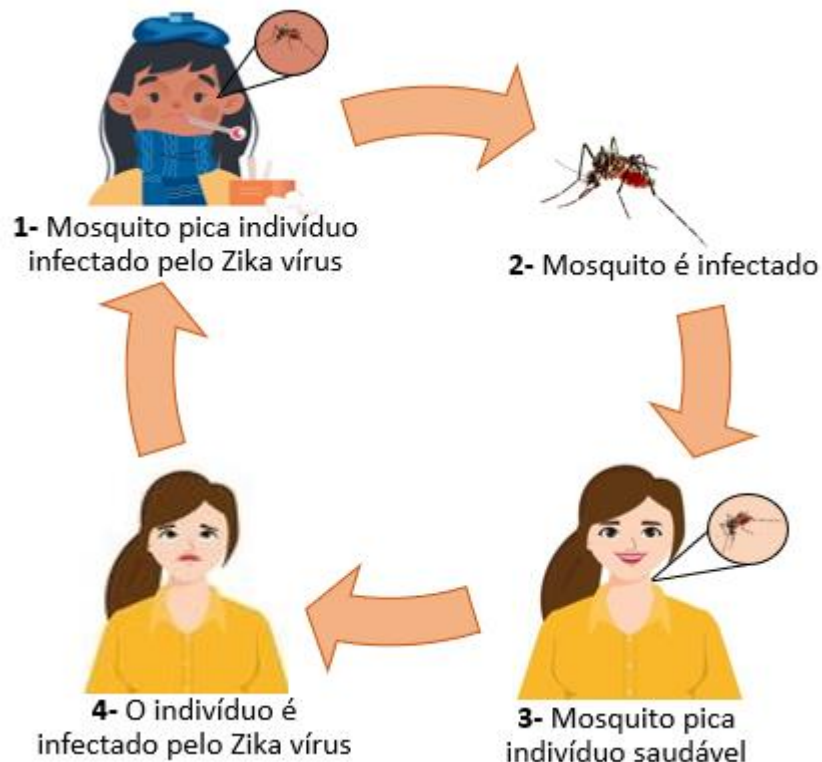
Estudos apontam que o ZIKV entrou no Brasil em 2013, mas a semelhança das manifestações clínicas com a dengue e chikungunya, dificultaram a identificação da doença no início do surto. O vírus se disseminou no território brasileiro, até que em 2016, a transmissão local atingiu todos os estados do país. A epidemia de zika no Brasil teve um número expressivo de infectados e atingiu predominantemente a região Nordeste do país (MAYER; TESH; VASILAKIS, 2017; PLOURDE; BLOCH, 2016; PROENCA-MODENA et al., 2018). De acordo com a Organização Pan-Americana de Saúde (OPAS), entre 2015 e 2020 foram registrados mais de 400.000 casos de zika no Brasil, sendo cerca de 60% das notificações somente no ano de 2016 (PAHO/WHO, 2021).

A transmissão local de ZIKV já foi registrada em 87 países e diversas regiões do mundo. Outros 61 países e territórios têm vetor competente estabelecido, mas não há registro de transmissão do ZIKV. Sendo assim, há um número expressivo de locais com potencial risco de disseminação do vírus. Além disso, de acordo com a Organização Mundial da Saúde (OMS) existe a possibilidade de reemergência da zika em todos os países com transmissões registradas anteriormente (WHO, 2019).

A transmissão do ZIKV ocorre principalmente pela picada dos mosquitos vetores *Aedes aegypti* e *Aedes albopictus*, mas também pode ocorrer através de transmissão sexual, materno-fetal, por transfusão sanguínea e transplante de órgãos

(KARKHAH et al., 2018). A Figura 1 ilustra o ciclo de transmissão vetorial, o qual se inicia quando o mosquito pica uma pessoa infectada. O vírus se multiplica no organismo do vetor, e atinge as glândulas salivares. Então, o mosquito infectado transmite o vírus ao picar outras pessoas saudáveis (PIELNAA et al., 2020).

Figura 1 – Ciclo de transmissão vetorial do Zika vírus



Fonte: A autora. Imagem elaborada com recursos do freepik.com.

Os principais sintomas da doença são febre, dor de cabeça, erupção cutânea, conjuntivite, dor muscular e articular (KARKHAH et al., 2018; SUKHRALIA et al., 2019).

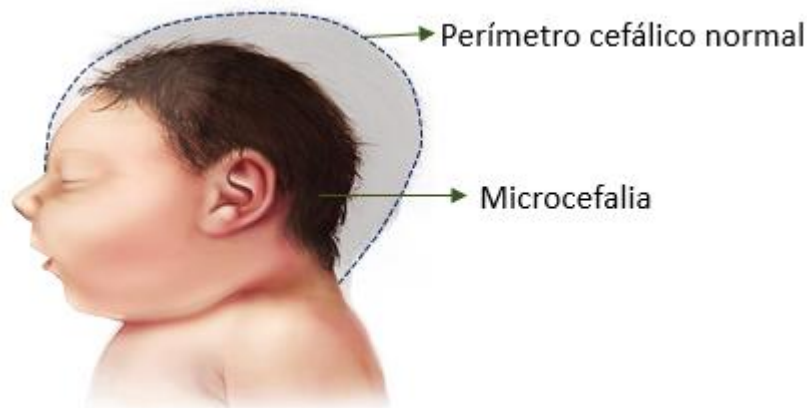
Apesar de alguns estudos apontarem fármacos promissores para o tratamento da zika, ainda não existem medicamentos disponíveis no mercado (BHARDWAJ et al., 2021). Desse modo, o tratamento é baseado na terapia de suporte, visando o alívio dos sintomas (YUN; LEE, 2017). Diversas vacinas contra o ZIKV estão em desenvolvimento, mas até o momento não há nenhum imunizante disponível (BHARDWAJ et al., 2021; SANYAOLU et al., 2019). Para prevenir a infecção é recomendado evitar a picada dos insetos vetores, através de medidas como o uso

de substâncias repelentes e barreiras físicas como roupas que protejam a maior parte do corpo e telas em janelas (KARKHAH et al., 2018).

1.1.1 Síndrome Congênita do Zika Vírus

O aumento dos casos de recém-nascidos com microcefalia no Brasil, especialmente no final de 2015, despertou a atenção dos profissionais de saúde. A ocorrência desses casos coincidente com as regiões com maior incidência de zika, levou especialistas a levantarem a hipótese de uma relação causal entre a infecção por ZIKV e a má formação. A microcefalia congênita é uma anomalia neurológica que provoca redução do perímetro cefálico inferior a dois desvios-padrão da média para sexo e idade gestacional. A Figura 2 ilustra a diferença entre, o perímetro cefálico normal e o microcefálico. Geralmente as principais causas dessa anomalia são desordens genéticas ou infecções ocorridas durante a gestação (DEPOUX et al., 2018; LOWE et al., 2018).

Figura 2 – Perímetro cefálico de recém-nascidos com e sem microcefalia



Fonte: Adaptado de *Centers for Disease Control and Prevention*.

Após o ZIKV ter sido detectado em líquido amniótico, tecido placentário, além de sangue e tecido cerebral de bebês natimortos, a correlação entre a infecção pelo vírus e a má formação foi estabelecida (AVELINO-SILVA; KALLAS, 2018; RASMUSSEN et al., 2016; ROMBI et al., 2020). Diversos países das Américas relataram complicações congênitas associadas a casos de zika. Desde então, a transmissão materno-fetal se tornou um problema de saúde pública, levando, em

2016, a OMS a declarar a zika uma emergência de saúde de interesse internacional em 2016 (LOWE et al., 2018; TALERO-GUTIÉRREZ et al., 2018).

A infecção por ZIKV durante a gestação pode afetar gravemente o desenvolvimento neurológico do bebê, além de acarretar complicações na gestação como aborto e parto prematuro. A transmissão materno-fetal pode ocorrer em qualquer período da gestação, mas as infecções no primeiro trimestre costumam ser mais preocupantes (MUSSO; KO; BAUD, 2019; RASMUSSEN; JAMIESON, 2020; TALERO-GUTIÉRREZ et al., 2018).

Além da microcefalia, a infecção pelo ZIKV provoca uma série alterações no organismo, denominadas de síndrome congênita do Zika vírus (CZS). Bebês expostos ao ZIKV durante a gestação e com perímetro cefálico dentro dos parâmetros normais ao nascer, podem apresentar alterações anatômicas e no desenvolvimento neurológico (CRANSTON et al., 2020; RASMUSSEN; JAMIESON, 2020). Crianças portadoras da CZS podem manifestar alterações oculares, perda auditiva, dificuldade de deglutição, deficiência motora, rigidez muscular, atraso no desenvolvimento neurológico, entre outras complicações (CRANSTON et al., 2020; MUSSO; KO; BAUD, 2019).

Dados do Ministério da Saúde brasileiro indicam que, entre 2015 e 2020, mais de 3000 casos de CZS foram confirmados em todo o país. As regiões de maior prevalência são Nordeste e Sudeste, com quase 70% dos casos. Apesar do período de emergência ter sido encerrado, casos de zika e CZS no Brasil ainda são recorrentes. Somente no ano de 2020, quase 600 casos de CZS foram confirmados (BRASIL, 2021). Diante deste cenário, com um número expressivo de pessoas afetadas pela CZS, conhecer os mecanismos envolvidos na infecção é de grande relevância. Compreender as alterações metabólicas decorrentes da zika, é muito importante para o desenvolvimento de tratamentos, e a metabolômica é uma ferramenta muito útil nesse processo.

1.2 METABOLÔMICA

Diversos progressos importantes relacionados à saúde humana ocorreram a partir do mapeamento do genoma humano, devido a compreensão de vias genéticas. As características multifatoriais das doenças, levaram ao surgimento de

outras ciências ômicas a fim de aprofundar o conhecimento e responder perguntas que a genômica não era capaz de atender (CALABRÒ et al., 2020).

As ciências ômicas procuram compreender o funcionamento de um sistema biológico através da identificação, caracterização e quantificação de moléculas biológicas. Nesse campo de estudo encontram-se a genômica, transcriptômica, proteômica e metabolômica. Estas áreas da ciência trabalham com o estudo de alterações do material genético, dos transcritos, das proteínas e dos metabólitos, respectivamente (CALABRÒ et al., 2020).

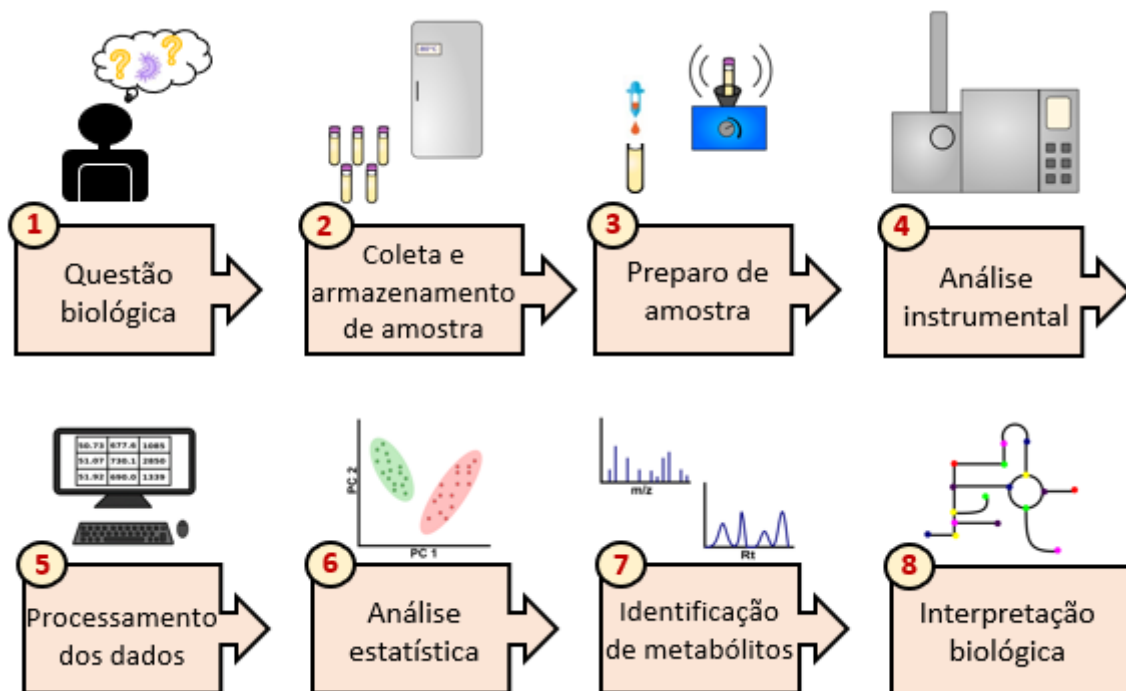
O termo metabolômica foi descrito no início dos anos 2000 por Oliver Fiehn, para designar a análise abrangente e quantitativa dos metabólitos presentes em um sistema biológico, sendo os metabólitos as substâncias geradas através dos processos celulares de um organismo (FIEHN, 2002). Os níveis desses compostos podem representar a última resposta de um sistema biológico às perturbações ocorridas, e indicar como o funcionamento do organismo está relacionado a condições externas como dieta e uso de medicamentos, por exemplo (YU et al., 2020).

As análises metabolômicas podem compreender uma análise abrangente do metaboloma, conjunto de todos os metabólitos produzidos e/ou alterados no organismo em estudo. Quando conduzida de forma abrangente é chamada de metabolômica global ou *untargeted* e compreendem análises semiquantitativas, onde determina-se a abundância relativa dos metabólitos entre grupos de estudo. Outra abordagem é a análise de compostos previamente definidos de acordo com o objetivo da investigação. Está é chamada de metabolômica alvo ou *targeted*, na qual a determinação quantitativa dos metabólitos é feita com padrões analíticos e construção de curvas de calibração. Independentemente da abordagem adotada, as medidas realizadas são sempre comparativas entre grupos de estudo (CANUTO et al., 2018; OLIVEIRA; SIMIONATO; CASS, 2021). Podem ser comparados os níveis de determinados compostos em um grupo submetido a alguma alteração (doença ou uso de medicamento, por exemplo), frente a um grupo que não sofreu a mesma alteração (como por exemplo, indivíduos saudáveis ou uso de placebo).

A metabolômica tem contribuído para a compreensão de diversos sistemas biológicos em diferentes campos de estudo, tais como: ambiental, toxicologia, farmacologia, nutrição, microbiologia, patologia, entre outros. Na área de patologia a

metabolômica clínica atua na compreensão de doenças, contribuindo para desenvolvimento de novos métodos de diagnóstico, novos tratamentos e identificação de biomarcadores (OLIVEIRA; SIMIONATO; CASS, 2021; VUCKOVIC, 2018). A seguir serão descritos os processos envolvidos no fluxo de trabalho da metabolômica (Figura 3).

Figura 3 – Fluxo de trabalho da metabolômica



Fonte: Adaptado de Nunes e Canuto (2020, p. 2104).

A primeira etapa do fluxo de trabalho da metabolômica consiste na definição do problema biológico. Para iniciar o estudo é necessário definir quais perguntas deverão ser respondidas ao final do processo. A partir dessas informações se define o tipo de abordagem (*targeted* ou *untargeted*), tipo de amostra adequada, número de amostras necessárias para uma amostragem representativa, as técnicas analíticas a serem empregadas e protocolos de preparo de amostras (TEBANI; AFONSO; BEKRI, 2018a).

A coleta e armazenamento das amostras é uma etapa extremamente importante para assegurar a integridade do material a ser analisado. As amostras necessitam ser submetidas ao *quenching* metabólico imediatamente após a coleta,

processo que consiste na interrupção imediata da atividade enzimática a fim de evitar alterações na composição do metaboloma. Para tal finalidade, a estratégia mais empregada consiste no resfriamento das amostras a baixas temperaturas, através de congelamento ou adição de solvente orgânico frio. Para evitar degradação ou ativação enzimática, as amostras devem ser armazenadas em baixas temperaturas (-40°C ou -80°C) (JACOB et al., 2019).

Os procedimentos de preparo de amostra variam de acordo com o tipo de amostra, abordagem e técnicas analíticas utilizadas. Podem ser realizadas rupturas de tecidos e células, extrações, diluições, precipitação de proteínas, derivatizações e etc. (GONG et al., 2017). Quando a abordagem *untargeted* é utilizada, o preparo de amostra é não seletivo, a fim de evitar perdas importantes na composição do metaboloma. Já para as análises *targeted*, o preparo de amostra objetiva a seleção e concentração dos metabólitos de interesse, portanto são usados procedimentos mais seletivos (TEBANI; AFONSO; BEKRI, 2018a).

