



UNIVERSIDADE DE SÃO PAULO



FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

Estudos metabolômicos de espécies brasileiras de Orchidaceae

Camila Capel Godinho

**Ribeirão Preto
2022**

CAMILA CAPEL GODINHO

Estudos Metabolômicos de Espécies Brasileiras de Orchidaceae

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas da Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP para obtenção do Título de Doutor em Ciências.

Área de Concentração: Produtos Naturais e Sintéticos

Orientador: Prof. Dr. Leonardo Gobbo Neto

Versão corrigida da Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas em 29/04/2022. A versão original encontra-se disponível na Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP.

Ribeirão Preto
2022

GODINHO, C.C.	Estudos Metabolômicos de Espécies Brasileiras de Orchidaceae		DOUTORADO FCFRPUSP 2022
------------------	---	--	-------------------------------

FICHA CATALOGRÁFICA

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO, PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Godinho, Camila Capel

Estudos metabolômicos de espécies brasileiras de Orchidaceae. Ribeirão Preto, 2022.

147 p.: il.; 30cm.

Tese de Doutorado, apresentada à Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP – Área de concentração: Produtos Naturais e Sintéticos

Orientador: Gobbo Neto, Leonardo

1.Metabolômica. 2.Orchidaceae. 3.Espectrometria de Massas. 4.Redes Moleculares. 5.Bioatividade. 6.Desreplicação. 7.Antioxidante

FOLHA DE APROVAÇÃO

Nome do autor: Camila Capel Godinho

Título do trabalho: Estudos metabolômicos de espécies brasileiras de Orchidaceae

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciências Farmacêuticas da Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP para obtenção do Título de Doutor em Ciências

Área de Concentração: Produtos Naturais e Sintéticos.

Aprovado em:

Banca Examinadora

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Prof. Dr. _____

Instituição: _____ Assinatura: _____

Aos meus ancestrais pesquisadores, curandeiros e benzedeiros, que sempre acreditaram no potencial da natureza para a cura integral do ser humano.

Agradecimentos

Estes seis anos de doutorado foram muito especiais, sendo dois deles durante a pandemia por COVID-19, devido ao grande crescimento profissional e pessoal. E neste processo, muitas pessoas foram de grande importância na minha vida. A elas, minha gratidão eterna e profunda.

Ao Prof. Dr. Leonardo Gobbo Neto, pela oportunidade de trabalhar sob sua supervisão e pelo crescimento que esse trabalho me proporcionou. Agradeço por compartilhar toda a experiência e conhecimento, e pelo incentivo a cada momento.

Agradeço aos companheiros do Núcleo de Pesquisa em Produtos Naturais e Sintéticos (NPPNS). Aos professores, pelos momentos de troca de conhecimento e saberes em nossas reuniões, cafés e bares, pois são nesses momentos de descontração que a criatividade e a experiência fluem. E em especial ao Prof. Dr. Norberto Peporine Lopes, por todo apoio, por disponibilizar seus recursos, tempo, conhecimento e conselhos quando mais precisei. Aos técnicos, Tomaz, Jacqueline e Izabel Cristina, por todo carinho, conhecimento e disposição em me ajudar quando eu mais precisava. A todos os pós-graduandos e ICs do grupo, mesmo que separados pelos blocos o carinho, companheirismo e o respeito sempre estiveram presentes.

Aos meus amigos e companheiros de laboratório, pelos grandes momentos de alegria, festas, gargalhadas, e também pelo apoio nos momentos de desespero e tristeza, por toda ajuda e companheirismo, disposição e crescimento pessoal impagáveis: Gabriela, Eduarda, Paula, Alan, Anelize, Ricardo, Arthur, Daniele, Matheus, Ariane, Pâmela, Maíra, Guilherme, Taynara, Marília. Vocês moram no meu coração.

À Faculdade de Ciências Farmacêuticas de Ribeirão Preto (FCFRPUSP), pela oportunidade e apoio durante todo o processo. Pelo apoio aos estudantes durante toda a pandemia, seja financeiro, estrutural ou psicológico. O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Código de Financiamento 001, e por ele agradeço.

Aos meus companheiros de coração, Eduarda, Gabriela, Arthur, Paula e Alan, pelos quais tenho gratidão infinita pelo carinho, companheirismo em todas as horas imagináveis e apoio incondicional.

Agradeço à minha família do coração e do Chão de Estrelas, que me acompanharam nos momentos de glória e nos de dificuldade na minha caminhada. A força de vocês e o amor me fizeram mais forte para ver a profissional, pesquisadora e doutora que hoje eu sou.

Agradeço, por fim, aos meus pais, Fátima e José, meus irmãos Carolina, Lucas e Pedro, e à toda minha família, parentes e ancestrais, base do meu ser e força pela qual posso estar aqui neste momento. Pelo carinho, apoio, amor incondicional, pela criação e educação que me deram e apoio em todas as escolhas da minha vida. Não há palavras que possa expressar a minha gratidão.

RESUMO

GODINHO, C.C. **Estudos metabolômicos de espécies brasileiras de Orchidaceae**. 2022. 147f. Tese (Doutorado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2022.

A metabolômica propõe um estudo integrativo, utilizando ferramentas clássicas e inovadoras para a mineração de dados e rápida anotação de compostos. A família Orchidaceae é uma das mais fascinantes devido à sua diversidade ecológica, mas também possui vasto uso medicinal no sudoeste asiático, Índia, China, Japão, Europa, África e Américas. No entanto, há poucos estudos sobre as orquídeas brasileiras, sendo interessante trazer um novo olhar para essas em um contexto biológico, ecológico e farmacológico. Portanto, este trabalho teve como objetivo aplicar ferramentas metabolômicas inovadoras à 130 extratos de espécies de Orchidaceae com o intuito de investigar sua diversidade de metabólitos secundários e avaliar o potencial bioativo de seus extratos quanto a ação moduladora da produção de espécies reativas de oxigênio (ERO) por neutrófilos humanos, atividade antioxidante e citotoxicidade, correlacionando-os com os dados químicos obtidos e propondo possíveis classes e substâncias relacionadas às atividades. Os extratos de média-baixa e média-alta polares foram analisados por GC-MS e LC-MS/MS, respectivamente. A partir desses dados, foi realizado um estudo de otimização do processamento de dados e da construção da rede molecular, resultando em redes robustas, anotações de alta qualidade e mais de 1700 compostos anotados. Além disso, a rede molecular demonstrou uma tendência de distribuição de metabólitos em grupos taxonômicos, filogenéticos e filogenômicos, a qual foi também observada por HCA e PCA à nível de subfamília e entre as tribos Cymbidieae e Epidendreae, e nas análises de enriquecimento de classe, demonstrando as classes mais abundantes em cada grupo. Os estudos sobre o potencial ativo dos extratos analisados resultaram em 27 espécies com atividade antioxidante e 75 com ação modulatória na produção de ERO, sendo que 14 espécies demonstraram ser duplamente ativas. As análises multivariadas indicaram as classes dos flavonóides-7-O-glicosilados, ácidos cumáricos e derivados, cumarinas glicosiladas, cromonas, ácidos hidroxicinâmicos, xantonas, sesquiterpenóides, triterpeno, acilglicosídeos graxos de mono e dissacarídeos e ácidos fenilpirúvicos e derivados como antioxidantes, e as classes dos 2'-Hidroxichalconas, flavanonas 8-preniladas, peptídeos cíclicos, flavonóides -7-O-glicosilados, lignanas glicosiladas, acilglicosídeos graxos de mono e dissacarídeos, quinolinas carboxamidas, diterpenos, sesquiterpenos, dentre outros, como moduladores da produção de ERO por neutrófilos humanos. Este é o primeiro estudo que avalia mais de 100 espécies brasileiras de Orchidaceae, contribuindo para o conhecimento dessa família tão diversa, demonstrando que as orquídeas brasileiras podem ser muito mais que ornamentais, e sim detentoras de possíveis novos fármacos.

Palavras-chave: Metabolômica, Orchidaceae, Rede Molecular, Antioxidante.

ABSTRACT

GODINHO, C.C. **Metabolomics studies of Brazilian species of Orchidaceae**. 2022. 147f. Thesis (Doctoral). School of Pharmaceutical Science of Ribeirão Preto – University de São Paulo, Ribeirão Preto, 2022.

Metabolomics proposes an integrative study, using classic and innovative tools for data mining and rapid annotation of compounds. The Orchidaceae family is one of the most fascinating due to its ecological diversity, but it also has vast medicinal uses in Southeast Asia, India, China, Japan, Europe, Africa and the Americas. However, there are few studies on Brazilian orchids, making it necessary to observe at these in a biological, ecological and pharmacological context. Therefore, this work aimed to apply innovative metabolomic tools to 130 extracts of Orchidaceae species in order to investigate their diversity of secondary metabolites and evaluate the bioactive potential of their extracts in terms of modulating the production of reactive oxygen species (ROS) by human neutrophils, antioxidant activity and cytotoxicity, correlating them with the chemical data obtained and proposing the substances related to the activities. Medium-low and medium-high polar extracts were analyzed by GC-MS and LC-MS/MS, respectively. From these data, an optimization study of the data processing and the construction of the molecular network was carried out, resulting in robust networks, high quality annotations and more than 1700 annotated compounds. Furthermore, the molecular network showed a tendency of metabolite distribution in taxonomic, phylogenetic and phylogenomic groups, which was also observed by HCA and PCA, at the subfamily level and between the Cymbidieae and Epidendreae tribes, and in the enrichment analyses, demonstrating the most abundant classes in each group. Studies on the active potential of the extracts resulted in 27 species with antioxidant activity and 75 with modulatory action in the production of ROS, and 14 species proved to be doubly active. Multivariate statistical analyzes indicated the classes of flavonoids-7-*O*-glycosylated, coumaric acids and derivatives, glycosylated coumarins, chromones, hydroxycinnamic acids, xanthones, sesquiterpenoids, triterpene, mono and disaccharide fatty acylglycosides and phenylpyruvic acids and derivatives as antioxidants, and the classes of 2'-Hydroxychalcones, 8-prenylated flavanones, cyclic peptides, 7-*O*-glycosylated flavonoids, glycosylated lignans, mono and disaccharide fatty acylglycosides, quinoline carboxamides, diterpenes, sesquiterpenes, among others, as modulators of ROS production by human neutrophils. This is the first study that evaluates more than 100 Brazilian species of Orchidaceae, contributing to the knowledge of of this very diverse family, demonstrating that Brazilian orchids can be much more than ornamental, but holders of new drugs.

Keywords: Metabolomics, Orchidaceae, Molecular Network, Antioxidant.

LISTA DE FIGURAS

INTRODUÇÃO GERAL

Figura 1. Esquema descrevendo como é criada e visualizada uma Rede Molecular. (Wang *et.al.*, 2016) 05

CAPÍTULO 01

Figura 1. Etapas que envolvem a aplicação de um estudo metabolômico em plantas. Fonte: 11
(Pilon *et.al.*2020).

Figura 1. Perfis cromatográficos das análises por UFLC-DAD-ESI(+)-MS/MS (micrOTOF-QII) das extrações realizadas na otimização do método de extração. **A.** Extrações feitas da espécie Cos131 variando a proporção de solução extratora Metanol/Água. *Vermelho:* (9:1). *Verde:* (7:3). *Azul:* (5:5). **B.** Extrações feitas da espécie Epu61 utilizando metanol (verde) e etanol (roxo) como solvente orgânico da solução extratora de proporção 7:3. **C.** Extrações da espécie Cos13 utilizando com Metanol/Água (7:3) com *clean up* (azul) e sem *clean up* (verde). **D.** Extrações da espécie Cos131 utilizando Metanol/Água 7:3 (verde) e Metanol/Água em 0,1% de ácido fórmico (7:3) (rosa). 27

Figura 3. Impressões digitais metabólicas das 131 espécies de Orchidaceae adquiridas no modo positivo por UFLC-DAD-ESI-MS/MS (micrOTOF-QII) após a otimização dos métodos de análise. Software de visualização utilizado: *MzMine™* (v. 2.51, BMC Bioinformatics, Reino Unido). 31

Figura 4. Impressões digitais metabólicas das 131 espécies de Orchidaceae adquiridas no modo negativo por UFLC-DAD-ESI-MS/MS (micrOTOF-QII) após a otimização dos métodos de análise. Software de visualização utilizado: *MzMine™* (v. 2.51, BMC Bioinformatics, Reino Unido). 31

Figura 5. Gráficos de Distribuição de cosseno de similaridade com o banco de dados espectrais do GNPS e comparação entre as redes moleculares geradas através de MN clássica e FBMN. **A.** Redes moleculares geradas através de MN clássica (LC-PP) de acordo com os modelos de otimização produzidos. **B.** Redes moleculares geradas através de MN clássica (LC-PP) de acordo com os modelos de otimização produzidos 45

Figura 6. Impressões digitais metabólicas das 131 espécies de Orchidaceae adquiridas por – GC-EI-MS QP2010 (Shimadzu). Software de visualização utilizado: *MzMine™* (v. 2.51, BMC Bioinformatics, Reino Unido). 52

Figura 7. Métodos de geração de redes moleculares da plataforma online GNPS a partir de dados de massa não direcionados. **(a)** Após a aquisição dos dados de LC-MS/MS (MS^1 e MS^2), dois métodos podem ser utilizados para gerar as redes moleculares: MN clássico e FBMN. **(b)** Gráficos demonstrando o número de análises realizadas. O gráfico em cima mostra o número de MN clássicos e FBMN realizados desde 2016 e o em baixo mostra o número de FBMN realizados desde sua criação e as variações de acordo com as melhoras de processos. Figura retirada de Nothias *et.al* (2019). 55

CAPÍTULO 02

Figura 1. Famílias moleculares, relacionadas a atividades biológicas, da rede molecular advinda dos extratos de Orchidaceae analisados por GC-MS, coloridas de acordo com as classes químicas. O tamanho do cluster (nodo) está relacionado ao número de espectros. 84

Figura 2. Algumas Famílias moleculares, relacionadas a atividades biológicas, da rede molecular advinda dos extratos de Orchidaceae analisados por LC-MS, coloridas de acordo com as classes químicas. O tamanho do cluster (nodo) está relacionado ao número de espectros. 93

Figure 3. Rede molecular do metaboloma apolar (GC-MS) de espécies de Orchidaceae, coloridas de acordo com a subfamília (A) e tribo (B) observando os padrões de correlação 97

com a classificação taxonômica e filogenética, segundo Givnish et al. (2015 e 2016) e Chase, et al. (2015)

Figure 4. Rede molecular do metaboloma semi-polar (LC-MS, modo positivo) de espécies de Orchidaceae, coloridas de acordo com a subfamília (A) e tribo (B) observando os padrões de correlação com a classificação taxonômica e filogenética, segundo Givnish et al. (2015 e 2016) e Chase, et al. (2015). 98

Figura 5. Análises de Componentes Principais (PCA) do Metaboloma apolar (A), semi-polar (B) e da combinação dos dois metabolomas (C) das espécies de Orchidaceae, observando tendências de agrupamento de acordo com a subfamília. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015). 100

Figura 6. Análises de Componentes Principais (PCA) do Metaboloma apolar (A), semi-polar (B) e da combinação dos dois metabolomas (C) das espécies de Orchidaceae, observando tendências de agrupamento de acordo com a tribo. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015). 101

Figura 7. Análise de Agrupamentos Hierárquicos (HCA) do metaboloma semi-polar das espécies de Orchidaceae, adquirido por LC-MS no modo de ionização positivo e negativo, observando tendências de agrupamento de acordo com a subfamília e tribo. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015). 102

Figura 8. Análise de Agrupamentos Hierárquicos (HCA) do metaboloma apolar das espécies de Orchidaceae, adquirido por GC-MS, observando tendências de agrupamento de acordo com a subfamília e tribo. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015). 102

Figura 9. Análise de Enriquecimento de Classes presentes nas subfamílias das espécies de Orchidaceae estudadas, destacando as 10 classes dentre as subfamílias. A. Epidendroideae. B. Orchidoideae. C. Vanilloideae.. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015). 104

Figura 10. Análise de Enriquecimento de Classes presentes nas tribos Cymbidieae e Epidendreae, avaliando a diferença entre elas à nível de classe. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015). 106

CAPÍTULO 03

Figura 1. Gráfico de dispersão de Score gerados pela análise estatística multivariada supervisionada PLS-DA (*Partial Least Square – Discriminant Analysis*) correlacionando as impressões digitais metabólicas com as propriedades antioxidantes dos extratos analisados de Orchidaceae brasileiras. R^2Y . 0,95. Q^2 .0,68 130

Figura 2. Gráfico de dispersão de Score gerados pela análise estatística multivariada supervisionada PLS-DA (*Partial Least Square – Discriminant Analysis*) correlacionando as impressões digitais metabólicas com as propriedades moduladoras da produção de ERO por neutrófilos humanos, dos extratos analisados de Orchidaceae brasileiras. R^2Y . 0,98. Q^2 .0,73. 151

LISTA DE TABELAS

CAPÍTULO 01

Tabela 1. Parâmetros avaliados na otimização do método de extração, se acordo com as proporções de líquido extrator (solvente orgânico/água): 9:1 (1) , 7:3 (2) e 5:5 (3) .	16
Tabela 2. Parâmetros e valores otimizados para a análise por UFLC-DAD-ESI-MS/MS (micrOTOF-QII) para a construção da Rede Molecular. Destacados em negrito e itálico: parâmetros selecionados.	29
Tabela 3. Modelos de otimização do pré-tratamento gerados pelo Mzmine e selecionados para construção das redes moleculares, com seus respectivos números de íons precursores, íons precursores filtrados e substâncias anotadas.	34
Tabela 4. Parâmetros utilizados no software <i>MzMine</i> TM para o processamento dos dados adquiridos no modo de ionização positivo por análise em UFLC-DAD-ESI-MS/MS (micrOTOF-QII).	35
Tabela 5. Parâmetros utilizados no software <i>MzMine</i> TM para o processamento dos dados adquiridos no modo de ionização negativo por análise em UFLC-DAD-ESI-MS/MS (micrOTOF-QII).	36
Tabela 6. Parâmetros e valores utilizados na otimização da construção das Redes Moleculares por MN clássico e FBMN.	40
Tabela 7. Características resultantes dos modelos na otimização da construção das Redes Moleculares para os dados brutos (LC-PR) das análises por LC-MS.	42
Tabela 8. Características resultantes dos modelos na otimização da construção das Redes Moleculares para os dados pré-tratados (LC-PP) das análises por LC-MS.	43
Tabela 9. Características resultantes dos modelos na otimização da construção das Redes Moleculares para os dados brutos (LC-NR) das análises por LC-MS.	46
Tabela 10. Características resultantes dos modelos na otimização da construção das Redes Moleculares para os dados pré-tratados NEG-1 (LC-NP) das análises por LC-MS.	47
Tabela 11. Características resultantes dos modelos na otimização da construção das Redes Moleculares para os dados pré-tratados NEG-5 (LC-NP) das análises por LC-MS.	48
Tabela 12. Parâmetros e valores utilizados na otimização da construção das Redes Moleculares a partir de dados de GC-MS.	53
Tabela 13. Características dos modelos da otimização da construção das Redes Moleculares para os dados deconvoluídos das análises por GC-MS.	54

CAPÍTULO 02

Tabela 1. Parâmetros utilizados no software <i>MzMine</i> TM para o processamento dos dados adquiridos no modo de ionização positivo por análise em UFLC-DAD-ESI-MS/MS (micrOTOF-QII).	73
Tabela 2. Parâmetros utilizados no software <i>MzMine</i> TM para o processamento dos dados adquiridos no modo de ionização negativo por análise em UFLC-DAD-ESI-MS/MS (micrOTOF-QII).	74
Tabela 3. Distribuição de algumas subclasses presentes nas amostras analisadas e de interesse farmacológico, de acordo com os compostos anotados nas redes moleculares advindos dos extratos de Orchidaceae analisados por GC-MS.	83
Tabela 4. Distribuição de algumas subclasses presentes nas amostras analisadas e de interesse farmacológico, de acordo com os compostos anotados nas redes moleculares advindos dos extratos de Orchidaceae analisados por LC-MS	92

CAPÍTULO 03

Tabela 1. Parâmetros utilizados no software <i>MzMine</i> TM para o processamento dos dados adquiridos no modo de ionização positivo por análise em UFLC-DAD-ESI-MS/MS (micrOTOF-QII).	120
Tabela 2. Parâmetros utilizados no software <i>MzMine</i> TM para o processamento dos dados adquiridos no modo de ionização negativo por análise em UFLC-DAD-ESI-MS/MS (micrOTOF-QII).	121
Tabela 3. Resultados das análises de ação antioxidante por DPPH de extratos hidrometanólicos de espécies brasileiras de Orchidaceae. *Maior que 70% de inibição.	123

LISTA DE ABREVIATURAS E SIGLAS

GC-MS	Cromatografia Gasosa acoplada a espectrometria de massas
LC	Cromatografia Líquida acoplada a espectrometria de massas
MS/MS	Análise de fragmentação por espectrometria de massas
GNPS	<i>Global Natural Products Social Molecular Networking</i>
MN	Rede Molecular (<i>Molecular Network</i>)
ESI-MS/MS	Eletrospray acoplado a análise e detecção de fragmentação por espectrometria de massas
ERO	Espécies reativas de oxigênio
RMN	Ressonância Magnética Nuclear
DAD	Arranjo de Diodos
EI	Impacto eletrônico
ESI	Eletrospray
PLS	<i>Partial Least Square</i> (Análise de Mínimos Quadrados Parciais)
OPLS	<i>Orthogonal Partial Least Square</i> (Análise de Mínimos Quadrados Parciais Ortogonais)
PCA	<i>Principal Componente Analysis</i> (Análise de Componentes Principais)
HCA	<i>Hierarchical Componente Analysis</i> (Análise de Componentes Hierárquicas)
R ²	Coeficiente de Correlação
Q ²	Coeficiente de Predição
UV	Ultra-Violeta
LES	Lúpus eritematoso sistêmico

LISTA DE SÍMBOLOS

α	Alfa
β	Beta
γ	Gama
TM	Trademark
®	Marca Registrada
©	Copyright

Sumário

RESUMO.....	i
ABSTRACT.....	ii
LISTA DE FIGURAS.....	iii
LISTA DE TABELAS.....	v
LISTA DE ABREVIATURAS E SIGLAS.....	vi
LISTA DE SÍMBOLOS.....	vii
INTRODUÇÃO GERAL.....	1
OBJETIVOS GERAIS.....	9
CAPÍTULO 1.....	10
1. INTRODUÇÃO.....	11
2. OBJETIVOS.....	14
3. MATERIAL E MÉTODOS.....	14
3.1. Material Vegetal.....	14
3.2. Processamento de Amostra.....	14
3.3. Extração.....	15
3.4. Instrumentação e Análise dos Extrato.....	17
3.5. Processamento dos Dados.....	20
3.6. Redes Moleculares (<i>Molecular Networking</i>).....	21
4. RESULTADOS E DISCUSSÕES.....	24
4.1. Análise Metabolômica - Compostos de média-alta polaridade.....	24
4.2. Análise do Metaboloma - Compostos de média-baixa polaridade.....	51
4.3. Importância do Processamento de Dados na Construção da Rede Molecular..	55
4.4. Importância da Distribuição de Cossenos de Similaridade na Construção da Rede molecular.....	58
5. CONCLUSÕES.....	60
CAPÍTULO 2.....	62
1. INTRODUÇÃO.....	63
2. OBJETIVOS.....	65
3. MATERIAL E MÉTODOS.....	65
3.1. Material Vegetal.....	65
3.2. Preparação dos Extratos.....	66
3.3. Cromatografia líquida de ultrarrápida eficiência acoplada a espectrometria de massas – UFLC-DAD-ESI-MS/MS (microTOF-QII).....	66

3.4. Cromatografia a gás acoplada a espectrometria de massas – GC-EI-MS QP2010 (Shimadzu).....	68
3.5. Processamento dos Dados	68
3.6. Redes Moleculares	71
3.7. Propagação de Anotações por Redes Espectrais in silico, Databases e Ranqueamento – Network Annotation Propagation (NAP)	73
3.8. Análises de Motivos (Motifs) ou anotação por Mass2Motifs (MS2LDA).....	74
3.9. MolNetEnhancer e ClassyFire	75
3.10. Análises de Enriquecimento de Classes Químicas.....	75
3.11. Análises Quimiométricas.....	76
4. RESULTADOS E DISCUSSÃO	77
4.1. Compostos de média-baixa polaridade.....	77
4.2. Compostos de média-alta polaridade.....	84
4.3. Padrões de Similaridade Química.....	92
5. CONCLUSÕES	103
 CAPÍTULO 3.....	 104
 1. INTRODUÇÃO	 105
2. OBJETIVOS.....	107
3. PROCEDIMENTOS EXPERIMENTAIS.....	108
3.1. Material Vegetal	108
3.2. Avaliação do Potencial Bioativo dos Extratos	109
3.3. Análises dos extratos por UFLC-DAD-ESI-MS/MS (micrOTOF-QII).....	112
3.4. Pré-processamento dos Dados.....	113
3.5. Análises de Correlação in silico	115
3.6. Anotações via Redes Moleculares	115
4. RESULTADOS E DISCUSSÃO	116
4.1. Atividade Antioxidante dos extratos brutos	116
4.2. Citotoxicidade dos extratos brutos	118
4.3. Efeitos sobre a Produção de ERO por Neutrófilos dos extratos brutos.....	119
4.4. Biomarcadores do Potencial Ativo de Orchidaceae Brasileiras	123
5. CONCLUSÕES	126
 CONCLUSÕES GERAIS e PERSPECTIVAS.....	 127
 REFERÊNCIAS GERAIS	 127
 ANEXOS	 136

INTRODUÇÃO GERAL

Análises Metabolômicas e suas Abordagens

O metaboloma é o conjunto de todos os metabólitos sintetizados por um organismo em um dado momento, consistindo de substâncias extremamente diversas e complexas (Sumner, 2003; Villas-Boas *et al.*, 2005), e uma vez que reflete o fenótipo do organismo em determinadas condições, fornece grande quantidade de informações no que diz respeito à compreensão dos sistemas biológicos. Devido à complexidade da matriz metabólica, torna-se necessária uma combinação de equipamentos sofisticados e estratégias elaboradas para analisar o maior número possível de metabólitos (Brunetti *et al.*, 2018). Uma análise que possibilita rapidez, eficácia e sensibilidade para o estudo do metaboloma é a análise metabolômica.

A metabolômica é a ciência que investiga a expressão fenotípica de maneira qualitativa e quantitativa do metabolismo primário e secundário de organismos, em um determinado momento sobre determinadas condições. Essa investigação ocorre através de análises comparativas de perfis metabólicos entre grupos determinados pelo estudo, sendo que esses foram sujeitos à diferentes condições (Pilon *et.al*, 2020). Devido à grande complexidade dos metabólitos e suas características, essa abordagem global é impraticável e dificilmente aplicável (Villas-Boas *et al.*, 2005; Wang *et al.*, 2016).

Com isso, as análises metabolômicas são divididas em estudos alvo (do inglês, *targeted*), nos quais são estudados metabólitos específicos, e estudos não-alvo (do inglês, *untargeted*), os quais avaliam o conjunto metabólico sem especificações. Os estudos não-alvo são separados em duas abordagens: as impressões digitais metabólicas (do inglês, *Metabolic Fingerprint*) e a Metabolômica de Plantas (do inglês, *Plant Metabolomics*). Os estudos alvo são separados em Perfil Metabólico (do inglês, *Metabolic Profile*) e Análises Alvo (do inglês, *Target Analysis*) (Pilon *et.al*, 2020).

A abordagem metabolômica através da impressão digital metabólica consiste na detecção detalhada do maior número de substâncias presentes no organismo sob determinada condição, produzindo um dado analítico que é característico daquela planta, como uma impressão digital (Hubert *et.al.*, 2017). Esta abordagem permite a comparação entre amostras através de ferramentas matemáticas e estatísticas para determinar padrões ou tendências dentro de um grupo de amostras e/ou correlacionar esses com

propriedades relacionadas, que possam agregar informações para o estudo, (Allwood & Goodacre., 2010; Bamba & Fukusaki, 2006; Chagas-Paula *et al.*, 2015; Moritz & Johansson, 2008; Nordström *et al.*, 2008).

A metabolômica de plantas é uma tentativa não enviesada de identificar e quantificar o maior número de compostos possível presentes no objeto de estudo (tecido, órgão, etc.). A obtenção do perfil metabólico visa à identificação e quantificação de um grupo de metabólitos de interesse responsáveis pelas diferenças observadas entre grupos estudados. Os metabólitos apontados como discriminantes da propriedade estudada podem servir de base para a discussão da hipótese defendida. A abordagem pela Análise Alvo visa a detecção ou descrição de grupos específicos que possuem propriedades químicas semelhantes ou rotas metabólicas em comum (Pilon *et.al*, 2020).

