



**UNIVERSIDADE ESTADUAL DE CAMPINAS**

**INSTITUTO DE QUÍMICA**

**HENRIQUE CARACHO RIBEIRO**

**INTEGRAÇÃO DE DADOS DE CIÊNCIAS ÔMICAS EMPREGANDO ESTRATÉGIAS  
BIOINFORMÁTICAS NO ESTUDO DO TRANSTORNO BIPOLAR**

**CAMPINAS**

**2023**

**HENRIQUE CARACHO RIBEIRO**

**INTEGRAÇÃO DE DADOS DE CIÊNCIAS ÔMICAS EMPREGANDO ESTRATÉGIAS  
BIOINFORMÁTICAS NO ESTUDO DO TRANSTORNO BIPOLAR**

Tese de Doutorado apresentada ao Instituto de Química da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Ciências.

**Orientadora: Profa. Dra. Alessandra Sussulini**

**O arquivo digital corresponde à versão final da Tese defendida pelo aluno Henrique Caracho Ribeiro e orientada pela Profa. Dra. Alessandra Sussulini.**

**CAMPINAS**

**2023**

Ficha catalográfica  
Universidade Estadual de Campinas  
Biblioteca do Instituto de Química  
Simone Luiz Alves - CRB 8/9094

R354i Ribeiro, Henrique Caracho, 1989-  
Integração de dados de ciências ômicas empregando estratégias  
bioinformáticas no estudo do transtorno bipolar / Henrique Caracho Ribeiro. –  
Campinas, SP : [s.n.], 2023.

Orientador: Alessandra Sussulini.  
Tese (doutorado) – Universidade Estadual de Campinas, Instituto de  
Química.

1. Transtorno bipolar. 2. Metabolômica. 3. Multi-ômicas. 4. Biologia de  
sistemas. 5. Esquizofrenia. I. Sussulini, Alessandra, 1981-. II. Universidade  
Estadual de Campinas. Instituto de Química. III. Título.

Informações Complementares

**Título em outro idioma:** Integration of omics data using bioinformatics approaches in the  
study of bipolar disorder

**Palavras-chave em inglês:**

Bipolar disorder  
Metabolomics  
Multi-omics  
Systems biology  
Schizophrenia

**Área de concentração:** Química Analítica

**Titulação:** Doutor em Ciências

**Banca examinadora:**

Leandro Wang Hantao  
Samara Flamini Kiihl  
Alexandra Christine Helena Frankland Sawaya  
Helder Takashi Imoto Nakaya  
Rafael Garrett da Costa

**Data de defesa:** 20-04-2023

**Programa de Pós-Graduação:** Química

**Identificação e informações acadêmicas do(a) aluno(a)**

- ORCID do autor: <https://orcid.org/0000-0002-0759-3895>

- Currículo Lattes do autor: <http://lattes.cnpq.br/3291324972012955>

## **BANCA EXAMINADORA**

Profa. Dra. Alessandra Sussulini (Orientadora)

Profa. Dra. Samara Flamini Kiihl (Universidade Estadual de Campinas – Unicamp)

Profa. Dra. Alexandra Christine Helena Frankland Sawaya (Universidade Estadual de Campinas – Unicamp)

Prof. Dr. Helder Takashi Imoto Nakaya (Universidade de São Paulo – USP)

Prof. Dr. Rafael Garrett da Costa (Universidade Federal do Rio de Janeiro – UFRJ)

A Ata da defesa assinada pelos membros da Comissão Examinadora, consta no SIGA/Sistema de Fluxo de Dissertação/Tese e na Secretaria do Programa da Unidade.

Este exemplar corresponde à redação final da Tese de Doutorado defendida pelo aluno **HENRIQUE CARACHO RIBEIRO**, aprovada pela Comissão Julgadora em 20 de abril de 2023.

*“Se ansiamos que nosso planeta seja importante, há algo que possamos fazer quanto a isso. Podemos fazer com que ele seja significativo com a coragem de nossas perguntas e a profundidade de nossas respostas.”*

**Cosmos**, de Carl Sagan

## AGRADECIMENTOS

Agradeço aos meus pais, José e Maria Helena, por todo o apoio incondicional ao longo de todos os anos, em especial nos anos de vida acadêmica, acreditando em mim e sempre me incentivando a ir além.

Aos meus irmãos, Erika e Gabriel, por todo o companheirismo, discussões (acadêmicas ou não), conversas, conselhos, desabafos, e até pelas brigas, afinal, existem irmãos que nunca brigaram?

Aos meus amigos da Química Unicamp 09, Rafael Scucuglia, Emerson, Rodrigo, Marina, Rafael Mendonça e Hugo, por todos os anos de amizade, estudos, provas, noites mal dormidas, relatórios, festas e conquistas que comemoramos juntos, às vezes nos separando devido às mazelas da vida adulta, mas nunca se esquecendo um do outro totalmente.

Aos meus amigos (também 09) Júlia Helena e Rafael de Santi, pela amizade que se estende desde o início da graduação e toda a caminhada juntos, nos dias bons e nos dias ruins, nessa montanha russa chamada Unicamp.

A todos os meus amigos da República do Patrão, em especial para Edison, Lucas Girello, Lucas Ryu, Daniel e Thales, Rafael, Danilo, Marcelo e à Tekila, pelos anos de convivência, dentro ou fora da república, cuja forma de amizade fez com que considerássemos uns aos outros como uma segunda família, sempre estando um para o outro para o que precisar.

Ao meu grande amigo Guilherme Bottino, junto comigo nesta caminhada acadêmica desde o início, pelos dias estudando juntos, pelas noites a fio terminando apresentações para serem entregues e apresentadas no dia seguinte (na última hora as ideias sempre fluem melhor, ou não), por suportar minhas (às vezes nem tanto) repentinas explosões, pelas conversas filosóficas, e pela amizade ao longo desses mais de 12 anos.

Aos meus amigos do Retrodorm de Turku, na Finlândia, em especial à Phillip e Kim, por fazerem ser mais tolerável estar em um país estrangeiro no meio de uma pandemia global, por todos os dias de conversa na sacada, de caminhadas na beira do rio Aura, partidas de LoL, e companheirismo.

A todos os funcionários do Biocity Turku, pela forma maravilhosa na qual me acolheram, pela simpatia e carinho, sempre prestativos para as mais diversas questões.

A Matej Orešič e Alex Dickens, por me receberem em Turku e me orientarem no meu período sanduíche, sempre dispostos a ajudar no que fosse necessário.

A todos do System Medicine Group, Annie, Peppi, Tuomas e Santosh, pelo suporte e disposição de ajudar nos mais diversos assuntos.

A Partho Sen, pela incrível paciência de me ajudar e ensinar toda a metodologia de bioinformática para o desenvolvimento da tese, conselhos, indicações, sugestões e códigos debugados.

A minha amiga Marina Amaral, por todas as viagens, conversas, pela parceria e amizade, sendo um pedaço do Brasil na Finlândia, e a Filipe Motta e Edu, pela amizade mútua nesse eixo Rio-São Paulo.

A Gustavo Duarte, pelo apoio e suporte experimental.

Ao Professor Anderson Rocha, do IC-Unicamp e a todos de seu grupo de pesquisa pela disponibilidade em ajudar no desenvolvimento das metodologias de aprendizado de máquina.

A todos os funcionários do IQ-Unicamp, pela prestatividade e disposição para ajudar no que estiver ao alcance.

A todos os meus amigos do LaBIOmics, Elisa, Flávia, Isabella, Donizete, Luidy, Luan, Taynara, Larissa, Mariana, Thales, Patrick, Bruna, Aline e Acácia, por todos os anos de convivência, cafés, happy hours, discussões científicas (ou não) e amizade dentro e fora do laboratório.

Finalmente, à minha orientadora Alessandra Sussulini, por todos os anos de trabalho juntos desde o mestrado, por acreditar em mim e nos meus projetos e me incentivar academicamente, me ensinando e ajudando a crescer como cientista ao longo desses quase 8 anos juntos, sempre respeitando princípios da ciência (que deveriam ser seguidos por todos da academia, mas infelizmente não são), e se norteados pela ética junto aos seus colegas de profissão.

Agradeço também ao Ministério da Educação da Finlândia e ao Prof. Matej Oresic e à Universidade de Turku pela bolsa de doutorado sanduíche.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001.

## RESUMO

Os transtornos mentais são definidos como um conjunto de condições clínicas que envolvem prejuízos no funcionamento social, modulados pela interação de fatores genéticos e ambientais. Dentre os diversos transtornos mentais elencados no Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-V), o transtorno bipolar (BD) e a esquizofrenia (SZ) são particularmente desafiadores, devido à sua alta ocorrência na população adulta mundial, com incidência de cerca de 1%, e devido à sua complexidade diagnóstica e fisiopatológica, especialmente por apresentarem similaridades sintomáticas. Diante desta problemática, a busca por testes clínicos que auxiliem o diagnóstico preciso de cada doença, bem como estudos que ajudem a melhorar a compreensão das alterações bioquímicas causadas por elas, são desejados pela comunidade médica e científica. Neste contexto, abordagens baseadas em metabolômica são particularmente interessantes, devido ao fato dos metabólitos indicarem o ponto final da expressão gênica e da atividade celular, fornecendo meios de se entender o fenótipo de um organismo. Nesta tese, foram utilizadas abordagens baseadas em metabolômica a fim de se obter mais informações quanto às fisiopatologias do BD e da SZ, auxiliadas por metodologias bioinformáticas distintas e, por meio destas propor metabólitos de interesse que possam também diferenciar pacientes com transtornos mentais de controles saudáveis. Dentre as metodologias bioinformáticas aplicadas estão os métodos baseados em análise de correlação parcial para determinar as interconexões entre as proteínas e metabólitos de pacientes com BD, o que revelou alterações relacionadas à homeostase. Utilizou-se também regressão logística do tipo *ridge* para a determinação de um painel de metabólitos com potencial de diferenciar pacientes com BD de indivíduos saudáveis. Expandindo a pesquisa para dois transtornos mentais, optou-se por avaliar a performance de duas metodologias de aprendizado de máquina na diferenciação multi-grupo de pacientes com BD e SZ em comparação com controles saudáveis, o que revelou possíveis alterações no sistema endocanabinoide. Por fim, espera-se que os resultados obtidos nesta Tese possam auxiliar na elucidação das bases bioquímicas do BD e da SZ, além de lançarem luz à importância da bioinformática dentro da metabolômica, reiterando a importância da interdisciplinaridade para a pesquisa científica, em especial no campo da química.

## ABSTRACT

Mental disorders are defined as a set of clinical conditions that involve impairments in social functioning, modulated by the interaction of genetic and environmental factors. Among the various mental disorders listed in the Diagnostic and Statistical Manual of Mental Disorders (DSM-V), bipolar disorder (BD) and schizophrenia (SZ) are particularly challenging due to their high occurrence in the global adult population, with an incidence of around 1%, and due to their diagnostic and pathophysiological complexity, especially as they present symptomatic similarities. Given this problem, the search for clinical tests that assist in the precise diagnosis of each disease, as well as studies that help to understand the biochemical alterations caused by them, are desired by the medical and scientific community. In this context, metabolomics-based approaches are particularly interesting, because metabolites indicate the endpoint of gene expression and cellular activity, providing a way to understand the phenotype of an organism. In this Thesis, we used metabolomics-based approaches to obtain more information on the pathophysiology of BD and SZ, assisted by different bioinformatics methodologies and, through these, metabolites that could differentiate patients with mental disorders from healthy controls were proposed. Among the bioinformatics methodologies applied, partial correlation analysis-based methods were used to determine the interconnections between proteins and metabolites in patients with BD, which revealed alterations related to hemostasis. Ridge logistic regression was also used to determine a panel of metabolites with the potential to differentiate BD patients from healthy individuals. Expanding the research to two mental disorders, the performance of two machine-learning methodologies in the multi-group differentiation of patients with BD and SZ compared to healthy controls were tested, which revealed possible alterations in the endocannabinoid system. Finally, the results obtained in this Thesis are expected to support the elucidation of the biochemical bases of BD and SZ, as well as shed light on the importance of bioinformatics within metabolomics, reiterating the importance of interdisciplinarity for scientific research, especially in the field of chemistry.

## LISTA DE FIGURAS

<b>Figura 1.1.</b> Características do transtorno bipolar tipo I (a) e tipo II (b) correlacionadas às suas oscilações de humor e ao tempo (adaptado de Vieta <i>et al.</i> , 2018).....	<b>25</b>
<b>Figura 1.2.</b> Representação esquemática de um espectrômetro de massas. Analisadores de massas podem ser simples (somente um analisador de massas) ou híbridos (contendo dois - ou mais - analisadores de massas, com uma célula de colisão entre eles) e, em alguns casos, o detector pode estar localizado junto com o analisador (como é o caso de analisadores baseados em transformada de Fourier, como Orbitrap e ressonância ciclotrônica de íons).	<b>27</b>
<b>Figura 1.3.</b> Os diferentes tipos de ciências ômicas e suas interações dentro da biologia de sistemas, também podendo ser conhecida como cascata ômica (adaptado de Santos <i>et al.</i> , 2020). .....	<b>31</b>
<b>Figura 1.4.</b> Publicações em multi-ômicas envolvendo dados proteômicos e metabolômicos ao longo dos anos na plataforma <i>Scopus</i> . .....	<b>34</b>
<b>Figura 1.5.</b> Gráfico de <i>boxplot</i> dos dados de metabolômica para cada amostra antes e após normalização, transformação logarítmica e escalamento.....	<b>46</b>
<b>Figura 1.6.</b> Gráfico de escores da análise de componentes principais (PCA) dos dados metabolômicos obtidos por GC-MS. ....	<b>47</b>
<b>Figura 1.7.</b> Gráfico de escores da PLS-DA multi-bloco dos dados metabolômicos (A), proteômicos (B), e os metabólitos e proteínas mais significativos (C e D), que serão utilizados na análise de correlação parcial. ....	<b>48</b>
<b>Figura 1.8.</b> Principais fatores de confusão nos conjuntos de dados de metabolômica e proteômica, demonstrando a ordem de influência na variância dos dados (Caso-Controle (CC) > tipo de medicação > IMC > gênero > idade). ....	<b>49</b>
<b>Figura 1.9.</b> Rede de correlação parcial ilustrando as interações entre metabólitos (M), proteínas (P) e os dados demográficos.....	<b>50</b>
<b>Figura 1.10.</b> <i>Beanplots</i> dos metabólitos selecionados na análise de correlação parcial e suas distribuições entre os grupos estudados. As linhas menores apresentam as observações individuais referentes a cada amostra, enquanto as áreas amarelas e verdes representam a área de distribuição total das amostras e a linha maior representa a média de tais valores individuais. ....	<b>57</b>
<b>Figura 1.11.</b> Painel de metabólitos construído utilizando modelagem de regressão logística [BD x HC] ajustado para idade, gênero e IMC. ....	<b>58</b>
<b>Figura 2.1.</b> Metodologias de aprendizado de máquina aplicadas a abordagens metabolômicas em transtornos mentais na plataforma Pubmed. ....	<b>66</b>
<b>Figura 2.2.</b> Forma de utilização da metodologia de aprendizado de máquina nos estudos metabolômicos aplicados a transtornos mentais na plataforma Pubmed. ....	<b>67</b>
<b>Figura 2.3.</b> Transtornos mentais estudados utilizando metodologias de aprendizado de máquina na plataforma Pubmed. ....	<b>67</b>

<b>Figura 2.4.</b> Diferenças de resolução entre equipamentos de baixa resolução (QqQ) e alta resolução (HR). Adaptado de Rochat, 2016.....	<b>71</b>
<b>Figura 2.5.</b> Resolução usual de diferentes tipos de espectrômetros de massa (Orbitrap, QTOF e QqQ) em diferentes <i>m/z</i> . Adaptado de Rochat, 2016.....	<b>71</b>
<b>Figura 2.6.</b> Estrutura e classe das diversas subclasses das amidas graxas, onde R <sub>1</sub> corresponde a um grupo acila com três ou mais átomos de carbono, e o radical R <sub>2</sub> dos N-acilaminoácidos caracteriza o aminoácido em questão. Adaptado de Farrell & Merkle, 2008. ....	<b>75</b>
<b>Figura 2.7.</b> Ordem de importância dos <i>features</i> selecionados na regressão LASSO para a classificação do grupo BD versus o resto dos grupos.....	<b>87</b>
<b>Figura 2.8.</b> Ordem de importância dos <i>features</i> selecionados na regressão LASSO para a classificação do grupo SZ versus o resto dos grupos. ....	<b>88</b>
<b>Figura 2.9.</b> Ordem de importância dos <i>features</i> selecionados na regressão LASSO para a classificação do grupo HC versus o resto dos grupos.....	<b>88</b>
<b>Figura 2.10.</b> Relação do transtorno bipolar com alterações inflamatórias e na cascata de ácido araquidônico. Possíveis atenuadores relacionados a receptores endocanabinóides dos processos indicados são ilustrados pelas linhas vermelhas. Adaptado de Arjmand et al., 2019. ....	<b>100</b>
<b>Figura 2.11.</b> Relação da esquizofrenia com alterações inflamatórias e produção de dopamina. Possíveis atenuadores relacionados a receptores endocanabinóides dos processos indicados são ilustrados pelas linhas vermelhas e potenciais compostos endocanabinóides relacionados à melhora dos sintomas psicóticos são ilustrados em verde. Adaptado de Arjmand et al., 2019 e Rohleder & Leweke, 2015).....	<b>102</b>

## LISTA DE TABELAS

<b>Tabela 1.1.</b> Dados clínicos e demográficos das amostras utilizadas na análise metabolômica, conjunto de dados completo (SD: desvio padrão; HC: controles saudáveis; BD: pacientes com transtorno bipolar).....	<b>39</b>
<b>Tabela 1.2.</b> Dados clínicos e demográficos das amostras pareadas nos estudos de metabolômica e proteômica (SD: desvio padrão; HC: controles saudáveis; BD: pacientes com transtorno bipolar).....	<b>40</b>
<b>Tabela 1.3.</b> Parâmetros utilizados no pré-processamento dos dados no software MS-DIAL	<b>42</b>
<b>Tabela 2.1.</b> Dados demográficos das amostras utilizadas na análise de metabólitos utilizando aprendizado de máquina (SD: desvio padrão; BD: pacientes com transtorno bipolar; SZ: pacientes com esquizofrenia; HC: controles saudáveis) .....	<b>81</b>
<b>Tabela 2.2.</b> Parâmetros utilizados no pré-processamento dos dados no software MS-DIAL	<b>83</b>
<b>Tabela 2.3.</b> Identificações preliminares dos <i>molecular features</i> considerados diferenciais nas metodologias LASSO, em ambas as metodologias, e na metodologia boruta, respectivamente .....	<b>91</b>

## LISTA DE ABREVIATURAS E ACRÔNIMOS

2-AG	2-Araquidonoilglicerol
AA	Ácido araquidônico
AEA	Anandamida
AIF	<i>All-ion fragmentation</i> (Fragmentação de todos os íons)
APCI	<i>Atmospheric pressure photoionization</i> (Fotoionização à pressão atmosférica)
APPI	<i>Atmospheric pressure chemical ionization</i> (Ionização química à pressão atmosférica)
AUC	<i>Area under curve</i> (Área sob a curva)
BD	<i>Bipolar disorder</i> (Transtorno bipolar)
BSTFA	N,O-Bis(trimetil)silil trifluoroacetamida
CB1	Receptor canabinoide tipo 1
CB2	Receptor canabinoide tipo 2
CBD	Canabidiol
CC	Caso-Controle
CE	<i>Capillary electrophoresis</i> (Eletroforese capilar)
CI	<i>Chemical ionization</i> (Ionização química)
COX-2	Ciclooxigenase-2
Da	Dalton
DCVs	Doenças cardiovasculares
DDA	<i>Data dependent acquisition</i> (Aquisição dependente de dados)
DESI	<i>Desorption electrospray ionization</i> (Ionização de dessorção por eletrospray)
DIA	<i>Data independent acquisition</i> (Aquisição independente de dados)
DSM-V	<i>Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition</i> (Manual Diagnóstico e Estatístico de Transtornos Mentais, Quinta Edição)
EDTA	Ácido etilenodiaminotetracético
EI	<i>Electron ionization</i> (Ionização por elétrons)
ERK1/2	<i>Extracellular signal-regulated kinase</i> (Quinases reguladas por sinal extracelular)
ESI	<i>Electrospray ionization</i> (Ionização por eletrospray)
FA	<i>Fatty acyls</i> (Acil graxos)
FAAH	<i>Fatty acid acyl hydrolase</i> (Acil hidrolase de ácido graxo)
FWHM	<i>Full width half maximum</i> (Largura total do pico na metade de sua altura máxima)
GABA	Ácido gama-aminobutírico
GBD	<i>Global Burden of Disease</i>
GC	<i>Gas chromatography</i> (Cromatografia em fase gasosa)
GC-MS	<i>Gas chromatography coupled to mass spectrometry</i> (Cromatografia em fase gasosa acoplada à espectrometria de massas)

GMD	<i>Golm Metabolome Database</i> (Banco de Dados de Metaboloma Golm)
GPR	<i>G-coupled receptors</i> (Receptores acoplados às proteínas G)
HAM-D	<i>Hamilton Depression Scale</i> (Escala de Depressão de Hamilton)
HC	<i>Healthy controls</i> (Controles saudáveis)
HDL	<i>High-density lipoprotein</i> (Lipoproteína de alta densidade)
HETE	<i>Hydroxyeicosatetraenoic acid</i> (Ácido hidroxeicosatetranóico)
HFBCF	1,1,1,2,2,3,3-Heptafluorobutil cloroformiato
HILIC	<i>Hydrophilic interaction liquid chromatography</i> (Cromatografia líquida de interação hidrofílica)
HMDB	<i>Human Metabolome Database</i> (Banco de Dados de Metaboloma Humano)
HPLC	<i>High performance liquid chromatography</i> (Cromatografia líquida de alta eficiência)
HR	<i>High resolution</i> (Alta resolução)
IMC	Índice de massa corpórea
IPA	<i>Ingenuity Pathway Analysis</i>
IT	<i>Ion trap</i> (Armadilha de íons)
KEGG	<i>Kyoto Encyclopedia of Genes and Genomes</i> (Enciclopédia de Genes e Genomas de Kyoto)
KNN	<i>K-nearest neighbors</i> (K-vizinhos mais próximos)
LASSO	<i>Least absolute shrinkage and selection operator</i>
LC-MS	<i>Liquid chromatography coupled to mass spectrometry</i> (Cromatografia líquida acoplada à espectrometria de massas)
LDL	<i>Low-density lipoprotein</i> (Lipoproteína de baixa densidade)
<i>m/z</i>	Relação massa/carga
MALDI	<i>Matrix-assisted laser desorption/ionization</i> (Dessorção/ionização à laser assistida por matriz)
MCF	Metilcloroformiato
MeOH	Metanol
miRNA	micro-RNA (Ácido ribonucléico)
MOX	Metoxiamina
MS	<i>Mass spectrometry</i> (Espectrometria de massas)
MS/MS	<i>Tandem mass spectrometry</i> (Espectrometria de massas sequencial)
MS-DIAL	<i>Mass Spectrometry – Data Independent AnaLysis</i>
MSI	<i>Metabolomics Standards Initiative</i>
MSTFA	N-metil-N-trimetilsililfluoroacetamida
NAA	<i>N-acyl aminoacids</i> (N-acil aminoácidos)
NAE	<i>N-acyl ethanolamines</i> (N-acil etanolaminas)
NAGABA	<i>N-Arachidonoyl-<math>\gamma</math>-aminobutyric acid</i> (N-araquidonoil- $\gamma$ -aminobutarato)
NAGly	<i>N-Arachidonoyl glycine</i> (N-araquidonoilglicina)

NAM	<i>N-acyl amides</i> (N-acil amidas)
NDA	<i>N-acyl dopamines</i> (N-acil dopaminas)
NIST	<i>National Institute of Standards and Technology</i> (Instituto Nacional de Padrões e Tecnologia)
NMDA	<i>N-methyl-D-aspartate</i> (N-metil-D-aspartato)
OEA	<i>Oleoylethanolamine</i> (Oleoiletanolamida)
PANSS	<i>Positive and Negative Syndrome Scale</i> (Escala de Síndromes Positivas e Negativas)
PC	<i>Phosphatidylcholine</i> (Fosfatidilcolina)
PCA	<i>Principal component analysis</i> (Análise de componentes principais)
PE-NMe	<i>Acylglycerophosphoethanolamines</i> (Acilglicerofosfoetanolaminas)
PES	Poliétersulfona
PFAM	<i>Primary fatty acid amides</i> (Amidas graxas primárias)
PG	<i>Phosphatidylglycerol</i> (Fosfatidilglicerol)
PGG <sub>2</sub> -G	<i>Prostaglandin G2-glycerol ester</i> (Prostaglandina G2 gliceril éster)
PGH <sub>2</sub> -Gly	<i>Prostaglandin H2-glycine</i> (Prostaglandina H2 glicina)
PI	<i>Phosphatidylinositol</i> (Fosfatidilinositol)
PLA2	<i>Phospholipase A2</i> (Fosfolipase A2)
PLS-DA	<i>Projection to latent structures discriminant analysis</i> (Análise discriminante por projeção a estruturas latentes)
PPAR $\alpha$	<i>Peroxisome proliferator-activated receptor</i> (Receptor ativado por proliferador de peroxissoma)
PR	Poder de resolução
PRM	<i>Parallel reaction monitoring</i> (Monitoramento paralelo de reações)
QC	<i>Quality control</i> (Controle de qualidade)
QIT	<i>Quadrupole-ion trap</i> (Quadrupolo-armadilha de íons)
QqQ	Triplo quadrupolo
QTOF	<i>Quadrupole-time of flight</i> (Quadrupolo-tempo de voo)
rCCA	<i>Regularized canonical correlation analysis</i> (Análise de correlação canônica regularizada)
ROC	<i>Receiver operator characteristics</i> (Caracterização de operação de receptor)
RSD	<i>Relative standard deviation</i> (Desvio padrão relativo)
SCID-P	<i>Structured Clinical Interview Axis I</i> (Entrevista Clínica Estruturada, Eixo I)
SD	<i>Standard deviation</i> (Desvio padrão)
SIM	<i>Selected ion monitoring</i> (Monitoramento de íon selecionado)
SIRIUS	<i>Sum formula Identification by Ranking Isotope patterns Using mass Spectrometry</i>
SMet	Síndrome metabólica
SPB	<i>Sphingoid base</i> (Base esfingóide)
sPLS-DA	<i>Sparse projection to latent structures discriminant analysis</i> (Análise discriminante por projeção esparsa a estruturas latentes)

SRM	<i>Selected reaction monitoring</i> (Monitoramento de reações selecionadas)
ST	<i>Sterol lipids</i> (Lipídios esteroides)
SVM	<i>Support vector machines</i> (Máquinas de vetor de suporte)
SWATH	<i>Sequential Window Acquisition of all THEoretical fragment-ion spectra</i>
SZ	<i>Schizophrenia</i> (Esquizofrenia)
TCMS	Trimetilclorosilano
TG	<i>Triacylglycerol</i> (Triacilglicerol)
THC	<i>Tetrahydrocannabinol</i> (Tetraidrocanabinol)
TIC	<i>Total ion chromatogram</i> (Cromatograma de íons totais)
TRPV1	<i>Vanilloid receptor type 1</i> (Receptores vanilóides tipo 1)
TXA2	<i>Thromboxane A2</i> (Tromboxano A2)
UHPLC	<i>Ultra high performance liquid chromatography</i> (Cromatografia líquida de ultra alta eficiência)
VIP	<i>Variable importance in projection</i> (Importância da variável na projeção)
WGCNA	<i>Weighted correlation network analysis</i> (Análise de redes de correlação por pesos)
YLD	<i>Years lived with disability</i> (Anos vividos com deficiências)
YMRS	<i>Young Mania Rating Scale</i> (Escala Classificatória de Mania de Young)

## SUMÁRIO

<b>Introdução geral e justificativa .....</b>	<b>19</b>
<b>CAPÍTULO 1</b>	
<b>Integração de dados metabolômicos e proteômicos de pacientes com transtorno bipolar e sua relação com a hemóstase.....</b>	<b>23</b>
1. Introdução.....	24
1.1. Transtorno bipolar e suas comorbidades.....	26
1.2. Espectrometria de massas e suas aplicações em metabolômica .....	23
1.3. Biologia de sistemas e integração de dados ômicos .....	30
2. Objetivos.....	37
2.1. Objetivos gerais .....	37
2.2. Objetivos específicos .....	37
3. Parte Experimental .....	38
3.1. Pacientes e amostragem.....	38
3.2. Análise metabolômica utilizando cromatografia gasosa acoplada à espectrometria de massas (GC-MS) .....	40
3.3. Proteômica.....	43
3.4. Identificação, análise estatística e integração de dados ômicos.....	43
4. Resultados e Discussão.....	46
4.1. Análise estatística e seleção de variáveis .....	46
4.2. Interações entre proteínas e metabólitos no BD utilizando correlação parcial em formato de rede.....	49
4.3. Painel de metabólitos como preditor do BD.....	56
5. Conclusões .....	62
<b>CAPÍTULO 2</b>	
<b>Avaliação de duas metodologias de aprendizado de máquina na diferenciação multi-grupo de pacientes com transtorno bipolar e esquizofrenia e sua relação com amidas graxas .....</b>	<b>63</b>
1. Introdução .....	64
1.1. Metodologias de aprendizado de máquina em estudos metabolômicos aplicados a transtornos mentais.....	64
1.2. Espectrometria de massas de alta resolução .....	69
1.3. Amidas graxas e sua relação com alterações cerebrais.....	74
2. Objetivos.....	79
2.1. Objetivos gerais .....	79
2.2. Objetivos específicos .....	79
3. Parte Experimental .....	80

3.1. Pacientes e amostragem.....	80
3.2. Análise metabolômica utilizando espectrometria de massas de alta resolução com infusão direta .....	81
3.3. Pré-tratamento e aplicação de metodologias de aprendizado de máquina.....	82
3.4. Identificação putativa dos metabólitos selecionados .....	84
4. Resultados e Discussão.....	86
4.1. Metodologias de aprendizado de máquina na diferenciação de pacientes com BD e SZ em comparação com controles saudáveis.....	86
4.2. Identificação putativa dos metabólitos selecionados .....	90
4.3. N-acilamidas, amidas graxas e sua relação com transtornos psiquiátricos .....	95
5. Conclusões .....	103
<b>Considerações finais.....</b>	<b>104</b>
<b>Referências .....</b>	<b>107</b>
<b>ANEXOS .....</b>	<b>137</b>

## Introdução geral e justificativa

Em 1990, a Organização Mundial da Saúde, juntamente com o Banco Mundial, criou um consórcio de pesquisadores para avaliar a carga de diferentes doenças para quantificar os efeitos de cada doença na redução da saúde mundial, tal qual seus efeitos na expectativa de vida da população, para que políticas de saúde pública pudessem ser ajustadas de acordo com a necessidade de cada país ou região, conhecido como GBD (do inglês, *Global Burden of Disease*) (Murray et al., 1994). Dentre as doenças analisadas pelo consórcio com a maior quantidade de anos vividos com deficiências (YLD, do inglês *years lived with disability*), transtornos mentais foram o segundo maior no ano de 2019 (GBD 2019 Mental Disorders Collaborators, 2022), causando um grande impacto na saúde pública mundial. Dentre os diversos transtornos mentais, o transtorno bipolar (BD) e a esquizofrenia (SZ) apresentam incidência de cerca de 1% na população mundial e, devido à complexidade diagnóstica e fisiopatológica intrínseca de cada doença, são particularmente desafiadoras para a saúde pública (McCutcheon et al., 2020; Vieta et al., 2018).

O transtorno bipolar, anteriormente conhecido como transtorno maníaco-depressivo, é caracterizado pela alternância entre períodos depressivos e maníacos ou hipomaníacos, dividido majoritariamente em dois subtipos, BD tipo I (presença de um ou mais episódios maníacos, com ou sem a presença de episódios depressivos) e BD tipo II (presença de um ou mais episódios hipomaníacos, com a presença de ao menos um episódio depressivo), com o início dos sintomas se desenvolvendo ao final da adolescência (Carvalho et al., 2020; Sussulini et al., 2009). O diagnóstico do BD ainda é um desafio para a comunidade médica, devido à intersecção dos sintomas com outros transtornos psiquiátricos, o que pode levar a um diagnóstico incorreto de depressão unipolar e conseqüentemente a tratamentos incorretos que podem desencadear episódios maníacos ou hipomaníacos (Sussulini et al., 2009), além da presença de diversas comorbidades em pacientes com BD que reduzem a qualidade de vida do indivíduo, aumentando o risco de morte devido a outras condições médicas, como doenças cardiovasculares, endócrinas e diabetes (Vieta et al., 2018).

A esquizofrenia (SZ) é caracterizada pela presença de sintomas positivos, como ilusões, alucinações, discurso desorganizado e comportamento catatônico; e

por sintomas negativos como expressão emocional reduzida, com a presença de ao menos dois sintomas pelo período de pelo menos um mês, sendo obrigatoriamente um deles caracterizado como sintoma positivo (Davison *et al.*, 2018; McCutcheon *et al.*, 2020). Assim como o BD, os sintomas da esquizofrenia acontecem ao final da adolescência, e está associada também à redução na expectativa de vida e a diferentes comorbidades ao longo da vida (McCutcheon *et al.*, 2020).

Apesar de ambos os transtornos mentais supracitados possuírem suas particularidades, eles apresentam fatores em comum relevantes para a pesquisa destas condições. Ambas não apresentam uma fisiopatologia clara e as bases bioquímicas que as regem ainda não são totalmente conhecidas, além de não possuírem um teste diagnóstico efetivo que possa precisar a condição do paciente, visto que o diagnóstico ainda é realizado por meio de observações clínicas. Para a resolução deste problema, cada vez mais se faz necessário o uso de ferramentas que possam auxiliar na detecção precoce e no entendimento da trajetória dos diferentes transtornos mentais com especificidade e sensibilidade, voltando-se para o desenvolvimento da chamada medicina personalizada, onde modelos preditivos treinados em um grande número de amostras de uma determinada condição podem ser utilizados para definir o estado de um determinado paciente (Bzdok & Meyer-Lindenberg, 2018; Rajula *et al.*, 2021). Neste contexto, a utilização da metabolômica na medicina personalizada é cada vez mais difundida, o que se deve ao fato dos metabólitos indicarem o ponto final da expressão gênica e da atividade celular, fornecendo meios de se entender o fenótipo de um organismo, o que é fundamental dentro de abordagens de biologia de sistemas (Klassen *et al.*, 2017).

Apesar da metabolômica ser uma ferramenta poderosa para a avaliação de perturbações causadas por doenças ou alterações exógenas (Boccio *et al.*, 2016), o caminho percorrido desde a análise da amostra até a interpretação biológica é extenso e exige o uso de diversas ferramentas para que informações relevantes sejam obtidas, em especial na metabolômica baseada em espectrometria de massas. Dados de espectrometria de massas possuem um peso computacional extenso, sendo necessárias abordagens computacionais e quimiométricas para a determinação de variáveis de interesse e construção de modelos preditivos que possam fornecer informações relevantes para a resolução da problemática proposta pelo estudo. Dentro deste escopo, o uso de abordagens supervisionadas e não supervisionadas

para observar padrões de classificação e agrupamento em conjuntos de amostras e o uso de técnicas de aprendizado de máquina vem sendo cada vez mais difundido para a resolução desta problemática, em especial em abordagens mais complexas, como na integração de dados ômicos, onde propõe-se analisar dois conjuntos de dados de diferentes ômicos de uma maneira integrada, a fim de ser obter informações relevantes quanto às interconexões entre as diferentes biomoléculas analisadas, visando trazer à tona informações que não poderiam ser observadas ao analisar cada ômica individualmente, provendo uma visão mais ampla das modulações em um sistema biológico sob efeito de uma determinada condição, como doenças e outros fatores exógenos (Santos *et al.*, 2020).

Este trabalho de Doutorado propõe o uso de diferentes técnicas bioinformáticas para a integração de dados ômicos e para a diferenciação de dados de espectrometria de massas obtidos após a análise de amostras biológicas de pacientes com diferentes transtornos mentais. No **Capítulo 1** utilizou-se uma abordagem de biologia de sistemas para integrar dados de proteômica previamente obtidos pelo grupo de pesquisa com dados metabolômicos obtidos por meio da análise do soro sanguíneo de pacientes com BD e de indivíduos saudáveis, fazendo uso de técnicas integrativas baseadas na análise de correlação parcial para a construção de uma rede de integração entre proteínas, metabólitos e dados clínicos. Ao se analisar um conjunto de dados expandido de metabolômica, obteve-se também um painel de metabólitos diferenciais com o potencial de diferenciar pacientes com BD de controles saudáveis. No **Capítulo 2**, utilizando espectrometria de massas de alta resolução, analisou-se o plasma de pacientes com BD e SZ em comparação com controles saudáveis e avaliou-se a performance de duas metodologias de aprendizado de máquina para a diferenciação dos dados metabolômicos em um modelo “um vs. o resto”, selecionando assim variáveis de interesse que foram subsequentemente identificadas utilizando dados de fragmentação de espectrometria de massas em sequência (MS/MS). Apesar de abordagens diferentes em cada capítulo, o uso de técnicas bioinformáticas na elucidação de potenciais informações obtidas por meio da metabolômica se faz visível em cada parte do trabalho, indicando a importância da computação científica em trabalhos com base na química. É relevante também citar a importância da produção e divulgação de plataformas de acesso aberto para a comunidade científica, visto que em diversos locais do mundo, inclusive no Brasil, o

custo de aquisição de plataformas pagas de tratamento de dados é extremamente elevado. Pensando nisso, a proposição deste trabalho também é de advogar em defesa destas plataformas, visto que todas as etapas computacionais desta tese foram realizadas utilizando plataformas de acesso aberto.

## CAPÍTULO 1

**INTEGRAÇÃO DE DADOS METABOLÔMICOS E PROTEÔMICOS DE  
PACIENTES COM TRANSTORNO BIPOLAR E SUA RELAÇÃO COM A  
HEMÓSTASE**

Os dados apresentados neste capítulo foram publicados na revista *Metabolomics* em agosto/2022. Para mais informações, acessar a referência:

Ribeiro, H. C., Sen, P., Dickens, A., Santa Cruz, E. C., Orešič, M., & Sussulini, A. (2022). Metabolomic and proteomic profiling in bipolar disorder patients revealed potential molecular signatures related to hemostasis. *Metabolomics*, 18(65). <https://doi.org/10.1007/s11306-022-01924-5>

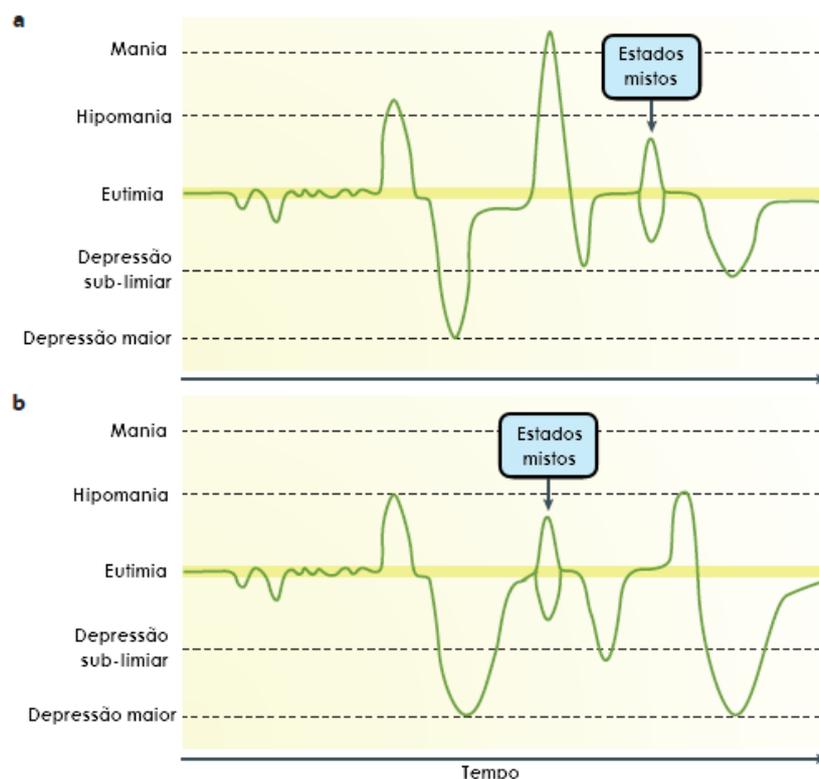
## 1. Introdução

### 1.1. Transtorno bipolar e suas comorbidades

As alterações de humor são comuns na vida cotidiana devido a situações estressantes ou prazerosas, entretanto, alterações severas e persistentes de humor podem ser causadas por transtornos afetivos, que incluem transtornos depressivos unipolares e transtornos bipolares (BD) do tipo I e II (Carvalho *et al.*, 2020). De acordo com a quinta edição do Manual Diagnóstico e Estatístico de Transtornos Mentais (DSM-V) (American Psychiatric Association, 2014), a categoria de BD e transtornos relacionados inclui os transtornos bipolares do tipo I, II e ciclotímico, sendo que a separação destes tipos de transtornos daqueles do tipo depressivo e esquizofrênico, e sua classificação dentro do manual na área entre os dois transtornos previamente mencionados, se dá devido ao reconhecimento do BD como uma ponte entre essas classes diagnósticas em termos de sintomatologia, genética e histórico familiar. A principal característica dos transtornos bipolares que os difere de outros transtornos afetivos é a presença de episódios maníacos ou hipomaníacos, que podem se alternar com episódios depressivos (Carvalho *et al.*, 2020). Os episódios maníacos podem se manifestar de diversas formas, nas quais o indivíduo pode ter casos de superconfiança, grandiosidade, desinibição extrema, irritabilidade, baixa necessidade de sono, humor altamente elevado, hiperatividade, sintomas psicóticos como ilusões ou alucinações, além de eventos que possam comprometer a atividade psicossocial, como gastos financeiros insensatos ou desenfreados e/ou indiscrições sexuais (American Psychiatric Association, 2014; Carvalho *et al.*, 2020; Vieta *et al.*, 2018).

BD do tipo I é caracterizado pela presença de episódios maníacos, com ou sem a presença de episódios depressivos, enquanto o BD tipo II é caracterizado pela presença de episódios de hipomania juntamente com a presença de pelo menos um episódio depressivo (American Psychiatric Association, 2014; Severus & Bauer, 2013; Sussulini *et al.*, 2009), e o transtorno ciclotímico pela presença de sintomas hipomaníacos e depressivos por pelo menos dois anos, sem atender aos critérios definidos para a caracterização de um episódio de mania ou hipomania e depressão

maior (American Psychiatric Association, 2014). Pode-se observar na **Figura 1.1** como estes episódios maníacos, hipomaníacos e depressivos se distribuem ao longo do tempo nos indivíduos acometidos pela doença nos casos de BD tipo I e II.



**Figura 1.1.** Características do transtorno bipolar tipo I (a) e tipo II (b) correlacionadas às suas oscilações de humor e ao tempo (adaptado de Vieta *et al.*, 2018).

A existência de comorbidades com outros transtornos psiquiátricos (transtornos relacionados ao abuso de substâncias, hiperatividade/déficit de atenção, transtornos de ansiedade e personalidade) faz com que o diagnóstico e o tratamento da doença sejam dificultados (Vieta *et al.*, 2018). Em um estudo em ampla população, Merikangas *et al.* (2011) mostraram que cerca de três quartos dos pacientes com BD também apresentaram outros tipos de transtornos psiquiátricos, como transtorno de ansiedade, transtornos comportamentais e de abuso de substâncias. Além da presença de comorbidades com outras doenças psiquiátricas, observou-se também uma alta incidência de comorbidades médicas gerais dentre os pacientes acometidos pelo BD. Diversas doenças foram reportadas em pacientes com BD, como doenças autoimunes, obesidade, diabetes, doenças cardiovasculares (Sayuri Yamagata *et al.*, 2017), síndrome metabólica, doenças respiratórias crônicas, endócrinas,

gastrointestinais, neurológicas, musculoesqueléticas, entre outras doenças (Sinha *et al.*, 2018). Além do risco de suicídio, acredita-se que a redução na expectativa de vida dentre pacientes com BD (em comparação com a população em geral) se dá primariamente devido às diversas comorbidades em tais pacientes (SayuriYamagata *et al.*, 2017). Estudos em médias e amplas populações também foram realizados para avaliar os impactos das comorbidades em pacientes bipolares, onde Sylvia *et al.* (2015) reportaram a presença de pelo menos uma comorbidade em 96,3% de um grupo de 482 pacientes diagnosticados com BD. Crump *et al.* (2013) observaram em um estudo populacional que a expectativa de vida de pacientes bipolares é reduzida em cerca de 9 anos em comparação com o resto da população, com risco de morte aumentado de 1,5 a 2,5 vezes devido a doenças cardiovasculares, 3 vezes para diabetes, 1,5 vezes para doença pulmonar obstrutiva crônica, 3 a 4 vezes para lesões não intencionais e 15 vezes para suicídio, além de risco aumentado também para doenças respiratórias e câncer. Apesar do risco de morte por suicídio ser extremamente elevado, observou-se neste estudo uma incidência de 5,4% e 9,6% nas mortes de pacientes bipolares mulheres e homens, respectivamente. Dentre as comorbidades descritas, acredita-se que doenças cardiovasculares são a principal causa de redução na expectativa de vida nos pacientes bipolares (Leboyer *et al.*, 2012; SayuriYamagata *et al.*, 2017; Sinha *et al.*, 2018), o que levanta a importância na identificação dos mecanismos moleculares e bioquímicos do BD, em especial a sua relação com vias e biomoléculas que estejam também relacionadas a alterações no sistema circulatório, hemóstase e em processos inflamatórios.