Para analisar as amostras é importante utilizar técnicas analíticas condizentes com os objetivos do estudo. As técnicas mais utilizadas são a espectroscopia de ressonância magnética nuclear (NMR) e a espectrometria de massas (MS), por infusão direta (DIMS) ou acoplada a técnicas de separação como a cromatografia gasosa (GC), cromatografia líquida (LC) e eletroforese capilar (CE). As amostras são analisadas de forma randomizada e intercaladas com amostras de controle de qualidade (QC), que são misturas de volumes iguais de todas as amostras do estudo. As amostras QC são utilizadas para avaliar a estabilidade instrumental durante as análises (CANUTO et al., 2018).

As análises metabolômicas produzem dados bastante complexos e por essa razão, é necessário o uso de ferramentas adequadas de processamento de dados, para extrair informações biológicas relevantes do sistema em estudo. Para extração dos dados são comumente utilizadas ferramentas de detecção de picos, alinhamento, correção do tempo de retenção, deconvolução espectral e etc. Análises estatísticas uni e multivariada são aplicadas para seleção de espécies discriminantes entre os grupos de amostras (TEBANI; AFONSO; BEKRI, 2018b).

Por fim, a identificação dos metabólitos pode ser realizada com o uso de bibliotecas de espectros comerciais ou construídas no laboratório de análise. Bases de dados online como METLIN e *Human Metabolome Database* (HMDB) também

são frequentemente utilizadas. A partir da identificação e, se possível, confirmação dos metabólitos alterados, é feita uma correlação com as rotas metabólicas, a fim de se obter respostas ou levantar hipóteses para compreensão do problema biológico em questão (PILON et al., 2020; TEBANI; AFONSO; BEKRI, 2018b).

1.2.1 Amostras Biológicas

A qualidade dos dados obtidos na metabolômica depende fortemente da execução criteriosa de cada uma das etapas mencionadas anteriormente, sendo necessário garantir que a amostra mantenha sua integridade até o momento da análise. Erros na coleta e armazenamento podem alterar a composição e comprometer os resultados analíticos. Dessa forma, a escolha do tipo de amostra, bem como a metodologia de coleta e armazenamento, são cruciais para obtenção de dados que representem corretamente as alterações do sistema biológico em estudo (DUDZIK et al., 2018).

A escolha do tipo de amostra depende de dois fatores principais: os objetivos do estudo e a disponibilidade das amostras. Diversas matrizes biológicas são utilizadas na metabolômica clínica, desde biofluidos (saliva, sangue, urina, líquido cefalorraquidiano, líquido amniótico e suor), até tecidos e cultura celulares. (DUDZIK et al., 2018; JACOB et al., 2019; TEBANI; AFONSO; BEKRI, 2018a). Cada tipo de amostra tem aspectos particulares acerca da coleta, armazenamento e manuseio, o que torna necessário um conhecimento mais aprofundado sobre a matriz de trabalho. Sendo assim, serão detalhados a seguir alguns aspectos específicos de fluidos biológicos com ampla utilização em metabolômica clínica: o plasma e o soro (HERNANDES; BARBAS; DUDZIK, 2017).

O sangue é uma matriz biológica de grande relevância para a metabolômica, pois apresenta sensibilidade aos efeitos de variação genética, dieta, ambiente e doenças (SOTELO-OROZCO et al., 2021). Para obtenção do plasma, é adicionado ao sangue uma substância anticoagulante e, em seguida, a amostra é centrifugada para separar o plasma das células sanguíneas. Já para a obtenção do soro não é adicionado o anticoagulante, e após a coagulação do sangue é feita a centrifugação para separar as células sanguíneas e o fibrinogênio do soro (HERNANDES; BARBAS; DUDZIK, 2017; SMITH et al., 2020).

A substância anticoagulante utilizada na obtenção do plasma é um fator a ser considerado durante o planejamento, observando os objetivos do estudo e as técnicas de análise empregadas. Isto porque a presença de substâncias como heparina, EDTA, citrato de sódio e fluoreto de sódio, podem interferir na detecção de alguns metabólitos presentes no plasma, devido a coeluição ou supressão de íons (PAGLIA et al., 2018; SOTELO-OROZCO et al., 2021).

O soro pode ser obtido sem a adição de outras substâncias, o que representa uma vantagem de não inserir interferentes no sistema analítico. Entretanto, alguns tubos de coleta possuem um revestimento polimérico com a função de ativar a coagulação. Este tipo de tubo não deve ser utilizado em metabolômica, pois já foram identificadas diferenças nos perfis de ácidos, aminoácidos, monoglicerídeos, entre outras classes de metabólitos (LÓPEZ-BASCÓN et al., 2016). Outro ponto delicado é o tempo de processamento da amostra, uma vez que a coagulação do sangue pode levar cerca de 60 minutos. Além de ser um processo mais lento, a coagulação é realizada em temperatura ambiente, o que posterga o *quenching* e pode acarretar em alterações no metaboloma (HERNANDES; BARBAS; DUDZIK, 2017; LIU et al., 2018).

O manuseio das amostras de sangue deve ser bastante cuidadoso desde a coleta para evitar a hemólise, que consiste no rompimento da membrana celular e consequente liberação de componentes intracelulares para o meio extracelular. A hemólise pode acarretar em alterações nos perfis metabólicos, tanto de plasma quanto de soro, levando ao aumento da concentração de substâncias. A hemólise pode ser detectada visualmente pela alteração na coloração da amostra de soro ou plasma (KAMLAGE et al., 2014; YIN et al., 2013). Para impedi-la é necessário evitar agitação vigorosa, centrifugação excessiva, aplicação prolongada de torniquete no paciente, acondicionamento da amostra em temperatura elevada, entre outros cuidados (HEIREMAN et al., 2017).

Amostras derivadas de sangue são influenciadas pelo tempo de intervalo entre o horário da última refeição e a coleta da amostra, devido a absorção dos componentes dietéticos (BARNES et al., 2016). A literatura aponta que o estado de jejum provoca alterações nos níveis metabólicos de diversas classes como lipídios, aminoácidos e hexoses. Os autores afirmam que, para muitos estudos, essas alterações não impactam significativamente, pois é possível identificar os

metabólitos sensíveis à essa variável e excluí-los da interpretação biológica. Além disso, outros metabólitos não são afetados pelo estado de jejum e, portanto, dependendo dos objetivos do estudo, essa variável pode não ter relevância (CARAYOL et al., 2015; JACOB et al., 2019; STEVENS et al., 2019; TOWNSEND et al., 2016). Em alguns casos, entretanto, como em trabalhos que avaliaram níveis de vitaminas hidrossolúveis, verificou-se a importância de se estabelecer um intervalo de jejum (TOWNSEND et al., 2016).

Assim como para as demais amostras biológicas utilizadas em metabolômica clínica, a recomendação é que amostras de plasma e soro sejam armazenadas a -80°C . Apesar de ser considerada uma temperatura adequada para armazenamento a longo prazo, alguns autores encontraram alterações nos níveis de aminoácidos, lipídios e hexoses em amostras de plasma armazenadas por mais de 5 anos (HAID et al., 2018; WAGNER-GOLBS et al., 2019). Idealmente, as amostras não devem ser submetidas a ciclos de congelamento-descongelamento, para evitar degradação de metabólitos (DUDZIK et al., 2018). Porém se houver necessidade, alguns estudos indicam que até 3 ciclos de congelamento-descongelamento não tem impacto significativo no perfil metabólico (ANTON et al., 2015; BREIER et al., 2014; GOODMAN et al., 2021).

1.2.2 Preparo de amostra

Assim como para a coleta e armazenamento, o preparo da amostra depende da matriz biológica, da técnica analítica empregada e da abordagem metabolômica. A complexidade das matrizes biológicas, com composição química bastante diversa e metabólitos em faixas de concentrações distintas, exige uma criteriosa execução do preparo de amostra para assegurar a qualidade dos dados. Esta é uma etapa que requer muito manuseio do analista, exige grande tempo de execução e é apontada como uma das mais suscetíveis a erros analíticos (CAUSON; HANN, 2016; DUDZIK et al., 2018).

Em trabalhos utilizando a abordagem *untargeted*, o preparo da amostra é focado na remoção de interferentes e os métodos de extração buscam a obtenção da maior quantidade de metabólitos possível. Procedimentos comuns são a remoção de sais, por meio da adição de solventes orgânicos, filtração e precipitação de proteínas. Para as análises *targeted*, além de remover interferentes, o objetivo do

preparo de amostra é extrair e pré-concentrar os metabólitos de interesse. Para isso, são frequentemente utilizadas extrações sólido-líquido, líquido-líquido e em fase sólida (CANUTO et al., 2018; JACOB et al., 2019).

Assim como outros fluidos biológicos, o soro possui elevada concentração de proteínas, o que provoca supressão de sinal analítico e prejudica a detecção dos metabólitos (NAGANA GOWDA; RAFTERY, 2014). Desse modo, é necessário recorrer a procedimentos para remoção desses interferentes. Um recurso muito aplicado na metabolômica é a precipitação de proteínas através da adição de solvente orgânico a frio e posterior centrifugação (BARNES et al., 2016). As proteínas também são precipitadas na presença de ácidos inorgânicos ou sob temperaturas elevadas. No entanto, esses procedimentos já mostraram prejudicar a detecção dos metabólitos e, portanto, não são recomendados (WANT et al., 2006). Em estudos de metabolômica *targeted*, além da desproteínização as amostras de soro podem ser submetidas a procedimentos mais seletivos, como a extração líquido-líquido e extração em fase sólida (QIU et al., 2021).

Frequentemente os estudos metabolômicos empregam NMR e MS acoplada a técnicas de separação como cromatografia e eletroforese capilar. O preparo do soro para análises por NMR envolve diluição da amostra em solução salina deuterada (BARNES et al., 2016; GIL; HEINZMANN, 2017). Para as análises por GC-MS é necessário evaporar a água do extrato e realizar o processo de derivatização, a fim de reduzir o ponto de ebulição de alguns metabólitos e assegurar a estabilidade destes sob as temperaturas empregadas na técnica (50°C a 350°C) (GIL; HEINZMANN, 2017; RATERINK et al., 2014). Usualmente, em trabalhos *untargeted* envolvendo LC-MS apenas a precipitação de proteínas é suficiente (CUI; LU; LEE, 2018). Ao empregar CE-MS, além da desproteínização, são removidos ácidos graxos e lipídios, visto que esta técnica é voltada para análise de compostos ionizados. Além disso, é necessário que as amostras sejam submetidas a evaporação do solvente orgânico devido a sua incompatibilidade com a eletroforese capilar (HIRAYAMA; WAKAYAMA; SOGA, 2014).

1.2.3 Técnicas analíticas

Como mencionado anteriormente, o metaboloma é composto por substâncias com propriedades físicas e químicas distintas, além de uma ampla faixa de

concentração. Tais características tornam desafiadora a análise *untargeted*, visto que não há uma única técnica analítica capaz de detectar toda essa variedade de compostos. Para obter uma ampla cobertura do metaboloma, é recomendada a utilização de mais de uma técnica de análise (PILON et al., 2020). Atualmente, as principais técnicas empregadas em metabolômica são a ressonância magnética nuclear (NMR) e a espectrometria de massas (MS), esta última tanto por meio da infusão direta, quanto hifenada a métodos de separação (cromatografia e eletroforese capilar) (TEBANI; AFONSO; BEKRI, 2018a).

A NMR está baseada na absorção de radiação por um núcleo em um campo magnético. A alteração do estado de spin por uma radiofrequência específica gera um sinal analítico no espectro de ressonância (GIL; HEINZMANN, 2017). A identificação dos metabólitos é feita a partir dos padrões espectrais gerados. Apesar da sua baixa sensibilidade, é uma técnica muito útil por ser reprodutível, robusta, não destrutível e ter um preparo de amostras simples (PILON et al., 2020; TEBANI; AFONSO; BEKRI, 2018a). Por manter a integridade da amostra, esta pode ser recuperada e analisada por cromatografia acoplada a espectrometria de massas, possibilitando a obtenção de uma maior quantidade de informações. Outra importante vantagem é viabilidade de analisar amostras biológicas sólidas, como tecidos, por exemplo (PUTRI et al., 2013).

Na MS os compostos são identificados através da relação *massa-carga* (m/z) dos íons em estado gasoso. Pode ser utilizada individualmente, por infusão direta no espectrômetro de massas, mas pode ter seu desempenho potencializado quando acoplada a técnicas de separação, pois reduz os efeitos da supressão de sinal analítico, resultante da complexidade das matrizes biológicas (BARNES et al., 2016; TEBANI; AFONSO; BEKRI, 2018a).

A eletroforese capilar (CE) é uma técnica de separação baseada na migração diferencial dos compostos sob a influência de um campo elétrico. A diferença de mobilidade decorre do tamanho do analito e da sua carga (ZHANG; HANKEMEIER; RAMAUTAR, 2017). Uma grande vantagem do acoplamento CE-MS para a metabolômica é o uso de pequenos volumes de amostra (na ordem de nL), um aspecto crítico em situações nas quais a disponibilidade de amostra é limitada. Outros aspectos positivos são a separação de compostos em meio aquoso, preparo de amostra simples, curto tempo de análise e alta resolução (GARCÍA et al., 2019;

ZHANG; HANKEMEIER; RAMAUTAR, 2017). A CE-MS possui as desvantagens de apresentar robustez e sensibilidade reduzida. A última em função do pequeno volume de amostra introduzido no sistema. Estes aspectos contribuem para a aplicação limitada da CE-MS em metabolômica (GARCÍA et al., 2019; ZHANG et al., 2019).

Na cromatografia líquida (LC) a separação ocorre através das diferenças de interações dos compostos com a fase móvel e a fase estacionária (PILON et al., 2020). Há dois principais métodos de separação: fase normal (NPLC), quando a fase estacionária é polar; e fase reversa (RPLC), quando a fase estacionária é apolar. Na metabolômica, a LC-MS é utilizada em diversos estudos, por ser capaz de detectar uma ampla gama de substâncias com polaridades distintas (TEBANI; AFONSO; BEKRI, 2018a). Além da versatilidade, é uma técnica com alta sensibilidade e não necessita de uma etapa de derivatização, o que reduz a complexidade do preparo de amostras. A principal desvantagem da LC-MS é o grande consumo de solventes. (CUI; LU; LEE, 2018).

A cromatografia gasosa (GC) é uma técnica cuja separação ocorre devido a diferenças de pressão de vapor e polaridade dos analitos. A temperatura do sistema é elevada gradualmente, podendo chegar a temperaturas acima dos 300°C. Desse modo, é imprescindível que os analitos sejam voláteis e termicamente estáveis nas condições de análise, tornando necessário a derivatização química quando as substâncias não apresentam tais condições (GIL; HEINZMANN, 2017). Esta etapa adicional, além de levar a um preparo de amostra mais prolongado, é uma dos principais fatores que afetam a repetibilidade das análises por acrescentar variabilidade analítica (MOROS et al., 2017; ZEKI et al., 2020).

Uma importante vantagem do acoplamento GC-MS é a obtenção de espectros de massas sob uma energia constante (em geral, 70 eV), em conjunto com informações de tempo de retenção. A associação desses dados possibilita o uso de bibliotecas para identificação dos metabólitos, além de assegurar reprodutibilidade ao método. Em metabolômica a fonte de ionização mais utilizada é por impacto de elétrons (EI), que fornece espectros de fragmentação molecular muito reprodutíveis, com informações estruturais importantes para identificação dos metabólitos (PILON et al., 2020; TEBANI; AFONSO; BEKRI, 2018a; ZEKI et al., 2020).

Em função da sua alta eficiência, as colunas capilares de GC-MS são as escolhidas para os trabalhos em metabolômica. O comprimento da coluna afeta a seletividade, eficiência e tempo de análise. A seleção do comprimento da coluna deve assegurar a melhor resolução no menor tempo de análise possível. De um modo geral, as amostras mais complexas requerem colunas mais longas e maior tempo de análise (ZEKI et al., 2020). Quanto a fase estacionária, é muito frequente o uso de colunas com revestimentos apolares (BARNES et al., 2016; ZEKI et al., 2020).

A GC-MS é uma poderosa ferramenta analítica para estudos em metabolômica. É uma técnica que propicia a identificação e quantificação de uma extensa gama de metabólitos. Conforme já mencionado anteriormente, suas limitações consistem na necessidade de analitos voláteis e termicamente estáveis devendo ser aplicada uma etapa de derivatização química, que adiciona variabilidade analítica e limita o número de amostras a serem analisadas (BARNES et al., 2016; KLASSEN et al., 2017).