Através dessas abordagens, é possível determinar similaridades e/ou dissimilaridades entre amostras, correlacionar informações de interesse, propor os metabólitos que discriminam tais propriedades, e direcionar para os próximos passos, sejam eles isolamentos direcionados, estudos de rotas biossintéticas ou estudos de atividade biológica mais aprofundados. Essas ferramentas são utilizadas frequentemente na resolução de problemas em classificações taxonômicas e distribuições biogeográficas (Martucci *et al.*, 2018; Chagas-Paula *et al.*, 2015; Kim *et al.*, 2010; Padilla-González, 2017), permite também a identificação de diferenças nas composições químicas de amostras, como extratos de plantas e amostras biológicas, como urina, plasma e sangue, sendo possível a identificação de biomarcadores de doenças (Zhang *et al.*, 2014), e na identificação de substâncias com potencial farmacológico, podendo anotá-las sem a necessidade do isolamento completo por fitoquímica clássica, o que demanda tempo e recursos financeiros (Villas-Boas *et al.*, 2005; Wolfender *et al.*, 2010; Weckwerth & Morgenthal, 2005).

Estudo Integrativo: Técnicas e Estratégias Metabolômicas

Atualmente não existe nenhuma técnica analítica que seja capaz de detectar toda a complexidade da amostra em um único experimento. Analisar matrizes complexas, com compostos em diferentes concentrações, e detectar as mínimas diferenças entre as amostras necessitam de técnicas analíticas hifenadas ou acopladas que possuam grande sensibilidade, eficiência e rapidez, tais como Ressonância Magnética Nuclear (RMN),

Espectrometria de Massas (MS), e cromatografia de alta e ultra alta eficiência, preferencialmente acoplada à detectores de UV e analisadores de massa que aumentam a performance da análise (Chagas-Paula *et al.*, 2015; Yuliana *et al.*, 2011a; Weckwerth e Morgenthal, 2005).

Uma técnica de escolha é a cromatografia líquida de alta ou ultra eficiência (HPLC ou UHPLC) acoplada a detector de ultravioleta com arranjo de diodos (UV-DAD) e espectrometria de massas de alta resolução (HRMS) – LC-UV-MS/MS, devido à abrangência de técnica e sua alta sensibilidade (Moritz e Johansson, 2008; Villas-Boas *et al.*, 2005, Van Der Kooy *et al.*, 2009; Wolfender *et al.*, 2010). Essa é uma das técnicas empregadas nas análises de extratos padronizados, de diferentes partes de uma planta, obtidos através de formas variadas, havendo a possibilidade de identificação de inúmeros constituintes químicos em concentrações ínfimas e sem a necessidade de seu isolamento. Além da rapidez, eficiência e sensibilidade, outra vantagem é poder utilizar os dados gerados na análise em estudos de correlação entre amostras através da mineração de dados (Katajamaa e Orešič, 2007; Nordström, 2008).

Os dados gerados nas análises são massivos, pois são conjuntos de vetores que possuem um grande número de informações, como os valores de tempo de retenção, m/z , intensidade e área do pico, além de espectros de UV e de massas para cada metabólito detectado. Além disso, devido à complexidade do metaboloma das plantas, os dados gerados por LC-MS e GC-MS podem apresentar coeluições que levam à sobreposição de perfis de fragmentação, além de possíveis fragmentações na fonte durante o processo em analisadores de massas (Bamba e Fukusaki, 2006; Tikunov *et al.*, 2005). Assim, é necessário o uso de softwares específicos e robustos que extraíam as informações úteis para as análises posteriores. Um processamento inadequado trará um viés aos dados químicos e biológicos inviabilizando a conclusão das análises.

O processamento dos dados, independente da abordagem, pode ser realizado utilizando os softwares *XC-MSTM* (METLIN, EUA), *MzMineTM* (BMC Bioinformatics, Reino Unido) dentre outros, com o intuito de filtrar os dados para diminuir o ruído, para realizar o alinhamento e detecção do picos, o qual envolve a deconvolução, retirada de isótopos, e a normalização dos dados. Dessa maneira, obtém-se dados de qualidade, organizados em matrizes de dados e arquivos de dados espectrais, que são utilizados nas análises de correlação e bioinformáticas.

A infinidade de informações geradas pelas abordagens metabolômicas podem ser utilizadas para diversas aplicabilidades, sendo esta etapa crucial para obter-se as respostas necessárias para deliberar a hipótese estudada. Os dados processados podem ser analisados por ferramentas estatísticas, matemáticas, bioinformáticas e de correlação por visualização de dados.

As análises quimiométricas, utilizando as análises estatísticas multivariadas, facilitam a visualização, a análise dos dados e a obtenção de discriminantes, e existem diferentes métodos para a modelagem dos dados. Os métodos usualmente empregados são os não-supervisionados, como HCA (*Hierarchical Cluster Analysis* – Análise de Agrupamento Hierárquico) e PCA (*Principal Component Analysis* - Análise de Componentes Principais), e os supervisionados, como PLS (*Partial Least Squares* - Análise por Mínimos Quadrados Parciais), OPLS (*Orthogonal Partial Least Squares* - Análise por Mínimos Quadrados Parciais Ortogonais), e SIMCA (*Soft Independent Modeling by Class Analogy*). Os métodos não-supervisionados oferecem uma visão global das correlações entre amostras para observar o comportamento químicos entre elas, buscando tendências ou padrões de similaridade ou dissimilaridade. Os supervisionados estão direcionados a propriedades ou classes atribuídas às amostras, observado os padrões de similaridade e buscando os discriminantes relacionados à cada propriedade, como moléculas bioativas de propriedades biológicas. (Pilon *et.al.*, 2020; Yuliana, 2011a; Weckwerth e Morgenthal, 2005; Lang, 2008; Katajamaa e Oresic, 2007; Trygg, 2007; Shyur, 2008).

Existem vários softwares estatísticos que podem ser empregados para a realização desses procedimentos de mineração de dados (em inglês, *Data Mining*), tais como The Unscrambler X (CAMO Software, Noruega), SIMCA (Umetrics) ou R (*R Project for Statistical Computing*).

A identificação dos metabólitos indicados na mineração de dados, chamados de biomarcadores ou *hits*, é preferencialmente feita dentro da abordagem metabolômica através de estratégias de anotação molecular. Pilon *et.al.* (2020) cita 5 níveis considerados pela *Metabolites Standard Initiaves* (MSI). O Nível 0 ou identificação Não-ambígua ocorre quando há a elucidação completa a partir do isolamento, geralmente realizada para moléculas inéditas ou quando os dados são insuficientes para a elucidação

em si. O Nível 1 ocorre pela comparação de dados espectroscópios e espectrométricos com um padrão autêntico.

Os níveis 2 e 3 ocorrem sem a comparação com padrão, sendo que para muitos estudos metabolômicos a determinação nestes níveis já é suficiente para prosseguir. A anotação em nível 2 possibilita a determinação da estrutura comparando com bases de dados e descrições na literatura. Já à nível 3 é possível determinar a classe da substância discriminada ou a estrutura mais provável, mas nada à nível de isômeros. À nível 4 fornece somente informações quanto à presença e ausência. As anotações podem ser realizadas em bases de dados online, bibliotecas *in house* e através de biblioteca espectrais de acesso livre, como o GNPS.

O *Global Natural Products Social Molecular Networking* (GNPS, Rede Mundial de Material Molecular de Produtos Naturais, disponível em <http://gnps.ucsd.edu>) é uma biblioteca espectral de acesso livre que permite a anotação de metabólitos em espectros em redes moleculares, a qual fornece ferramentas para tratamento e compartilhamento de dados de espectrometria de massas com base em dados de eliminações neutras obtidas por ativação colisional em equipamentos de ESI-MS/MS (Bouslimani *et.al.*, 2014; Demarque, *et.al.*, 2016; Wang *et.al.*, 2016). Através de suas ferramentas computacionais, é possível construir o *Molecular Network*.

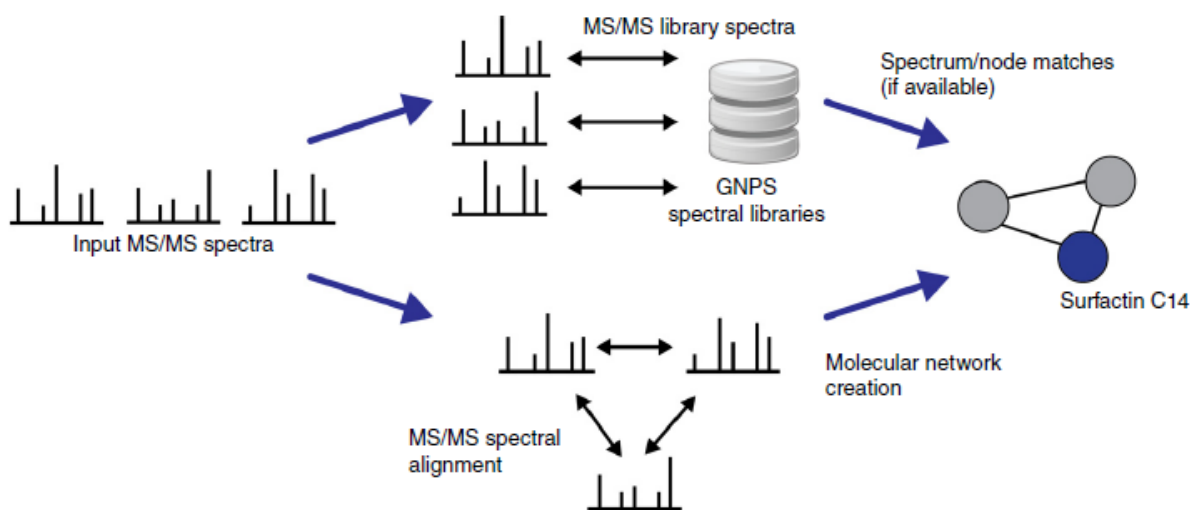


Figura 1. Esquema descrevendo com é criada e visualizada uma Rede Molecular. (Wang *et.al.*, 2016).

A elaboração de redes moleculares (do inglês, *Molecular Networking*) é uma abordagem computacional empregada no tratamento de dados de MS/MS utilizados na

análise metabolômica para anotação de substâncias e estudos de correlação. Nestas redes, os espectros de cada substância são representados como nodos (*clusters*) e uma série de nodos conectados (família molecular) indica moléculas que apresentam semelhanças es espectrais. Assim, substâncias de uma mesma classe são agrupadas dentro de uma mesma rede (Yang *et.al.*, 2013; Duncan *et.al.*, 2015).

Esta abordagem permite que sejam realizadas comparações de alto rendimento e consiste em um método rápido para anotar novos metabólitos. Sendo assim, representa um importante avanço no campo de pesquisa de produtos naturais, sendo possível dar um passo significativo em direção à elucidação das estruturas desconhecidas ou, ao menos, guiar seu isolamento monitorado por espectrometria de massas, para posterior análise por meio de técnicas integradas (RMN, IV, UV) (Bouslimani *et.al.*, 2014; Krug & Muller, 2014; Yang *et.al.*, 2013).

A família Orchidaceae, sua Diversidade Metabólica e Potencial Ativo

A Orchidaceae é segunda maior família dentre as angiospermas, detentoras de flores e frutos, ficando atrás da Asteraceae e a maior dentre as monocotiledôneas em número de espécies. É considerada uma das mais fascinantes devido a sua diversidade ecológica e diversidade e especialização da estrutura floral, características tais que a faz uma família de grande interesse ornamental. (Chase *et al.*, 2015; Dressler *et al.*, 2005; Freudenstein *et al.*, 2004; Wong & Sun, 1999).

De acordo com Chase *et al.* (2015), são conhecidas atualmente em torno de 25.000 espécies e 736 gêneros de Orchidaceae, e desde a última classificação foram descritos por volta de 150 novos gêneros. As orquídeas estão contidas em cinco grandes subfamílias, de acordo principalmente com o número e a posição das anteras, sendo essas: Apostasioideae, Cypripedioideae, Epidendroideae, Orchidoideae e Vanilloideae. A subfamília Epidendroideae é de longe a mais vasta, contendo cerca de 580 gêneros e 15.000 espécies, possuindo número de gêneros e espécies superior ao que soma todas as outras subfamílias em conjunto (Bastos e Van Der Berg, 2012; Cameron *et al.* 1999; Chase *et al.*, 2015; Pridgeon *et al.* 2009 Stern e Judd, 2001; Silva, 1998; Endress, 2016).

Em território brasileiro, a família ocorre em todos os tipos de formações vegetais, estando entre os países que detém uma das maiores diversidades de orquídeas do

continente americano e do mundo, junto com a Colômbia e Chile, ocorrendo 2420 espécies distribuídas em 235 gêneros e 4 subfamílias (Barros *et al.* 2010).

Devido à sua beleza exótica e diversidade de formas, tamanhos e cores muito atrativas, as orquídeas são cultivadas como ornamentais e possuem grande valor comercial devido a sua longa vida de prateleira. Devido ao interesse econômico, há um vasto conhecimento e investimento no cultivo e manutenção dessas plantas, além de estudos de técnicas que os otimizam, proporcionando vantagens na comercialização. Como exemplo há a multiplicação em larga escala de orquídeas utilizando técnicas de cultura de tecidos, especialmente de híbridos e espécies ameaçadas, que tem ajudado a família a ocupar uma posição entre as dez melhores flores de cultivo ornamental. (Baker *et al.*, 2014).

Além da ampla utilização ornamental, as orquídeas são utilizadas na Medicina Tradicional Chinesa, Medicina Ayurveda (indiana) e na medicina popular de diversos outros lugares do mundo, como no sudoeste asiático, Índia, China, Japão, Europa, África e Américas. Em sua revisão, Hossain (2011) descreve alguns empregos na medicina popular para diversos tipos de doenças e situações, como tuberculose, afecções do estômago, dores no peito, artrite, sífilis, icterícia, eczema, tumores, furúnculos, inflamações, dores musculares, hepatite, dispepsia, enjoos, fraturas ósseas, reumatismo, asma, malária, dores de ouvido, feridas e úlceras, Além disso, também são usadas como antiemético e no tratamento de ferimentos por escorpião e cobra.

Estudos sobre as atividades biológicas e o potencial medicinal dos metabólitos de orquídeas vêm cada vez mais sendo realizados. Os metabólitos das preparações e extratos foram principalmente estudados nas espécies exóticas, indicando principalmente a presença de flavonóides e alcalóides, além de antocianinas, terpenos, quinonas, fenantrenos, cumarinas, sesquiterpenóides, dentre outros (Cota *et.al.*, 2008; Gutiérrez *et.al.*, 2010; Ramos *et.al.*, 2013). No entanto, estudos sobre as espécies nativas do Brasil e América do Sul permanecem escassos (Gutiérrez, 2010; Ramos *et.al.*, 2012; Silva *et.al.*, 2016).

O *Cyrtopodium macrobulbon* tem sido muito utilizado na medicina tradicional mexicana para o tratamento de doenças urinárias dolorosas ("mal de Orin"). O estudo realizado por Morales-Sánchez *et al.* (2014) demonstrou que os compostos gigantol, batatasin III e eucalyptol presentes nos extratos possuem atividade antinociceptiva,

apoiando o uso popular desta orquídea. O estudo utilizando orquídeas brasileiras de Porte *et.al.* (2014) promoveu um estudo bioquímico contra o fungo *Cladosporium herbarum* buscando substâncias ativas advindas da orquídea brasileira *Miltonia flavescens* Lindl.

Estudos *in vitro* de ação antimicrobiana e antifúngica realizados em espécies exóticas demonstraram grande potencial bioativo, tanto de extratos quanto de substâncias isoladas. Nos estudos de Lyudmyla *et.al.* (2017) e Singh *et.al.* (2017) foram realizados o fracionamento dos extratos brutos de *Coelogyne speciosa* e *Eria alba*, respectivamente, e ambos demonstraram grande atividade antimicrobiana. Já em Yoshikawa *et.al.* (2014), dois fenantrenos e um fenilpropanóide (ephemeranthoquinona C, marylaurencinols C e D) foram isolados das raízes de *Cymbidium sp.* e testados contra *Bacillus subtilis*, *Klebsiella pneumoniae*, and *Trichophyton rubrum*, sendo o primeiro relato de derivados fenantrenos de uma fonte natural mostrando atividade antimicrobiana contra *T. rubrum*.

Em Jie Ren *et.al.* (2016), dois novos glicosídeos fenantrenos (nudolglicosídeos A e B) e trinta compostos conhecidos foram isolados de *Liparis regnierii* Finet. Seis derivados fenantreno conhecidos exibiram atividade antibacteriana significativa contra *S. agalactiae* e *B. subtilis*. Em Yanbei Tu et al. (2018), substâncias presentes no extrato etanólico (Etanol/Água 95:5) dos tubérculos de *Cremastra appendiculata* demonstraram atividade anticolinesterásica pela inibição de butirilcolinesterase (BChE), atividade antioxidante e inibição da agregação de peptídios β -amiloides, sugerindo que esses poderiam ser utilizados como agente multifuncionais para o desenvolvimento de medicamento contra a doença de Alzheimer.

A grande diversidade de classes e compostos, observadas nos estudos químicos de espécies exóticas de Orchidaceae, e o potencial bioativo observado em estudos etnofarmacológicos e de atividade biológica mostram a grande riqueza química e bioativa desta fascinante família. Esses proporcionam um pilar sólido científico para pesquisarmos o potencial das espécies de um dos territórios de maior e diversidades de orquídeas do planeta, estudos os quais começaram a crescer nos últimos 10 anos (Cota *et al.* 2008; Gutiérrez, 2010; Ramos *et al.*, 2012; Silva, 2003; Xochitl, 2011).

OBJETIVOS GERAIS

Os objetivos gerais deste estudo foram investigar a diversidade de metabólitos secundários presentes em espécies brasileiras de Orchidaceae e avaliar o potencial bioativo de seus extratos quanto a atividade moduladora da produção de espécies reativas de oxigênio (ERO) por neutrófilos humanos, atividade antioxidante, e citotoxicidade. Após isso, correlacionar os dados obtidos e propor as substâncias relacionadas às atividades através de ferramentas metabolômicas.

Esta tese está apresentada em 3 capítulos, com o intuito de discutir os resultados de maneira clara e estruturada. A seguir, os objetivos de cada capítulo.

Capítulo 1. Obtenção das impressões digitais metabólicas e construção de redes moleculares (*Molecular Networking*) para a análise dos perfis químicos das espécies de Orchidaceae estudadas.

Capítulo 2. Estudo Integrativo dos metabólitos secundários de Orchidaceae brasileiras usando análises metabolômicas e bioinformáticas para desreplicação e análises de correlação.

Capítulo 3. Orquídeas brasileiras como fontes de metabólitos moduladores da produção de espécies reagentes de oxigênio (ERO) por neutrófilos humanos e sua ação antioxidante.

CAPÍTULO 1

1. INTRODUÇÃO

A metabolômica surgiu nos últimos anos como uma ferramenta indispensável para a análise de matrizes complexas como extratos de produtos naturais brutos, levando a uma mudança de paradigma na pesquisa de medicamentos de produtos naturais. (Salem *et.al*, 2020). Além disso, a metabolômica tem desempenhado um papel central em diversas áreas do conhecimento, e dentre as relacionadas à estudos de plantas temos a agricultura, taxonomia, quimiotaxonomia e a ecologia química (Pilon *et.al.*2020).

Acessar a rica diversidade de matrizes complexas estudadas na metabolômica de plantas requer a seleção de métodos adequados às particularidades das amostras e aos objetivos do estudo, os quais envolvem diversas metodologias e abordagens analíticas que delimitam as etapas posteriores (Pilon *et.al.*2020). Dessa forma, a construção de fluxos de trabalho baseia-se em protocolos estabelecidos pela comunidade metabolômica e estudos metabolômicos, obtendo assim um conjunto de dados de alta qualidade e, conseqüentemente, análises quimiométricas e bioinformáticas que expressam a real informação buscada no metaboloma estudado. (Brunetti *et.al.*, 2018; Salem *et.al.*, 2020; Pilon *et.al.*,2020).



Figura 1. Etapas que envolvem a aplicação de um estudo metabolômico em plantas. Fonte: (Pilon *et.al.*2020).

O planejamento experimental deve estar contextualizado com a hipótese do trabalho e com todo o conjunto de etapas realizadas no estudo. Ele é o primeiro passo que concerne o estudo e deve ser realizado com cautela, pois interfere nos resultados e nas interpretações. Primeiramente determina-se o conjunto amostral, a coleta das plantas estudadas e a preparação da amostra, baseando-se na hipótese e objetivos referente ao estudo através protocolos metabólicos. A partir disso, o planejamento consiste em observar os parâmetros que mais influenciam no processo extrativo e/ou no método de análise empregado e realizar etapas de otimização dessas variáveis e fatores (Pilon et.al.2020).

Protocolos de extração em análise metabólicas têm sido desenvolvidos e revisados, e diferentes técnicas têm sido discutidas para a produção de perfis metabólicos representativos. Essa etapa determina a matriz metabólica a ser investigada, considerando a polaridade e solubilidade dos metabólitos e as técnicas analíticas utilizadas, conforme a abordagem metabólica na qual se situa o estudo (Pilon et.al.,2020; Schripsema, 2010; Salem et.al., 2020).Uma vez que os estudos metabólicos trabalham com uma grande quantidade de amostras e matrizes complexas, é de grande importância que a extração seja rápida, simples, com poucas etapas, reprodutível e garanta a extração da maior quantidade possível de metabólitos, e sem degradação. Uma cobertura razoável do metabólica pode ser realizada com dois tipos: Extração apolar (ex. clorofórmio, diclorometano, acetato-hexano) e extração polar (ex. água, água-metanol, água-etanol) (Creydt, et.al., 2018; Pilon et.al.,2020; Schripsema, 2010; Salem et.al., 2020).

A extração e análise das amostras tem um papel importante nos estudos metabólicos, direcionados ou não-direcionados. Não há uma técnica somente que permite analisar o metaboloma completo de um organismo, devido a sua complexidade química e físico-química, assim como a ampla gama de concentrações em que os compostos podem estar presentes, e para analisar essa matriz tão complexa diferentes combinações de técnicas e estratégias são utilizadas, assim como tratamento dos dados obtidos e análises de correlação (Brunetti et al., 2018; Creydt et.al., 2018) .

Métodos analíticos, como espectrometria de massa (MS) e ressonância magnética nuclear (NMR) são os mais comuns na análise metabólica, usados para detectar e

anotar de forma abrangente os constituintes metabólicos de produtos naturais advindos de plantas, e a construção das matrizes metabólicas, de acordo com a abordagem metabolômica (Salem *et.al*, 2020). Dentre essas estão a cromatografia líquida (LC-MS) e cromatografia gasosa (GC-MS), a injeção direta em espectrometria de massas (ID-MS) e a ressonância magnética nuclear (RMN). Outras técnicas também podem ser usadas, como a eletroforese capilar acoplada a espectrometria de massas (EC-MS), cromatografia em camada delgada de alta eficiência (HPTLC - *High performance thin-layer chromatography*) e a cromatografia líquida com detecção por arranjos de diodos (LC-DAD). (Salem *et.al*, 2020, Pilon *et.al*.2020).

O processamento de dados é uma etapa intermediária e essencial para as análises metabolômicas. Ele visa a melhora na qualidade dos sinais como um todo e os estrutura de maneira a alinhá-los, os tornando comparáveis entre si. Após o processamento, os dados são organizados em matrizes robustas que permitem a utilização por diferentes softwares e para diferentes fins, abrindo o leque de possibilidades, interpretações e análise dos dados (Pilon *et.al*.2020, Salem *et.al*.2020).

O GNPS, além de ser uma biblioteca espectral online de acesso livre, é também uma plataforma que agregou e construiu ferramentas que auxiliam desde a aquisição até a análise de um conjunto de dados e compará-los com dados disponíveis publicamente, realizando as anotações. Uma das abordagens é a construção de redes moleculares, que permite a visualização do conteúdo químico detectado pelas análises de MS, comparando-o à nível metabólico de acordo com características ou propriedades, e a anotação dos metabólitos através da busca na base de dados espectral (Wang *et.al.*, 2016).

As ferramentas para a construção das redes moleculares são o *Molecular Networking*, o qual utiliza os dados de MS/MS não tratados na construção da rede, e o *Feature-based Molecular Networking*, o qual usa os dados processados e exportados de software como *MZmine*, *OpenMS*, *MS-DIAL*, *MetaboScape*, *XCMS*, *Progenesis QI* e o *mzTab-M format* (Wang *et.al.*, 2016). De acordo com Nothias *et.at.* (2020), ao utilizar os dados processados, as redes moleculares são capazes de distinguir isômeros, facilitar as anotações e possibilitar a incorporação de informações quantitativas que permitem uma análise estatística robusta, proporcionando uma evolução da capacidade analítica, nas correlações e anotações.

2. OBJETIVOS

Este capítulo teve como objetivos principais a obtenção das impressões digitais metabólicas de 130 espécies de Orchidaceae, dentre elas brasileiras e exóticas, o processamento dos dados obtidos e a construção de redes moleculares (*molecular networking*) de alta qualidade, fornecendo informações que contribuam para a desreplicação e análises de correlação *in silico*. Os objetivos específicos foram:

- a. Coleta das espécies brasileiras e exóticas de Orchidaceae;
- b. Obtenção dos extratos de média-baixa e média-alta polaridades;
- c. Análise dos extratos de média-baixa polaridade por GC-EI-MS QP2010 (Shimadzu);
- d. Análise dos extratos de média-alta polaridade por UFLC-DAD-ESI-MS/MS (micrOTOF-QII);
- e. Processamento dos dados adquiridos por LC-MS e GC-MS;
- f. Otimização e construção das redes moleculares utilizando a plataforma GNPS e suas ferramentas.

3. MATERIAL E MÉTODOS

3.1. Material Vegetal

As 130 espécies de Orchidaceae deste estudo foram fornecidas pelo Prof. Emerson Ricardo Pasarin, coordenador e responsável pelo orquidário do Laboratório de Biologia Molecular e Biosistemática de Plantas (LBMBP-USP) da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (FFCLRP-USP). Foram também selecionadas algumas espécies exóticas como grupo externo. Os respectivos identificadores, nomes científicos, Vouchers e informações sobre as espécies estão no Apêndice I.

3.2. Processamento de Amostra

As partes utilizadas neste estudo foram folhas e pseudobulbos saudáveis e completamente desenvolvidos. As espécies foram coletadas e transferidas para tubos *falcon* de 50 mL, mergulhadas em nitrogênio líquido e liofilizadas a -90°C até eliminação total de água. Após desidratadas, as amostras foram trituradas e o tamanho de partículas

foi padronizado em tamiz de malha granulométrica 45 mesh. Finalmente o material vegetal foi armazenado em frascos de vidro vedados em congelador -26 °C.

3.3. Extração

3.3.1. Compostos de média-alta polaridade

Os extratos contendo os compostos de média-alta polaridade foram produzidos com o intuito de avaliá-los quanto às atividades antioxidante, moderadora da produção de ERO por neutrófilos humanos e citotóxica. Estudos de atividade biológica envolvendo culturas de células, meios tamponados, baixa tolerância à solventes orgânicos, dentre outros, exigem que os extratos sejam solúveis em seus meios para garantir a qualidade dos resultados obtidos. Com isso, foi realizada a seleção da solução extratora e avaliada a necessidade de limpeza dos componentes graxos dos extratos. Os métodos de extração foram otimizados baseando-se nos estudos de De Vos et al. (2007), Gallon et al. (2018), Liu et.al. (2017), Martucci et.al. (2018), Mushtaq et.al. (2014), Padilla-González et al. (2017), Sedio et.al. (2018) e Kim & Veeport (2010).

O procedimento de extração utilizou de 10,0 mg de cada amostra previamente pulverizada e 1,0 mL de solução extratora (Bloco A, B, C e D - Tabela 1) hidroalcolica em 3 diferentes proporções, conforme descrito na Tabela 1. A extração foi realizada em banho de ultrassom por 10 minutos, seguida de centrifugação à 15000 rpm (26 °C) por 10 minutos, filtração em filtro de membrana PTFE (0,22 µm) e secagem em centrífuga à vácuo à temperatura ambiente, com posterior pesagem para cálculo de rendimento em massa.

Cinco espécies foram selecionadas por sorteio para serem utilizadas na otimização: *Vanilla paulista* (Vpa16), *Cyrtopodium gigas* (Cgi52), *Epidendrum puniceoluteum* (Epu61), *Epidendrum denticulatum* (Ede59) e *Catasetum osculatum* (Cos131). Os parâmetros avaliados foram o tipo de solvente extrator (metanol ou etanol), proporção da mistura hidroalcolica do solvente extrator (9:1 **(1)**, 7:3 **(2)** e 5:5 **(3)**), execução ou não de procedimento de *clean up* (particionamento utilizando *n*-hexano) e acidificação ou não da solução extratora utilizando ácido fórmico (0,1%). As extrações foram realizadas em quatro blocos, conforme descrito na Tabela 1.