## **1.2. Espectrometria de massas e suas aplicações em metabolômica**

A espectrometria de massas (MS) é uma técnica analítica que consiste em realizar medidas da relação massa/carga ( $m/z$ ) de íons e suas respectivas abundâncias relativas para determinar estruturas moleculares e também para quantificar compostos. A estrutura do equipamento é composta por uma fonte de ionização, um analisador de massas e um detector, que por sua vez pode estar junto com o próprio analisador de massas, dependendo do caso. A amostra é inicialmente inserida na fonte de ionização para que as moléculas dos analitos sejam ionizadas;

em seguida, os íons são resolvidos pelo analisador de massas em função de suas  $m/z$  e, por fim, estas são medidas por um detector (Han, 2016). Uma representação esquemática de um espectrômetro de massas é mostrada na **Figura 1.2**.



**Figura 1.2.** Representação esquemática de um espectrômetro de massas. Analisadores de massas podem ser simples (somente um analisador de massas) ou híbridos (contendo dois - ou mais - analisadores de massas, com uma célula de colisão entre eles) e, em alguns casos, o detector pode estar localizado junto com o analisador (como é o caso de analisadores baseados em transformada de Fourier, como Orbitrap e ressonância ciclotrônica de íons).

Dentre as diversas fontes de ionização utilizadas, podem ser citadas a ionização por elétrons (EI, do inglês *electron ionization*) (Nier, 1947), a ionização química (CI, do inglês *chemical ionization*) (Munson & Field, 1966), ionização química à pressão atmosférica (APCI, do inglês *atmospheric pressure chemical ionization*) (Horning *et al.*, 1973), a fotoionização à pressão atmosférica (APCI, do inglês *atmospheric pressure photoionization*) (Robb *et al.*, 2000), a ionização por eletrospray (ESI, do inglês *electrospray ionization*) (Fenn *et al.*, 1989), a dessorção/ionização a laser assistida por matriz (MALDI, do inglês *matrix-assisted laser desorption/ionization*) (Han, 2016) e a ionização de dessorção por eletrospray (DESI, do inglês *desorption electrospray ionization*) (Eberlin *et al.*, 2011) como as técnicas mais utilizadas para estudos metabolômicos, sendo as duas últimas muito utilizadas para fins de imageamento. Vale ressaltar que as diferentes fontes de ionização necessitam de diferentes níveis de pressão, variando desde níveis considerados como alto vácuo (como é o caso da EI) até a pressão ambiente, como é o caso das metodologias de ionização à pressão atmosférica, como APPI e APCI.

Dentre os diversos analisadores de massas disponíveis no mercado, pode-se citar como sendo os mais utilizados para análises metabolômicas globais, ou seja,

a análise de todos os analitos possíveis de serem medidos incluindo compostos desconhecidos (Roberts *et al.*, 2012), os analisadores híbridos, como o Orbitrap e o quadrupolo/tempo-de-voo (QTOF, do inglês *quadrupole time-of-flight*), o triplo quadrupolo (QqQ) e o *quadrupolo/armadilha de ions* (QIT, do inglês *quadrupole ion trap*) (Han, 2016). Todos estes analisadores podem realizar experimentos de espectrometria de massas em sequência (MS/MS), que são extremamente úteis para fins de identificação, sendo possível a seleção de íons precursores e sua posterior fragmentação, para que se obtenha o perfil de fragmentação dos analitos, auxiliando na identificação dos compostos.

Em boa parte dos experimentos de metabolômica e proteômica utilizam-se técnicas de separação antes da injeção da amostra no espectrômetro de massas, em especial, a cromatografia líquida de alta eficiência e de ultra alta eficiência (HPLC e UHPLC), a cromatografia gasosa (GC) e, em alguns casos, a eletroforese capilar (CE), permitindo, assim, a separação dos diversos compostos de uma amostra, reduzindo efeitos de supressão iônica e a detecção de isóbaros, melhorando a eficiência da análise. Na metabolômica, em geral, recomenda-se o uso de diferentes fases estacionárias em HPLC e em GC, onde pode-se citar o uso de colunas octadecilsilano (C18) para a análise de metabólitos hidrofóbicos e o uso de colunas de interação hidrofílica (HILIC, do inglês *hydrophilic interaction liquid chromatography*) para metabólitos polares (Engskog *et al.*, 2016), no caso da LC. No que diz respeito à GC, há dois fatores que devem ser levados em consideração, que são a derivatização e a coluna utilizada. A derivatização consiste na modificação química de um composto para alterar sua estrutura química e, por consequência, suas propriedades físico-químicas (Qi *et al.*, 2014), aumentando sua volatilidade, estabilidade térmica e melhorando a sensibilidade e resposta do detector (Beale *et al.*, 2018). A mais utilizada para a análise de metabólitos polares é a derivatização em duas etapas com metoxiamina (MOX) juntamente com um agente sililante (BSTFA, MSTFA, entre outros). Há relatos também de outros métodos de derivatização utilizando metilcloroformiato (MCF) para análises de aminoácidos e de 1,1,1,2,2,3,3-heptafluorobutil cloroformiato (HFBCF) para ácidos orgânicos (Begou *et al.*, 2017). Quanto à coluna a ser utilizada, há uma diversa gama de colunas, como a ZB-XLB e as colunas HP-5MS e DB5-MS (compostas por 5% difenilpolisiloxano e 95% dimetilpolisiloxano) (Begou *et al.*, 2017), muito utilizadas devido à sua compatibilidade

não só com a biblioteca NIST, mas com a biblioteca espectral FiehnLib (Kind *et al.*, 2009), que requer o uso de condições similares às condições de aquisição da biblioteca, incluindo também o uso do mesmo tipo de derivatização (MOX/MSTFA).

Para fins de identificação, na GC-MS a maioria das metodologias em metabolômica utiliza a EI devido à disponibilidade e compatibilidade com diversas bibliotecas espectrais com este método de ionização (Beale *et al.*, 2018). A EI fragmenta as moléculas analisadas de forma reprodutível na própria fonte de ionização, onde os fragmentos obtidos podem ser considerados como uma “impressão digital” de cada molécula, mesmo em equipamentos de diferentes fabricantes, contanto que a energia de ionização se mantenha em 70 eV, padrão utilizado nas diversas bibliotecas espectrais disponíveis na literatura (Beale *et al.*, 2018). Na LC-MS, as metodologias para a análise de metabólitos são baseadas majoritariamente na utilização de ESI. Na ionização por eletrospray, os analitos são carregados eletricamente em um capilar metálico sob o efeito de uma corrente elétrica, podendo assim carregar os íons positivamente e negativamente dependendo da polaridade do capilar. Em seguida, os íons são repelidos do capilar em um formato de cone (chamado de cone de Taylor), produzindo um jato fino de líquido que se fragmenta em gotículas menores carregadas, cujo solvente passa por um processo de evaporação rápida, reduzindo o tamanho das gotículas até o ponto cuja tensão superficial da gotícula não suporta mais as repulsões causadas pelas cargas presentes na própria gotícula, fazendo com que esta se divida em gotículas menores, em um processo conhecido como “fissão de Coulomb”, que acontece repetidamente até o ponto que as gotículas atinjam volumes extremamente pequenos (Raio ~1 nm). Estas gotículas são por sua vez atraídas para o analisador de massas (Banerjee & Mazumdar, 2012). Por se tratar de uma técnica de ionização branda (que não produz extensa fragmentação, gerando majoritariamente íons intactos) (Fenn *et al.*, 1989), é necessária a aquisição de dados de fragmentação por meio de experimentos de MS/MS para a identificação putativa dos analitos nos bancos de dados disponíveis. Atualmente, para que um metabólito seja considerado como identificado inambiguamente em LC-MS, é necessário ainda a comparação do espectro de MS/MS obtido experimentalmente com o espectro de MS/MS de um padrão analisado sob as mesmas condições do metabólito em questão (Domingo-Almenara *et al.*, 2018). Visto que o número e as classes de compostos possíveis de serem ionizados e detectados

são vastos, recomenda-se também que as análises por ESI sejam realizadas em ambos os modos de ionização (positivo e negativo), para que se tenha uma abrangência maior em relação à identificação de biomoléculas (Kim & Heyman, 2018).

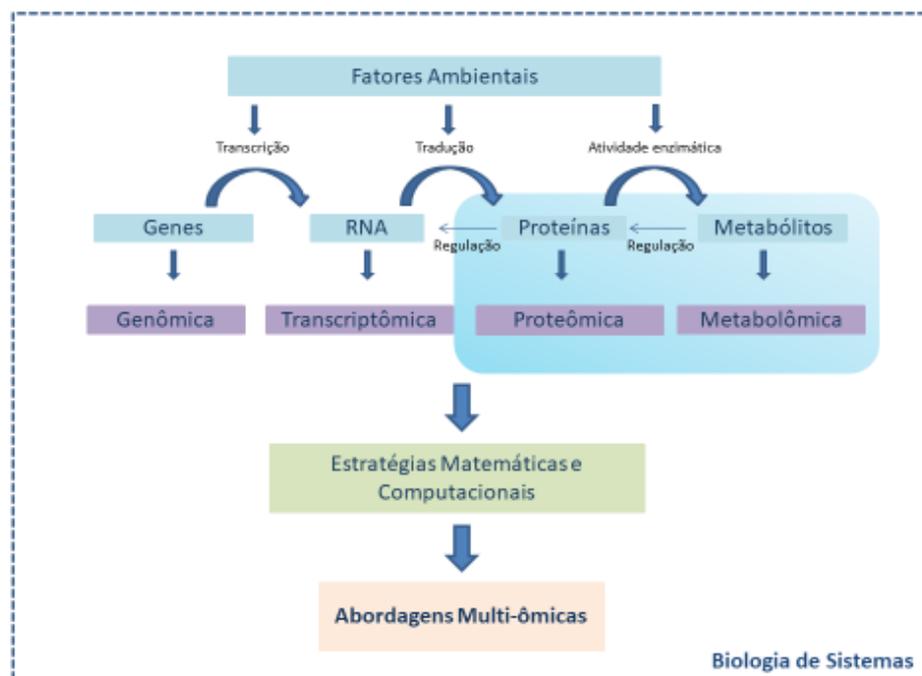
Ambas as técnicas possuem suas particularidades e suas limitações. Apesar da identificação de metabólitos analisados via GC-MS ser mais precisa e simples devido às bibliotecas espectrais, já bem estabelecidas na literatura, a GC-MS é utilizada para a análise de compostos de baixa massa molecular (50-600 Da) e compostos voláteis (ou volatilizáveis via derivatização) (Beale *et al.*, 2018). A LC-MS ainda é a técnica analítica mais utilizada para a análise de compostos não voláteis e compostos que possuem maior massa molecular, como lipídios por exemplo (Alves *et al.*, 2021). Entretanto, para se analisar todo o metaboloma, é necessário um maior número de experimentos em LC-MS para ter uma cobertura completa dos compostos presentes na amostra, além de experimentos adicionais para fins de identificação. Além disso, as bibliotecas espectrais de LC-MS não são tão bem estabelecidas e reproduzíveis como as encontradas em GC-MS, onde cerca de somente 10% dos metabólitos conhecidos nos bancos de dados disponíveis possuem informação espectral experimental (Domingo-Almenara *et al.*, 2018).

Com a procura de marcadores moleculares cada vez mais em voga no cenário científico, ferramentas como a MS são de grande importância para a prospecção e descoberta destes marcadores que, além de trazer novas perspectivas na área diagnóstica, auxiliam também no entendimento das alterações bioquímicas causadas por doenças e condições ambientais, norteando e construindo uma base sólida para ensaios futuros.

### **1.3. Biologia de sistemas e integração de dados ômicos**

Com o advento das abordagens ômicas nos anos 1990, novas ferramentas de pesquisa na área de ciências da natureza foram desenvolvidas e revolucionaram a forma na qual podemos observar e entender um sistema biológico, tal como estabelecer uma correlação entre genótipo e fenótipo (Hyötyläinen & Orešič, 2014; Misra *et al.*, 2019; Santos *et al.*, 2020). Entretanto, dados ômicos individualizados não

são suficientes para explicar como diferentes processos biológicos se correlacionam, devido à sua complexidade (Santos *et al.*, 2020). Assim, a biologia de sistemas é essencial não somente para lidar com a complexidade dos dados ômicos em si, mas também para analisar como diferentes componentes moleculares contribuem como partes de um sistema complexo relacionado a um fenótipo específico ou função biológica (Hyötyläinen & Orešič, 2014). A biologia de sistemas é um campo de pesquisa interdisciplinar que agrega estratégias computacionais e modelagens matemáticas aplicadas a dados biológicos a fim de se obter um maior entendimento dos sistemas biológicos, bem como identificar as relações entre diferentes processos biológicos (Alves *et al.*, 2021; Pinu *et al.*, 2019; Santos *et al.*, 2020). Deste modo, a integração de diferentes ômicas permite uma visão mais detalhada a nível celular, auxiliando na estratificação de indivíduos saudáveis versus doentes (estudos caso-controle), assim como observar a influência de fatores externos que possam interagir causando alterações no sistema biológico (Alves *et al.*, 2021; Hyötyläinen & Orešič, 2014; Zierer *et al.*, 2015). É possível observar as interações e denominações dentre as diferentes ômicas dentro da biologia de sistemas ilustrada na **Figura 1.3**.



**Figura 1.3.** Os diferentes tipos de ciências ômicas e suas interações dentro da biologia de sistemas, também podendo ser conhecida como cascata ômica (adaptado de Santos *et al.*, 2020).

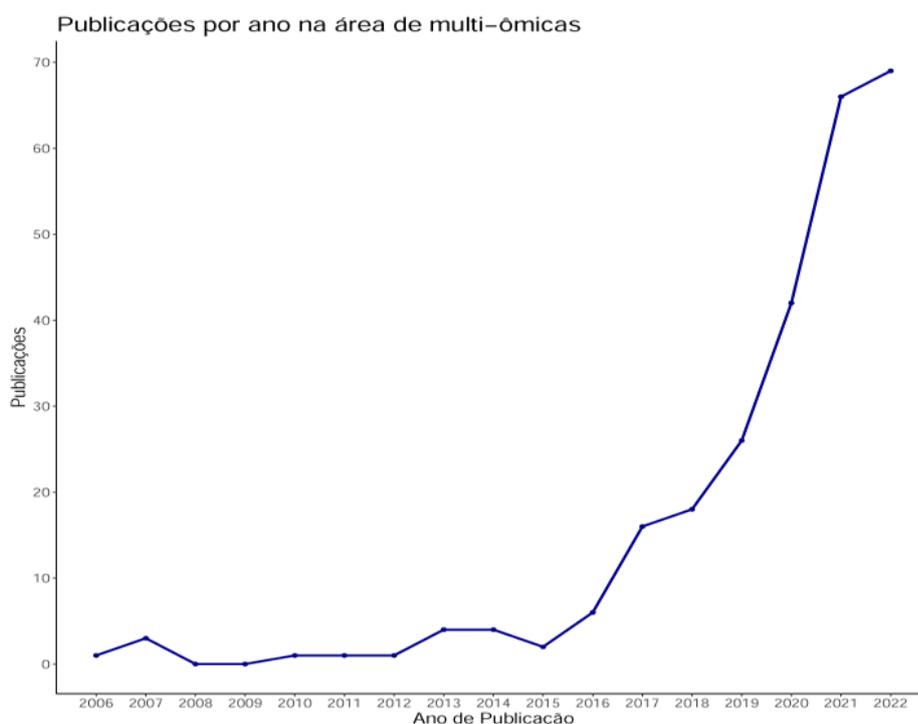
A integração de dados ômicos vem sendo utilizada há mais de uma década (Kim & Tagkopoulos, 2018), sendo que trabalhos de integração de dados de proteômica e metabolômica vêm apresentando relevância, com diferentes abordagens de tratamento de dados, que incluem o uso de programas baseados em R ([www.r-project.org](http://www.r-project.org)), Python ([www.python.org](http://www.python.org)) e MatLab ([www.mathworks.com](http://www.mathworks.com)), que utilizam correlação estatística, aprendizado de máquina, regressão logística e construção de redes para a integração dos dados, gerando diferentes abordagens e resultados possíveis mesmo para conjuntos de dados iguais (Fondi & Liò, 2015; Misra *et al.*, 2019; Wanichthanarak *et al.*, 2015), além de programas comerciais como por exemplo o *Ingenuity Pathway Analysis* (IPA).

Os dados ômicos obtidos por diferentes plataformas podem ser integrados de três formas utilizando a combinação de duas camadas diferentes, dados (por exemplo, concentração de metabólitos, níveis de expressão de proteínas e genes, variantes genéticas) e informação (por exemplo, reações de metabólitos com determinadas enzimas, sítios de fosforilação conhecidos de determinadas proteínas, funções de determinados genes ou *splicings* alternativos de determinados genes em uma determinada posição): dados – dados, dados – informação e informação – informação (Kim & Tagkopoulos, 2018). Dentro destas três formas de integração, deve-se reconhecer que, para diferentes objetivos, distintas metodologias de análise de dados devem ser utilizadas. No tocante ao tratamento dos dados, existem também três diferentes formas de análise: i) integração via *pathway*, que integra os resultados obtidos com bancos de dados já existentes de vias metabólicas e bioquímicas; ii) integração via redes biológicas, onde conexões complexas entre diferentes componentes celulares (genes, proteínas, metabólitos) são utilizadas para integrar e mapear múltiplos dados ômicos, sem depender de vias bioquímicas pré-definidas (quando utilizada sozinha não fornece muitas informações, utiliza-se juntamente com correlações empíricas ou comparações com outras espécies); iii) integração via correlação empírica, que é utilizada quando não se tem o conhecimento bioquímico total do fenômeno estudado, sendo necessário ter algum conhecimento prévio do mesmo para estimar interações entre variáveis (Wanichthanarak *et al.*, 2015).

Além das três formas supracitadas de análise, é possível combinar as abordagens para obter um resultado mais conciso das interações entre as diferentes ômicas. Uma forma natural de complementar métodos baseados em

correlação/covariância é a sua apresentação em forma de rede, que é uma forma interessante de visualizar e entender as interações no sistema biológico, onde não só as contribuições individuais de cada componente da rede são levadas em consideração, mas também as suas interações com os outros componentes e sua função como um todo (Rosato *et al.*, 2018). Em metabolômica, a similaridade entre os metabólitos normalmente é avaliada utilizando métodos de correlação de Pearson ou de Spearman (Rosato *et al.*, 2018); entretanto, quando passamos a utilizar diferentes conjuntos de dados de diferentes ômicas, outras formas de correlação podem ser úteis para integrar dados biológicos e outros meta-dados (informações clínicas e demográficas, por exemplo) (Wanichthanarak *et al.*, 2015). Metodologias como análise de correlação canônica (*regularized canonical correlation analysis – rCCA*) baseadas em análise discriminante por projeção esparsa a estruturas latentes (sPLS-DA, do inglês *Projection to sparse latent structures discriminant analysis*) (Singh *et al.*, 2017), análise de redes de correlação genômica por pesos (WGCNA, do inglês *Weighted gene correlation network analysis*) (Langfelder & Horvath, 2008) ou análise de correlação parcial (Roverato & Castelo, 2017), por exemplo, podem ser utilizadas para correlacionar informações entre dois conjuntos de dados de alta dimensionalidade como dados proteômicos, genômicos ou transcriptômicos. Devido à natureza sistêmica do controle metabólico, normalmente as correlações entre metabólitos são pequenas, mesmo com dois metabólitos sendo vizinhos dentro de uma via metabólica, pois a variância nas enzimas que os controlam pode afetar seus níveis para diferentes direções (Camacho *et al.*, 2005), o que pode levar a correlações indiretas entre metabólitos distantemente correlacionados (Lee, 2006). Uma forma de observar esse fenômeno pode ser feita utilizando análise de correlação parcial (Krusmiek *et al.*, 2011; Roverato & Castelo, 2017), por meio de um modelo gráfico gaussiano que estima a dependência condicional entre variáveis, onde altas correlações parciais geralmente correspondem a vias metabólicas conhecidas. Entretanto, associações em uma rede de correlação podem não corresponder a reações metabólicas conhecidas, visto que estas abordagens não dependem exclusivamente de funções conhecidas de uma molécula, uma vez que diversos genes e proteínas podem apresentar diferentes atividades e funções em diferentes sistemas biológicos, onde moléculas podem apresentar funções desconhecidas até então no sistema analisado (Misra *et al.*, 2019; Rosato *et al.*, 2018).

Dentro da integração de dados entre proteômica e metabolômica, é visível o interesse da comunidade científica neste tipo de abordagem. Nos últimos anos, o número de publicações de abordagens multi-ômicas envolvendo dados de proteômica e metabolômica cresceu vertiginosamente. Observando os resultados da pesquisa na base de dados *Scopus* (utilizando os termos *proteomics AND metabolomics AND multi-omics OR "data integration" AND NOT transcriptomics AND NOT genomics*) é possível observar a clara tendência de crescimento, especialmente a partir do ano de 2019, como ilustrado na **Figura 1.4**.



**Figura 1.4.** Publicações em multi-ômicas envolvendo dados proteômicos e metabolômicos ao longo dos anos na plataforma *Scopus*.

Nos últimos anos, trabalhos de integração de dados ômicos de proteômica e metabolômica em amostras humanas vêm apresentando resultados interessantes. Utilizando o programa IPA, Boccio *et al.* (2016) obtiveram uma rede de correlação entre proteínas e metabólitos em pacientes com esclerose múltipla, sendo possível observar também as diferentes vias metabólicas alteradas devido à doença. Alcazar *et al.* (Alcazar *et al.*, 2021) utilizaram uma abordagem similar com o programa IPA obtendo uma rede de integração entre proteínas, metabólitos, miRNA e lipídios,

adquirindo assim uma rede quadra-ômica para a identificação de assinatura integrada de biomarcadores em pacientes com diabetes tipo I. Por meio de uma abordagem diferente, mais focada à análise correlacional e na criação de modelos de regressão e classificação, Cambiaghi *et al.* (2018) observaram a progressão de choque séptico em pacientes, e a sua relação com lipídios e proteínas da cascata de coagulação. Semelhantemente, Kiebish *et al.* (2020) utilizaram regressão logística para construir um painel integrado de biomarcadores entre proteínas e metabólitos para avaliar a progressão de câncer de próstata em pacientes.

Chen *et al.* (2020) apresentaram resultados interessantes em relação à síndrome metabólica, uma das muitas comorbidades relativas ao BD, em uma abordagem de integração de dados de metabolômica, proteômica e peptidômica, obtendo o perfil molecular de seis diferentes tipos de doenças dentro da síndrome metabólica, tal como o perfil de cada um dos grupos de pacientes analisados, obtendo assim as suas assinaturas biológicas, mostrando diferentes *clusters* para cada um dos diferentes tipos, além de construírem uma rede de integração entre as três diferentes ômicas analisadas. Além disso, por meio da construção de uma curva ROC (do inglês, *receiver operating characteristic*), foi possível observar que a integração de dados possui uma capacidade de predição maior que cada uma das ômicas separadas, mostrando a importância das abordagens integrativas no âmbito diagnóstico/prognóstico. Apesar dos avanços na área de biologia de sistemas e integração ômica, o número de publicações na área da psiquiatria é muito limitado. Dentro dos últimos cinco anos, só houveram dois trabalhos que fizeram uso de estratégias integrativas de proteômica e metabolômica para avaliar transtornos psiquiátricos. Em 2018, Gui *et al.* (2018) utilizaram o programa IPA para integrar dados de proteômica e metabolômica obtidos por meio da análise do plasma sanguíneo de pacientes com transtorno depressivo maior, descobrindo alterações em vias metabólicas relacionada ao metabolismo de fosfolipídios, transporte de lipídios, inflamação e sistema imune. Em seguida, Madrid-Gambin *et al.* (2019) analisaram o plasma de crianças de 12 anos que só desenvolveram sintomas psicóticos aos 18 anos, com o intuito de descobrir marcadores precoces de experiências psicóticas, utilizando análise de correlação canônica regularizada para correlacionar dados de proteômica e metabolômica. A partir dessa abordagem, os autores obtiveram um conjunto de lipídios alterados entre os grupos controle e psicótico e uma rede de

integração mostrando uma forte correlação entre os lipídios associados com experiências psicóticas e plasminogênio, indicando que o lipidoma e o proteoma de tais pacientes que apresentaram sintomas psicóticos aos 18 anos já estava alterado aos 12 anos, o que pode indicar que alterações metabólicas podem estar relacionadas à possível vulnerabilidade a episódios psicóticos.

Apesar do número de trabalhos de multi-ômicas estar em constante crescimento, trabalhos que de fato integram os dados de diferentes ômicas, em especial metabolômica e proteômica, é limitado. Muitos autores denominam seus trabalhos como integrativos, porém, analisam separadamente cada conjunto de dados ômicos, utilizando os resultados de ambos juntos somente na interpretação biológica, o que não é a estratégia mais adequada, visto que a integração de dados consiste na combinação simultânea de dados ômicos individuais a fim de entender a interação das diferentes moléculas, funcionando como uma ponte entre duas ou mais diferentes ômicas, para assim facilitar a ligação entre o genótipo e o fenótipo (Subramanian *et al.*, 2020).

Visto a necessidade de se obter informações relevantes quanto aos mecanismos bioquímicos relativos ao BD que possam revelar potenciais alterações causadas pela doença, optou-se por utilizar uma abordagem de integração de dados proteômicos e metabolômicos obtidos por meio da análise do soro sanguíneo de pacientes com BD em comparação com controles saudáveis neste trabalho, fazendo uso de estratégias bioinformáticas de correlação canônica e regressão logística para a integração de tais dados.

## **2. Objetivos**

### **2.1. Objetivos gerais**

O objetivo deste trabalho consistiu na avaliação das potenciais conexões entre proteínas e metabólitos em amostras de soro sanguíneo de pacientes com transtorno bipolar (BD) em comparação com controles saudáveis (HC) e suas possíveis correlações com comorbidades relacionadas ao BD, bem como a análise de um painel de metabólitos com o potencial de diferenciar pacientes com BD e HC.

### **2.2. Objetivos específicos**

- Determinação e identificação dos metabólitos presentes no soro sanguíneo de pacientes com BD e HC;
- Aplicação de análises estatísticas para integrar dados demográficos e clínicos, dados metabolômicos e dados proteômicos;
- Construção e análise de uma rede de correlação entre proteínas, metabólitos e dados demográficos e clínicos;
- Proposição de um painel de metabólitos com o maior potencial de diferenciação entre os grupos BD e HC.

### 3. Parte Experimental

#### 3.1. Pacientes e amostragem

As amostras de soro sanguíneo provenientes de indivíduos saudáveis (grupo controle) e de pacientes com BD foram adquiridas em parceria com o Prof. Claudio Banzato (Departamento de Psiquiatria, Faculdade de Ciências Médicas, Unicamp). O projeto foi aprovado no Comitê de Ética em Pesquisa da Unicamp (protocolo 775/2010), sendo que todos os pacientes assinaram um termo de consentimento antes da coleta das amostras.

Os participantes da pesquisa não poderiam possuir nenhuma outra doença concomitante, como AIDS, hepatite, doenças endócrinas ou metabólicas, ou fazer uso de substâncias (exceto nicotina), além de não estar em período de gravidez ou pós-parto. Os voluntários do grupo controle foram selecionados dentre a comunidade e, para serem elegíveis, os indivíduos deveriam ter idade entre 18 e 65 anos e não possuir histórico de qualquer condição psiquiátrica (atual ou anterior), bem como nunca ter feito uso de medicamentos psiquiátricos. Além disso, somente voluntários sem histórico familiar de transtornos mentais maiores (transtornos no espectro da esquizofrenia, transtornos de humor e suicídio) foram incluídos, além dos critérios de exclusão utilizados para os pacientes do grupo de pacientes com BD. O sangue foi coletado em um tubo Vacutainer, deixado em temperatura ambiente por 30 minutos e depois colocado em gelo para que se efetuassem a coagulação do sangue. Em seguida, as amostras foram centrifugadas a 3500 g por 15 minutos a 4 °C. O soro sanguíneo obtido foi então transferido para microtubos e guardado a -80 °C para posterior análise.

Este estudo envolveu 14 pacientes com transtorno bipolar (grupo BD) e 14 controles saudáveis (grupo HC) como o conjunto de dados completo, entretanto, para cada parte deste estudo, um número diferente de amostras foi analisado. Para a integração dos dados ômicos, analisou-se 5 pacientes do grupo BD e 8 controles do grupo HC, visando integrar os dados proteômicos e metabolômicos de um mesmo indivíduo. Para a construção do painel de metabólitos utilizando regressão logística

como preditor do BD, utilizou-se o conjunto de dados completo, com 14 pacientes do grupo BD e 14 controles do grupo HC.

Os dados clínicos e demográficos dos indivíduos cujos soros sanguíneos foram analisados empregando a estratégia metabolômica são apresentados na **Tabela 1.1** (conjunto de dados completo) e na **Tabela 1.2** é descrito o conjunto de amostras pareado para a integração de dados metabolômica-proteômica. Observa-se que a média de idade entre os grupos bipolar e controle são parecidos, bem como a porcentagem de mulheres/homens incluídos no estudo e o índice de massa corpórea (IMC) dos mesmos.

**Tabela 1.1.** Dados clínicos e demográficos das amostras utilizadas na análise metabolômica, conjunto de dados completo (SD: desvio padrão; HC: controles saudáveis; BD: pacientes com transtorno bipolar)

Fatores	BD (n = 14)	HC (n = 14)	p-valor
<b>Idade (anos); média (SD)</b>	40 (19)	43 (15)	0,724 <sup>a</sup>
<b>Gênero F/M; n (%)</b>	9/5 (64/36)	10/4 (71/28)	0,686 <sup>b</sup>
<b>IMC; kg m<sup>-2</sup>; média (SD)</b>	26,4 (5,4)	25,0 (1,9)	0,369 <sup>a</sup>
<b>Tipo de medicação; n (%)</b>			
<b>Anticolinérgico</b>	1 (2,9)		
<b>Anticonvulsivo</b>	8 (23,5)		
<b>Estabilizador de humor</b>	12 (35,3)		
<b>Antipsicótico</b>	11 (32,3)		
<b>Antidepressivo</b>	2 (5,9)		

<sup>a</sup> Teste-t de Student, <sup>b</sup> Pearson Chi-Square

**Tabela 1.2.** Dados clínicos e demográficos das amostras pareadas nos estudos de metabolômica e proteômica (SD: desvio padrão; HC: controles saudáveis; BD: pacientes com transtorno bipolar)

Fatores	BD (n = 5)	HC (n = 8)	p-valor
<b>Idade (anos); média (SD)</b>	34 (11)	38 (10)	0,504 <sup>a</sup>
<b>Gênero F/M; n (%)</b>	3/2 (60/40)	5/3 (63/38)	0,928 <sup>b</sup>
<b>IMC; kg m<sup>-2</sup>; média (SD)</b>	25,4 (5,3)	24,7 (1,9)	0,793 <sup>a</sup>
<b>Tipo de medicação; n (%)</b>			
<b>Anticolinérgico</b>	1 (9,1)		
<b>Anticonvulsivo</b>	2 (18,2)		
<b>Estabilizador de humor</b>	4 (36,4)		
<b>Antipsicótico</b>	4 (36,4)		

<sup>a</sup> Teste-*t* de Student, <sup>b</sup> Pearson Chi-Square

### 3.2. Análise metabolômica utilizando cromatografia gasosa acoplada à espectrometria de massas (GC-MS)

Os metabólitos das amostras de soro sanguíneo foram extraídos utilizando um processo simples de precipitação de proteínas, utilizando 50 µL de soro sanguíneo e 150 µL de MeOH. A mistura foi submetida à agitação em vortex e em seguida centrifugada. O sobrenadante da extração foi separado, e em seguida seco em um concentrador centrífugo a vácuo, sendo que as amostras secas foram então derivatizadas utilizando um processo composto por duas etapas, com a adição de uma solução de 10 µL de metoxiamina em piridina (40 mg mL<sup>-1</sup>), que foi mantida a 30 °C durante 90 minutos. Em seguida, adicionou-se 90 µL de N-metil-N-trimetilsililfluoroacetamida (MSTFA) com 1 % (v/v) de trimetilclorosilano (TCMS), mantendo as amostras incubadas a 37 °C durante 30 minutos. Em seguida, 60 µL de uma solução de ácido pentadecanóico em heptano (10 µg mL<sup>-1</sup>) foi adicionada como padrão interno em cada amostra. Para observar possíveis fontes de variação nas

análises das amostras, um terceiro grupo de amostras, chamado de controle de qualidade (QC) foi também analisado. Amostras de controle de qualidade consistem em um *pool* (mistura) de todas as amostras, por meio da alíquotagem de 5  $\mu\text{L}$  de cada uma das amostras. Três brancos também foram preparados e analisados para a remoção de picos cromatográficos não correspondentes às amostras, onde o processo de extração das amostras QC e dos brancos foram realizadas utilizando o mesmo protocolo de preparo das outras amostras.

Após o preparo das amostras, as mesmas foram inseridas em *vials* para a análise via GC-MS, onde utilizou-se um cromatógrafo a gás HP 5890 Series II acoplado a um espectrômetro de massas HP 5970 Series Quadrupole Mass Selective Detector (Agilent, Santa Clara, CA, EUA). Para a análise, utilizou-se uma coluna Agilent HP5-MS 30 (+10) m x 0,25 mm x 0,25  $\mu\text{m}$ , com uma vazão do gás de arraste (He) de 1  $\text{mL min}^{-1}$ , temperatura do injetor de 250  $^{\circ}\text{C}$  e utilizando o modo de injeção *splitless*, injetando 1  $\mu\text{L}$  utilizando uma seringa de 10  $\mu\text{L}$ . A corrida cromatográfica consistiu em uma rampa de aquecimento, com temperatura inicial de 60  $^{\circ}\text{C}$  mantida por 1 minuto, com uma taxa de aquecimento de 10  $^{\circ}\text{C min}^{-1}$ , até alcançar uma temperatura final de 325  $^{\circ}\text{C}$  que foi mantida por mais 10 minutos, até o tempo final da corrida de 38 minutos.

Para facilitar a posterior identificação dos metabólitos, preferiu-se utilizar o protocolo e as condições de análise descritas por Fiehn e colaboradores (Kind *et al.*, 2009), por ser compatível com a utilização da biblioteca espectral FiehnLib que, além da informação espectral, possui também as informações de tempo de retenção, extremamente importantes para aumentar a precisão das posteriores identificações. Os dados adquiridos foram exportados em formato .AIA no *software* GC Chemstation, sendo em seguida transformados em formato .d utilizando o *software* Agilent GC-MS Translator, sendo posteriormente convertidos em formato .abf utilizando o *software* MS-DIAL ABF File Converter. Os arquivos em formato .abf foram então analisados no *software* MS-DIAL (Tsugawa *et al.*, 2015) para a identificação dos picos cromatográficos, deconvolução e identificação. Os parâmetros utilizados na etapa de pré-processamento são apresentados na **Tabela 1.3**.

**Tabela 1.3.** Parâmetros utilizados no pré-processamento dos dados no software MS-DIAL

<b>Parâmetro</b>	<b>Valor</b>
<b>Parâmetros de coleta de dados</b>	
<i>Retention time begin</i>	0
<i>Retention time end</i>	100
<i>Mass range begin</i>	0
<i>Mass range end</i>	1000
<b>Parâmetros de detecção de picos</b>	
<i>Smoothing method</i>	LinearWeightedMovingAverage
<i>Smoothing level</i>	3
<i>Average peak width</i>	20
<i>Minimum peak height</i>	1000
<i>Mass slice width</i>	0,5
<i>Mass accuracy</i>	0,5
<b>Parâmetros de Deconvolução</b>	
<i>Sigma window value</i>	0,5
<i>Amplitude cut off</i>	10
<b>Identificação</b>	
<i>Retention time tolerance</i>	100
<i>Retention index tolerance</i>	100
<i>EI similarity library tolerance</i>	70
<i>Identification score cut off</i>	70
<i>Use retention information for scoring</i>	True
<i>Use retention information for filtering</i>	False
<i>Use quant masses defined in MSP format file</i>	False
<b>Parâmetros de alinhamento</b>	
<i>Retention index tolerance</i>	20
<i>Retention time tolerance</i>	0,075
<i>EI similarity tolerance</i>	70
<i>Retention time factor</i>	0,5
<i>EI similarity factor</i>	0,5
<i>Identification after alignment</i>	False
<i>Gap filling by compulsion</i>	True
<i>Basepeak mz selected as the representative quant mass</i>	False
<b>Ajustes de filtragem</b>	
<i>Peak count filter</i>	0
<i>Remove feature based on peak height fold-change</i>	True
<i>Sample max / blank average</i>	5
<i>Sample average / blank average</i>	5
<i>Keep identified and annotated metabolites</i>	True
<i>Keep removable features and assign the tag for checking</i>	True

A tabela de dados contendo as identificações feitas através da biblioteca FiehnLib (Kind *et al.*, 2009) e as suas respectivas intensidades no cromatograma de íons totais (TIC) foram analisadas na plataforma estatística R (Versão 3.6.0) (*R Core Team*, 2018) para a avaliação dos desvios padrões relativos (RSD) dos metabólitos nos controles de qualidade (QCs). Caso o metabólito apresentasse um valor de RSD

acima de 30 %, este era excluído de posteriores análises estatísticas, por ser considerado inconsistente ao longo das corridas cromatográficas.

Optou-se por fazer uma segunda curadoria dos metabólitos identificados e considerados consistentes ( $RSD < 30\%$ ) usando o banco de dados Golm Metabolome Database (GMD) (Hummel *et al.*, 2010), onde devido a possíveis oscilações no equipamento durante as análises, foi observada uma variação de cerca de  $\pm 0,2$  minutos em relação ao observado na biblioteca. Para a anotação definitiva do metabólito, utilizou-se um nível de similaridade  $\geq 70\%$  na biblioteca FiehnLib e uma confirmação do mesmo metabólito na biblioteca GMD. Os metabólitos que apresentaram inconsistências entre as identificações nas duas plataformas foram considerados como desconhecidos (*Unknown*), sendo que alguns destes foram identificados somente com os seus grupos funcionais, por meio de umas das funcionalidades do banco de dados GMD.

### **3.3. Proteômica**

Mais informações relativas à descrição experimental e tratamento de dados das análises proteômicas está presente no artigo publicado por Santa Cruz *et al.* (2021), não sendo objeto de estudo dentro deste trabalho. As amostras de soro sanguíneo foram depletadas utilizando um sistema Äkta Start FPLC (GE Healthcare, Alemanha), com uma coluna HiTrap de afinidade com albumina e IgG. Os peptídeos provenientes da digestão de proteínas com tripsina foram analisados utilizando um sistema de UHPLC Ulimite3000 acoplado a um espectrômetro de massas LTQ-Orbitrap Elite (Thermo Fisher Scientific, Bremen, Alemanha). Utilizou-se o software MaxQuant com o banco de dados *Human Uniprot* para a análise de dados e identificação das proteínas. A lista de proteínas proveniente dos pacientes que foram analisados conjuntamente nas análises metabolômicas foi utilizada neste estudo.

### **3.4. Identificação, análise estatística e integração de dados ômicos**

Para realizar a integração dos dados metabolômicos com os dados de proteômica obtidos previamente por nosso grupo (Santa Cruz *et al.*, 2021), a análise

estatística foi feita separadamente para cada um dos dados ômicos, a fim de determinar os metabólitos e proteínas diferenciais entre o grupo de pacientes com transtorno bipolar (BD) e controles saudáveis (HC).

Os dados de proteômica e metabolômica foram submetidos à transformação logarítmica ( $\log_2$ ), e os dados de metabolômica foram normalizados utilizando as intensidades do padrão interno para cada amostra. A homogeneidade das amostras foi analisada utilizando a análise de componentes principais (PCA). Utilizou-se a linguagem estatística R (versão 3.60) para a análise dos dados. A PCA foi realizada utilizando a função *'prcomp'* incluída no pacote *'stats'*. Os pacotes *'boxplot'*, *'beanplot'*, *'gplot'*, *'ggplot2'* e *'igraph'* foram utilizados para visualização dos dados.

O efeito dos diferentes fatores clínicos e demográficos como idade, gênero, IMC, status da doença, uso de antipsicóticos e estabilizadores de humor foram avaliados nos *datasets* de proteômica e metabolômica individualmente. Ambos os conjuntos de dados foram centrados na média e autoescalados, e em seguida a contribuição relativa de cada fator (caso-controle, IMC, idade, gênero e uso de medicação) para a variância total no conjunto de dados foi estimada por meio de um modelo de regressão linear onde as intensidades normalizadas dos metabólitos e proteínas foram sujeitas à regressão em relação ao fator de interesse, obtendo assim valores de  $R^2$  para cada variável, que são ranqueadas em seguida em relação a seu valor de mediana de  $R^2$ , mostrando quais variáveis explicam a maior parte da variância dos dados (McCarthy *et al.*, 2017). Para este procedimento, utilizou-se os pacotes *'SingleCellExperiment'* (Amezquita *et al.*, 2020) e *'scater'* (McCarthy *et al.*, 2017).

Os dados de proteômica e metabolômica foram analisados utilizando Análise de Correlação Canônica Generalizada com o pacote *'mixOmics'* (Singh *et al.*, 2017). O modelo foi construído utilizando a função *'block.splsda'*, onde os principais contribuidores (*loadings* utilizados para maximizar a covariância entre X e Y) foram exportados, a fim de proceder com a análise de correlação parcial. Todos os metabólitos e proteínas que apresentaram um coeficiente de contribuição maior que  $|0,15|$  foram selecionados para a análise de correlação parcial, bem como os dados clínicos e demográficos apresentados na **Tabela 1.2**.

A correlação de Spearman foi feita utilizando a função *'rcorr'* do pacote *'Hmisc'*. O pacote *'qgraph'* foi utilizado para a correlação parcial de primeira ordem. A função *'qpAvgNrr'* foi utilizada para avaliar a taxa média de não-rejeição, onde um limiar de 0,4 foi utilizado para filtrar possíveis correlações espúrias entre diferentes variáveis na rede de correlação. A rede de correlação parcial foi construída e visualizada utilizando o pacote *'igraph'* (Csárdi & Nepusz, 2006).

Para identificar as diferenças nos níveis de proteínas e metabólitos nos pacientes, foi realizado um teste *t*. As intensidades diferenciais dos metabólitos e proteínas ( $p$ -valor < 0,05) entre os grupos BD e HC foram ilustrados na forma de *beanplots* (Kampstra, 2008), sendo que um *script* próprio foi utilizado para observar a dispersão das intensidades entre as amostras. Os metabólitos significativos foram então selecionados para modelagem via regressão logística iterativa.

Modelos de regressão logística *ridge* foram desenvolvidos para realizar a estratificação entre os grupos BD e HC. Os diferentes metabólitos medidos neste estudo foram utilizados tanto individualmente quanto em uma combinação para a modelagem. Um esquema recursivo de eliminação foi implementado para a seleção dos metabólitos, onde estes foram incorporados ou removidos de maneira iterativa. Os modelos foram ajustados para os dados demográficos (IMC, gênero e idade) e sua precisão foi determinada através da área sob a curva (AUC, do inglês *area under curve*), que foi estimada através de *bootstrapping* (1000 vezes). O modelo com a maior AUC foi considerado como o melhor modelo, sendo então construída uma curva ROC utilizando o pacote *'pROC'*. Os modelos de regressão *ridge* utilizando a função *'cv.glmnet'* necessitam de um hiper-parâmetro ' $\lambda$ ', onde o ' $\lambda_{\min}$ ' corresponde ao erro mínimo da validação cruzada (10-*fold* CV).