1.2.4 Tratamento de dados

As análises metabolômicas *untargeted* fornecem uma extensa e complexa quantidade de dados. Para chegar à identificação dos metabólitos e suas alterações no organismo em estudo, é preciso recorrer a ferramentas adequadas de tratamento dos dados (PILON et al., 2020). Diversos softwares de livre acesso para processamento de dados obtidos por MS estão disponíveis, a exemplo do XCMS (SMITH et al., 2006), MZmine (KATAJAMAA; MIETTINEN; OREŠIČ, 2006) e MetAlign (DE VOS et al., 2007). Existem também os softwares comerciais, como Profiling Solution (Shimadzu), MassLynx (Waters) e Mass Profiler Professional (Agilent Technologies).

Para dados obtidos no MS, o processamento inclui as etapas de remoção de ruído, correção da linha de base, detecção e alinhamento de picos, e deconvolução espectral. O objetivo da remoção de ruído é diferenciar o sinal analítico de sinais gerados por efeito de matriz ou interferência instrumental. Estas interferências podem também provocar distorções na linha de base, que precisam ser ajustadas durante o processamento dos dados. A detecção dos picos por sua vez, consiste na identificação e quantificação dos sinais correspondentes aos metabólitos (TEBANI;

AFONSO; BEKRI, 2018b; YI et al., 2016). O alinhamento dos dados é uma etapa utilizada para corrigir diferenças nas posições (em tempo) dos picos entre as amostras. Essas alterações podem ocorrer em função de efeito de matriz, alargamento ou degradação da fase estacionária da cromatografia (GIL; HEINZMANN, 2017). A deconvolução, etapa essencial no processamento dos dados de GC-MS, é necessária para separar os espectros de picos sobrepostos (TEBANI; AFONSO; BEKRI, 2018b). Um dos softwares largamente utilizado para este fim é o *Automated Mass Spectral Deconvolution and Identification* (AMDIS) (KIND et al., 2009). Por fim, é necessário realizar uma etapa de normalização, para remover variações experimentais provocadas por divergências analíticas e biológicas, na qual podem ser utilizadas estratégias baseadas em modelos estatísticos ou a partir de abordagens químicas, como o uso de padrão interno (ZEKI et al., 2020).

Após obter a matriz de dados bidimensionais, esta deve ser submetida à análise estatística. Métodos multivariados são muito aplicados em metabolômica devido a grande quantidade de dados, pois permite que o conjunto de dados com extensa quantidade de variáveis seja reduzido ao menor número de variáveis possíveis. A Análise de Componentes Principais (PCA) é um método não-supervisionado bastante útil na identificação de diferenças entre os grupos de estudo (GIL; HEINZMANN, 2017; ZEKI et al., 2020). A PCA projeta cada amostra em uma nova matriz ortogonal, preservando a variância máxima dos dados originais. Um método supervisionado, também utilizado frequentemente na análise de dados metabolômicos, é a Análise Discriminante de Mínimos Quadrados Parciais (PLS-DA), que busca a variação ligada a uma classe específica de amostra, possibilitando a melhor separação entre os grupos. Análises univariadas, como teste t de Student e Análise de Variância (ANOVA), podem ser aplicadas para complementar as informações sobre metabólitos discriminantes, bem como corroborar as respostas das análises multivariadas (GIL; HEINZMANN, 2017; YI et al., 2016).

A complexidade dos dados metabolômicos requer aperfeiçoamento constante das ferramentas de tratamento de dados, para extração de mais informações. Nos últimos anos, os algoritmos de *machine learning* (ML) tem se mostrado um importante aliado à análise dos dados metabolômicos. As ferramentas de ML utilizam modelos estatísticos treinados para realizar previsões sobre dados desconhecidos, o que a torna importante para desenvolvimento de métodos

diagnósticos baseado em biomarcadores (LIEBAL et al., 2020). A capacidade da ML de processar uma extensa quantidade de dados e realizar previsões, se mostra interessante para o enriquecimento dos conhecimentos acerca das alterações metabólicas provocadas por patógenos, e leva a metabolômica a importantes progressos.

1.3 ZIKA E METABOLÔMICA

A zika e a CZS se tornaram um problema de saúde pública nos últimos anos. Desse modo, conhecer o impacto da infecção no metabolismo é essencial para compreensão da patogênese da zika e o desenvolvimento de tratamentos. No entanto, encontram-se poucos trabalhos na literatura com estudos metabolômicos sobre a zika e suas sequelas neurológicas (NUNES; CANUTO, 2020; ROMBI et al., 2020).

Recentemente duas revisões sobre metabolômica no estudo de arboviroses, incluindo a zika, foram publicadas. Byers e colaboradores (2019), compilaram informações acerca das alterações metabólicas encontradas em estudos com amostras de humanos. Já Nunes e Canuto (2020), além de apresentarem os dados sobre os metabólitos alterados, destacam aspectos analíticos dos trabalhos revisados, como o tamanho amostral, técnicas analíticas e nível de identificação, entre outros. Nesta revisão foram englobados estudos envolvendo tanto amostras humanas, quanto estudos com modelos animais. Ambas as revisões reúnem informações importantes acerca das contribuições da metabolômica nos estudos de arboviroses.

Em um interessante trabalho de metabolômica aplicada à zika, Melo e colaboradores (2017) analisaram amostras de soro humano (adultos) por MS. Os pesquisadores encontraram alterações nos níveis de angiotensina, metabólito envolvido na redução da replicação viral, e os fosfatidilinositol fosfatos (PIP), que fazem parte do processo de sinalização celular para a autofagia. As alterações nos níveis de gangliosídeo GM2 foram relacionadas com a síndrome de Guillain-Barré. O gangliosídeo GM2 funciona como marcador de células infectadas para o sistema imunológico. A sua presença nos neurônios, pode levar o sistema imunológico a atacá-los além de atacar as células infectadas. Este mesmo grupo de pesquisa publicou outro trabalho voltado para o desenvolvimento de um método de

diagnóstico da zika (MELO et al., 2018). Os autores desenvolveram um algoritmo, supervisionado por Random Forest, para introduzir os dados de MS e obter um conjunto de *features*, que são com uma “impressão digital” para a infecção. Para obter essa “impressão digital” os autores analisaram amostras de soro humano (adultos), entre pacientes saudáveis e infectados por ZIKV. Os *features* característicos para os pacientes acometidos pela zika corresponderam a seis metabólitos: pentapeptídeo, tetrapeptídeo, divalente e monovalente gangliosídeo, cardiolipina e um bisfosfoglicerol. Os resultados fornecidos pelo algoritmo para o diagnóstico foram concordantes com os resultados de RT-PCR, método convencional para o diagnóstico da doença. A ML também foi aplicada no trabalho de Jarrin e colaboradores (2019), no qual algoritmos de Regressão Logística e Máquina de Vetor de Suporte foram utilizados para identificar a infecção por ZIKV.

Um estudo com células microgliais infectadas por ZIKV, encontrou alterações em aminoácidos, glicerofosfolipídios, esfingolipídios ácidos graxos, ácidos carboxílicos, aminofosfolipídios e aminas. Através da análise por LC-MS os autores identificaram o aumento nos níveis de lisofosfatilcolina (LPC), esfingomielina e ceramida, metabólitos relacionados à neuroinflamação (DIOP et al., 2018).

Em 2019, Oliveira e colaboradores (2019), publicaram um estudo visando identificar metabólitos envolvidos no dano neurológico responsável pela microcefalia. Os autores analisaram por MS amostras da saliva de bebês com até dois anos de idade, e encontraram alterações em metabólitos relacionados com inflamação neurológica. Derivados da prostaglandina e hepoxilinas foram encontrados aumentados no grupo com a síndrome, o que segundo os autores, reflete uma resposta anti-inflamatória para lesões neuronais.

Em um estudo multiômico, publicado recentemente, amostras de tecido cerebral de camundongos foram analisadas por LC-MS. Os autores encontraram alterações nos níveis de arginina, em intermediários do ciclo do ácido tricarboxílico (TCA), do metabolismo de purina e pirimidina, nicotinamida adenina dinucleotídeo (NAD⁺) e triptofano-quinurenina. Os autores destacam que a diminuição de NAD⁺ pode estar relacionada com a morte de células neuronais (PANG et al., 2021).

Utilizando a metabolômica em fluxo, Thaker e colaboradores (2019) compararam as alterações no metabolismo da glicose, em células humanas e em células do *Aedes albopictus*. De acordo com os autores, as alterações metabólicas

encontradas modificam os níveis de nucleotídeos, e podem levar a morte celular para as células humanas.

Apesar de escassos, há uma tendência crescente na quantidade de trabalhos da metabolômica aplicada à zika. Os artigos encontrados na literatura foram publicados nos últimos quatro anos, período após a descoberta dos danos do ZIKV para a gestação e o recém-nascido. A maioria dos estudos utilizou a abordagem *untargeted*, indicando que as investigações ainda se concentram na caracterização do metaboloma. Isto demonstra que as pesquisas nessa área carecem de aprofundamento em metabólitos específicos ou identificação de biomarcadores. No entanto, relevantes informações preliminares foram publicadas, possibilitando aprofundar o conhecimento acerca da infecção por ZIKV e suas sequelas. Um resumo das principais informações sobre os estudos publicados, está apresentado na Tabela 1.

Este trabalho visa obter novas informações acerca das alterações metabólicas provocadas pelo ZIKV em recém-nascidos, e suas relações com sequelas neurológicas. Este é o primeiro trabalho a caracterizar por GC-MS, o metaboloma do soro de recém-nascidos infectados pelo ZIKV. Conforme mencionado no tópico 1.2.1, o soro é uma matriz capaz de fornecer informações extremamente relevantes por apresentar um instantâneo do metabolismo em investigação. Juntamente com a alta reprodutibilidade da GC-MS, e a ampla gama de metabólitos que podem ser identificados por esta técnica, o presente estudo apresenta importantes informações para o estudo metabolômico da Zika.

Tabela 1 – Compilado das informações de destaque dos trabalhos de metabolômica aplicada à zika

Amostra	Grupos de estudo	Abordagem metabolômica	Técnica analítica	Metabólitos (ou classes) alterados	Referência
Soro humano	ZIKV x sadio	<i>Untargeted</i>	MS	Angiotensina (1-7), fosfatidilinositol fosfatos (PIP), fosfatidilinositol bifosfatos (PIP2), Angiotensina I, e gangliosideo GM2.	Melo et al. (2017)
Soro humano	ZIKV x sadio	<i>Untargeted</i>	MS	Pentapeptideo, tetrapeptideos, divalente e monovalente gangliosideos, cardiopina, e bisfosfoglicerol.	Melo et al. (2018)
Células microgliais	ZIKV infectados x não infectados (6, 12, 24, e 28 h.p.i.)	<i>Untargeted</i>	LC-MS	Ácidos carboxílicos, glicerofosfolipídios, esfingolipídios, aminoácidos, ácidos graxos, aminofosfolipídios e aminas.	Diopet et al. (2018)
Saliva humana	ZIKV x sadio	<i>Untargeted</i>	MS	Hepoxilina A ₃ e B ₃ , 15-deoxi- δ -12,14-prostaglandina J ₂ , derivados das prostaglandinas E/D.	Oliveira et al. (2019)
Células humanas e células de <i>A. albopictus</i>	Controle x células humanas infectadas por ZIKV x células (C6/36) do <i>A. albopictus</i> infectadas por ZIKV (24, 36, and 48 h.p.i.)	<i>Untargeted</i>	LC-MS	Nucleotideo trifosfates, intermediários do ciclo TCA e aminoácidos.	Thaker et al. (2019)
Não informado	ZIKV x sadio	<i>Untargeted</i>	LC-MS	Não informado	Jarrin et al. (2019)
Tecido cerebral de camundongos	ZIKV x controle	<i>Targeted</i>	LC-MS	Arginina, intermediários do ciclo TCA, do metabolismo de purina, pirimidina, nicotinamida adenina dinucleotídeo (NAD ⁺) e triptofano-quinurenina	Pang et al. (2021)

2 OBJETIVOS

2.1 OBJETIVO GERAL

Investigar as alterações metabólicas provocadas pela infecção por ZIKV associada ou não a microcefalia em recém-nascidos, através da análise de soro sanguíneo por GC-MS utilizando abordagem metabolômica *untargeted*.

2.2 OBJETIVOS ESPECÍFICOS

- Otimizar os parâmetros de processamento dos dados em ferramentas dedicadas à análises metabolômicas.
- Caracterizar o perfil metabólico do soro sanguíneo de recém-nascidos diagnosticados com zika.
- Buscar metabólitos discriminantes e correlacionar sua função biológica com os efeitos do ZIKV no organismo dos recém-nascidos.

3 PARTE EXPERIMENTAL

3.1 MATERIAIS E REAGENTES

Isopropanol (P. A., Merck, Alemanha), cloridrato de metoxiamina (99%, Sigma Aldrich, Alemanha), piridina (Merck, Alemanha), N,O-Bis(trimetilsilil)trifluoroacetamida (BSTFA) com 1% de trimetilclorosilano (TMCS) (Cerilliant, Alemanha), heptano (grau LC-MS, Merck, Alemanha) e água deionizada.

Agitador Vórtex (Phoenix, AP56, Brasil), microcentrífuga, SpeedVac (Christ, RVC 2-18 CD plus, Alemanha) e estufa (Fabbe, 119, Brasil) foram utilizados neste trabalho.

3.2 DELINEAMENTO EXPERIMENTAL E COLETA DAS AMOSTRAS

Participaram deste estudo bebês nascidos no ano de 2016 em uma maternidade pública localizada na cidade de Salvador, Bahia. Amostras de 26 recém-nascidos foram divididas em 3 grupos: 9 considerados saudáveis, com diagnóstico negativo para Zika vírus e sem microcefalia (grupo ZNMN), 7 com diagnóstico positivo para Zika vírus e sem microcefalia (grupo ZPMN) e 10 portadores do Zika vírus e com microcefalia (grupo ZPMP). As amostras de sangue foram coletadas do cordão umbilical no momento do nascimento, sendo o soro a fração utilizada neste estudo.

O diagnóstico do Zika vírus foi realizado por avaliação conjunta dos resultados de reação em cadeia da polimerase por transcriptase reversa quantitativa (RT-qPCR) e imunoglobulina M (IgM) por ensaio de imunoabsorção enzimática (ELISA). O diagnóstico de microcefalia foi realizado a partir da medida do perímetro cefálico dos recém-nascidos. A classificação da microcefalia foi definida como um perímetro cefálico inferior a 2 desvios-padrão da média para o sexo e idade, de acordo com os gráficos do *International Fetal and Newborn Growth Consortium for the 21st Century (INTERGROWTH-21st)* (VILLAR et al., 2014).

Esse estudo foi aprovado pelo Comitê de Ética em Pesquisa do Instituto Gonçalo Muniz, Fundação Oswaldo Cruz (IGM-FIOCRUZ, protocolo nº 1.935.854/2016) e Laboratório de Flavivírus, IOC/ FIOCRUZ, Rio de Janeiro, Brasil (CAAE: 0249218.6.1001.5248). Os responsáveis legais de todos os recém-nascidos participantes assinaram um termo de consentimento livre e esclarecido.

3.3 PREPARO DA AMOSTRA

3.3.1 Precipitação das proteínas

As amostras de soro foram descongeladas em banho de gelo e em seguida homogeneizadas com auxílio de um vórtex. Uma alíquota de 100 µL foi transferida para um microtubo e 300µL de isopropanol a frio foram adicionados para realização da etapa de precipitação de proteínas. As amostras foram agitadas em vórtex e mantidas em freezer por 1 hora. Após esse período, foram centrifugadas a 13500 rpm durante 10 minutos. Do sobrenadante foram retiradas duas alíquotas: uma de 200 µL, que foi transferida para um vial com insert e o solvente foi evaporado em SpeedVac a 35°C; e outra de 70 µL foi transferida para um tubo Falcon, onde foi misturada com alíquotas de mesmo volume de todas as amostras para preparo das amostras de controle de qualidade (QC). Dessa mistura foram retiradas 15 alíquotas de 200 µL e os sobrenadantes foram evaporados em SpeedVac. Uma amostra branco foi preparada utilizando 100 µL de água deionizada substituindo o soro, e esta foi submetida aos mesmos procedimentos aplicados às amostras.