Tabela 1. Parâmetros avaliados na otimização do método de extração, se acordo com as proporções de líquido extrator (solvente orgânico/água): 9:1 **(1)**, 7:3 **(2)** e 5:5 **(3)**.

Bloco	Parâmetros	Espécies				
		Vpa16	Cgi52	Epu61	Ede 59	Cos131
A	Metanol/Água	A1 A2 A3	A1 A2 A3	A1 A2 A3	A1 A2 A3	A1 A2 A3
B	Metanol/Água com ácido fórmico (0,1%)	B1 B2 B3	B1 B2 B3	B1 B2 B3	B1 B2 B3	B1 B2 B3
C	<i>Clean up</i> (<i>n</i> -hexano) e Metanol/Água	C1 C2 C3	C1 C2 C3	C1 C2 C3	C1 C2 C3	C1 C2 C3
D	Etanol/água	D1 D2	D1 D2	D1 D2	D1 D2	D1 D2

Os extratos produzidos foram analisados por UFLC-DAD-ESI(+)-MS/MS (microTOF-QII) (Bruker Daltonics) utilizando um método abrangente e generalista. Para tanto, utilizou-se CH₃CN-HCO₂H (0,1%) (v/v) (B) e H₂O-HCO₂H (0,1%) (v/v) (A) como fase móvel, a uma vazão de 0,8 mL/min, volume de injeção 15 µL e o seguinte gradiente de eluição: 0 – 2 min, 3% B; 2 – 20 min, 3 - 100% B; 20 – 23 min, 100% B; 23 – 26 min, 100 - 3% B, permanecendo em 3% de B durante 2 minutos após a análise sem aquisição dos dados de massas (*Prerun*). Utilizou-se uma coluna C18 (Kinetex™, Phenomenex), de tamanho de partícula de 2,7 µm (150 x 4,6 mm). O detector de UV(DAD) foi programado para detectar os comprimentos de onda entre 100 e 900 nm. A temperatura do forno foi de 35 °C.

As análises por espectrometria de massas foram realizadas nos modos de ionização positivo, separadamente, utilizando a taxa de espectros (*spectra rate*) de 2,0Hz e faixa de massa de 50 a 1300 *m/z*. Os parâmetros da fonte de ionização foram: voltagem do *End Plate Offset* de 450V; voltagem do capilar, 3,0 kV (modo negativo) e 3,5 kV (modo positivo); nebulizador a 4,5 bar; vazão do gás de secagem de 9,0 L/min; temperatura do gás de secagem, 200 °C. Os parâmetros para a fragmentação MS/MS foram Auto MS/MS com a seleção dos 5 íons precursores mais intensos para a fragmentação, excluindo-o para seleção após detecção de três espectro do mesmo íon ou após 0,60min, reconsiderando sua seleção se a intensidade do íon precursor for o dobro da anterior. Para as fragmentações dos íons precursores, foi utilizada uma rampa de energia de colisão de 20 a 60 eV.

De modo a obter massas de alta resolução, o instrumento e os experimentos foram calibrados a partir de uma solução de padrão interno de Na-TFA (400 mg/100 mL) ao final

de cada análise. Os dados foram analisados usando o software DataAnalysis 4.3 (Bruker, Daltonik). A cada dez amostras, um branco contendo por solução extratora foi analisado.

As amostras foram preparadas a partir dos extratos secos utilizando 1,0 mg de extrato em 1,0 mL do respectivo solvente de cada extração, foram filtradas com filtro de seringa com membrana filtrante de PTFE (0,22 μm), acondicionada em *vials* de vidro de 2,0 mL e, em seguida, analisadas. Após a otimização, os extratos foram obtidos e foram acondicionados em frascos pesados e vedados, colocados em freezer à -26°C .

3.3.2. Compostos de média-baixa polaridade

A obtenção das impressões digitais metabólicas dos extratos de Orchidaceae relacionadas à compostos de baixa polaridade nos proporcionou uma visão mais ampla sobre o metaboloma desta família. Através dessas, foi possível correlacioná-las com a filogenia e taxonomia (descrita no Capítulo 2 desta Tese), além de permitir a comparação com a literatura disponível sobre os compostos presentes em espécies exóticas e suas propriedades. O método de extração foi baseado nos estudos e protocolos de metabolômica de plantas de Beale *et.al.* (2018), Kanani *et.al.* (2008), Lisec *et.al.*, (2006), Liu *et.al.* (2017), Martucci *et.al.* (2018), Mushtaq *et.al.* (2014) e Wang *et.al.* (2018), com algumas modificações, observando-se parâmetros como quantidade de material vegetal, solução extratora, tempo de homogeneização, extração e centrifugação, processo de filtração e secagem das amostras. Abaixo está o procedimento completo de extração.

As amostras foram preparadas adicionando 2,5 mL de CH_2Cl_2 à 50,0 mg de droga vegetal, e foram homogeneizadas por 10 segundos em vórtex e colocadas em banho de ultrassom durante 20 minutos. Em seguida, o extrato foi centrifugado por 10 minutos à 2000 rpm, filtrado com filtro de membrana PTFE (0,22 μm) resistente ao solvente para um *vial* de vidro e o solvente foi evaporado em temperatura ambiente. Os extratos secos foram acondicionadas em freezer à -26°C até o momento da análise, no qual os extratos foram ressuspensos com CH_2Cl_2 para uma concentração de $5,0 \text{ mg.mL}^{-1}$.

3.4. Instrumentação e Análise dos Extrato

Os métodos utilizando para análise por cromatografia e espectrometria de massas foram baseados em protocolos envolvendo métodos exploratórios utilizados na

Metabolômica não-direcionada (*Untarget Metabolomics*) para análises por GC-MS e LC-MS, e estudos metabolômicos envolvendo plantas (*Plant Metabolomics*).

As análises foram realizadas em colaboração com o Prof. Norberto Peporine Lopes, da Faculdade de Ciências Farmacêuticas de Ribeirão Preto (FCFRP-USP), na Central de Espectrometria de Massas de Micromoléculas Orgânicas (CEMMO).

3.4.1. Cromatografia líquida de ultrarrápida eficiência acoplada a espectrometria de massas – UFLC-DAD-ESI-MS/MS (*micrOTOF-QII*)

O método de análise por LC-MS/MS foi otimizado baseando-se nos estudos de Bauermeister et.al, (2019), Ernst et.al., 2014, De Vos et al. (2007), Liu et.al. (2017) e Sedio et.al. (2018). Cinco extratos foram selecionados para a otimização do método de análise: Vpa16, Cgi52, Epu61, Ede59 e Cos131, os mesmos selecionados na otimização do método de extração. A este grupo foram adicionados dois extratos de espécies do gênero *Dendrobium* (Dmo54 e Dfi53) e uma amostra constituída de uma fração alcaloídica enriquecida de Harmina e Harmalina (Fr.A.Ar), alcalóides da classe das β -carbolinas, solubilizada na mesma solução extratora das outras espécies. Essa fração foi cedida pela Profa. Dra. Larissa Rolim da Universidade Federal do Vale do Rio São Francisco, advinda do extrato metanólico obtido por maceração exaustiva da semente de *Peganum harmala* L. A adição desses na otimização foi realizada para analisar capacidade de detecção do método à alcalóides.

Os extratos foram analisados em um sistema UFLC-DAD-ESI-MS/MS (*micrOTOF-QII*) (Bruker Daltonics), que consiste em um cromatógrafo líquido de ultrarrápida eficiência acoplado a um espectrômetro de massas com fonte de ionização por *electrospray* (ESI) e analisador de massas tipo híbrido *micrOTOF/Duplo* Quadrupolo, da Bruker Daltonics, que utiliza o Tempo de Vôo (*Time Of Flight - TOF*) para a separação dos íons, e Duplo Quadrupolo (QII) para seleção e fragmentação dos íons selecionados.

As análises para a otimização do método foram realizadas no modo de ionização positivo, no qual foram detectados os íons precursores e seus perfis de fragmentação, gerados através da análise por MS/MS. A coluna utilizada foi uma C18 (Kinetex™, Phenomenex) de tamanho de partícula de 2,7 μm (150 x 4,6 mm). O detector de UV (DAD) foi programado na faixa de comprimentos de onda entre 100 e 900 nm. A temperatura do forno foi de 35 °C.

Os parâmetros observados na otimização do método cromatográfico foram o gradiente de eluição, volume de injeção, fluxo de solvente e fase móvel, e os parâmetros da fonte de ionização, voltagem do *End plate offset*, voltagem do capilar, pressão do nebulizador, temperatura e fluxo do gás de secagem.

Os parâmetros do espectrômetro de massas foram testados e otimizados com auxílio de outros integrantes do grupo de pesquisa NPPNS e técnicos da Bruker. A taxa de espectros (*spectra rate*) de 2,0Hz e faixa de massa de 50 a 1300 *m/z* foram mantidas durante toda a análise. Os parâmetros de MS/MS foram Auto MS/MS com a seleção dos íons precursores em ciclos de 3,0 s, excluindo a seleção após detecção de 2 espectros do mesmo íon ou após 0,60min, reconsiderando sua seleção se a intensidade do íon precursor for o dobro da anterior. Para as fragmentações dos íons precursores, foi utilizada uma rampa de energia de colisão de 20 a 50 eV.

A calibração do analisador foi realizada antes das análises e a calibração interna ocorreu em todas as análises, utilizando como calibrante o ácido trifluoroacético (TFA). Os dados foram gravados usando o software DataAnalysis 4.3 (Bruker, Daltonik).

O método de análise no modo de ionização negativo utilizou os valores da otimização do modo positivo, havendo ajustes somente na voltagem do *End plate offset* e voltagem do capilar. Em seguida, todos os extratos de média-alta polaridade foram analisados, obtendo as impressões digitais metabólicas com as informações necessárias para a construção das Redes Moleculares e utilização das ferramentas do GNPS e Carcará, além dos espectros de UV, valores de *m/z* e tempo de retenção, os quais auxiliam na anotação dos metabólitos no processo de desreplicação.

3.4.2. Cromatografia a gás acoplada a espectrometria de massas – GC-EI-MS QP2010 (Shimadzu)

O método de análise por GC-MS foi baseado nos estudos de Lisec *et.al.*, (2006), Kanani *et.al.* (2008), Liu *et.al.* (2017), Martucci *et.al.* (2018), Mushtaq *et.al.* (2014) e Wang *et.al.* (2018), com algumas modificações.

As análises dos extratos de média-baixa polaridade foram realizadas por GC-EI-MS QP2010 (Shimadzu), um cromatográfico a gás Shimadzu acoplado a um espectrômetro de massas com modo de ionização por impacto eletrônico (IE) e analisador do tipo quadrupolo (Shimadzu). Utilizou-se uma coluna capilar ZB-5MS (J &

W Agilent) (30 m x 0,25 mm x 0,25 mm, 0,25 μm) acoplada a uma pré-coluna Duraguard – 10 m, hélio como gás de arraste a 1,0 mL.min⁻¹, fonte de ionização à 250 °C, e volume de injeção 1,0 μL à 5,0 mg/mL.

As condições experimentais utilizadas foram: injeção split 1:5; temperatura do filamento, 230 °C e injetor, 270 °C; corrente de emissão de interferência, 60 μA e ionização por impacto de elétrons, 70 eV; programa de aquecimento da coluna: 0 – 3,0 min, 60 °C, rampa de temperatura até 330 °C (taxa de aquecimento, 10 °C.min⁻¹), e mantendo-se nesta temperatura por 20 min; $\Delta m/z$ de 40-600. Os dados obtidos foram analisados utilizando o software CGSolution e as bibliotecas NIST, Wiley e FFNSC. As impressões digitais metabólicas adquiridas estão no Apêndice II.

3.5. Processamento dos Dados

3.5.1. Impressões digitais metabólicas – Compostos média-alta polaridade

Os dados adquiridos nas análises por LC-MS/MS no modo de ionização positivo e negativo foram convertidos, separadamente, em arquivos .mzxml utilizando o software livre ProteoWizard - MSconvert (v.3, Proteowizard Software Foundation, Palo Alto, CA, EUA) (Chambers et.al. , 2012). Posteriormente, foram processados utilizando o software *MzMineTM* (v. 2.51, BMC Bioinformatics, Reino Unido) (Pluskal et.al., 2010). Os parâmetros utilizados no processamento foram otimizados a fim de obter cromatogramas alinhados, deconvoluídos, com mínima influência de ruído e isótopos, e que representem ao máximo os extratos analisados.

A otimização envolveu principalmente parâmetros de construção dos cromatogramas (*Chromatogram builder*) e deconvolução dos picos (*Chromatogram deconvolution*), observando as tendências de variação no número de íons precursores detectados (MS¹) e número dos íons precursores detectados que possuem perfil de fragmentação (MS²) de acordo com o parâmetro alterado.

Dentro dos modelos gerados, os mais interessantes foram selecionados para a construção de Redes Moleculares, com o intuito de avaliar a quantidade de íons anotados e a influência do parâmetro selecionado na construção de uma rede. Os valores utilizados para a construção da rede foram os indicados pela plataforma para uma análise exploratória, de acordo com o equipamento utilizado.

Após o processamento, foram exportados 3 arquivos de dados: arquivo .mgf, com as informações dos espectros de MS², um arquivo .csv, com a matriz de dados contendo os íons precursores selecionados para fragmentação (MS²) durante as análises de LC-MS/MS, as espécies em que aparecem, as áreas dos picos, razões massa-carga (*m/z*) e tempos de retenção, e outro arquivo .csv com uma matriz de dados contendo os íons precursores totais (selecionados ou não para a fragmentação) (MS¹) com os mesmo descritores do arquivo contendo MS².

Os dois primeiros foram enviados para a plataforma online do GNPS (<https://gnps.ucsd.edu>, Wang et al., 2016) através do software WinSCP (<https://winscp.net/eng/index.php>) e utilizados para a construção das redes moleculares. O terceiro arquivo foi utilizado das análises estatísticas multivariadas supervisionadas e não-supervisionadas.

3.5.2. Impressões digitais metabólicas – Compostos média-baixa polaridade

Os dados adquiridos nas análises por GC-MS foram transformados para o formato .cdf, importados para a plataforma GNPS e processados através da ferramenta “GC-MS *El Data Analysis – Data Processing – Deconvolution*”, (Aksenov et.al, 2021), a qual promove a deconvolução e alinhamento espectrais em todo o conjunto de dados utilizando o processamento através do algoritmo MSHub. Os parâmetros de processamento são ajustados automaticamente e internamente pelo MSHub, como os relacionados a simetria de pico e ajustes na linha de base. Os parâmetros envolvendo alinhamento dos cromatogramas em “*Advanced Clustering*” foram ajustados em: *Cluster_Spectra*, NO; *Time_Unit*, MIN; *RT_tolerance*, 120. Os dados produzidos foram inspecionados.

3.6. Redes Moleculares (*Molecular Networking*)

O processo de otimização da construção das redes moleculares avaliou cinco blocos de dados:

- **Grupo LC-PR:** dados brutos adquiridos por LC-MS/MS(+) (compostos de média-alta polaridade);
- **Grupo LC-PP:** dados adquiridos por LC-MS/MS(+) e processados pelo *MzMine*TM (compostos de média-alta polaridade);

- **Grupo LC-NR:** dados brutos adquiridos por LC-MS/MS(-) (compostos de média-alta polaridade);
- **Grupo LC-NP:** dados adquiridos por LC-MS/MS(-) e processados pelo *MzMineTM* (compostos de média-alta polaridade);
- **Grupo GC:** dados adquiridos por GC-MS e processados na plataforma GNPS (composto de média-baixa polaridade).

Os dados brutos das análises por LC-MS/MS (Grupo LC-PR e LC-NR) foram convertidos para o formato .mzXML usando software MSconvert (ProteoWizard) e enviados para a plataforma online do GNPS (<https://gnps.ucsd.edu>, Wang et al., 2016) pelo do software WinSCP (<https://winscp.net/eng/index.php>) para a construção das redes moleculares clássicas através da ferramenta “*Data Analysis – Molecular Networking*”.

Os arquivos .mgf e .csv (MS²), adquiridos após processamento dos dados (Grupo LC-PP e LC-NP), foram enviados para a plataforma online do GNPS pelo do software WinSCP, para a construção das redes moleculares utilizando a ferramenta “*Feature-Based Molecular Networking (FBMN)*” (Nothias et al., 2019).

Os dados de GC-MS processados pela plataforma GNPS (Grupo GC) foram exportados diretamente do processamento para a construção da rede molecular através da opção “*Continue Processing – Search Spectral Library and Molecular Network*”, em “*Job Status*”.

A construção das redes, em todos os grupos citados, utilizou um arquivo .csv de metadados, contendo informações sobre as espécies, como subfamília, tribo, subtribo, gênero, local de coleta, bioma e bioatividade, para a realização de análises de correlação.

Os parâmetros que foram utilizados na otimização das redes moleculares geradas a partir de dados de LC-MS estão citados a seguir:

- *Basic Options*
 - *Precursor Ion Mass Tolerance (PIMT)*: tolerância de massa de íons precursores;
 - *Fragment Ion Mass Tolerance (FIMT)*: tolerância de massa de íons fragmento.
- *Advanced Network Option*

- *Min Pairs Cos*: valor mínimo de cosseno entre um par de espectros consenso MS/MS (nodo);
- *Minimum Matched Fragment Ion (Min Matched Peaks)*: número mínimo de íons fragmento em comuns, que são compartilhados por dois espectros consenso MS/MS (nodo), para ser considerado espectros correspondentes;
- *Advanced Library Search Options*
 - *Library Search Min Matched Peaks*: número mínimo de íons fragmento compatíveis com a pesquisa na biblioteca espectral,
 - *Score Threshold*: valor mínimo de cosseno de similaridade que os espectros de MS/MS devem possuir em correspondência espectral com as bibliotecas espectrais para serem considerados uma anotação.

Além desses parâmetros, tanto na construção das redes clássicas quando pelo FBMN, as bordas (*edges*) entre dois nodos somente foram mantidas para os 10 nodos mais semelhantes (*Network TopK*), o tamanho máximo de uma família molecular (*cluster*) foi definido como 130 (*Maximum Connected Component Size*) e o tamanho mínimo como 2 (*Minimum Cluster Size*). A máxima diferença de massa entre os precursores foi de 500Da. As análises realizadas pela ferramenta de Rede Molecular Clássica utilizavam do algoritmo MS-Cluster, o qual realizou a combinação dos espectros de MS/MS repetidos por similaridade para criação de espectros consenso. E a ferramenta DEREPLICATOR foi utilizado nas análises dos FBMN para anotar os espectros de MS/MS (Mahimani et.al., 2018).

E, a seguir, os parâmetros ajustados na otimização das redes moleculares geradas a partir de dados de GC-MS:

- *Search Options*:
 - *Fragment Ion Mass Tolerance (FIMT)*: tolerância de massa de íons fragmento.
 - *Minimum Matched Peaks*: número mínimo de íons de fragmento em comum que um espectro de MS/MS deve possuir para ser considerado uma anotação de biblioteca espectral.
 - *Score Threshold*: valor mínimo de cosseno de similaridade para ser considerado uma anotação na biblioteca espectral.

- *Advanced Search Options*
 - *Top Hits Per Spectrum*: Número de resultados principais, classificados pela pontuação de cosseno, devem ser incluídos nos resultados de anotações finais. É recomendado pelos desenvolvedores manter esse valor relativamente alto (por exemplo, 10 para manter as 10 principais correspondências com a biblioteca) para aumentar a possibilidade de uma anotação correta, caso não esteja indicada na primeira ocorrência.

Nessas redes, as bordas (*edges*) entre dois nós foram mantidas na rede somente se fosse os 10 nodos mais semelhantes (*Network TopK*). O tamanho máximo da família molecular foi definido como 100 (*Maximum Connected Component Size*), e as bordas que possuíam valores menores foram removidas das famílias moleculares até que o tamanho estivesse abaixo desse limite. Além disso, foi selecionada a opção para cálculo do índice de Kovatz.

Os resultados avaliados na otimização para comparação entre modelos foram: número de nodos, número de nodos anotados e distribuição dos valores de cosseno de similaridade com a biblioteca (*Score Threshold*), sendo os dois últimos os mais importantes.

A visualização das redes moleculares foi realizada no Cytoscape 3.2 (Shannon et al., 2003), o qual permitiu a visualização na forma de uma rede de agrupamentos de nodos (íons precursores) e arestas (valores de cosseno de similaridade entre análogos). Cada agrupamento de nodos corresponde a uma família molecular (*cluster*).

4. RESULTADOS E DISCUSSÕES

4.1. Análise Metabolômica - Compostos de média-alta polaridade

4.1.1 Otimização do Método de Extração

Estudos sobre de espécies da família Orchidaceae não são recentes (Hossain, 2011). Os metabólitos secundários foram principalmente descritos em espécies exóticas, ex. *Dendrobium sp.*, e foi identificada a presença de flavonóides, alcalóides, antocianinas, terpenos, quinonas, fenatrenos, cumarinas, sesquiterpenóides, dentre outros (Cota et al. 2008; Gutiérrez et al., 2010; Ramos et al., 2012; Silva et al., 2003). No entanto, estudos sobre as espécies nativas do Brasil e América do Sul permanecem escassos.

Com o objetivo de obter um extrato com grande diversidade metabólica, foi realizada a otimização do método de extração e foram avaliados os rendimentos das extrações e os perfis químicos adquiridos por LC-MS.

Foram avaliados quatro parâmetros: proporção de solvente (9:1 (1), 7:3 (2) e 5:5 (3)), particionamento do com *n*-hexano, utilização de ácido no líquido extrator (0,1% de ácido fórmico) e tipo de solvente (metanol ou etanol). Após secagem e pesagem dos extratos, foram obtidos os rendimentos de extração. O gráfico 1 indica os valores de média e desvio padrão dos rendimentos de extração das 5 espécies estudadas de acordo com cada bloco de extração.

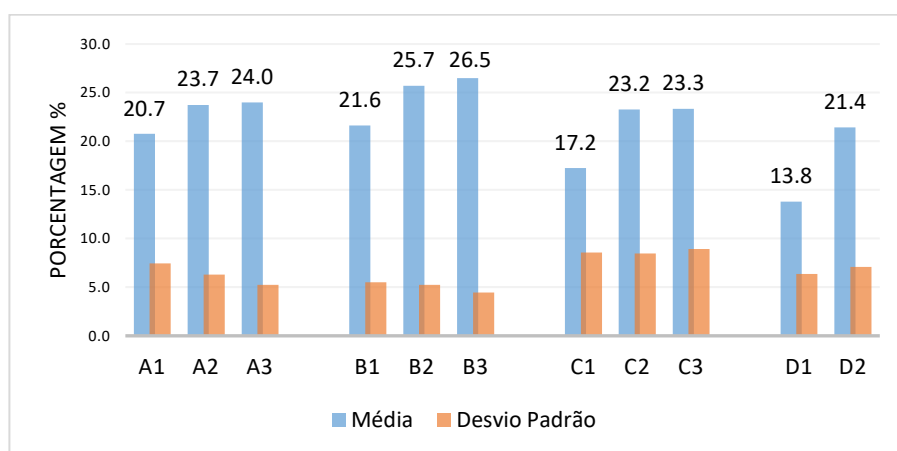


Gráfico 1. Média e desvio padrão dos rendimentos de extração das 5 espécies estudadas de acordo com o tipo de extração realizada. **A.** Metanol/Água. **B.** Metanol/Água em 0,1% de ácido fórmico. **C.** Metanol/Água e *clean up* (*n*-hexano). **D.** Etanol/Água. **Proporção de solvente:** (1) 9:1, (2) 7:3 e (3) 5:5.

As soluções hidroalcoólicas possuem a capacidade de extrair substâncias de ampla faixa de polaridade, sendo o metanol o solvente mais utilizado nos protocolos e estudos metabolômicos (De Vos *et al.*, 2007; Liu *et al.*, 2017; Mushtaq *et al.*, 2014; Sedio *et al.*, 2018; Kim & Veeport, 2010). Em De Vos *et al.*, 2007, há a indicação de etanol como um possível solvente para extrações metabolômicas, além disso este é um solvente de menor toxicidade. Por isso, foi inserida a extração com etanol para comparação. Observando os valores de rendimento, todas as extrações realizadas com metanol obtiveram um rendimento maior do que a extração utilizando etanol. Os extratos produzidos com metanol demonstraram possuir uma maior quantidade e intensidade de picos, incluindo íons de massa par dentre os tempos de retenção de 2 a 5 minutos, e houve maior resolução dos picos cromatográficos (Figura 1.B).

Dentre as três proporções dos solventes metanol:água utilizadas nos blocos de extração A, B e C, as que resultaram maiores rendimentos foram 7:3 (2) e 5:5 (3), com discreta diferença entre eles (Gráfico 1). Ao avaliar a diversidade de picos cromatográficos, a Figura 1.A mostra que o extrato de proporção 5:5 (azul) possui menos picos e intensidades menores a partir de 12,5 minutos do que os outros. As proporções 7:3 (verde) e 9:1(vermelho) possuem os mesmos picos durante toda a análise, com poucas diferenças na intensidade. Dessa maneira, a proporção 7:3 foi selecionada, por possuir melhor rendimento, diversidade e intensidade de picos no perfil cromatográfico, além de ser uma proporção a qual auxilia a extrair de alcaloides e flavonoides ionizados pela a maior proporção de água.

Comparando as extrações realizadas com e sem o particionamento utilizando *n*-hexano (*clean up*), não houve diferenças significativas no rendimento de extração, principalmente entre as proporções 7:3 e 5:5 de solvente. A diferença está no rendimento utilizando a proporção 9:1, sendo que essa proporção permite a extração de componentes mais apolares e que podem ter sido arrastados para a partição hexânica (Figura 1.C).

Foi observado um leve aumento no rendimento da extração que utilizou 0,1% de ácido fórmico, mas de grandeza muito próxima do desvio padrão. Analisando os perfis das amostras extraídas com solução acidificada (marrom) e não acidificada (verde), não houve mudança no perfil cromatográfico (Figura 1.D). Devido à ausência de diferenças significativas utilizando *clean up* e solução extratora ácida, optou-se por não realizar o *clean up* e utilizar solução extratora sem ácido, prevenindo perdas de substâncias de menor polaridade para a fase *n*-hexânica e favorecimento da extração de moléculas ionizadas, além da diminuição da extração de moléculas pouco ionizáveis.

Portanto, o método selecionado para a extração das espécies de Orchidaceae deste estudo foi a extração utilizando solução extratora hidrometanólica (7:3), sem adição de ácido e sem realizar particionamento com *n*-hexano (*clean up*).