Os scripts utilizados neste capítulo para as análises estatísticas, utilizando linguagem de programação R, bem como o fluxo de trabalho utilizado estão disponíveis publicamente na página: [https://github.com/Hcarachoribeiro/Metabolomics\\_08\\_22](https://github.com/Hcarachoribeiro/Metabolomics_08_22).

## 4. Resultados e discussão

### 4.1. Anotação dos metabólitos, análise estatística e seleção de variáveis

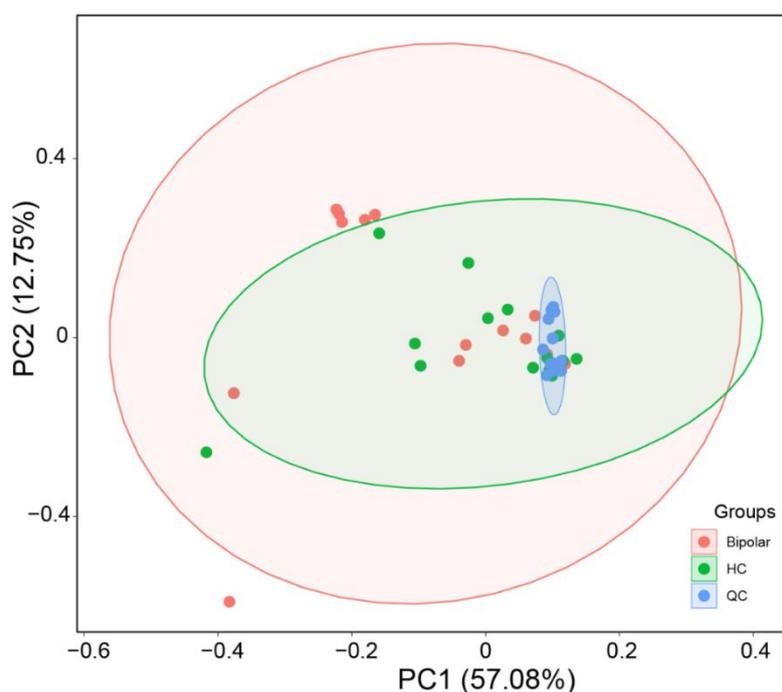
Antes de proceder com a análise, verificou-se a distribuição das amostras frente às variáveis no conjunto completo dos dados metabolômicos, antes e após os processos de normalização, transformação logarítmica e escalamento, na forma de um gráfico *boxplot* (Figura 1.5).



**Figura 1.5.** Gráfico de *boxplot* dos dados de metabolômica para cada amostra antes e após normalização, transformação logarítmica e escalamento.

Após o pré-processamento e a verificação da distribuição dos dados de metabolômica, procedeu-se com a etapa de anotação dos metabólitos presentes nas amostras. Inicialmente identificou-se 527 *molecular features* e, após a anotação utilizando a biblioteca FiehnLib e a biblioteca GMD, 69 metabólitos foram selecionados para a posterior análise estatística. A tabela contendo os tempos de retenção observados, a identificação inicial na biblioteca FiehnLib e a identificação na biblioteca GMD está presente nos **Anexos**. A tabela de dados total contendo os metabólitos anotados e seus níveis, bem como os dados brutos utilizados neste capítulo, estão presentes na plataforma MetabolomicsWorkbench (Sud *et al.*, 2016), sob o número de acesso ST001946.

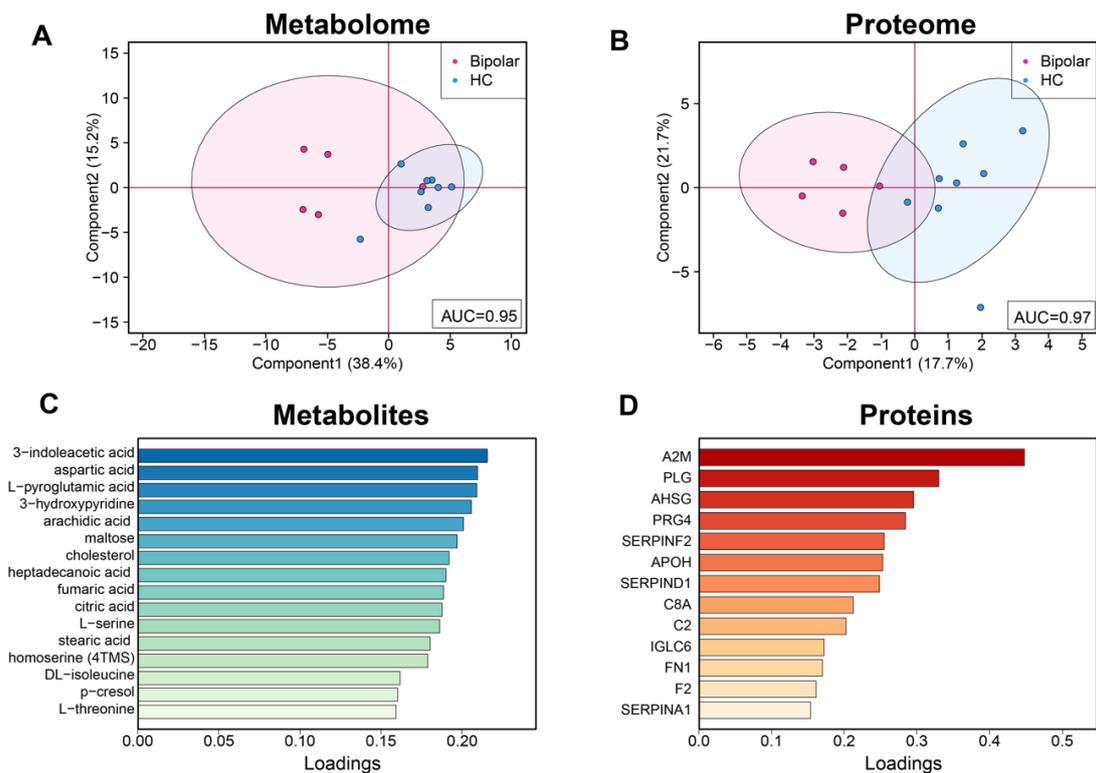
Os metabólitos obtidos das amostras dos grupos BD e HC foram avaliados usando análise multivariada. O gráfico de escores da análise de componentes principais apresentou baixa variância das amostras de QC (**Figura 1.6**), demonstrando a qualidade dos dados, visto que a baixa variância dos QCs indica que possíveis variações instrumentais não foram significantes estatisticamente.



**Figura 1.6.** Gráfico de escores da análise de componentes principais (PCA) dos dados metabolômicos obtidos por GC-MS.

Os gráficos de escores da análise discriminante por mínimos quadrados parciais (PLS-DA) dos dados de metabolômica apresentam uma tendência de separação entre os grupos de amostras nas duas componentes (**Figura 1.7A**). O teste

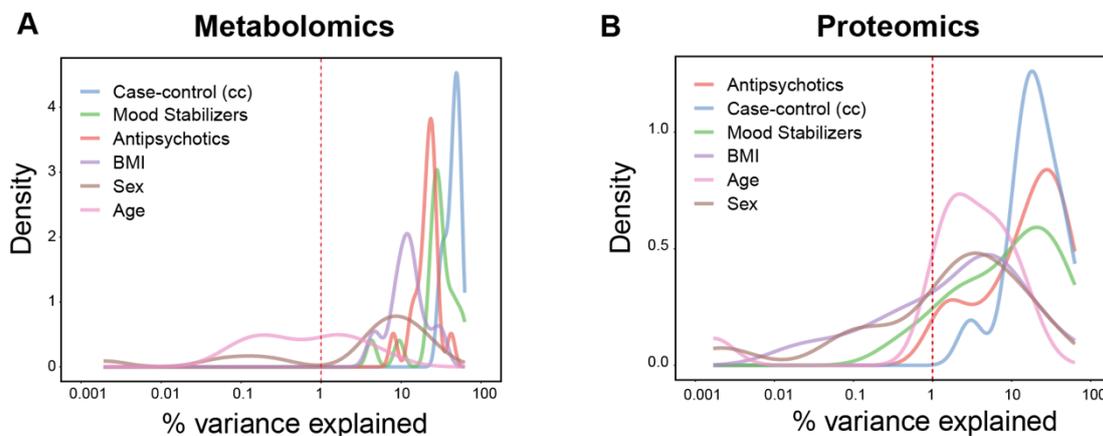
ROC multivariado mostrou uma AUC de 0,95, confirmando a validade do modelo multivariado. Os testes ROC multivariados podem ser utilizados como parâmetro de validação cruzada para modelos PLS-DA e apresentam melhor performance do que a utilização de parâmetros normalmente utilizados como  $R^2$  e  $Q^2$  para determinar pequenas diferenças entre grupos. Os valores de AUC vão de 0,5 (onde considera-se que o modelo proposto não possui capacidade discriminatória) até 1 (valor que expressa a capacidade do modelo de discriminar perfeitamente entre os grupos) (Szymańska *et al.*, 2012). O gráfico de escores dos dados de proteômica apresentou separação consistente na primeira componente entre os dois grupos (**Figura 1.7B**), com um valor de AUC de 0,97 nos testes de performance do modelo. Em seguida, utilizando a função 'circosPlot' do pacote 'mixOmics', foi realizada uma análise de correlação, onde os principais contribuidores para as análises de correlação foram selecionados utilizando um limiar de  $|0,15|$  (**Figuras 1.7C e 1.7D**).



**Figura 1.7.** Gráfico de escores da PLS-DA multi-bloco dos dados metabolômicos (A), proteômicos (B), e os metabólitos e proteínas mais significativos (C e D), que serão utilizados na análise de correlação parcial.

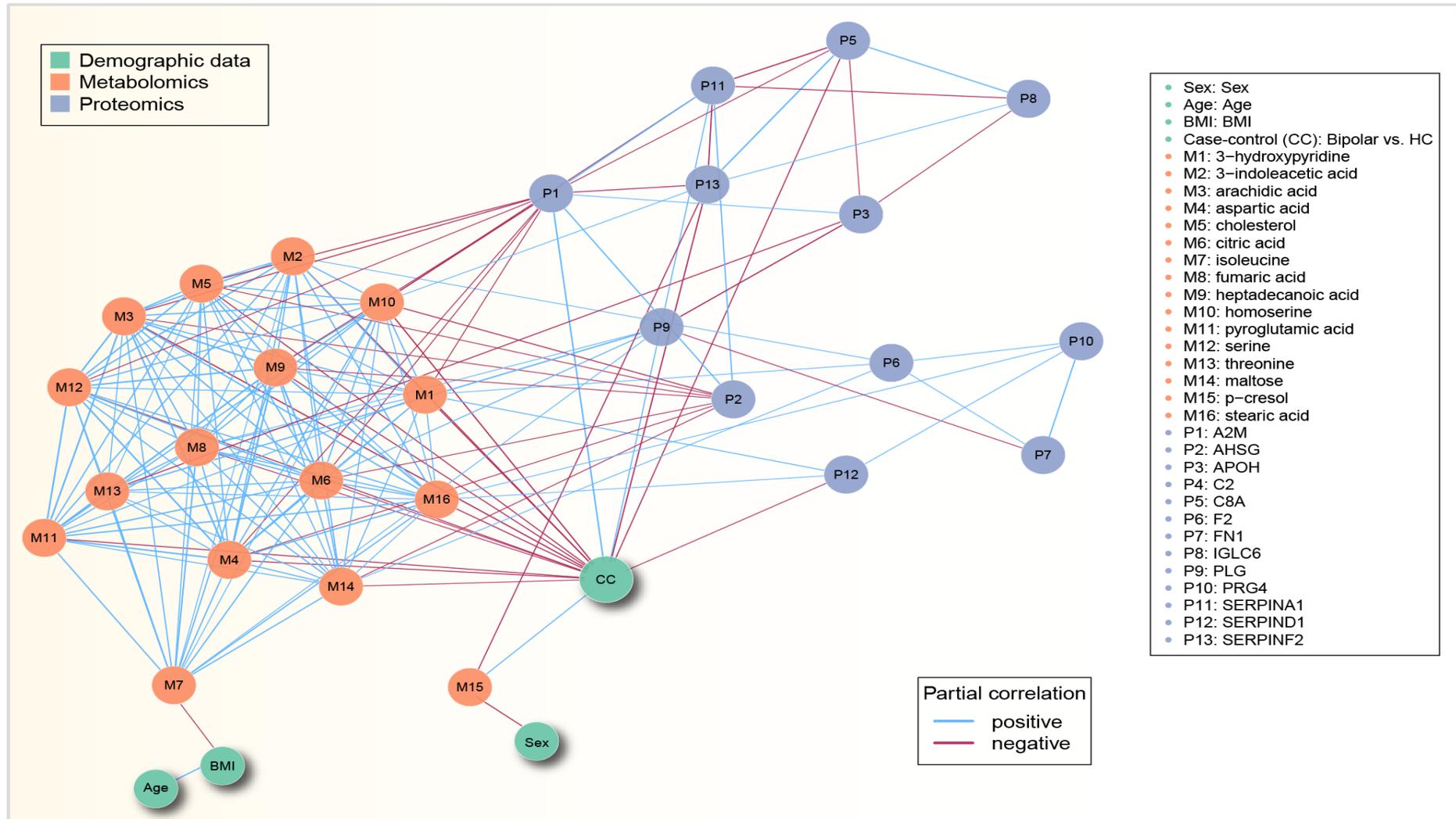
## 4.2. Interações entre proteínas e metabólitos no BD utilizando correlação parcial em formato de rede

Para observar as interações entre as proteínas e os metabólitos no BD, utilizou-se os principais contribuidores de cada conjunto de dados e os dados demográficos apresentados na **Tabela 1.2** para construir uma rede de correlação parcial. Como esperado, a diferenciação Caso-Controle (CC) foi a variável que explicou a maior parte da variância dos dados, tanto na metabolômica, quanto na proteômica (**Figura 1.8**).



**Figura 1.8.** Principais fatores de confusão nos conjuntos de dados de metabolômica e proteômica, demonstrando a ordem de influência na variância dos dados (Caso-Controle (CC) > tipo de medicação > IMC > gênero > idade).

Na rede de correlação parcial (**Figura 1.9**), é visível a importância do nodo CC como o *hub* principal da rede, com 13 metabólitos e 5 proteínas ligados a ele. Os dados mostraram uma intensa tendência de agrupamento dentro de cada conjunto de dados, porém, as interações abordadas serão as que ocorrem entre os dois conjuntos de dados.



**Figura 1.9.** Rede de correlação parcial ilustrando as interações entre metabólitos (M), proteínas (P) e os dados demográficos.

Somente os metabólitos e proteínas que mostraram ligação direta com o nodo Caso-Controle (CC) foram selecionados para discussão, somente se a proteína e o metabólito apresentassem ambos a ligação com o nodo CC.

Os processos biológicos das proteínas foram determinados utilizando a plataforma Reactome (Jassal *et al.*, 2020), usando o modo *Analysis*. Dentre as proteínas que apresentaram ligação com o nodo CC (A2M, SERPINF2, SERPIND1, SERPINA1 e C8A)\*<sup>1</sup>, somente a C8A não estava relacionada com processos de hemóstase. Entretanto, a C8A está relacionada à cascata de coagulação devido à sua participação na cascata complemento (Ogata *et al.*, 1999), o que pode explicar sua associação com processos de hemóstase. Como a C8A não apresentou ligação com nenhum metabólito na rede de correlação, foi retirada de posterior análise. Dez metabólitos apresentaram as interações requeridas (3-hidroxipiridina, ácido araquídico, ácido aspártico, colesterol, ácido cítrico, ácido heptadecanóico, homoserina, serina, *p*-cresol e ácido esteárico). Estes metabólitos serão discutidos a seguir, a respeito de suas interações com as proteínas e suas possíveis relações com o BD.

Dentre as comorbidades comumente reportadas em pacientes com BD, doenças cardiovasculares (DCVs) estão entre as mais discutidas na literatura. Em comparação com indivíduos saudáveis, pacientes com BD apresentam um risco maior de apresentarem síndrome metabólica (SMet) e desenvolvimento de DCV, ambas relacionadas com alterações na hemóstase e em processos inflamatórios (Goldstein *et al.*, 2009; Hoirisch-Clapauch *et al.*, 2014; Leboyer *et al.*, 2012; SayuriYamagata *et al.*, 2017). Como não há uma visão clara de diversos aspectos da fisiopatologia do BD, a busca por um melhor entendimento das alterações bioquímicas causadas pela doença e seu tratamento é um objeto de estudo relevante. Logo, novas descobertas são cruciais para confirmar ou corroborar hipóteses levantadas anteriormente. Uma destas hipóteses correlaciona o BD como uma doença multissistêmica, causando alterações em processos inflamatórios e alterações cardiovasculares (Goldstein *et al.*, 2009; Hoirisch-Clapauch *et al.*, 2014; Leboyer *et al.*, 2012).

---

\*<sup>1</sup> A2M – alfa-2-macroglobulina; SERPINF2 - alfa-2-antiplasmina; SERPIND1 - cofator heparina II  
SERPINA1 - alfa-1-antitripsina; C8A – componente complemento C8 cadeia alfa

Apesar de ser possível observar que o efeito da doença é o responsável pela maior parte da variância dos dados, como mostrado na **Figura 1.8**, ainda é possível observar que as medicações utilizadas pelos pacientes contribuem para a variância entre BD e HC. Dado a vasta gama de medicações que podem ser utilizadas no tratamento do BD, observou-se apenas os efeitos das medicações utilizadas pelos pacientes neste estudo, em especial, estabilizadores de humor e antipsicóticos. Por exemplo, observa-se na literatura que o tratamento com risperidona levou a uma redução na absorção da glicose em adipócitos e promoveu a lipogênese em modelos animais (possível aumento da obesidade) (Vestri *et al.*, 2007), alterações em metabólitos relacionados ao metabolismo energético, metabolismo de neurotransmissores, metabolismo de ácidos graxos e fosfolipídios, além de níveis aumentados de glicose após o tratamento (Xuan *et al.*, 2011), o que pode levar a um maior risco de desenvolver SMet e outras DCVs. No trabalho de Paredes *et al.* (Paredes *et al.*, 2014) é possível observar que os níveis de insulina em pacientes esquizofrênicos tratados com medicações consideradas de alto risco para o desenvolvimento de SMet é significativamente maior que os de médio e baixo risco, além disso, é possível observar alterações em di e triacilgliceróis em pacientes sob tratamento com medicações de médio risco (quetiapina e risperidona), onde é sugerido uma possível alteração na produção de energia, desfavorecendo a quebra dos lipídios para o ciclo do ácido cítrico, favorecendo a lipogênese, reduzindo metabólitos relacionados à produção de energia e aumentando o nível de triacilgliceróis depositados nos tecidos, aumentando assim a gordura corporal e conseqüentemente o ganho de peso, bem como metabólitos de degradação do colesterol. Lenski *et al.* (2021) em seu trabalho analisaram o efeito de drogas psicotrópicas em pacientes psiquiátricos após 1 mês e 3 meses de uso, e observaram alterações em diversas vias metabólicas, além de alterações mitocondriais e disfunção na beta-oxidação de ácidos graxos, tal como a disfunção de enzimas responsáveis pela síntese de carnitinas, responsáveis pelo transporte de ácidos graxos na mitocôndria, levando a uma redução da lipólise, promovendo o aumento do estoque de ácidos graxos nos adipócitos e hepatócitos, e conseqüentemente, ao aumento de peso dos pacientes. Em um estudo de meta-análise de pacientes com esquizofrenia e outros transtornos psiquiátricos relacionados, Mitchell *et al.* (2013) observaram que pacientes após

tratamento com antipsicóticos apresentam risco aumentado de desenvolver SMet, apesar destes possuírem já um risco aumentado devido à sua condição psiquiátrica, o que levanta a questão da reavaliação das metodologias de tratamento de pacientes acometidos por alguma condição psiquiátrica levando em conta este risco de desenvolvimento de SMet, que pode no futuro levar à DCVs mais graves. Apesar dos estudos previamente citados indicarem alterações relacionadas ao metabolismo de glicose, lipídios e conseqüentemente sua relação direta com riscos do desenvolvimento de diabetes devido à medicação, existem resultados conflitantes na literatura, que apontam o contrário. Newcomer (2005) observou em seu estudo de revisão que o tratamento com risperidona e quetiapina não estão relacionados a efeitos de dislipidemia ou possíveis riscos de aumento do risco de desenvolver diabetes, tal como alterações nos níveis de glicose e de lipídios, exceto quando misturados com outras medicações que produzem esse efeito colateral. Ward *et al.* também observaram resultados semelhantes em seu estudo, concluindo que o tratamento com antipsicóticos podem não ser o fator primário que indica alterações nos níveis de insulina em pacientes esquizofrênicos, apesar de poder se observar o aumento em alguns ácidos graxos (Ward *et al.*, 2018).

A partir da análise da rede de integração ômica, a proposta foi analisar as interações entre os metabólitos, proteínas e dados demográficos a nível de vias, em um formato *upstream* (de metabólitos para proteínas). Dentre os metabólitos selecionados, que apresentaram uma ligação direta com o nodo CC, o *p*-cresol foi o único que apresentou uma correlação positiva, o que indica níveis aumentados deste metabólito nos pacientes com BD em relação aos controles. O *p*-cresol apresenta uma ligação com a proteínas SERPINF2 (alfa-2-antiplasmina), proteína relacionada à hemóstase e à cascata complemento, e que possui efeito supressor da agregação de plaquetas através da inibição de espécies reativas de oxigênio, assim inibindo a ativação de ERK1/2, TXA<sub>2</sub> e fosforilação p38 (Chang *et al.*, 2011). O *p*-cresol também foi reportado na literatura por causar disfunção da função endotelial, de reparos de ferimentos e resposta a citocinas (Cerini *et al.*, 2004; Chang *et al.*, 2014). Este metabólito é produzido no corpo humano como um produto final da quebra de proteínas por bactérias intestinais e está intimamente relacionado à microbiota intestinal. Há

diversas evidências na literatura de disfunções na microbiota intestinal de pacientes com BD, correlacionando estas alterações com alterações na ativação da resposta imune, inflamação e função cerebral (Hu *et al.*, 2019; Lu *et al.*, 2019), apesar dos mecanismos dessa correlação não estarem claramente elucidados. O ácido cítrico apresentou uma correlação negativa com o nodo CC e à A2M (alfa-2-macroglobulina), que também está relacionada à hemóstase. Scaravilli *et al.* (2018) estudaram os efeitos do ácido cítrico na coagulação, indicando que o citrato possui atividade anticoagulante, apresentando um efeito acidificante no sangue semelhante ao ácido láctico. Além disso, alterações no ciclo do ácido cítrico podem afetar a produção de plaquetas, levando a alterações na hemóstase (Kulkarni *et al.*, 2019).

Três ácidos graxos também apresentaram correlação direta com o diferenciador caso-controle e com ao menos uma proteína ligada a este nodo. O ácido araquídico está ligado à proteína A2M, enquanto o ácido esteárico se liga à SERPIND1 (cofator heparina II), e o ácido heptadecanóico à SERPINA1 (alfa-1-antitripsina). O ácido araquídico foi reportado na literatura por ter uma correlação significativa com o dímero-D, um marcador de coagulação sanguínea (Scarabin *et al.*, 2001), que pode ser utilizado como marcador de trombogênese (Lip & Lowe, 1995). Quanto ao ácido esteárico, sua relação com hemóstase e coagulação é deveras conflitante. Lange *et al.* (2019) indicaram que um aumento nos níveis de fibrinogênio está correlacionado com os níveis de ácido esteárico. Entretanto, os autores pontuaram que essa relação não deve ser levada completamente em consideração para qualquer tipo de conclusão referente à relação entre níveis de ácidos graxos e alterações na coagulação sanguínea, visto que outros estudos são conflitantes. Outros trabalhos reportam que não há uma relação clara entre os níveis de ácido esteárico e alterações na hemóstase e coagulação (Gebauer *et al.*, 2014; Thijssen *et al.*, 2005; Tholstrup, 2005), sendo necessários mais estudos para concluir algo sobre esta correlação. O ácido heptadecanóico, por sua vez, apresentou uma associação inversa com leptina e o inibidor-1 do ativador de plasminogênio, que vem sendo associado a DCVs (Kurotani *et al.*, 2017). Além disso, níveis aumentados de ácido heptadecanóico se correlacionam com a redução nos riscos de diversas

condições clínicas, como diabetes tipo 2, DCVs, SMet e processos inflamatórios (Jenkins *et al.*, 2015; Pfeuffer & Jaudszus, 2016; Venn-Watson *et al.*, 2020).

A proteína A2M também apresentou ligação com o colesterol na rede de integração, onde este é conhecido por influenciar a ativação de plaquetas (Wang & Tall, 2016), assim como possuir uma relação direta com a lipoproteína de alta densidade (HDL do inglês *high density lipoprotein*), que possui efeitos antitrombóticos (van der Stoep *et al.*, 2014; Wang & Tall, 2016). O colesterol apresenta uma correlação negativa com o nodo CC, o que significa que ele possui níveis reduzidos nos pacientes com BD em comparação com controles. Isso pode ser explicado pelo fato de os pacientes deste estudo estarem todos em estado eutímico após o tratamento. Pode-se supor que devido aos níveis reduzidos de colesterol, os níveis de HDL e lipoproteína de baixa densidade (LDL, do inglês *low density lipoprotein*) nos pacientes também será menor devido a menor disponibilidade de colesterol no organismo. Entretanto, uma relação clara entre tais interações não pode ser concluída sem um estudo aprofundado dos perfis lipídicos dos pacientes.

Três aminoácidos apresentaram as interações indicadas como requisitos neste estudo. Ácido aspártico (ligado à A2M), serina (ligado à A2M) e homoserina (ligado à A2M e SERPINF2). Os metabólitos D-ácido aspártico e D-serina são estudados devido às suas ligações com receptores de NMDA (N-metil-D-aspartato) e sua relação com transtornos mentais como BD e esquizofrenia (Errico *et al.*, 2018; Gardoni *et al.*, 2021; Ghasemi *et al.*, 2014; Hashimoto, 2016; Yoshimi *et al.*, 2016). Apesar dos aminoácidos mais predominantes na natureza e em processos metabólicos serem os L-aminoácidos (Errico *et al.*, 2018; Fontanarosa *et al.*, 2017), os D-aminoácidos também são encontrados em mamíferos. A D-serina por exemplo, é sintetizada no cérebro a partir da L-serina através da enzima serina racemase (Ghasemi *et al.*, 2014; Hashimoto, 2016; Wolosker *et al.*, 1999). Errico *et al.* (2018) sugerem em seu estudo que o D-aspartato pode ser produzido no cérebro com um mecanismo similar ao qual a D-serina é produzida. Além disso, os autores sugerem, baseados em estudos anteriores, que a serina racemase pode contribuir parcialmente para esta conversão (Ito *et al.*, 2016) e que esta enzima pode ter servido como um precursor evolutivo para a aspartato racemase

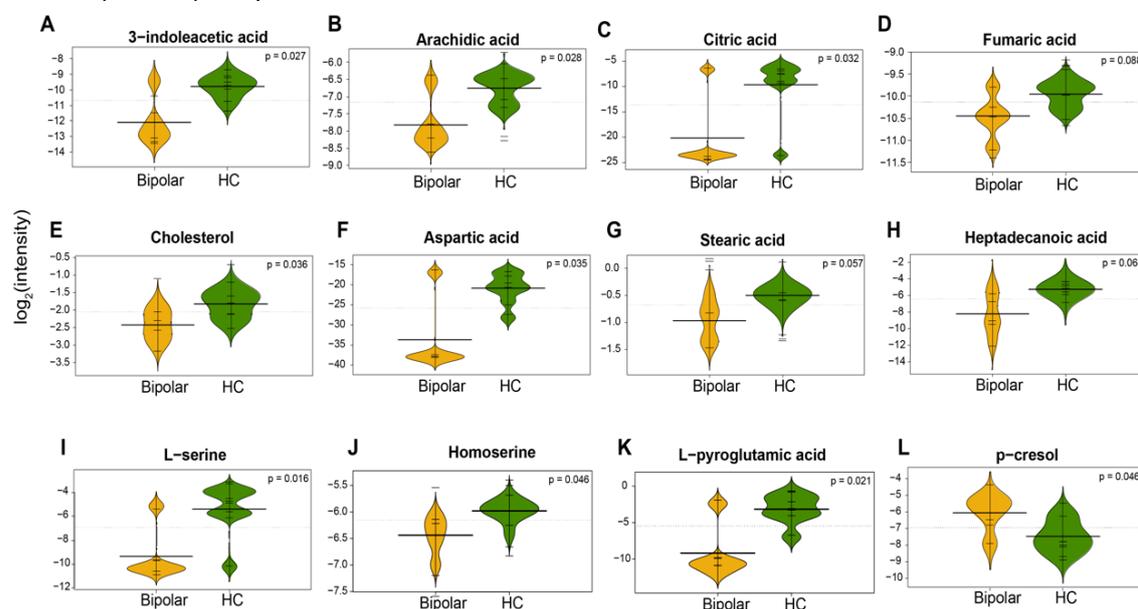
(Graham *et al.*, 2019). O NMDA pode induzir neuroinflamação e morte neuronal com um estímulo anormalmente intenso (Errico *et al.*, 2018), sendo que o aumento na neuroinflamação foi previamente reportado em pacientes com BD (Rao *et al.*, 2010). A homoserina pode participar dos mesmos processos que o ácido aspártico, visto que ela pode ser convertida em aspartato através do intermediário aspartato semialdeído (Liu *et al.*, 2014), participando indiretamente desta interação. Entretanto, observou-se em um estudo de revisão envolvendo modelos animais tratados com diversos estabilizadores de humor que o lítio, a carbamazepina e a lamotringina afetam o metabolismo de ácido araquidônico e os receptores de NMDA na célula (Rapoport, 2014), o que leva a crer que efeitos da medicação também podem contribuir para essa alteração observada.

#### 4.3. Painel de metabólitos como preditor do BD

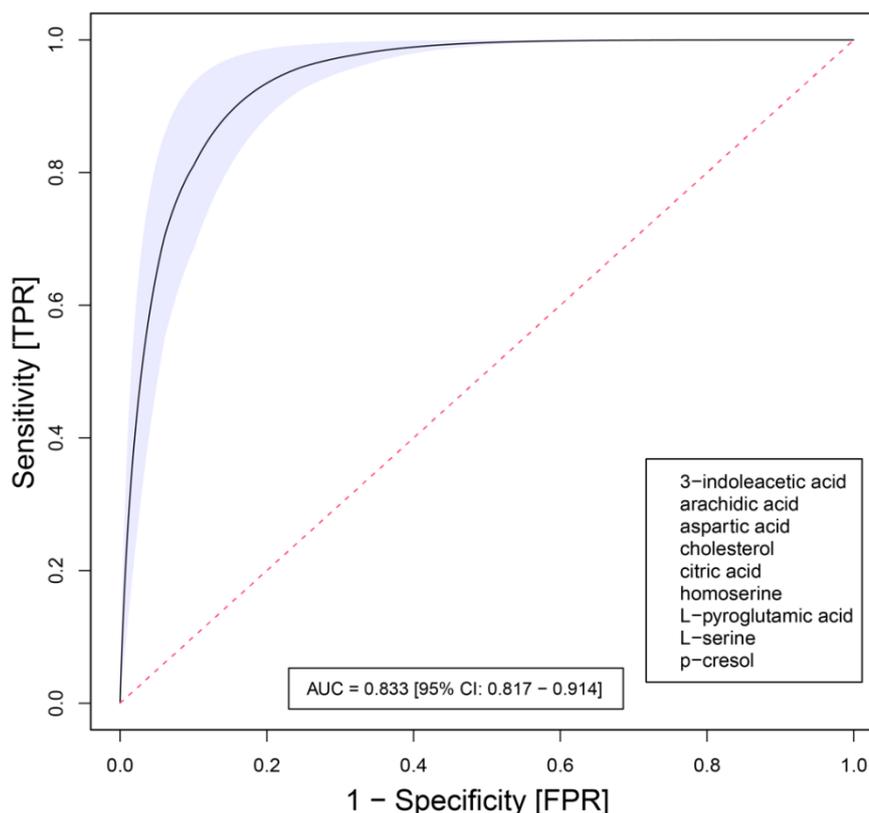
Dentre os métodos de regressão logística mais utilizados (*ridge*, LASSO e *elastic-net*), existem diferenças entre as formas de regressão. Enquanto a regressão *ridge* se utiliza de uma penalidade baseada na multiplicação de um hiper-parâmetro  $\lambda$  com a soma dos quadrados dos coeficientes de regressão, a regressão LASSO utiliza os valores absolutos multiplicados ao hiper-parâmetro  $\lambda$ . A regressão *elastic-net* é uma combinação das regressões *ridge* e LASSO, utilizando ambas as penalidades, com um hiper-parâmetro  $\lambda$  para cada uma das regressões incluídas nessa modalidade, onde os hiper-parâmetros  $\lambda$  são obtidos por meio de validação cruzada (Acharjee *et al.*, 2013). Em metabolômica, os conjuntos de dados não seguem um padrão definido, entretanto, o ideal é que o número de amostras seja o maior possível para melhorar a capacidade de predição do modelo. Além disso, é necessário avaliar o nível de correlação entre as variáveis para selecionar um método de regressão. Dentro do conjunto de dados proposto neste estudo, é possível observar um alto índice de correlação entre as variáveis, como observado no *cluster* de metabólitos na **Figura 1.9**. Dentre os métodos de regressão, o método LASSO possui performance menor quando comparado com os outros dois métodos quando o conjunto de dados apresenta um alto valor de correlação

entre as variáveis quando comparado com o método *ridge*. Em relação ao método *elastic-net*, este apresenta performance menor em comparação com o método *ridge* em números de amostras pequenos (Bujak *et al.*, 2016; Kirpich *et al.*, 2018; Neto *et al.*, 2014), o que é o caso do conjunto de dados analisado neste trabalho. Baseado nos principais contribuidores da análise multi-bloco, plotou-se os *beanplots* de cada metabólito para avaliar a diferença em suas intensidades nos pacientes com BD e nos controles, conforme ilustrado na **Figura 1.10**. Todos os metabólitos, exceto o *p*-cresol, apresentaram níveis reduzidos nos pacientes com BD em comparação com os controles. Dos metabólitos selecionados na análise multi-bloco, fez-se uma regressão logística *ridge*, ajustando-se o modelo para IMC, idade e gênero, para assim obter um painel de metabólitos para a diferenciação do BD em comparação com os controles (**Figura 1.11**).

Obteve-se um painel com 9 metabólitos que possuem o maior potencial de diferenciação (ácido 3-indolacético, ácido araquídico, ácido aspártico, colesterol, ácido cítrico, homoserina, serina, *p*-cresol e ácido L-piroglutâmico), com uma AUC na curva ROC de 0,833 (intervalo de confiança a 95 %: 0,817-0,914).



**Figura 1.10.** *Beanplots* dos metabólitos selecionados na análise de correlação parcial e suas distribuições entre os grupos estudados. As linhas menores apresentam as observações individuais referentes a cada amostra, enquanto as áreas amarelas e verdes representam a área de distribuição total das amostras e a linha maior representa a média de tais valores individuais.



**Figura 1.11.** Painel de metabólitos construído utilizando modelagem de regressão logística [BD x HC] ajustado para idade, gênero e IMC.

Analisando os resultados obtidos, observou-se níveis reduzidos de ácido cítrico nos pacientes com BD, assim como o ácido aspártico, que se correlaciona ao ciclo do ácido cítrico através do 2-oxoglutarato (Araújo *et al.*, 2014). O ciclo do ácido cítrico é a principal rota energética do metabolismo humano, ocorrendo dentro da mitocôndria. Uma das hipóteses levantadas quanto aos mecanismos bioquímicos do BD relaciona as alterações de humor com disfunções na produção de energia na mitocôndria (Gubert *et al.*, 2013; T. Kato, 2017; T. M. Kato *et al.*, 2018; Pereira *et al.*, 2018), que pode também ser correlacionada à hipótese de que o BD é uma doença multissistêmica, afetando diversas funções no corpo humano, inclusive a produção de energia (Leboyer *et al.*, 2012). A serina e homoserina também foram consideradas como estatisticamente significantes para a diferenciação dos pacientes com BD dos controles. A L-serina participa de diversas reações e vias metabólicas, como o metabolismo de glicerofosfolídeos através do intermediário fosfatidilserina, e do ciclo do ácido cítrico, através de sua transformação em piruvato pela enzima

serina deaminase (Ogata *et al.*, 1999). Pålsson *et al.* (2015) analisaram o fluido cerebrospinal de pacientes com BD, que apresentaram níveis reduzidos de diversos metabólitos em comparação com controles saudáveis, entre eles a L-serina. Yoshimi *et al.* (2016) também observaram níveis reduzidos de L-serina no soro de pacientes com BD, corroborando com os resultados obtidos neste estudo. Níveis reduzidos de aminoácidos em pacientes com BD também podem explicar os níveis reduzidos de ácido piroglutâmico nestes pacientes. Visto que o ácido piroglutâmico é produzido na via de metabolismo de glutamato através da transformação de L-aminoácidos em L-ácido piroglutâmico pelas enzimas gama-glutamilttransferase 6 e gama-glutamiltciclotransferase (Ogata *et al.*, 1999), a deficiência em L-aminoácidos pode implicar em níveis reduzidos de ácido piroglutâmico. O ácido araquídico também foi considerado como um metabólito diferencial, com níveis reduzidos em pacientes com BD e é transformado em eicosanoil-CoA, sofrendo  $\beta$ -oxidação para produzir acetil-CoA ao final do ciclo do ácido cítrico (Ogata *et al.*, 1999). Yang *et al.* (Yang *et al.*, 2017) observaram um aumento na  $\beta$ -oxidação em ácidos graxos de cadeia longa em pacientes com esquizofrenia e Al Awam *et al.* (Al Awam *et al.*, 2015) reportaram níveis reduzidos de ácido araquídico em pacientes com esquizofrenia, mostrando que o metabolismo de ácidos graxos pode estar alterado em pacientes com transtornos psiquiátricos, corroborando com os resultados obtidos. Houve também alterações nos níveis de colesterol nos pacientes com BD. Em um estudo de meta-análise, Pillinger *et al.* (Pillinger *et al.*, 2017) compararam os resultados de diversos estudos em relação aos níveis de colesterol, LDL e triglicérides, indicando uma possível correlação entre o colesterol e medicamentos utilizados para o tratamento de transtornos psiquiátricos como lítio (Joshi *et al.*, 2019), haloperidol, clozapina e aripiprazol (Korade *et al.*, 2017). Devido ao fato de os pacientes neste estudo estarem medicados, é possível que os níveis reduzidos de colesterol observados possam estar relacionados a efeitos colaterais da medicação. Entre os metabólitos selecionados em nosso painel, o *p*-cresol foi o único que apresentou níveis aumentados nos pacientes com BD em comparação com os controles. Além de sua função na hemóstase e inflamação, descrita anteriormente, o *p*-cresol pode reduzir Na<sup>+</sup> e K-ATPase (Calderón-Guzmán *et al.*, 2005). Alterações no sistema dopaminérgico, testados em modelos animais (Goodhart *et al.*, 1987; Pascucci *et al.*, 2020), indicam que a administração de *p*-

cresol pode aumentar comportamentos que mimetizam transtornos de ansiedade e aumentar os níveis de dopamina, o que pode correlacionar o aumento dos níveis desse metabólito com as alterações causadas no cérebro causadas pelo BD.

Embora seja possível observar na literatura alterações correlacionadas entre o BD e os metabólitos selecionados no painel da **Figura 1.11**, observa-se também que alterações em diversas vias metabólicas, incluindo do metabolismo energético, também estão relacionadas à possíveis efeitos da medicação, benéficos ou maléficos. Jauhar e Young (2019) descreveram os efeitos medicamentosos de diversos antipsicóticos e estabilizadores de humor em seu trabalho, bem como os efeitos adversos causados por estes medicamentos. No caso do lítio, comumente utilizado em tratamentos para BD, as hipóteses do mecanismo de ação consistem na regulação das propriedades da membrana celular, como transporte e distribuição de íons, na regulação de neurotransmissores e sinalização intracelular. Entre os efeitos colaterais do lítio estão declínio na função renal, hipotireoidismo e hipercalcemia, em especial a longo prazo (Shine *et al.*, 2015). Em se tratando de marcadores periféricos, observa-se que o lítio apresenta ação na atividade mitocondrial e no estresse oxidativo, bem como a modulação de cálcio, fatores neurotróficos e outros mecanismos de sinalização celular, inflamação e metabolismo de glutamato (Machado-Vieira, 2018). Apesar dos efeitos colaterais causados pela medicação, Roux *et al.* relacionaram o tratamento de pacientes com BD com lítio pode estar relacionado a mudanças nos níveis do ácido N-acetilglutâmico e serina, estando relacionado também a outros metabólitos ligados ao metabolismo energético, indicando que o lítio pode atuar melhorando a função mitocondrial (Roux & Dosseto, 2017).

Outras medicações também apresentaram evidências na literatura que indicam efeitos que podem estar relacionados a alterações metabólicas, em especial no metabolismo energético, como o haloperidol (Mcclay *et al.*, 2015), risperidona (Almeida *et al.*, 2020; Cai *et al.*, 2012; Paredes *et al.*, 2014; Xuan *et al.*, 2011) e outras drogas psicotrópicas (Lenski *et al.*, 2021). Além de alterações no metabolismo energético, antipsicóticos de segunda geração já foram reportados como possíveis alteradores da microbiota intestinal, aumentando a

abundância de determinadas bactérias intestinais, onde o perfil bacteriano se assemelhou ao observado em indivíduos com obesidade e disfunções metabólicas, o que pode ser evitado com a indicação de probióticos juntamente com a medicação (Skonieczna-Żydecka *et al.*, 2020). Entretanto, Cai *et al.* (2012) observaram alterações na flora intestinal e em metabólitos do ciclo do ácido cítrico em pacientes não tratados em relação a controles, o que indica uma possível disfunção do metabolismo energético e outros sistemas relacionados antes do tratamento, o que indica que mais pesquisas são necessárias para se precisar quais alterações metabólicas são ocasionadas pela doença e quais são causadas por efeitos da medicação, tanto pela interação medicamentosa, quanto por possíveis efeitos colaterais causados pela medicação.

## 5. Conclusões

A utilização de estratégias de integração ômica e de biologia de sistemas para o estudo de doenças é uma ferramenta poderosa para não somente avaliar as correlações entre as diferentes ômicas, mas também para analisar as possíveis alterações bioquímicas e metabólicas que possam ser causadas por fatores externos. Utilizando uma abordagem focada na correlação canônica foi possível obter uma rede de integração entre proteínas, metabólitos e informações demográficas dos pacientes com BD, o que indicou uma correlação importante entre a doença e processos de hemóstase, que está diretamente ligado a comorbidades como DCVs, que são a principal causa de morte dentre os pacientes com BD. Ao obtermos mais informações relativas a essas interações, podemos entender melhor a fisiopatologia da doença e, assim, contribuir para futuros estudos na área que possam melhorar a qualidade de vida de tais pacientes. Além disso, visto que o diagnóstico do BD é realizado através de observações médicas e sujeita a erros, a proposição de uma assinatura molecular que possa estratificar os pacientes bipolares de indivíduos saudáveis pode contribuir também para estudos futuros em relação ao diagnóstico da doença. A hipótese de que o BD se trata de um transtorno multissistêmico envolvendo diferentes aspectos do metabolismo ainda precisa ser mais estudada para que alguma conclusão concreta quanto a como os sistemas estão interligados. A partir dos resultados obtidos, foi possível demonstrar que, além de processos de hemóstase, a produção de energia também pode estar correlacionada ao BD, bem como alterações na microbiota intestinal.