3.3.2 Derivatização

Os extratos secos das amostras, QCs e branco foram derivatizados em duas etapas: oximação e sililação. Para a oximação, foram adicionados 20 µL de cloridrato de metoxiamina em piridina (15 mg.mL⁻¹) ao extrato seco e a mistura submetida a agitação vigorosa em vórtex. Em seguida, as amostras foram colocadas em local escuro e a reação ocorreu à temperatura ambiente por 90 minutos (CANUTO et al., 2017). Para a etapa de sililação foram adicionados 20µL de N,O-Bis(trimetilsilil)trifluoroacetamida (BSTFA) com 1% de trimetilclorosilano (TMCS) (v/v). As misturas foram agitadas em vórtex e mantidas a 40°C por 30 minutos (CANUTO et al., 2017). Após as duas etapas de derivatização e resfriamento das amostras, 120 µL de heptano foram adicionados aos extratos.

3.4 ANÁLISE POR GC-MS

As amostras foram analisadas de forma randomizada e intercaladas com injeção de QCs a cada 5 amostras. O equipamento utilizado foi um cromatógrafo gasoso (modelo 7820A, Agilent Technologies) acoplado a um espectrômetro de massas do tipo quadrupolo (modelo 5977E MSD, Agilent Technologies). A separação dos analitos ocorreu em coluna DB5-MS (fenil arileno equivalente a (5%-

fenil)-dimetilpolisiloxano) (60 m de comprimento, 0,32 mm de diâmetro interno, filme de 0,25 μm , Agilent Technologies). O gás de arraste utilizado foi Hélio no fluxo de 1,0 mL/min. A temperatura do injetor foi mantida a 250°C e as amostras foram injetadas com split 1:10 a 10 mL/min de Hélio. O gradiente de temperatura do forno foi de 50°C a 325°C a uma taxa de 5°C/min. A linha de transferência do detector, o filamento da fonte e o quadrupolo foram mantidos à 290°C, 230°C e 150°C, respectivamente. A fonte por impacto de elétrons utilizou -70eV de energia e o espectrômetro de massas foi operado no modo scan na faixa de 50 – 600 m/z . O software Mass Hunter Qualitative Analysis B.05.00 (Agilent Technologies) foi utilizado para operação do equipamento e aquisição dos dados.

3.5 PROCESSAMENTOS DOS DADOS

Os dados obtidos foram convertidos para *.mzML no software ProteoWizard (versão 3.0) (HOLMAN; TABB; MALLICK, 2014), para posterior processamento no software de livre acesso XCMS (versão 3.6.1) (SMITH et al., 2006), executado em plataforma R (versão 3.6.0). A matriz de dados do GC-MS foi extraída utilizando o método "matched filter", empregando os seguintes parâmetros: largura do pico (fwhm = 2), relação sinal ruído (snthresh = 1.5), largura de banda (bw = 3 e 1.5, para o primeiro e segundo agrupamento, respectivamente), número máximo de picos por cromatograma de íon extraído (max = 100), largura das faixas sobrepostas de m/z 's (mzwid = 0.25) e fração mínima de amostras necessárias em ao menos um grupo para ser considerada válida (minfrac = 0.5). O alinhamento foi realizado pelo método "retcor" com alinhamento não linear e grau de suavização para ajuste de regressão polinomial local (span = 0.5). A função "fillPeaks" foi aplicada para reduzir os *missing values*. Para os demais parâmetros de processamento de dados, foram utilizados os valores padrão do XCMS. Os dados foram extraídos em termos da intensidade dos picos.

3.6 IDENTIFICAÇÃO DOS METABÓLITOS

Os metabólitos foram identificados com auxílio do software *Automated Mass Spectral Deconvolution and Identification System* (AMDIS, versão 2.73) usando a biblioteca de espectros Fiehn RT library (KIND et al., 2009). A identificação dos metabólitos se deu através da comparação do perfil de fragmentação espectral e

tempo de retenção dos analitos, após as análises do índice de retenção (RI) e correção do tempo de retenção (RT). Ambas as análises foram realizadas aplicando um fator de correspondência mínimo igual a 60, modo reverso e largura do componente de 15. Para a análise por RT foi utilizada uma janela de tempo de retenção de 0,05 minutos

3.7 ANÁLISE ESTATÍSTICA

Para buscar metabólitos discriminantes entre os grupos foram utilizadas análises estatísticas multivariada e univariada, juntamente com *machine learning* (ML). A análise estatística multivariada foi realizada no software MetaboAnalyst 4.0 e 5.0 (CHONG; WISHART; XIA, 2019), utilizando modelos de PCA (*Principal Component Analysis*) e PLS-DA (*Partial Least Squares Discriminant Analysis*). Os dados foram normalizados pela mediana. Escalonamento Pareto e transformação logarítmica também foram aplicados.

Na análise multivariada os metabólitos discriminantes foram determinados através da análise de VIP score (*Variable Importance Projection*) nos modelos PLS-DA, selecionando aqueles com $VIP > 1,0$. Na análise univariada a seleção dos metabólitos estatisticamente significativos ocorreu por meio da ANOVA (*Analysis of Variance*) seguida do teste de Tukey com corte de p -valor $< 0,05$ (Statistica 10).

Análise por ML também foi conduzida neste estudo, para corroborar os achados da estatística multivariada e univariada e complementar a busca por metabólitos discriminantes entre os grupos. Essas análises foram realizadas pelo Grupo de Pesquisa do Laboratório de Genética Celular e Molecular da Universidade Federal de Minas Gerais, na qual foi utilizado o algoritmo de classificação XGboost (CHEN; GUESTRIN, 2016; FRIEDMAN, 2001; SCHAPIRE, 2003). O ajuste dos parâmetros do modelo foi realizado por validação cruzada k-fold e a avaliação do desempenho em um conjunto de dados de teste. A estrutura SHAP (*SHapley Additive exPlanations*) foi utilizada para interpretação dos resultados (LUNDBERG; LEE, 2017)

3.8 CORRELAÇÃO COM EFEITOS BIOLÓGICOS

Após a identificação dos metabólitos discriminantes, estes foram correlacionados com rotas metabólicas, através da busca em bases de dados públicas como a Kyoto Encyclopedia of Genes and Genomes KEGG (www.genome.jp/kegg/) e a BioCyc Database Collection (www.biocyc.org/). A interpretação biológica foi feita em conjunto com a equipe de investigadores da FIOCRUZ e do Instituto Multidisciplinar em Saúde da UFBA, que são especialistas em infecção e genômica de arboviroses.

4 RESULTADOS E DISCUSSÕES

As amostras de soro de bebês nascidos em 2016, foram divididas em três grupos: ZNMN, com diagnóstico negativo para ZIKV e sem microcefalia; ZPMN, com diagnóstico positivo para ZIKV e com microcefalia; e ZPMP, positivo para ZIKV e com a microcefalia. A Tabela 2 apresenta dados clínicos dos pacientes, compreendendo os resultados dos testes sorológicos e moleculares para ZIKV, o perímetro cefálico e sua classificação InterGrowth (VILLAR et al., 2014), comorbidades maternas existentes, a ocorrência de complicações na gestação e os casos de erupção cutânea durante a gestação (sintoma característico da zika).

Para este trabalho, utilizamos as amostras de soro dos recém-nascidos. As frações sanguíneas são amplamente utilizadas, em métodos diagnósticos e em pesquisa na área da saúde, pois são biofluidos que agregam informações de diferentes áreas do corpo e, portanto, são adequadas para obter informações instantâneas do metabolismo em investigação (TEBANI; AFONSO; BEKRI, 2018b; TORELL et al., 2015). A literatura aponta que os perfis metabólicos gerados por amostras de soro e plasma, através de análises por NMR ou MS, não diferem significativamente entre si, indicando que ambas as frações sanguíneas são adequadas para estudos metabolômicos. No entanto, trabalhos recentes demonstram que o soro fornece uma quantidade maior de metabólitos, além de níveis mais elevados para diversos aminoácidos, ácidos graxos, lipídios, aminas e carboidratos (LIU et al., 2018; NISHIUMI et al., 2018; PAGLIA et al., 2018; SOTELO-OROZCO et al., 2021; STEVENS et al., 2019). Desse modo, utilizar amostras de soro propiciaria a detecção de um maior número de metabólitos, aspecto importante na caracterização do perfil metabólico (abordagem *untargeted*), como realizado neste trabalho. Vale ressaltar que, as condições de temperatura e tempo de coagulação para obtenção do soro devem ser cuidadosamente controladas, pois durante a coagulação reações enzimáticas podem gerar alterações metabólicas (GONZÁLEZ-DOMÍNGUEZ et al., 2020).

Tabela 2 – Dados clínicos dos pacientes em estudo.

Grupo	Amostra	Classificação InterGrowth	Anti-ZIKV IgM recém-nascido	RT-PCR ZIKV recém-nascido	Sexo	Perímetro cefálico (cm)	Comorbidade materna	Complicação durante a gestação	Erupção cutânea materna (trimestre gestacional)
ZNMN	1	Normocefálico	Negativo	Negativo	M	34,5	Não	Não	Sim (1º)
	2	Normocefálico	Negativo	Negativo	F	32,0	Não	Não	Não
	3	Normocefálico	Negativo	Negativo	M	32,0	Não	Não	Não
	4	Normocefálico	Negativo	Negativo	M	32,0	Não	Não	Não
	5	Normocefálico	Negativo	Negativo	M	34,5	Não	Não	Sim (1º)
	6	Normocefálico	Negativo	Negativo	M	34,0	Não	Não	Sim (3º)
	7	Normocefálico	Negativo	Negativo	F	34,0	Não	Não	Não
	8	Normocefálico	Negativo	Negativo	M	34,0	Não	Não	Sim (1º)
	9	Normocefálico	Negativo	Negativo	F	32,0	Não	Não	Sim (1º)
ZPMN	10	Normocefálico	Positivo	Positivo	M	33,0	Não	Não	Sim (2º)
	11	Normocefálico	Positivo	Negativo	F	29,0	Não	Não	Sim (NI)
	12	Normocefálico	Positivo	Negativo	F	37,0	Não	Não	Sim (1º)
	13	Normocefálico	Negativo	Positivo	M	37,0	HA	Não	Sim (1º)
	14	Normocefálico	Positivo	Negativo	M	34,5	Não	Não	Não
	15	Normocefálico	Negativo	Positivo	M	35,6	Não	Não	Sim (1º)
	16	Normocefálico	Negativo	Positivo	F	34,0	Não	Não	Sim (1º)
ZPMP	17	Microcefalia grave	Negativo	Positivo	F	25,5	Não	Não	Sim (NI)
	18	Microcefálico	Negativo	Positivo	F	31,0	HA	Não	Não

19	Microcefálico	Negativo	Positivo	M	31,0	Não	Não	Não
20	Microcefálico	Positivo	Positivo	M	32,0	Não	Não	Não
21	Microcefalia grave	Positivo	Negativo	M	27,0	Não	Não	Não
22	Microcefálico	Positivo	Negativo	F	31,0	TE	Não	Não
23	Microcefálico	Positivo	Negativo	F	30,0	HA	PE	Não
24	Microcefálico	Positivo	Negativo	F	31,0	Não	Não	Sim (NI)
25	Microcefálico	Positivo	Negativo	F	31,0	Não	Não	Sim (3º)
26	Microcefalia grave	Positivo	Negativo	F	27,0	Não	Não	Sim (NI)

Legenda: F- feminino; M- masculino; HA- hipertensão arterial; TE- trombocitemia essencial; PE- Pré-eclâmpsia; NI- não informado.

Visando uma análise abrangente do metaboloma, o preparo de amostra foi não seletivo, voltado para remoção de interferentes e adequação dos analitos à técnica de análise. Amostras de soro são ricas em proteínas, que se configuram em um interferente relevante para a análise de metabólitos. Devido a sua elevada concentração, podem suprimir o sinal analítico de compostos de interesse, além de acarretar em prejuízos ao desempenho instrumental por acúmulo na coluna cromatográfica podendo levar a obstrução da coluna capilar e efeitos de carry over (GIL; HEINZMANN, 2017; RATERINK et al., 2014; SITNIKOV; MONNIN; VUCKOVIC, 2016). Alguns trabalhos avaliaram a eficiência da desproteínização com diferentes solventes orgânicos, puros e misturas destes. O isopropanol puro foi apontado como o mais eficiente para este fim (DEI CAS et al., 2020; SARAFIAN et al., 2014). Sendo assim, neste trabalho as proteínas foram removidas por precipitação, com isopropanol a frio, na proporção de 1:3 soro:isopropanol (v/v). O solvente foi adicionado ao soro e após 1 hora de precipitação no freezer as amostras foram centrifugadas. Do sobrenadante foram retiradas duas alíquotas: uma para preparo da amostra e a outra para o preparo das amostras de controle de qualidade (QC), que são mistura de volumes iguais de todas as amostras utilizadas no estudo. As amostras QCs são essenciais para avaliar a estabilidade instrumental durante a aquisição de dados (TEBANI; AFONSO; BEKRI, 2018a). Após a precipitação das proteínas, o sobrenadante das amostras de estudo e das QCs foram completamente evaporados em SpeedVac a 35°C. A obtenção do extrato seco é indispensável para garantir a efetividade das reações de derivatização, visto que a presença de água provoca hidrólise dos reagentes utilizados (MASTRANGELO et al., 2015). Uma amostra branco foi preparada utilizando água deionizada em substituição ao soro, e esta foi submetida aos mesmos processos aplicados às amostras.

As reações de derivatização são necessárias para análises por GC-MS, pois diversos metabólitos são termicamente lábeis sob as temperaturas utilizadas nas análises (50 a 325°C). Diversos métodos de derivatização são aplicados na metabolômica, como a fluoroesterificação (KARAMANI et al., 2013), reações com trifluoroetanol, anidrido trifluoroacético e ácido trifluoroacético (ZAMPOLLI et al., 2006). No entanto, o método mais utilizado na metabolômica é o de oximação seguido de sililação. A ampla aplicação deste método se dá devido a sua versatilidade e disponibilidade de bibliotecas espectrais contendo numerosos

espectros dos adutos formados (MASTRANGELO et al., 2015). Neste trabalho, a etapa de oximação foi realizada adicionando o cloridrato de metoxiamina em piridina aos extratos secos. As misturas foram agitadas em vórtex e mantidas a temperatura ambiente em local escuro, para evitar degradação da piridina, por 90 minutos (CANUTO et al., 2017). Nesta reação ocorre a proteção de grupos carbonílicos de aldeídos e cetonas, além da inibição da ciclização de açúcares, importante para redução da quantidade de isômeros (BARNES et al., 2016; HE et al., 2020). Para a sililação pode ser utilizado o N-terc-butildimetilsilil-N-metiltrifluoroacetamida (MBDST), o N-metil-N-trimetilsililtrifluoroacetamida (MSTFA), ou o N,O-bis(trimetilsilil)trifluoroacetamida (BSTFA), que é o mais reativo entre estes (MASTRANGELO et al., 2015). Sendo assim, após a oximação, foram adicionados BSTFA com 1% de trimetilclorosilano (TMCS) (v/v), para a catálise da reação. As misturas foram agitadas em vórtex e mantidas a 40°C por 30 minutos (CANUTO et al., 2017). A reação de sililação é responsável pela diminuição da temperatura de ebulição e aumento da estabilidade térmica dos analitos, isto ocorre devido a substituição dos hidrogênios ácidos por grupos silanóis ($\text{Si}(\text{CH}_3)_3$) (MOROS et al., 2017; PILON et al., 2020). Após o resfriamento, as amostras derivatizadas foram diluídas em heptano e encaminhadas para análise. É importante ressaltar que, além das amostras, QCs e o branco também foram submetidos às reações de derivatização.