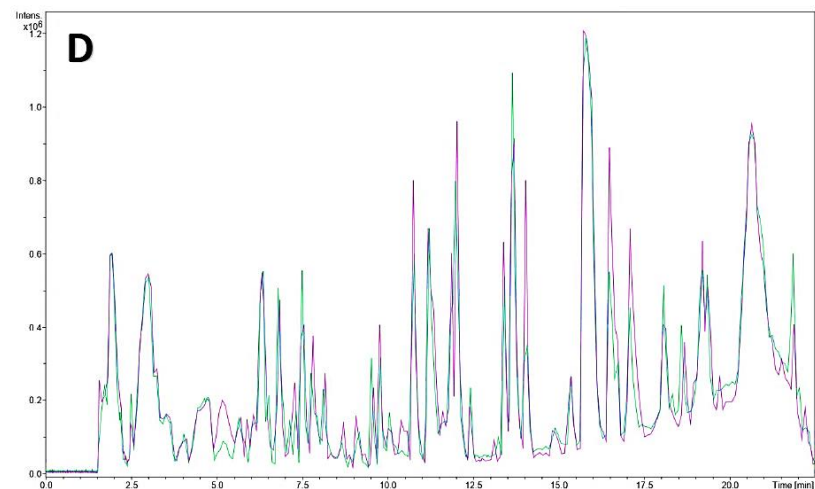
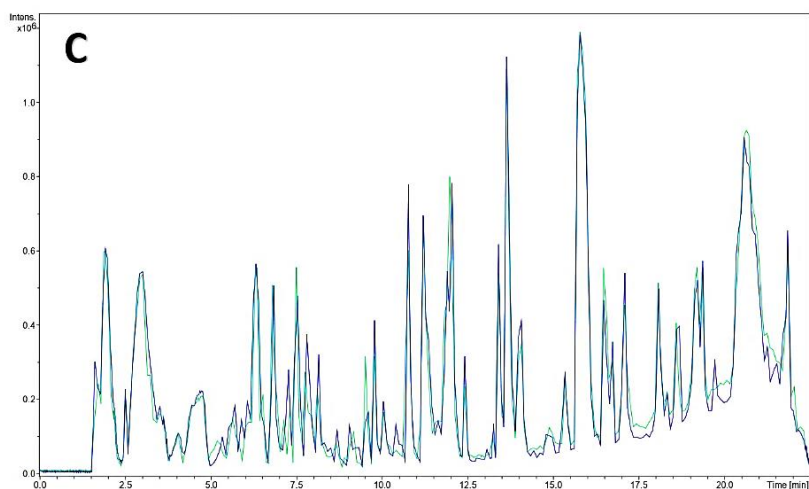
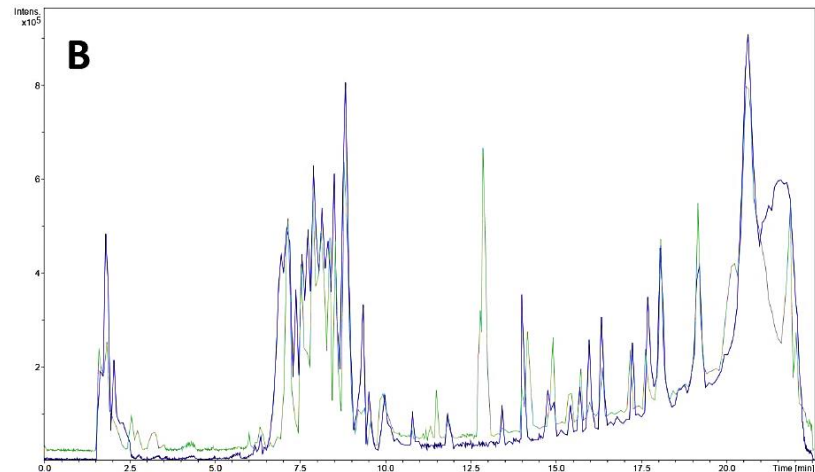
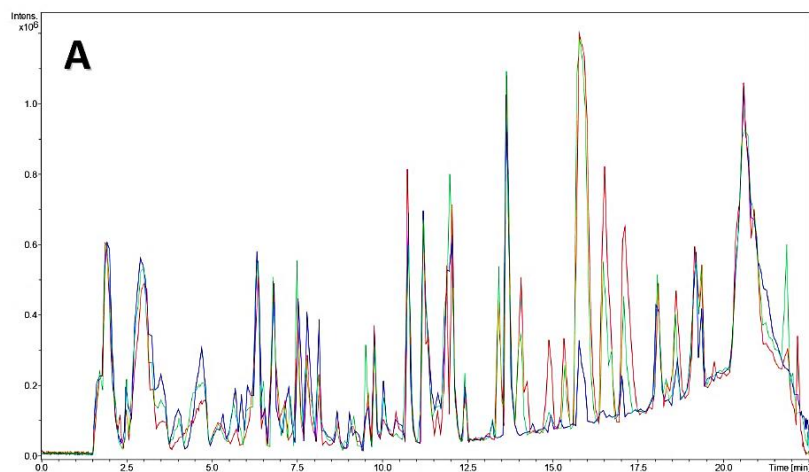


Figura 1. Perfis cromatográficos das análises por UFLC-DAD-ESI(+)-MS/MS (microTOF-QII) das extrações realizadas na otimização do método de extração. **A.** Extrações feitas da espécie Cos131 variando a proporção de solução extratora Metanol/Água. *Vermelho:* (9:1). *Verde:* (7:3). *Azul:* (5:5). **B.** Extrações feitas da espécie Epu61 utilizando metanol (verde) e etanol (roxo) como solvente orgânico da solução extratora de proporção 7:3. **C.** Extrações da espécie Cos13 utilizando com Metanol/Água (7:3) com *clean up* (azul) e sem *clean up* (verde). **D.** Extrações da espécie Cos131 utilizando Metanol/Água 7:3 (verde) e Metanol/Água em 0,1% de ácido fórmico (7:3) (rosa).

4.1.2. Otimização do Método de Análise dos Extratos por UFLC-DAD-ESI-MS (micrOTOF-QII)

A otimização do método cromatográfico e dos parâmetros da análise por espectrometria de massas foi realizada visando boa resolução cromatográfica e separação dos picos, a detecção do maior número de substâncias possível, a detecção do maior número de moléculas protonadas e da capacidade em se produzir espectros de fragmentação fidedignos. Além disso, priorizou-se o uso de um método exploratório e abrangente, que fosse possível utilizá-lo na análise de todas as espécies a serem estudadas, incluindo a possível presença de alcalóides nos extratos.

As amostras Dmo54 e Dfi53 foram inseridas neste estudo por pertencerem ao gênero *Dendrobium*, conhecido por deterem principalmente metabólitos como sesquiterpenos, alcalóides, compostos fenólicos, polissacarídeos bioativos e mais de 19 tipos de alcalóides denominados as dendrobinas (Dan Yang *et al.*, 2018). Utilizando essas amostras e a Fr.H.Hr, foi possível avaliar a aptidão do método em detectar a possível presença de alcaloides.

Os parâmetros do método cromatográfico, como gradiente de eluição, volume de injeção, fluxo de solvente e fase móvel, e da fonte de ionização, como a voltagem do *End plate offset*, voltagem do capilar, pressão do nebulizador, temperatura e fluxo do gás de secagem, foram otimizados e se encontram na tabela 2, sendo que os valores selecionados na otimização seguem destacados em negrito.

Utilizar o gradiente crescente de eluição, de 3 a 100% de CH₃CN, permitiu uma boa separação dos picos, diminuindo as coeluições, e foi possível detectar compostos em uma ampla faixa de polaridade em um mesmo extrato, principalmente utilizando o gradiente em um tempo de 40 minutos (Fig. 2.B). Por serem amostras complexas e oriundas de uma família de grande diversidade química, fez-se necessário um método abrangente e mais longo para promover uma boa separação e resolução dos picos cromatográficos, auxiliando assim na desreplicação via *Molecular Network* pela plataforma GNPS e pela biblioteca de compostos *in house*.

Durante a otimização foi observado que para a concentração de 1,0 mg/mL dos extratos era necessário um ajuste no volume de injeção. Considerando que o volume de injeção correspondente a 0,1% do volume total da coluna utilizada é de 25 µL, foram testados valores próximos a esse. Observou-se que com um volume de injeção de 25µL da amostra preparada à concentração de 1,0 mg/mL, foi possível obter um

aumento de intensidade que proporcionava boa resolução, boa intensidade, além de favorecer a detecção dos íons fragmento.

Tabela 2. Parâmetros e valores otimizados para a análise por UFLC-DAD-ESI-MS/MS (microTOF-QII) para a construção da Rede Molecular. Destacados em negrito e itálico: parâmetros selecionados.

Parâmetros do Método Cromatográfico			
Gradiente de eluição	0-2 min: 3% B	<i>0-2 min: 3% B</i>	0-2 min: 3% B
	2-20 min: 3-100% B	<i>2-30 min: 3-100% B</i>	2-40 min: 3-100% B
	20-44 min: 100% B	<i>30-34 min: 100% B</i>	40-44 min: 100% B
	24-27 min: 100-3% B	<i>34-37 min: 100-3% B</i>	44-47 min: 100-3% B
	27-30 min: 3% CH ₃ CN	<i>37-40 min: 3% B</i>	47-50 min: 3% B
	B: CH ₃ CN	<i>B: CH₃CN</i>	B: CH ₃ CN
	Calibrante: 27-30 min	<i>Calibrante: 37-40 min</i>	Calibrante: 47-50 min
Tempo de Análise	30min, <i>40min</i> , 50 min		
Volume de injeção	5 µL, 10 µL, 15 µL, <i>25µL</i> , 30 µL		
Fluxo de solvente	0.5 mL/min, 0.8 mL/min, <i>1.0 mL/min</i>		
Fase móvel	CH ₃ CN e H ₂ O CH ₃ CN e H ₂ O-HCO ₂ H (0,1%) <i>CH₃CN-HCO₂H (0,1%) e H₂O-HCO₂H (0,1%)</i>		
Padrões internos	Prolina, isoleucina, valina, <i>hidrocortizona</i>		
Parâmetros da Fonte de Ionização (ESI)			
Voltagem do <i>End Plate Offset</i>	<i>450 V</i> , 500 V		
Voltagem do Capilar	3400 V, 3500 V, <i>3600 V</i>		
Nebulizador	4,0 bar, 4,5 bar, <i>5,0 bar</i>		
Vazão do gás de secagem	9,0 L/min, <i>10,5 L/min</i>		
Temperatura do gás de secagem	200 °C, <i>220 °C</i>		

A utilização da fase móvel acidificada auxiliou na detecção de grande quantidade de íons durante a toda a análise, além da boa resolução dos picos. Na Figura 2.C é possível observar que na análise utilizando CH₃CN-HCO₂H 0,1% e H₂O-HCO₂H 0,1% foram detectados um número maior de picos e com melhor resolução, comparando-se com as análises utilizando CH₃CN e H₂O-HCO₂H (0,1%) e CH₃CN e H₂O.

A vazão da fase móvel estabelecida em 1,0 mL/min permitiu uma boa separação dos picos cromatográficos, além de uma boa resolução e distribuição ao

longo do gradiente. Em função desta vazão, foram variados os valores de pressão do gás nebulizador, vazão e temperatura do gás de secagem, de acordo com as recomendações do fabricante. Através dessas alterações, foi possível observar uma boa ionização e detecção dos íons, baixo ruído na linha de base, e foram evitadas as fragmentações na fonte de ionização.

A seleção do padrão interno a ser utilizado nas análises buscou padrões que tivessem boa ionização, localização de eluição e intensidade adequadas, para que não suprimisse a ionização de outros metabólitos, e boa resolução de pico para servir como referência para as análises estatísticas. Dentre os padrões internos analisados, a hidrocortisona, a 10 µg/mL, mostrou melhor ionização, intensidade de pico e detecção que os outros padrões, sendo que seu tempo de retenção em 13 minutos não coincidiu com outros íons observados neste gradiente de análise.

O método otimizado demonstrou ser adequado para a detecção de alcalóides, já que foram detectados íons de massa par em todas as amostras (grande indício da detecção de certas classes de alcaloides). Os alcaloides de massa ímpar presentes na Fr.H.Hr, como a Harmina (m/z 213,1131) e Harmalina (m/z 215,1305) foram detectados em 9,7 min e 9,9 min respectivamente, não havendo perdas no volume morto da coluna ou baixa ionização que comprometeria a detecção destas e de outras substâncias desta classe (Figura 2.D).

Após a otimização do método, as análises das 131 espécies de Orchidaceae foram realizadas. A Figura 3 e 4 representam as impressões digitais metabólicas adquiridas por UFLC-DAD-ESI-MS/MS (micrOTOF-QII) no modo positivo e negativo, respectivamente, após a otimização, e foi observada uma maior quantidade de picos de alta intensidade na região entre 8 a 15 minutos. Também foi observado que metabólitos foram detectados ao longo de toda a análise, demonstrando diversidade química e boa representatividade química do extrato na impressão digital metabólica. As impressões digitais metabólicas das espécies estudadas estão no Apêndice III.

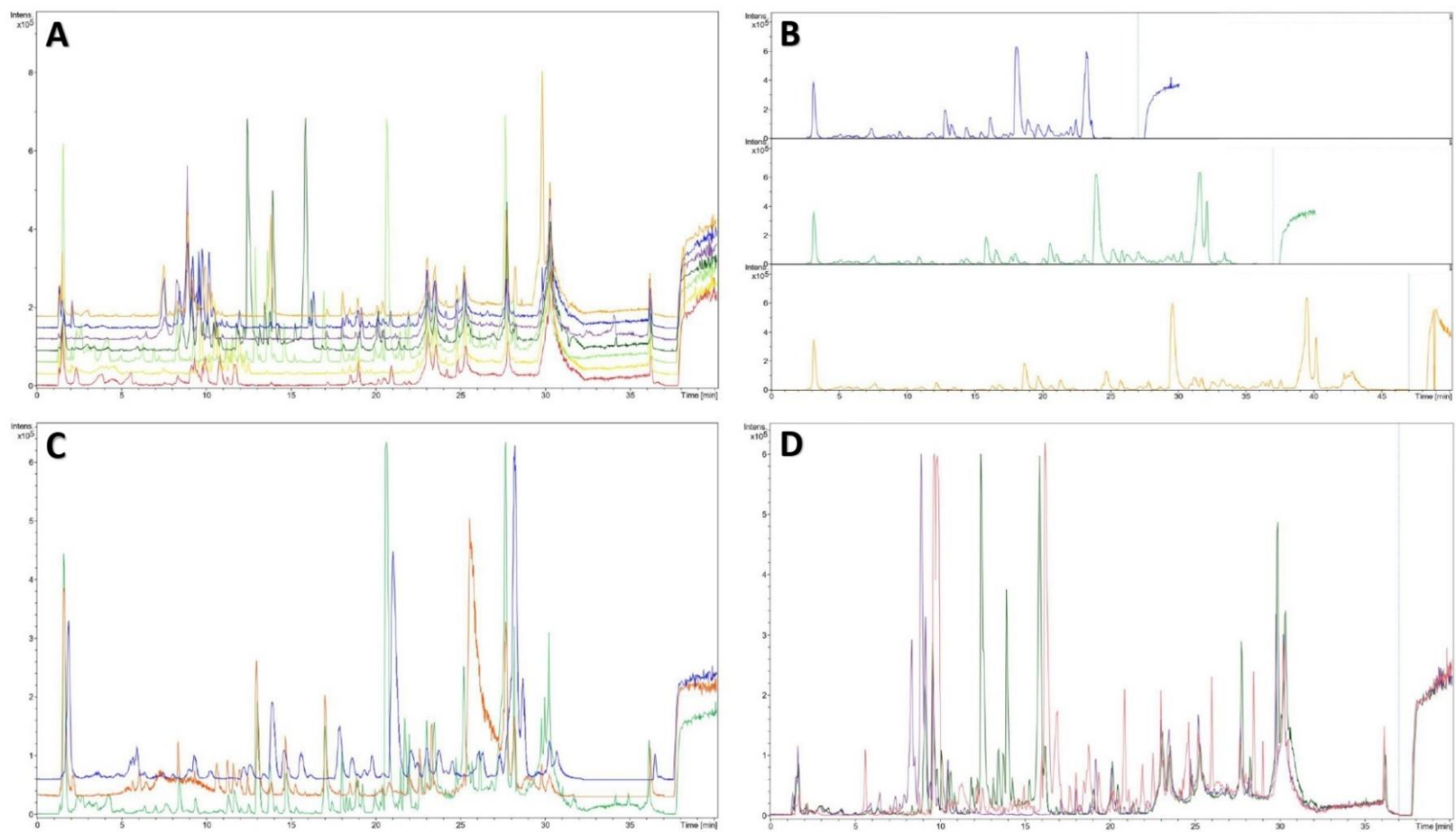


Figura 2. Perfis cromatográficos obtidos na otimização do método de análise por UFLC-DAD-ESI-MS/MS (microTOF-QII) para a construção da Rede Molecular. **A.** Perfis obtidos após a otimização do método de análise. Vermelho: Vpa16. Azul: Epu61. Laranja: Ede59. Roxo: Dmo64. Verde escuro: Dfi53. Verde Claro: Cos131. Amarelo: Cgi52. **B.** Perfis obtidos na análise de Cos131 em diferentes gradientes de eluições e tempos de análise: 50 minutos (laranja), 40 minutos (verde) e 30 minutos (azul). **C.** Perfis obtidos na análise de Cos131 em três diferente tipos de fase móvel. Azul: CH₃CN e H₂O-HCO₂H (0,1%). Laranja: CH₃CN e H₂O. Verde: CH₃CN-HCO₂H (0,1%) e H₂O-HCO₂H (0,1%) **D.** Perfis obtidos das amostras Dfi53 (verde), Dmo54 (roxo) e FrHHr (rosa).

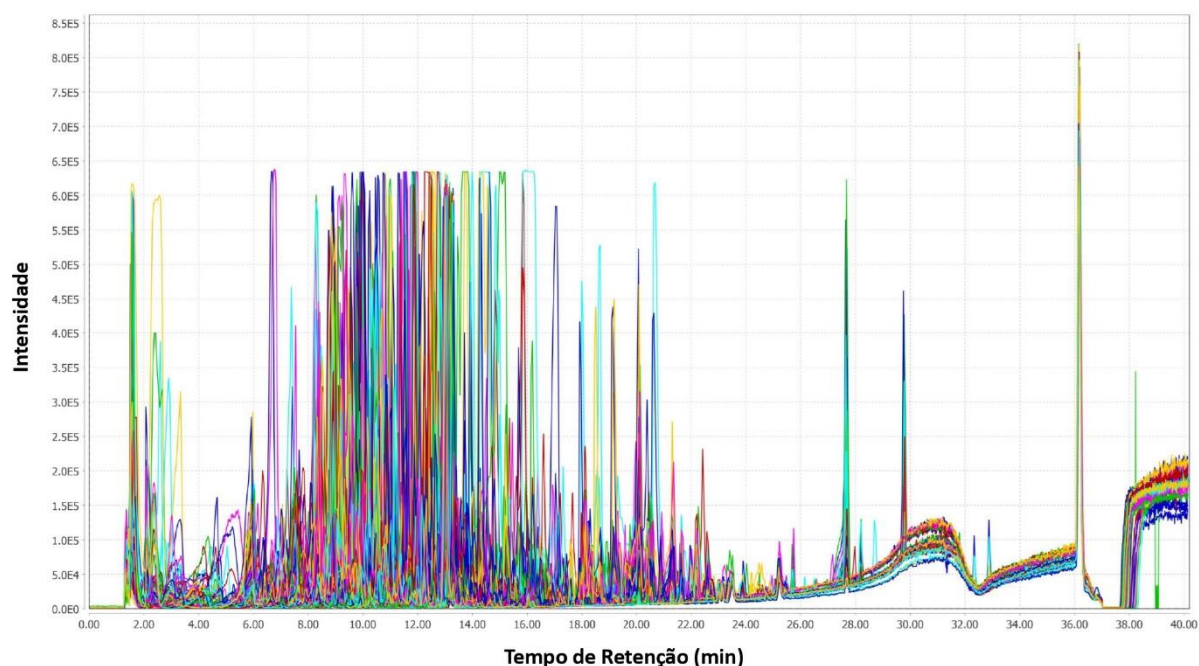


Figura 3. Impressões digitais metabólicas das 131 espécies de Orchidaceae adquiridas no modo positivo por UFLC-DAD-ESI-MS/MS (microTOF-QII) após a otimização dos métodos de análise. Software de visualização utilizado: *MzMine*TM (v. 2.51, BMC Bioinformatics, Reino Unido).

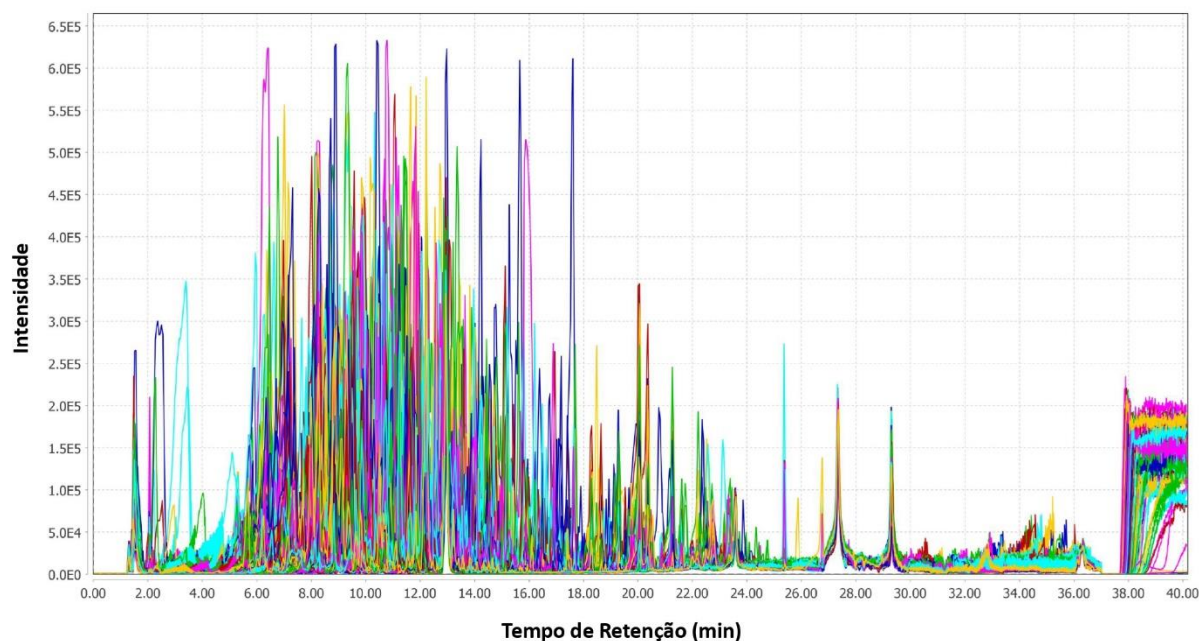


Figura 4. Impressões digitais metabólicas das 131 espécies de Orchidaceae adquiridas no modo negativo por UFLC-DAD-ESI-MS/MS (microTOF-QII) após a otimização dos métodos de análise. Software de visualização utilizado: *MzMine*TM (v. 2.51, BMC Bioinformatics, Reino Unido).

4.1.3. *Processamento das Impressões Digitais Metabólicas - LC-MS*

O processamento de dados é um dos passos mais importantes para as análises de dados. É através dele que se obtém a matriz de dados contendo os íons precursores presentes no grupo de amostras e os respectivos valores de área do pico, massa-carga e tempo de retenção, além dos espectros de fragmentação gerados pelas análises de MS/MS. Esses dados são utilizados na construção das redes moleculares através da ferramenta FBMN, além de outras ferramentas disponíveis na plataforma GNPS, e em análises de correlação *in silico* através de estatísticas multivariadas supervisionadas e não-supervisionadas.

O processamento dos dados adquiridos por LC-MS, no modo de ionização positivo e negativo, foi otimizado envolvendo principalmente os parâmetros de construção dos cromatogramas e deconvolução dos picos. Foram gerados 45 modelos diferentes, comparando os modelos em relação à quantidade de íons precursores detectados, número de íons precursores detectados com perfil de fragmentação, e o aspecto dos picos e dos sinais ruído. Dentre esses, 29 modelos de processamento (dados de ionização no modo positivo e negativo) foram selecionados para a construção de redes moleculares e avaliação da quantidade de nodos anotados (Tabela 3).

Os modelos de processamento POS-7 e NEG-5 foram selecionados devido à grande quantidade de íons precursores totais e de íons com perfis de fragmentação, e ao maior número de nodos anotados, com alta compatibilidade com a biblioteca de espectros da plataforma GNPS. Os parâmetros utilizados em cada modelo selecionados após a otimização estão nas Tabelas 4 e 5.

Tabela 3. Modelos de otimização do pré-tratamento gerados pelo Mzmine e selecionados para construção das redes moleculares, com seus respectivos números de íons precursores, íons precursores filtrados e substâncias anotadas.

<i>Modelo</i>	<i>Nº de Íons Precursores</i>	<i>Nº de Íons Precursores Filtrados*</i>	<i>Nº de Nodos Anotados</i>
POS-1	7796	3185	456
POS-2	4109	2178	411
POS-3	18064	8828	895
POS-4	13022	6766	799
POS-5	9105	4649	635
POS-6	4459	2306	387
POS-7	24518	10617	921
POS-8	17432	8308	801
NEG-1	17819	2648	15
NEG-2	11559	1768	10
NEG-3	9023	1336	11
NEG-4	17097	2469	13
NEG-5	17070	2473	15
NEG-6	12753	1829	13
NEG-7	10665	1562	12
NEG-8	10667	1562	12
NEG-9	9015	1337	12
NEG-10	12766	1833	13
NEG-11	17491	2711	14
NEG-12	36319	4851	10
NEG-13	24984	3878	11
NEG-14	18188	2755	11
NEG-15	32908	4134	12
NEG-16	18163	2748	11
NEG-17	18188	2760	11
NEG-18	23576	3427	11
NEG-19	32881	4123	11
NEG-20	19818	2617	10
NEG-21	19812	2614	10

*Íons detectados que foram selecionados para fragmentação e possuem, portanto, perfis de fragmentação.

Tabela 4. Parâmetros utilizados no software *MzMine*TM para o processamento dos dados adquiridos no modo de ionização positivo por análise em UFLC-DAD-ESI-MS/MS (microTOF-QII).

Raw data methods/ Peak detection	<p>Mass detection: Scans: Retention Time: 0-33 min; MS level: MS1 Mass Detector: Centroid - Noise level= 1.0E3</p> <p>Mass detection: Scans: Retention Time: 0-33 min; MS level: MS2 Mass Detector: Centroid - Noise level= 1.0E1</p> <p>Chromatogram builder: Scans: Retention Time: 0-33 min; MS level: MS1 Min time span (min)= 0.01; Min height= 3.0E3; m/z tolerance= 20,0 ppm</p>
Peak list methods	<p>Chromatogram deconvolution: <u>Baseline cut-off:</u> Min peak height = 3.0E3; peak duration= 0.01 – 4.0; baseline level= 3.0E3 <i>m/z</i> range for MS2 scan pairing (Da): 0.01; RT range for MS2 scan pairing (min): 0.2</p> <p>Isotopic peaks grouper: <i>m/z</i> tolerance= 20 ppm; Retention time tolerance= 0,2 min; Maximum charge= 3; Representative isotope= most intense.</p> <p>Alignment: <u>Join aligner:</u> <i>m/z</i> tolerance= 20 ppm; Weight for <i>m/z</i>= 75; Retention time tolerance= 0.4 min; weight for RT= 25</p> <p>Filtering: <u>Peak list rows filter:</u> Keep only peaks with MS2 scan (GNPS): OK Reset the peak number ID: OK</p>

Legenda: min, minutos; *m / z*, massa/carga; ppm, partes por milhão.

Tabela 5. Parâmetros utilizados no software *MzMine*TM para o processamento dos dados adquiridos no modo de ionização negativo por análise em UFLC-DAD-ESI-MS/MS (microTOF-QII).

Raw data methods/ Peak detection	<p>Mass detection: Scans: Retention Time: 0-33 min; MS level: MS1 Mass Detector: Centroid - Noise level= 1.0E3</p> <p>Mass detection: Scans: Retention Time: 0-33 min; MS level: MS2 Mass Detector: Centroid - Noise level= 1.0E1</p> <p>ADAP Chromatogram builder: Scans: Retention Time: 0-33 min; MS level: MS1 Min group size in # of scan= 5 ; Group intensity threshold= 1.0E3; Min highest intensity= 1.5E3; m/z tolerance= 20,0 ppm</p>
Peak list methods	<p>Chromatogram deconvolution: <u>Wavelets (ADAP):</u> S/N threshold= 10; S/N estimator= Intensity window S/N; min feature height= 2.0E3; coefficient/area threshold= 10; Peak duration range= 0.01 – 4.0; RT wavelet range= 0.001-0.1 m/z center calculation: MEDIAN; m/z range for MS2 scan paring (Da): 0.01; RT range for MS2 scan paring (min): 0.2</p> <p>Isotopic peaks grouper: m/z tolerance= 20 ppm; Retention time tolerance= 0,2 min; Maximum charge= 3; Representative isotope= most intense.</p> <p>Alignment: <u>Join aligner:</u> m/z tolerance= 20 ppm; Weight for m/z= 75; Retention time tolerance= 0.2 min; weight for RT= 25</p> <p>Filtering: Peak list rows filter: Keep only peaks with MS2 scan (GNPS): OK Reset the peak number ID: OK</p>

Legenda: min, minutos; *m / z*, massa/carga; ppm, partes por milhão.

Pode-se observar na Tabela 3 que a quantidade de íons precursores selecionados para fragmentação (filtrados) modo de ionização negativo é menor. Isso se deve pela menor intensidade dos íons contidas nos extratos no modo negativo, havendo assim um menor número de íons disponíveis na intensidade mínima para serem selecionados para fragmentação.

Também observamos a grande discrepância entre o número de substâncias anotadas para o modo positivo e negativo, ocorrida devido à detalhes técnicos relacionados à biblioteca de espectros da plataforma GNPS. Nesta biblioteca, a maioria das substâncias a compõe foram analisadas no modo de ionização positivo,

havendo uma gama muito maior de perfis de fragmentação desse modo para a análise por compatibilidade de espectros.

Nesta otimização, foram utilizados 2 tipos de parâmetros para a construção dos cromatogramas: *ADAP Chromatogram builder* e *Chromatogram builer*. Dependendo da versão utilizada do software. Neste caso, módulo ADAP inclui parâmetros como o tamanho mínimo do grupo em números de scans (*Min group size in # of scan*) e a intensidade da linha de corte do grupo (*Group intensity threshold*).