**APLICAÇÃO DE DUAS METODOLOGIAS DE APRENDIZADO DE MÁQUINA  
NA DIFERENCIAÇÃO MULTI-GRUPO DE PACIENTES COM TRANSTORNO  
BIPOLAR E ESQUIZOFRENIA E SUA RELAÇÃO COM AMIDAS GRAXAS**

## 1. Introdução

### 1.1. Metodologias de aprendizado de máquina em estudos metabolômicos aplicados a transtornos mentais

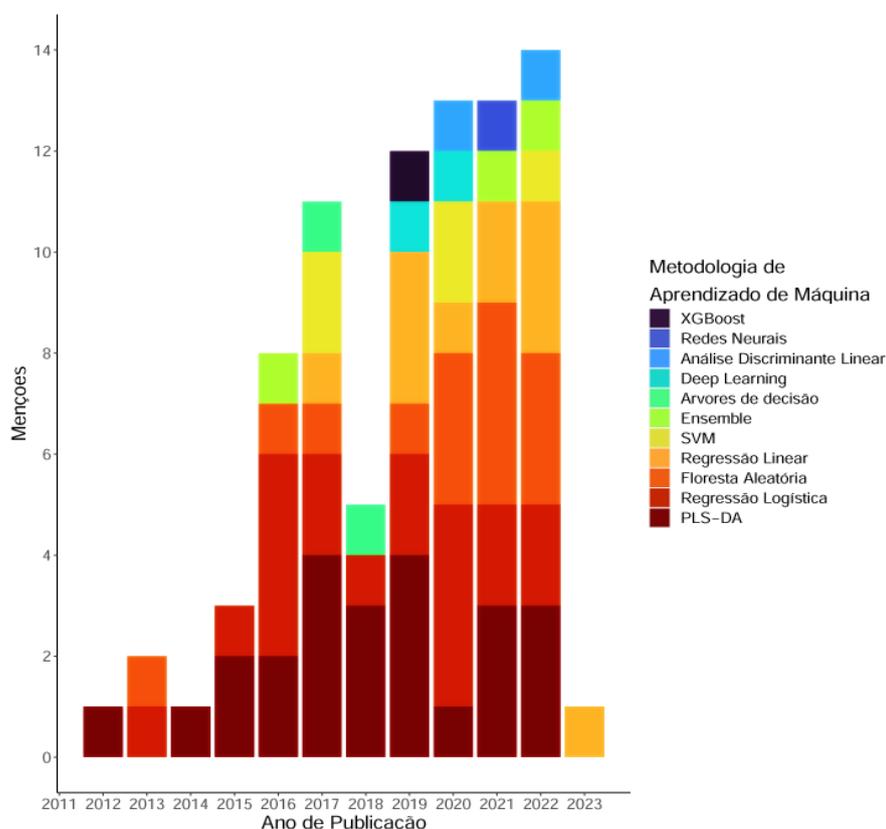
Na última década, poucos desenvolvimentos científicos na área computacional geraram tanta expectativa quanto a expansão de metodologias de aprendizado de máquina. Com a melhora na capacidade de processamento dos computadores e o advento de metodologias de análise que produzem quantidades de dados grandes e complexas, conhecido como “*big data*” (Tandon & Tandon, 2019), se tornou necessário o uso de métodos de análise de dados que consigam lidar com esta vasta quantidade de informação. Dentro da ciência de dados, duas diferentes metodologias competem entre si: a estatística clássica, que visa formalizar relações entre variáveis com base em modelos matemáticos, a fim de entregar inferências estatísticas para as variáveis medidas, assumindo que estas são uma amostra de uma população maior; e metodologias de aprendizado de máquina, que se baseiam em algoritmos computacionais que otimizam de forma iterativa sua capacidade de predição sem precisarem necessariamente de conjecturas estatísticas para tal (Mendez et al., 2019).

No campo das ciências biomédicas, o crescimento das ciências ômicas ao longo dos anos resultou na aquisição de diversas assinaturas moleculares para o desenvolvimento da medicina personalizada (Koppad et al., 2021), entretanto, a vasta quantidade de dados gerados dentro destas metodologias faz com que seja necessário o uso de abordagens bioinformáticas avançadas, em especial em dados cujo número de variáveis supera o número de amostras, como é o caso da metabolômica baseada em espectrometria de massas (Koppad et al., 2021; Liebal et al., 2020). Além disso, o tratamento de dados metabolômicos de espectrometria de massas possui outros desafios, como a presença de ruídos de fundo, efeitos de batelada durante análises e *missing values* (Liebal et al., 2020), que podem ser minimizados com a utilização de técnicas já desenvolvidas para estes problemas, como a aquisição de amostras de fundo (e a sua subsequente utilização para subtrair os sinais de

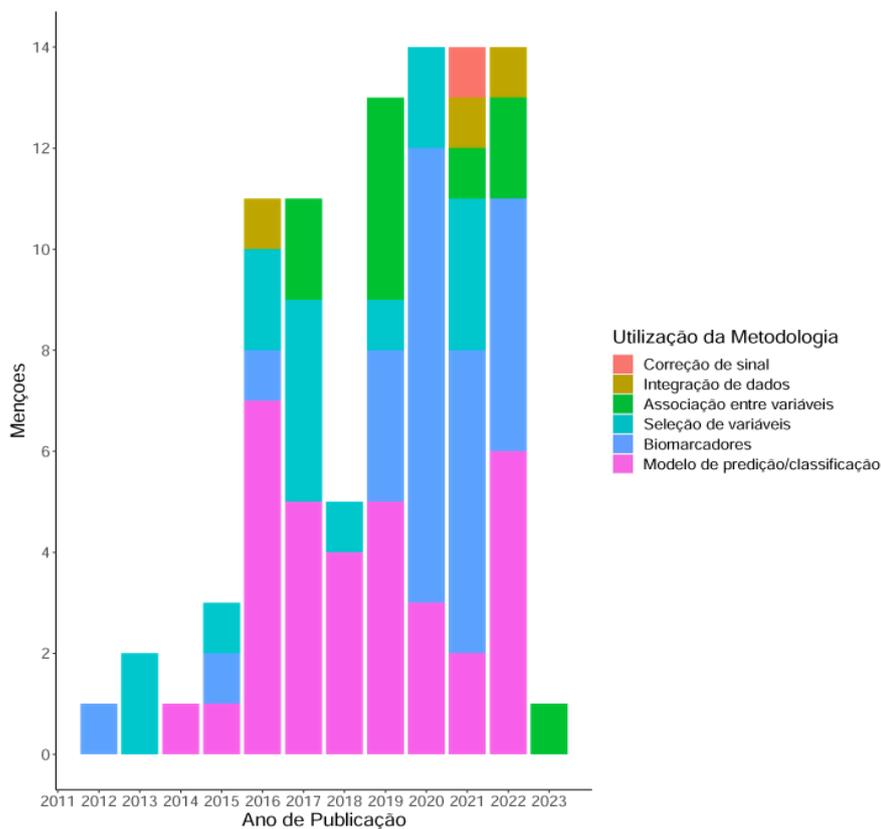
fundo das amostras teste), utilização de amostras de controle de qualidade para verificar a presença ou não de efeitos de batelada (Broadhurst et al., 2018), e a substituição dos *missing values* na etapa de pré-processamento por valores pequenos diferentes de zero, como o valor do limite de detecção do equipamento, ou valores estimados utilizando metodologias de estatística multivariada como os K-vizinhos mais próximos (KNN), por exemplo (Do et al., 2018). Mesmo com metodologias já estabelecidas para lidar com diversos problemas relacionados à complexidade dos dados metabolômicos, novas abordagens computacionais ainda são necessárias para lidar com problemas ainda pendentes, como redução de variáveis e modelos de classificação e regressão, por exemplo.

Tratando-se de abordagens metabolômicas, a metodologia de aprendizado de máquina mais utilizada é a projeção a estruturas latentes (PLS, do inglês *projection to latent structures*) e sua extensão de análise discriminante (PLS-DA). Isso se deve ao fato da PLS projetar dados altamente multivariados em um espaço coordenado menor (variáveis latentes), que permite que dados com mais variáveis que amostras possam ser modeladas sem a necessidade de pré-filtrar essas variáveis (Mendez et al., 2019), o que vai ao encontro à estrutura dos dados metabolômicos, que naturalmente possuem mais variáveis do que amostras. Além disso, a sua extensão de análise discriminante (PLS-DA) fornece informações importantes para interpretação dos resultados obtidos, como a importância dos pesos (*loadings*) e a importância da variável na projeção (VIP), que são utilizados para determinar as variáveis mais importantes na construção do modelo de classificação (Gromski et al., 2015). Além do uso da PLS-DA, estudos metabolômicos fazem uso de diversas outras metodologias para a construção de modelos de classificação e regressão, redução de variáveis, processamento de picos, integração multi-ômica e análise de vias, que utilizam algoritmos como floresta aleatória, árvores de decisão, máquinas de vetor de suporte (SVM), redes neurais, *ensemble*, algoritmos do tipo *boosting*, *deep learning* e regressões lineares e logísticas (Galal et al., 2022; Liebal et al., 2020; Mendez et al., 2019). Na área de metabolômica aplicada a transtornos mentais, a aplicação de metodologias de aprendizado de máquina vem sendo cada vez mais utilizada para a resolução de problemas, construção de modelos, análise

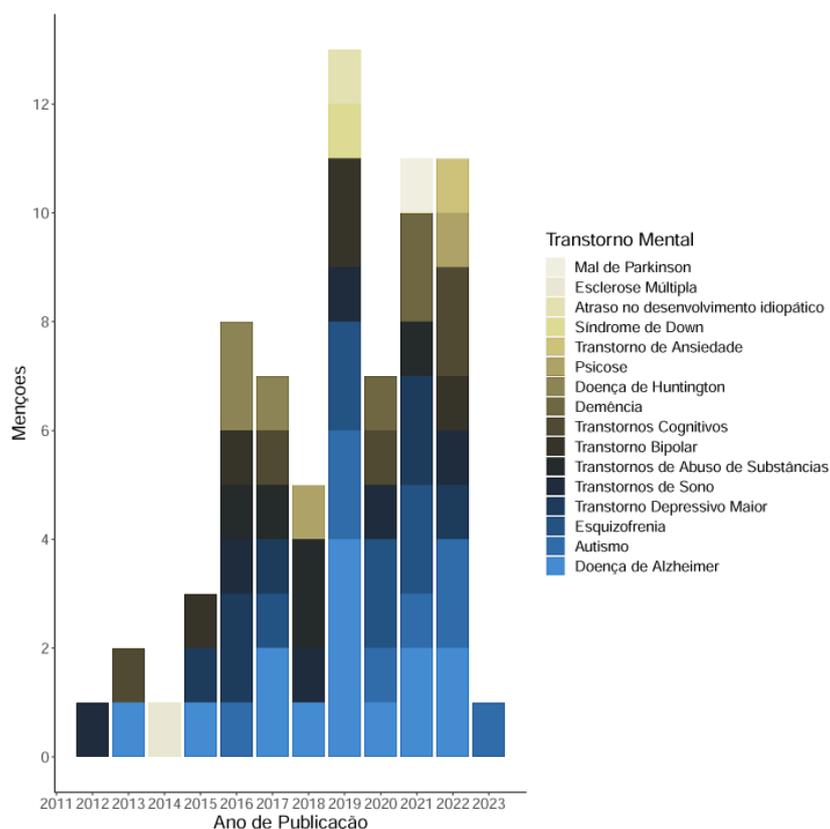
de interações entre variáveis, e prospecção de biomarcadores, além de também ser utilizada para a integração de dados ômicos de diferentes plataformas. Utilizando a plataforma PubMed, é possível observar a utilização de tais metodologias para o tratamento de dados metabolômicos aplicados a transtornos mentais em trabalhos originais, ou seja, removendo-se artigos de revisão, através do uso dos operadores booleanos *((metabolomics[Title/Abstract]) AND ("PLS-DA" OR "ensemble" OR "boosting algorithm" OR "linear discriminant analysis" OR "decision trees" OR "K-means" OR "Naive Bayes" OR "linear regression" OR "logistic regression" OR "random forest" OR "SVM" OR "neural networks" OR "deep learning") NOT (Review[Publication Type])) AND (mental illness)*. Assim, levantou-se não somente as metodologias de aprendizado de máquina utilizadas, mas também a aplicação da metodologia e o objeto de estudo da publicação. Tais levantamentos são ilustrados nas **Figuras 2.1, 2.2 e 2.3**.



**Figura 2.1.** Metodologias de aprendizado de máquina aplicadas a abordagens metabolômicas em transtornos mentais na plataforma Pubmed.



**Figura 2.2.** Forma de utilização da metodologia de aprendizado de máquina nos estudos metabolômicos aplicados a transtornos mentais na plataforma Pubmed.



**Figura 2.3.** Transtornos mentais estudados utilizando metodologias de aprendizado de máquina na plataforma Pubmed.

Avaliando os dados obtidos, observa-se na **Figura 2.1** que a metodologia mais utilizada para a análise de dados metabolômicos aplicados a transtornos mentais ao longo dos anos ainda é a PLS-DA, entretanto, ao se observar a tendência de crescimento observada nos últimos anos, a utilização desta metodologia se estagnou, enquanto outras abordagens ganharam força, como a floresta aleatória. Além disso, outras abordagens como *deep learning*, redes neurais e SVM começaram a serem utilizadas dentro de estudos metabolômicos. Em relação à forma na qual o aprendizado de máquina é utilizado em cada estudo, é evidente na **Figura 2.2** a utilização massiva de tais metodologias para a construção de modelos e para a prospecção de biomarcadores, entretanto, é importante pontuar a utilização de tais abordagens na integração de dados ômicos, que por sua natureza são dados extremamente complexos e multidimensionais, sendo necessário assim um algoritmo robusto para efetuar com sucesso tal demanda. Por sua vez, na **Figura 2.3** é interessante observar a grande variedade de transtornos mentais que são estudados utilizando metabolômica e aprendizado de máquina. Doenças neurodegenerativas como Alzheimer, Parkinson, Huntington, demência e esclerose múltipla ainda são muito estudadas dentro do universo dos transtornos mentais e da medicina diagnóstica, visto que suas bases moleculares ainda não foram completamente elucidadas, especialmente se tratando de marcadores diagnósticos, que não possuem precisão perfeita até o momento (Dugger & Dickson, 2017). Dentro do espectro dos transtornos de humor, como depressão, esquizofrenia, transtorno bipolar, psicoses e ansiedade, o estudo da esquizofrenia e do transtorno depressivo maior ainda se mantém como o grande protagonista, seguido pelo transtorno bipolar. Assim como para as doenças neurodegenerativas, não se possui um entendimento claro da fisiopatologia dos transtornos de humor supracitados, fazendo com que a busca por biomarcadores de tais doenças seja cada vez mais procurada pela comunidade científica. Na **Figura 2.3** é possível observar também a presença de transtornos menores como distúrbios do sono e transtornos por uso de substâncias, assim como o extenso número de publicações envolvendo o estudo de doenças do desenvolvimento como o autismo e transtornos cognitivos, majoritariamente também para a busca de biomarcadores diagnósticos, visto que o seu

diagnóstico ainda demanda a avaliação de diversos fatores, como o histórico do desenvolvimento do indivíduo, análise de fala e habilidades intelectuais (West et al., 2014).

## 1.2. Espectrometria de massas de alta resolução

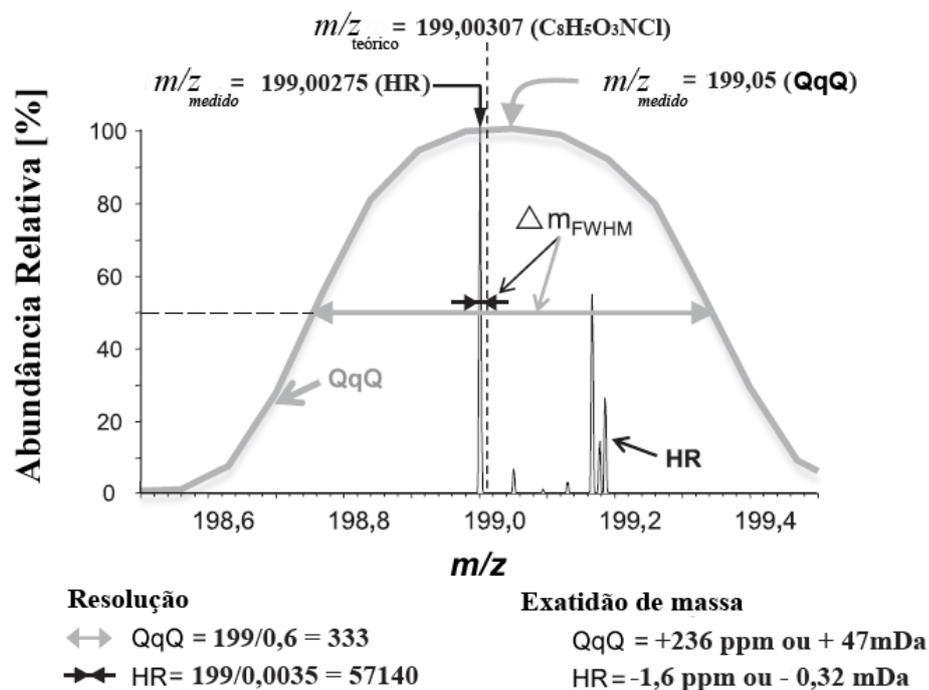
Dentre os muitos avanços na espectrometria de massas desde sua criação, há mais de um século atrás, a melhoria no poder de resolução dos espectrômetros pode ser considerada uma (se não a mais) importante evolução no que diz respeito a criação de diferentes possibilidades de aplicação, contribuindo para o desenvolvimento de metodologias que necessitavam deste tipo de diferenciação molecular, como por exemplo a proteômica e a petroleômica (Marshall & Hendrickson, 2008). Metodologias como a metabolômica também se beneficiam de equipamentos de alta resolução, visto que diferentes adutos (no caso do uso de metodologias baseadas em LC-MS) e compostos isóbaros podem apresentar interferências na identificação do composto de interesse. No passado, era comum associar o uso de equipamentos com analisadores de massas de alta resolução somente com análises exploratórias (*fingerprint*), enquanto equipamentos como o triplo quadrupolo (QqQ) e o *ion trap* (IT) eram associados com análises quantitativas. Apesar de nos dias atuais o QqQ e o IT ainda serem utilizados como equipamentos de escolha para análises quantitativas (em especial para experimentos como o monitoramento de íon selecionado - SIM - e monitoramento de reações selecionadas - SRM), os equipamentos de alta resolução como o *quadrupole-time-of-flight* (QTOF) e o *Orbitrap* sofreram avanços significativos em suas tecnologias a partir de 2010, sendo capazes também de efetuar estudos quantitativos com qualidade (Rochat, 2016).

Dentro da espectrometria de massas, duas principais figuras de mérito são responsáveis por explicar a detecção seletiva de diferentes compostos com  $m/z$  similares: a resolução (**Eq. 1**) e a exatidão de massa (**Eq. 2**) (Dias et al., 2016; Lai & Wang, 2022; Perez de Souza et al., 2021; Rochat, 2016).

$$Resolução = \frac{m}{\Delta m_{FWHM}} \quad (1)$$

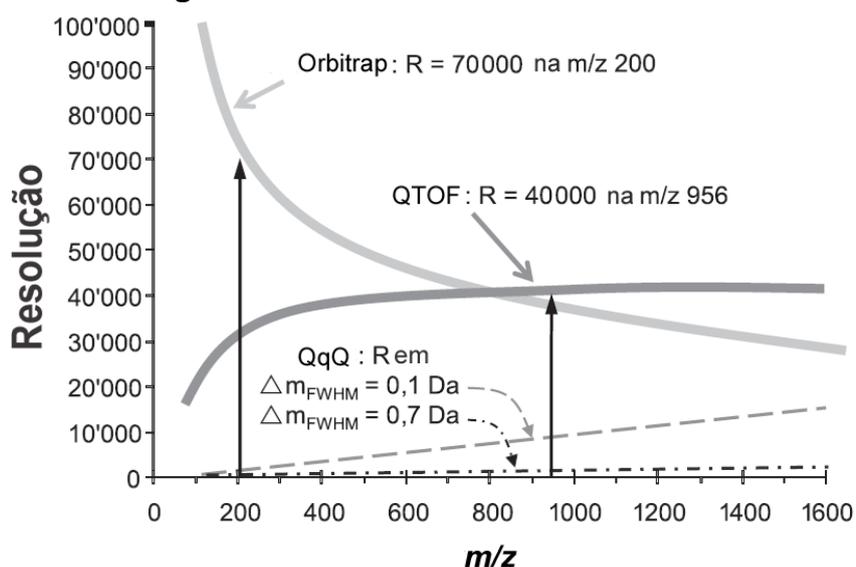
$$Exatidão_{(ppm)} = \frac{m/z_{medida} - m/z_{teórica}}{m/z_{teórica}} \times 10^6 \quad (2)$$

A resolução se refere à capacidade do equipamento de separar íons, e é dada pela  $m/z$  do analito dividida pelo delta da distribuição de  $m/z$  na largura total do pico na metade de sua altura máxima (FWHM, do inglês *full width half maximum*), e não deve ser confundido com poder de resolução (PR), que é a habilidade de um equipamento de fornecer um valor específico de resolução de massas (Perez de Souza et al., 2021; Vessecchi et al., 2010). A exatidão de massa, por sua vez, dita a diferença entre a  $m/z$  experimental e a  $m/z$  teórica, e é utilizada especialmente para a determinação da fórmula molecular do composto de interesse, restringindo assim as possíveis opções de compostos possíveis na etapa de identificação nos bancos de dados (Perez de Souza et al., 2021). Os equipamentos normalmente considerados como espectrômetros de massa de alta resolução são aqueles que conseguem rotineiramente alcançar um PR acima de 10000, que faz com que seja possível identificar dois compostos com a mesma massa nominal, porém com composições elementares diferentes. Na **Figura 2.4** podemos visualizar a diferença entre um espectrômetro de massas de alta resolução (HR) e um de baixa resolução (QqQ), a resolução alcançada pelo equipamento na  $m/z$  selecionada e sua exatidão de massa.



**Figura 2.4.** Diferenças de resolução entre equipamentos de baixa resolução (QqQ) e alta resolução (HR). Adaptado de Rochat, 2016.

Apesar de ambos serem considerados como equipamentos de alta resolução, o QTOF e o Orbitrap possuem particularidades em relação às suas resoluções. Enquanto o QTOF apresenta um aumento da resolução em massas mais altas, o Orbitrap apresenta tendência contrária, sendo extremamente alto em massas baixas e decaindo com o aumento da  $m/z$  (Rochat, 2016), como pode ser observado na **Figura 2.5**.



**Figura 2.5.** Resolução usual de diferentes tipos de espectrômetros de massas (Orbitrap, QTOF e QqQ) em diferentes  $m/z$ . Adaptado de Rochat, 2016.

A utilização de espectrômetros de massa de HR em metabolômica vai ao encontro de auxiliar um dos maiores gargalos no fluxo de trabalho de análises metabolômicas, que é a etapa de identificação dos metabólitos analisados. Reduzir o número de possíveis metabólitos dentro de uma  $m/z$  que pode ser considerada diferencial (através de análises estatísticas) é essencial para facilitar a análise dos dados obtidos e evitar assim falsos positivos, entretanto, mesmo identificando a fórmula molecular de um composto de interesse não é possível se obter a totalidade de sua informação estrutural, devido à grande quantidade de possíveis compostos isóbaros e/ou adutos que possuam uma  $m/z$  dentro da faixa de erro do equipamento. Dentro dos padrões acordados pela *Metabolomics Standards Initiative* para a identificação de metabólitos (Sumner et al., 2007; Viant et al., 2017), quatro níveis de identificação podem ser obtidos em análises metabolômicas, sendo o nível 1 o mais alto e o nível 4 o mais baixo:

Nível 1: Compostos totalmente identificados e verificados por duas diferentes técnicas analíticas ortogonais, utilizando padrões comerciais e fornecendo informações apropriadas, como massa exata e tempo de retenção, dados de fragmentação e padrão isotópico;

Nível 2: Compostos putativamente identificados, sem o uso de padrões comerciais, porém, baseado nas propriedades físico-químicas e/ou na similaridade espectral com bibliotecas espectrais;

Nível 3: Compostos caracterizados putativamente por sua classe, baseado em informações físico-químicas de uma determinada classe de compostos ou através da similaridade espectral com compostos de uma determinada classe;

Nível 4: Compostos desconhecidos, que apresentem um sinal discernível, e que possam ser detectados e quantificados com reprodutibilidade.

Assim, para se obter um nível elevado de confiança na identificação de metabólitos, deve-se fazer uso de técnicas complementares para a elucidação de sua estrutura, como por exemplo o uso de experimentos de

espectrometria de massas em sequência (MS/MS), que envolvem a fragmentação das moléculas (Dias et al., 2016). Para a obtenção da informação estrutural dos metabólitos através da fragmentação, diferentes modos de aquisição de dados podem ser efetuados, como a aquisição dependente de dados (DDA, do inglês *data dependent acquisition*), a aquisição independente de dados (DIA, do inglês *data independent acquisition*) ou a análise *targeted* de íons de interesse, que comprime metodologias como o monitoramento de reações selecionadas (SRM, do inglês *selective reaction monitoring*) e monitoramento paralelo de reações (PRM, do inglês *parallel reaction monitoring*), por exemplo.

No modo de aquisição DDA, o equipamento automaticamente alterna entre o modo de aquisição MS e MS/MS quando os íons precursores atingem um limiar pré-estabelecido de intensidade (Fenaille et al., 2017), fragmentando assim uma parte dos íons analisados no modo *full scan*. Neste modo, é possível adquirir espectros de MS/MS dentro de uma janela de isolamento pequena (cerca de 1 Da), levando a espectros de fragmentação mais limpos e de melhor qualidade. O problema deste modo de aquisição é o fato de o equipamento selecionar somente os íons precursores de maior intensidade, o que pode levar à seleção de íons biologicamente irrelevantes (Hilaire et al., 2020). No modo de aquisição DIA, por sua vez, todos os íons dentro de uma janela de isolamento são fragmentados e a informação do íon precursor é perdida (Dias et al., 2016). Dentro da aquisição DIA, duas metodologias são utilizadas, a MS<sup>E</sup> (também conhecida por AIF, ou *All-Ion Fragmentation*), que fragmenta todos os íons, fornecendo um espectro de fragmentação com os fragmentos de todos os íons na amostra (Zhu et al., 2014) e a SWATH (*Sequential Window Acquisition of all Theoretical fragment-ion spectra*), que ao invés de fragmentar todos os íons precursores juntos, utiliza uma janela de isolamento menor (entre 20-50 Da) e fragmenta todos os íons dentro da janela de isolamento selecionada (Fenaille et al., 2017; Hilaire et al., 2020). Nas abordagens *targeted*, as metodologias SRM e PRM são muito similares, porém, diferem na forma na qual os íons precursores são escolhidos. Na abordagem SRM, cada transição entre íon precursor e íon produto é monitorada de cada vez, enquanto na abordagem PRM todos os íons produtos derivados de um precursor são analisados com alta resolução e exatidão de massa (Rauniyar, 2015). A abordagem SRM normalmente é

realizada em equipamentos com analisadores de massas de baixa resolução, como QqQ e IT, enquanto a abordagem PRM é exclusiva de equipamentos que possuam um quadrupolo acoplado a um analisador de alta resolução (QOrbitrap e QTOF), visto que todos os íons produtos são analisados simultaneamente no analisador de alta resolução após a fragmentação (Peterson et al., 2012).

### 1.3. Amidas graxas e sua relação com alterações cerebrais

Dentro das variadas classes de lipídios endógenos, a classe das amidas graxas vem ganhando importância ao longo dos anos, especialmente após a descoberta da anandamida (N-araquidonoiletanolamina, ou AEA) como um composto com propriedades endocanabinóides (Devane et al., 1992). Amidas graxas são definidas como compostos que possuem um núcleo comum contendo  $R_1-CO-N-R_{(2/3)}$ , onde o radical  $R_1$  contém normalmente três ou mais átomos de carbono (Waluk et al., 2014), e os radicais  $R_2$  e  $R_3$  ligados ao átomo de nitrogênio podem variar das mais variadas formas para diferentes subclasses de moléculas. De acordo com a notação proposta pela plataforma LIPID MAPS (Liebisch et al., 2020), as amidas graxas se encontram dentro do grupo geral de acil graxos (FA, do inglês *Fatty Acyls*) e contêm as subclasses das amidas graxas primárias, N-acil amidas, N-acil etanolaminas e homoserinalactonas de amidas graxas. Entretanto, há outros compostos conhecidos que possuem uma denominação própria, porém, não são classificados por seus subgrupos dentro da plataforma, como no caso os N-acil aminoácidos e as N-acil dopaminas (que são classificados somente como N-acilamidas). Na **Figura 2.6** é possível observar as estruturas padrão de cada classe de amidas graxas e suas denominações.

Classe de amida graxa	Estrutura
N-acilamida (NAM)	
N-acilaminoácidos (NAA)	
N-acildopamina (NDA)	
N-aciletanolamina (NAE)	
Amidas graxas primárias (PFAM)	

**Figura 2.6.** Estrutura e classe das diversas subclasses das amidas graxas, onde  $R_1$  corresponde a um grupo acila com três ou mais átomos de carbono, e o radical  $R_2$  dos N-acilaminoácidos caracteriza o aminoácido em questão. Adaptado de Farrell & Merkler, 2008.

Os N-acilaminoácidos, também chamados de ácidos elmíricos, são moléculas sinalizadoras endógenas que regulam processos de dor e inflamação (Burstein, 2018), que inicialmente foram encontradas em meados dos anos 2000, e que vêm sendo estudadas ao longo dos anos por sua potencial atividade anti-inflamatória e analgésica. O primeiro N-acilaminoácido sintetizado e estudado foi a N-araquidonoilglicina, por possuir similaridades estruturais com a anandamida, sendo em seguida identificado endogenamente no cérebro, onde este apresentou atividade no receptor GPR18 (Arul Prakash & Kamlekar, 2021), hoje considerado como um potencial receptor secundário do sistema endocanabinóide. Dentre os N-acilaminoácidos, destacam-se a presença das N-acilglicinas, N-aciltaurinas e N-acilserinas. Relatos na literatura apontam a atividade das N-acilglicinas em funções como regulação inflamatória, mediação sensorial e da neurotransmissão, como é o caso da N-araquidonoilglicina (Burstein, 2018), redução de atividade motora e respostas cerebrais de abstinência, para a N-oleoilglicina e efeitos de sinalização celular e ativação dos canais de  $Ca^{2+}$  para a N-palmitoilglicina (Arul Prakash & Kamlekar, 2021). No

caso das N-aciltaurinas, é observado ação nos receptores TRPV1, outro receptor que é considerado como potencial receptor secundário do sistema endocanabinóide (Ezzili et al., 2010), além de ação inibitória da enzima FAAH (*fatty acid acyl hydrolase*) e regulação dos canais de  $Ca^{2+}$  no cérebro e no pâncreas (Arul Prakash & Kamlekar, 2021). As N-acilserinas por sua vez, mostraram ação na proliferação, proteção e diferenciação de celular progenitoras neurais, indicando uma potencial capacidade neuroprotetiva, além de melhorar funções neurocomportamentais e cognitivas (Arul Prakash & Kamlekar, 2021). Observa-se também ação no sistema nervoso central de diversos outras N-acilaminoácidos como a N-araquidonoil- $\gamma$ -aminobutarato (NAGABA), N-linoleilalanina, N-linoleiltirosina e derivados de N-estearoilaminoácidos (tirosina, serina e treonina) (Arul Prakash & Kamlekar, 2021; Burstein, 2018; Waluk et al., 2014).

Dentre as N-acildopaminas, há relatos da existência de algumas poucas dentro do sistema nervoso de mamíferos como endógenas. Sabe-se que a N-araquidonoildopamina é um ativador do receptor endocanabinóide  $CB_1$  e agonista do receptor TRPV1 (este ativado também pela N-oleoildopamina) (Ezzili et al., 2010; Farrell & Merkler, 2008; Leishman & Bradshaw, 2015; Waluk et al., 2014), além de ambos apresentarem aumento no controle de dor relacionada a temperatura (Divito & Cascio, 2013) e potencial ação pró-convulsiva relacionada à N-oleoildopamina (Cristino et al., 2020). Outras N-acildopaminas como a N-palmitoildopamina e a N-esteaorildopamina também foram encontradas em diferentes regiões do cérebro, entretanto, as suas funções biológicas ainda são desconhecidas (Ezzili et al., 2010; Farrell & Merkler, 2008; Waluk et al., 2014).

O grupo das N-aciletanolaminas são sem sombra de dúvida os compostos mais estudados dentre as amidas graxas, especialmente devido à sua potencial atividade endocanabinóide, em especial a anandamida, que possui atividade nos receptores  $CB_1$  e  $CB_2$ , além de regular diversos processos metabólicos relacionados à cascata do ácido araquidônico (Bisogno & Marzo, 2007) e na produção de prostaglandinas (Urquhart et al., 2015). Além da anandamida, diversas outras N-aciletanolaminas apresentam atividade endocanabinóide, como a N-linoleiletanolamina (atividade nos receptores  $CB_1$ ,  $CB_2$  e  $GPR_{18}$ ), a N-palmitoiletanolamina (atividade nos receptores  $CB_2$  e  $GPR_{55}$ )

(Cristino et al., 2020; Leishman & Bradshaw, 2015) e a N-oleoiletanolamina (atividade nos receptores TRPV1 e GPR<sub>119</sub>) (Cristino et al., 2020). Além de apresentarem atividade endocanabinóide, as N-aciletanolaminas possuem diversas funções metabólicas, como por exemplo regulação da temperatura corporal, locomoção e saciedade e percepção de dor, medo e ansiedade, no caso da anandamida (Farrell & Merkler, 2008), efeitos anti-inflamatórios e propriedades que simulam a anorexia, no caso da N-esteoriletanolamina (Leishman & Bradshaw, 2015), efeitos analgésicos, antiepiléticos e neuroprotetivos, no caso da N-palmitoiletanolamina (Tsuboi et al., 2018). Observou-se também possíveis efeitos neuroprotetivos para a N-linoleoiletanolamina e N-lauroiletanolamina, e potenciais funções relacionadas ao controle de apetite, melhora da memória e regulação do sono e do sistema de alerta no caso da N-oleoiletanolamina (Waluk et al., 2014). Outra N-aciletanolamina que não possui atividade endocanabinóide, porém possui relevância no sistema nervoso central é a docosahexanoiletanolamina (NAE C22:6), que possui ação neurogênica e sinaptogênica (Tsuboi et al., 2018). Além disso, seus compostos de oxidação possuem potencial anti-inflamatório, efeitos na agregação de plaquetas e proteção de órgãos (Urquhart et al., 2015).

As amidas graxas primárias, por sua vez, são majoritariamente lipídios identificados como moléculas sinalizadoras (Divito & Cascio, 2013), entretanto, o interesse nessa classe de moléculas surgiu a partir do momento em que se descobriu a função da oleiamida como potencial indutora de sono em modelos animais (Cravatt et al., 1995), dando início ao estudo constante de diferentes amidas graxas primárias e suas funções no metabolismo. A oleamida, além de possuir função no controle de sono, mostrou capacidade de controle no fluxo de Ca<sup>2+</sup>, controle da locomoção, regulação de processos de memória, modulação de receptores de medicação no sistema nervoso central e ativação de receptores GABA e de serotonina (Farrell & Merkler, 2008). A linoleiamida por sua vez, também possui funções na regulação do fluxo de Ca<sup>2+</sup>, além de apresentar funções no bloqueio da junção comunicante, e juntamente com a elaidamida (C18:1) apresenta função inibitória das enzimas epóxido hidrolase e fosfolipase A<sub>2</sub> (Waluk et al., 2014). Outras amidas graxas primárias apresentam poucos estudos relacionados às suas funções no metabolismo, como por

exemplo a erucamida, que apresenta função angiogênica e de regulação de fluídos, além de apresentar potencial como antidepressivo e ansiolítico (Ezzili et al., 2010; Li et al., 2017); araquidonamida e palmitoleiamida, que apresentam relação com o controle da junção comunicante (Divito & Cascio, 2013; Ezzili et al., 2010; Waluk et al., 2014); e palmitamida, que possui ação moderada anticonvulsivante (Waluk et al., 2014). É interessante também citar que, mesmo as amidas graxas primárias não atuando em nenhum receptor do sistema endocanabinóide, elas são hidrolisadas pela enzima FAAH, juntamente com outras classes de amidas graxas como as N-aciletanolaminas e N-acilaminoácidos (Ueda et al., 2000) e que sofre ação inibitória pelo canabidiol (CBD) (Maione et al., 2011), um dos componentes da cannabis.

Apesar de diversas amidas graxas não possuírem ligação direta com os receptores endocanabinóides (CB<sub>1</sub> e CB<sub>2</sub>) ou com receptores considerados como parte do “sistema endocanabinóide expandido”, como os receptores da classe GPR e TRPV (Cristino et al., 2020), a grande maioria destes compostos participam de forma indireta, seja por ação de enzimas que também atuam em compostos endocanabinóides, seja por serem substratos para a síntese de compostos com atuação nos receptores supracitados. Além da relação com o sistema endocanabinóide, diversas amidas graxas apresentam funções no metabolismo cerebral, na hemóstase e em processos de sinalização, processos estes relacionados a condições psiquiátricas diversas, como o transtorno bipolar e a esquizofrenia. Neste capítulo, observou-se a presença de diversas amidas graxas como metabólitos diferenciais em pacientes bipolares e esquizofrênicos em comparação com controles saudáveis, que foram selecionados através do uso de duas diferentes metodologias de aprendizado de máquina em dados de espectrometria de massas de alta resolução, e a relação destas duas condições com tais metabólitos e conseqüentemente com o sistema endocanabinóide.

## 2. Objetivos

### 2.1. Objetivos Gerais

O objetivo deste estudo visou avaliar duas diferentes metodologias de aprendizado de máquina e de seu potencial de diferenciação do perfil metabólico do plasma de pacientes *drug-naïve* com transtorno bipolar (BD) e esquizofrenia (SZ) comparados com controles saudáveis (HC).

### 2.2. Objetivos Específicos

- Análise dos metabólitos presentes no plasma sanguíneo de pacientes com BD, SZ e HC;
- Aplicação de metodologias de aprendizado de máquina para a determinação de um painel de metabólitos com o maior potencial de diferenciação dentre cada grupo analisado;
- Identificação e análise dos metabólitos selecionados.

### 3. Parte experimental

#### 3.1. Pacientes e amostragem

Este estudo foi conduzido em parceria com o Instituto de Psiquiatria da Universidade de São Paulo, consistindo em 79 controles e 77 pacientes *drug-naïve* (46 SZ e 31 BD), ou seja, pacientes que nunca fizeram uso de nenhuma medicação psiquiátrica até o momento da coleta das amostras. Este estudo foi aprovado pelo Comitê de Ética da Universidade de São Paulo (CAPPesq N° 943.883) e realizado de acordo com os protocolos da Declaração de Helsinki. Todos os participantes assinaram um termo de consentimento antes da inclusão deles no estudo. O diagnóstico de SZ foi realizado de acordo com o Manual Estatístico de Doenças Mentais (DSM-IV), sendo confirmado em seguida utilizando o SCID-P (*Structured Clinical Interview Axis I*, versão 2.0), para o DSM-IV. A psicopatologia dos pacientes foi avaliada utilizando a escala PANSS (*Positive and Negative Syndrome Scale*), incluindo as subescalas positivas e negativas e psicopatologia geral. Os sintomas maníacos e depressivos dos pacientes com BD foram avaliados utilizando a Escala de Depressão de Hamilton (HAM-D) e a Escala Classificatória de Mania de Young (YMRS). Pacientes com outros transtornos psiquiátricos ou neurológicos foram excluídos do estudo. Amostras de sangue dos participantes foram coletadas após jejum de 8 horas em tubos Vacutainer revestidos com EDTA, específicos para a separação do plasma sanguíneo. As amostras foram em seguida centrifugadas a 20 °C e 1800 g durante 15 minutos e estocadas em biofreezer a -80 °C. Os dados sócio demográficos dos pacientes e controles são apresentados na **Tabela 2.1**.

**Tabela 2.1.** Dados demográficos das amostras utilizadas na análise de metabólitos utilizando aprendizado de máquina (SD: desvio padrão; BD: pacientes com transtorno bipolar; SZ: pacientes com esquizofrenia; HC: controles saudáveis)

Fatores	BD (n = 31)	SZ (n = 46)	HC (n = 79)	p-valor
<b>Idade (anos); média (SD)</b>	26 (4,6)	26 (6,9)	27 (4,5)	0,699 <sup>a</sup>
<b>Gênero M/F; n (%)</b>	11/20 (36/64)	30/16 (65/35)	51/28 (65/35)	0,012 <sup>b</sup>
<b>PANSS; Média (SD)</b>		62 (26)		
<b>PANSS - Sintomas positivos; Média (SD)</b>		14 (7)		
<b>PANSS - Sintomas negativos; Média (SD)</b>		15 (8)		
<b>HAM-D; Média (SD)</b>	10 (8)			
<b>YMRS; Média (SD)</b>	8 (12)			

<sup>a</sup> ANOVA, fator único, <sup>b</sup> Pearson Chi-Square

### 3.2. Análise metabolômica utilizando espectrometria de massas de alta resolução com infusão direta

Os metabólitos das amostras de plasma sanguíneo foram extraídos utilizando uma metodologia simples de precipitação de proteínas, utilizando 50 µL de plasma sanguíneo e 300 µL de metanol, sendo em seguida submetidas à agitação em vórtex por 30 segundos, centrifugadas a 5 minutos a 11200 g e 4°C e filtradas em filtros PES de 0,22 µm para a remoção de possíveis materiais particulados nas amostras. Em seguida, as amostras filtradas foram transferidas para *vials* de vidro para posterior análise. Amostras de controle de qualidade (QC) foram preparadas e analisadas a fim de se observar possíveis variações instrumentais ao longo das análises. Os QC consistem em um *pool* de todas as amostras, preparadas utilizando 5 µL de cada uma das amostras de plasma. Três brancos foram preparados e analisados antes das amostras para a

avaliação e posterior remoção de possíveis picos de fundo na etapa de pré-processamento. O preparo das amostras de QC e brancos foram realizadas utilizando a mesma metodologia das amostras teste.

Após o preparo das amostras, estas foram analisadas utilizando um espectrômetro de massas Orbitrap QExactive (Thermo Fisher, Bremen, Alemanha) em sistema de infusão direta, onde injetou-se 20 µL de cada amostra, utilizando um solvente de injeção composto por 50 % (v/v) metanol e 50 % (v/v) água com 0,1% (m/v) de ácido fórmico como aditivo para favorecer a ionização dos compostos. O tempo total da análise de cada amostra foi de 3 minutos, para se certificar que não houvesse nenhuma coeluição nas amostras subsequentes. As amostras foram analisadas em modo ESI (+), com uma faixa de massas de 100-1500 Da, utilizando uma resolução de 140000 em modo *full scan*. Três brancos e 5 amostras QC foram analisados antes da injeção das amostras teste, que foram injetadas de forma randomizada. Durante a corrida, amostras QC foram injetadas uma vez a cada 5 amostras teste e 5 vezes no final da corrida, para avaliar se houveram possíveis variações instrumentais ao longo das análises.

### **3.3. Pré-tratamento e aplicação de metodologias de aprendizado de máquina**

Os dados adquiridos foram exportados em formato .raw e analisados utilizando o programa MS-DIAL (versão 4.70) (Tsugawa et al., 2015) para o pré-processamento, identificação dos sinais dos íons presentes, remoção de sinais de fundo e alinhamento. Os parâmetros utilizados na etapa de pré-processamento são apresentados na **Tabela 2.2**.

**Tabela 2.2.** Parâmetros utilizados no pré-processamento dos dados no software MS-DIAL

<b>Parâmetro</b>	<b>Valor</b>
<b>Projeto</b>	
<i>MS1 Data type</i>	Profile
<i>MS2 Data type</i>	Profile
<i>Ion Mode</i>	Positive
<b>Parâmetros de coleta de dados</b>	
<i>Retention time begin</i>	0
<i>Retention time end</i>	5
<i>Mass range begin</i>	100
<i>Mass range end</i>	1500
<b>Parâmetros de detecção de picos</b>	
<i>Smoothing method</i>	LinearWeightedMovingAverage
<i>Smoothing level</i>	3
<i>Minimum peak height</i>	1000000
<i>Mass slice width</i>	0,05
<i>Minimum peak width</i>	5
<b>Parâmetros de Deconvolução</b>	
<i>Sigma window value</i>	0,5
<i>Amplitude cut off</i>	10
<i>Exclude after precursor</i>	True
<i>Keep isotope until</i>	0,5
<i>Keep original precursor isotopes</i>	False
<b>Parâmetros de alinhamento</b>	
<i>Retention time tolerance</i>	0,05
<i>MS1 tolerance</i>	0,005
<i>Retention time factor</i>	0,5
<i>MS1 factor</i>	0,5
<i>Gap filling by compulsion</i>	True
<b>Ajustes de filtragem</b>	
<i>Peak count filter</i>	60
<i>Remove feature based on peak height fold-change</i>	True
<i>Sample max / blank average</i>	5
<i>Sample average / blank average</i>	5
<i>Keep removable features and assign the tag for checking</i>	True

Após a etapa de pré-processamento, obteve-se um conjunto de dados com 484 valores de relações massa/carga ( $m/z$ ). Em seguida, separou-se os dados em conjuntos de treino e teste, onde obteve-se 124 exemplos para treino e 32 exemplos para o teste final dos modelos.