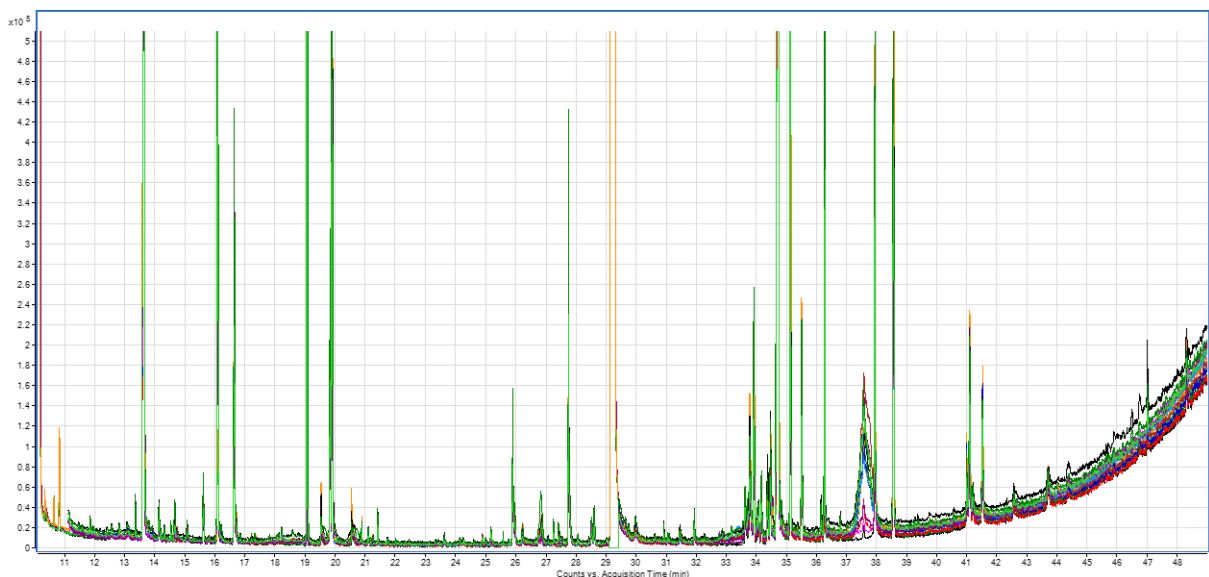
As amostras foram analisadas de forma randomizada e alternadas com as amostras QC, que foram injetadas no início, no fim e ao longo das injeções das amostras. A randomização das amostras tem a função de evitar que uma possível variabilidade analítica, altere os sinais em um determinado grupo de estudo e conduza a uma alteração metabólica equivocada. Procedendo deste modo, todos os grupos serão impactados por eventuais variações instrumentais (KENNEDY et al., 2018; TEBANI; AFONSO; BEKRI, 2018b).

A técnica de escolha para análise metabolômica foi a GC-MS, pois esta é uma técnica que possibilita a detecção de metabólitos de diversas classes químicas, com elevada sensibilidade (HE et al., 2020). Além disso, a técnica possibilita precisão adequada na identificação de metabólitos, em função dos espectros de fragmentação característicos das moléculas, gerados pela ionização por impacto de elétrons, associados às informações de tempo de retenção (MOROS et al., 2017;

ZEKI et al., 2020). Desse modo, na análise por GC-MS fez-se o uso de uma biblioteca espectral, Fiehn RT library (KIND et al., 2009), para a identificação dos metabólitos. Os parâmetros de separação utilizados no cromatógrafo foram adaptados a partir do método indicado para esta biblioteca.

Após a aquisição dos dados cromatográficos, foi realizada uma inspeção visual em todos os cromatogramas, e nenhum perfil discrepante foi encontrado como *outlier*. Como todas as QCs possuem a mesma composição, uma instabilidade nas condições instrumentais poderia ser detectada no perfil cromatográfico obtido. Desse modo, foi realizada uma avaliação comparativa dos cromatogramas obtidos das análises das QCs (Figura 4), e foi verificada a ausência de variações expressivas no perfil e nas intensidades dos picos. Sendo assim, é possível afirmar que não houve variabilidade instrumental significativa durante a análise das amostras.

Figura 4 – Sobreposição dos cromatogramas obtidos das análises das amostras QCs.



Os metabólitos foram identificados no software AMDIS (versão 2.73), utilizando a biblioteca de espectros a Fiehn RT library (KIND et al., 2009). O AMDIS realiza a deconvolução dos sinais e calcula a similaridade entre os espectros resultantes das análises com os espectros da biblioteca (BARNES et al., 2016). Para tanto, os dados obtidos na análise cromatográfica foram inicialmente submetidos a análise de índice de retenção (RI), a qual realiza a comparação de cada pico detectado na amostra com os espectros da biblioteca e indica a similaridade entre ambos. Para a análise

por RI foi utilizada uma lista de calibração composta por metabólitos detectados nas amostras de soro, com fator de correspondência espectral (Net) acima de 85% (Tabela 3). Em seguida foi realizada a análise por tempo de retenção (RT), que confronta os tempos de retenção dos metabólitos detectados nas amostras com os tempos de retenção da biblioteca para os respectivos metabólitos. Como ocorreu uma adaptação do método original, os tempos de retenção obtidos foram corrigidos. A utilização das duas análises, por RI e RT, eleva o nível de confiabilidade da identificação. Os metabólitos identificados no branco foram excluídos da lista de compostos identificados, visto que estes podem ser contaminantes dos reagentes ou dos instrumentos de análise, como compostos de análises anteriores retidos na coluna cromatográfica (MASTRANGELO et al., 2015). Por fim, foram identificados 77 metabólitos.

Tabela 3 – Lista de calibração para análise por RI

Metabólito	Net	Íon mais intenso
Ácido láctico	100	147
DL-isoleucina	87	86
Glicerol	100	73
Ácido glicérico	95	73
D-treitol	91	73
Fenilalanina	91	120
Tetradecanoato de metila	88	74
D-glicose	93	73
Ácido palmítico	99	73
Ácido esteárico	99	73

Paralelamente a identificação, foi realizado o processamento dos dados para extração da matriz de dados brutos. Neste processamento ocorre a correção de diferenças nas larguras e formatos dos picos cromatográficos, geradas por fatores analíticos, como ruído e efeito de matriz (TEBANI; AFONSO; BEKRI, 2018a). Desse modo, inicialmente foi feita a conversão dos arquivos para o formato *.mzML, que é compatível com o software de processamento, utilizando o software Prete Wizard (HOLMAN; TABB; MALLICK, 2014). Em seguida, os dados brutos foram

processados no XCMS (SMITH et al., 2006) para obtenção da matriz de dados contendo os *molecular features* (intensidade de sinal m/z em determinado tempo de retenção). A primeira etapa do processamento no XCMS envolveu a detecção dos picos, na qual foi utilizado o método “*matched filter*”. Em seguida foi realizado o agrupamento dos *molecular features* detectados em amostras individuais, para comparação da intensidade relativa de íons e cálculos dos desvios nos tempos de retenção. Os tempos de retenção foram corrigidos utilizando o comando *retcor* e, devido ao desagrupamento ocorrido nesta etapa, foi necessário realizar o agrupamento dos *molecular features* com os tempos de retenção corrigidos (SMITH et al., 2006). Os parâmetros de detecção, agrupamento e alinhamento dos picos foram otimizados em termos da quantidade de *molecular features* e porcentagem de *missing values*. A Tabela 2 apresenta os valores dos parâmetros avaliados nesta otimização: relação sinal/ruído (*snthresh*), largura do pico (*fwhm*), correção da largura de banda (*bw*) e grau de suavização (*span*). Durante a etapa de detecção de picos, alguns destes podem não terem sido detectados pelo algoritmo. Para preenchimento destes *missing values*, ao final do processamento, foi aplicada a função “*fillPeaks*”. Este comando realiza a integração de picos em regiões do cromatograma onde dados podem ter sido perdidos. O software realiza esse procedimento com base em dados estatísticos obtidos para a posição dos picos, durante as etapas de detecção e alinhamento (SMITH et al., 2006). Conforme mostrado na Tabela 4, a condição “*E*” gerou uma matriz de dados com maior quantidade de *molecular features* (1947), e a porcentagem de *missing values* (10,5%) dentro do limite tolerado (20% a 25%) (ARMITAGE et al., 2015; BIJLSMA et al., 2006). Os demais parâmetros utilizados foram o *default* do software XCMS.

Após obtenção da matriz de dados, os *molecular features* foram correlacionados com os metabólitos identificados no AMDIS, de acordo com os tempos de retenção e os valores de m/z dos fragmentos de maior intensidade do espectro de massas. Assim, dos 77 metabólitos identificados inicialmente, 20 foram correlacionados com os *molecular features*, dentre os quais estão aminoácidos, ácidos graxos, carboidratos, ácidos orgânicos e inorgânicos.

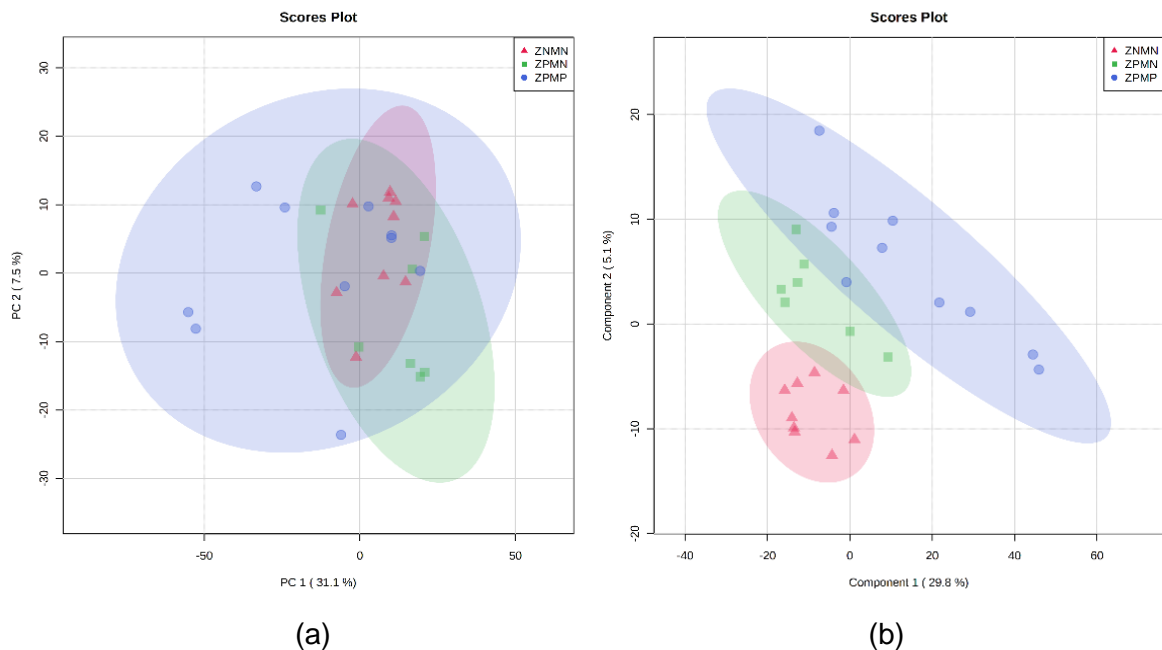
Tabela 4 – Parâmetros do XCMS avaliados para processamento dos dados.

	Detecção de picos		Agrupamento		Alinhamento	Nº <i>Molecular features</i>	% <i>Missing values</i>
	<i>snthresh</i>	<i>fwhm</i>	<i>bw (1)</i>	<i>bw (2)</i>	<i>span</i>		
A	2	4	5	3	0.5	1227	6,36
B	2	5	5	3	0.7	1166	5,72
C	1.5	3	3	1.5	0.7	1670	7,86
D	2	4	4	1.5	0.5	1190	5,80
E	1.5	2	3	1.5	0.5	1947	10,5
F	3	5	5	2	0.6	1939	10,8

Legenda: Parâmetros do XCMS apresentados com separador de decimal da língua inglesa, conforme demandado pelo software.

Para encontrar os metabólitos discriminantes e avaliar os efeitos da infecção por ZIKV no metabolismo dos recém-nascidos, além da sua relação com a microcefalia, foram utilizados métodos estatísticos multivariados e univariados. As análises multivariadas foram realizadas no software MetaboAnalyst (CHONG; WISHART; XIA, 2019). Inicialmente, a matriz de dados foi normalizada pela mediana, para eliminar as variações experimentais provocadas por divergências analíticas e biológicas. Na normalização pela mediana, o software executa a divisão da intensidade de uma amostra pela mediana das intensidades de todas as amostras do estudo (CHONG; WISHART; XIA, 2019; HENDRIKS et al., 2007). O escalonamento Pareto foi aplicado aos dados com a finalidade de ajustar a importância dos compostos no modelo gerado, de modo que metabólitos com abundâncias distintas tenham a mesma contribuição (GIL; HEINZMANN, 2017; MASTRANGELO et al., 2015). A matriz de dados normalizados, contendo todos os grupos envolvidos no estudo, foi então submetida a análise multivariada por meio dos modelos de Análise de Componentes Principais (PCA), e Análise Discriminante de Mínimos Quadrados Parciais (PLS-DA). Os modelos de análise multivariada estão ilustrados na Figura 5. O gráfico de escores da PCA, análise que preserva a variância máxima entre os dados, indica sobreposição entre os grupos avaliados e semelhança entre os perfis metabólicos. No entanto, ao aplicar o método supervisionado PLS-DA, o qual busca a variação relacionada a uma classe específica de amostra, o modelo gerado apresenta uma separação entre as amostras de cada grupo. Tal resultado indica que há diferença entre os perfis metabólicos avaliados.

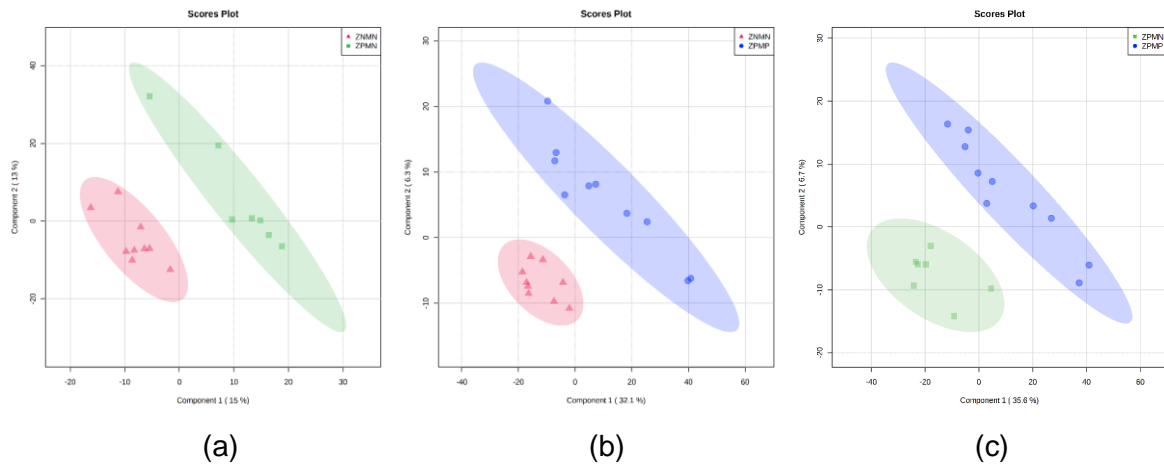
Figura 5 – Modelos de análise multivariada de todos os grupos de estudo. (a) Gráfico de escores do modelo de PCA e (b) Gráfico de escores do modelo de PLS-DA. Dados normalizados pela mediana e escalonados por Pareto.



Legenda: ▲ amostras controle (ZNMN); ■ amostras positivas para Zika, sem microcefalia (ZPMN); ● amostras positiva para Zika, com microcefalia (ZPMP).

Novos modelos de PLS-DA foram construídos comparando os grupos entre si (Figura 6). Os gráficos de escores demonstram clara distinção entre os grupos de estudo para todas as comparações. A qualidade dos modelos de PLS-DA foi avaliada em termos de correlação linear entre os dados, indicada pelos valores de R^2 , e previsibilidade do modelo, indicada pelo Q^2 . Para estudos metabômicos, são indicados valores de $R^2 > 0,8$ e de $Q^2 > 0,6$ (WESTERHUIS et al., 2008). A grande variabilidade apresentada por amostras biológicas de humanos e o pequeno número de amostras neste estudo, contribuíram para a redução da previsibilidade do modelo. Os metabólitos discriminantes foram determinados através da análise de *VIP score* (*Variable Importance Projection*), pois este parâmetro indica a contribuição das variáveis para o modelo. Sendo assim, foram selecionados como significativos, os metabólitos que apresentaram $VIP > 1,0$ (QUINTÁS et al., 2012). A partir dessa análise foram encontrados 12 metabólitos discriminantes entre as comparações dos grupos.

Figura 6 – Gráficos de escores para os modelos de PLS-DA. (a) ZPMN vs ZNMN ($R^2=0,9593$; $Q^2=0,00055$), (b) ZPMP vs ZNMN ($R^2=0,9568$; $Q^2=0,04034$) e (c) ZPMP vs ZPMN ($R^2=0,9377$; $Q^2=0,1362$). Dados normalizados pela mediana e escalonados por Pareto.



Legenda: ▲ amostras controle (ZNMN); ■ amostras positivas para Zika, sem microcefalia (ZPMN); ● amostras positiva para Zika, com microcefalia (ZPMP).