Esta etapa constrói cromatogramas de íons extraídos para os íons que foram detectados continuamente durante as análises por espectrometria de massas. No módulo ADAP, deve ser selecionado pelo menos um número de scans sequenciais (*Min group size in # of scan*) com íons de intensidade acima do limite da linha de corte definida para o grupo (*Group intensity threshold*). Mesmo que o valor ideal dependa da configuração do sistema cromatográfico, a melhor maneira de determina-lo é observar os dados brutos e determinar qual é o intervalo de tempo típico dos picos cromatográficos. Este parâmetro é muito similar ao Menor valor de intervalo de tempo dos picos (*Min time span (min)*) no módulo *Chromatogram builder*. (Myers et.al., 2017)

A otimização da deconvolução dos cromatogramas também utilizou dois tipos de parâmetros: Corte da linha de Base (*Baseline cut-off*) e o algoritmo *Wavelets (ADAP)*. O Corte da linha de base permite determinar a altura mínima de pico e o valor da linha de base para os íons precursores, o qual elimina o ruído e sujeira, além da duração dos picos em minutos. Além disso é possível determinar a faixa de m/z e tempo de retenção para o pareamento dos íons fragmento. O *Wavelets (ADAP)* permite aprofundar na eliminação do ruído, pois determina uma razão entre ruído e pico para considerar aquele sinal como um pico verdadeiro (Myers et.al., 2017).

Os valores selecionados para os parâmetros acima permitiram observar variações no número de íons detectados, comparando com o número de íons selecionados para fragmentação e até o número de nodos anotados. Assim, foi possível observar se no alto número de íons precursores detectados não havia uma grande porcentagem de sinais ruído.

Nenhum dos módulos, com o seu ADAP, foram excluídos pelos fluxos de trabalho indicados na plataforma GNPS na sessão de documentação (*Documentation*). Com isso, foram feitos testes nesses dois módulos para ambos os modos de ionização. Para o modo negativo, esse aprofundamento na eliminação do

ruído que o módulo ADAP permite a produção de cromatogramas e listas de íons muito mais fidedignas aos extratos, devido à baixa intensidade de íons precursores e, portanto, baixa intensidade dos íons fragmento. Para o modo positivo, utilizar a clássica construção e deconvolução trouxe boas listas de íons com alto nível de compatibilidade com a biblioteca.

4.1.4. Construção das Redes Moleculares

Devido à grande complexidade química das matrizes biológicas, ferramentas quimiométricas e computacionais vem sendo desenvolvidas para auxiliarem no processo de anotação molecular, promovendo a extração de informações dos dados adquiridos e as comparando com bancos de dados espectrais de substâncias, sendo a qualidade dos dados e quantidade de espectros disponíveis nos bancos fatores importantes para uma anotação fidedigna (Pilon et.al.,2020; Carnevale Neto et.al.,2016). Uma dessas ferramentas é a Rede Molecular, e sua construção e outras ferramentas associadas e essa procedeu na plataforma online e de livre acesso da *Global Natural Products Social Networks – GNPS* (<http://www.gnps.ucsd.edu>).

Organizar os dados em redes moleculares facilita a observação de grandes conjuntos de dados espectrais baseado na similaridade nos perfis de fragmentação (espectros MS/MS) de diferentes íons precursores, agrupando classes e metabólitos que partilham de perfis espectrais semelhantes. Dois conjuntos de dados podem ser utilizados para a construção das redes: dados processados e não-processados.

A construção e redes moleculares utilizando dados não-processados é realizada através da ferramenta *Classical Molecular Networking* (MN clássico), que inicialmente realiza o agrupamento de espectros pertencentes a um mesmo pico e/ou mesmo composto à nível de similaridade maior que 99%, utilizando o algoritmo *MS-Cluster* (Frank et.al., 2008). Esse analisa todos os espectros de MS/MS resultantes de íons que se enquadram na tolerância de massa de íons precursores definida e mescla os espectros de MS/MS quase idênticos (acima da pontuação de cosseno selecionada) em um único espectro MS/MS, chamado espectro consenso. Geralmente, cada espectro consenso consiste em vários espectros MS/MS de advindos dos vários arquivos de dados que compõe de amostras do conjunto (Wang et.al., 2016).

A construção das redes utilizando os dados processados é realizada através do *Feature-Based Molecular Networking* (FBMN), uma ferramenta complementar ao MN clássico de análise dentro da infraestrutura que a Rede Global Molecular Social de Produtos Naturais (GNPS) que se baseia em ferramentas de detecção e alinhamento de recursos cromatográficos, como Mzmine.

Após o processamento é obtida uma lista de íons precursores onde seus dados estão alinhados, deconvoluídos, livres de ruído e redundâncias de moléculas MS/MS, com resolução de compostos isômeros e que oferece valores de área de picos (dados para uma quantificação mais precisa) e tempo de retenção (formato .csv), e os respectivos espectros de fragmentação (formato .mgf). O FBMN aproveita toda essa informação gerada pelo software de processamento e aperfeiçoa o MN clássico, gerando redes que podem distinguir isômeros de perfis de fragmentação semelhantes que são somente diferenciados por separação cromatográfica, e que podem ter sido agrupados no MN clássico, facilitando a anotação espectral e incorporando informações quantitativas que permitem uma análise estatística robusta (Nothias et.al, 2020).

Uma vasta otimização foi realizada na construção das redes moleculares, utilizando o MN clássico e FBMN e os dados obtidos nos modos de ionização positivo e negativo, com o intuito de obter redes que melhor representem o metaboloma das espécies estudadas, que anotem o máximo de nodos possível e que possua uma distribuição de valores de cosseno de similaridade com a biblioteca do GNPS adequada, garantindo maior confiabilidade de anotações. Através dela foi possível observar a influência do uso do processamento dos dados na construção da rede, na compatibilidade com os espectros depositados na biblioteca e na correlação dos *clusters* (agrupamentos) formados.

A otimização da construção das redes moleculares foi realizada separadamente para as 4 matrizes de dados geradas: LC-PR, LC-PP, LC-NR e LC-NP. Para cada grupo de dados foram construídos 29 modelos de otimização, cujos parâmetros e os valores utilizados estão na tabela 6.

Os resultados observados para comparação entre modelos e seleção foram: número de nodos, número de nodos anotados e distribuição dos valores de cosseno (*Score Threshold*) de similaridade com a biblioteca, sendo os dois últimos os mais importantes.

Tabela 6. Parâmetros e valores utilizados na otimização da construção das Redes Moleculares por MN clássico e FBMN.

Modelos	<i>Basic Options</i>		<i>Advanced Network Options</i>		<i>Advanced Library Search Options</i>	
	<i>Precursor Ion Mass Tolerance</i>	<i>Fragment Ion Mass Tolerance</i>	<i>Min Pairs Cos</i>	<i>Minimum Matched Fragment Ions</i>	<i>Library Search Min Matched Peaks</i>	<i>Score Threshold</i>
otm01	0.10	0.02	0.7	3	3	0.7
otm02	0.08	0.02	0.7	3	3	0.7
otm03	0.05	0.02	0.7	3	3	0.7
otm04	0.02	0.02	0.7	3	3	0.7
otm05	0.02	0.05	0.7	3	3	0.7
otm06	0.02	0.08	0.7	3	3	0.7
otm07	0.02	0.1	0.7	3	3	0.7
otm08	0.02	0.02	0.7	4	3	0.7
otm09	0.02	0.02	0.7	5	3	0.7
otm10	0.02	0.02	0.7	6	3	0.7
otm11	0.02	0.02	0.65	3	3	0.7
otm12	0.02	0.02	0.65	4	3	0.7
otm13	0.02	0.02	0.65	5	3	0.7
otm14	0.02	0.02	0.65	6	3	0.7
otm15	0.02	0.02	0.6	3	3	0.7
otm16	0.02	0.02	0.6	4	3	0.7
otm17	0.02	0.02	0.6	5	3	0.7
otm18	0.02	0.02	0.6	6	3	0.7
otm19	0.02	0.02	0.7	3	4	0.7
otm20	0.02	0.02	0.7	3	5	0.7
otm21	0.02	0.02	0.7	3	6	0.7
otm22	0.02	0.02	0.7	3	3	0.65
otm23	0.02	0.02	0.7	3	4	0.65
otm24	0.02	0.02	0.7	3	5	0.65
otm25	0.02	0.02	0.7	3	6	0.65
otm26	0.02	0.02	0.7	3	3	0.6
otm27	0.02	0.02	0.7	3	4	0.6
otm28	0.02	0.02	0.7	3	5	0.6
otm29	0.02	0.02	0.7	3	6	0.6

4.1.4.1. MN Clássico x FBMN – Dados do Modo de Ionização Positivo

A otimização da rede construída a partir de dados de LC-MS no modo positivo foi realizada para dois grupos de dados: dados de Auto-MS/MS (LC-PR)) e dados pré-processados (LC-PP), separadamente. As tabelas 7 e 8 mostram os valores das análises de otimização dos dois grupos.

Observando essas tabelas, é nítida a diferença entre os grupos em número de nodos, de pares (componentes conectados) e de nodos anotações, havendo uma compatibilidade muito maior com a biblioteca da plataforma GNPS usando os dados pré-processados a partir do *Mzmine*. Esse resultado mostra que utilizar os dados obtidos a partir do processamento para este modo de ionização trouxe um aprimoramento da análise. Portanto, foram selecionados 6 modelos do grupo LC-PP que obtiveram o maior número de nodos anotados e proporcionam uma visualização da distribuição de cossenos de similaridade de acordo com as mudanças de parâmetros: LC-PP – otm01, LC-PP – otm02, LC-PP – otm03, LC-PP – otm04 LC-PP – otm22 e LC-PP – otm26.

Tabela 7. Características resultantes dos modelos na otimização da construção das Redes Moleculares para os dados brutos (LC-PR) das análises por LC-MS.

Modelos	Nº de Nodos	Nº de Pares	Nº de Nodos Anotados
otm01	7671	5349	568
otm02	7746	5423	562
otm03	7931	5466	539
otm04	8270	5489	465
otm05	9614	5582	473
otm06	10155	5508	506
otm07	10332	5362	492
otm08	8270	5439	465
otm09	8270	5177	465
otm10	8270	4930	465
otm11	8270	5713	465
otm12	8270	5654	465
otm13	8270	5351	465
otm14	8270	5207	465
otm15	8270	6032	465
otm16	8270	5958	465
otm17	8270	5696	465
otm18	8270	5476	465
otm19	8270	5489	427
otm20	8270	5489	359
otm21	8270	5489	285
otm22	8270	5489	559
otm23	8270	5489	514
otm24	8270	5489	430
otm25	8270	5489	349
otm26	8270	5489	666
otm27	8270	5489	599
otm28	8270	5489	495
otm29	8270	5489	408

Tabela 8. Características resultantes dos modelos na otimização da construção das Redes Moleculares para os dados pré-tratados (LC-PP) das análises por LC-MS.

Modelos	Nº de Íons Precusores	Nº de Nodos	Nº de Arestas	Nº de Componentes de Conectados	Nº de Nodos Anotados
otm01	24518	10617	11870	1826	1411
otm02	24518	10617	11870	1826	1360
otm03	24518	10617	11870	1826	1210
otm04	24518	10617	11870	1826	921
otm05	24518	10617	11951	2251	1024
otm06	24518	10617	11963	2369	1090
otm07	24518	10617	11937	2440	1124
otm08	24518	10617	11864	1808	921
otm09	24518	10617	11853	1797	921
otm10	24518	10617	11842	1789	921
otm11	24518	10617	12022	2194	921
otm12	24518	10617	12098	2201	921
otm13	24518	10617	12080	2191	921
otm14	24518	10617	12065	2179	921
otm15	24518	10617	12116	2461	921
otm16	24518	10617	12153	2451	921
otm17	24518	10617	12142	2442	921
otm18	24518	10617	12130	2430	921
otm19	24518	10617	11870	1826	791
otm20	24518	10617	11870	1826	682
otm21	24518	10617	11870	1826	610
otm22	24518	10617	11870	1826	1329
otm23	24518	10617	11870	1826	1142
otm24	24518	10617	11870	1826	981
otm25	24518	10617	11870	1826	883
otm26	24518	10617	11870	1826	1815
otm27	24518	10617	11870	1826	1581
otm28	24518	10617	11870	1826	1366
otm29	24518	10617	11870	1826	1224

A distribuição dos valores de cosseno de similaridade informa a faixa de distribuição dos valores de cosseno das anotações realizadas pela biblioteca de dados, o que nos permite visualizar se os modelos possuem maior ou menor compatibilidade espectral com o banco de dados espectrais do GNPS. É de grande interesse obter modelos que trazem grande número de moléculas anotadas com altos valores de cosseno.

A avaliação da distribuição dos valores de cosseno de similaridade também foi realizada para o grupo LC-PR, para conferir a qualidade das anotações realizadas, já

que a qualidade das anotações independe do número de anotações realizadas na rede. Foram selecionados os mesmos modelos de rede para este grupo: LC-PR – otm01, LC-PR – otm02, LC-PR – otm03, LC-PR – otm04 LC-PR – otm22 e LC-PR – otm26.

O modelo LCPP - otm26 possui muitos íons anotados, mas possui uma distribuição de cossenos em faixas onde há maior risco de falsas anotações, assim como otm22 (Figura 5.A). Assim, os dois modelos de maior número de íons anotados e de valores de cossenos mais elevados são LCPP - otm01 e LCPP - otm02. Esses dois modelos resultaram nos mesmos valores de número de precursores, nodos, arestas e componentes conectado, havendo somente diferença nos números de nodos anotados e na distribuição de cossenos. Essa diferença se dá pela maior tolerância de massa entre íons precursores do modelo LCPP - otm01, sendo compatível com mais compostos do banco de dados espectrais. Essa tendência também ocorre nos modelos do grupo LC-PR, já que possuem os mesmos parâmetros. A avaliação da distribuição relacionada aos modelos de LC-PR trouxe um fato curioso. Mesmo com um menor número de anotações, esse grupo apresentou um maior número de anotações nas faixas de cosseno maiores que 0,91, sendo que só esse grupo possui valores de cosseno na faixa de 0,96 a 0,98 (Figura 5.A e B).

Assim, foram selecionados os dois modelos, LCPR - otm02 e LCPP - otm02, visando aumentar a possibilidade de uma anotação correta. Os dois modelos foram selecionados para serem utilizados nos estudos metabolômicos para desreplicação e somente o modelo LC-PP – otm02 foi selecionado para as análises metabolômicas de correlação via rede molecular e análises estatísticas multivariadas supervisionadas e não-supervisionadas, descritas no capítulo 2 e 3 desta tese.

A Figura 5 mostra que LC-PP – otm02 possui grande número de anotações com valores de cosseno entre 0,81 e 0,9, os quais são interessantes e não devem ser descartados, além das 938 anotações com valores de cosseno entre 0.7 e 0.8, valor maior do que o total de anotações de LC-PR - otm02. Mesmo que possua poucas anotações entre 0,91 e 0,98, é um modelo que vale a pena ser explorado e ter suas anotações conferidas. Através dos dois modelos, é realizada a conferência das anotações, comparando manualmente os perfis de fragmentação dos dados com os da biblioteca, e montada uma biblioteca de anotações da família Orchidaceae.

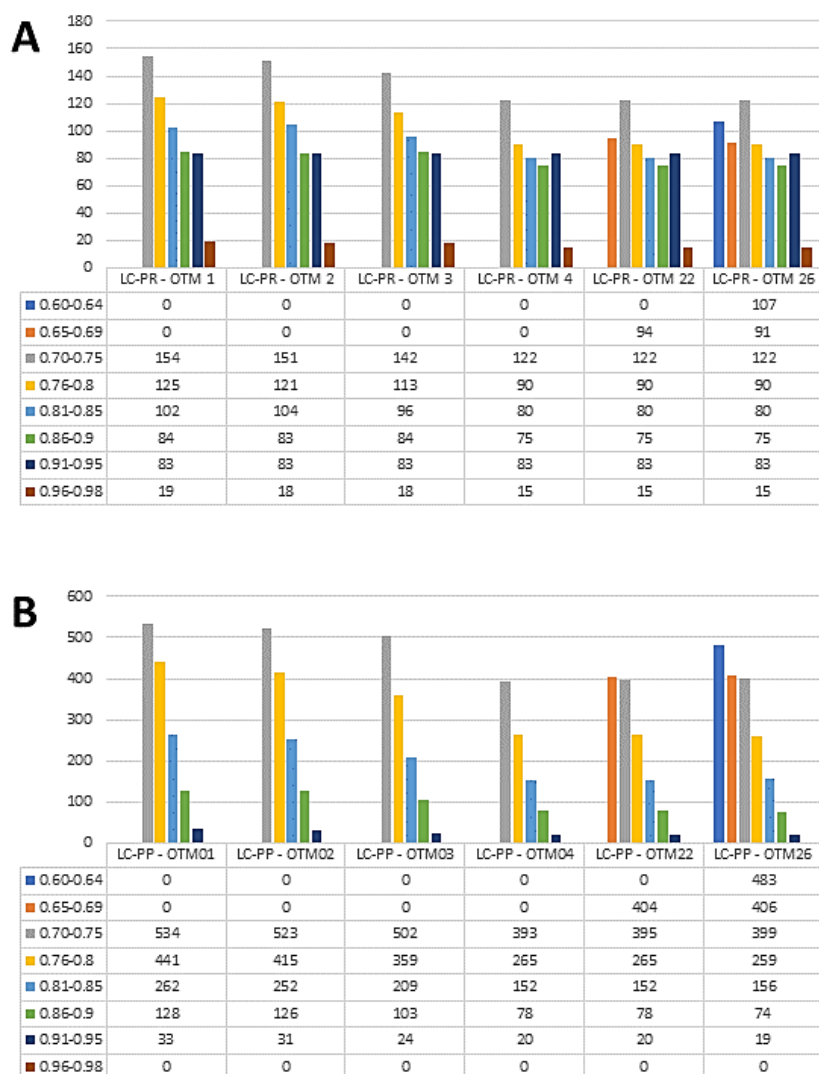


Figura 5. Gráficos de Distribuição de cosseno de similaridade com o banco de dados espectrais do GNPS e comparação entre as redes moleculares geradas através de MN clássica e FBMN. **A.** Redes moleculares geradas através de FBMN (LC-PR) de acordo com os modelos de otimização produzidos. **B.** Redes moleculares geradas através de MN clássica (LC-PP) de acordo com os modelos de otimização produzidos.

4.1.4.2. MN Clássico x FBMN – Dados do Modo de Ionização Negativo

A otimização da rede molecular construída a partir de dados de LC-MS no modo negativo foi realizada com os dados de Auto-MS/MS (LC-NR) e dados pré-processados (LC-NP) advindos de dois modelos otimizados pelo Mzmine: NEG-1 e NEG-5, separadamente. Esses dois modelos possuem o mesmo número de *hits*, números de íons precursor e íons selecionados para a fragmentação muito próximos. Com a análise da distribuição dos cossenos temos mais certeza de qual é o modelo

que contribui para maior número de hits com maiores valores de cosseno. As tabelas 9, 10 e 11 mostram os valores das análises de otimização dos três grupos.

Tabela 9. Características resultantes dos modelos na otimização da construção das Redes Moleculares para os dados brutos (LC-NR) das análises por LC-MS.

Modelos	Nº de Nodos	Nº de Pares	Nº de Nodos Anotados
otm01	2826	1425	156
otm02	2853	1426	152
otm03	2892	1448	151
otm04	3009	1507	83
otm05	3729	1408	87
otm06	3992	1465	93
otm07	4066	1332	89
otm08	3009	1653	83
otm09	3009	1517	83
otm10	3009	1476	83
otm11	3009	1568	83
otm12	3009	1667	83
otm13	3009	1488	83
otm14	3009	1491	83
otm15	3009	1533	83
otm16	3009	1683	83
otm17	3009	1529	83
otm18	3009	1510	83
otm19	3009	1507	71
otm20	3009	1507	55
otm21	3009	1507	30
otm22	3009	1507	99
otm23	3009	1507	84
otm24	3009	1507	64
otm25	3009	1507	38
otm26	3009	1507	114
otm27	3009	1507	94
otm28	3009	1507	75
otm29	3009	1507	45

Tabela 10. Características resultantes dos modelos na otimização da construção das Redes Moleculares para os dados pré-tratados NEG-1 (LC-NP) das análises por LC-MS.

Modelos	Nº de Íons Precusores	Nº de Nodos	Nº de Arestas	Nº de Componentes Conectados	Nº de Nodos Anotados
otm01	17819	2648	2660	143	41
otm02	17819	2648	2660	143	39
otm03	17819	2648	2660	143	33
otm04	17819	2648	2660	143	15
otm05	17819	2648	2689	161	15
otm06	17819	2648	2702	171	18
otm07	17819	2648	2711	176	18
otm08	17819	2648	2660	143	15
otm09	17819	2648	2660	143	15
otm10	17819	2648	2660	143	15
otm11	17819	2648	2701	163	15
otm12	17819	2648	2701	163	15
otm13	17819	2648	2701	163	15
otm14	17819	2648	2701	163	15
otm15	17819	2648	2719	196	15
otm16	17819	2648	2719	196	15
otm17	17819	2648	2719	196	15
otm18	17819	2648	2719	196	15
otm19	17819	2648	2660	143	12
otm20	17819	2648	2660	143	11
otm21	17819	2648	2660	143	10
otm22	17819	2648	2660	143	31
otm23	17819	2648	2660	143	26
otm24	17819	2648	2660	143	23
otm25	17819	2648	2660	143	20
otm26	17819	2648	2660	143	49
otm27	17819	2648	2660	143	39
otm28	17819	2648	2660	143	32
otm29	17819	2648	2660	143	29

Tabela 11. Características resultantes dos modelos na otimização da construção das Redes Moleculares para os dados pré-tratados NEG-5 (LC-NP) das análises por LC-MS.

Modelos	Nº de Íons Precusores	Nº de Nodos	Nº de Arestas	Nº de Componentes Conectados	Nº de Nodos Anotados
otm01	17070	2473	2502	119	40
otm02	17070	2473	2502	119	39
otm03	17070	2473	2502	119	34
otm04	17070	2473	2502	119	15
otm05	17070	2473	2523	139	15
otm06	17070	2473	2536	151	17
otm07	17070	2473	2540	157	17
otm08	17070	2473	2502	119	15
otm09	17070	2473	2502	119	15
otm10	17070	2473	2502	119	15
otm11	17070	2473	2536	144	15
otm12	17070	2473	2536	144	15
otm13	17070	2473	2536	144	15
otm14	17070	2473	2536	144	15
otm15	17070	2473	2553	175	15
otm16	17070	2473	2553	175	15
otm17	17070	2473	2553	175	15
otm18	17070	2473	2553	175	15
otm19	17070	2473	2502	119	12
otm20	17070	2473	2502	119	11
otm21	17070	2473	2502	119	10
otm22	17070	2473	2502	119	32
otm23	17070	2473	2502	119	27
otm24	17070	2473	2502	119	24
otm25	17070	2473	2502	119	20
otm26	17070	2473	2502	119	49
otm27	17070	2473	2502	119	41
otm28	17070	2473	2502	119	34
otm29	17070	2473	2502	119	30

Observando as tabelas dos dados processados e não processados é nítida a diferença entre eles em número de anotações, havendo uma compatibilidade muito maior com a biblioteca da plataforma GNPS usando os dados não processados.

No caso do das análises com o modo negativo, não podemos fazer a mesma afirmação que foi feita para os dados no positivo, em que houve um aprimoramento da análise. Relembrando que a grande maioria dos espectros depositados no banco de dados espectral do GNPS são de análises no modo de ionização positivo, quando utilizamos o MN clássico é possível que haja muitas compatibilidades falsas com a biblioteca. Ao ponto que utilizamos os dados advindos do Mzmine, temos maior confiança de que há uma menor quantidade de falsos *hits*.

Com o intuito de confirmar se os hits possuem certa qualidade em compatibilidade, foram construídos os gráficos de distribuição do cosseno para os três modelos: dados brutos, processados NEG-1, processados NEG-5.

Os modelos de LC-NR foram os que obtiveram maior número de nodos anotados (Tabela 9, Gráfico 3) e foram selecionados para a análise da distribuição dos valores de cosseno.

A distribuição de cosseno possui poucas diferenças entre os modelos otm01, otm02 e otm03 do grupo LC-NR. O modelo otm02 possui uma maior distribuição dos cossenos na faixa de 0,91 a 0,98 que os outros, possui uma tolerância de massa aceitável de acordo com o equipamento utilizado nas análises, e diminuir a possibilidade de falsas anotações.

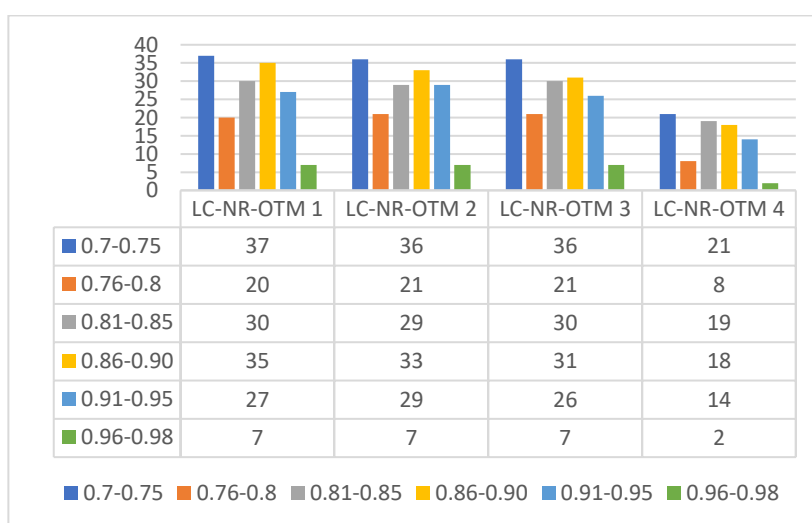


Gráfico 3 Distribuição dos valores de cosseno dos nodos anotados dentro de cada modelo selecionado da análise de LC-NR.

Os modelos de LC-NP obtiveram menor número de nodos anotados (Tabela 11, Gráfico 4 e 5), mas foram selecionados para a análise da distribuição dos valores de cosseno para observarmos se seus valores de cosseno correspondem à um bom modelo ou melhor que o de LC-NR.

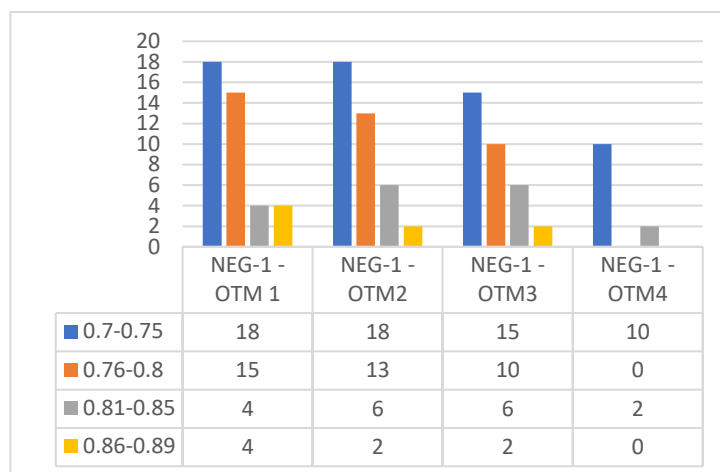


Gráfico 4. Distribuição dos valores de cosseno dos nodos anotados dentro de cada modelo selecionado da análise de LC-NP/NEG-1.

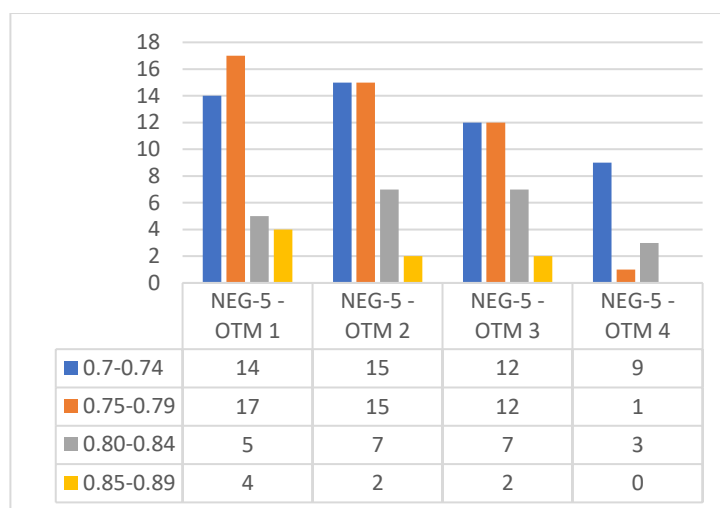


Gráfico 5. Distribuição dos valores de cosseno dos nodos anotados dentro de cada modelo selecionado da análise de LC-NP/NEG-5.

Comparando os gráficos 3,4 e 5, é nítido que as anotações do modelo LC-NR estão mais distribuídas para os cossenos mais próximos de 1,0. Ter essa distribuição de anotados em cossenos mais elevados e a maior quantidade de íons anotados faz este modelo muito interessante para ser utilizado para as próximas etapas.