Para determinar as  $m/z$  mais importantes para a diferenciação dos grupos, utilizou-se duas metodologias de aprendizado de máquina distintas, baseadas em regressão LASSO (Tibshirani, 1996) e no algoritmo boruta (Kursa & Rudnicki, 2010). Nos dois métodos, transformamos o problema de classificação multi-classe em 3 problemas binários de um grupo *versus* o resto.

Após a aplicação das metodologias, obteve-se 10 variáveis mais relevantes para a diferenciação entre os grupos na metodologia LASSO e todas as variáveis foram consideradas relevantes na metodologia boruta.

### 3.4. Identificação putativa dos metabólitos selecionados

Após a seleção das variáveis mais importantes na diferenciação dos grupos através das metodologias de aprendizado de máquina, utilizou-se a plataforma CEU Mass Mediator (Gil-De-La-Fuente et al., 2019) para a anotação putativa das  $m/z$  frente a diferentes bancos de dados, utilizando uma tolerância de 10 ppm e os adutos mais comumente encontrados em LC-MS ( $[M+H]^+$ ,  $[M+2H]^{2+}$ ,  $[M+Na]^+$ ,  $[M+K]^+$  e  $[M+H-H_2O]^+$ ). Para determinar as identidades dos metabólitos selecionados, optou-se por realizar um experimento de MS/MS adicional utilizando as amostras de QC, representativas do conjunto total de amostras, utilizando as  $m/z$  consideradas diferenciais. Para tal, optou-se por um experimento de PRM, utilizando um espectrômetro de massas Orbitrap QExactive (Thermo Fisher, Bremen, Alemanha) e as mesmas condições de infusão direta descritas na seção 3.2, alterando-se as condições do espectrômetro de massas para a realização do experimento de PRM. As massas consideradas diferenciais foram selecionadas em duas listas de inclusão (cada uma com 18 *features* selecionados), sendo em seguida fragmentadas em uma célula de colisão com diferentes energias de ionização (10, 20 e 30 eV), e posteriormente analisadas no Orbitrap para a determinação das  $m/z$  dos fragmentos em alta resolução. Os dados foram adquiridos em modo *profile*, com uma resolução de 17000 e uma janela de isolamento de 0,4  $m/z$ , com um *offset* de  $m/z$  0,3, população de íons (AGC) de  $1 \times 10^5$  e tempo máximo de injeção dos íons de 50 ms, analisando os íons na janela de 0,08 – 0,3 minutos de injeção da amostra, o que levou a cerca de 8 experimentos de MS/MS por  $m/z$  selecionada.

Após a aquisição dos espectros de MS/MS, eles foram transformados em formato mgf utilizando o software Proteowizard 3.0.21175 (Chambers et al., 2012), sendo analisados posteriormente na plataforma SIRIUS 4.0 (Dührkop et al., 2019) para a confirmação de suas identidades. Utilizou-se os bancos de dados HMDB (Wishart et al., 2018) e KEGG (Ogata et al., 1999), bem como o recurso da própria plataforma que imputa as identidades moleculares baseado nas  $m/z$  dos fragmentos, dando assim maior confiabilidade nas identificações obtidas.

## 4. Resultados e discussão

### 4.1. Metodologias de aprendizado de máquina na diferenciação de pacientes com BD e SZ em comparação com controles saudáveis

Após a utilização de duas metodologias de aprendizado de máquina para a diferenciação dos grupos, diferentes resultados foram obtidos para cada um dos métodos. Optou-se por segregar os grupos em um esquema de um *versus* o resto, onde calculou-se a acurácia e o *score* F1, que é a capacidade de predição do método. Os valores de acurácia e de *score* F1 são métricas utilizadas para avaliar a performance de um modelo de aprendizado de máquina, representados nas **Equações 3 e 4**, e são baseados nos números de amostras classificadas como verdadeiros positivos (VP), verdadeiros negativos (VN), falsos positivos (FP) e falsos negativos (FN).

$$Acc = \frac{\# \text{ amostras classificadas corretamente}}{\# \text{ todas as amostras}} = \frac{VP+VN}{VP+FP+VN+FN} \quad (3)$$

$$F_1 = 2 \frac{P \cdot R}{P+R} \quad (4)$$

Os valores de precisão (P) e revocação (*recall*) (R) são dados pelas **Equações 5 e 6**, e representam qual a proporção dos dados classificados como verdadeiros positivos (quais elementos selecionados são relevantes) e qual a proporção classificada como positiva dentro do conjunto (quantos elementos relevantes foram selecionados) (Hicks *et al.*, 2022).

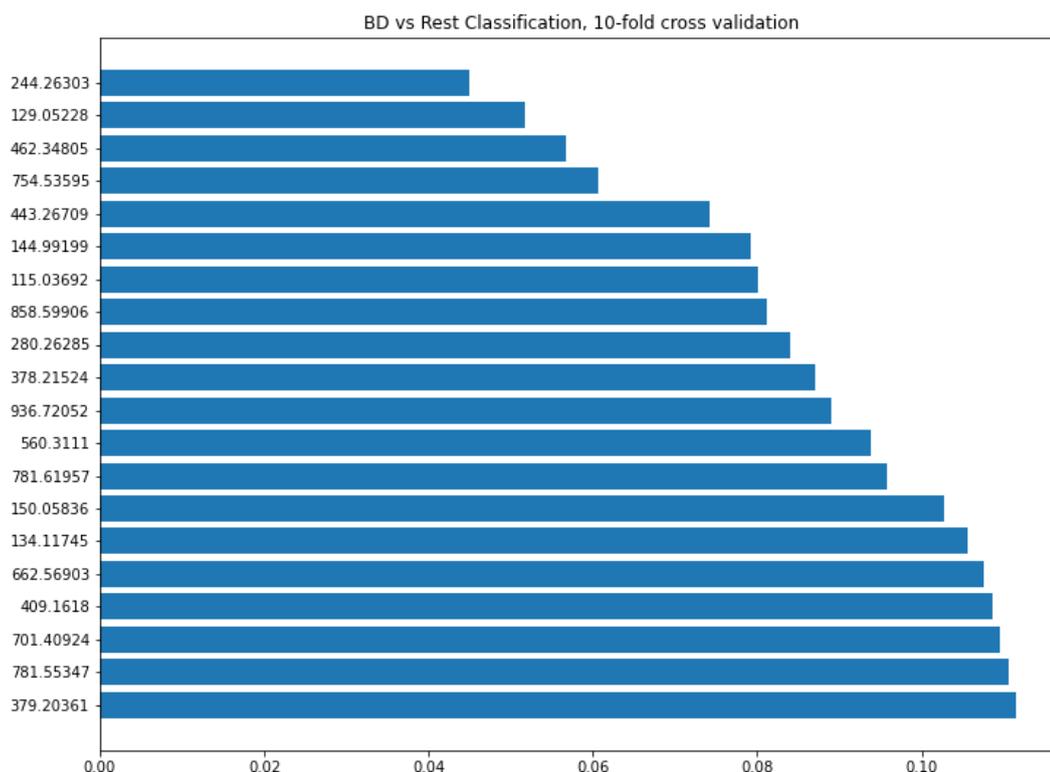
$$P = \frac{\# \text{ predições positivas corretas}}{\# \text{ amostras positivas}} = \frac{VP}{VP+FP} \quad (5)$$

$$R = \frac{\# \text{ amostras VP}}{\# \text{ amostras positivas}} = \frac{VP}{VP+FN} \quad (6)$$

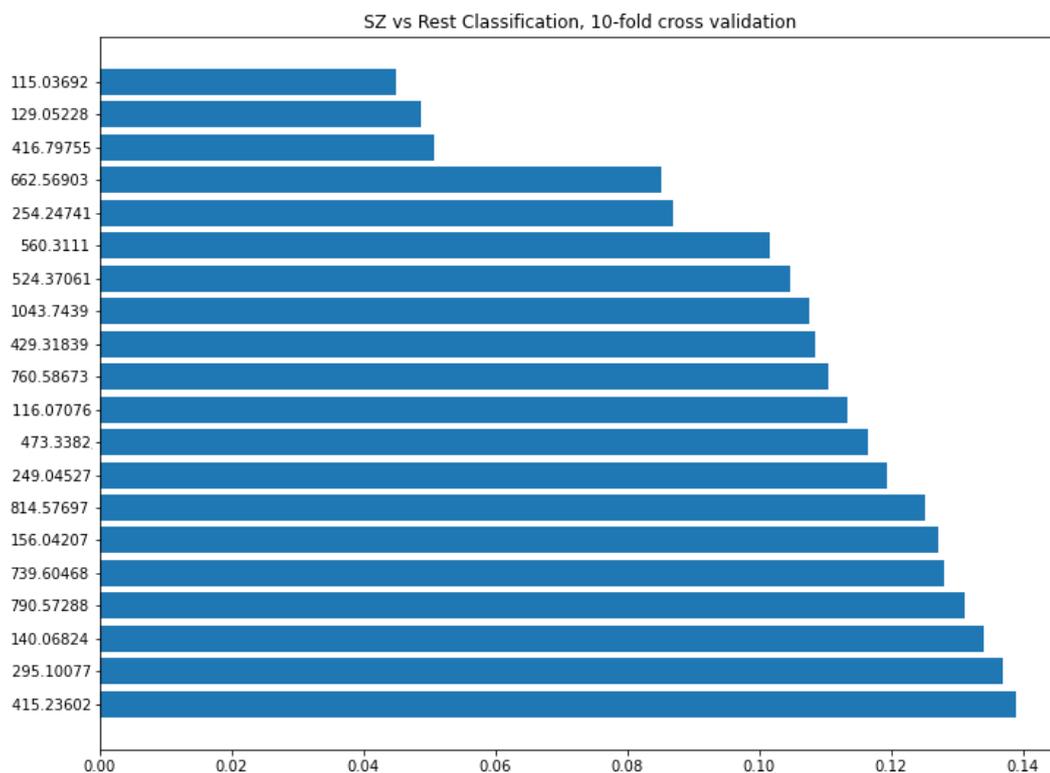
Utilizando a regressão LASSO, obteve-se diferentes resultados de predição para cada um dos casos, após validação cruzada:

- Precisão e score F1 para BD vs. Resto, 10-fold CV:  
Acc:  $0,79 \pm 0,07$   
F1-score:  $0,72 \pm 0,09$
- Precisão e score F1 para SZ vs. Resto, 10-fold CV:  
Acc:  $0,87 \pm 0,07$   
F1-score:  $0,8 \pm 0,1$
- Precisão e score F1 para HC vs. Resto, 10-fold CV:  
Acc:  $0,98 \pm 0,05$   
F1-score:  $0,98 \pm 0,05$

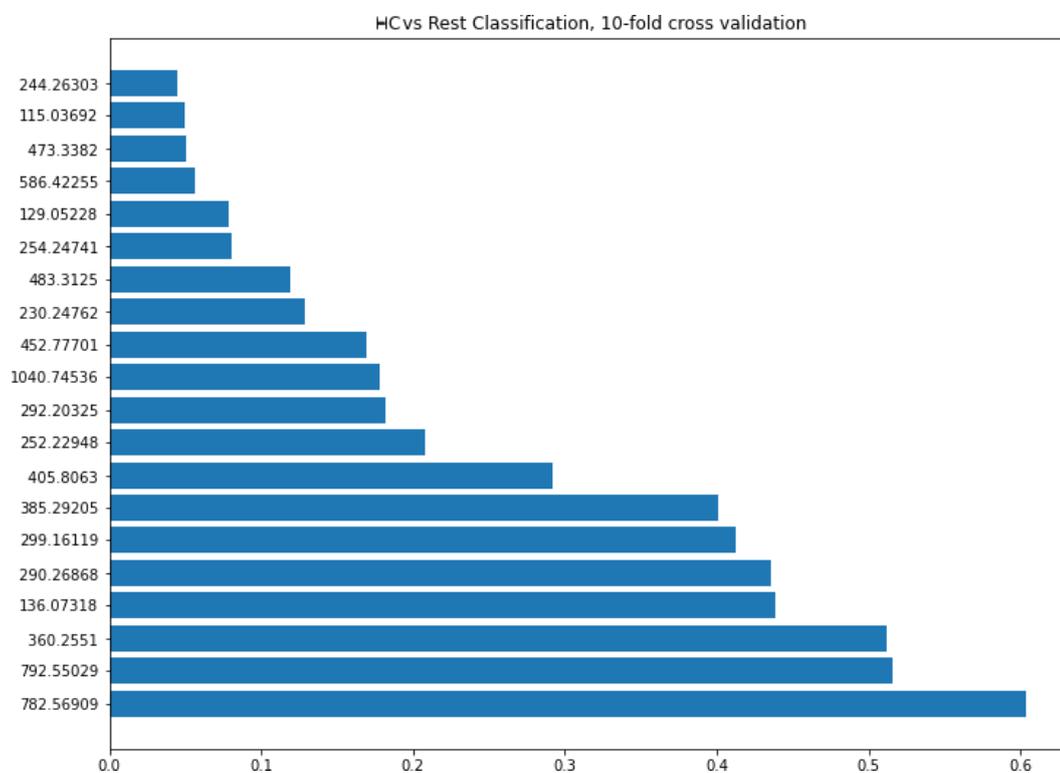
Utilizando a regressão LASSO, escolheu-se os 10 *features* com a melhor capacidade de predição dentro do modelo, onde é possível observar a ordem de importância de cada um deles dentro dos modelos de regressão nas **Figuras 2.7, 2.8 e 2.9.**



**Figura 2.7.** Ordem de importância dos *features* selecionados na regressão LASSO para a classificação do grupo BD versus o resto dos grupos.



**Figura 2.8.** Ordem de importância dos *features* selecionados na regressão LASSO para a classificação do grupo SZ versus o resto dos grupos.



**Figura 2.9.** Ordem de importância dos *features* selecionados na regressão LASSO para a classificação do grupo HC versus o resto dos grupos.

A LASSO é uma regressão logística com uma penalização  $L^1$ . O objetivo do modelo é minimizar o erro que é composto de um erro da classificação, onde utilizou-se a entropia cruzada mais a penalização apresentada na **Eq. 7**:

$$\lambda \|\beta\|_1 = \lambda \left( \sum_i |\beta_i| \right) \quad (7)$$

onde  $\beta = (\beta_1, \beta_2, \dots, \beta_n)$  são os coeficientes da regressão linear e  $\lambda$  é um hiperparâmetro de penalização dos coeficientes. Se o valor de  $\lambda$  é grande, a solução do modelo é dada por  $\beta = 0$  e, então, temos o modelo constante. Se  $\lambda$  é zero, então a solução do modelo é a de uma regressão logística sem penalização, com todos os coeficientes não nulos. Assim, variando  $\lambda$  de zero até um valor onde todos os coeficientes são nulos, os coeficientes que não se anulam primeiro são considerados os mais importantes. Isso determina uma ordem de importância das  $m/z$  para cada classe versus o resto (Tibshirani, 1996).

Na metodologia boruta, não há uma ordem de importância de cada *feature* individualizado e sim de um conjunto de variáveis que formam um painel com a maior capacidade preditiva de classificação, utilizando o mesmo modelo da metodologia LASSO, de um grupo versus o resto. Na validação cruzada 10-*fold* da floresta aleatória, obteve-se um *score* F1 de  $0,6 \pm 0,1$ .

O algoritmo boruta é um método baseado em florestas aleatórias que tem como objetivo não apenas selecionar as variáveis mais importantes, mas encontrar todas as variáveis que são relevantes ao problema. Uma cópia de cada coluna dos dados embaralhando seus valores é feita, chamadas de variáveis sombra. Em seguida, o modelo é treinado juntando as variáveis originais com as variáveis sombras e a importância de cada uma delas é calculada por um método padrão de florestas aleatórias, no caso, o método Gini de importância. As variáveis originais que possuem importância maior que todas as variáveis sombra são consideradas relevantes, enquanto que as que apresentam importância significativamente menor são consideradas não relevantes, sendo

removidas em seguida. Todas as variáveis sombra são removidas e o processo é repetido até um limite determinado ou até que tenhamos associado importância a todas as variáveis (Kursa & Rudnicki, 2010).

#### 4.2. Identificação putativa dos metabólitos selecionados

Após a seleção das  $m/z$  mais importantes para a construção dos modelos de classificação, estas foram introduzidas na plataforma CEU Mass Mediator (Gil-De-La-Fuente et al., 2019), buscando assim quais são as moléculas referentes a cada  $m/z$  dentro de diversos bancos de dados. As identificações preliminares de cada  $m/z$ , seus adutos, erro de massas e fórmula molecular são apresentadas na **Tabela 2.3**. As identificações dos *features* selecionados somente na metodologia LASSO são apresentados na tabela de cor branca, enquanto os que se mostraram relevantes em ambas as metodologias são apresentados na tabela de cor cinza e os que se mostraram relevantes somente na metodologia boruta na cor azul. Os *features* que possuíram identificação positiva pelo espectro de MS/MS foram considerados como nível MSI 2, enquanto os que possuíram inconsistências e/ou não possuíram informações suficientes para a determinação de suas identidades moleculares foram considerados como nível MSI 3, seguindo os protocolos do *Metabolomics Standard Initiative* (Sumner et al., 2007). Os metabólitos identificados como lipídios foram classificados por sua classe e pela soma dos átomos de carbono presentes em sua estrutura e pelo número de duplas ligações, por exemplo, PC (18:1) corresponde a uma fosfatidilcolina com 18 carbonos somados em sua estrutura, com a presença de uma dupla ligação. Os lipídios cujas identidades foram confirmadas por MS/MS foram considerados como nível MSI 2, por ser possível a identificação de sua classe, apesar de não ser possível determinar a composição completa de sua estrutura (Liebisch et al., 2013, 2020). Os espectros de fragmentação dos metabólitos com identificação parcial, mas considerados nível MSI 3, e dos metabólitos considerados nível MSI 2 identificados pela plataforma SIRIUS estão disponíveis nos **Anexos**.

**Tabela 2.3.** Identificações preliminares dos *molecular features* considerados diferenciais nas metodologias LASSO, em ambas as metodologias, e na metodologia boruta, respectivamente

BD x Resto	SZ x Resto	HC x Resto	m/z	Aduto	Erro (ppm)	Putative ID	Fórmula molecular	Nível MSI
			230,2476	[M+H] <sup>+</sup>	1	SPB 14:0,O	C <sub>14</sub> H <sub>31</sub> NO	3
			280,2629	[M+H] <sup>+</sup>	2	Linoleamide	C <sub>18</sub> H <sub>33</sub> NO	3*
				[M+H-H <sub>2</sub> O] <sup>+</sup>	2	Palmitoleoyl-EA	C <sub>18</sub> H <sub>35</sub> NO <sub>2</sub>	
			378,2152	[M+2H] <sup>2+</sup>	3	Schidigerasaponin C2	C <sub>39</sub> H <sub>62</sub> O <sub>14</sub>	3
				[M+2H] <sup>2+</sup>	3	Agavoside B	C <sub>39</sub> H <sub>62</sub> O <sub>14</sub>	
				[M+2H] <sup>+</sup>	3	Alliospiroside D	C <sub>39</sub> H <sub>62</sub> O <sub>14</sub>	
			429,3184	[M+H] <sup>+</sup>	6	Sorbitan oleate	C <sub>24</sub> H <sub>44</sub> O <sub>6</sub>	3
				[M+H] <sup>+</sup>	6	Glyceril lactooleate	C <sub>24</sub> H <sub>44</sub> O <sub>6</sub>	
				[M+2H] <sup>2+</sup>	3	PG(42:3)	C <sub>48</sub> H <sub>89</sub> O <sub>10</sub> P	
				[M+H-H <sub>2</sub> O] <sup>+</sup>	0	N8,N'8-Citryl-bis(spermidine)	C <sub>20</sub> H <sub>42</sub> N <sub>6</sub> O <sub>5</sub>	
			443,2671	[M+H] <sup>+</sup>	7	Prostaglandin G2 2-glyceril Ester	C <sub>23</sub> H <sub>38</sub> O <sub>8</sub>	3
			452,7770			NI		
			483,3125	[M+H] <sup>+</sup>	3	26-hydroxycholesterol 3-sulfate	C <sub>27</sub> H <sub>46</sub> O <sub>5</sub> S	3
				[M+2H] <sup>2+</sup>	7	PI(44:7)	C <sub>53</sub> H <sub>89</sub> O <sub>13</sub> P	
				[M+H-H <sub>2</sub> O] <sup>+</sup>	4	Ganolucidic acid A	C <sub>30</sub> H <sub>44</sub> O <sub>6</sub>	
				[M+H-H <sub>2</sub> O] <sup>+</sup>	4	Ganolucidic acid D	C <sub>30</sub> H <sub>44</sub> O <sub>6</sub>	
				[M+H-H <sub>2</sub> O] <sup>+</sup>	4	Physalolactone B	C <sub>30</sub> H <sub>44</sub> O <sub>6</sub>	
				[M+H-H <sub>2</sub> O] <sup>+</sup>	4	Ganoderic acid beta	C <sub>30</sub> H <sub>44</sub> O <sub>6</sub>	
				[M+H-H <sub>2</sub> O] <sup>+</sup>	4	11-Deoxocucurbitacin I	C <sub>30</sub> H <sub>44</sub> O <sub>6</sub>	
				[M+H-H <sub>2</sub> O] <sup>+</sup>	4	Hippuristanolide	C <sub>30</sub> H <sub>44</sub> O <sub>6</sub>	
			524,3706	[M+H] <sup>+</sup>	1	PC(18:0)	C <sub>26</sub> H <sub>54</sub> NO <sub>7</sub> P	2
			560,3111	[M+K] <sup>+</sup>	0	PC(18:1)	C <sub>26</sub> H <sub>52</sub> NO <sub>7</sub> P	2
			754,5359	[M+Na] <sup>+</sup>	0	PC(32:1)	C <sub>40</sub> H <sub>78</sub> NO <sub>8</sub> P	2
			760,5867	[M+H] <sup>+</sup>	2	PC(34:1)	C <sub>42</sub> H <sub>82</sub> NO <sub>8</sub> P	2
			858,5991	[M+Na] <sup>+</sup>	1	PC(40:5)	C <sub>48</sub> H <sub>86</sub> NO <sub>8</sub> P	2

			1040,7450			NI		
			1043,7439	[M+K] <sup>+</sup>	3	TG(64:13)	C <sub>67</sub> H <sub>104</sub> O <sub>6</sub>	3
			115,0369	[M+Na] <sup>+</sup>	4	Glycerol	C <sub>3</sub> H <sub>8</sub> O <sub>3</sub>	3
			129,0523	[M+Na] <sup>+</sup>	1	(2R*,3R*)-1,2,3-Butanetriol	C <sub>4</sub> H <sub>10</sub> O <sub>3</sub>	3
				[M+Na] <sup>+</sup>	1	Trihydroxybutane	C <sub>4</sub> H <sub>10</sub> O <sub>3</sub>	
				[M+Na] <sup>+</sup>	1	Diethylene glycol	C <sub>4</sub> H <sub>10</sub> O <sub>3</sub>	
			144,9920	[M+Na] <sup>+</sup>	8	3-Mercaptolactate	C <sub>3</sub> H <sub>6</sub> O <sub>3</sub> S	3
				[M+Na] <sup>+</sup>	8	1,3-Propane sultone	C <sub>3</sub> H <sub>6</sub> O <sub>3</sub> S	
				[M+Na] <sup>+</sup>	8	3-Mercaptolactic acid	C <sub>3</sub> H <sub>6</sub> O <sub>3</sub> S	
			244,2630			NI		
			254,2474	[M+H] <sup>+</sup>	2	Palmitoleamide	C <sub>16</sub> H <sub>31</sub> NO	3*
				[M+H-H <sub>2</sub> O] <sup>+</sup>	2	Myristoyl-EA	C <sub>16</sub> H <sub>33</sub> NO <sub>2</sub>	
			416,7976	[M+2H] <sup>2+</sup>	3	PC(40:7)	C <sub>48</sub> H <sub>82</sub> NO <sub>8</sub> P	3
				[M+2H] <sup>2+</sup>	3	PE-NMe(42:7)	C <sub>48</sub> H <sub>82</sub> NO <sub>8</sub> P	
			462,3481	[M+2H] <sup>2+</sup>	6	PI(O-41:0)	C <sub>50</sub> H <sub>99</sub> O <sub>12</sub> P	3
			473,3382	[M+H] <sup>+</sup>	7	Methyl (7Z,9Z,9'Z)-6'-apo-γ-caroten-6'-oate	C <sub>33</sub> H <sub>44</sub> O <sub>2</sub>	3
				[M+Na] <sup>+</sup>	2	Vitamin K1	C <sub>31</sub> H <sub>46</sub> O <sub>2</sub>	
				[M+K] <sup>+</sup>	2	ST 28:0;O3	C <sub>28</sub> H <sub>50</sub> O <sub>3</sub>	
			586,4226			NI		
			662,5690			NI		
			162,0501	[M+2H] <sup>2+</sup>	5	Vulpinic Acid	C <sub>19</sub> H <sub>14</sub> O <sub>5</sub>	3
				[M+2H] <sup>2+</sup>	5	4-(3,4-Dihydroxyphenyl)-2,3-dihydro-2,3-dihydroxy-1H-phenalen-1-one	C <sub>19</sub> H <sub>14</sub> O <sub>5</sub>	
			226,9513	[M+H] <sup>+</sup>	5	Ethyl 3-iodo-2E-acrylate	C <sub>5</sub> H <sub>7</sub> IO <sub>2</sub>	3
				[M+H] <sup>+</sup>	2	2,5-Dichloro-4-oxohex-2-enedioate	C <sub>6</sub> H <sub>4</sub> Cl <sub>2</sub> O <sub>5</sub>	
				[M+K] <sup>+</sup>	2	2-Chloro-5-methyl-cis-dienelactone	C <sub>7</sub> H <sub>5</sub> ClO <sub>4</sub>	
				[M+K] <sup>+</sup>	2	3-Chloro-2-methyldienelactone	C <sub>7</sub> H <sub>5</sub> ClO <sub>4</sub>	
				[M+H-H <sub>2</sub> O] <sup>+</sup>	5	Thicyofen	C <sub>8</sub> H <sub>5</sub> CIN <sub>2</sub> OS <sub>2</sub>	
			228,2324	[M+H] <sup>+</sup>	1	Myristamide	C <sub>14</sub> H <sub>29</sub> NO	2
			242,9252			NI		

			250,2140	[M+Na] <sup>+</sup>	1	Myristamide	C <sub>14</sub> H <sub>29</sub> NO	2
			260,2582	[M+H] <sup>+</sup>	1	SPB 15:0,O2	C <sub>15</sub> H <sub>33</sub> O <sub>2</sub>	3
			336,2890	[M+Na] <sup>+</sup>	5	Margaroyl-EA	C <sub>19</sub> H <sub>39</sub> NO <sub>2</sub>	2
			338,3401	[M+H] <sup>+</sup>	5	Erucamide	C <sub>22</sub> H <sub>43</sub> NO	2
			348,2658	[M+2H] <sup>2+</sup>	3	PG(O-31:0)	C <sub>37</sub> H <sub>75</sub> O <sub>9</sub> P	3
				[M+2H] <sup>2+</sup>	3	Mannosyl-1beta-phosphomycoketide C31	C <sub>37</sub> H <sub>75</sub> O <sub>9</sub> P	
				[M+K] <sup>+</sup>	2	Oleoyl Ethyl Amide	C <sub>20</sub> H <sub>39</sub> NO	
			360,2551	[M+H-H <sub>2</sub> O] <sup>+</sup>	4	15-HETE-Gly	C <sub>22</sub> H <sub>35</sub> NO <sub>4</sub>	3
				[M+H-H <sub>2</sub> O] <sup>+</sup>	4	12-HETE-Gly	C <sub>22</sub> H <sub>35</sub> NO <sub>4</sub>	
				[M+H-H <sub>2</sub> O] <sup>+</sup>	4	N-margaroyl taurine	C <sub>19</sub> H <sub>39</sub> NO <sub>4</sub> S	
			364,2721	[M+2H] <sup>2+</sup>	2	1-O-beta-D-Glucopyranosyl-2,3-di-O-(8-hexadecenoyl)glycerol	C <sub>41</sub> H <sub>74</sub> O <sub>10</sub>	3
			376,2973	[M+2H] <sup>2+</sup>	3	PG(O-35:0)	C <sub>41</sub> H <sub>83</sub> O <sub>9</sub> P	3
				[M+K] <sup>+</sup>	1	N-Cyclohexanecarbonylpentadecylamine	C <sub>22</sub> H <sub>43</sub> NO	
				[M+K] <sup>+</sup>	1	Erucamide	C <sub>22</sub> H <sub>43</sub> NO	
				[M+H-H <sub>2</sub> O] <sup>+</sup>	3	N-palmitoyl histidine	C <sub>22</sub> H <sub>39</sub> N <sub>3</sub> O <sub>3</sub>	

### Legenda

	Presença do <i>feature</i> na metodologia LASSO, modelo BD x Resto
	Presença do <i>feature</i> na metodologia boruta, modelo BD x Resto
	Presença do <i>feature</i> na metodologia LASSO, modelo SZ x Resto
	Presença do <i>feature</i> na metodologia boruta, modelo SZ x Resto
	Presença do <i>feature</i> na metodologia LASSO, modelo HC x Resto
	Presença do <i>feature</i> na metodologia boruta, modelo HC x Resto

NI: Não identificado para os critérios selecionados

\*: Identificação parcial, duas moléculas com fragmentação similar e com scores significativos na identificação

Após os experimentos de MS/MS, foi possível identificar algumas das identidades dos compostos presentes utilizando a plataforma SIRIUS (Dührkop et al., 2019). Os espectros de MS/MS contendo as identidades dos compostos determinados como nível MSI 2, bem como a estrutura proposta são apresentadas no **Anexo 1**.

Dentre as  $m/z$  selecionadas na metodologia LASSO, foi possível observar a presença de algumas fosfatidilcolinas ( $m/z$  524,3706; 560,3111; 754,5359; 760,5867 e 858,5991), que foram identificadas devido à presença do íon 184,07 no espectro de fragmentação, o que corresponde à parte polar da molécula de fosfatidilcolina em modo positivo de ionização. Além das fosfatidilcolinas previamente citadas, observou-se um espectro de fragmentação consistente com a erucamida ( $m/z$  338,3104), e diversas outras N-acilamidas e amidas graxas apresentaram fragmentação consistente com os bancos de dados. No caso da miristamida ( $m/z$  228,2324 para o aduto  $[M+H]^+$  e 250,2140 para o aduto  $[M+Na]^+$ ) a fragmentação mais consistente foi considerada para efeitos de identificação. Pela plataforma SIRIUS foi possível observar outras amidas graxas com a mesma fórmula molecular que a miristamida, entretanto, esta foi a única que estava presente na busca preliminar na plataforma CEU Mass Mediator e que era considerada uma amida graxa endógena relacionada aos seres humanos. Foi observada também a presença conflitante na fragmentação de N-acilamidas e amidas graxas em duas  $m/z$ , 280,2629 e 254,2474. Pelo fato da fragmentação destas amidas serem extremamente parecidas, optou-se por manter os seus níveis de identificação como nível MSI 3, devido à falta de clareza na determinação da identificação. Apesar disso, por todos os casos se tratarem de classes similares de moléculas, pode-se dizer com certa confiabilidade que as  $m/z$  em questão se tratam desta classe de moléculas, logo, estas serão discutidas a seguir.

Apesar de todas as  $m/z$  determinadas nas metodologias de aprendizado de máquina terem sido analisadas no experimento de MS/MS, não foi possível encontrar informações suficientes nos espectros de fragmentação para precisar a identidade de outros metabólitos além do nível MSI 3. Isso se dá possivelmente à metodologia analítica utilizada para as análises, que foram feitas em modo de infusão direta, sem o uso de técnicas de separação acopladas ao espectrômetro de massas, o que leva a formação de espectros “quiméricos”, ou seja, a combinação da fragmentação de duas ou mais moléculas em um único espectro de fragmentação. Visto que a janela de isolamento do quadrupolo é de 1 Da, existe a possibilidade de diferentes moléculas

com a  $m/z$  dentro desta janela serem selecionadas pelo método em questão. No caso do experimento de MS/MS nesse estudo, utilizou-se uma janela de isolamento de 0,4 Da com um *offset* de + 0,3 Da, o que abre a possibilidade de íons com  $m/z$  entre - 0,4 Da e + 0,7 Da em relação ao valor inserido na lista de inclusão serem selecionados para a fragmentação.

### 4.3. N-acilamidas, amidas graxas e sua relação com transtornos psiquiátricos

Dentre os resultados obtidos, observou-se a presença de N-acilamidas e amidas graxas nas identificações putativas em todos os três modelos. No modelo BD *versus* o resto, a  $m/z$  280,2629 apresentou similaridades com a linoleiamida e a palmitoiletanolamida, enquanto a  $m/z$  336,2890 apresentou similaridades com a margaroiletanolamida, além da identificação putativa da prostaglandina G2 2-gliceril éster ( $m/z$  443,2671). No modelo SZ *versus* o resto, a  $m/z$  338,3401 foi identificada como a erucamida (cis-13-docosenamida). Além da erucamida, observou-se nas identificações putativas com nível MSI 3 a possível presença da oleoiletanolamida ( $m/z$  348,2658) e da 12-HETE-Gly, ou 15-HETE-Gly ou N-margaroiltaurina ( $m/z$  360,2551). Observou-se também a possível presença da palmitoilamida ou miristoiletanolamida ( $m/z$  254,2474) nos modelos SZ *versus* o resto e HC *versus* o resto. No grupo HC *versus* o resto obteve-se também a identificação putativa da miristamida no espectro de MS/MS, além de diversas fosfatidilcolinas identificadas nos grupos BD e SZ *versus* o resto.

Dentre os resultados obtidos, o que mais chamou a atenção foi a presença marcante de diversos compostos bioativos relacionados ao sistema endocanabinóide, ao sistema nervoso central e ao sistema de sinalização celular. As diversas amidas encontradas nas identificações putativas se reúnem dentro de um grande grupo classificatório (Liebisch et al., 2013, 2020) chamado de amidas graxas, na qual estão dentro deste grupo as N-acilamidas (na qual se incluem os N-acilaminoácidos), N-aciletanolaminas e amidas graxas primárias. Dentre as supracitadas, as N-aciletanolaminas são a classe de moléculas mais estudadas e conhecidas, dado as suas conexões com o sistema endocanabinóide. O sistema endocanabinóide consiste em metabólitos endógenos com atividade endocanabinóide, receptores e enzimas responsáveis pela síntese e degradação destes metabólitos. Estes lipídios

conseguem se ligar a receptores específicos dentro do cérebro, causando alterações comportamentais similares ao causado pelos componentes psicoativos da cannabis, em especial o tetraidrocanabinol (THC) (Lu & MacKie, 2016). O sistema endocanabinóide apresenta importantes interações em diversas vias fisiológicas e patológicas dentro do sistema nervoso central e em funções de sinalização relacionadas à alimentação, sistema imune e inflamatório, metabolismo de lipídios e glicose, metabolismo do fígado, metabolismos muscular e de células adiposas, bem como a comunicação entre o trato intestinal e o sistema nervoso central (Lu & MacKie, 2016; Moreno et al., 2021; Röhrig et al., 2019). As amidas graxas primárias, por sua vez, são lipídios bioativos de sinalização, que apresentam funções no controle do sono, angiogênese, relações com receptores de serotonina, receptores endocanabinóides, receptores GABA, transportadores de glicina e sinalização de  $Ca^{2+}$  (Divito & Cascio, 2013; Ezzili et al., 2010; Farrell & Merkler, 2008), enquanto N-acilaminoácidos apresentam diversas interações biológicas, especialmente em reações de sinalização, incluindo neuroproteção e neurogênese, inibição de dor, homeostase e modulação de diversas reações biológicas. Além disso, discute-se a sua relação direta com eicosanoides e receptores relacionados ao sistema endocanabinóide e ao sistema nervoso, como por exemplo os receptores GPR18 e GPR92 (Arul Prakash & Kamlekar, 2021; Burstein, 2018; Farrell & Merkler, 2008; Hanuš et al., 2014).

A linoleiamida, encontrada no grupo BD versus o resto, é um lipídio endógeno que constitui os neurônios e pode ser encontrada no cérebro, possuindo relações com o fluxo de  $Ca^{2+}$  (Huang & Jan, 2001) e alterações no fluxo de correntes iônicas em células neuroendócrinas (Liu & Wu, 2003), além de ser conhecido como um indutor de sono (Cravatt et al., 1995).

Observou-se que a palmitoiletanolamida, presente no cérebro, fígado e músculo esquelético, apresenta atividade antialérgica e anti-inflamatória, bem como efeitos neuroprotetivos, agindo no sistema nervoso central e sensorial e em células do sistema imune (Ezzili et al., 2010; Lambert et al., 2012; M. M. Li et al., 2017; Petrosino & Di Marzo, 2009; Saturnino et al., 2017). Há indícios também da palmitoiletanolamida apresentar atividade endocanabinóide, atuando nos receptores PPAR $\alpha$  e GPR55 (Ezzili et al., 2010; Farrell & Merkler, 2008), além de ser reportada como alterada em pacientes esquizofrênicos (Leweke et al., 1999).

A palmitoleiamida, por sua vez, encontrada no grupo SZ versus o resto, possui atividade bloqueadora da junção neuronal, dificultando a conexão entre astrócitos e afetando os receptores serotoninérgicos no cérebro (Boger et al., 1998), além de estar correlacionado também a hipertriglicemia, um fator de risco para doenças cardiovasculares (Lee et al., 2015), enquanto que a miristoiletanolamida foi encontrada aumentada no plasma de pacientes com risco de desenvolver infartos, juntamente com diversas fosfatidilcolinas (Jové et al., 2015), além de apresentar redução na capacidade de migração de células na medula óssea, indicando que o sistema endocanabinoide pode possuir outras funções além do sistema nervoso central (Patinkin et al., 2008).

A erucamida não possui ligações com nenhum receptor do sistema endocanabinoide, apesar de apresentar funções como a criação de vasos sanguíneos e modulação de fluidos (Ezzili et al., 2010; Waluk et al., 2014). Apesar de não possuir nenhuma ligação com receptores endocanabinóides até o momento, a erucamida vem sendo testada como uma alternativa de antidepressivo, proposto no trabalho de Li *et al.* (2017), onde a erucamida foi administrada em modelos animais em comparação a fluoxetina, um tradicional antidepressivo, mostrando que a erucamida pode possuir efeito regulador do eixo hipotálamo-hipófise-adrenal, relacionado à resposta ao estresse.

É reportado na literatura a ação da oleoiletanolamida (OEA) como possuidora de atividade endocanabinoide, ligando-se ao receptor PPAR $\alpha$  e ao receptor GPR<sub>119</sub>, envolvido em diversas vias metabólicas como oxidação de lipídios e lipólise, processos inflamatórios e controle de apetite (Laleh et al., 2018; Ramírez-Orozco et al., 2019). Estudos indicam níveis reduzidos de OEA em pacientes psicóticos em primeiro surto não tratados (Dickens et al., 2020), entretanto, em pacientes tratados esta tendência não se manteve. Em um estudo com pacientes esquizofrênicos tratados, Potvin *et al.* mostraram níveis aumentados de OEA (Potvin et al., 2020) e Leweke *et al.* (2012) indicaram que aumentos nos níveis de OEA juntamente com outros endocanabinóides podem estar relacionados à melhora nos sintomas psicóticos, o que leva a crer que o tratamento padrão para SZ pode estar relacionado a um aumento nos níveis deste endocanabinoide e em uma possível desregulação não somente do sistema nervoso central, mas também do sistema nervoso periférico (Dickens et al., 2020).

Apesar de estarem selecionados como diferenciais em grupos distintos, a prostaglandina G<sub>2</sub> gliceril éster (PGG<sub>2</sub>-G) e o 12 (ou 15)-HETE-Gly estão intimamente ligados dentro da via de metabolismo do ácido araquidônico. Ambos são produtos de dois compostos endocanabinóides (2-araquidonoilglicerol, ou 2-AG e N-araquidonoilglicina, ou NAGly, respectivamente), e são produzidos pela enzima ciclooxigenase-2 (COX-2) (Burstein, 2018; Kozak et al., 2002; Musee & Marnett, 2012; Prusakiewicz et al., 2002). No caso da PGG<sub>2</sub>-G, esta é produzida como um intermediário da PGH<sub>2</sub>-G, que em seguida é transformada em diversos outros produtos, como outras prostaglandinas e tromboxanos. A HETE-Gly, por sua vez, é produzida juntamente com a PGH<sub>2</sub>-Gly pela COX2 tendo como substrato o NAGly (Prusakiewicz et al., 2002).

Em relação à margaroiletanolamida (heptadecanoiletanolamida), esta foi encontrada em estudos quantitativos de N-aciletanolamidas (Chen et al., 2021; McGurk et al., 2021), porém, não se encontrou estudos que indiquem a sua relação direta com qualquer função biológica. Entretanto, N-aciletanolamidas em geral sofrem uma reação de hidrólise pela enzima FAAH, gerando ácidos graxos e etanolamina (McGurk et al., 2021). No caso da margaroiletanolamida, obtém-se o ácido heptadecanóico, que possui ligações com DCVs (Kurotani et al., 2017), diabetes, SMet e processos inflamatórios (Jenkins et al., 2015; Pfeuffer & Jaudszus, 2016; Venn-Watson et al., 2020), bem como correlação entre o transtorno bipolar e hemóstase (Ribeiro et al., 2022), descrita anteriormente no Capítulo 1.

A miristamida, encontrada no grupo BD versus o resto, foi quantificada em estudos metabolômicos reportados na literatura (Castillo-Peinado et al., 2019; Divito et al., 2016), porém, a sua função biológica ainda não é clara. Estudos apontam que a miristamida apresenta certa atividade antimicrobiana (Narasimhan et al., 2006), além de funcionar como um substrato para a enzima FAAH (Ueda et al., 2000), porém, não há nenhuma relação com outras vias biológicas e/ou com o sistema endocanabinóide.

Além das N-acilamidas encontradas, observou-se também a presença de diversas fosfatidilcolinas nos grupos BD e SZ versus o resto. Além de possuírem diversas funções biológicas, as fosfatidilcolinas também são precursoras de compostos com atividade endocanabinóide como o 2-AG e a anandamida, além de

outras N-acilamidas sem atividade endocanabinóide (Mulder & Cravatt, 2006; Tsuboi et al., 2018; Waluk et al., 2014).

Não é surpreendente a relação do sistema endocanabinóide com o BD e a SZ, visto que está intimamente ligado com o sistema nervoso central (Mechoulam & Parker, 2013). Além disso, observou-se também a sua relação com alterações inflamatórias (Lu & MacKie, 2016), o que vai ao encontro com algumas hipóteses sobre a fisiopatologia de doenças psiquiátricas estarem relacionadas à neuroinflamação (Najjar et al., 2013), além da interação com outros sistemas, levantando a possibilidade do BD e a SZ serem transtornos multissistêmicos (Leboyer et al., 2012; Mitchell & Dinan, 2010).