A análise estatística univariada foi realizada no software Statistica (versão 10), aplicando a ANOVA seguida do teste de Tukey com corte de p-valor $< 0,05$. Foram encontrados três metabólitos com diferenças significativas entre os grupos, todos estes com significância estatística também verificada pela análise multivariada. A fim de ampliar as informações obtidas, foi realizada uma análise de *machine learning* (ML) utilizando o algoritmo de classificação XGboost (CHEN; GUESTRIN, 2016). Além de corroborar com as informações obtidas nas análises multivariada e univariada, quatro novos metabólitos foram encontrados como discriminantes. Na Tabela 5 estão listados todos os metabólitos com significância estatística, suas respectivas fórmulas químicas e massa monoisotópica. Além disso, a Tabela 5 apresenta os resultados da análise estatística, que incluem valores de VIP score, fold change e p-valor, em cada uma das comparações.

Tabela 5 – Resultados estatísticos dos metabólitos (ou classes químicas) alterados.

Metabólito	Massa monoisotópica	Fórmula Química	ZPMP vs. ZNMN			ZPMN vs. ZNMN			ZPMP vs. ZPMN		
			VIP score	p-valor	FC	VIP score	p-valor	FC	VIP score	p-valor	FC
DL-isoleucina	131,0946	C ₆ H ₁₃ NO ₂				1,20		0,97			
L-serina	105,0426	C ₃ H ₇ NO ₃				1,88		0,71	1,10		1,21
L-treonina ou L-alotreonina*	119,0582	C ₄ H ₉ NO ₃				1,28		0,75	1,04		1,37
L-valina	117,0790	C ₅ H ₁₁ NO ₂				1,21		0,88			
Ácido N-metilglutâmico	161,0688	C ₆ H ₁₁ NO ₄									0,72
Aldohexose (D-glicose ou D-manose)#	180,0634	C ₆ H ₁₂ O ₆	1,18	0,0334	0,60				1,20	0,0417	0,28
Álcool de açúcar (D-manitol ou D-sorbitol)#	182,0790	C ₆ H ₁₄ O ₆							1,16		0,43
Aldohexose (D-manose ou D-altrose)#	180,0634	C ₆ H ₁₂ O ₆	1,10	0,0286	0,61	1,06	0,0286	2,30	1,13	0,0320	0,26
Ácido glucoheptônico ou gama-lactona de ácido ribônico ou glucono-delta-lactona	226,0689 / 148,0372 / 178,0477	C ₇ H ₁₄ O ₈ / C ₅ H ₈ O ₅ / C ₆ H ₁₀ O ₆						0,88	1,05		1,67
Metil β-D-galactopiranosídeo	194,0790	C ₇ H ₁₄ O ₆	1,12	0,0397	0,52			0,84	1,01	0,0320	0,62
Hexose (tagatose ou L-sorbose)# ou D-lixosilamina (hexoamina)	180,0634 / 149,0688	C ₆ H ₁₂ O ₆ / C ₅ H ₁₁ NO ₄			0,88						
Ácido palmítico	256,2402	C ₁₆ H ₃₂ O ₂	1,07		0,62						
Ácido esteárico	284,2715	C ₁₈ H ₃₆ O ₂	1,05		0,64			0,84			
Ácido fosfórico	97,9769	H ₃ PO ₄			0,75			0,69			
Ácido pirúvico	88,0160	C ₃ H ₄ O ₃									1,31
Ureia	60,0324	CH ₄ N ₂ O	1,31		0,50	2,41		0,41	2,15		1,21

Legenda: ZNMN: grupo controle, ZPMN: ZIKV positivo, sem microcefalia, ZPMP: ZIKV positivo, com microcefalia. FC: fold change (calculado pela razão do primeiro grupo pelo segundo grupo).

*Isômeros indistinguíveis. #Metabólitos indistinguíveis por padrão espectral e tempo de retenção sob análise de GC-MS.

Para obter a magnitude da diferenciação dos metabólitos discriminantes, foi calculado o *fold change* (FC) através da Equação 1, descrita a seguir:

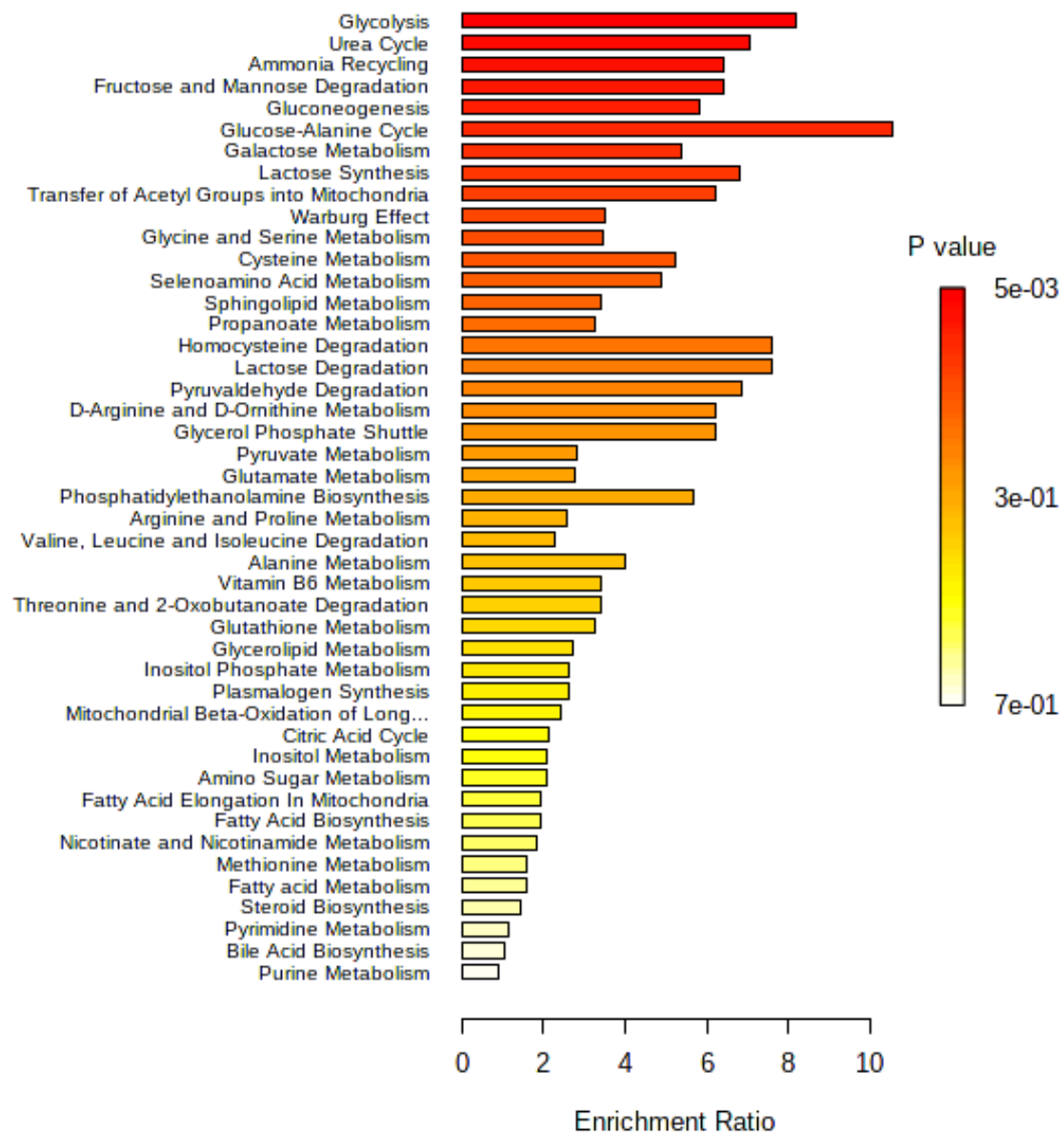
$$\textit{Fold change} = \frac{\textit{Intensidade média do metabólito no grupo 1}}{\textit{Intensidade média do metabólito no grupo 2}} \quad (\text{Eq.1})$$

Como pode ser observado na Equação 1, o FC é determinado pela razão entre a intensidade média do metabólito em um determinado grupo e a intensidade média do mesmo metabólito no grupo de comparação. Desse modo, valores de FC > 1 indicam que o metabólito está aumentado no grupo 1, enquanto que FC < 1 indicam que o metabólito está reduzido no grupo 1 (DUDZIK et al., 2018). Sendo assim, na comparação ZPMP vs ZPMN os valores de FC indicam aumento ou diminuição da intensidade dos metabólitos no grupo ZPMP em relação ao grupo ZPMN. O mesmo raciocínio foi aplicado às demais comparações apresentadas na Tabela 5.

Na Tabela 5 é possível observar também que, em alguns casos, há mais de uma possibilidade de identificação para o composto. Isto ocorre devido a uma limitação do GC-MS para distinguir alguns isômeros (Anexo A). A maioria destes casos envolve os carboidratos e seus derivados, que além dos espectros de fragmentação destes compostos serem muito semelhantes, os tempos de retenção também são muito próximos (MASTRANGELO et al., 2015). Por esta razão não foi possível diferenciá-los e estes foram classificados pelo tipo de carboidrato.

Neste estudo foram observadas alterações em metabólitos de diversas classes como aminoácidos, carboidratos, ácidos graxos e ácidos orgânicos. Para auxiliar na compreensão de quais vias metabólicas foram mais afetadas por estas alterações, foi realizada uma análise de enriquecimento no software MetaboAnalyst (CHONG; WISHART; XIA, 2019). O resultado desta análise está apresentado na Figura 7, na qual as vias metabólicas significativamente afetadas foram o Ciclo de Glicose-Alanina, Glicólise, Ciclo de Ureia, Reciclagem da Amônia e Degradação da Frutose e Manose.

Figura 7 – Análise de enriquecimento dos resultados obtidos na análise metabolômica.



Os resultados deste estudo indicam níveis reduzidos de uma aldohexose (D-glicose ou D-manose) no grupo de recém-nascidos portadores da microcefalia provocada pela infecção por Zika vírus (ZPMP), tanto ao ser comparado com o grupo controle (ZPMP vs ZNMN, FC = 0,60), quanto ao ser comparado com o grupo positivo para Zika sem microcefalia (ZPMP vs ZPMN, FC = 0,28). Esse metabólito não apresentou alterações significativas apontadas pelas análises estatísticas, ao avaliar a infecção por Zika sem microcefalia (ZPMN) com o grupo controle (ZNMN). Essas informações sugerem que os níveis da aldohexose estejam correlacionados com o desenvolvimento da microcefalia. Na glicólise, a glicose é metabolizada em piruvato, que é convertido em acetil-CoA e por fosforilação oxidativa gera adenosina

trifosfato (ATP). A quantidade de ATP gerada por cada unidade de glicose na fosforilação oxidativa é maior do que na glicólise. Estudos já demonstraram que alguns vírus são capazes de induzir a ativação glicolítica e possivelmente usar a ATP produzida pela glicólise. A literatura aponta ainda que a glicólise induzida pelo Zika vírus é necessária para a replicação viral (FONTAINE et al., 2015; TIWARI et al., 2017).

A glicose é uma molécula importante durante o desenvolvimento fetal e é utilizada pelos neurônios para a geração de ATP por fosforilação oxidativa. Trabalhos recentes apontaram que o Zika vírus infecta as células-tronco neurais do feto, o que provoca uma perturbação no metabolismo do hospedeiro. As malformações cerebrais podem ocorrer como resultado dos níveis de glicose se tornarem insuficientes para o desenvolvimento cerebral adequado, devido a ativação glicolítica. O aumento nos níveis de ácido pirúvico, verificado no grupo ZPMP em relação ao ZPMN (FC = 1,31), também é um indicativo de indução da glicólise pela infecção, uma vez que esse metabólito é produto da glicólise. Ou seja, se há a ocorrência de glicólise induzida pelo Zika vírus, é esperado o aumento nos níveis do produto dessa reação (GILBERT-JARAMILLO et al., 2019; MINER; DIAMOND, 2017). É importante ressaltar que a microcefalia é uma condição provavelmente de natureza multifatorial, envolvendo fatores genéticos do hospedeiro ou variações genéticas das cepas virais.

Também foram verificadas alterações nos níveis de aminoácidos e seus derivados. Os aminoácidos fornecem fontes de energia e desempenham funções biológicas em diversos tecidos, como o cérebro. Nos mamíferos, as formas enantioméricas L são predominantes. A L-serina foi encontrada aumentada nos recém-nascidos do grupo ZPMP em relação ao ZPMN (FC = 1,21). Esse resultado é esperado em condições de indução da glicólise, pois a serina é sintetizada a partir de intermediários envolvidos nessa via metabólica (PLOUX; FRERET; BILLARD, 2021; YAMAMOTO; YAKU; NAKAGAWA, 2021).

Além da microcefalia, as malformações oculares são complicações importantes observadas em recém-nascidos infectados pelo Zika vírus durante a gestação (PACHECO et al., 2021). Ainda não está claro na literatura se essa condição está relacionada com o desenvolvimento da microcefalia, ou se é provocada diretamente pelo ZIKV (MACEDO-DA-SILVA et al., 2020). Os resultados apresentados na Tabela

5 indicam redução nos níveis de um álcool de açúcar (D-manitol ou D-sorbitol) no grupo ZPMP em relação ao grupo ZPMN (FC = 0,43). O sorbitol é um metabólito relacionado com problemas oculares como a catarata diabética, provocada pelo acúmulo intracelular deste composto. Sendo assim, os resultados desse estudo sugerem que as lesões oculares provavelmente não estão ligadas ao acúmulo de sorbitol, visto que os recém-nascidos com microcefalia apresentam menores níveis deste metabólito (KIZILTOPRAK et al., 2019). É importante ressaltar que as discussões aqui apresentadas para o sorbitol necessitam de mais estudos para confirmação dessa hipótese.

Estudos dessa natureza ainda são escassos para ZIKV. Neste trabalho obtivemos resultados inéditos sobre as alterações metabólicas em recém-nascidos infectados pelo ZIKV durante a gestação e que desenvolveram a microcefalia. As hipóteses aqui levantadas fornecem informações iniciais relevantes para a compreensão das sequelas neurológicas provocadas pelo ZIKV, fornecendo subsídios para desenvolvimento de tratamentos personalizados aos pacientes acometidos pela doença, visando melhor qualidade de vida para esses bebês e suas famílias.

5 CONCLUSÃO

A síndrome congênita do Zika vírus (CZS) atinge um grande número de pessoas no Brasil. No entanto, os mecanismos envolvidos na infecção ainda são pouco conhecidos. A compreensão das alterações metabólicas provocadas por esta infecção é muito importante para o desenvolvimento de tratamentos, e a metabolômica é uma ferramenta eficaz para obtenção dessas informações. Neste trabalho, foi realizado um estudo *untargeted* com amostras de soro sanguíneo de recém-nascidos, infectados por ZIKV através da transmissão materno-fetal, com e sem microcefalia. As análises metabolômicas foram realizadas por GC-MS e para a identificação dos metabólitos foi utilizada uma biblioteca dedicada à metabolômica (Fiehn RT library). Os *molecular features* extraídos foram correlacionados com os metabólitos identificados no qual foram assinalados 20 metabólitos, dentre os quais estão aminoácidos, ácidos graxos, carboidratos, ácidos orgânicos e inorgânicos. As análises estatísticas multivariada e univariada, juntamente com *machine learning*, apontaram como discriminantes entre os grupos estudados, metabólitos como: serina, valina, ácido pirúvico. A análise de enriquecimento, realizada no software MetaboAnalyst, contribuiu para a compreensão de que as vias metabólicas mais afetadas pelas alterações nos níveis destes metabólitos foram Ciclo de Glicose-Alanina, Glicólise, Ciclo de Ureia, Reciclagem da Amônia e Degradação da Frutose e Manose. Os resultados indicam que as alterações nos níveis de uma aldohexose, que poderia ser glicose ou manose, possivelmente estejam correlacionados com o desenvolvimento da microcefalia. Foram obtidas informações inéditas sobre o perfil metabólico de recém-nascidos que desenvolveram microcefalia em decorrência da infecção por ZIKV durante a gestação. As hipóteses levantadas neste estudo ainda necessitam de uma investigação mais aprofundada, mas apresentam informações iniciais relevantes para um problema de saúde ainda pouco compreendido. Para prosseguimento deste estudo, amostras de soro podem ser analisadas por outras técnicas analíticas, a fim de alcançar maior cobertura do metaboloma e buscar informações complementares sobre o impacto do ZIKV no metabolismo dos recém-nascidos.

REFERÊNCIAS

ANTON, G. et al. Pre-analytical sample quality: Metabolite ratios as an intrinsic marker for prolonged room temperature exposure of serum samples. **PLoS ONE**, v. 10, n. 3, p. 1–19, 2015.