O modelo otm 02 de LC-NP/NEG-5 possui uma maior distribuição de cosseno entre 0,75 e 0,89, comparado com seus outros modelos e com o grupo LC-NP/NEG-1. Mesmo assim, a distribuição de cosseno do grupo LC-NR é muito mais interessante para as anotações. Sendo assim, o modelo selecionado para ser utilizado nas próximas etapas deste trabalho foi o modelo otm02 do grupo LC-NR.

4.2. Análise do Metaboloma - Compostos de média-baixa polaridade

4.2.1. Processamento das Impressões Digitais Metabólicas - GC-MS

As impressões digitais metabólicas adquiridas por GC-MS (Figura 6) foram processadas através da ferramenta “*GC-MS EI Data Analysis – Data Processing – Deconvolution*”, disponível na plataforma GNPS. Essa ferramenta possui uma abordagem de aprendizado de máquina (*Machine learning*) algorítmica que promove a auto-deconvolução e alinhamento espectral de dados de GC-MS, chamada *MShub*. Essa parceria entre *MShub* e GNPS promove a auto-deconvolução de padrões de fragmentação dos compostos detectados via matrix de fatorização não-negativa não-supervisionada e quantifica a reprodutibilidade dos padrões de fragmentação através das amostras (Aksenov et.al, 2021).

De acordo com Aksenov et.al. (2021), o *MShub* possui um desempenho tão bom quanto, ou melhor, que outras ferramentas de deconvolução. Devido a relação linear entre a eficiência e número de arquivos utilizados para a deconvolução, é a única ferramenta passível de operação em escala de repositório em sua forma atual.

Após a deconvolução, são disponibilizados para download os dados espectrais no formato .mgf, a tabela com as informações sobre os picos deconvoluídos, a tabela com os dados integrais do *MShub* e os dados espectrais gerados pelo *MShub*. Também possibilita continuar o processamento dos dados através da construção da rede molecular e Anotações pela busca na biblioteca espectral do GNPS.

O processamento das impressões digitais metabólicas das espécies de Orchidaceae através da ferramenta citada acima resultou em uma tabela de dados contendo 1633 íons precursores deconvoluídos.

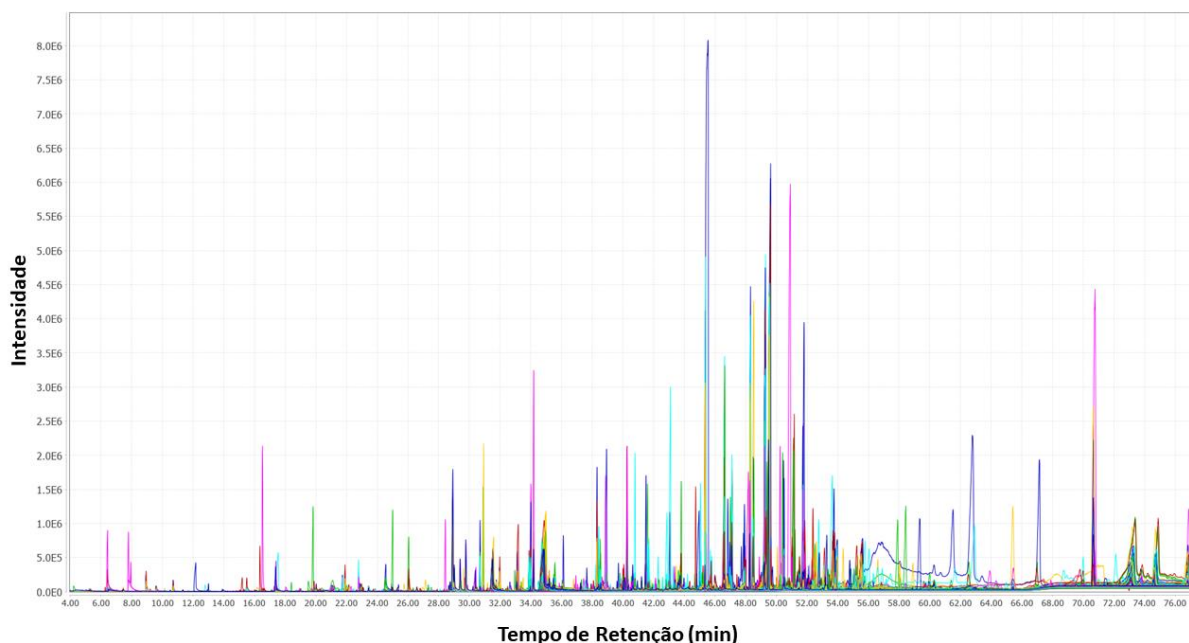


Figura 6. Impressões digitais metabólicas das 131 espécies de Orchidaceae adquiridas por – GC-EI-MS QP2010 (Shimadzu). Software de visualização utilizado: *MzMineTM* (v. 2.51, BMC Bioinformatics, Reino Unido).

4.2.2. Construção das Redes Moleculares

A construção de redes moleculares a partir de dados obtidos por de GC-MS foi somente realizada através da utilização de dados deconvoluídos, diferentemente dos dados de LC-MS que podem ser construídas redes com os dados brutos.

Após o processamento de dados realizado na plataforma GNPS através do MSHub, os arquivos .mgf, contendo os espectros de massas por impacto eletrônico deconvoluídos, e .cdf, contendo a tabela de quantificação de recursos, foram utilizados na otimização da construção das redes moleculares.

A otimização da construção da rede molecular permitiu obter uma rede que melhor represente o metaboloma estudado e que, através da busca espectral na biblioteca GNPS, possua grande número de anotações com valores de cosseno de similaridade altos. Assim como nas análises por LC-MS, esta otimização possibilitou observar a influência dos parâmetros na construção da rede e na anotação dos compostos. Foram construídos 15 modelos de otimização, cujos parâmetros e os valores utilizados estão na tabela12.

Tabela 12. Parâmetros e valores utilizados na otimização da construção das Redes Moleculares a partir de dados de GC-MS.

Modelos	Search Options		Advanced Search Options	
	Fragment Ion Mass Tolerance	Minimum Matched Peaks	Score Threshold	Top Hits Per Spectrum
otm01	1	6	0.85	1
otm02	0.8	6	0.85	1
otm03	0.6	6	0.85	1
otm04	0.4	6	0.85	1
otm05	0.02	9	0.85	1
otm06	1	12	0.85	1
otm07	1	15	0.85	1
otm08	1	6	0.8	1
otm09	1	6	0.75	1
otm10	1	6	0.7	1
otm01 – TOP2	1	6	0.85	2
otm01 – TOP4	1	6	0.85	4
otm01 – TOP6	1	6	0.85	6
otm01 – TOP8	1	6	0.85	8
otm01 – TOP10	1	6	0.85	10
otm08 – TOP10	1	6	0.8	10

A comparação entre os modelos foi realizada observando o número total de correspondências com a biblioteca espectral, o número anotações de acordo com o número de correspondências por espectro (*Top Hits Per Spectrum*) e a distribuição dos valores de cosseno (*Score Threshold*) de similaridade com a biblioteca. A tabela 13 apresenta os valores obtidos nas análises de otimização.

Observando a tabela 13 é evidente que ao haver uma diminuição na tolerância de massa dos íons fragmento (otm01 a otm05) e um aumento no número mínimo de íons correspondentes, há uma diminuição no número de anotações totais e principais. Ao passo que ao diminuir os valores de cosseno de similaridade, aumentou o número de anotações (otm08 a otm10). Além disso, ao aumentar o número de correspondências por espectro (*Top Hits Per Spectrum*), houve um aumento significativo comparado com otm01 (padrão indicado pelos desenvolvedores).

Para confirmar se os altos valores de anotações observados têm significância, foram realizadas as distribuições de cosseno, comparando os modelos otm01, otm08, otm09, otm10, otm08-TOP10 (Gráfico 6).

Tabela 13. Características dos modelos da otimização da construção das Redes Moleculares para os dados deconvoluídos das análises por GC-MS.

Modelos	Anotações Totais	Anotações Principais (Top Hits)	% Anotações Totais	% Anotações Principais	Razão Anotações principais/totais
otm01	1087	400	66,56	24,49	36,80
otm02	633	226	38,76	13,84	35,70
otm03	628	223	38,46	13,66	35,51
otm04	626	223	38,33	13,66	35,62
otm05	1016	378	62,22	23,15	37,20
otm06	948	353	58,05	21,62	37,24
otm07	904	332	55,36	20,33	36,73
otm08	1881	637	115,19	39,01	33,86
otm09	2740	836	167,79	51,19	30,51
otm10	3541	998	216,84	61,11	28,18
otm01 – TOP2	1944	461	66,56	28,23	23,71
otm01 – TOP4	3293	483	119,04	29,58	14,67
otm01 – TOP6	4324	491	201,65	30,07	11,36
otm01 – TOP8	5208	492	264,79	30,13	9,45
otm01 – TOP10	5974	495	318,92	30,31	8,29
otm08 – TOP10	12672	796	365,83	48,745	6,28

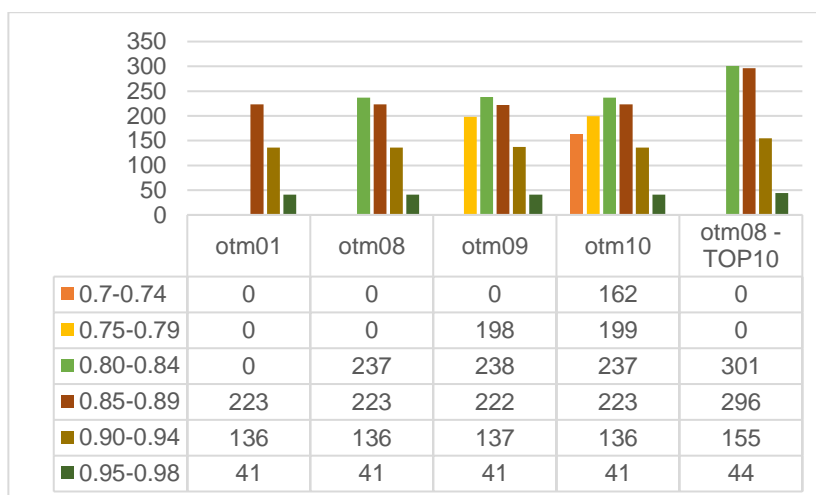


Gráfico 6. Distribuição dos valores de cosseno de acordo com o número de nodos anotados nas redes moleculares construídas a partir dos dados de GC-MS.

Observando os modelos otm01, otm08, otm09 e otm10 percebe-se que a diminuição dos valores de cosseno de similaridade com a biblioteca não auxilia a melhorar os modelos em relação à distribuição dos valores de cosseno. Essa permaneceu a mesma com a alteração de parâmetro, havendo somente um crescente de número de anotações em faixas inferiores de cosseno. Dentre todos, o que possuiu

um aumento na distribuição dos valores de cosseno em faixas superiores foi o modelo otm08-TOP10.

Ao aumentar o número de correspondências por espectro, foi possível obter anotações de alto valor de cosseno, importantes para auxiliar na desreplcação do extrato. Com isso, o modelo selecionado para as análises futuras é o otm08-TOP10.

4.3. Importância do Processamento de Dados na Construção da Rede Molecular

A inovação trazida por Nothias e colaboradores (2019) ao aproveitar os dados processados produzidos por softwares bem-estabelecidos de processamento de dados complementando o MN clássico trouxe uma melhora na estrutura da rede molecular e na qualidade e número de anotações. Abaixo está uma figura retirada de Nothias *et.al.*(2019) que detalha os métodos de geração das redes moleculares a partir de MN Clássico e FBMN (a) e mostra o número de estudos realizados no GNPS entre MN clássicos e FBMN desde 2016 e o número de FBMN produzidos desde seu início e eventos que auxiliaram e facilitaram seu uso.

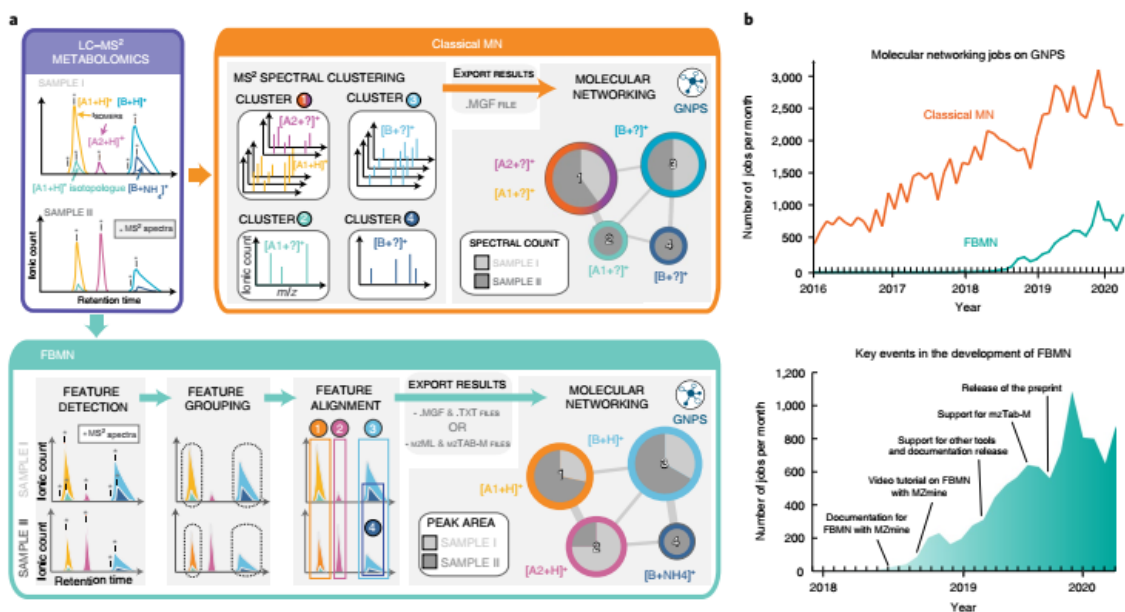


Figura 7. Métodos de geração de redes moleculares da plataforma online GNPS a partir de dados de massa não direcionados. **(a)** Após a aquisição dos dados de LC-MS/MS (MS¹ e MS²), dois métodos podem ser utilizados para gerar as redes moleculares: MN clássico e FBMN. **(b)** Gráficos demonstrando o número de análises realizadas. O gráfico em cima mostra o número de MN clássicos e FBMN realizados desde 2016 e o em baixo mostra o número de FBMN realizados desde sua criação e as variações de acordo com as melhoras de processos. Figura retirada de Nothias *et.al.* (2019).

Os dois modelos apresentados na figura 7, MN clássico e FBMN, utilizam o agrupamento (*cluster*) de espectros para formação de um espectro representativo somente, e utilizá-lo nas buscas nas bases de dados. O MN clássico utiliza o algoritmo MS-Cluster, o qual gera um espectro consenso baseando-se na similaridade entre espectros do mesmo íon precursor descrita pelo cosseno de similaridade (*threshold* τ). O algoritmo consiste em várias rodadas de agrupamentos com valores decrescentes de cosseno de similaridade, sendo que os agrupamentos mais semelhantes são mesclados primeiro. (Figura 7).

O FBMN utiliza dos dados deconvoluídos e alinhados produzidos pelo software Mzmine, o qual produz um espectro representativo para o íon precursor referente a determinado tempo de retenção e razão massa/carga, de acordo com a tolerância de massa definida para íons precursores e fragmento.

A otimização da construção da rede molecular utilizando os dados de LC-MS confirmou os benefícios da utilização da FBMN no aumento dos números anotações em cada modelo, principalmente nas redes que utilizaram os dados do modo de ionização positivo. Comparando-os na tabela 7 e 8, é nítida a diferença em número de nodos, número de pares (ou componentes conectados) e número de nodos anotados. O modelo LC-PP otm02 possui aproximadamente 2,4 vezes mais anotações utilizando FBMN que a MN clássico.

Avaliando a distribuição de cosseno de similaridade entre os íons anotados utilizando o fluxo e trabalho por FBMN (LC-PP – otm02) e MN clássico (LC-PR – otm02), foi observada uma distribuição de valores de cosseno que não era a esperada: distribuição em valores mais elevados no modelo do MN Clássico do que no de FBMN. Através disso nota-se que avaliar uma rede molecular somente pelo número de anotações produzidas pode resultar em falsas anotações ou até anotações duplicadas.

Um dos aspectos das redes que pode diminuir a qualidade das anotações e da rede são a produção de agrupamentos fragmentados, ou seja, dois espectros consenso (agrupamento) referentes à um íon precursor. Esse ocorre no MN clássico quando há em grandes agrupamentos de espectro, devido às diferenças de intensidade e número de íons fragmento entre espectros (Frank *et.al.*, 2008). Além disso, pode ocorrer quando há o ajuste inadequado da tolerância de massa de íons precursor ou de íons fragmento, além da coeluição de íons isobáricos e fragmentados

juntos, fazendo com que o MS-Cluster não os classifique como o mesmo íon (Nothias et.al., 2019). Um dos grupos fragmentados pode possuir espectros de baixa qualidade e, conseqüentemente, falsas anotações ou baixos valores de cosseno de similaridade com a base de dados. O FBMN permite diminuir o número de nodos duplicados que podem ocorrer no MN clássico, devido à deconvolução e alinhamento prévio dos dados, gerando um nodo como um espectro de fragmentação representativo.

No caso deste estudo, a separação dos isômeros no FBMN, e/ou a baixa intensidade de íons fragmento ou diferente número de íons fragmento, pôde ter gerado espectros representativos de menor cosseno de similaridade com a biblioteca, havendo anotações com valores baixos de cosseno de similaridade. Isso pode ser observado no aumento dos números de anotados com valores de cosseno de similaridade na faixa de 0,7-0,8 nas análises por FBMN (Figura 5, gráfico 3-5). Ao separar os isômeros, pode haver a formação de agrupamentos com consenso de baixa razão sinal/ruído, com íons fragmento de intensidades e quantidades diferentes, havendo uma menor compatibilidade com a base de dados e, conseqüentemente, anotações com menores valores de cosseno.

Diferente das análises por MN clássico, que não fazem essa distinção entre isômeros, fazendo com que espectros de isômeros, de baixa e alta qualidade, participassem do mesmo agrupamento. Esse tipo de análise possui mais espectros disponíveis para a formação do consenso, e dentre esses pode haver espectros de alta qualidade trazendo maior peso no agrupamento, construindo um consenso com maior compatibilidade com a biblioteca. Frank *et.al.* (2008) afirma em seu trabalho que em alguns casos o espectro consenso melhora a razão sinal/ruído comparando com o melhor espectro individual do agrupamento, o que leva à novas identificações, isso devido à consenso com maior número de íons fragmento e espectro com sinal de maior intensidade como um todo.

Os dois modelos selecionados na otimização (LCPR - otm02 e LCPP - otm02) serão utilizados nos estudos metabolômicos para desrepliação, sendo realizada a conferência de cada anotação observando os perfis de fragmentação como um todo, para confirmar a dada anotação. Assim, utilizamos o melhor dos dois modelos na desrepliação e construção da biblioteca de anotações da família Orchidaceae.

4.4. Importância da Distribuição de Cossenos de Similaridade na Construção da Rede molecular

Após a construção de uma rede molecular, na busca nas bibliotecas de espectros e para as análises de correlação de acordo com a hipótese estudada, sempre deve ser realizada a confirmação das anotações afim de eliminar possíveis anotações falsas.

Como uma estratégia para diminuir os casos de falsas anotações e retirar a influência dessas nas análises de correlação, uma triagem preliminar é fundamental, e deve ser realizada preferencialmente de maneira rápida e em grande escala. A avaliação do número de anotações combinada com a avaliação da distribuição dos valores de cosseno de similaridade é uma estratégia rápida e de fácil visualização.

A triagem preliminar realizada na otimização da construção das redes moleculares clássica e por FBMN permitiu observar que os parâmetros dentro de *Basic Options* e *Advanced Library Search Options* influenciam na quantidade de nodos anotados através da compatibilidade entre íons precursores e entre espectros de fragmentação, relacionada ao cosseno de similaridade. Foi observado também que os parâmetros de *Advanced Network Options* influenciam somente na formação e configuração dos clusters (agrupamentos de análogos ou famílias moleculares), e conexão entre análogos.

De acordo com as tabelas 7 a 11, os parâmetros que geraram um aumento no número de anotações com a variação de seus valores foi a tolerância de massa de íon precursor (*Precursor Ion Mass Tolerance*) e o valor mínimo de cosseno de similaridade (*Score Threshold*), mas eles possuem diferentes influências na distribuição dos valores de cosseno dos íons anotados.

O aumento da tolerância de massa de Íons precursores proporcionou o aumento do número de anotações, e essas foram distribuídas entre todas as faixas de cosseno (Gráficos 2 a 5), havendo uma diferenciação pouco significativa dentre os modelos de otimização. Neste caso, é interessante que seja levado em consideração para a seleção do parâmetro o equipamento utilizado nas análises (alta ou baixa resolução, indicadas pelo desenvolvedor no fluxo de trabalho), e a intensão de produzir o mínimo de falsas anotações possível.

A diminuição do valor mínimo de cosseno de similaridade (*Score Threshold*) proporcionou um grande aumento das anotações, mas uma pobre distribuição de

valores de cosseno. O Gráfico 2 mostra a diferença entre os modelos otm04, otm22 e otm26, em não houve distribuição das anotações em todas as faixas de cosseno acima de 0.7, mas sim foram adicionadas às faixas menores conforme diminuía o valor mínimo de cosseno. Esses modelos trazem um maior risco de inserir anotações falsas na rede molecular pela baixa similaridade espectral.

A triagem preliminar realizada na otimização da construção das redes moleculares utilizando dados de GC-MS permitiu observar que os parâmetros dentro de *Search Options* e *Advanced Search Options* foram os que mais influenciam no aumento da quantidade de nodos anotados.

Observando os parâmetros de valores de tolerância de massa de íons fragmento (*Fragment Ion Mass Tolerance*) e o número mínimo de íons de fragmento em comum (*Minimum Matched Peaks*), não foi observada uma diferença significativa no número de anotações ao variar os valores dos parâmetros. Já a variação do valor mínimo de cosseno de similaridade (*Score Threshold*) proporcionou um grande aumento nas anotações totais e principais (*Top Hits*) (Tabela 11).

A avaliação da distribuição de valores de cosseno de acordo com o valor mínimo de cosseno de similaridade indicou a mesma tendência das rede moleculares clássicas e FBMN geradas com dados de LC-MS: ao diminuir os valores mínimos de cosseno, as anotações geradas foram adicionadas às faixas inferiores de cosseno de similaridade, distribuição observada no Gráfico 6 comparando os modelos otm01, otm08, otm09 e otm10.

Foi avaliada também a variação dos valores do parâmetro *Top Hits Per Spectrum*, relacionado às principais correspondências com a biblioteca espectral, de acordo com a pontuação de cosseno. É recomendado utilizar valores altos para aumentar a possibilidade de uma anotação correta, caso não esteja indicada na primeira ocorrência. O aumento no número de anotações totais e principais é nítido, conforme aumentamos o número de correspondências. Comparando os modelos otm08 e otm08-TOP10, foi observado que as anotações geradas foram distribuídas em todas as faixas de cosseno, havendo até maior número de anotações em faixas superiores. Através disso, utilizar altos valores desse parâmetro pode proporcionar anotações com menor probabilidade de falsos positivos.

Através da triagem baseada na avaliação da distribuição dos cossenos de similaridade entre espectros de fragmentação dos íons precursores e a biblioteca de

espectros do GNPS, foi possível observar que diminuir o valor mínimo de cossenos de similaridade (*Score Threshold*) pode aumentar o número de anotações, mas essas terão baixa similaridade espectral com a biblioteca. Já o aumento da tolerância de massa de íons precursores proporciona um aumento das anotações e essas são distribuídas por todas as faixas de cosseno, sendo muito mais interessante para análises com grande número de dados. Além disso, nas redes moleculares utilizando dados de GC-MS, é de grande valia utilizar altos valores no parâmetro *Top Hits Per Spectrum*, pois além de aumentar a possibilidade de uma anotação correta e aumentar o número de anotações, essas são distribuídas em todas as baixas de cosseno.

5. CONCLUSÕES

As atividades de pesquisa descritas neste capítulo demonstraram os passos utilizados, e necessários, para se construir a base do estudo metabolômico das plantas brasileiras de Orchidaceae estudadas, relacionados não somente obtenção das impressões digitais, matrizes de dados e redes moleculares, mas também à importância do processamento de dados na construção das redes moleculares e da avaliação da qualidade das redes geradas.

A utilização de dados processados pelo software Mzmine na construção da rede molecular proporcionou uma grande evolução da rede, havendo um aumento no número de íons precursores, anotações e nodos conectados comparados com as redes clássicas. Além disso, a avaliação criteriosa dos parâmetros utilizados nos fluxos de trabalho na construção das redes moleculares, principalmente avaliando o número de anotações e a distribuição do cosseno de similaridade, permitiu selecionar redes moleculares com espectros consenso estruturados e de alta qualidade, elevando o nível de compatibilidade com a base de dados da plataforma GNPS.

As redes construídas através da abordagem clássica demonstraram ser interessantes ao estudo, devido a sua contribuição através das anotações de alto valor de cosseno de similaridade. Com isso, observou-se a grande contribuição do processamento de dados e da construção criteriosa das redes moleculares para a construção do alicerce dos estudos metabolômicos.

Portanto, este estudo permitiu a obtenção de impressões digitais metabólicas de alta qualidade, estando essas relacionadas aos compostos de média-baixa e

média-alta polaridades, e a construção dados robustos para, assim, realizar os estudos metabolômicos sobre a química de espécies brasileiras da família Orchidaceae e o estudo dos metabólitos das orquídeas brasileiras como fonte de metabólitos moduladores da produção de espécies reagentes de oxigênio (ERO) por neutrófilos humanos e antioxidantes.

CAPÍTULO 2

1. INTRODUÇÃO

A família Orquidaceae é um dos maiores grupos florísticos, detentora de grande diversidade ecológica e um dos grupos de plantas florísticas mais difundidos pelo mundo, compreendendo espécies terrestres, saprófitas e epífitas. O seu uso medicinal é descrito por eras dentro da Medicina Tradicional Chinesa, incluindo lendas chinesas do uso de espécies de *Dendrobium* em sua Materia Medica, datadas do século 28 a.C (CHANDRA DE et al., 2014; SUT; MAGGI; DALL'ACQUA, 2017). Além disso, é também utilizada na Medicina Ayurveda (indiana) e na medicina popular do sudoeste asiático, Japão, Europa, África e Américas (HOSSAIN, 2011).

Devido ao uso medicinal, várias classes de metabólitos já foram isoladas de orquídeas utilizadas terapêuticamente, como derivados fenólicos, alcaloides, terpenos, estilbenos e diidroestilbenóides derivados, fenantropiranos, fenantrenos e diidrofenantrenos e diidrofenantrenos, dentre outros, mostrando a grande diversidade química da família (SUT; MAGGI; DALL'ACQUA, 2017).

A diversidade de classes e compostos observadas nos estudos químicos de espécies de Orchidaceae, além do potencial demonstrado em estudos etnofarmacológicos e de atividade biológica, mostram uma grande oportunidade de encontrar novos metabólitos nesta família tão vasta e fascinante, principalmente metabólitos inéditos para as espécies brasileiras.

A Metabolômica de Plantas (*Plant Metabolomics*) é uma abordagem não-alvo rápida, eficaz e otimizada na investigação, qualitativa e quantitativa, de metabólitos presentes em tecidos e órgãos vegetais ou na planta como um todo, afim de identificar e/ou quantificar o maior número de metabólitos possível, além de permitir a correlação dos metabólitos com propriedades atribuídas às amostras (AL., 2020).

A identificação de metabólitos secundários em estudos de produtos naturais é realizada a partir do isolamento do composto advindos de fontes biológicas e elucidação estrutural tridimensional, incluindo, se necessário, a definição de centros estereogênicos. Esse processo é manual, laborioso, complexo e demorado, que demanda tempo e experiência do espectroscopista, para ser utilizado em estudos que envolvem conjuntos de dados massivos (*Big data*) e análise de grande quantidade de amostras. A análise metabolômica utiliza da desreplicação como ferramenta de anotação molecular, uma rápida caracterização dos compostos visando anotar

aqueles já relatados na literatura, evitando assim o isolamento de produtos naturais conhecidos (PILON et al., 2020; CARNEVALE-NETO et al., 2016)

Ela pode ser realizada comparando-se os dados obtidos experimentalmente com dados presentes em bancos de dados on line, como SciFinder Scholar® (©American Chemical Society), *Dictionary of Natural Products*® (DNP) (©Taylor & Francis Group), PubChem (©National Center for Biotechnology Information), dentre outros; bibliotecas *in house*, construídas a partir de buscas e revisões bibliográficas, permanecendo fechadas ao grupo de pesquisa, e através de bibliotecas espectrais de acesso livre, como o GNPS.