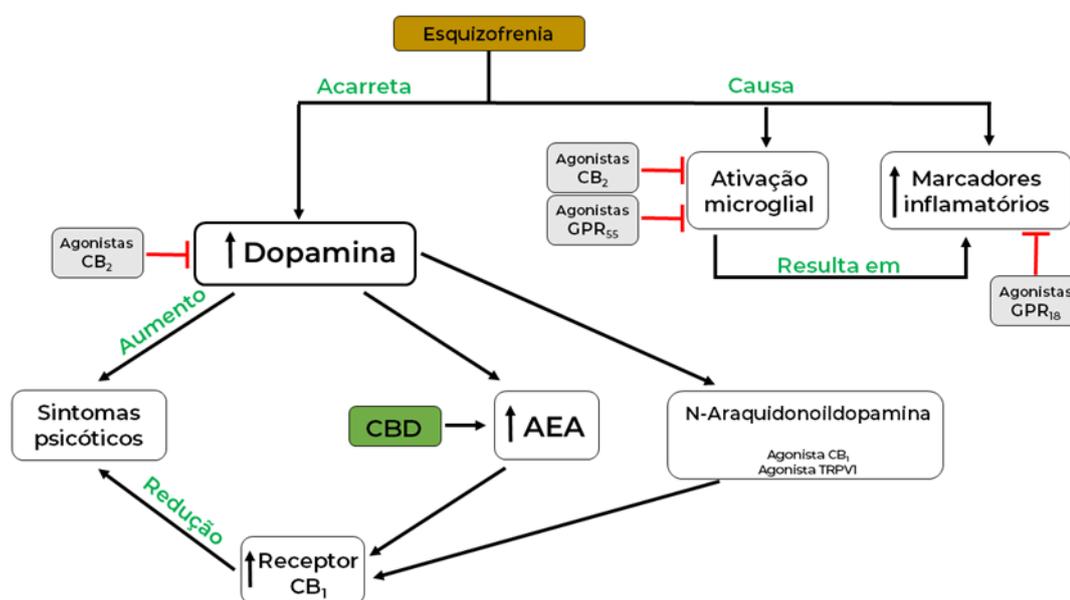
Tratando-se do BD, existem diversas evidências na literatura que correlacionam a doença com alterações no ácido araquidônico (McNamara et al., 2008, 2015; Pomponi et al., 2013; Rapoport, 2014), metabólito essencial para o funcionamento do sistema endocanabinóide, visto que é produto dos dois principais compostos endocanabinóides estudados, a anandamida e o 2-araquidonoilglicerol (2-AG) (Lu & MacKie, 2016). Estes compostos sofrem degradação sob a ação de diferentes enzimas, sendo transformados em ácido araquidônico pela ação da enzima FAAH no caso da anandamida e pela enzima monoacilglicerol lipase (MAGL) no caso do 2-AG (Bisogno & Marzo, 2007), e em prostaglandinas-gliceriléter e prostaglandinas-etanolamidas pela ação da enzima COX-2 (Arjmand et al., 2019). Além disso, a anandamida também pode ser transformada em NAGly pela ação das enzimas álcool desidrogenase e aldeído desidrogenase (Burstein, 2018) e pela conjugação do ácido araquidônico com a glicina. A NAGly em particular, apresenta atividade inibitória da enzima FAAH e também pode agir como um agonista do receptor GPR<sub>18</sub>, cuja ativação está relacionada a efeitos anti-inflamatórios (Fabisiak et al., 2021). É interessante citar também a ação de outros receptores endocanabinóides putativos, tal como o GPR<sub>18</sub> na cascata de ácido araquidônico. Os receptores GPR<sub>55</sub>, GPR<sub>119</sub> e TRPV1 vêm sendo estudados como potenciais receptores do sistema endocanabinóide, dado a sua ativação com compostos derivados da cannabis. O receptor GPR<sub>55</sub> está relacionado à ativação microglial (Fernández-Ruiz et al., 2017), que por sua vez está relacionada à ativação de citocinas inflamatórias (Arjmand et al., 2019), atuando em conjunto com receptores CB<sub>2</sub>. O receptor GPR<sub>119</sub> por sua vez, está relacionado ao controle de apetite, juntamente com



termos de similaridade com o BD, observa-se que a SZ também apresenta indícios do aumento de marcadores inflamatórios (Najjar et al., 2013), com aumento na secreção de citocinas inflamatórias (Trindade et al., 2022) e ativação microglial (de Almeida & Martins-de-Souza, 2018), esta mediada por receptores CB<sub>2</sub> e GPR<sub>55</sub> (de Almeida & Martins-de-Souza, 2018; Fernández-Ruiz et al., 2017), assim como o BD. Portanto, é possível também relacionar o receptor GPR<sub>18</sub> com a atenuação de sintomas inflamatórios (Fabisiak et al., 2021) na SZ, visto que este efeito não é exclusivo do BD, mas do sistema endocanabinóide em geral. Apesar de haver indícios de alterações na regulação do ácido araquidônico na SZ, a relação do AA com a doença está mais relacionada com a resposta à niacina e à alterações na membrana celular, em especial nas enzimas fosfolipase A2 (PLA2) e ciclo-oxigenase (COX), que são anormalmente estimuladas pela niacina em pacientes esquizofrênicos (Yang et al., 2021), efeito este que começou a ser utilizado como um teste clínico para SZ nos anos 2000 (Smesny et al., 2007), mas ainda discutido até os dias de hoje devido a sua alta especificidade para SZ quando comparado com outras doenças psiquiátricas como transtorno depressivo maior ou BD (Messamore, 2018).

Em termos de diferenças com o BD, a SZ apresenta mais correlações dentro da literatura com alterações no sistema dopaminérgico, hipótese esta discutida desde os anos 1960, propondo que as vias metabólicas dopaminérgicas são superativadas em pacientes esquizofrênicos (Seeman, 1987). Além disso, é observado também que antipsicóticos de primeira geração como haloperidol e clopromazina, por exemplo, atuam no sistema dopaminérgico, levando a reduções nos sintomas positivos da esquizofrenia, porém, causando efeitos colaterais como efeitos motores extrapiramidais, hiperprolactinemia e alterações cognitivas (Li et al., 2016). A hipótese da atuação da dopamina na fisiopatologia da SZ e sua interação com o sistema endocanabinóide vem sendo estudado juntamente com ação da anandamida, onde propõe-se que o sistema endocanabinóide age em um processo de *feedback* negativo para se opor a liberação de dopamina (Rohleder & Leweke, 2015). Basicamente, os receptores de dopamina D<sub>2</sub> estimulam a liberação de anandamida, que por sua vez ativa os receptores CB<sub>1</sub> do sistema endocanabinóide, reduzindo assim os sintomas psicóticos (Giuffrida et al., 1999; Giuffrida et al., 2004), levando a crer que o aumento da anandamida em pacientes esquizofrênicos age em reposta ao aumento da neurotransmissão de dopamina. Além disso, observou-se também a ação

do canabidiol (CBD) na redução de sintomas psicóticos em pacientes esquizofrênicos, ao se reduzir a desativação de anandamida, aumentando os níveis deste metabólito (Leweke et al., 2012) e, assim, ativando os receptores CB<sub>1</sub>. Observou-se também evidências da ação do receptor CB<sub>2</sub> como um possível agente para a redução da hiperativação de dopamina, além deste receptor também agir na ativação microglial, reduzindo a neuroinflamação (Cortez et al., 2020). Outro fator que mostra a interação do sistema endocanabinóide com a dopamina é a interação da N-araquidonoildopamina, produzida a partir da dopamina, que possui ação agonista dos receptores CB<sub>1</sub> e TRPV1. A interação do sistema endocanabinóide discutido acima pode ser ilustrado na **Figura 2.11**.



**Figura 2.11.** Relação da esquizofrenia com alterações inflamatórias e produção de dopamina. Possíveis atenuadores relacionados a receptores endocanabinóides dos processos indicados são ilustrados pelas linhas vermelhas e potenciais compostos endocanabinóides relacionados à melhora dos sintomas psicóticos são ilustrados em verde. Adaptado de Arjmand *et al.*, 2019 e Rohleder & Leweke, 2015.

## 5. Conclusões

Dentro das ferramentas disponíveis para a análise de dados complexos como os obtidos por meio da espectrometria de massas, a utilização de metodologias de aprendizado de máquina apresenta um potencial imenso para a resolução de problemas complexos como os apresentados na metabolômica, em especial na seleção de variáveis de interesse para a discriminação de grupos diversos. Utilizando duas abordagens de aprendizado de máquina frente a dados metabolômicos, foi possível obter variáveis de interesse com potencial para diferenciar um grupo de amostras de outros dois, dentre transtorno bipolar, esquizofrenia e controles saudáveis. Após a identificação das variáveis de interesse, foi possível observar a presença de N-acilamidas e outros compostos do sistema endocanabinóide, cuja importância na fisiopatologia de transtornos mentais é conhecida e estudada há décadas. Observou-se também um potencial fator em comum nas duas doenças, em relação à neuroinflamação, o que fortalece a hipótese de que não somente o transtorno bipolar se trata de uma doença multissistêmica, mas a esquizofrenia também, afirmando, assim, a importância de se pesquisar mais a fundo as interações de ambas as doenças com os diferentes sistemas do metabolismo humano.

## Considerações finais

Juntamente com o avanço das técnicas analíticas, o avanço de metodologias estatísticas e computacionais são essenciais para o desenvolvimento de novas abordagens dentro da química analítica, e com a metabolômica não é diferente. Ao longo dos anos, diversas atualizações e melhorias na espectrometria de massas, em especial no desenvolvimento de novos equipamentos, fontes de ionização, e na sensibilidade das metodologias analíticas fizeram com que o campo da metabolômica se desenvolvesse enormemente em diversas áreas do conhecimento, seja na análise de amostras animais, vegetais, de cultura de células e até na análise de outros organismos vivos como algas e fungos. Entretanto, tais avanços analíticos não poderiam por si só alavancar a metabolômica a patamares vistos atualmente sem o suporte da área computacional e bioinformática.

Vemos presentemente diversos esforços colaborativos entre a medicina e a química e biologia na busca da chamada medicina de precisão, e entende-se no campo científico que abordagens deste tipo são o futuro da medicina nos próximos anos, não somente pelo esforço da comunidade científica na prospecção de biomarcadores nas mais diversas condições, mas também no melhor entendimento do metabolismo humano, de suas vias metabólicas, da interconexão entre as diferentes biomoléculas existentes no corpo humano, das alterações no fenótipo de um indivíduo devido à exposição a diferentes agentes exógenos e da interação entre outros organismos conosco, como por exemplo, as bactérias de nossa microbiota intestinal. Nesse contexto, cada vez mais se faz necessária a análise do organismo humano de forma conectada, como é o proposto pela biologia de sistemas, em especial em condições que apresentam diversos fatores conflitantes, como transtornos mentais.

Assim, de forma a tentar se obter maiores informações quanto às alterações bioquímicas causadas por transtornos mentais, propôs-se duas abordagens diferentes com base na metabolômica.

No Capítulo 1, ao analisarmos as interconexões entre proteínas e metabólitos de pacientes com BD em comparação com indivíduos saudáveis, foi possível

observar alterações na hemóstase relacionadas ao BD, que vai ao encontro às diversas comorbidades na qual estes pacientes são acometidos, em especial doenças cardiovasculares e síndrome metabólica. Também foi possível obter um painel de metabólitos com potencial para diferenciar pacientes com BD de controles saudáveis, onde os metabólitos majoritariamente estavam relacionados ao metabolismo energético.

No Capítulo 2, optou-se por expandir o escopo dentro dos transtornos mentais ao se incluir também pacientes com SZ. Utilizando duas diferentes metodologias de aprendizado de máquina, foi possível obter variáveis de interesse com o potencial de diferenciar cada um dos grupos frente aos outros dois, em uma abordagem multi-grupo. Após a identificação putativa dos metabólitos diferenciais, observou-se a presença de diversas amidas graxas, na qual se correlacionam com o sistema endocanabinoide, que está intimamente ligado ao sistema nervoso central e ao eixo intestino-cerebral. Alterações nos receptores endocanabinóides ajudam a explicar diversas hipóteses relativas à fisiopatologia do BD e da SZ, onde pontuou-se alterações na cascata do ácido araquidônico no caso do BD e na cascata dopaminérgica, no caso da SZ.

Uma das hipóteses defendidas quanto à bioquímica de alguns transtornos mentais, como o BD e a SZ, é de que tais doenças não são somente doenças regidas pelo sistema nervoso central, mas sim, doenças multisistêmicas que envolvem diversas alterações no corpo humano. Esta hipótese também ajuda a explicar o fato de pacientes com transtornos mentais terem chance aumentada de desenvolverem outras condições, como diabetes, doenças cardiovasculares, endócrinas, gastrointestinais, respiratórias e renais. A presença de comorbidades nos pacientes psiquiátricos é responsável por reduzir a qualidade de vida dos indivíduos acometidos pelos diferentes transtornos mentais e, inclusive, de se aumentar o risco de morte, reduzindo a expectativa de vida por volta de 10 anos em comparação com a população em geral.

Claramente, os estudos apresentados nesta Tese possuem suas limitações, em especial quanto ao número de indivíduos incluídos nos estudos. Apesar do número reduzido no Capítulo 1, foi possível obter uma rede de integração de dados proteômicos e metabolômicos com qualidade, funcionando assim como uma prova de

conceito, onde esperamos que esta metodologia desenvolvida possa ser replicada sob diferentes condições no futuro e, assim, validada. No Capítulo 2, um número maior de amostras foi analisado em comparação com o primeiro capítulo, entretanto, quando falamos de metodologias de aprendizado de máquina, é essencial haver um grande número de amostras, pois quanto maior o número, melhor será o resultado da metodologia. Além disso, a utilização de espectrometria de massas por infusão direta faz com que a qualidade dos espectros de fragmentação não seja ideal, apesar da rapidez do método em comparação a metodologias que fazem uso de cromatografia.

Por fim, faz-se necessário pontuar que este trabalho fez uso somente de plataformas computacionais de acesso aberto. É necessário um esforço conjunto da comunidade científica para a popularização e o desenvolvimento de tais plataformas, assim contribuindo para que metodologias promissoras possam ser replicadas em diversos lugares do mundo, que muitas vezes não possuem os fundos necessários para a aquisição de programas pagos de tratamento de dados. Espero que as metodologias desenvolvidas neste trabalho possam ser de bom uso para outros cientistas, não somente para a elucidação de problemas relacionados a transtornos mentais, mas para outras condições clínicas e ambientais que possam fazer uso da integração de dados ômicos para os mais variados fins.

## Referências

- Acharjee, A., Finkers, R., Visser, R. G., & Maliepaard, C. (2013). Comparison of Regularized Regression Methods for Omics Data. *Metabolomics*, 3, 3. <https://doi.org/10.4172/2153-0769.1000126>
- Al Awam, K., Haußleiter, I. S. ibylle, Dudley, E., Donev, R., Brüne, M., Juckel, G., & Thome, J. (2015). Multiplatform metabolome and proteome profiling identifies serum metabolite and protein signatures as prospective biomarkers for schizophrenia. *Journal of Neural Transmission*, 122, S111–S122. <https://doi.org/10.1007/s00702-014-1224-0>
- Alcazar, O., Hernandez, L. F., Nakayasu, E. S., Nicora, C. D., Ansong, C., Muehlbauer, M. J., Bain, J. R., Myer, C. J., Bhattacharya, S. K., Buchwald, P., & Abdulreda, M. H. (2021). Parallel multi-omics in high-risk subjects for the identification of integrated biomarker signatures of type 1 diabetes. *Biomolecules*, 11(3), 383. <https://doi.org/10.3390/biom11030383>
- Almeida, V. De, Alexandrino, G. L., Aquino, A., Gomes, A. F., Murgu, M., Dobrowolny, H., Guest, P. C., Steiner, J., & Martins-de Souza, D. (2020). Changes in the blood plasma lipidome associated with effective or poor response to atypical antipsychotic treatments in schizophrenia patients. *Progress in Neuropsychopharmacology & Biological Psychiatry*, 101, 109945. <https://doi.org/10.1016/j.pnpbp.2020.109945>
- Alves, M. A., Lamichhane, S., Dickens, A., McGlinchey, A., Ribeiro, H. C., Sen, P., Wei, F., Hyötyläinen, T., & Orešič, M. (2021). Systems biology approaches to study lipidomes in health and disease. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, 1866, 2. <https://doi.org/10.1016/j.bbalip.2020.158857>
- American Psychiatric Association. (2014). Manual Diagnóstico e Estatístico de Transtornos Mentais - DSM-5, estatísticas e ciências humanas: inflexões sobre normalizações e normatizações, Artmed, 5. ed., Porto Alegre.
- Amezquita, R. A., Lun, A. T. L., Becht, E., Carey, V. J., Carpp, L. N., Geistlinger, L., Marini, F., Rue-Albrecht, K., Risso, D., Sonesson, C., Waldron, L., Pagès, H.,

- Smith, M. L., Huber, W., Morgan, M., Gottardo, R., & Hicks, S. C. (2020). Orchestrating single-cell analysis with Bioconductor. *Nature Methods*, *17*(2), 137–145. <https://doi.org/10.1038/s41592-019-0654-x>
- Araújo, W. L., Martins, A. O., Fernie, A. R., & Tohge, T. (2014). 2-oxoglutarate: Linking TCA cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. *Frontiers in Plant Science*, *5*, 552. <https://doi.org/10.3389/fpls.2014.00552>
- Arjmand, S., Behzadi, M., Kohlmeier, K. A., Mazhari, S., Sabahi, A., & Shabani, M. (2019). Bipolar disorder and the endocannabinoid system. *Acta Neuropsychiatrica*, *31*(4), 193-201. <https://doi.org/10.1017/neu.2019.21>
- Arul Prakash, S., & Kamlekar, R. K. (2021). Function and therapeutic potential of N-acyl amino acids. *Chemistry and Physics of Lipids*, *239*, 105114. <https://doi.org/10.1016/j.chemphyslip.2021.105114>
- Banerjee, S., & Mazumdar, S. (2012). Electrospray Ionization Mass Spectrometry: A Technique to Access the Information beyond the Molecular Weight of the Analyte. *International Journal of Analytical Chemistry*, *2012*, 282574. <https://doi.org/10.1155/2012/282574>
- Beale, D. J., Pinu, F. R., Kouremenos, K. A., Poojary, M. M., Narayana, V. K., Boughton, B. A., Kanojia, K., Dayalan, S., Jones, O. A. H., & Dias, D. A. (2018). Review of recent developments in GC – MS approaches to metabolomics-based research. *Metabolomics*, *14*, 152. <https://doi.org/10.1007/s11306-018-1449-2>
- Begou, O., Gika, H. G., Wilson, I. D., & Theodoridis, G. (2017). Hyphenated MS-based targeted approaches in metabolomics. *Analyst*, *142*(17), 3079–3100. <https://doi.org/10.1039/c7an00812k>
- Benítez-Angeles, M., Morales-Lázaro, S. L., Juárez-González, E., & Rosenbaum, T. (2020). TRPV1: Structure, endogenous agonists, and mechanisms. *International Journal of Molecular Sciences*, *21*(10), 3421. <https://doi.org/10.3390/ijms21103421>
- Bisogno, T., & Marzo, V. Di. (2007). Short- and long-term plasticity of the endocannabinoid system in neuropsychiatric and neurological disorders.

*Pharmacological Research*, 56, 428–442.  
<https://doi.org/10.1016/j.phrs.2007.09.002>

- Boccio, P. Del, Rossi, C., Cicalini, I., Sacchetta, P., & Pieragostino, D. (2016). Integration of metabolomics and proteomics in multiple sclerosis: From biomarkers discovery to personalized. *Proteomics Clinical Applications*, 10, 470–484. <https://doi.org/10.1002/prca.201500083>
- Boger, D. L., Patterson, J. E., & Jin, Q. (1998). Structural requirements for 5-HT<sub>2A</sub> and 5-HT<sub>1A</sub> serotonin receptor potentiation by the biologically active lipid oleamide. *Proceedings of the National Academy of Sciences of the United States of America*, 95(8), 4102–4107. <https://doi.org/10.1073/pnas.95.8.4102>
- Broadhurst, D., Goodacre, R., Reinke, S. N., Kuligowski, J., Wilson, I. D., Lewis, M. R., & Dunn, W. B. (2018). Guidelines and considerations for the use of system suitability and quality control samples in mass spectrometry assays applied in untargeted clinical metabolomic studies. *Metabolomics*, 14, 72. <https://doi.org/10.1007/s11306-018-1367-3>
- Bujak, R., Dagher-Wojtkowiak, E., Kaliszan, R., & Markuszewski, M. J. (2016). PLS-based and regularization-based methods for the selection of relevant variables in non-targeted metabolomics data. *Frontiers in Molecular Biosciences*, 3, 35. <https://doi.org/10.3389/fmolb.2016.00035>
- Burstein, S. H. (2018). N-Acyl Amino Acids (Elmiric Acids): Endogenous signaling molecules with therapeutic potential. *Molecular Pharmacology*, 93(3), 228–238. <https://doi.org/10.1124/mol.117.110841>
- Bzdok, D., & Meyer-Lindenberg, A. (2018). Machine Learning for Precision Psychiatry: Opportunities and Challenges. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(3), 223–230. <https://doi.org/10.1016/j.bpsc.2017.11.007>
- Cai, H., Li, H., Yan, X., Sun, B., Zhang, Q., Yan, M., Zhang, W., Jiang, P., Zhu, R., Liu, Y., Fang, P., Xu, P., Yuan, H., Zhang, X., Hu, L., Yang, W., & Ye, H. (2012). Metabolomic Analysis of Biochemical Changes in the Plasma Schizophrenia Patients and Urine of First-Episode Neuroleptic-Naive after Treatment with Risperidone. *Journal of Proteome Research*, 11, 4338–4350. <https://doi.org/10.1021/pr300459d>

- Calderón-Guzmán, D., Hernández-Islas, J. L., Espítia Vázquez, I. R., Barragán-Mejía, G., Hernández-García, E., Del Angel, D. S., & Juárez-Olguín, H. (2005). Effect of toluene and cresols on Na<sup>+</sup>,K<sup>+</sup>-ATPase, and serotonin in rat brain. *Regulatory Toxicology and Pharmacology*, 41(1), 1–5. <https://doi.org/10.1016/j.yrtph.2004.09.005>
- Camacho, D., de la Fuente, A., & Mendes, P. (2005). The origin of correlations in metabolomics data. *Metabolomics*, 1(1), 53–63. <https://doi.org/10.1007/s11306-005-1107-3>
- Cambiaghi, A., Díaz, R., Martínez, J. B., Odena, A., Brunelli, L., Caironi, P., Masson, S., Baselli, G., Ristagno, G., Gattinoni, L., De Oliveira, E., Pastorelli, R., & Ferrario, M. (2018). An Innovative Approach for the Integration of Proteomics and Metabolomics Data in Severe Septic Shock Patients Stratified for Mortality. *Scientific Reports*, 8, 6681. <https://doi.org/10.1038/s41598-018-25035-1>
- Carvalho, A. F., Firth, J., & Vieta, E. (2020). Bipolar disorder. *New England Journal of Medicine*, 383, 58–66. <https://doi.org/10.1056/NEJMra1906193>
- Castillo-Peinado, L. S., López-Bascón, M. A., Mena-Bravo, A., Luque de Castro, M. D., & Priego-Capote, F. (2019). Determination of primary fatty acid amides in different biological fluids by LC–MS/MS in MRM mode with synthetic deuterated standards: Influence of biofluid matrix on sample preparation. *Talanta*, 193, 29–36. <https://doi.org/10.1016/j.talanta.2018.09.088>
- Cerini, C., Dou, L., Anfosso, F., Sabatier, F., Moal, V., Glorieux, G., Smet, R. De, Vanholder, R., Dignat-George, F., Sampol, J., Berland, Y., & Brunet, P. (2004). P-cresol, a uremic retention solute, alters the endothelial barrier function in vitro. *Thrombosis and Haemostasis*, 92, 140–150. <https://doi.org/10.1160/TH03-07-0491>
- Chambers, M. C., MacLean, B., Burke, R., Amodei, D., Ruderman, D. L., Neumann, S., Gatto, L., Fischer, B., Pratt, B., Egertson, J., Hoff, K., Kessner, D., Tasman, N., Shulman, N., Frewen, B., Baker, T. A., Brusniak, M. Y., Paulse, C., Creasy, D., ... Mallick, P. (2012). A cross-platform toolkit for mass spectrometry and proteomics. *Nature Biotechnology*, 30(10), 918–920. <https://doi.org/10.1038/nbt.2377>

- Chang, M., Chang, H., Chan, C., & Yeung, S. (2014). p-Cresol Affects Reactive Oxygen Species Generation , Cell Cycle Arrest , Cytotoxicity and Inflammation / Atherosclerosis-Related Modulators Production in Endothelial Cells and Mononuclear Cells. *PLoS ONE*, 9(12), e114446. <https://doi.org/10.1371/journal.pone.0114446>
- Chang, M., Wang, T., Yeung, S., Jeng, P., Liao, C., Lin, T., Lin, C., Lin, B., & Jeng, J. (2011). Antiplatelet effect by p -cresol , a uremic and environmental toxicant , is related to inhibition of reactive oxygen species , ERK / p38 signaling and thromboxane A<sub>2</sub> production. *Atherosclerosis*, 219(2), 559–565. <https://doi.org/10.1016/j.atherosclerosis.2011.09.031>
- Chen, D., Zhao, X., Sui, Z., Niu, H., Chen, L., Hu, C., Xuan, Q., Hou, X., Zhang, R., Zhou, L., Li, Y., Yuan, H., Zhang, Y., Wu, J., Zhang, L., Wu, R., Piao, H. L., Xu, G., & Jia, W. (2020). A multi-omics investigation of the molecular characteristics and classification of six metabolic syndrome relevant diseases. *Theranostics*, 10(5), 2029–2046. <https://doi.org/10.7150/thno.41106>
- Chen, Y. Q., Shen, H., Yang, R. J., & Wan, J. B. (2021). Accurate quantification of endogenous N-acylethanolamides by chemical isotope labeling coupled with liquid chromatography-tandem mass spectrometry. *Analytica Chimica Acta*, 1179, 338839. <https://doi.org/10.1016/j.aca.2021.338839>
- Cortez, I. L., Rodrigues da Silva, N., Guimarães, F. S., & Gomes, F. V. (2020). Are CB2 Receptors a New Target for Schizophrenia Treatment? *Frontiers in Psychiatry*, 11, 587154. <https://doi.org/10.3389/fpsy.2020.587154>
- Cravatt, B. F., Prospero-Garcia, O., Siuzdak, G., Gilula, N. B., Henriksen, S. J., Dale L. Boger, & Lerner, R. A. (1995). Chemical Characterization of a Family of Brain Lipids That Induce Sleep. *Science*, 268(5216), 1506–1509. <https://doi.org/10.1007/BF00322863>
- Cristino, L., Bisogno, T., & Di Marzo, V. (2020). Cannabinoids and the expanded endocannabinoid system in neurological disorders. *Nature Reviews Neurology*, 16(1), 9–29. <https://doi.org/10.1038/s41582-019-0284-z>
- Crump, C., Sundquist, K., Winkleby, M. A., & Sundquist, J. (2013). Comorbidities and mortality in bipolar disorder: A Swedish national cohort study. *JAMA Psychiatry*,

- 70(9), 931–939. <https://doi.org/10.1001/jamapsychiatry.2013.1394>
- Csárdi, G., & Nepusz, T. (2006). The igraph software package for complex network research. In *InterJournal Complex Systems* (Vol. 1695, Issue 5, pp. 1–9). <https://doi.org/10.5281/zenodo.3630268>
- Davison, J., O’Gorman, A., Brennan, L., & Cotter, D. R. (2018). A systematic review of metabolite biomarkers of schizophrenia. *Schizophrenia Research*, *195*, 32–50. <https://doi.org/10.1016/j.schres.2017.09.021>
- de Almeida, V., & Martins-de-Souza, D. (2018). Cannabinoids and glial cells: possible mechanism to understand schizophrenia. *European Archives of Psychiatry and Clinical Neuroscience*, *268*(7), 727–737. <https://doi.org/10.1007/s00406-018-0874-6>
- de Lange, Z., Kahler, B., Smuts, C. M., & Pieters, M. (2019). Plasma phospholipid fatty acids are associated with altered fibrin clot properties in a population-based setting. *Prostaglandins Leukotrienes and Essential Fatty Acids*, *143*, 1–7. <https://doi.org/10.1016/j.plefa.2019.02.001>
- Devane, W. A., Hanus, L., Breuer, A., Pertwee, R. G., Stevenson, L. A., Griffin, G., Gibson, D., Mandelbaum, A., Etinger, A., & Mechoulam, R. (1992). Isolation and Structure of a Brain Constituent That Binds to the Cannabinoid Receptor. *Science*, *258*(5090), 1946–1949. <https://doi.org/10.1039/CT9120102209>
- Dias, D. A., Jones, O. A. H., Beale, D. J., Boughton, B. A., Benheim, D., Kouremenos, K. A., Wolfender, J. L., & Wishart, D. S. (2016). Current and future perspectives on the structural identification of small molecules in biological systems. *Metabolites*, *6*, 46. <https://doi.org/10.3390/metabo6040046>
- Dickens, A., Borgan, F., Laurikainen, H., Lamichhane, S., Marques, T., Rönkkö, T., Veronese, M., Lindeman, T., Hyötyläinen, T., Howes, O., Hietala, J., & Orešič, M. (2020). Links between central CB1-receptor availability and peripheral endocannabinoids in patients with first episode psychosis. *Npj Schizophrenia*, *6*(21), 664086. <https://doi.org/10.1101/664086>
- Divito, E. B., & Cascio, M. (2013). Metabolism, physiology, and analyses of primary fatty acid amides. *Chemical Reviews*, *113*(10), 7343–7353.

<https://doi.org/10.1021/cr300363b>

- Divito, E. B., Kroniser, K. M., & Cascio, M. (2016). Multidimensional liquid chromatography coupled with tandem mass spectrometry for identification of bioactive fatty acyl derivatives. *Frontiers in Physiology*, 7, 608. <https://doi.org/10.3389/fphys.2016.00608>
- Do, K. T., Wahl, S., Raffler, J., Molnos, S., Laimighofer, M., Adamski, J., Suhre, K., Strauch, K., Peters, A., Gieger, C., Langenberg, C., Stewart, I. D., Theis, F. J., Grallert, H., Kastenmüller, G., & Krumsiek, J. (2018). Characterization of missing values in untargeted MS-based metabolomics data and evaluation of missing data handling strategies. *Metabolomics*, 14(10), 1–18. <https://doi.org/10.1007/s11306-018-1420-2>
- Domingo-Almenara, X., Montenegro-Burke, J. R., Benton, H. P., & Siuzdak, G. (2018). Annotation: A Computational Solution for Streamlining Metabolomics Analysis. *Analytical Chemistry*, 90, 480–489. <https://doi.org/10.1021/acs.analchem.7b03929>
- Dugger, B. N., & Dickson, D. W. (2017). Pathology of neurodegenerative diseases. *Cold Spring Harbor Perspectives in Biology*, 9(7), 1–22. <https://doi.org/10.1101/cshperspect.a028035>
- Dührkop, K., Fleischauer, M., Ludwig, M., Aksenov, A. A., Melnik, A. V., Meusel, M., Dorrestein, P. C., Rousu, J., & Böcker, S. (2019). SIRIUS 4: a rapid tool for turning tandem mass spectra into metabolite structure information. *Nature Methods*, 16(4), 299–302. <https://doi.org/10.1038/s41592-019-0344-8>
- Eberlin, L. S., Ferreira, C. R., Dill, A. L., Ifa, D. R., & Cooks, G. R. (2011). Desorption Electrospray Ionization Mass Spectrometry for Lipid Characterization and Biological Tissue Imaging. *Biochimica et Biophysica Acta*, 1811(11), 946–960. <https://doi.org/10.1016/j.bbailip.2011.05.006>
- Engskog, M. K. R., Haglöf, J., Arvidsson, T., & Pettersson, C. (2016). LC–MS based global metabolite profiling: the necessity of high data quality. *Metabolomics*, 12(7). <https://doi.org/10.1007/s11306-016-1058-x>
- Errico, F., Nuzzo, T., Carella, M., Bertolino, A., & Usiello, A. (2018). The emerging role

- of altered d-aspartate metabolism in schizophrenia: New insights from preclinical models and human studies. *Frontiers in Psychiatry*, 9, 559. <https://doi.org/10.3389/fpsy.2018.00559>
- Ezzili, C., Otrubova, K., & Boger, D. L. (2010). Fatty acid amide signaling molecules. *Bioorganic and Medicinal Chemistry Letters*, 20(20), 5959–5968. <https://doi.org/10.1016/j.bmcl.2010.08.048>
- Fabisiak, A., Fabisiak, N., Mokrowiecka, A., Malecka-Panas, E., Jacenik, D., Kordek, R., Zielińska, M., Kieć-Kononowicz, K., & Fichna, J. (2021). Novel selective agonist of GPR18, PSB-KK-1415 exerts potent anti-inflammatory and anti-nociceptive activities in animal models of intestinal inflammation and inflammatory pain. *Neurogastroenterology and Motility*, 33, e14003. <https://doi.org/10.1111/nmo.14003>
- Farrell, E. K., & Merkle, D. J. (2008). Biosynthesis, degradation and pharmacological importance of the fatty acid amides. *Drug Discovery Today*, 13(13–14), 558–568. <https://doi.org/10.1016/j.drudis.2008.02.006>
- Fenaille, F., Barbier Saint-Hilaire, P., Rousseau, K., & Junot, C. (2017). Data acquisition workflows in liquid chromatography coupled to high resolution mass spectrometry-based metabolomics: Where do we stand? *Journal of Chromatography A*, 1526, 1–12. <https://doi.org/10.1016/j.chroma.2017.10.043>
- Fenn, J. B., Mann, M., Meng, C. K. A. I., Wong, S. F., & Whitehouse, C. M. (1989). Electrospray Ionization for Mass Spectrometry of Large Biomolecules. *Science*, 246(4926), 64–71. <https://doi.org/10.1126/science.2675315>
- Fernández-Ruiz, J., Gómez-Ruiz, M., García, C., Hernández, M., & Ramos, J. A. (2017). Modeling Neurodegenerative Disorders for Developing Cannabinoid-Based Neuroprotective Therapies. *Methods in Enzymology*, 593, 175–198. <https://doi.org/10.1016/bs.mie.2017.06.021>
- Fondi, M., & Liò, P. (2015). Multi -omics and metabolic modelling pipelines : Challenges and tools for systems microbiology. *Microbiological Research*, 171, 52–64. <https://doi.org/10.1016/j.micres.2015.01.003>
- Fontanarosa, C., Pane, F., Sepe, N., Pinto, G., Trifuoggi, M., Squillace, M., Errico, F.,

- Usiello, A., Pucci, P., & Amoresano, A. (2017). Quantitative determination of free D-Asp, L-Asp and N-methyl-D-aspartate in mouse brain tissues by chiral separation and Multiple Reaction Monitoring tandem mass spectrometry. *PLoS ONE*, *12*(6), e0179748. <https://doi.org/10.1371/journal.pone.0179748>
- Galal, A., Talal, M., & Moustafa, A. (2022). Applications of machine learning in metabolomics: Disease modeling and classification. *Frontiers in Genetics*, *13*, 1017340. <https://doi.org/10.3389/fgene.2022.1017340>
- Gardoni, F., Stanic, J., Scheggia, D., Benussi, A., Borroni, B., & Di Luca, M. (2021). NMDA and AMPA Receptor Autoantibodies in Brain Disorders: From Molecular Mechanisms to Clinical Features. *Cells*, *10*, 77. <https://doi.org/10.3390/cells10010077>
- GBD 2019 Mental Disorders Collaborators. (2022). Global, regional, and national burden of 12 mental disorders in 204 countries and territories, 1990–2019: a systematic analysis for the Global Burden of Disease Study 2019. *The Lancet Psychiatry*, *9*(2), 137–150. [https://doi.org/10.1016/S2215-0366\(21\)00395-3](https://doi.org/10.1016/S2215-0366(21)00395-3)
- Gebauer, S. K., Tracy, R. P., & Baer, D. J. (2014). Impact of stearic acid and oleic acid on hemostatic factors in the context of controlled diets consumed by healthy men. *European Journal of Clinical Nutrition*, *68*, 1072–1074. <https://doi.org/10.1038/ejcn.2014.62>
- Ghasemi, M., Phillips, C., Trillo, L., Miguel, Z. De, Das, D., & Salehi, A. (2014). The role of NMDA receptors in the pathophysiology and treatment of mood disorders. *Neuroscience and Biobehavioral Reviews*, *47*, 336–358. <https://doi.org/10.1016/j.neubiorev.2014.08.017>
- Gil-De-La-Fuente, A., Godzien, J., Saugar, S., Garcia-Carmona, R., Badran, H., Wishart, D. S., Barbas, C., & Otero, A. (2019). CEU Mass Mediator 3.0: A Metabolite Annotation Tool. *Journal of Proteome Research*, *18*(2), 797–802. <https://doi.org/10.1021/acs.jproteome.8b00720>
- Giuffrida, A., Parsons, L. H., Kerr, T. M., Rodríguez De Fonseca, F., Navarro, M., & Piomelli, D. (1999). Dopamine activation of endogenous cannabinoid signaling in dorsal striatum. *Nature Neuroscience*, *2*(4), 358–363. <https://doi.org/10.1038/7268>

- Giuffrida, Andrea, Leweke, F. M., Gerth, C. W., Schreiber, D., Koethe, D., Faulhaber, J., Klosterkötter, J., & Piomelli, D. (2004). Cerebrospinal anandamide levels are elevated in acute schizophrenia and are inversely correlated with psychotic symptoms. *Neuropsychopharmacology*, 29(11), 2108–2114. <https://doi.org/10.1038/sj.npp.1300558>
- Godlewski, G., Offertáler, L., Wagner, J. A., & Kunos, G. (2009). Receptors for acylethanolamides-GPR55 and GPR119. *Prostaglandins and Other Lipid Mediators*, 89(3–4), 105–111. <https://doi.org/10.1016/j.prostaglandins.2009.07.001>
- Goldstein, B. I., Kemp, D. E., Soczynska, J. K., & McIntyre, R. S. (2009). Inflammation and the phenomenology, pathophysiology, comorbidity, and treatment of bipolar disorder: A systematic review of the literature. *Journal of Clinical Psychiatry*, 70(8), 1078–1090. <https://doi.org/10.4088/JCP.08r04505>
- Goodhart, P. J., DeWolf Jr., W. E., & Kruse, L. I. (1987). Mechanism-Based Inactivation of Dopamine beta-Hydroxylase by p-Cresol and Related Alkylphenols. *Biochemistry*, 26, 2576–2583. <https://doi.org/10.1021/bi00383a025>
- Graham, D. L., Beio, M. L., Nelson, D. L., & Berkowitz, D. B. (2019). Human serine racemase: Key residues/active site motifs and their relation to enzyme function. *Frontiers in Molecular Biosciences*, 6, 8. <https://doi.org/10.3389/fmolb.2019.00008>
- Gromski, P. S., Muhamadali, H., Ellis, D. I., Xu, Y., Correa, E., Turner, M. L., & Goodacre, R. (2015). A tutorial review: Metabolomics and partial least squares-discriminant analysis - a marriage of convenience or a shotgun wedding. *Analytica Chimica Acta*, 879, 10–23. <https://doi.org/10.1016/j.aca.2015.02.012>
- Gubert, C., Stertz, L., Pfaffenseller, B., Schilling, B., Tezza, G., Massuda, R., & Kunz, M. (2013). Mitochondrial activity and oxidative stress markers in peripheral blood mononuclear cells of patients with bipolar disorder, schizophrenia, and healthy subjects. *Journal of Psychiatric Research*, 47(10), 1396–1402. <https://doi.org/10.1016/j.jpsychires.2013.06.018>
- Gui, S.-W., Liu, Y.-Y., Zhong, X.-G., Liu, X., Zheng, P., Pu, J.-C., Zhou, J., Chen, J.-J., Zhao, L.-B., Liu, L.-X., Xu, G., & Xie, P. (2018). Plasma disturbance of

- phospholipid metabolism in major depressive disorder by integration of proteomics and metabolomics. *Neuropsychiatric Disease and Treatment*, *14*, 1451–1461. <https://doi.org/10.2147/NDT.S164134>
- Han, X. (2016). Mass spectrometry for lipidomics. In: *Lipidomics: Comprehensive Mass Spectrometry of Lipids*, John Wiley and Sons, Hoboken.
- Hanuš, L., Shohami, E., Bab, I., & Mechoulam, R. (2014). N-Acyl amino acids and their impact on biological processes. *BioFactors*, *40*(4), 381–388. <https://doi.org/10.1002/biof.1166>
- Hashimoto, K. (2016). Serine enantiomers as diagnostic biomarkers for schizophrenia and bipolar disorder. *European Archives of Psychiatry and Clinical Neuroscience*, *266*(1), 83–85. <https://doi.org/10.1007/s00406-015-0602-4>
- Hicks, S. A., Strümke, I., Thambawita, V., Hammou, M., Riegler, M. A., Halvorsen, P., & Parasa, S. (2022). On evaluation metrics for medical applications of artificial intelligence. *Scientific Reports*, *12*, 5979. <https://doi.org/10.1038/s41598-022-09954-8>
- Hilaire, P. B. Saint, Rousseau, K., Seyer, A., Dechaumet, S., Damont, A., Junot, C., & Fenaille, F. (2020). Comparative evaluation of data dependent and data independent acquisition workflows implemented on an orbitrap fusion for untargeted metabolomics. *Metabolites*, *10*, 4. <https://doi.org/10.3390/metabo10040158>
- Hoirisch-Clapauch, S., Nardi, A. E., Gris, J.-C., & Brenner, B. (2014). Coagulation and Mental Disorders. *Rambam Maimonides Medical Journal*, *5*(4). <https://doi.org/10.5041/rmmj.10170>
- Horning, E. C., Horning, M. G., Carroll, D. I., Dzidic, I., & Stillwell, R. N. (1973). New picogram detection system based on a mass spectrometer with an external ionization source at atmospheric pressure. *Analytical Chemistry*, *45*(6), 936–943. <http://pubs.acs.org/doi/abs/10.1021/ac60328a035>
- Hu, S., Li, A., Huang, T., Lai, J., Li, J., Sublette, M. E., Lu, H., Lu, Q., Du, Y., Hu, Z., Ng, C. H., Zhang, H., Lu, J., Mou, T., Lu, S., Wang, D., Duan, J., Hu, J., & Huang, M. (2019). Gut Microbiota Changes in Patients with Bipolar Depression. *Advanced*

- Science*, 6, 1–11. <https://doi.org/10.1002/adv.201900752>
- Huang, J. K., & Jan, C. R. (2001). Linoleamide, a brain lipid that induces sleep, increases cytosolic Ca<sup>2+</sup> levels in MDCK renal tubular cells. *Life Sciences*, 68(9), 997–1004. [https://doi.org/10.1016/S0024-3205\(00\)01002-X](https://doi.org/10.1016/S0024-3205(00)01002-X)
- Hummel, J., Strehmel, N., Selbig, J., Walther, D., & Kopka, J. (2010). Decision tree supported substructure prediction of metabolites from GC-MS profiles. *Metabolomics*, 6(2), 322–333. <https://doi.org/10.1007/s11306-010-0198-7>
- Hyötyläinen, T., & Orešič, M. (2014). Systems biology strategies to study lipidomes in health and disease. *Progress in Lipid Research*, 55(1), 43–60. <https://doi.org/10.1016/j.plipres.2014.06.001>
- Ito, T., Hayashida, M., Kobayashi, S., Muto, N., Hayashi, A., Yoshimura, T., & Mori, H. (2016). Serine racemase is involved in D-aspartate biosynthesis. *Journal of Biochemistry*, 160(6), 345–353. <https://doi.org/10.1093/jb/mvw043>
- Jassal, B., Matthews, L., Viteri, G., Gong, C., Lorente, P., Fabregat, A., Sidiropoulos, K., Cook, J., Gillespie, M., Haw, R., Loney, F., May, B., Milacic, M., Rothfels, K., Sevilla, C., Shamovsky, V., Shorser, S., Varusai, T., Weiser, J., ... D'Eustachio, P. (2020). The reactome pathway knowledgebase. *Nucleic Acids Research*, 48(D1), D498–D503. <https://doi.org/10.1093/nar/gkz1031>
- Jauhar, S., & Young, A. H. (2019). Controversies in bipolar disorder ; role of second - generation antipsychotic for maintenance therapy. *International Journal of Bipolar Disorders*, 7, 10. <https://doi.org/10.1186/s40345-019-0145-0>
- Jenkins, B., West, J. A., & Koulman, A. (2015). A Review of Odd-Chain Fatty Acid Metabolism and the Role of Pentadecanoic Acid (C15:0) and Heptadecanoic Acid (C17:0) in Health and Disease. *Molecules*, 20, 2425–2444. <https://doi.org/10.3390/molecules20022425>
- Joshi, M. B., Pai, S., Balakrishnan, A., Bhat, M., Kotambail, A., Sharma, P. S. V. N., & Satyamoorthy, K. (2019). Evidence for perturbed metabolic patterns in bipolar disorder subjects associated with lithium responsiveness. *Psychiatry Research*, 273, 252–259. <https://doi.org/10.1016/j.psychres.2019.01.031>
- Jové, M., Mauri-Capdevila, G., Suárez, I., Cambray, S., Sanahuja, J., Quílez, A., Farré,