ARMITAGE, E. G. et al. Missing value imputation strategies for metabolomics data. **Electrophoresis**, v. 36, n. 24, p. 3050–3060, 2015.

AVELINO-SILVA, V. I.; KALLAS, E. G. Untold stories of the Zika virus epidemic in Brazil. **Reviews in Medical Virology**, v. 28, n. 6, p. 1–9, 2018.

BARNES, S. et al. Training in metabolomics research. I. Designing the experiment, collecting and extracting samples and generating metabolomics data. **Journal of Mass Spectrometry**, p. 461–475, 2016.

BHARDWAJ, U. et al. Gist of Zika Virus pathogenesis. **Virology**, v. 560, n. May, p. 86–95, 2021.

BIJLSMA, S. et al. Large-scale human metabolomics studies: A strategy for data (pre-) processing and validation. **Analytical Chemistry**, v. 78, n. 2, p. 567–574, 2006.

BRASIL. Boletim Epidemiológico - Situação epidemiológica da síndrome congênita associada à infecção pelo vírus Zika, 2015 a 2020. **Secretaria de vigilância em Saúde, Ministério da Saúde.**, v. 52, p. 1–24, 2021.

BREIER, M. et al. Targeted metabolomics identifies reliable and stable metabolites in human serum and plasma samples. **PLoS ONE**, v. 9, n. 2, p. 1–11, 2014.

BYERS, N. M. et al. Metabolomic insights into human arboviral infections: Dengue, chikungunya, and zika viruses. **Viruses**, v. 11, n. 3, p. 1–30, 2019.

CALABRÒ, G. E. et al. Citizens' Attitudes, Knowledge, and Educational Needs in the Field of Omics Sciences: A Systematic Literature Review. **Frontiers in Genetics**, v. 11, n. October, 2020.

CANUTO, G. A. B. et al. New insights into the mechanistic action of methyldehydrodieugenol B towards *Leishmania (L.) infantum* via a multiplatform based untargeted metabolomics approach. **Metabolomics**, v. 13, n. 5, p. 1–14, 2017.

CANUTO, G. A. B. et al. Metabolômica: definições, estado-da-arte e aplicações representativas. **Química Nova**, v. 41, n. 1, p. 75–91, 1 jan. 2018.

CARAYOL, M. et al. Reliability of serum metabolites over a two-year period: A targeted metabolomic approach in fasting and non-fasting samples from EPIC. **PLoS ONE**, v. 10, n. 8, p. 1–10, 2015.

CAUSON, T. J.; HANN, S. Review of sample preparation strategies for MS-based metabolomic studies in industrial biotechnology. **Analytica Chimica Acta**, v. 938, p. 18–32, 28 set. 2016.

CHEN, T.; GUESTRIN, C. **XGBoost: A Scalable Tree Boosting System**. 2016

CHONG, J.; WISHART, D. S.; XIA, J. Using MetaboAnalyst 4.0 for Comprehensive and Integrative Metabolomics Data Analysis. **Current Protocols in Bioinformatics**, v. 68, n. 1, 2019.

CRANSTON, J. S. et al. Association Between Antenatal Exposure to Zika Virus and Anatomical and Neurodevelopmental Abnormalities in Children. **JAMA network open**, v. 3, n. 7, p. e209303, 2020.

CUI, L.; LU, H.; LEE, Y. H. Challenges and emergent solutions for LC-MS/MS based untargeted metabolomics in diseases. **Mass Spectrometry Reviews**, v. 37, n. 6, p. 772–792, 2018.

DE OLIVEIRA, D. N. et al. Inflammation markers in the saliva of infants born from Zika-infected mothers: exploring potential mechanisms of microcephaly during fetal development. **Scientific Reports**, v. 9, n. 1, p. 1–7, 2019.

DE VOS, R. C. H. et al. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. **Nature Protocols**, v. 2, n. 4, p. 778–791, 2007.

DEI CAS, M. et al. A straightforward LC-MS/MS analysis to study serum profile of short and medium chain fatty acids. **Journal of Chromatography B**, v. 1154, p. 121982, 1 out. 2020.

DEPOUX, A. et al. A multi-faceted pandemic: A review of the state of knowledge on the Zika virus. **Public Health Reviews**, v. 39, n. 1, p. 1–12, 2018.

DIOP, F. et al. Zika virus infection modulates the metabolomic profile of microglial cells. **PLoS ONE**, v. 13, n. 10, p. 1–16, 2018.

DUDZIK, D. et al. Quality assurance procedures for mass spectrometry untargeted metabolomics. a review. **Journal of Pharmaceutical and Biomedical Analysis**, v. 147, p. 149–173, 2018.

FIEHN, O. Metabolomics - The link between genotypes and phenotypes. **Plant**

Molecular Biology, v. 48, n. 1–2, p. 155–171, 2002.

FONTAINE, K. A. et al. Dengue Virus Induces and Requires Glycolysis for Optimal Replication. **Journal of Virology**, v. 89, n. 4, p. 2358–2366, 2015.

FRIEDMAN, J. H. Greedy function approximation: A gradient boosting machine. **Annals of Statistics**, v. 29, n. 5, p. 1189–1232, 2001.

GARCÍA, A. et al. Capillary electrophoresis mass spectrometry as a tool for untargeted metabolomics. **Bioanalysis**, v. 1978, p. 55–77, 2019.

GIL, R. B.; HEINZMANN, S. Metabolomics of Body Fluids. In: **Integration of Omics Approaches and Systems Biology for Clinical Applications**. [s.l.: s.n.]. p. 173–195.

GILBERT-JARAMILLO, J. et al. The potential contribution of impaired brain glucose metabolism to congenital Zika syndrome. **Journal of Anatomy**, v. 235, n. 3, p. 468–480, 2019.

GONG, Z. G. et al. The Recent Developments in Sample Preparation for Mass Spectrometry-Based Metabolomics. **Critical Reviews in Analytical Chemistry**, v. 47, n. 4, p. 325–331, 4 jul. 2017.

GONZÁLEZ-DOMÍNGUEZ, R. et al. Recommendations and best practices for standardizing the pre-analytical processing of blood and urine samples in metabolomics. **Metabolites**, v. 10, n. 6, p. 1–18, 2020.

GOODMAN, K. et al. Assessment of the effects of repeated freeze thawing and extended bench top processing of plasma samples using untargeted metabolomics. **Metabolomics**, v. 17, n. 3, p. 1–11, 2021.

HAID, M. et al. Long-Term Stability of Human Plasma Metabolites during Storage at -80 °C. **Journal of Proteome Research**, v. 17, n. 1, p. 203–211, 5 jan. 2018.

HE, Z. et al. Extensive evaluation of sample preparation workflow for gas chromatography-mass spectrometry-based plasma metabolomics and its application in rheumatoid arthritis. **Analytica Chimica Acta**, v. 1131, p. 136–145, 22 set. 2020.

HEIREMAN, L. et al. Causes, consequences and management of sample hemolysis in the clinical laboratory. **Clinical Biochemistry**, v. 50, n. 18, p. 1317–1322, 2017.

HENDRIKS, M. M. W. B. et al. How to distinguish healthy from diseased? Classification strategy for mass spectrometry-based clinical proteomics. **Proteomics**, v. 7, n. 20, p. 3672–3680, 2007.

HERNANDES, V. V.; BARBAS, C.; DUDZIK, D. A review of blood sample handling and pre-processing for metabolomics studies. **Electrophoresis**, v. 38, n. 18, p. 2232–2241, 1 set. 2017.

HIRAYAMA, A.; WAKAYAMA, M.; SOGA, T. Metabolome analysis based on capillary electrophoresis-mass spectrometry. **TrAC Trends in Analytical Chemistry**, v. 61, p. 215–222, 1 out. 2014.

HOLMAN, J. D.; TABB, D. L.; MALLICK, P. Employing ProteoWizard to convert raw mass spectrometry data. **Current Protocols in Bioinformatics**, n. SUPPL.46, p. 1–9, 2014.

JACOB, M. et al. Metabolomics toward personalized medicine. **Mass Spectrometry Reviews**, v. 38, n. 3, p. 221–238, 2019.

JARRIN, E. P. et al. A Machine Learning-Based algorithm for the assessment of clinical metabolomic fingerprints in Zika virus disease. **2019 IEEE Latin American Conference on Computational Intelligence, LA-CCI 2019**, p. 8–13, 2019.

KAMLAGE, B. et al. Quality markers addressing preanalytical variations of blood and plasma processing identified by broad and targeted metabolite profiling. **Clinical Chemistry**, v. 60, n. 2, p. 399–412, 2014.

KARAMANI, A. A. et al. Fluoroacetylation/fluoroethylesterification as a derivatization approach for gas chromatography–mass spectrometry in metabolomics: Preliminary study of lymphohyperplastic diseases. **Journal of Chromatography A**, v. 1302, p. 125–132, 9 ago. 2013.

KARKHAH, A. et al. Zika virus: epidemiology, clinical aspects, diagnosis, and control of infection. **European Journal of Clinical Microbiology and Infectious Diseases**, v. 37, n. 11, p. 2035–2043, 2018.

KATAJAMAA, M.; MIETTINEN, J.; OREŠIČ, M. MZmine: Toolbox for processing and visualization of mass spectrometry based molecular profile data. **Bioinformatics**, v. 22, n. 5, p. 634–636, 2006.

KENNEDY, A. D. et al. Metabolomics in the clinic: A review of the shared and unique features of untargeted metabolomics for clinical research and clinical testing. **Journal of Mass Spectrometry**, v. 53, n. 11, p. 1143–1154, 2018.

KIND, T. et al. FiehnLib: Mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. **Analytical Chemistry**, v. 81, n. 24, p. 10038–10048, 2009.

KIZILTOPRAK, H. et al. Cataract in diabetes mellitus. **World Journal of Diabetes**, v. 10, n. 3, p. 140–153, 2019.

KLASSEN, A. et al. Metabolomics: Definitions and Significance in Systems Biology. In: **Metabolomics: From Fundamentals to Clinical Applications**. [s.l.: s.n.]. v. 965p. 3–17.

LIEBAL, U. W. et al. Machine learning applications for mass spectrometry-based metabolomics. **Metabolites**, v. 10, n. 6, p. 1–23, 2020.

LIU, X. et al. Serum or plasma, what is the difference? Investigations to facilitate the sample material selection decision making process for metabolomics studies and beyond. **Analytica Chimica Acta**, v. 1037, p. 293–300, 2018.

LÓPEZ-BASCÓN, M. A. et al. Influence of the collection tube on metabolomic changes in serum and plasma. **Talanta**, v. 150, p. 681–689, 2016.

LOWE, R. et al. The zika virus epidemic in brazil: From discovery to future implications. **International Journal of Environmental Research and Public Health**, v. 15, n. 1, 2018.

LUNDBERG, S. M.; LEE, S. I. A unified approach to interpreting model predictions. **arXiv**, n. Section 2, p. 1–10, 2017.

MACEDO-DA-SILVA, J. et al. Serum Proteomics Reveals Alterations in Protease Activity, Axon Guidance, and Visual Phototransduction Pathways in Infants With In Utero Exposure to Zika Virus Without Congenital Zika Syndrome. **Frontiers in Cellular and Infection Microbiology**, v. 10, n. November, p. 1–16, 2020.

MASTRANGELO, A. et al. From sample treatment to biomarker discovery: A tutorial for untargeted metabolomics based on GC-(EI)-Q-MS. **Analytica Chimica Acta**, v. 900, p. 21–35, 2015.

MAYER, S. V; TESH, R. B.; VASILAKIS, N. The emergence of arthropod-borne viral diseases: A global prospective on dengue, chikungunya and zika fevers. **Acta Tropica**, v. 166, n. February, p. 155–163, 2017.

MELO, C. F. O. R. et al. Serum metabolic alterations upon ZIKA infection. **Frontiers in Microbiology**, v. 8, n. OCT, p. 1–10, 2017.

MELO, C. F. O. R. et al. A machine learning application based in random forest for integrating mass spectrometry-based metabolomic data: A simple screening method for patients with Zika virus. **Frontiers in Bioengineering and Biotechnology**, v. 6, n. APR, 2018.

MINER, J. J.; DIAMOND, M. S. Zika Virus Pathogenesis and Tissue Tropism. **Cell Host and Microbe**, v. 21, n. 2, p. 134–142, 2017.

- MOROS, G. et al. Investigation of the derivatization conditions for GC-MS metabolomics of biological samples. **Bioanalysis**, v. 9, n. 1, p. 53–65, 2017.
- MUSSO, D.; KO, A. I.; BAUD, D. Zika Virus Infection — After the Pandemic. **New England Journal of Medicine**, v. 381, n. 15, p. 1444–1457, 2019.
- NAGANA GOWDA, G. A.; RAFTERY, D. Quantitating metabolites in protein precipitated serum using NMR spectroscopy. **Analytical Chemistry**, v. 86, n. 11, p. 5433–5440, 2014.
- NISHIUMI, S. et al. Differences in metabolite profiles caused by pre-analytical blood processing procedures. **Journal of Bioscience and Bioengineering**, v. 125, n. 5, p. 613–618, 2018.
- NUNES, E. DA C.; CANUTO, G. A. B. Metabolomics applied in the study of emerging arboviruses caused by *Aedes aegypti* mosquitoes: A review. **Electrophoresis**, v. 41, n. 24, p. 2102–2113, 2020.
- OLIVEIRA, R. V; SIMIONATO, A. V. C.; CASS, Q. B. Enantioselectivity Effects in Clinical Metabolomics and Lipidomics. **Molecules**, v. 26, n. 17, 2021.
- PACHECO, O. et al. Neurodevelopmental findings in children 20-30 months of age with postnatal Zika infection at 1-12 months of age, Colombia, September-November 2017. **Paediatric and Perinatal Epidemiology**, v. 35, n. 1, p. 92–97, 2021.
- PAGLIA, G. et al. Influence of collection tubes during quantitative targeted metabolomics studies in human blood samples. **Clinica Chimica Acta**, v. 486, n. July, p. 320–328, 2018.
- PAHO/WHO. **PAHO/WHO Data - Weekly Report**. Disponível em: <<https://www.paho.org/data/index.php/en/mnu-topics/chikv-en/550-chikv-weekly-en.html>>. Acesso em: 1 ago. 2021.
- PANG, H. et al. Aberrant NAD⁺ metabolism underlies Zika virus–induced microcephaly. **Nature Metabolism**, v. 3, n. 8, p. 1109–1124, 2021.
- PATTERSON, J.; SAMMON, M.; GARG, M. Dengue, zika and chikungunya: Emerging arboviruses in the new world. **Western Journal of Emergency Medicine**, v. 17, n. 6, p. 671–679, 2016.
- PIELNAA, P. et al. Zika virus-spread, epidemiology, genome, transmission cycle, clinical manifestation, associated challenges, vaccine and antiviral drug development. **Virology**, v. 543, p. 34–42, 1 abr. 2020.
- PILON, A. C. et al. Metabolômica de plantas: métodos e desafios. **Química Nova**, v.

43, n. 3, p. 329–354, 2020.

PLOURDE, A. R.; BLOCH, E. M. A literature review of zika virus. **Emerging Infectious Diseases**, v. 22, n. 7, p. 1185–1192, 2016.

PLOUX, E.; FRERET, T.; BILLARD, J. M. d-serine in physiological and pathological brain aging. **Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics**, v. 1869, n. 1, p. 140542, 1 jan. 2021.

PROENCA-MODENA, J. L. et al. Zika virus: lessons learned in Brazil. **Microbes and Infection**, v. 20, n. 11–12, p. 661–669, 2018.

PUTRI, S. P. et al. Current metabolomics: Technological advances. **Journal of Bioscience and Bioengineering**, v. 116, n. 1, p. 9–16, 1 jul. 2013.

QIU, S. et al. Sample Preparation Method for Mass Spectrometry-Based Metabolomics. In: **Mass Spectrometry-Based Metabolomics in Clinical and Herbal Medicines**. [s.l: s.n.]. p. 33–43.

QUINTÁS, G. et al. Chemometric approaches to improve PLSDA model outcome for predicting human non-alcoholic fatty liver disease using UPLC-MS as a metabolic profiling tool. **Metabolomics**, v. 8, n. 1, p. 86–98, 2012.

RASMUSSEN, S. A. et al. Zika Virus and Birth Defects — Reviewing the Evidence for Causality. **New England Journal of Medicine**, v. 374, n. 20, p. 1981–1987, 2016.

RASMUSSEN, S. A.; JAMIESON, D. J. Teratogen update: Zika virus and pregnancy. **Birth Defects Research**, v. 112, n. 15, p. 1139–1149, 2020.