O *Global Natural Products Social Molecular Networking* (GNPS, Rede Mundial de Material Molecular de Produtos Naturais, disponível em <http://gnps.ucsd.edu>) é uma plataforma (ou ecossistema) online de dados espectrais que tem como objetivo prover o conhecimento em livre acesso para a comunidade científica e compartilhamento de dados. Para isso, ele possui ferramentas que auxiliam na anotação de compostos, tratamento de dados e análises de correlação (DEMARQUE et al., 2016; WANG et al., 2016).

Uma das principais abordagens da plataforma é a rede molecular (do inglês, *Molecular Network*), uma ferramenta de visualização do espaço químico na qual é apresentada utilizando espectros de massas, e as respectivas anotações realizadas pela compatibilidade espectral com a biblioteca do GNPS. Nestas redes, os espectros de cada substância são representados como nodos (*clusters*) e uma série de nodos conectados (*molecular families*) indica moléculas que apresentam semelhanças químicas e, portanto, em seus espectros de fragmentação. Assim, substâncias de uma mesma classe são agrupadas dentro de uma mesma rede (WANG et al., 2016). Além disso, os nodos podem ser complementados com dados de propriedades e classes utilizando metadados, promovendo a visualização de correlações e tendências.

Uma vez construída a rede, estão disponíveis ferramentas avançadas de análise para evoluir nas anotações, correlações e observação de padrões, como o NAP (*Network Annotation Propagation*) (DA SILVA et al., 2018) que permite a propagação *in silico* das anotações realizadas na rede molecular através de para auxiliar na desreplificação de análogos; o MS2LDA, que decompõe a fragmentação de dados moleculares e mostra padrões de fragmentação em massa e/ou perdas neutras que frequentemente representam subestruturas moleculares; MolNetEnhancer , que

promove a compilação dos dados e realiza a classificação das moléculas anotadas, com auxílio do ClassyFire, dentre outras.

2. OBJETIVOS

Este capítulo teve como objetivos principais analisar as impressões digitais metabólicas de 130 espécies de Orchidaceae, dentre elas brasileiras e exóticas, quanto seus metabólitos secundários e a diversidade química, através de ferramentas metabolômicas de anotação e correlação de dados, afim de contribuir para o conhecimento dos metabólitos das espécies brasileiras de Orchidaceae.

Os objetivos específicos foram:

- a. Coleta das espécies brasileiras e exóticas de Orchidaceae;
- b. Obtenção dos extratos de média-baixa e média-alta polaridades;
- c. Análise dos extratos de média-baixa polaridade por GC-EI-MS QP2010 (Shimadzu);
- d. Análise dos extratos de média-alta polaridade por UFLC-DAD-ESI-MS/MS (micrOTOF-QII);
- e. Processamento das impressões digitais metabólicas;
- f. Construção das redes moleculares pela plataforma GNPS;
- g. Análises metabolômicas utilizando ferramentas disponíveis na plataforma GNPS, como a Rede Molecular, NAP, MS2LDA e MolNetEnhancer, além de análises de Enriquecimento de Classe via pyClassRich, para aperfeiçoar as anotações e discussões;
- h. Análises de correlação in silico buscando padrões de similaridade química entre as espécies estudadas.

3. MATERIAL E MÉTODOS

Os procedimentos experimentais utilizados neste capítulo foram otimizados e descritos no Capítulo 1, desde a coleta até a construção da rede molecular. Abaixo estão descritos os métodos otimizados, e as análises bioinformáticas, de correlação e de enriquecimento de classe.

3.1. Material Vegetal

As 130 espécies de Orchidaceae foram fornecidas pelo Prof. Emerson Ricardo Pasarin, responsável pelo orquidário do Laboratório de Biologia Molecular e

Biossistemática de Plantas (LBMBP-USP) da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (FFCLRP-USP). Foram também selecionadas algumas espécies exóticas como grupo externo. Os respectivos identificadores, nomes científicos, Vouchers e informações sobre as espécies estão no Apêndice I.

Foram utilizadas as folhas e pseudobulbos que se apresentaram saudáveis e desenvolvidas. As amostras foram transferidas para tubos *falcon* de 50 mL, mergulhadas em nitrogênio líquido e liofilizadas a -90°C . Em seguida, foram pulverizadas e o tamanho de partículas foi padronizado utilizando tamiz de malha granulométrica 45 mesh. O material pulverizado foi armazenado em congelador -26°C em frascos de vidro vedados.

3.2. Preparação dos Extratos

3.2.1. Compostos de média-alta polaridade

As amostras foram preparadas utilizando 10,0 mg de droga vegetal pulverizada e extraídas com 1,0 mL de solução MeOH:H₂O (7:3, v/v) em banho de ultrassom por 10 minutos, seguida de centrifugação à 15000 rpm (26°C) por 10 minutos, filtração em filtro de membrana PTFE (0,22 μm) e secagem em centrífuga à vácuo à temperatura ambiente, com posterior pesagem para cálculo de rendimento em massa. Os extratos secos foram acondicionados em freezer à -26°C até o momento da análise.

3.2.2. Compostos de média-baixa polaridade

As amostras foram preparadas adicionando 2,5 mL de CH₂Cl₂ à 50,0 mg de droga vegetal, que foram homogeneizadas por 10 segundos em vórtex e colocadas em banho de ultrassom durante 20 minutos. Em seguida, o extrato foi centrifugado por 10 minutos à 2000 rpm, filtrado com filtro de membrana PTFE (0,22 μm) resistente ao solvente para um *via*/de vidro e o solvente foi evaporado em temperatura ambiente. Os extratos secos foram acondicionados em freezer à -26°C até o momento da análise, no qual os extratos foram ressuspensos com CH₂Cl₂ para uma concentração de 5,0 mg.mL⁻¹.

3.3. Cromatografia líquida de ultrarrápida eficiência acoplada a espectrometria de massas – UFLC-DAD-ESI-MS/MS (micrOTOF-QII)

Os extratos de média-alta polaridade foram analisados em um sistema UFLC-DAD-ESI-MS/MS (micrOTOF-QII) (Bruker Daltonics), utilizando uma coluna C18 (Kinetex™, Phenomenex) de tamanho de partícula de 2,7 µm (150 x 4,6 mm). O detector de UV (DAD) foi programado na faixa de comprimentos de onda entre 100 e 900 nm e a temperatura do forno foi de 35 °C. Foi utilizado CH₃CN-HCO₂H (0,1%) (v/v) (B) e H₂O-HCO₂H (0,1%) (v/v) (A) como fase móvel, a uma vazão de 1,0 mL.min⁻¹, volume de injeção 25 µL e o seguinte gradiente de eluição: 0 – 2 min, 3% B; 2 – 30 min, 3 - 100% B; 30 – 34 min, 100% B; 34 – 37 min, 100 - 3% B, 37 – 40 min, 3% B, sendo o calibrante analisado no intervalo de 37 – 40 min.

As análises por espectrometria de massas foram realizadas nos modos de ionização positivo e negativo, separadamente, utilizando a taxa de espectros (*spectra rate*) de 2,0Hz e faixa de massa de 50 a 1300 *m/z*. Os parâmetros da fonte de ionização foram: voltagem do *End Plate Offset* de 450V; voltagem do capilar, 3,0 kV (modo negativo) e 3,6 kV (modo positivo); nebulizador a 5,0 bar; vazão do gás de secagem de 10,5 L/min; temperatura do gás de secagem, 220 °C.

Os parâmetros para a fragmentação foram Auto MS/MS com a seleção dos íons precursores em ciclos de 3,0 s, excluindo a seleção após detecção de 2 espectros do mesmo íon ou após 0,60min, reconsiderando sua seleção se a intensidade do íon precursor for o dobro da anterior. Para as fragmentações dos íons precursores, foi utilizada uma rampa de energia de colisão de 20 a 50 eV.

De modo a obter massas de alta resolução, o instrumento e os experimentos foram calibrados a partir de uma solução de padrão interno de Na-TFA (400 mg/100 mL) ao final de cada análise. Os dados foram analisados usando o software DataAnalysis 4.3 (Bruker, Daltonik). A cada dez amostras, um branco contendo por solução extratora e um QC-Pool foi analisado. O QC-Pool é uma mistral realizada para controle de qualidade das análises realizadas, contendo 10,0 µl de cada amostra a ser analisada.

As amostras foram preparadas solubilizando 1,0 mg de extratos secos com 1,0 mL de solução MeOH:H₂O (7:3, v/v), foram filtradas com filtro de seringa com membrana filtrante de PTFE (0,22 µm), acondicionada em *vials* de vidro de 2,0 mL e, em seguida, analisadas.

3.4. Cromatografia a gás acoplada a espectrometria de massas – GC-EI-MS QP2010 (Shimadzu)

As análises dos extratos de média-baixa polaridade foram realizadas por GC-EI-MS QP2010 (Shimadzu), e utilizou-se uma coluna capilar ZB-5MS (J & W Agilent) (30 m x 0,25 mm x 0,25 mm, 0,25 μm) acoplada a uma pré-coluna Duraguard – 10 m, hélio como gás de arraste a 1,0 mL.min⁻¹, fonte de ionização à 250°C, e volume de injeção 1,0 μL à 5,0 mg/mL.

As condições experimentais utilizadas foram: injeção split 1:5; temperatura do filamento, 230 °C e injetor, 270 °C; corrente de emissão de interferência, 60 μA e ionização por impacto de elétrons, 70 eV; programa de aquecimento da coluna: 0 – 3,0 min, 60 °C, rampa de temperatura até 330 °C (taxa de aquecimento, 10 °C.min⁻¹), e mantendo-se nesta temperatura por 20 min; $\Delta m/z$ de 40-600. Os dados obtidos foram analisados utilizando o software CGSolution e as bibliotecas NIST, Wiley e FFNSC. As impressões digitais metabólicas adquiridas estão no Apêndice II.

3.5. Processamento dos Dados

3.5.1. Compostos de média-alta polaridade

Os dados adquiridos nas análises por LC-MS/MS no modo de ionização positivo e negativo foram convertidos, separadamente, em arquivos .mzxml utilizando o software livre ProteoWizard - MSconvert (v.3, Proteowizard Software Foundation, Palo Alto, CA, EUA) (Chambers et.al. , 2012). Posteriormente, foram processados utilizando o software *MzMine*TM (v. 2.51, BMC Bioinformatics, Reino Unido) (Pluskal et.al., 2010)., com o intuito de obter cromatogramas alinhados, deconvoluídos, com mínima influência de ruído e isótopos, e que representem ao máximo os extratos analisados. Os parâmetros utilizados no processamento dos dados adquiridos no modo positivo e negativo estão nas Tabelas 1 e 2, respectivamente.

Após o processamento, foram exportados 3 arquivos de dados: arquivo .mgf, com as informações dos espectros de MS², um arquivo .csv, com a matriz de dados contendo os íons precursores selecionados para fragmentação (MS²), as espécies em que aparecem, as áreas dos picos, razões massa-carga (m/z) e tempos de retenção; e outro arquivo .csv com uma matriz de dados contendo os íons precursores totais (selecionados ou não para a fragmentação) (MS¹) com os mesmo descritores do arquivo contendo MS².

Tabela 1. Parâmetros utilizados no software *MzMine*TM para o processamento dos dados adquiridos no modo de ionização positivo por análise em UFLC-DAD-ESI-MS/MS (microTOF-QII).

Raw data methods/ Peak detection	<p>Mass detection: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS1 <u>Mass Detector:</u> Centroid - Noise level= 1.0E3</p> <p>Mass detection: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS2 <u>Mass Detector:</u> Centroid - Noise level= 1.0E1</p> <p>Chromatogram builder: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS1 Min time span (min)= 0.01; Min height= 3.0E3; <i>m/z</i> tolerance= 20,0 ppm</p>
Peak list methods	<p>Chromatogram deconvolution: <u>Baseline cut-off:</u> Min peak height = 3.0E3; peak duration= 0.01 – 4.0; baseline level= 3.0E3 <i>m/z</i> range for MS2 scan pairing (Da): 0.01; RT range for MS2 scan pairing (min): 0.2</p> <p>Isotopic peaks grouper: <i>m/z</i> tolerance= 20 ppm; Retention time tolerance= 0,2 min; Maximum charge= 3; Representative isotope= most intense.</p> <p>Alignment: <u>Join aligner:</u> <i>m/z</i> tolerance= 20 ppm; Weight for <i>m/z</i>= 75; Retention time tolerance= 0.4 min; weight for RT= 25</p> <p>Filtering: <u>Peak list rows filter:</u> Keep only peaks with MS2 scan (GNPS): OK Reset the peak number ID: OK</p>

Legenda: min, minutos; *m / z*, massa/carga; ppm, partes por milhão.

Tabela 2. Parâmetros utilizados no software *MzMine*TM para o processamento dos dados adquiridos no modo de ionização negativo por análise em UFLC-DAD-ESI-MS/MS (microTOF-QII).

Raw data methods/ Peak detection	<p>Mass detection: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS1 <u>Mass Detector:</u> Centroid - Noise level= 1.0E3</p> <p>Mass detection: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS2 <u>Mass Detector:</u> Centroid - Noise level= 1.0E1</p> <p>ADAP Chromatogram builder: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS1 Min group size in # of scan= 5 ; Group intensity threshold= 1.0E3; Min highest intensity= 1.5E3; <i>m/z</i> tolerance= 20,0 ppm</p>
Peak list methods	<p>Chromatogram deconvolution: <u>Wavelets (ADAP):</u> S/N threshold= 10; S/N estimator= Intensity window S/N; min feature height= 2.0E3; coefficient/area threshold= 10; Peak duration range= 0.01 – 4.0; RT wavelet range= 0.001-0.1 <i>m/z</i> center calculation: MEDIAN; <i>m/z</i> range for MS2 scan paring (Da): 0.01; RT range for MS2 scan paring (min): 0.2</p> <p>Isotopic peaks grouper: <i>m/z</i> tolerance= 20 ppm; Retention time tolerance= 0,2 min; Maximum charge= 3; Representative isotope= most intense.</p> <p>Alignment: <u>Join aligner:</u> <i>m/z</i> tolerance= 20 ppm; Weight for <i>m/z</i>= 75; Retention time tolerance= 0.2 min; weight for RT= 25</p> <p>Filtering: Peak list rows filter: Keep only peaks with MS2 scan (GNPS): OK Reset the peak number ID: OK</p>

Legenda: min, minutos; *m / z*, massa/carga; ppm, partes por milhão.

3.5.2. Compostos de média-baixa polaridade

Os dados adquiridos nas análises por GC-MS foram transformados para o formato .cdf, importados para a plataforma GNPS e processados através da ferramenta “*GC-MS EI Data Analysis – Data Processing – Deconvolution*”, (Aksenov et.al, 2021), a qual promove a deconvolução e alinhamento espectrais em todo o conjunto de dados utilizando o processamento através do algoritmo MSHub. Os parâmetros de processamento são ajustados automaticamente e internamente pelo MSHub, como os relacionados a simetria de pico e ajustes na linha de base. Os parâmetros envolvendo alinhamento dos cromatogramas em “*Advanced Clustering*” foram ajustados em: *Cluster_Spectra*, NO; *Time_Unit*, MIN; *RT_tolerance*, 120. Os dados produzidos foram inspecionados.

3.6. Redes Moleculares

3.6.1. Compostos média-alta polaridade – Modo de Ionização Positivo

Os dados brutos das análises por LC-MS/MS no modo de ionização positivo foram convertidos para o formato .mzXML usando software MSconvert (ProteoWizard) e enviados para a plataforma online do GNPS (<https://gnps.ucsd.edu>, Wang et al., 2016) pelo do software WinSCP (<https://winscp.net/eng/index.php>) para a construção das redes moleculares clássicas através da ferramenta “*Data Analysis – Molecular Networking*”, utilizando um fluxo de trabalho baseado na descrição fornecida (<https://ccms-ucsd.github.io/GNPSDocumentation/>) pela plataforma GNPS.

Os espectros de MS/MS foram convertidos em vetores e os repetidos foram combinados utilizando o algoritmo MS-Cluster para criar o espectro consenso. Para formação da rede, foi considerado os valores de tolerância de massa de íons precursores e de íons fragmento de, respectivamente, 0,08 Da e 0,02 Da. Na construção dos agrupamentos (*cluster*), o valor mínimo de cosseno de similaridade para que fosse formada uma borda (*edge*) e conectasse dois espectros consenso (nodos) foi de 0.7, com no mínimo 3 íons fragmento compatível no espectro. As bordas foram mantidas entre dois nodos somente se cada nodo está entre os 10 mais similares (*Network TopK*) e o tamanho máximo de uma família molecular (*cluster*) foi definido como 130 (*Maximum Connected Component Size*), e o tamanho mínimo como 2 (*Minimum Cluster Size*). A máxima diferença de massa entre os precursores foi de 500 Da. Logo depois, os espectros consenso da rede foram pesquisados nas bibliotecas espectrais do GNPS, considerando compatíveis os espectros que possuíam cosseno de similaridade maiores que 0,7 e pelo menos 3 íons fragmento em comum.

A visualização das redes moleculares foi realizada no Cytoscape 3.2 (Shannon et al., 2003), o qual permitiu a visualização na forma de uma rede de agrupamentos de nodos (íons precursores) e arestas (valores de cosseno de similaridade entre análogos). Cada agrupamento de nodos corresponde a uma família molecular (*cluster*).

Os arquivos .mgf e .csv (MS²), adquiridos após processamento dos dados das análises por LC-MS/MS no modo de ionização positivo foram enviados para a plataforma online do GNPS (<https://gnps.ucsd.edu>, Wang et al., 2016) pelo do

software WinSCP (<https://winscp.net/eng/index.php>), para a construção das redes moleculares utilizando a ferramenta “*Feature-Based Molecular Networking (FBMN)*” (NOTHIAS et al., 2019).

A formação da rede considerou os valores de tolerância de massa de íons precursores e de íons fragmento de, respectivamente, 0,08 Da e 0,02 Da. Na construção dos agrupamentos (*cluster*), o valor mínimo de cosseno de similaridade para que fosse formada uma borda (*edge*) e conectasse dois espectros consenso (nodos) foi de 0.7, com no mínimo 3 íons fragmento compatível no espectro. As bordas foram mantidas entre dois nodos somente se cada nodo está entre os 10 mais similares (*Network TopK*) e o tamanho máximo de uma família molecular (*cluster*) foi definido como 100 (*Maximum Connected Component Size*). A máxima diferença de massa entre os precursores foi de 500Da.

Os espectros da rede foram pesquisados nas bibliotecas espectrais do GNPS (Wang et al., 2016; Horai et al., 2010) considerando compatíveis os espectros que possuíam cosseno de similaridade maiores que 0,7 e pelo menos 3 íons fragmento em comum. A ferramenta DEREPLICATOR foi utilizada nas análises dos FBMN para anotar os espectros de MS/MS (Mahimani et al., 2018) e a rede foi visualizada usando o Cytoscape 3.2 (SHANNON et al., 2003).

3.6.2. Compostos média-alta polaridade – Modo de Ionização Negativo

Os dados brutos das análises por LC-MS/MS no modo de ionização negativo foram convertidos para o formato .mzXML usando software MSconvert (ProteoWizard) e enviados para a plataforma online do GNPS (<https://gnps.ucsd.edu>, Wang et al., 2016) pelo do software WinSCP (<https://winscp.net/eng/index.php>) para a construção das redes moleculares clássicas através da ferramenta “*Data Analysis – Molecular Networking*”, utilizando um fluxo de trabalho baseado na descrição fornecida (<https://ccms-ucsd.github.io/GNPSDocumentation/>) pela plataforma GNPS. A construção e visualização da rede molecular ocorreu de maneira idêntica aos dados brutos analisados no modo de ionização positiva, descritos no item anterior (item 3.6.1).

Os arquivos .mgf e .csv (MS²), adquiridos após processamento dos dados das análises por LC-MS/MS no modo de ionização negativo foram enviados para a plataforma online do GNPS (<https://gnps.ucsd.edu>, Wang et al., 2016) pelo do

software WinSCP (<https://winscp.net/eng/index.php>), para a construção das redes moleculares utilizando a ferramenta “*Feature-Based Molecular Networking (FBMN)*” (Nothias et al., 2019). A construção e visualização da rede molecular ocorreu de maneira idêntica aos dados brutos analisados no modo de ionização positiva, descritos no item anterior (item 3.6.1). A visualização das redes moleculares foi realizada no Cytoscape 3.2 (SHANNON et al., 2003).

3.6.3. Compostos média-baixa polaridade

As impressões digitais metabólicas adquiridas por GC-MS foram processadas através da ferramenta “*GC-MS EI Data Analysis – Data Processing – Deconvolution*”, disponível na plataforma GNPS e os dados processados foram exportados diretamente do processamento para a construção da rede molecular através da opção “*Continue Processing – Search Spectral Library and Molecular Network*”, em “*Job Status*”. Este fluxo de trabalho foi baseado na descrição fornecida (<https://ccms-ucsd.github.io/GNPSDocumentation/>) pela plataforma GNPS.

A formação da rede considerou os valores de tolerância de massa de íons precursores e de íons fragmento de, respectivamente, 20000 Da e 1 Da. Na construção dos agrupamentos (*cluster*), o valor mínimo de cosseno de similaridade foi de 0.7, com no mínimo 6 íons fragmento compatível no espectro. As bordas foram mantidas entre dois nodos somente se cada nodo está entre os 10 mais similares (*Network TopK*) e o tamanho máximo de uma família molecular (*cluster*) foi definido como 100 (*Maximum Connected Component Size*). A máxima diferença de massa entre os precursores foi de 500Da.

Todas as correspondências entre os espectros da rede e os espectros da biblioteca deveriam ter um valor acima de 0,80 e pelo menos 6 íons fragmento correspondentes. A visualização das redes moleculares foi realizada no Cytoscape 3.2 (SHANNON et al., 2003)

3.7. Propagação de Anotações por Redes Espectrais in silico, Databases e Ranqueamento – Network Annotation Propagation (NAP)

Os dados gerados pelas redes moleculares utilizando os dados de LC-MS (modo de ionização positivo e negativo, separadamente) foram posteriormente analisados pela ferramenta Rede de Propagação de Anotações (DA SILVA et al.,

2018) (Network Annotation Propagation - NAP). O NAP utiliza de ferramentas de fragmentação *in silico* (MetFrag) juntamente com bases espectrais do GNPS, dicionário de produtos naturais (Dictionary of Natural Products - DNP), Super Natural II e NuBBEDB na tentativa de classificação e ranqueamento de possíveis candidatos espectrais, ou seja, anotações moleculares.

Todos os íons precursores foram considerados como moléculas protonadas $[M+H]^+$ ou cationizadas $[M+Na]^+$ e a exatidão de massa para os candidatos foi configurada a 10 ppm. As pontuações de fusão e consenso foram calculadas baseado nos 10 primeiros candidatos na fase da rede de propagação.

3.8. Análises de Motivos (Motifs) ou anotação por Mass2Motifs (MS2LDA)

Os dados gerados pelas redes moleculares utilizando os dados de LC-MS (modo de ionização positivo e negativo, separadamente) foram posteriormente analisados pela ferramenta MS2LDA (VAN DER HOOFT et al., 2016; WANG et al., 2016). Após a construção da rede molecular, os arquivos com os espectros de fragmentação e os pares obtidos na rede molecular foram exportados diretamente através da opção “*Analyze with MS2LDA*” em “*Advanced Views – Experimental Views*” em “*Job Status*”. Este fluxo de trabalho foi baseado na descrição fornecida (<https://ccms-ucsd.github.io/GNPSDocumentation/>) pela plataforma GNPS.

A matriz de dados no formato .csv contendo os dados de quantificação dos dados filtrados advindos do Mzmine foi adicionado para realizar as correlações. Em “*Advanced MS2LDA parameters*”, quatro parâmetros foram ajustados. A janela de tolerância de massa (*Bin Width*) selecionada foi a compatível com dados de TOF 0.01 Da, o número de interações LDA foi de 1000, a intensidade mínima dos picos de MS2, parâmetro o qual permite filtrar o ruído base dos espectros de MS/MS, foi de 100 e o número de Mass2Motifs livres que são adicionados na parte superior dos Mass2Motifs anotados com prefixo das bases de dados (*MotifDB*), chamado de “*LDA Free Motifs*” foi de 300. Todas as bases de dados Motif foram incluídas na análise, em “*MotifDB Selection*”.

A análise gera um novo arquivo utilizado para enriquecer a rede molecular criada anteriormente, incluindo distribuições Mass2Motif através das famílias moleculares. Para isso, é necessário fornecer algumas informações, situadas em “*Advanced MS2LDA Parameters for Networking File Output*”. Os valores de limite de

sobreposição e probabilidade de sobreposição foram ajustados em 0.3 e 0.1, respectivamente. O TopX nos nodos foi ajustado como 5, o qual corresponde ao Mass2Motifs com o maior valor de limite de sobreposição que foram listados no novo arquivo contando novos valores de borda.

3.9. MolNetEnhancer e ClassyFire

Os dados gerados pelas redes moleculares utilizando os dados de LC-MS (modo de ionização positivo e negativo, separadamente) foram posteriormente analisados pela ferramenta MolNetEnhancer (ERNST et al., 2019). Após a construção da rede molecular, os arquivos com os espectros de fragmentação e os pares obtidos na rede molecular foram exportados diretamente através da opção “*Enhance with MolNetEnhancer*” em “*Advanced Views – Experimental Views*” em “*Job Status*”. Este fluxo de trabalho foi baseado na descrição fornecida (<https://ccms-ucsd.github.io/GNPSDocumentation/>) pela plataforma GNPS.

Esta ferramenta possibilita a combinação dos dados exportados das análises realizadas previamente (FBMN, NAP, DEREPLICATOR, MS2LDA) e a classificação química automatizada realizada através da plataforma ClassyFire (FEUNANG et al., 2016). Dessa maneira, a rede molecular é enriquecida e fornece uma visão química mais abrangente dos dados metabolômicos (ERNST et al., 2019). No campo “*Molecular Network Enhancer*” foram adicionados os identificadores de cada análise realizada na plataforma do GNPS.

O ClassyFire é uma ferramenta baseada em estruturas químicas capaz de atribuir automaticamente uma taxonomia químico-estrutural composta por mais 4800 categorias. Esta taxonomia química é subdividida em 11 níveis diferentes (Reino > SuperClasse > Classe > Subclasse, etc.) com cada uma das categorias definidas por regras estruturais computacionais não ambíguas (FEUNANG et al., 2016).

Os compostos anotados nas pelas redes moleculares geradas a partir dos dados de GC-MS foram classificados diretamente pelo ClassyFire (<http://classyfire.wishartlab.com/>), uma vez que não há a opção de análises pelo MolNetEnhancer utilizando os dados de GC-MS.

3.10. Análises de Enriquecimento de Classes Químicas

A análise do enriquecimento de classes químicas foi realizada em colaboração com o Dr. Ricardo R. da Silva, Cientista Jovem Pesquisador FAPESP, responsável pelo Laboratório de Biologia e Química Computacional da Faculdade de Ciências Farmacêuticas de Ribeirão Preto – FCFRP-USP.

A análise foi realizada através do *pyClassRich*, o qual realiza a classificação de cada composto anotado utilizando a ferramenta ClassyFire, submete cada composto anotado e classificado à estatísticas univariadas obtendo sua importância dentro do táxon, realiza a análise de enriquecimento de classes utilizando a ferramenta ChemRICH. Os dados utilizados nas análises foram os das redes moleculares dos dados de LC-MS no modo positivo, as anotações obtidas pelo pareamento espectral do GNPS, a predição *in silico* utilizando o NAP e o arquivo de *metadata*, com as informações sobre as classes a serem analisadas.

3.11. Análises Quimiométricas

As matrizes de dados produzida após processamento de dados de LC-MS e GC-MS foram utilizadas para a realização das análises estatísticas multivariadas supervisionadas e não-supervisionadas. As análises não-supervisionada foram realizadas com o intuito de observar como se comportam as amostras através de padrões e tendências de agrupamentos em relação à similaridade química, sendo essas as análises realizadas por HCA e PCA.

As Análises de Agrupamentos Hierárquicos (*Hierarchical Cluster Analysis* - HCA) foram realizadas no software R (*R Project for Statistical Computing*) através do pacote *dendextend*, utilizando como algoritmo de agrupamento Distância Camberra com método de ligação do tipo Ward. As Análises de Componentes Principais (*Principal Component Analysis* - PCA) foram realizadas no software SIMCA 13.0.3.0[®] (Umetrics[®]) utilizando a transformação por log em todas as matrizes.

As análises supervisionadas foram realizadas para propor as substâncias responsáveis pelos padrões de similaridade observados nas análises anteriores. As Análises Discriminantes por Mínimos Quadrados Parciais (*Partial Least Square – Discriminant Analysis* – PLS-DA) e as Análises Discriminantes por Mínimos Quadrados Parciais Ortogonais (*Orthogonal Partial Least Square – Discriminant Analysis* – OPLS-DA) foram realizadas utilizando o software SIMCA 13.0.3.0[®] (Umetrics) utilizando a transformação por Log em todas as matrizes. A partir desses,

foram obtidas as variáveis (metabólitos) discriminantes para cada grupo, as quais foram selecionadas a partir do gráfico VIP, que indica a importância da variável para a projeção do grupo no modelo estatístico, e o gráfico de coeficientes, os quais são utilizados para interpretar a influência das variáveis X (metabólitos) em Y (táxon).