- J., Benabdelhak, I., Pamplona, R., Portero-Otín, M., & Purroy, F. (2015). Metabolomics predicts stroke recurrence after transient ischemic attack. *Neurology*, *84*(1), 36–45. <https://doi.org/10.1212/WNL.0000000000001093>
- Kampstra, P. (2008). Beanplot: A Boxplot Alternative for Visual Comparison of Distributions. *Journal of Statistical Software*, *28*(1), 1–9. <https://doi.org/10.18637/jss.v028.c01>
- Kato, T. (2017). Neurobiological basis of bipolar disorder : Mitochondrial dysfunction hypothesis and beyond. *Schizophrenia Research*, *187*, 62–66. <https://doi.org/10.1016/j.schres.2016.10.037>
- Kato, T. M., Noriko, M. K., Saitow, F. F., Fuke, S., Masuda, A., Itohara, S., Suzuki, H., & Kato, T. (2018). Ant1 mutant mice bridge the mitochondrial and serotonergic dysfunctions in bipolar disorder. *Molecular Psychiatry*, *25*, 2039–2049. <https://doi.org/10.1038/s41380-018-0251-x>
- Kiebish, M. A., Cullen, J., Mishra, P., Ali, A., Milliman, E., Rodrigues, L. O., Chen, E. Y., Tolstikov, V., Zhang, L., Panagopoulos, K., Shah, P., Chen, Y., Petrovics, G., Rosner, I. L., Sesterhenn, I. A., McLeod, D. G., Granger, E., Sarangarajan, R., Akmaev, V., ... Dobi, A. (2020). Multi-omic serum biomarkers for prognosis of disease progression in prostate cancer. *Journal of Translational Medicine*, *18*, 10. <https://doi.org/10.1186/s12967-019-02185-y>
- Kim, M., & Tagkopoulos, I. (2018). Data integration and predictive modeling methods for multi-omics datasets. *Molecular Omics*, *14*(1), 8–25. <https://doi.org/10.1039/C7MO00051K>
- Kim, Y. M., & Heyman, H. M. (2018). Mass spectrometry-based metabolomics. *Methods in Molecular Biology*, *1775*(1), 107–118. [https://doi.org/10.1007/978-1-4939-7804-5\\_10](https://doi.org/10.1007/978-1-4939-7804-5_10)
- Kind, T., Wohlgemuth, G., Lee, D. Y., Lu, Y., Palazoglu, M., Shahbaz, S., & Fiehn, O. (2009). FiehnLib: Mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Analytical Chemistry*, *81*(24), 10038–10048. <https://doi.org/10.1021/ac9019522>
- Kirpich, A., Ainsworth, E. A., Wedow, J. M., Newman, J. R. B., Michailidis, G., &

- McIntyre, L. M. (2018). Variable selection in omics data: A practical evaluation of small sample sizes. *PLoS ONE*, *13*(6), e0197910. <https://doi.org/10.1371/journal.pone.0197910>
- Klassen, A., Faccio, A. T., Canuto, G. A. B., da Cruz, P. L. R., Ribeiro, H. C., Tavares, M. F. M., & Sussulini, A. (2017). Metabolomics: Definitions and Significance in Systems Biology. In: A. Sussulini (Ed.), *Metabolomics: From Fundamentals to Clinical Applications*, Springer International Publishing, Cham.
- Koppad, S., B, A., Gkoutos, G. V., & Acharjee, A. (2021). Cloud Computing Enabled Big Multi-Omics Data Analytics. *Bioinformatics and Biology Insights*, *15*(2). <https://doi.org/10.1177/11779322211035921>
- Korade, Ž., Liu, W., Warren, E. B., Armstrong, K., Porter, N. A., & Konradi, C. (2017). Effect of psychotropic drug treatment on sterol metabolism. *Schizophrenia Research*, *187*, 74–81. <https://doi.org/10.1016/j.schres.2017.02.001>
- Kozak, K. R., Crews, B. C., Morrow, J. D., Wang, L. H., Ma, Y. H., Weinander, R., Jakobsson, P. J., & Marnett, L. J. (2002). Metabolism of the endocannabinoids, 2-arachidonylglycerol and anandamide, into prostaglandin, thromboxane, and prostacyclin glycerol esters and ethanolamides. *Journal of Biological Chemistry*, *277*(47), 44877–44885. <https://doi.org/10.1074/jbc.M206788200>
- Krumsiek, J., Suhre, K., Illig, T., Adamski, J., & Theis, F. J. (2011). Gaussian graphical modeling reconstructs pathway reactions from high-throughput metabolomics data. *BMC Systems Biology*, *5*, 21. <https://doi.org/10.1186/1752-0509-5-21>
- Kulkarni, P. P., Tiwari, A., Singh, N., Gautam, D., Sonkar, V. K., Agarwal, V., & Dash, D. (2019). Aerobic glycolysis fuels platelet activation: Small-molecule modulators of platelet metabolism as anti-thrombotic agents. *Haematologica*, *104*(4), 806–818. <https://doi.org/10.3324/haematol.2018.205724>
- Kurotani, K., Sato, M., Yasuda, K., Kashima, K., Tanaka, S., Hayashi, T., Shirouchi, B., Akter, S., Kashino, I., Hayabuchi, H., & Mizoue, T. (2017). Even- and odd-chain saturated fatty acids in serum phospholipids are differentially associated with adipokines. *PLoS ONE*, *12*(5), e0178192. <https://doi.org/10.1371/journal.pone.0178192>

- Kursa, M. B., & Rudnicki, W. R. (2010). Feature selection with the boruta package. *Journal of Statistical Software*, *36*(11), 1–13. <https://doi.org/10.18637/jss.v036.i11>
- Lai, Y. H., & Wang, Y. S. (2022). Advances in high-resolution mass spectrometry techniques for analysis of high mass-to-charge ions. *Mass Spectrometry Reviews*, *2022*, e21790. <https://doi.org/10.1002/mas.21790>
- Laleh, P., Yaser, K., Abolfazl, B., Shahriar, A., Mohammad, A. J., Nazila, F., & Alireza, O. (2018). Oleoylethanolamide increases the expression of PPAR-A and reduces appetite and body weight in obese people: A clinical trial. *Appetite*, *128*, 44–49. <https://doi.org/10.1016/j.appet.2018.05.129>
- Lambert, D., Vandevoorde, S., Jonsson, K.-O., & Fowler, C. (2012). The Palmitoylethanolamide Family: A New Class of Anti-Inflammatory Agents? *Current Medicinal Chemistry*, *9*(6), 663–674. <https://doi.org/10.2174/0929867023370707>
- Langfelder, P., & Horvath, S. (2008). WGCNA: An R package for weighted correlation network analysis. *BMC Bioinformatics*, *9*, 559. <https://doi.org/10.1186/1471-2105-9-559>
- Leboyer, M., Soreca, I., Scott, J., Frye, M., Henry, C., Tamouza, R., & Kupfer, D. J. (2012). Can bipolar disorder be viewed as a multi-system inflammatory disease? *Journal of Affective Disorders*, *141*, 1–10. <https://doi.org/10.1016/j.jad.2011.12.049>
- Lee, S. Y., Kim, M., Jung, S., Lee, S. H., & Lee, J. H. (2015). Altered plasma lysophosphatidylcholines and amides in non-obese and non-diabetic subjects with borderline-to-moderate hypertriglyceridemia: A case-control study. *PLoS ONE*, *10*(4), e0123306. <https://doi.org/10.1371/journal.pone.0123306>
- Lee, W. N. P. (2006). Characterizing phenotype with tracer based metabolomics. *Metabolomics*, *2*(1), 31–39. <https://doi.org/10.1007/s11306-006-0017-3>
- Leishman, E., & Bradshaw, H. B. (2015). N-Acyl Amides: Ubiquitous Endogenous Cannabimimetic Lipids That Are in the Right Place at the Right Time. In: *The Endocannabinoidome: The World of Endocannabinoids and Related Mediators* (pp. 33–48). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-420126-2.00003-1>

- Lenski, M., Sidibé, J., Gholam, M., Hennart, B., Dubath, C., Augsburger, M., Gunten, A. Von, Conus, P., Allorge, D., Thomas, A., & Eap, C. B. (2021). Metabolomic alteration induced by psychotropic drugs : term metabolite profile as a predictor of weight gain evolution. *Clinical and Translational Science*, *14*, 2544–2555. <https://doi.org/10.1111/cts.13122>
- Leweke, F. M., Giuffrida, A., Wurster, U., Emrich, H. M., & Piomelli, D. (1999). Elevated endogenous cannabinoids in schizophrenia. *NeuroReport*, *10*(8), 1665–1669.
- Leweke, F. M., Piomelli, D., Pahlisch, F., Muhl, D., Gerth, C. W., Hoyer, C., Klosterkötter, J., Hellmich, M., & Koethe, D. (2012). Cannabidiol enhances anandamide signaling and alleviates psychotic symptoms of schizophrenia. *Translational Psychiatry*, *2*, e94. <https://doi.org/10.1038/tp.2012.15>
- Li, M. M., Jiang, Z. er, Song, L. Y., Quan, Z. S., & Yu, H. L. (2017). Antidepressant and anxiolytic-like behavioral effects of erucamide, a bioactive fatty acid amide, involving the hypothalamus–pituitary–adrenal axis in mice. *Neuroscience Letters*, *640*, 6–12. <https://doi.org/10.1016/j.neulet.2016.12.072>
- Li, P., L. Snyder, G., & E. Vanover, K. (2016). Dopamine Targeting Drugs for the Treatment of Schizophrenia: Past, Present and Future. *Current Topics in Medicinal Chemistry*, *16*(29), 3385–3403. <https://doi.org/10.2174/1568026616666160608084834>
- Liebal, U. W., Phan, A. N. T., Sudhakar, M., Raman, K., & Blank, L. M. (2020). Machine learning applications for mass spectrometry-based metabolomics. *Metabolites*, *10*, 243. <https://doi.org/10.3390/metabo10060243>
- Liebisch, G., Fahy, E., Aoki, J., Dennis, E. A., Durand, T., Ejsing, C. S., Fedorova, M., Feussner, I., Griffiths, W. J., Köfeler, H., Merrill, A. H., Murphy, R. C., O'Donnell, V. B., Oskolkova, O., Subramaniam, S., Wakelam, M. J. O., & Spener, F. (2020). Update on LIPID MAPS classification, nomenclature, and shorthand notation for MS-derived lipid structures. *Journal of Lipid Research*, *61*(12), 1539–1555. <https://doi.org/10.1194/jlr.S120001025>
- Liebisch, G., Vizcaíno, J. A., Köfeler, H., Trötz Müller, M., Griffiths, W. J., Schmitz, G., Spener, F., & Wakelam, M. J. O. (2013). Shorthand notation for lipid structures derived from mass spectrometry. *Journal of Lipid Research*, *54*(6), 1523–1530.

<https://doi.org/10.1194/jlr.M033506>

- Lip, G. Y. H., & Lowe, G. D. O. (1995). Fibrin D-dimer: A useful clinical marker of thrombogenesis? *Clinical Science*, 89(3), 205–214. <https://doi.org/10.1042/cs0890205>
- Liu, M. L., Zheng, P., Liu, Z., Xu, Y., Mu, J., Guo, J., Huang, T., Meng, H. Q., & Xie, P. (2014). GC-MS based metabolomics identification of possible novel biomarkers for schizophrenia in peripheral blood mononuclear cells. *Molecular BioSystems*, 10(9), 2398–2406. <https://doi.org/10.1039/c4mb00157e>
- Liu, Y. C., & Wu, S. N. (2003). Block of erg current by linoleoylamide, a sleep-inducing agent, in pituitary GH3 cells. *European Journal of Pharmacology*, 458(1–2), 37–47. [https://doi.org/10.1016/S0014-2999\(02\)02728-0](https://doi.org/10.1016/S0014-2999(02)02728-0)
- Lu, H. C., & MacKie, K. (2016). An introduction to the endogenous cannabinoid system. *Biological Psychiatry*, 79(7), 516–525. <https://doi.org/10.1016/j.biopsych.2015.07.028>
- Lu, Q., Lai, J., Lu, H., Ng, C., Huang, T., Zhang, H., Ding, K., Wang, Z., Jiang, J., Hu, J., Lu, J., Lu, S., Mou, T., Wang, D., Du, Y., Xi, C., Lyu, H., Chen, J., Xu, Y., ... Hu, S. (2019). Gut Microbiota in Bipolar Depression and Its Relationship to Brain Function: An Advanced Exploration. *Frontiers in Psychiatry*, 10, 784. <https://doi.org/10.3389/fpsy.2019.00784>
- Machado-Vieira, R. (2018). Lithium , Stress , and Resilience in Bipolar Disorder : Deciphering this key homeostatic synaptic plasticity regulator. *Journal of Affective Disorders*, 233, 92–99. <https://doi.org/10.1016/j.jad.2017.12.026>
- Madrid-Gambin, F., Föcking, M., Sabherwal, S., Heurich, M., English, J. A., O’Gorman, A., Suvitaival, T., Ahonen, L., Cannon, M., Lewis, G., Mattila, I., Scaife, C., Madden, S., Hyötyläinen, T., Orešič, M., Zammit, S., Cagney, G., Cotter, D. R., & Brennan, L. (2019). Integrated Lipidomics and Proteomics Point to Early Blood-Based Changes in Childhood Preceding Later Development of Psychotic Experiences: Evidence From the Avon Longitudinal Study of Parents and Children. *Biological Psychiatry*, 86(1), 25–34. <https://doi.org/10.1016/j.biopsych.2019.01.018>

- Maione, S., Piscitelli, F., Gatta, L., Vita, D., De Petrocellis, L., Palazzo, E., De Novellis, V., & Di Marzo, V. (2011). Non-psychoactive cannabinoids modulate the descending pathway of antinociception in anaesthetized rats through several mechanisms of action. *British Journal of Pharmacology*, *162*(3), 584–596. <https://doi.org/10.1111/j.1476-5381.2010.01063.x>
- Marshall, A. G., & Hendrickson, C. L. (2008). High-resolution mass spectrometers. *Annual Review of Analytical Chemistry*, *1*(1), 579–599. <https://doi.org/10.1146/annurev.anchem.1.031207.112945>
- McCarthy, D. J., Campbell, K. R., Lun, A. T. L., & Wills, Q. F. (2017). Scater: Pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. *Bioinformatics*, *33*(8), 1179–1186. <https://doi.org/10.1093/bioinformatics/btw777>
- Mcclay, J. L., Vunck, S. A., Batman, A. M., Crowley, J. J., Vann, R. E., Beardsley, P. M., & Oord, E. J. van Den. (2015). Neurochemical Metabolomics Reveals Disruption to Sphingolipid Metabolism Following Chronic Haloperidol Administration. *Journal of Neuroimmune Pharmacology*, *10*, 425–434. <https://doi.org/10.1007/s11481-015-9605-1>
- McCutcheon, R. A., Reis Marques, T., & Howes, O. D. (2020). Schizophrenia - An Overview. *JAMA Psychiatry*, *77*(2), 201–210. <https://doi.org/10.1001/jamapsychiatry.2019.3360>
- McGurk, K. A., Williams, S. G., Guo, H., Watkins, H., Farrall, M., Cordell, H. J., Nicolaou, A., & Keavney, B. D. (2021). Heritability and family-based GWAS analyses of the N-acyl ethanolamine and ceramide plasma lipidome. *Human Molecular Genetics*, *30*(6), 500–513. <https://doi.org/10.1093/hmg/ddab002>
- McNamara, R. K., Jandacek, R., Rider, T., Tso, P., Stanford, K. E., Hahn, C. G., & Richtand, N. M. (2008). Deficits in docosahexaenoic acid and associated elevations in the metabolism of arachidonic acid and saturated fatty acids in the postmortem orbitofrontal cortex of patients with bipolar disorder. *Psychiatry Research*, *160*(3), 285–299. <https://doi.org/10.1016/j.psychres.2007.08.021>
- McNamara, R. K., Jandacek, R., Tso, P., Blom, T. J., Welge, J. A., Strawn, J. R., Adler, C. M., DeBello, M. P., & Strakowski, S. M. (2015). First-episode bipolar disorder

- is associated with erythrocyte membrane docosahexaenoic acid deficits: Dissociation from clinical response to lithium or quetiapine. *Psychiatry Research*, 230(2), 447–453. <https://doi.org/10.1016/j.psychres.2015.09.035>
- Mechoulam, R., & Parker, L. A. (2013). The endocannabinoid system and the brain. *Annual Review of Psychology*, 64, 21–47. <https://doi.org/10.1146/annurev-psych-113011-143739>
- Mendez, K. M., Reinke, S. N., & Broadhurst, D. I. (2019). A comparative evaluation of the generalised predictive ability of eight machine learning algorithms across ten clinical metabolomics data sets for binary classification. *Metabolomics*, 15, 150. <https://doi.org/10.1007/s11306-019-1612-4>
- Merikangas, K. R., Jin, R., He, J. P., Kessler, R. C., Lee, S., Sampson, N. A., Viana, M. C., Andrade, L. H., Hu, C., Karam, E. G., Ladea, M., Medina-Mora, M. E., Ono, Y., Posada-Villa, J., Sagar, R., Wells, J. E., & Zarkov, Z. (2011). Prevalence and correlates of bipolar spectrum disorder in the World Mental Health Survey Initiative. *Archives of General Psychiatry*, 68(3), 241–251. <https://doi.org/10.1001/archgenpsychiatry.2011.12>
- Messamore, E. (2018). The niacin response biomarker as a schizophrenia endophenotype: A status update. *Prostaglandins Leukotrienes and Essential Fatty Acids*, 136, 95–97. <https://doi.org/10.1016/j.plefa.2017.06.014>
- Misra, B. B., Langefeld, C., Olivier, M., & Cox, L. A. (2019). Integrated omics : tools , advances and future approaches. *Journal of Molecular Endocrinology*, 62(1), R21–R45. <https://doi.org/10.1530/JME-18-0055>
- Mitchell, A. J., & Dinan, T. G. (2010). Schizophrenia: a multisystem disease? *Journal of Psychopharmacology (Oxford, England)*, 24(4 Suppl), 5–7. <https://doi.org/10.1177/1359786810382059>
- Mitchell, A. J., Vancampfort, D., Sweers, K., Winkel, R. Van, Yu, W., & Hert, M. De. (2013). Prevalence of Metabolic Syndrome and Metabolic Abnormalities in Schizophrenia and Related Disorders — A Systematic Review and Meta-Analysis. *Schizophrenia Bulletin*, 39(2), 306–318. <https://doi.org/10.1093/schbul/sbr148>
- Moreno, E., Cavic, M., & Canela, E. I. (2021). Functional fine-tuning of metabolic

- pathways by the endocannabinoid system—implications for health and disease. *International Journal of Molecular Sciences*, 22, 7. <https://doi.org/10.3390/ijms22073661>
- Mulder, A. M., & Cravatt, B. F. (2006). Endocannabinoid metabolism in the absence of fatty acid amide hydrolase (FAAH): Discovery of phosphorylcholine derivatives of N-acyl ethanolamines. *Biochemistry*, 45(38), 11267–11277. <https://doi.org/10.1021/bi061122s>
- Munson, M. S. B., & Field, F. H. (1966). Chemical Ionization Mass Spectrometry. I. General Introduction. *Journal of the American Academy Chemical Society*, 88(12), 2621–2630. <https://doi.org/10.1038/156283a0>
- Murray, C. J. L., Lopez, A. D., & Jamison, D. T. (1994). The global burden of disease in 1990: Summary results, sensitivity analysis and future directions. *Bulletin of the World Health Organization*, 72(3), 495–509.
- Musee, J., & Marnett, L. J. (2012). Prostaglandin H synthase-2-catalyzed oxygenation of 2-arachidonoylglycerol is more sensitive to peroxide tone than oxygenation of arachidonic acid. *Journal of Biological Chemistry*, 287(44), 37383–37394. <https://doi.org/10.1074/jbc.M112.381202>
- Najjar, S., Pearlman, D. M., Alper, K., Najjar, A., & Devinsky, O. (2013). Neuroinflammation and psychiatric illness. *Journal of Neuroinflammation*, 10(43), 816. <https://doi.org/10.1186/1742-2094-10-43>
- Narasimhan, B., Mourya, V., & Dhake, A. (2006). Design, synthesis, antibacterial, and QSAR studies of myristic acid derivatives. *Bioorganic and Medicinal Chemistry Letters*, 16(11), 3023–3029. <https://doi.org/10.1016/j.bmcl.2006.02.056>
- Neto, E. C., Bare, J. C., & Margolin, A. A. (2014). Simulation studies as designed experiments: The comparison of penalized regression models in the "large p, small n" setting. *PLoS ONE*, 9(10), e107957. <https://doi.org/10.1371/journal.pone.0107957>
- Newcomer, J. W. (2005). Second-Generation ( Atypical ) Antipsychotics and Metabolic Effects A Comprehensive Literature Review. *CNS Drugs*, 19(Suppl 1), 1–93.
- Nier, A. O. (1947). A mass spectrometer for isotope and gas analysis. *Review of*

- Scientific Instruments*, 18(6), 398–411. <https://doi.org/10.1063/1.1740961>
- Ogata, H., Goto, S., Sato, K., Fujibuchi, W., Bono, H., & Kanehisa, M. (1999). KEGG: Kyoto encyclopedia of genes and genomes. *Nucleic Acids Research*, 27(1), 29–34. <https://doi.org/10.1093/nar/27.1.29>
- Pålsson, E., Jakobsson, J., Södersten, K., Fujita, Y., Sellgren, C., Ekman, C. J., Ågren, H., Hashimoto, K., & Landén, M. (2015). Markers of glutamate signaling in cerebrospinal fluid and serum from patients with bipolar disorder and healthy controls. *European Neuropsychopharmacology*, 25(1), 133–140. <https://doi.org/10.1016/j.euroneuro.2014.11.001>
- Paredes, R. M., Quinones, M., Marballi, K., Gao, X., Valdez, C., Ahuja, S. S., Velligan, D., & Walss-bass, C. (2014). Metabolomic profiling of schizophrenia patients at risk for metabolic syndrome. *International Journal of Neuropsychopharmacology*, 17, 1139–1148. <https://doi.org/10.1017/S1461145714000157>
- Pascucci, T., Colamartino, M., Fiori, E., Sacco, R., Coviello, A., Ventura, R., Puglisi-Allegra, S., Turriziani, L., & Persico, A. M. (2020). P-cresol alters brain dopamine metabolism and exacerbates autism-like behaviors in the BTBR mouse. *Brain Sciences*, 10(4), 233. <https://doi.org/10.3390/brainsci10040233>
- Patinkin, D., Milman, G., Breuer, A., Fride, E., & Mechoulam, R. (2008). Endocannabinoids as positive or negative factors in hematopoietic cell migration and differentiation. *European Journal of Pharmacology*, 595(1–3), 1–6. <https://doi.org/10.1016/j.ejphar.2008.05.002>
- Pereira, C., Chavarria, V., Vian, J., Ashton, M., Berk, M., Bhealthsci, W. M., May, O., & Hons, D. (2018). Mitochondrial Agents for Bipolar Disorder. *International Journal of Neuropsychopharmacology*, 21(6), 550–569. <https://doi.org/10.1093/ijnp/pyy018>
- Perez de Souza, L., Alseekh, S., Scossa, F., & Fernie, A. R. (2021). Ultra-high-performance liquid chromatography high-resolution mass spectrometry variants for metabolomics research. *Nature Methods*, 18(7), 733–746. <https://doi.org/10.1038/s41592-021-01116-4>
- Peterson, A. C., Russell, J. D., Bailey, D. J., Westphall, M. S., & Coon, J. J. (2012).

- Parallel reaction monitoring for high resolution and high mass accuracy quantitative, targeted proteomics. *Molecular and Cellular Proteomics*, 11(11), 1475–1488. <https://doi.org/10.1074/mcp.O112.020131>
- Petrosino, S., & Di Marzo, V. (2009). Anandamide and Other Acylethanolamides. In: *Handbook of Neurochemistry and Molecular Neurobiology*. <https://doi.org/10.1007/978-0-387-30378-9>
- Pfeuffer, M., & Jaudszus, A. (2016). Pentadecanoic and heptadecanoic acids: Multifaceted odd-chain fatty acids. *Advances in Nutrition*, 7(4), 730–734. <https://doi.org/10.3945/an.115.011387>
- Pillinger, T., Beck, K., Stubbs, B., & Howes, O. D. (2017). Cholesterol and triglyceride levels in first-episode psychosis: Systematic review and meta-analysis. *British Journal of Psychiatry*, 211(6), 339–349. <https://doi.org/10.1192/bjp.bp.117.200907>
- Pinu, F. R., Beale, D. J., Paten, A. M., Kouremenos, K., Swarup, S., Schirra, H. J., & Wishart, D. (2019). Systems Biology and Multi-Omics Integration: Viewpoints from the Metabolomics Research Community. *Metabolites*, 9(4), 76. <https://doi.org/10.3390/metabo9040076>
- Pomponi, M., Janiri, L., La Torre, G., Di Stasio, E., Di Nicola, M., Mazza, M., Martinotti, G., Bria, P., Lippa, S., Natili, R., & Pomponi, M. F. L. (2013). Plasma levels of n-3 fatty acids in bipolar patients: Deficit restricted to DHA. *Journal of Psychiatric Research*, 47(3), 337–342. <https://doi.org/10.1016/j.jpsychires.2012.11.004>
- Potvin, S., Mahrouche, L., Assaf, R., Chicoine, M., Giguère, C. É., Furtos, A., & Godbout, R. (2020). Peripheral Endogenous Cannabinoid Levels Are Increased in Schizophrenia Patients Evaluated in a Psychiatric Emergency Setting. *Frontiers in Psychiatry*, 11, 628. <https://doi.org/10.3389/fpsy.2020.00628>
- Prusakiewicz, J. J., Kingsley, P. J., Kozak, K. R., & Marnett, L. J. (2002). Selective oxygenation of N -arachidonylglycine by cyclooxygenase-2. *Biochemical and Biophysical Research Communications*, 296, 612–617.
- Qi, B. L., Liu, P., Wang, Q. Y., Cai, W. J., Yuan, B. F., & Feng, Y. Q. (2014). Derivatization for liquid chromatography-mass spectrometry. *TrAC - Trends in*

- Analytical Chemistry*, 59, 121–132. <https://doi.org/10.1016/j.trac.2014.03.013>
- R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at <https://www.R-project.org/>. (2018).
- Rajula, H. S. R., Manchia, M., Carpiniello, B., & Fanos, V. (2021). Big data in severe mental illness: The role of electronic monitoring tools and metabolomics. *Personalized Medicine*, 18(1), 75–90. <https://doi.org/10.2217/pme-2020-0033>
- Ramírez-Orozco, R. E., García-Ruiz, R., Morales, P., Villalón, C. M., Villafán-Bernal, J. R., & Marichal-Cancino, B. A. (2019). Potential metabolic and behavioural roles of the putative endocannabinoid receptors GPR18, GPR55 and GPR119 in feeding. *Current Neuropharmacology*, 17(10), 947–960. <https://doi.org/10.2174/1570159x17666190118143014>
- Rao, J. S., Harry, G. J., Rapoport, S. I., & Kim, H. W. (2010). Increased excitotoxicity and neuroinflammatory markers in postmortem frontal cortex from bipolar disorder patients. *Molecular Psychiatry*, 18, 384–392. <https://doi.org/10.1038/mp.2009.47>
- Rapoport, S. I. (2014). Lithium and the Other Mood Stabilizers Effective in Bipolar Disorder Target the Rat Brain Arachidonic Acid Cascade. *ACS Chemical Neuroscience*, 5, 459–467.
- Rauniyar, N. (2015). Parallel reaction monitoring: A targeted experiment performed using high resolution and high mass accuracy mass spectrometry. *International Journal of Molecular Sciences*, 16(12), 28566–28581. <https://doi.org/10.3390/ijms161226120>
- Ribeiro, H. C., Sen, P., Dickens, A., Santa Cruz, E. C., Orešič, M., & Sussulini, A. (2022). Metabolomic and proteomic profiling in bipolar disorder patients revealed potential molecular signatures related to hemostasis. *Metabolomics*, 18, 65. <https://doi.org/10.1007/s11306-022-01924-5>
- Robb, D. B., Covey, T. R., & Bruins, A. P. (2000). Atmospheric Pressure Photoionization: An Ionization Method for Liquid Chromatography–Mass Spectrometry. *Analytical Chemistry*, 72(15), 3653–3659. <http://pubs.acs.org/doi/abs/10.1021/ac0001636>

- Roberts, L. D., Souza, A. L., Gerszten, R. E., & Clish, C. B. (2012). Targeted metabolomics. In: *Current Protocols in Molecular Biology* (Vol. 1, Issue SUPPL.98, pp. 30.2.1-30.2.24). <https://doi.org/10.1002/0471142727.mb3002s98>
- Rochat, B. (2016). From targeted quantification to untargeted metabolomics: Why LC-high-resolution-MS will become a key instrument in clinical labs. *TrAC - Trends in Analytical Chemistry*, *84*, 151–164. <https://doi.org/10.1016/j.trac.2016.02.009>
- Rohleder, C., & Leweke, F. M. (2015). Cannabinoids and schizophrenia. In *Cannabinoids in Neurologic and Mental Disease* (pp. 191–204). Elsevier Inc. <https://doi.org/10.1016/B978-0-12-417041-4.00008-4>
- Röhrig, W., Achenbach, S., Deutsch, B., & Pischetsrieder, M. (2019). Quantification of 24 circulating endocannabinoids, endocannabinoid-related compounds, and their phospholipid precursors in human plasma by UHPLC-MS/MS. *Journal of Lipid Research*, *60*(8), 1475–1488. <https://doi.org/10.1194/jlr.D094680>
- Rosato, A., Tenori, L., Cascante, M., De Atauri Carulla, P. R., Martins dos Santos, V. A. P., & Saccenti, E. (2018). From correlation to causation: analysis of metabolomics data using systems biology approaches. *Metabolomics*, *14*, 37. <https://doi.org/10.1007/s11306-018-1335-y>
- Roux, M., & Dosseto, A. (2017). From direct to indirect lithium targets: a comprehensive review of omics data. *Metallomics*, *9*(10), 1326–1351. <https://doi.org/10.1039/C7MT00203C>
- Roverato, A., & Castelo, R. (2017). The networked partial correlation and its application to the analysis of genetic interactions. *Journal of the Royal Statistical Society. Series C: Applied Statistics*, *66*(3), 647–665. <https://doi.org/10.1111/rssc.12166>
- Santa Cruz, E. C., Zandonadi, F. da S., Fontes, W., & Sussulini, A. (2021). A pilot study indicating the dysregulation of the complement and coagulation cascades in treated schizophrenia and bipolar disorder patients. *Biochimica et Biophysica Acta (BBA) - Proteins and Proteomics*, *1869*(8), 140657. <https://doi.org/10.1016/j.bbapap.2021.140657>
- Santos, E. A. F., Santa Cruz, E. C., Ribeiro, H. C., Barbosa, L. D., Zandonadi, F. S., & Sussulini, A. (2020). Multi-omics: An Opportunity to Dive into Systems Biology.

*Brazilian Journal of Analytical Chemistry*, 7(29), 18–44.  
<https://doi.org/10.30744/brjac.2179-3425.rv-03-2020>

- Saturnino, C., Popolo, A., Ramunno, A., Adesso, S., Pecoraro, M., Plutino, M. R., Rizzato, S., Albinati, A., Marzocco, S., Sala, M., Iacopetta, D., & Sinicropi, M. S. (2017). Anti-inflammatory, antioxidant and crystallographic studies of N-Palmitoyl-ethanol amine (PEA) derivatives. *Molecules*, 22(4), 616. <https://doi.org/10.3390/molecules22040616>
- SayuriYamagata, A., Brietzke, E., Rosenblat, J. D., Kakar, R., & McIntyre, R. S. (2017). Medical comorbidity in bipolar disorder: The link with metabolic-inflammatory systems. *Journal of Affective Disorders*, 211, 99–106. <https://doi.org/10.1016/j.jad.2016.12.059>
- Scarabin, P. Y., Aillaud, M. F., Luc, G., Lacroix, B., Mennen, L., Amouyel, P., Evans, A., Ferrières, J., Arveiler, D., & Juhan-Vague, I. (2001). Haemostasis in relation to dietary fat as estimated by erythrocyte fatty acid composition: The prime study. *Thrombosis Research*, 102(4), 285–293. [https://doi.org/10.1016/S0049-3848\(01\)00245-6](https://doi.org/10.1016/S0049-3848(01)00245-6)
- Scaravilli, V., Di Girolamo, L., Scotti, E., Busana, M., Biancolilli, O., Leonardi, P., Carlin, A., Lonati, C., Panigada, M., Pesenti, A., & Zanella, A. (2018). Effects of sodium citrate, citric acid and lactic acid on human blood coagulation. *Perfusion*, 33(7), 577–583. <https://doi.org/10.1177/0267659118777441>
- Seeman, P. (1987). Dopamine Receptors and the Dopamine Hypothesis of Schizophrenia. *Synapse*, 1, 133–152.
- Severus, E., & Bauer, M. (2013). Diagnosing bipolar disorders in DSM-5. *International Journal of Bipolar Disorders*, 1, 14. <https://doi.org/10.1186/2194-7511-1-14>
- Shine, B., Mcknight, R. F., Leaver, L., & Geddes, J. R. (2015). Long-term effects of lithium on renal, thyroid, and parathyroid function: A retrospective analysis of laboratory data. *The Lancet*, 386(9992), 461–468. [https://doi.org/10.1016/S0140-6736\(14\)61842-0](https://doi.org/10.1016/S0140-6736(14)61842-0)
- Singh, A., Cao, K. L., & Rohart, F. (2017). mixOmics : An R package for ‘ omics feature selection and multiple data integration. *PLoS Computational Biology*, 13(11),

e1005752. <https://doi.org/10.1371/journal.pcbi.1005752>

- Sinha, A., Shariq, A., Said, K., Sharma, A., Jeffrey Newport, D., & Salloum, I. M. (2018). Medical Comorbidities in Bipolar Disorder. *Current Psychiatry Reports*, 20, 5. <https://doi.org/10.1007/s11920-018-0897-8>
- Skonieczna-Żydecka, K., Łoniewski, I., Stachowska, E., Marlicz, W., & Correll, C. U. (2020). Current and Novel Approaches to Mitigate Cardiometabolic Adverse Effects of Second-Generation Antipsychotics. *International Journal of Neuropsychopharmacology*, 23(8), 491–495. <https://doi.org/10.1093/ijnp/pyaa026>
- Smesny, S., Rosburg, T., Baur, K., Rudolph, N., & Sauer, H. (2007). Cannabinoids influence lipid-arachidonic acid pathways in schizophrenia. *Neuropsychopharmacology*, 32(10), 2067–2073. <https://doi.org/10.1038/sj.npp.1301343>
- Subramanian, I., Verma, S., Kumar, S., Jere, A., & Anamika, K. (2020). Multi-omics Data Integration , Interpretation , and Its Application. *Bioinformatics and Biology Insights*, 14, 1–24. <https://doi.org/10.1177/1177932219899051>
- Sud, M., Fahy, E., Cotter, D., Azam, K., Vadivelu, I., Burant, C., Edison, A., Fiehn, O., Higashi, R., Nair, K. S., Sumner, S., & Subramaniam, S. (2016). Metabolomics Workbench: An international repository for metabolomics data and metadata, metabolite standards, protocols, tutorials and training, and analysis tools. *Nucleic Acids Research*, 44(D1), D463–D470. <https://doi.org/10.1093/nar/gkv1042>
- Sumner, L. W., Amberg, A., Barrett, D., Beale, M. H., Beger, R., Daykin, C. A., Fan, T. W. M., Fiehn, O., Goodacre, R., Griffin, J. L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R., Kopka, J., Lane, A. N., Lindon, J. C., Marriott, P., Nicholls, A. W., ... Viant, M. R. (2007). Proposed minimum reporting standards for chemical analysis: Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). *Metabolomics*, 3(3), 211–221.
- Sussulini, A., Prando, A., Maretto, D. A., Poppi, R. J., Tasic, L., Banzato, C. E. M., & Arruda, M. A. Z. (2009). Metabolic profiling of human blood serum from treated patients with bipolar disorder employing <sup>1</sup>H NMR spectroscopy and chemometrics. *Analytical Chemistry*, 81(23), 9755–9763.

<https://doi.org/10.1021/ac901502j>

- Sylvia, L. G., Shelton, R. C., Kemp, D. E., Bernstein, E. E., Friedman, E. S., Brody, B. D., Mcelroy, S. L., Singh, V., Tohen, M., Bowden, C. L., Ketter, T. A., Deckersbach, T., Thase, M. E., Reilly-Harrington, N. A., Nierenberg, A. A., Rabideau, D. J., Kinrys, G., Kocsis, J. H., Bobo, W. V., ... Calabrese, J. R. (2015). Medical burden in bipolar disorder: Findings from the Clinical and Health Outcomes Initiative in Comparative Effectiveness for Bipolar Disorder study (Bipolar CHOICE). *Bipolar Disorders*, *17*(2), 212–223. <https://doi.org/10.1111/bdi.12243>
- Szymańska, E., Saccenti, E., Smilde, A. K., & Westerhuis, J. A. (2012). Double-check: Validation of diagnostic statistics for PLS-DA models in metabolomics studies. *Metabolomics*, *8*, S3–S16. <https://doi.org/10.1007/s11306-011-0330-3>
- Tandon, N., & Tandon, R. (2019). Using machine learning to explain the heterogeneity of schizophrenia. Realizing the promise and avoiding the hype. *Schizophrenia Research*, *214*, 70–75. <https://doi.org/10.1016/j.schres.2019.08.032>
- Thijssen, M. A. M. A., Hornstra, G., & Mensink, R. P. (2005). Stearic , Oleic , and Linoleic Acids Have Comparable Effects on Markers of Thrombotic Tendency in Healthy Human Subjects. *The Journal of Nutrition*, *135*, 2805–2811. <https://doi.org/10.1093/jn/135.12.2805>
- Tholstrup, T. (2005). Influence of Stearic Acid on Hemostatic Risk Factors in Humans. *Lipids*, *40*(12), 1229–1235. <https://doi.org/10.1007/s11745-005-1490-1>
- Tibshirani, R. (1996). Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*, *58*(1), 267–288. <https://doi.org/10.1111/j.2517-6161.1996.tb02080.x>
- Trindade, P., Nascimento, J. M., Casas, B. S., Monteverde, T., Gasparotto, J., Ribeiro, C. T., Devalle, S., Sauma, D., Moreira, J. C. F., Gelain, D. P., Porciuncula, L. O., Palma, V., Martins-de-Souza, D., & Rehen, S. K. (2022). Induced pluripotent stem cell-derived astrocytes from patients with schizophrenia exhibit an inflammatory phenotype that affects vascularization. *Molecular Psychiatry*, *28*, 871–882 <https://doi.org/10.1038/s41380-022-01830-1>
- Tsuboi, K., Uyama, T., Okamoto, Y., & Ueda, N. (2018). Endocannabinoids and related

- N-acylethanolamines: Biological activities and metabolism. *Inflammation and Regeneration*, 38, 28. <https://doi.org/10.1186/s41232-018-0086-5>
- Tsugawa, H., Cajka, T., Kind, T., Ma, Y., Higgins, B., Ikeda, K., Kanazawa, M., Vandergheynst, J., Fiehn, O., & Arita, M. (2015). MS-DIAL: Data-independent MS/MS deconvolution for comprehensive metabolome analysis. *Nature Methods*, 12(6), 523–526. <https://doi.org/10.1038/nmeth.3393>
- Ueda, N., Puffenbarger, R. A., Yamamoto, S., & Deutsch, D. G. (2000). Fatty acid amide hydrolase (FAAH). *Chemistry and Physics of Lipids*, 108, 107–121. <https://doi.org/10.1016/B978-008055232-3.64098-X>
- Urquhart, P., Nicolaou, A., & Woodward, D. F. (2015). Endocannabinoids and their oxygenation by cyclo-oxygenases, lipoxygenases and other oxygenases. *Biochimica et Biophysica Acta - Molecular and Cell Biology of Lipids*, 1851(4), 366–376. <https://doi.org/10.1016/j.bbalip.2014.12.015>
- van der Stoep, M., Korporaal, S. J. A., & Van Eck, M. (2014). High-density lipoprotein as a modulator of platelet and coagulation responses. *Cardiovascular Research*, 103(3), 362–371. <https://doi.org/10.1093/cvr/cvu137>
- Venn-Watson, S., Lumpkin, R., & Dennis, E. A. (2020). Efficacy of dietary odd-chain saturated fatty acid pentadecanoic acid parallels broad associated health benefits in humans: could it be essential? *Scientific Reports*, 10, 8161. <https://doi.org/10.1038/s41598-020-64960-y>
- Vessecchi, R., Lopes, N. P., Gozzo, F. C., Dörr, F. A., Murgu, M., Lebre, D. T., Abreu, R., Bustillos, O. V., & Riveros, J. M. (2010). Nomenclaturas de espectrometria de massas em língua portuguesa. *Química Nova*, 34(10), 18875–1877.
- Vestri, H. S., Maianu, L., Moellering, D. R., & Garvey, W. T. (2007). Atypical Antipsychotic Drugs Directly Impair Insulin Action in Adipocytes: Effects on Glucose Transport, Lipogenesis, and Antilipolysis. *Neuropsychopharmacology*, 32, 765–772. <https://doi.org/10.1038/sj.npp.1301142>
- Viant, M. R., Kurland, I. J., Jones, M. R., & Dunn, W. B. (2017). How close are we to complete annotation of metabolomes? *Current Opinion in Chemical Biology*, 36, 64–69. <https://doi.org/10.1016/j.cbpa.2017.01.001>

- Vieta, E., Berk, M., Schulze, T. G., Carvalho, A. F., Suppes, T., Calabrese, J. R., Gao, K., Miskowiak, K. W., & Grande, I. (2018). Bipolar disorders. *Nature Reviews Disease Primers*, 4, 18008. <https://doi.org/10.1038/nrdp.2018.8>
- Waluk, D. P., Battistini, M. R., Dempsey, D. R., Farrell, E. K., Jeffries, K. A., Mitchell, P., Hernandez, L. W., McBride, J. C., Merkler, D. J., & Hunt, M. C. (2014). Mammalian Fatty Acid Amides of the Brain and CNS. In *Omega-3 Fatty Acids in Brain and Neurological Health* (pp. 87–107). Elsevier Inc. <https://doi.org/10.1016/b978-0-12-410527-0.00009-0>
- Wang, N., & Tall, A. R. (2016). Cholesterol in platelet biogenesis and activation. *Blood*, 127(16), 1949–1953. <https://doi.org/10.1182/blood-2016-01-631259>
- Wanichthanarak, K., Fahrman, J. F., & Grapov, D. (2015). Genomic , Proteomic , and Metabolomic Data Integration Strategies. *Biomarker Insights*, 10(S4), 1–6. <https://doi.org/10.4137/BMI.S29511>
- Ward, K. M., Yeoman, L., Mchugh, C., Kraal, A. Z., Stephanie, A., Rothberg, A. E., Karnovsky, A., Das, A. K., Ellingrod, V. L., & Stringer, K. A. (2018). Atypical Antipsychotic Exposure May Not Differentiate Metabolic Phenotypes of Patients with Schizophrenia. *Pharmacotherapy*, 38(6), 638–650. <https://doi.org/10.1002/phar.2119>
- West, P. R., Amaral, D. G., Bais, P., Smith, A. M., Egnash, L. A., Ross, M. E., Palmer, J. A., Fontaine, B. R., Conard, K. R., Corbett, B. A., Cezar, G. G., Donley, E. L. R., & Burrier, R. E. (2014). Metabolomics as a tool for discovery of biomarkers of autism spectrum disorder in the blood plasma of children. *PLoS ONE*, 9(11), e112445. <https://doi.org/10.1371/journal.pone.0112445>
- Wishart, D. S., Feunang, Y. D., Marcu, A., Guo, A. C., Liang, K., Vázquez-Fresno, R., Sajed, T., Johnson, D., Li, C., Karu, N., Sayeeda, Z., Lo, E., Assempour, N., Berjanskii, M., Singhal, S., Arndt, D., Liang, Y., Badran, H., Grant, J., ... Scalbert, A. (2018). HMDB 4.0: The human metabolome database for 2018. *Nucleic Acids Research*, 46(D1), D608–D617. <https://doi.org/10.1093/nar/gkx1089>
- Wolosker, H., Blackshaw, S., & Snyder, S. H. (1999). Serine racemase: A glial enzyme synthesizing D-serine to regulate glutamate-N-methyl-D-aspartate neurotransmission. *PNAS*, 96(23), 13409–13414.