RATERINK, R. J. et al. Recent developments in sample-pretreatment techniques for mass spectrometry-based metabolomics. **TrAC - Trends in Analytical Chemistry**, v. 61, p. 157–167, 2014.

ROMBI, F. et al. The journey of Zika to the developing brain. **Molecular Biology Reports**, v. 47, n. 4, p. 3097–3115, 2020.

SANYAOLU, A. et al. Changing Epidemiology, Treatment, and Vaccine Update on Chikungunya, Dengue, and Zika Viruses. **Current Tropical Medicine Reports**, v. 6, n. 3, p. 145–159, 2019.

SARAFIAN, M. H. et al. Objective set of criteria for optimization of sample preparation procedures for ultra-high throughput untargeted blood plasma lipid profiling by ultra performance liquid chromatography-mass spectrometry. **Analytical Chemistry**, v. 86, n. 12, p. 5766–5774, 2014.

SCHAPIRE, R. E. The Boosting Approach to Machine Learning: An Overview. p. 149–171, 2003.

SITNIKOV, D. G.; MONNIN, C. S.; VUCKOVIC, D. Systematic Assessment of Seven Solvent and Solid-Phase Extraction Methods for Metabolomics Analysis of Human Plasma by LC-MS. **Scientific Reports**, v. 6, n. December, p. 1–11, 2016.

SMITH, C. A. et al. XCMS: Processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. **Analytical Chemistry**, v. 78, n. 3, p. 779–787, 2006.

SMITH, L. et al. Important considerations for sample collection in metabolomics studies with a special focus on applications to liver functions. **Metabolites**, v. 10, n. 3, 2020.

SOTELO-OROZCO, J. et al. A Comparison of Serum and Plasma Blood Collection Tubes for the Integration of Epidemiological and Metabolomics Data. **Frontiers in Molecular Biosciences**, v. 8, n. July, 2021.

STEVENS, V. L. et al. Pre-analytical factors that affect metabolite stability in human urine, plasma, and serum: A review. **Metabolites**, v. 9, n. 8, 2019.

SUKHRALIA, S. et al. From dengue to Zika: the wide spread of mosquito-borne arboviruses. **European Journal of Clinical Microbiology and Infectious Diseases**, v. 38, n. 1, p. 3–14, 2019.

TALERO-GUTIÉRREZ, C. et al. Zika virus epidemiology: From Uganda to world pandemic, an update. **Epidemiology and Infection**, v. 146, n. 6, p. 673–679, 2018.

TEBANI, A.; AFONSO, C.; BEKRI, S. Advances in metabolome information retrieval: turning chemistry into biology. Part I: analytical chemistry of the metabolome. **Journal of Inherited Metabolic Disease**, v. 41, n. 3, p. 379–391, 2018a.

TEBANI, A.; AFONSO, C.; BEKRI, S. Advances in metabolome information retrieval: turning chemistry into biology. Part II: biological information recovery. **Journal of Inherited Metabolic Disease**, v. 41, n. 3, p. 393–406, 2018b.

THAKER, S. K. et al. Differential Metabolic Reprogramming by Zika Virus Promotes Cell Death in Human versus Mosquito Cells. **Cell Metabolism**, v. 29, n. 5, p. 1206–1216.e4, 2019.

TIWARI, S. K. et al. Zika virus infection reprograms global transcription of host cells to allow sustained infection. **Emerging Microbes and Infections**, v. 6, n. 4, p. e24–10, 2017.

TORELL, F. et al. Multi-organ contribution to the metabolic plasma profile using hierarchical modelling. **PLoS ONE**, v. 10, n. 6, p. 1–18, 2015.

TOWNSEND, M. K. et al. Impact of pre-analytic blood sample collection factors on metabolomics. **Cancer Epidemiology Biomarkers and Prevention**, v. 25, n. 5, p. 823–829, 2016.

VILLAR, J. et al. International standards for newborn weight, length, and head circumference by gestational age and sex: the Newborn Cross-Sectional Study of the INTERGROWTH-21st Project. **The Lancet**, v. 384, n. 9946, p. 857–868, 6 set. 2014.

VUCKOVIC, D. Improving metabolome coverage and data quality: Advancing metabolomics and lipidomics for biomarker discovery. **Chemical Communications**, v. 54, n. 50, p. 6728–6749, 2018.

WAGNER-GOLBS, A. et al. Effects of long-term storage at –80 °C on the human plasma metabolome. **Metabolites**, v. 9, n. 5, 1 maio 2019.

WANT, E. J. et al. Solvent-dependent metabolite distribution, clustering, and protein extraction for serum profiling with mass spectrometry. **Analytical Chemistry**, v. 78, n. 3, p. 743–752, 2006.

WESTERHUIS, J. A. et al. Assessment of PLS-DA cross validation. **Metabolomics**, v. 4, n. 1, p. 81–89, 2008.

WHO. Zika Epidemiology Update. **WHO**, n. July, p. 1–14, 2019.

YAMAMOTO, T.; YAKU, K.; NAKAGAWA, T. Simultaneous measurement of amino acid enantiomers in aged mouse brain samples by lc/ms/ms combined with derivatization using α -(5-fluoro-2,4-dinitrophenyl)-l-leucinamide (l-fdla). **Metabolites**, v. 11, n. 1, p. 1–13, 2021.

YI, L. et al. Chemometric methods in data processing of mass spectrometry-based metabolomics: A review. **Analytica Chimica Acta**, v. 914, p. 17–34, 31 mar. 2016.

YIN, P. et al. Preanalytical aspects and sample quality assessment in metabolomics studies of human blood. **Clinical Chemistry**, v. 59, n. 5, p. 833–845, 2013.

YU, P. et al. Metabolomics profiles in umbilical cord blood. **Journal of King Saud University - Science**, v. 32, n. 8, p. 3213–3216, 2020.

YUN, S. I.; LEE, Y. M. Zika virus: An emerging flavivirus. **Journal of Microbiology**, v. 55, n. 3, p. 204–219, 2017.

ZAMPOLLI, M. et al. GC-MS analysis of amino acid enantiomers as their N(O,S)-

perfluoroacyl perfluoroalkyl esters: Application to space analysis. **Chirality**, v. 18, n. 4, p. 279–295, 2006.

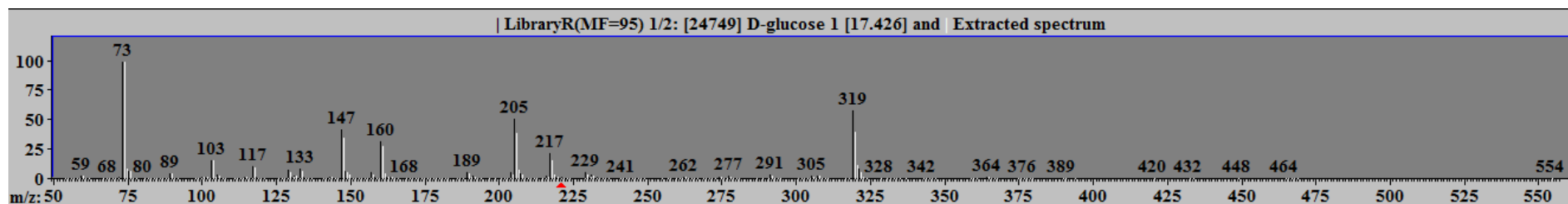
ZEKI, Ö. C. et al. Integration of GC–MS and LC–MS for untargeted metabolomics profiling. **Journal of Pharmaceutical and Biomedical Analysis**, v. 190, p. 113509, 25 out. 2020.

ZHANG, W. et al. Assessing the suitability of capillary electrophoresis-mass spectrometry for biomarker discovery in plasma-based metabolomics. **Electrophoresis**, v. 40, n. 18–19, p. 2309–2320, 2019.

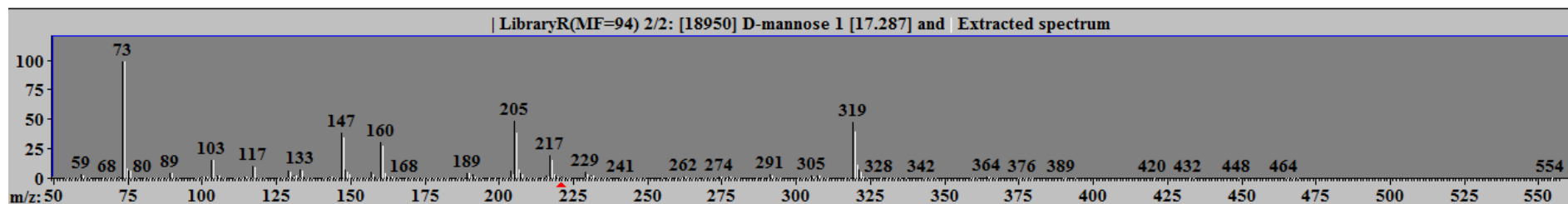
ZHANG, W.; HANKEMEIER, T.; RAMAUTAR, R. Next-generation capillary electrophoresis–mass spectrometry approaches in metabolomics. **Current Opinion in Biotechnology**, v. 43, p. 1–7, 1 fev. 2017.

ANEXO A – Espectros de massas dos isômeros

Espectro de massas da D-glicose (A) vs espectro de massas da D-manose (B)



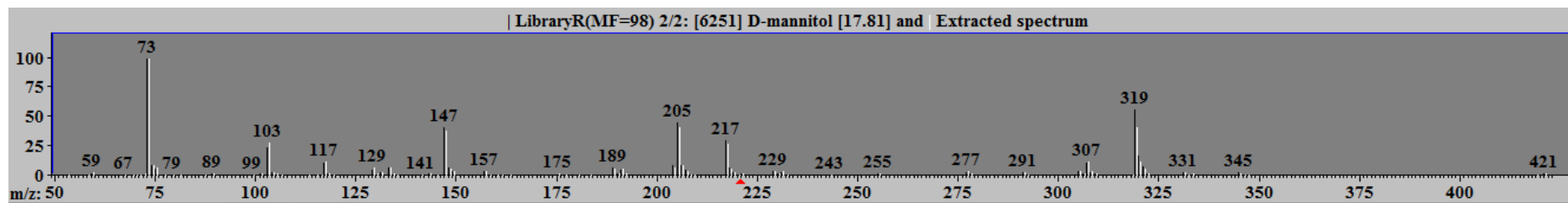
A



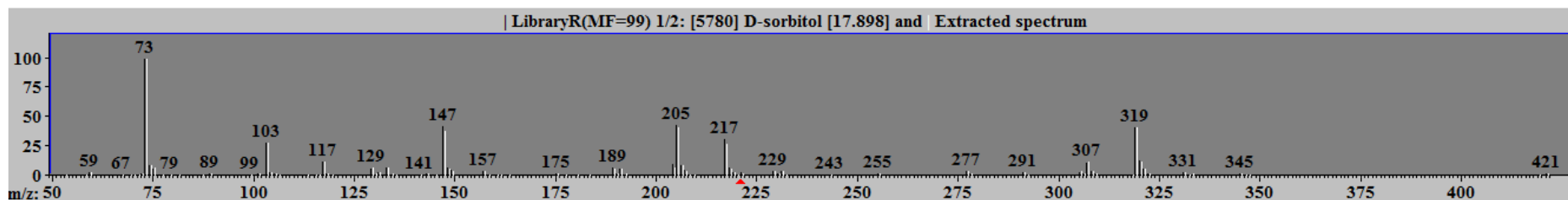
B

Legenda: Os picos em preto representam o espectro de massas da biblioteca Fiehn RT library, e os picos em branco representam o espectro de massas gerado pelo analito.

Espectro de massas do D-manitol (A) vs espectro de massas do D-sorbitol (B)



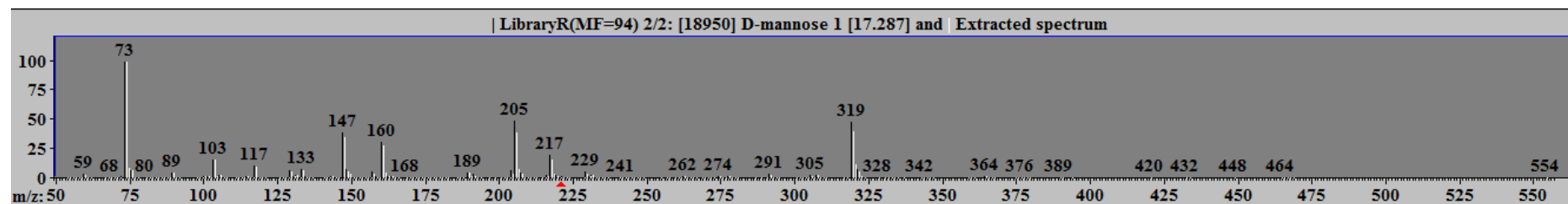
A



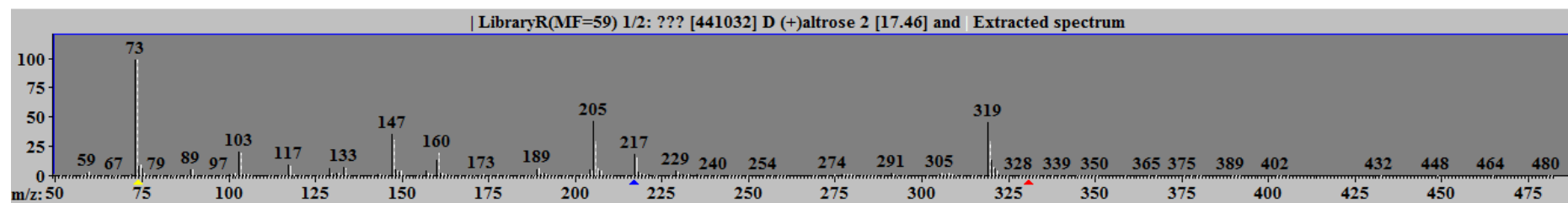
B

Legenda: Os picos em preto representam o espectro de massas da biblioteca Fiehn RT library, e os picos em branco representam o espectro de massas gerado pelo analito.

Espectro de massas da D-manose (A) vs espectro de massas da D-altrose (B)



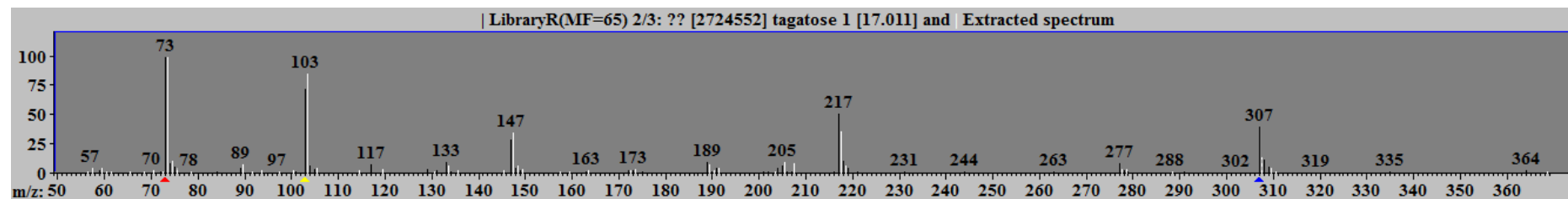
A



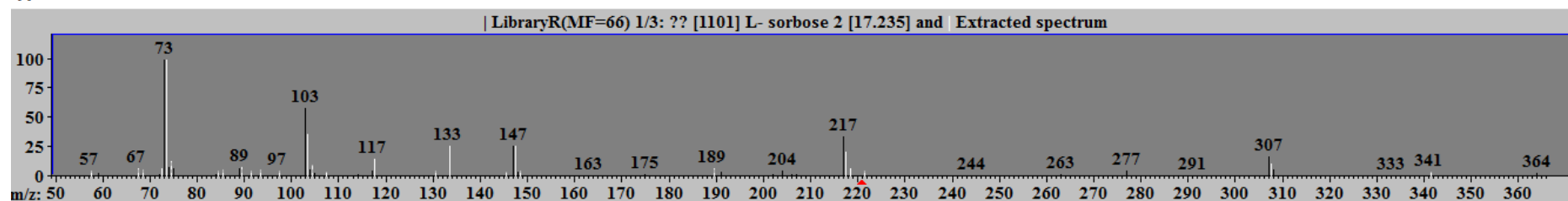
B

Legenda: Os picos em preto representam o espectro de massas da biblioteca Fiehn RT library, e os picos em branco representam o espectro de massas gerado pelo analito.

Espectro de massas da tagatose (A) vs espectro de massas da L-sorbose (B)



A



B

Legenda: Os picos em preto representam o espectro de massas da biblioteca Fiehn RT library, e os picos em branco representam o espectro de massas gerado pelo analito.