<https://doi.org/10.1073/pnas.96.23.13409>

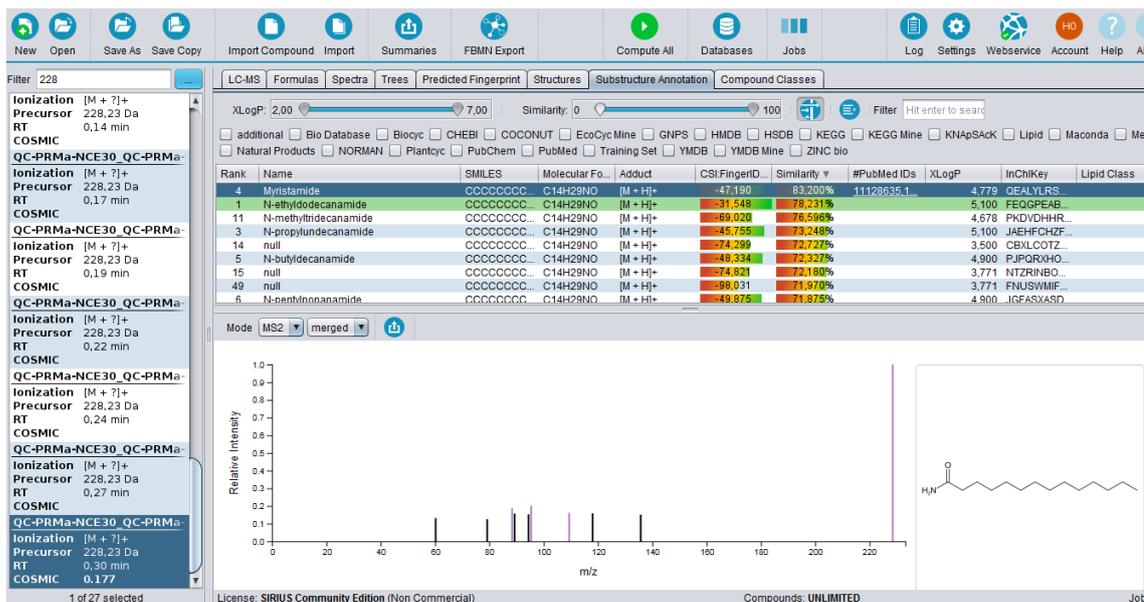
- Xuan, J., Pan, G., Qiu, Y., Yang, L., Su, M., Chen, J., Feng, G., Fang, Y., Jia, W., Xing, Q., & He, L. (2011). Metabolomic profiling to identify potential serum biomarkers for schizophrenia and risperidone action. *Journal of Proteome Research*, 10, 5433–5443. <https://doi.org/10.1021/pr2006796>
- Yang, X., Sun, L., Zhao, A., Hu, X., Qing, Y., Jiang, J., Yang, C., Xu, T., Wang, P., Liu, J., Zhang, J., He, L., Jia, W., & Wan, C. (2017). Serum fatty acid patterns in patients with schizophrenia: a targeted metabolomics study. *Translational Psychiatry*, 7(7), e1176. <https://doi.org/10.1038/tp.2017.152>
- Yang, Xuhan, Li, M., Jiang, J., Hu, X., Qing, Y., Sun, L., Yang, T., Wang, D., Cui, G., Gao, Y., Zhang, J., Li, X., Shen, Y., Qin, S., & Wan, C. (2021). Dysregulation of phospholipase and cyclooxygenase expression is involved in Schizophrenia. *EBioMedicine*, 64, 103239. <https://doi.org/10.1016/j.ebiom.2021.103239>
- Yoshimi, N., Futamura, T., Kakumoto, K., Salehi, A. M., Sellgren, C. M., Holmén-Larsson, J., Jakobsson, J., Pålsson, E., Landén, M., & Hashimoto, K. (2016). Blood metabolomics analysis identifies abnormalities in the citric acid cycle, urea cycle, and amino acid metabolism in bipolar disorder. *BBA Clinical*, 5, 151–158. <https://doi.org/10.1016/j.bbacli.2016.03.008>
- Zhu, X., Chen, Y., & Subramanian, R. (2014). Comparison of information-dependent acquisition, SWATH, and MSAll techniques in metabolite identification study employing ultrahigh-performance liquid chromatography-quadrupole time-of-flight mass spectrometry. *Analytical Chemistry*, 86(2), 1202–1209. <https://doi.org/10.1021/ac403385y>
- Zierer, J., Menni, C., Kastenmu, G., & Spector, T. D. (2015). Integration of 'omics' data in aging research: from biomarkers to systems biology. *Aging Cell*, 14, 933–944. <https://doi.org/10.1111/accel.12386>

**Anexos**

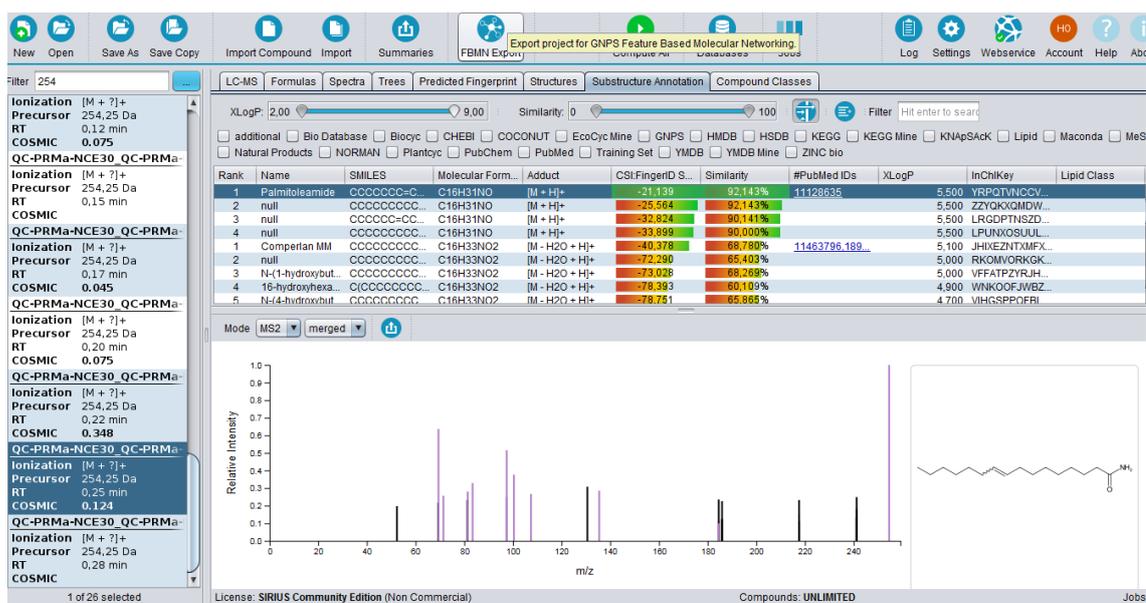
**Anexo 1.** Tabela de curadoria dos dados de metabolômica contendo as identificações da FiehnLib, GMD, o tempo de retenção experimental (Average Rt) e o tempo de retenção anotado na biblioteca FiehnLib (Rt FiehnLib).

<b><u>Alignment ID</u></b>	<b><u>Average Rt (min)</u></b>	<b><u>Rt FiehnLib (min)</u></b>	<b><u>Metabolite name</u></b>	<b><u>GOLM Match</u></b>
359	17.394	17.625	D-glucose 2	Galactitol(6TMS)
426	20.559	20.675	stearic acid	OK
405	19.642	19.8	heptadecanoic acid	OK
394	18.746	18.846	palmitic acid	OK
268	13.26	17.295	acetylisatin 5	no matches
342	16.733	16.887	myristic acid	OK
153	9	9.041	2-ketoisocaproic acid 2	OK
316	15.338	15.614	L-(-)-fucose 1	OK
454	22.206	22.367	arachidic acid	OK
420	20.258	20.399	linoleic acid	OK
299	14.591	14.789	lauric acid	OK
171	9.565	9.594	benzoic acid	OK
250	12.726	16.365	L-allothreonine 2	Unknown Carboxilic acid
403	19.514	19.663	Methyl Stearate	OK
321	15.655	29.185	tetratriacontane	Octadecane
193	10.243	10.456	glycine	OK
180	9.772	9.941	glycerol	OK
352	17.195	17.426	D-glucose 1	OK
118	8.216	8.211	p-cresol	OK
133	8.515	8.541	2-ketoisocaproic acid 1	OK
390	18.526	18.728	palmitoleic acid	OK
354	17.204	17.278	D-allose 1	D-Glucose
341	16.707	17.584	talose 2	Unknown sugar
406	19.661	19.44	methyl oleate	OK
422	20.3	20.504	oleic acid	OK
252	12.788	12.954	D-threitol	OK
489	24.433	24.915	maltose 2	OK
374	17.95	18.089	3-indoleacetic acid	OK
336	16.363	16.615	citric acid	OK
208	10.824	10.94	fumaric acid	OK
397	19.086	17.245	allo-inositol	Myo-innositol
199	10.524	10.735	glyceric acid	OK
312	15.107	17.766	sedoheptulose anhydride monohydrate	Glucose, 1,6-anhydro-, beta- (3TMS)
344	16.91	17.011	tagatose 1	Unknown sugar

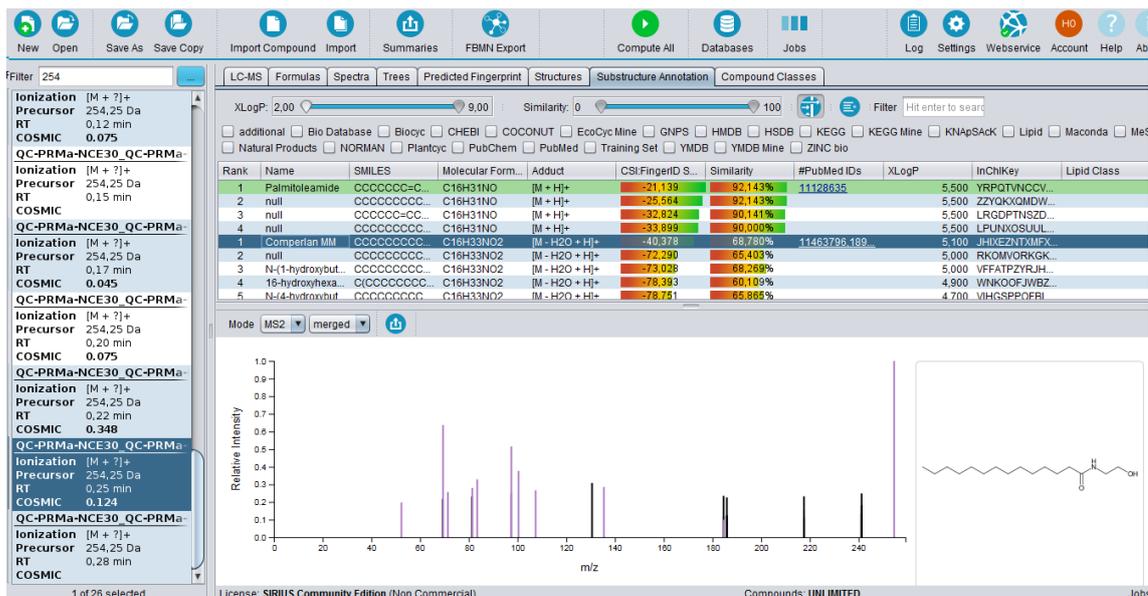
217	11.235	11.464	L-threonine 2	OK
145	8.797	9.041	2-ketoisocaproic acid 1	OK
302	14.741	22.224	5-hydroxy-L-tryptophan 1	Homoserine (4TMS)
105	8.078	8.047	3-hydroxypyridine	OK
47	6.999	6.851	L-(+) lactic acid	OK
514	27.394	27.555	cholesterol	OK
173	9.595	9.706	L-serine 1	OK
347	17.019	17.288	fructose 2	OK
393	18.741	9.81	caprylic acid	palmitic acid
288	14.279	12.359	L-homoserine 2	Phenylalanine (2TMS)
440	21.393	18.985	glucoheptonic acid 1	Unknown lactone
485	24.07	24.386	lactose 1	OK
170	9.554	9.599	urea	OK
240	12.341	14.885	DL-threo-beta-hydroxyaspartic acid	Unknown alpha aminoacid
327	15.941	14.955	N-acetyl-L-aspartic acid 2	Unknown carboxilic acid derivative
181	9.787	9.966	phosphoric acid	OK
417	20.228	21.868	N-acetyl-D-tryptophan 2	Tryptophan (2TMS)
407	19.67	19.82	uric acid 2	No matches on GOLM, consistent match on MS-DIAL
346	16.957	23.163	dioctyl phthalate	Unknown
120	8.269	8.298	L-leucine 1	Isoleucine (1TMS), by RT
463	22.659	15.377	acetol 3	Unknown 1,2-diol
231	11.815	12.002	aspartic acid 1	OK
49	7.044	13.247	4-isopropylbenzoic acid	Unknown carboxilic acid
196	10.43	10.509	succinic acid	OK
89	7.805	7.852	2-hydroxybutyric acid	OK
229	11.704	11.835	L-methionine 1	OK
35	6.772	7.883	oxalic acid	Glycolic acid
399	19.247	17.521	D-allose 2	Unknown primary alcohol
188	10.078	10.224	L-threonine 1	OK
135	8.562	8.576	DL-isoleucine 1	OK
262	13.037	13.218	L-pyroglutamic acid	OK
119	8.265	8.919	malonic acid 1	Unknown carboxilic acid derivative
260	12.982	13.207	aspartic acid 2	OK
130	8.476	7.504	2-ketobutyric acid	Unknown
512	27.283	27.379	alpha tocophereol	OK



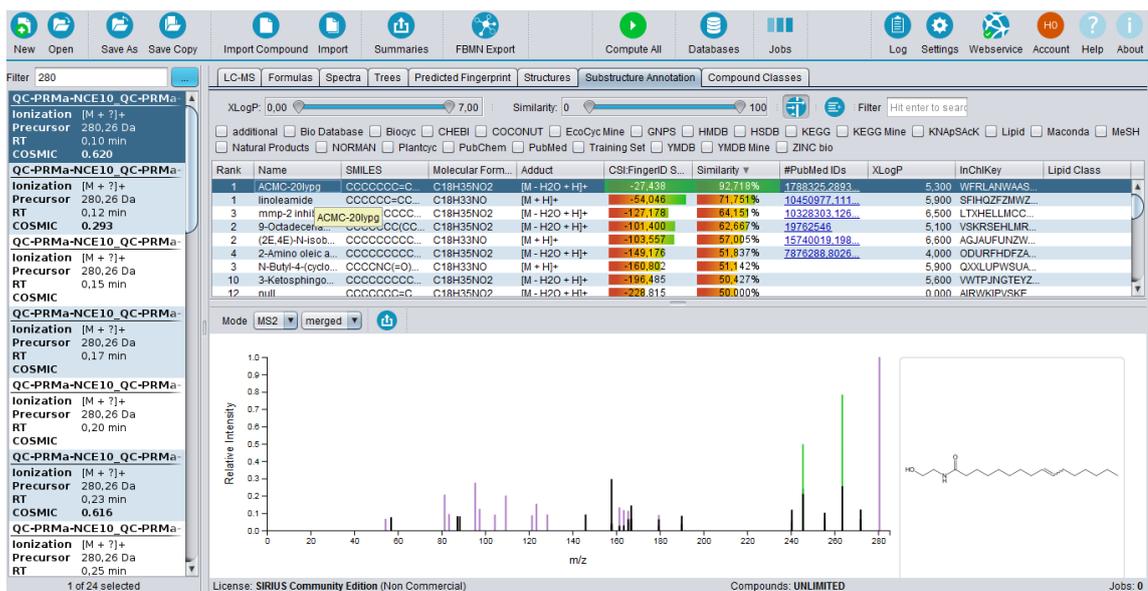
Anexo 2. Espectro de fragmentação do íon de  $m/z$  228,2234.



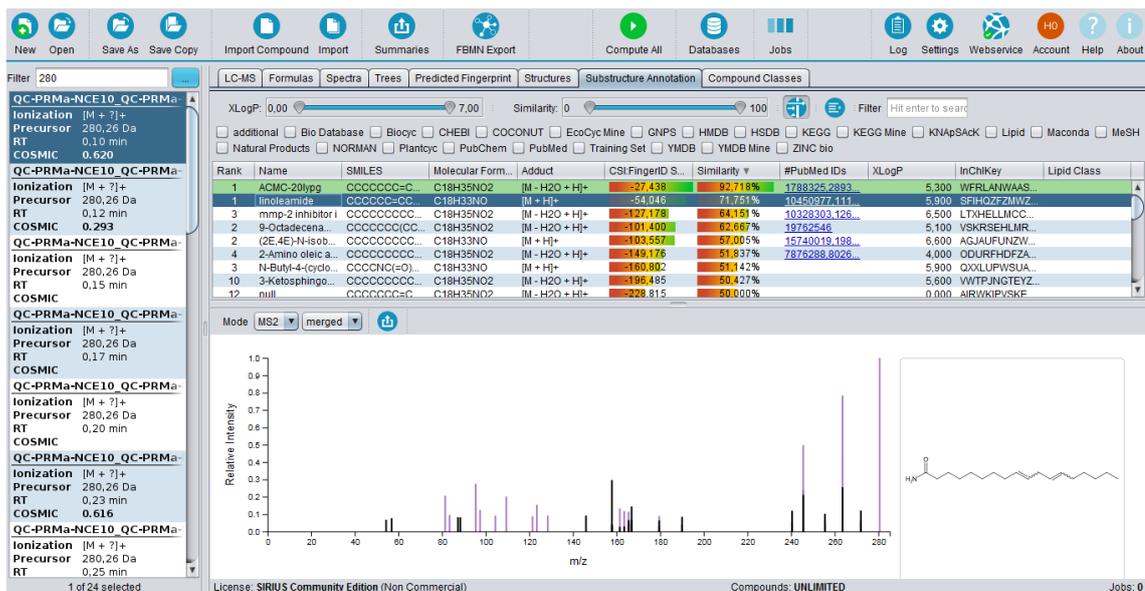
Anexo 3. Espectro de fragmentação do íon de  $m/z$  254,2474, referente à palmitoleiamida.



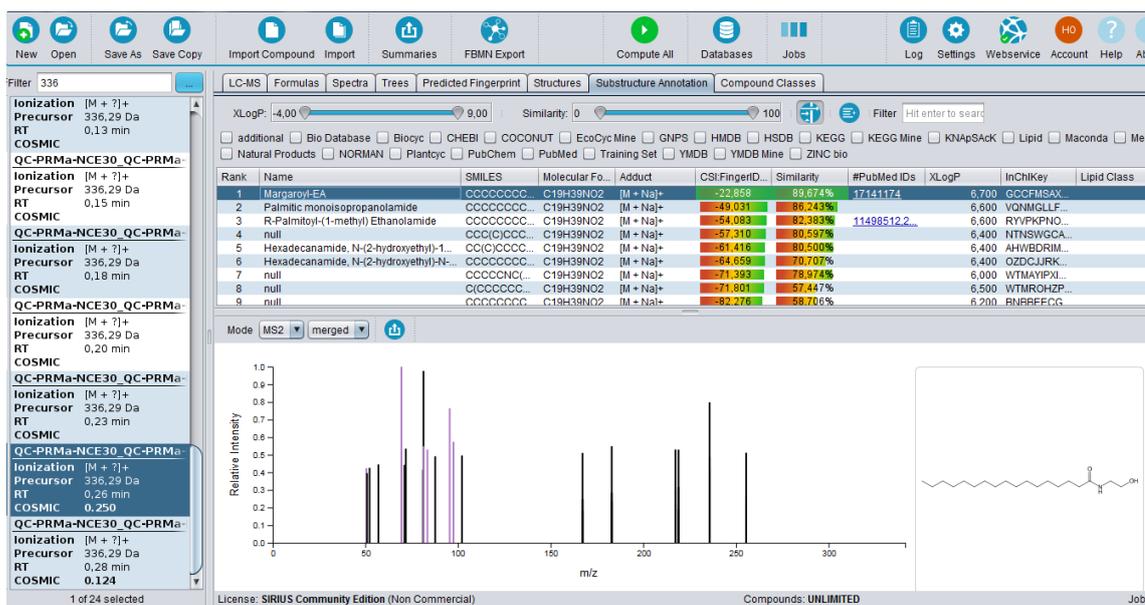
Anexo 4. Espectro de fragmentação do íon de  $m/z$  254,2474, referente à miristoiletanolamina.



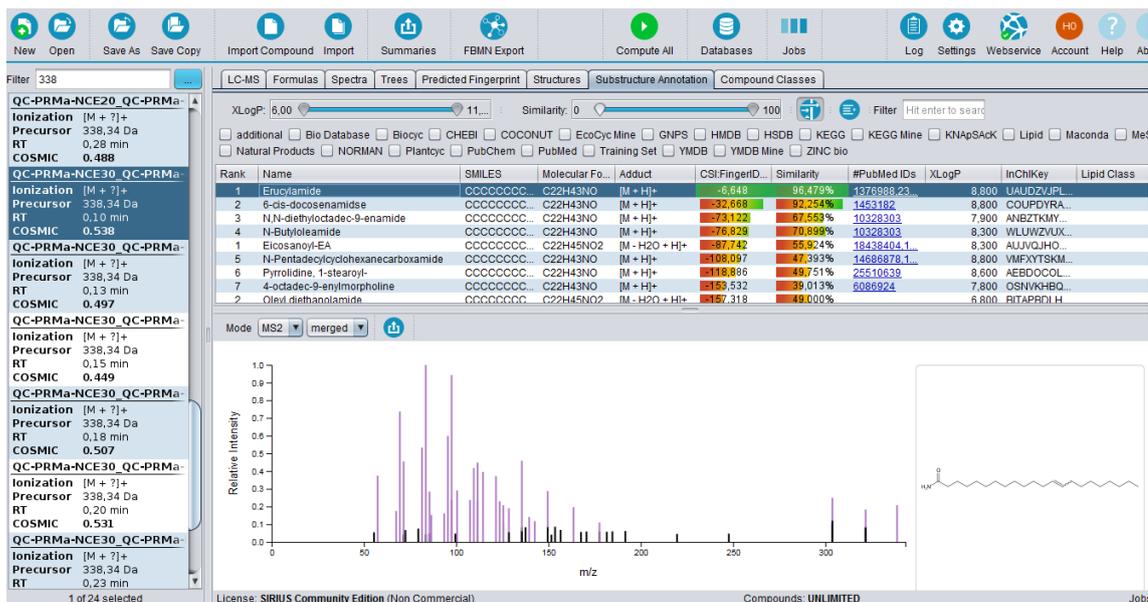
Anexo 5. Espectro de fragmentação do íon de  $m/z$  280,2629, referente à palmitoleiletanolamina.



Anexo 6. Espectro de fragmentação do íon de  $m/z$  280,2629, referente à linoleiamida.



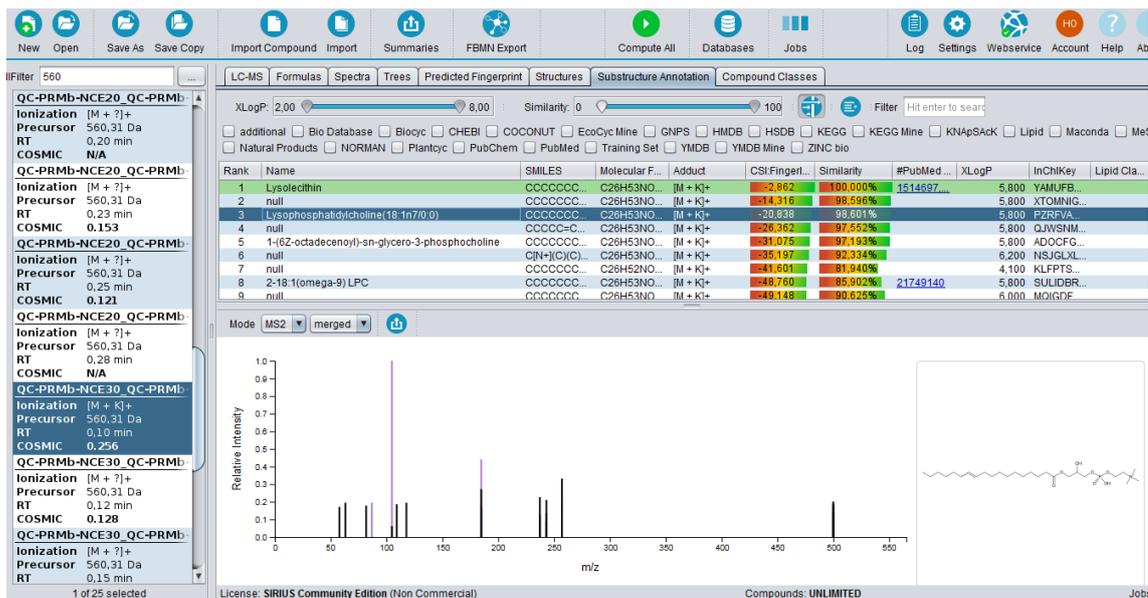
Anexo 7. Espectro de fragmentação do íon de  $m/z$  336,2890, referente à margaroletanolamina.



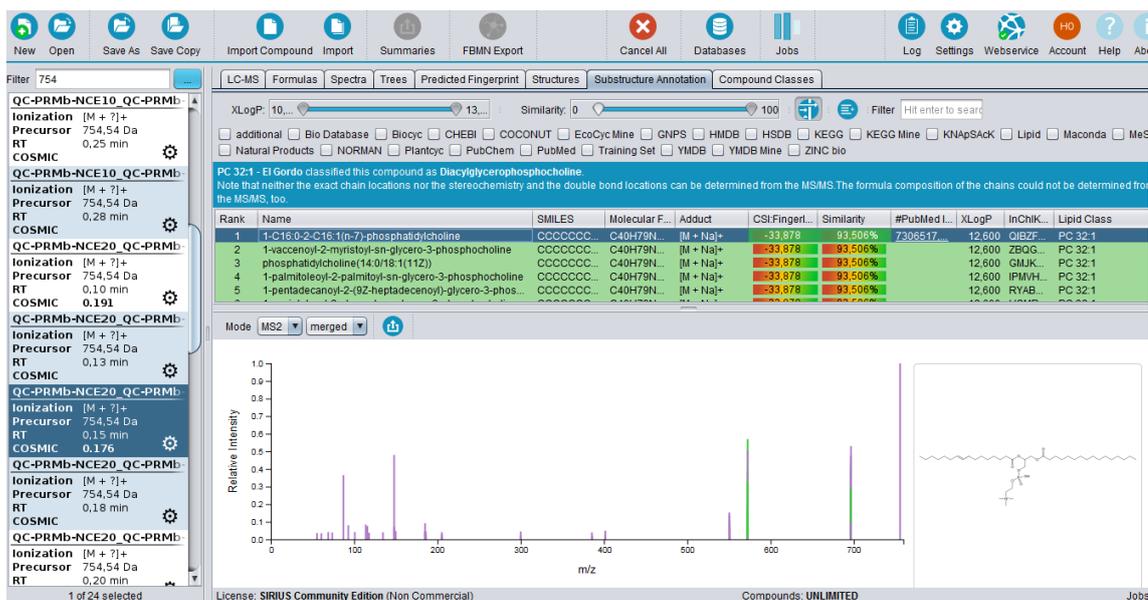
Anexo 8. Espectro de fragmentação do íon de  $m/z$  338,3401, referente à erucamida.



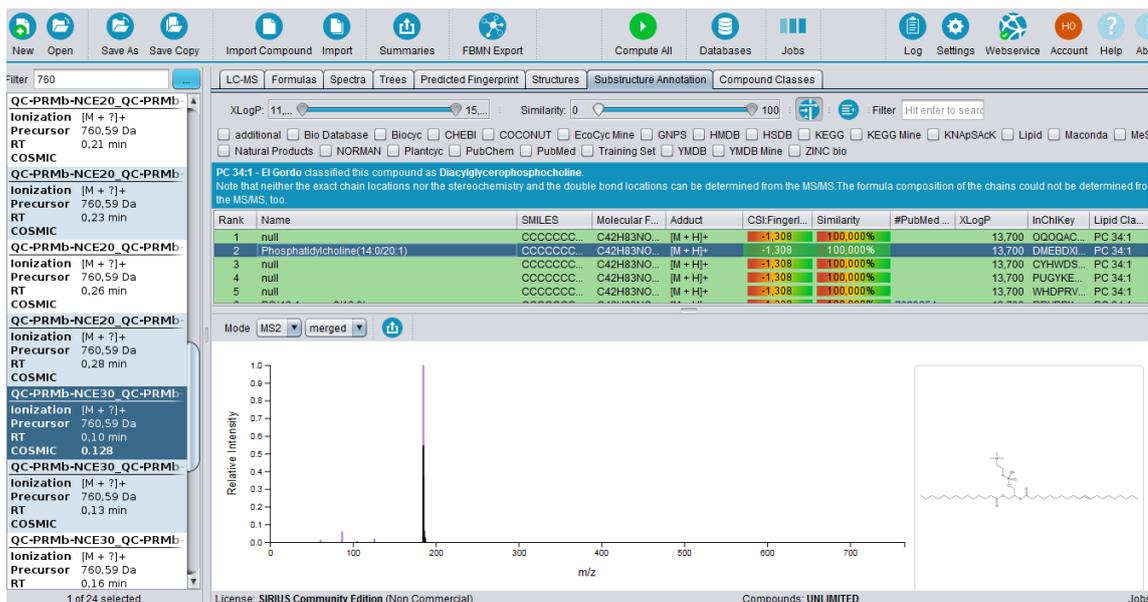
Anexo 9. Espectro de fragmentação do íon de  $m/z$  524,3706, referente à PC(18:0).



Anexo 10. Espectro de fragmentação do íon de  $m/z$  560,3111, referente à PC(18:1).



Anexo 11. Espectro de fragmentação do íon de  $m/z$  754,5359, referente à PC(32:1).



Anexo 12. Espectro de fragmentação do íon de  $m/z$  760,5867, referente à PC(34:1).



Anexo 13. Espectro de fragmentação do íon de  $m/z$  858,5991, referente à PC(40:5).

**Anexo 14.** Parecer do comitê de ética em pesquisa referente ao protocolo 775/2010



FACULDADE DE CIÊNCIAS MÉDICAS  
COMITÊ DE ÉTICA EM PESQUISA

[www.fcm.unicamp.br/pesquisa/etica/index.html](http://www.fcm.unicamp.br/pesquisa/etica/index.html)

CEP: 12/11/10  
(Grupo II)

**PARECER CEP:** Nº 775/2010 (Este nº deve ser citado nas correspondências referente a este projeto).  
**CAAE:** 0605.0.146.000-10

#### I - IDENTIFICAÇÃO:

**PROJETO:** "APLICAÇÃO DE ESTRATÉGIAS PROTEÔMICAS E METALÔMICAS NA IDENTIFICAÇÃO DE POSSÍVEIS BIOMARCADORES PARA O TRANSTORNO AFETIVO BIPOLAR".

**PESQUISADOR RESPONSÁVEL:** Herbert de Sousa Barbosa

**INSTITUIÇÃO:** Hospital das Clínicas/UNICAMP

**APRESENTAÇÃO AO CEP:** 11/08/2010

**APRESENTAR RELATÓRIO EM:** 12/11/11 (O formulário encontra-se no site acima).

#### II - OBJETIVOS

Investigar alterações no proteoma e no metaloma do soro sanguíneo de indivíduos com transtorno afetivo bipolar, com o intuito de identificar possíveis biomarcadores para essa doença mental.

#### III - SUMÁRIO

Serão avaliados 100 sujeitos exclusivamente neste centro, todos maiores de idade, compreendendo grupo vulnerável (portadores de transtorno mental). Os procedimentos incluem a análise de marcadores em amostras de sangue de sujeitos com transtorno afetivo bipolar (TAB) e o preenchimento de um questionário de coleta de dados. Serão formados três grupos de sujeitos, um com pacientes diagnosticados com essa condição e que não estejam em tratamento farmacológico, outro de pacientes em uso de lítio e o terceiro de pacientes em tratamento com outros fármacos que não o lítio. Serão formados, também quatro grupos controle: um de pacientes com esquizofrenia, outro de pacientes sem TAB ou esquizofrenia que fazem uso de lítio por outra indicação, um terceiro de parentes em primeiro grau dos indivíduos investigados e sem transtorno mental e o quarto de indivíduos não afetados que residam próximo aos pacientes com TAB. Serão excluídas as amostras de indivíduos que possuírem certas doenças (cânceres, AIDS, doenças hepáticas e doenças metabólicas e endocrinológicas em geral, como diabetes e hipo ou hipertireoidismo) (pg. 22). O cronograma prevê duração de quatro semestres.

#### IV - COMENTÁRIOS DOS RELATORES

Após analisar as respostas às pendências encaminhadas em 27/10/10, todas as questões solicitadas pelo Comitê de Ética em Pesquisa foram respondidas, com isso o projeto de pesquisa encontra-se adequadamente redigido e de acordo com a Resolução CNS/MS 196/96 e suas complementares, bem como o novo Termo de Consentimento Livre e Esclarecido.

Comitê de Ética em Pesquisa - UNICAMP  
Rua: Tessália Vieira de Camargo, 126  
Caixa Postal 6111  
13083-887 Campinas - SP

FONE (019) 3521-8936  
FAX (019) 3521-7187  
cep@fcm.unicamp.br



## V - PARECER DO CEP

O Comitê de Ética em Pesquisa da Faculdade de Ciências Médicas da UNICAMP, após acatar os pareceres dos membros-relatores previamente designados para o presente caso e atendendo todos os dispositivos das Resoluções 196/96 e complementares, resolve aprovar sem restrições o Protocolo de Pesquisa, o Termo do Consentimento Livre e Esclarecido, bem como todos os anexos incluídos na pesquisa supracitada.

O conteúdo e as conclusões aqui apresentados são de responsabilidade exclusiva do CEP/FCM/UNICAMP e não representam a opinião da Universidade Estadual de Campinas nem a comprometem.

## VI - INFORMAÇÕES COMPLEMENTARES

O sujeito da pesquisa tem a liberdade de recusar-se a participar ou de retirar seu consentimento em qualquer fase da pesquisa, sem penalização alguma e sem prejuízo ao seu cuidado (Res. CNS 196/96 – Item IV.1.f) e deve receber uma cópia do Termo de Consentimento Livre e Esclarecido, na íntegra, por ele assinado (Item IV.2.d).

Pesquisador deve desenvolver a pesquisa conforme delineada no protocolo aprovado e descontinuar o estudo somente após análise das razões da descontinuidade pelo CEP que o aprovou (Res. CNS Item III.1.z), exceto quando perceber risco ou dano não previsto ao sujeito participante ou quando constatar a superioridade do regime oferecido a um dos grupos de pesquisa (Item V.3.).

O CEP deve ser informado de todos os efeitos adversos ou fatos relevantes que alterem o curso normal do estudo (Res. CNS Item V.4.). É papel do pesquisador assegurar medidas imediatas adequadas frente a evento adverso grave ocorrido (mesmo que tenha sido em outro centro) e enviar notificação ao CEP e à Agência Nacional de Vigilância Sanitária – ANVISA – junto com seu posicionamento.

Eventuais modificações ou emendas ao protocolo devem ser apresentadas ao CEP de forma clara e sucinta, identificando a parte do protocolo a ser modificada e suas justificativas. Em caso de projeto do Grupo I ou II apresentados anteriormente à ANVISA, o pesquisador ou patrocinador deve enviá-las também à mesma junto com o parecer aprovatório do CEP, para serem juntadas ao protocolo inicial (Res. 251/97, Item III.2.e).

Relatórios parciais e final devem ser apresentados ao CEP, de acordo com os prazos estabelecidos na Resolução CNS-MS 196/96.

## VI - DATA DA REUNIÃO

Homologado na VIII Reunião Ordinária do CEP/FCM, em 24 de agosto de 2010.

**Prof. Dr. Carlos Eduardo Steiner**  
PRESIDENTE do COMITÊ DE ÉTICA EM PESQUISA  
FCM / UNICAMP

## Anexo 15. Parecer do comitê de ética em pesquisa referente ao protocolo CAPPesq N° 943.883

Anexo 3



### PARECER CONSUBSTANCIADO DO CEP

#### DADOS DO PROJETO DE PESQUISA

**Título da Pesquisa:** Instituto Nacional de Biomarcadores em Neuropsiquiatria (INBION)

**Pesquisador:** Wagner Farid Gattaz

**Área Temática:** Genética Humana:

(Trata-se de pesquisa envolvendo Genética Humana que não necessita de análise ética por parte da CONEP);

**Versão:** 5

**CAAE:** 66092117.0.1001.0068

**Instituição Proponente:** Hospital das Clínicas da Faculdade de Medicina da USP

**Patrocinador Principal:** FUNDAÇÃO DE AMPARO A PESQUISA DO ESTADO DE SÃO PAULO  
MINISTERIO DA CIENCIA, TECNOLOGIA E INOVACAO

#### DADOS DO PARECER

**Número do Parecer:** 2.450.834

#### Apresentação do Projeto:

O projeto é apresentado como um protocolo que reúne vários projetos de natureza multidisciplinar e multiinstitucional com a co-participação da UNICAMP, que prevê a coleta de material biológico dos participantes para o desenvolvimento de projetos específicos, que interagem por meio da comparação dos diferentes resultados, na busca de substratos biológicos das doenças neuropsiquiátricas como pré-requisito para o desenvolvimento de terapias mais eficazes e, sobretudo, de estratégias preventivas através do diagnóstico precoce.

Grupos de pacientes das diferentes categorias diagnósticas (Esquizofrenia, o Transtorno Bipolar, Doença de Alzheimer, indivíduos com Comprometimento Cognitivo Leve, idosos com queixas de memória) serão estudados simultaneamente em 6 dimensões: neuroquímica, neuroimagem, neuromodulação, neurocognição/psicopatologia, genômica e proteômica. Os achados nestes grupos serão comparados com um grupo controle de indivíduos saudáveis. Adicionalmente, serão também investigados nestas 6 dimensões indivíduos com transtornos subclínicos, que apresentam sintomas psiquiátricos esporádicos (por exemplo alucinações ou ideias paranóides) sem todavia desenvolver a doença psiquiátrica completa. Este grupo intermediário permitirá investigar estádios de transição para esclarecer o continuum entre saúde e doença nas dimensões clínicas e neurodiagnósticas.

**Endereço:** Rua Ovídio Pires de Campos, 225 5º andar  
**Bairro:** Cerqueira Cesar **CEP:** 05.403-010  
**UF:** SP **Município:** SAO PAULO  
**Telefone:** (11)2661-7585 **Fax:** (11)2661-7585 **E-mail:** cappesq.adm@hc.fm.usp.br



Continuação do Parecer: 2.450.834

#### **Objetivo da Pesquisa:**

O projeto conforme apresentado reúne diferentes projetos voltados para o estudo de aspectos etiológicos comuns e marcadores diferenciais entre a Esquizofrenia, o Transtorno Bipolar e a Doença de Alzheimer. Esses pacientes serão estudados simultaneamente em 6 dimensões, a saber: 1. Neuroquímica; 2. Neuroimagem; 3. Neuromodulação; 4. neurocognição e psicopatologia; 5. genômica; 6. Proteômica, sendo que cada uma delas apresenta objetivos específicos.

#### **Avaliação dos Riscos e Benefícios:**

Os desconfortos e riscos esperados são específicos para cada módulo de atuação. De maneira geral:- para os processos da coleta de sangue o risco é o de apresentar dor, desconforto ou hematoma.- Com relação à coleta de líquor, o risco envolve apresentação de dor de cabeça ao levantar, sensações cutâneas subjetivas (ex., frio, calor, formigamento, pressão etc.), dor local, infecção, meningite, sangramento e excepcionalmente óbito.- Para o procedimento da Estimulação magnética transcraniana, a colocação de touca de tecido na cabeça para mapeamento magnético, que não traz dor, mas pode trazer algum desconforto. Dor de cabeça é possível acontecer, mas é de leve intensidade e incomum. Pode ocorrer sensação de contração muscular nos membros durante o exame.- Durante o exame de ressonância magnética, o desconforto é o do tempo em que é preciso ficar no aparelho e o barulho forte que o aparelho faz.

- Nos exames de PET ou SPECT, há o desconforto de uma picada de injeção, para que seja injetado o material do exame e a necessidade de ficar imóvel. Em relação aos benefícios, não há benefício direto para o participante; trata-se de estudo experimental. É possível, mas não garantido, que os pacientes com estas doenças possam ser beneficiados pelos resultados deste estudo no futuro, bem como seus familiares. O estudo visa maior esclarecimento dos mecanismos biológicos responsáveis pelo aparecimento de transtorno psiquiátricos. De posse destes conhecimentos, espera-se conseguir, no futuro, melhores estratégias para o tratamento do referido transtorno.

#### **Comentários e Considerações sobre a Pesquisa:**

O projeto conforme apresentado reúne iniciativas de projetos múltiplos e fará a coleta de material biológico e outros dados dos pacientes com diferentes doenças psiquiátricas. É um projeto de 6 anos que fará acompanhamento desses pacientes para estudo da evolução dos respectivos quadros clínicos.

#### **Considerações sobre os Termos de apresentação obrigatória:**

O pesquisador esclareceu que foram acatadas todas as sugestões da CAPPEsq e que a carta encaminhada anteriormente permanecia com o conteúdo completo, apenas com o intuito de

Endereço: Rua Ovídio Pires de Campos, 225 5º andar  
 Bairro: Cerqueira Cesar CEP: 05.403-010  
 UF: SP Município: SAO PAULO  
 Telefone: (11)2661-7585 Fax: (11)2661-7585 E-mail: cappesq.adm@hc.fm.usp.br

## Anexo 3



Continuação do Parecer: 2.450.834

salientar a importância desse tipo de pesquisa no âmbito internacional. Foram encaminhados também o Regulamento do Biorrepositório referente, especificamente, à essa pesquisa e o pedido de encerramento do Projeto 1293/09, juntamente com o relatório final.

**Recomendações:**

Não há.

**Conclusões ou Pendências e Lista de Inadequações:**

Não há mais pendências.

**Considerações Finais a critério do CEP:**

Em conformidade com a Resolução CNS nº 466/12 – cabe ao pesquisador: a) desenvolver o projeto conforme delineado; b) elaborar e apresentar relatórios parciais e final; c) apresentar dados solicitados pelo CEP, a qualquer momento; d) manter em arquivo sob sua guarda, por 5 anos da pesquisa, contendo fichas individuais e todos os demais documentos recomendados pelo CEP; e) encaminhar os resultados para publicação, com os devidos créditos aos pesquisadores associados e ao pessoal técnico participante do projeto; f) justificar perante ao CEP interrupção do projeto ou a não publicação dos resultados.

**Este parecer foi elaborado baseado nos documentos abaixo relacionados:**

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_858575.pdf	15/12/2017 16:34:49		Aceito
Outros	Carta_RelatorioFinalTematico_.pdf	15/12/2017 16:33:44	Wagner Farid Gattaz	Aceito
Outros	RespostaParecer_2315487.pdf	15/12/2017 16:20:11	Wagner Farid Gattaz	Aceito
Declaração de Manuseio Material Biológico / Biorrepositório / Biobanco	RegulamentoFuncional_Biorrepositorio.pdf	15/12/2017 16:18:45	Wagner Farid Gattaz	Aceito
Outros	adendoprojeto_nov_2017.doc	27/11/2017 13:47:31	Wagner Farid Gattaz	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TCLInBion_nov_2017.docx	27/11/2017 13:47:00	Wagner Farid Gattaz	Aceito
Projeto Detalhado / Brochura	INBioN.doc	14/06/2017 11:31:45	Wagner Farid Gattaz	Aceito

Endereço: Rua Ovídio Pires de Campos, 225 5º andar  
 Bairro: Cerqueira Cesar CEP: 05.403-010  
 UF: SP Município: SAO PAULO  
 Telefone: (11)2661-7585 Fax: (11)2661-7585 E-mail: cappesq.adm@hc.fm.usp.br



Continuação do Parecer: 2.450.834

Investigador	INBioN.doc	14/06/2017 11:31:45	Wagner Farid Gattaz	Aceito
Folha de Rosto	2248_0001folhaderosto.pdf	22/03/2017 14:02:54	Wagner Farid Gattaz	Aceito

**Situação do Parecer:**

Aprovado

**Necessita Apreciação da CONEP:**

Não

SAO PAULO, 20 de Dezembro de 2017

---

**Assinado por:**  
**ALFREDO JOSE MANSUR**  
**(Coordenador)**

**Endereço:** Rua Ovídio Pires de Campos, 225 5º andar  
**Bairro:** Cerqueira Cesar **CEP:** 05.403-010  
**UF:** SP **Município:** SAO PAULO  
**Telefone:** (11)2661-7585 **Fax:** (11)2661-7585 **E-mail:** cappesq.adm@hc.fm.usp.br