4. RESULTADOS E DISCUSSÃO

4.1. Compostos de média-baixa polaridade

O processamento dos dados gerados a partir das análises por GC-MS, realizado na plataforma GNPS, gerou uma matriz de dados químicos com 1633 íons precursores entre amostras, brancos e padrões, e dentre esses somente 1056 são íons referentes às amostras. Dentre esses, 796 foram anotados e, após conferir a compatibilidade dos espectros de fragmentação dos anotados com os da biblioteca espectral no GNPS, 539 anotações foram obtidas, resultando em 33% de anotação do metaboloma detectado. As anotações foram classificadas por superclasse, classe e subclasse usando o ClassyFire (Feunang *et.al.*, 2016).

Dentre as anotações estão presentes 11 superclasses, 66 classes e 111 subclasses químicas. O Gráfico 1 traz as 11 superclasses e abundância dessas dentre as moléculas anotadas. Os Gráficos 2 e 3 mostram as 10 classe e subclasses, respectivamente, mais abundantes dentre os compostos anotados na rede molecular.

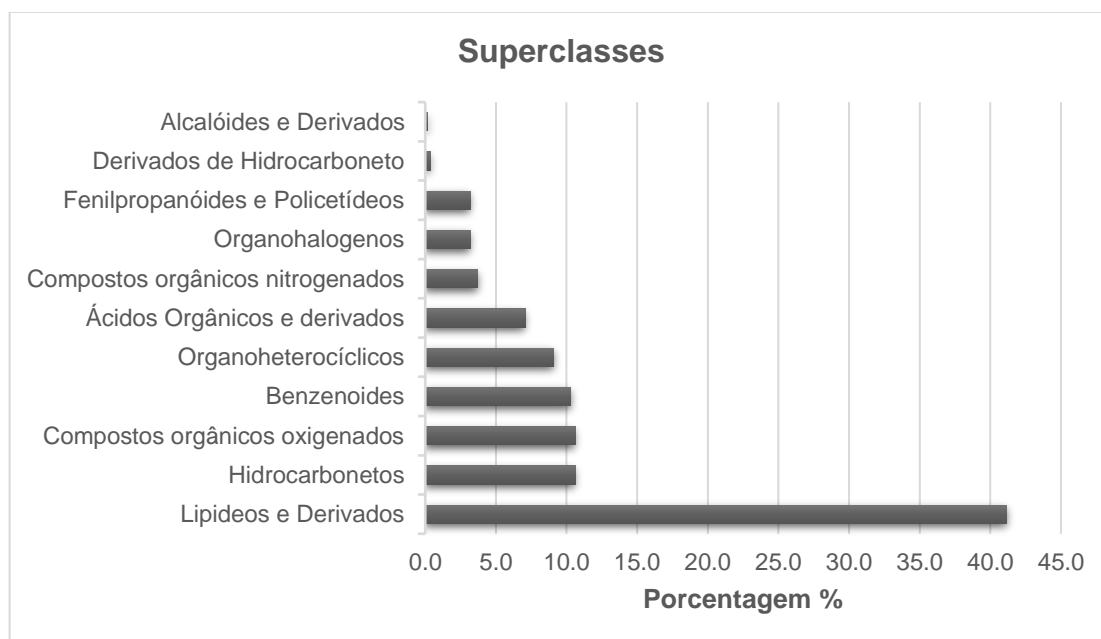


Gráfico 1. Superclasses dos compostos anotados na rede molecular de compostos de média-baixa polaridade advindos dos extratos de Orchidaceae e analisados por GC-MS.

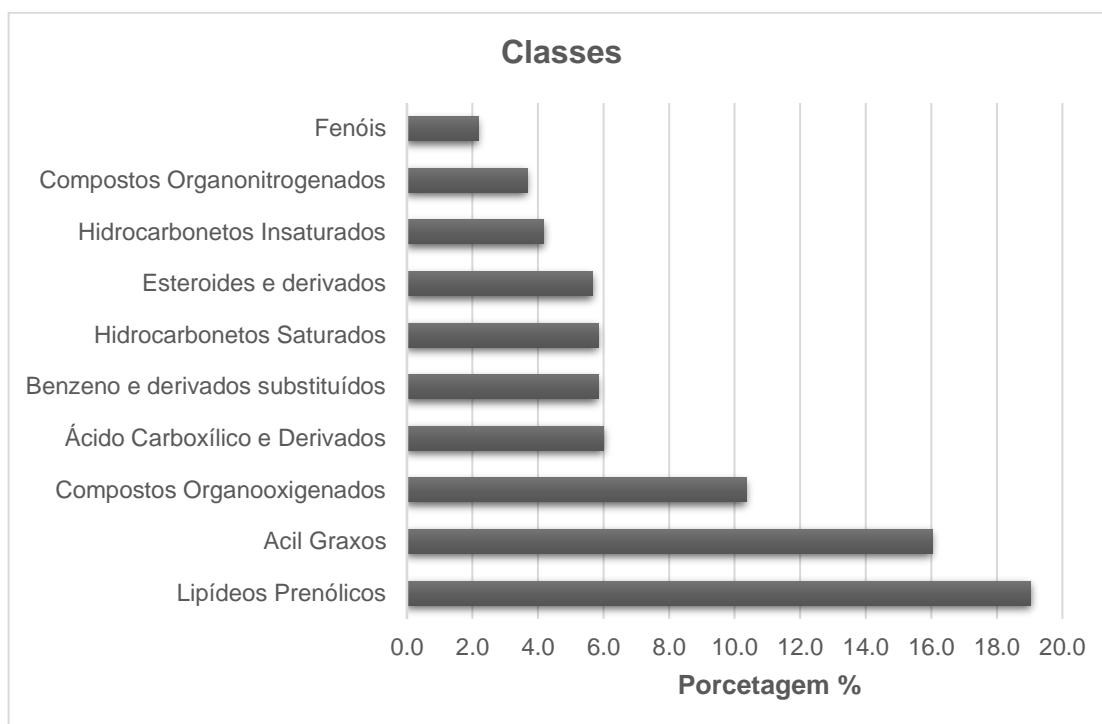


Gráfico 2. Dez classes mais abundantes dentre os compostos anotados na rede molecular de compostos de média-baixa polaridade advindos dos extratos de Orchidaceae e analisados por GC-MS.

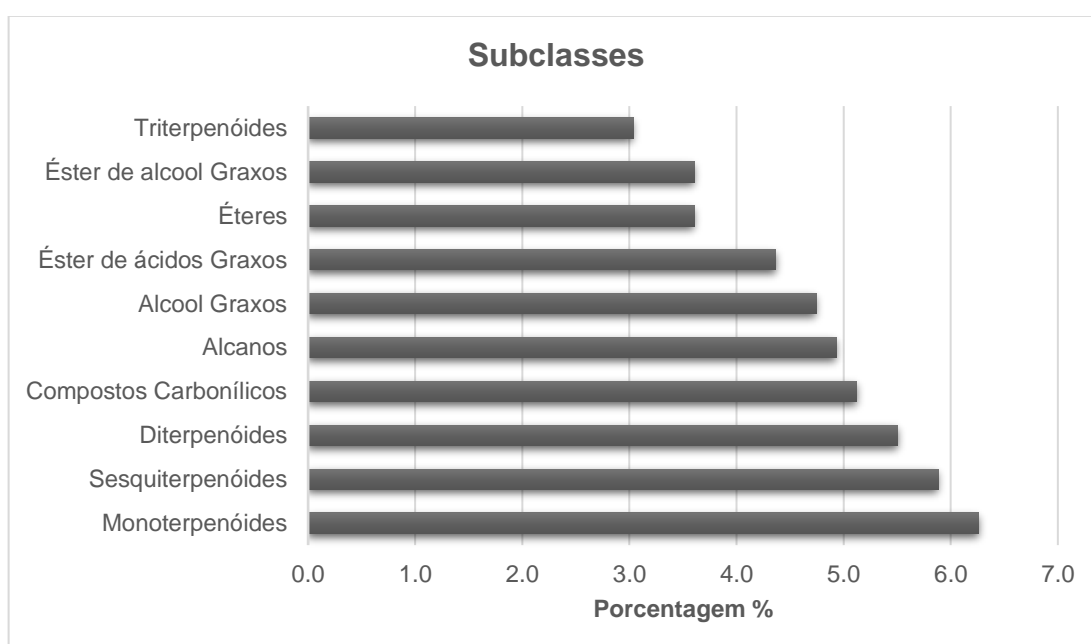


Gráfico 3. Dez subclasses mais abundantes dentre os compostos anotados na rede molecular de compostos de média-baixa polaridade advindos dos extratos de Orchidaceae e analisados por GC-MS.

Dentre os compostos anotados, 8 subclasses de compostos foram detectadas em todas as amostras: alcanos, chalconas e dihidrochalconas, esteroides derivados de ergostano, quinonas e hidroquinonas lipídicas, sesquiterpenos, ácidos tetracarboxílicos e derivados, e triterpenos. Em acima de 95% das amostras foram detectadas também antraquinonas, monoterpenos, diterpenos, sesterterpenos, estilbenos, esteroides derivados de colestano, lactonas esteroidais, aminas, derivados de ácidos linoleicos, macrolidos e análogos, ácidos graxos e conjugados, éster de ácidos graxos, álcoois graxos, dentre outros. A tabela 3 a seguir mostra a distribuição das algumas subclasses dentre as 132 espécies estudadas.

Tabela 3. Distribuição de algumas subclasses presentes nas amostras analisadas e de interesse farmacológico, de acordo com os compostos anotados nas redes moleculares advindos dos extratos de Orchidaceae analisados por GC-MS.

Subclasse (ClassyFire)	Porcentagem %	Subclasse (ClassyFire)	Porcentagem %
Chalconas e dihidrochalconas	100	Lactonas terpênicas	85
Esteróides derivados de Ergostano	100	Benzilaminas	83
Quinona e hidroquinona lipídica	100	Esteróides derivados de Estrano	76
Sesquiterpenos	100	Hidroxiesteróides	76
Triterpenos	100	Quinolizinas	72
Antraquinonas	99	Flavanas	71
Diterpenos	99	Dialquiltioeteres	69
Monoterpenos	99	Esteres de esteróides	69
Sesterterpenos	99	Flavonas	64
Estilbenos	99	Sesquaterpenos	47
Esteróides derivados de Colestano	98	Oxepanos	37
Lactonas esteroidais	97	Hidroquinolinas	34
Aminas	96	Cumarinas e derivados	33
Estigmastanos e derivados	93	Alcaloides tropânicos	13

%. Porcentagem

Esta abordagem metabolômica utilizada permitiu a anotação de diferentes classes químicas que revelam as espécies brasileiras de Orchidaceae como uma

fonte rica de compostos de interesse farmacológico. Abaixo estão algumas dessas classes, assim como na figura 1.

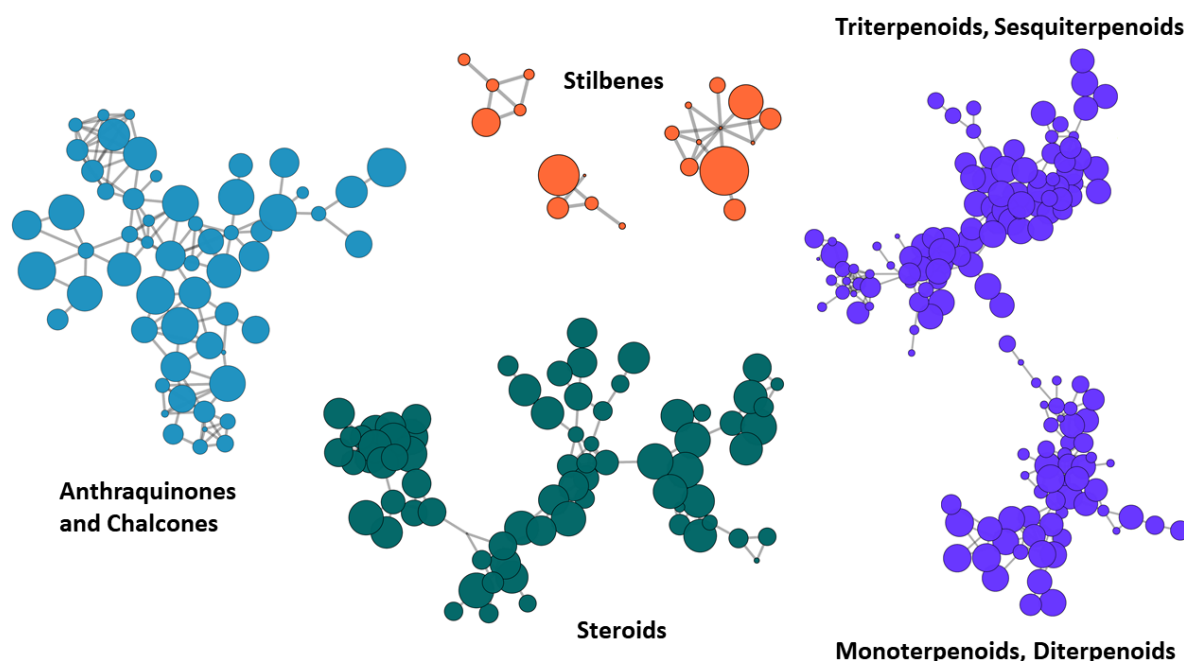


Figura 1. Famílias moleculares, relacionadas a atividades biológicas, da rede molecular advinda dos extratos de Orchidaceae analisados por GC-MS, coloridas de acordo com as classes químicas. O tamanho do cluster (nodo) está relacionado ao número de espectros.

- *Terpenos*

Os terpenos são a maior classe de metabólitos secundários dentro do reino vegetal, incluindo mais de 40 mil compostos diferentes. Eles possuem múltiplos papéis fisiológicos e ecológicos, são essenciais para o crescimento e desenvolvimento da planta, para a atração de polinizadores e para a defesa, além de possuir propriedades inseticidas e serem utilizados como agentes farmacológicos (HUANG et al., 2021)

Dentre os terpenos anotados, os sesquiterpenos e triterpenos estão presentes em 100% das amostras analisadas e os monoterpênicos, diterpenos e sesterterpenos, em mais de 95%. Outras subclasses também foram anotadas, como sesquiterpenos e lactonas terpênicas.

Diversos sesquiterpenos já foram descritos para Orchidaceae, principalmente para o gênero *Dendrobium*, como o sesquiterpeno amino éter wardianumina A (ZHANG et al., 2017), os sesquiterpenos dos tipos picrotoxano, emmotina e *allo*-aromadentrano (BOLYEN et al., 2019; LAM et al., 2015; ZHAO et al., 2003), amontina, amoenina, dendronobilina A (LAM et al., 2015), sesquiterpenos glicosídeos como

dendrosídeo D, E, F e G (YE; QIN; ZHAO, 2002), dentre outros. Ye e colaboradores (2002) isolaram sesquiterpenos glicosídicos e avaliaram a proliferação de células T e B em ratos. Foi observado que os dendrosídeos D-G promoveram a proliferação celular e estimulação das células T e/ ou linfócitos B de rato, mostrando a ação imunomodulatória.

Alguns dos sesquiterpenos anotados foram nerolidol, cis-tujopseno, cariofileno, β -cariofileno, t-muurolol, acetato de cis-nerolidil, γ -elemeno, 6,10,14-trimetilpentadecan-2-ona, dentre outros. O nerolidol possui um papel ativo no sistema de defesa da planta contra a herbivoria, e várias atividades farmacológicas já foram relatadas para ele, como a ação antioxidante, antiparasitária, antimicrobiana, antifúngica, antiobiofilme, antinociceptiva, anti-inflamatória, anticancerígena dentre outras (CHAN et al., 2016). No estudo de Nunes e colaboradores (NUNES et al., 2017) o γ -elemeno demonstrou potente atividade leishmanicida, pois gerou uma grande inibição do crescimento de promastigotas de *Leishmania amazonenses* e estimulou a produção de ERO (reativos de oxigênio) por macrófagos quando incubados com as formas promastigotas.

A ocorrência dos triterpenos na família Orchidaceae foi descrita em diversas espécies, como em *Agrostophyllum brevipes*, *Agrostophyllum callosum*, *Cimbidium giganteum*, *Cremastra appendiculata*, *Pleione formosana*, *Epidendrum mosenii*, dentre outras (DAHMEÑ; LEANDER, 1978; MAJUMDER; BANDYOPADHYAY, 2010; SHIAO; CHEN; LIN, 2009; XIAO et al., 2008). Os triterpenos possuem propriedades farmacológicas como atividades hipoglicêmicas, antiparasitárias e imunomoduladoras pela inibição da elastase de neutrófilos (BOLYEN et al., 2019). Além disso, já foram descritas ações anti-inflamatórias, antifúngicas, antibacteriana, antiviral, antiulcerogênico, hepatoprotetor, analgésico, dentre outras (GONZÁLEZ-COLOMA et al., 2011).

Alguns dos triterpenos anotados foram: esqualeno, acetato de lupan-3ol, acetato de sitosterol, lanosterina, dihidrolanosterol, serratenol, 21-episerratenediol, betulina, dentre outros. O esqualeno é um triterpeno precursor na biossíntese de colesterol e hormônios esteróis, e é utilizado de maneira extensa como excipiente na indústria farmacêutica e no controle e tratamento de doenças. Além disso, atua como agente protetor e demonstrou auxiliar na diminuição dos efeitos colaterais induzidos pela quimioterapia. A betulina é um triterpeno pentacíclico distribuído em muitas

plantas, e possui vários estudos que mostram o amplo espectro de suas propriedades farmacológicas e biológicas, como antibacteriana, antifúngica, antiviral, além de anticancer, diminuindo a proliferação de células cancerígenas de vários tipos (KRÓL et al., 2015).

As lactonas terpênicas foram reportadas em sua maioria em plantas, e são os bioativos principais dos extratos padronizados de *Ginkgo biloba*, utilizados de maneira ampla como uma terapia complementar ao tratamento de doenças cardiovasculares e cerebrovasculares isquêmicas (ISHMURATOV et al., 2015; LIU et al., 2018).

Monoterpenos, diterpenos e sesquiterpenos foram muito estudados nas espécies de Orchidaceae em estudos relacionados à ecologia química, à composição da essência floral e seus compostos voláteis, às rotas biossintéticas desses terpenos e a regulação da transcrição gênica para sua produção (RAMYA et al., 2018). Monoterpenos e sesquiterpenos compõe a maioria dos componentes dos óleos essenciais, sendo produzidos tanto para a atração de insetos para a polinização, como para proteção de diversas doenças de raiz fúngica ou bacteriana. Também podem ser usados com fim medicinal, como mucolíticos ou expectorantes, para afecções do trato gastrointestinal, e até para melhorar a circulação sanguínea, dentre outros (MÜLLER; BUCHBAUER, 2011).

Alguns dos monoterpenos anotados foram: β -bergamoteno, carano, citral, fenchona, acetato de geranyl, neral, norbonano, norcamphor, ácido perílico, acetato de p-menth-8-en-3-ol, cis-crisantemol, citronitrila, dentre outros. O ácido perílico, um metabólito do álcool perílico, teve seu potencial explorado em diversas áreas do conhecimento, mas foi amplamente estudado em relação ao seu potencial anticâncer e imunomodulatório. No estudo de Haber et.al. (HABER et al., 2019) foi avaliada a influência do β -bergamoteno na atração de polinizadores e observou-se que este impulsiona as preferências dos polinizadores independente de recompensas de pólen.

- *Chalconas e Dihidrochalconas*

As chalconas e dihidrochalconas possuem um esqueleto químico comum de 1,3-diaril-2-propen-1-ona, existindo os isômeros *cis* e *trans*, sendo este último o mais estável termodinamicamente (ZHUANG et al., 2017). Esta subclasse está presente em todas as espécies estudadas, e foram anotados três: trans-3',4',5'-trimetoxi-4-

(metiltio)chalcona, trans-4'-dimetilamino-4-(metiltio)chalcona e trans-4'-fluoro-4-(metiltio)chalcona.

A utilização das chalconas para fins terapêuticos data de milhares de anos, através do uso de plantas e produtos naturais para diferentes distúrbios, como câncer, inflamação e diabetes. Algumas chalconas foram aprovadas para testes clínicos, como a metochalcona, com ação colerética, e sofalcona, usada como medicamento antiúlcera e mucoprotetor (ZHUANG et al., 2017)

Sua importância para a família Orchidaceae é tanta que estudos foram desenvolvidos para analisar a evolução molecular, a especialização funcional da enzima Chalcona Sintase (CHS), o gene da enzima em si e a relação com o taxon e o sucesso evolutivo da família, uma vez que essa enzima é a chave para a biossíntese de chalconas e flavonoides, metabólitos importantes para a adaptação da planta.(CHAN et al., 2016; HUANG et al., 2021; JIA et al., 2016; LIU et al., 2012) (Han et.al., 2006a; Han et.al., 2006b; Liu et.al., 2012; Kuo et.al.2019).

- *Estilbenos e Derivados*

Os estilbenos são uma classe de polifenóis vegetais conhecidos por sua estrutura complexa e amplas atividades biológicas, no entanto ocorre em um número limitado de famílias no reino vegetal, e Orchidaceae é uma delas (RIVIÈRE; PAWLUS; MÉRILLON, 2012; SHEN; WANG; LOU, 2009). Algumas espécies de Orchidaceae são conhecidas por acumular estilbenoides, utilizados como meio de proteção contra predação, como o acúmulo de cyrtopodinona, cyrtopodinol e coeludols A e B em raízes de *Cyrtopodium paniculatum* (AUBERON et al., 2017), phragmidimer A e B em algumas espécies de *Phragmipedium* 31] e thunalbeno em vários gêneros de Orchidaceae, como *Phragmipedium*, *Thunia* e *Pholidota* (RIVIÈRE; PAWLUS; MÉRILLON, 2012).

Além da sua ocorrência e o papel que exerce nos sistemas de defesa das plantas, os estilbenos têm demonstrado diversas atividades biológicas e potencial farmacológico, como anticancerígeno, antimicrobiano, antidiabético, osteoporose, cardioprotetor, antienvelhecimento, antioxidante, anti-inflamatório, entre outros(TEKA et al., 2022).

- *Esteróis*

Os esteróis vegetais apresentaram diversas atividades biológicas, como atividades anticarcinogênicas, anti-inflamatórias, antibacterianas e antifúngicas. Os esteróides e derivados anotados neste estudo pertencem às subclasses estigmastanos e derivados, lactonas esteroidais, ésteres de esteroides, oxosteróides, hidroxisteróides, esteróides derivados de estranos, ergostanos e androstanos. A Tabela 3 mostra a distribuição de alguns desses esteróides, mostrando o quão ampla é a ocorrência entre as espécies brasileiras estudadas.

Muitos dos produtos naturais esteroidais com esqueleto ergostano possuem atividades biológicas (DUECKER; REUSS; HERETSCH, 2019). O campesterol está presente em todas as espécies analisadas e é um dos esteroides vegetais mais comuns na natureza, e possui evidências de efeitos quimiopreventivos contra câncer de próstata, pulmão e mama. Os esteroides estigmastanos são conhecidos por sua atividade anti-inflamatória e as lactonas esteroidais são conhecidos pelos whitanolideos, como a Withaferina A que tem efeito inibitório contra vários tipos diferentes de células cancerígenas (ROY et al., 2013)

- *Antraquinonas*

As antraquinonas (9,10-dioxoantracenos) são uma importante classe de produtos naturais com uma ampla gama de aplicações e atividades biológicas, incluindo ação antioxidante, contra a esclerose múltipla, câncer, artrite e constipação. A antraquinona anotada emodin, (1,3,8-triidroxi-6-metil-antraquinona) é um composto amplamente estudado e sua função biológica está relacionada à proteção contra herbivoria, patógenos e fatores abióticos como a radiação solar. É um derivado natural utilizado como fármaco anticancerígeno ativo contra células de câncer de pulmão, mama, fígado e ovário, e tem sido estudado seu uso em combinação com quimioterápicos e medicamentos alopáticos para aumentar a eficácia e minimizar a toxicidade do tratamento. Além de sua atividade antineoplásica, apresenta atividade antimicrobiana, anti-inflamatória, antiúlcero-gênica, hepatoprotetora, imunossupressora e outras (SEMWAL et al., 2021).

4.2. Compostos de média-alta polaridade

A análise dos compostos de média-alta polaridade e anotações desses metabólitos secundários foram realizadas utilizando duas redes moleculares geradas

a partir dos dados analisados por LC-MS no modo positivo, LC-PR (otm02) e LC-PP (otm2), como descrito no primeiro capítulo desta tese. Assim, foram obtidas anotações de altos valores de cosseno de similaridade, demonstrando alta qualidade de anotação.

A rede molecular LC-PR (otm02), a qual utilizou os dados brutos, possui 7746 íons precursores e 562 compostos anotados, e, após conferir a compatibilidade dos espectros de fragmentação dos anotados com os da biblioteca espectral no GNPS, 394 anotações foram obtidas, resultando em 5,085 % de anotação do total detectado. A rede LC-PP (otm2), a qual utilizou os dados tratados, possui 10617 íons precursores e 1360 compostos anotados, e, após conferir a compatibilidade dos espectros de fragmentação dos anotados com os da biblioteca espectral no GNPS, 759 anotações foram obtidas, resultando em 7,14 % de anotação do total detectado.

As redes de propagação de anotações *in silico* (NAP) desses modelos proporcionaram as anotações de mais 168 análogos, não anotados anteriormente. Assim, foram anotados 1321 compostos a partir das redes moleculares geradas dos dados de LC-MS no modo positivo.

A rede molecular gerada a partir dos dados analisados por LC-MS no modo negativo (LC-NR (otm2)) possui 2853 íons precursores e 152 compostos anotados, e, após conferir a compatibilidade dos espectros de fragmentação dos anotados com os da biblioteca espectral no GNPS, 140 anotações foram obtidas, resultando em 4.9% de anotação do total detectado. As análises realizadas pelo NAP adicionaram à essas anotações 43 compostos. Assim, foram anotados 183 compostos a partir das redes moleculares geradas dos dados de LC-MS no modo negativo.

Assim, dentre anotações geradas no modo positivo e negativo, retirando as anotações duplicadas nos dois modos de ionização, totalizou-se em 1205 compostos anotados dentre os compostos de média-alta polaridade.

Através das análises pela ferramenta MolNetEnhancer, foram obtidas as anotações à nível de classes moleculares. Assim, foram observadas 12 superclasses, 67 classes e 99 subclasses. O Gráfico 4 traz as superclasses e abundância dessas dentre as moléculas anotadas, e os Gráficos 5 e 6 mostram as classe e subclasses, respectivamente, mais abundantes dentre os compostos anotados nas redes moleculares.

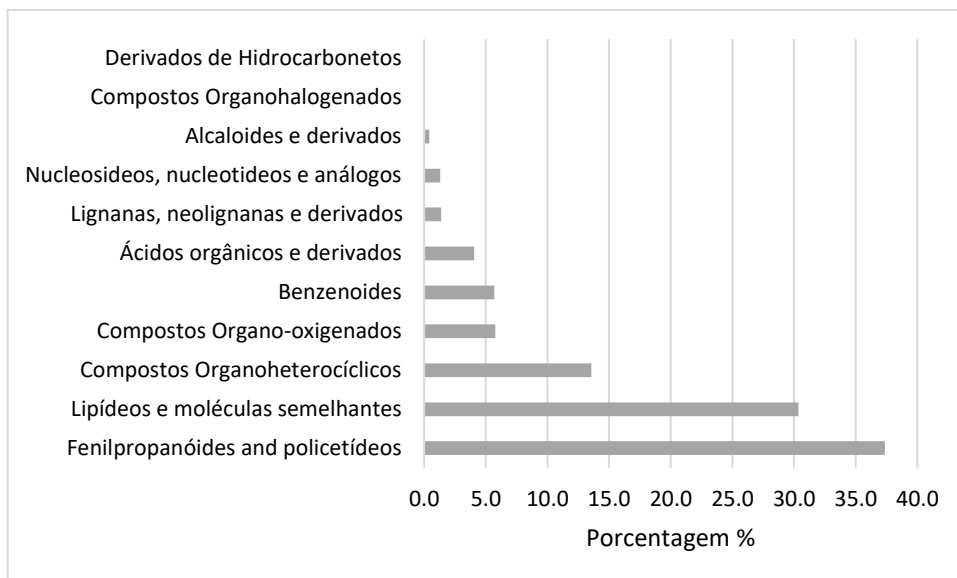


Gráfico 4. Superclasses dos compostos anotados na rede molecular de compostos de média-alta polaridade advindos dos extratos de Orchidaceae e analisados por LC-MS.

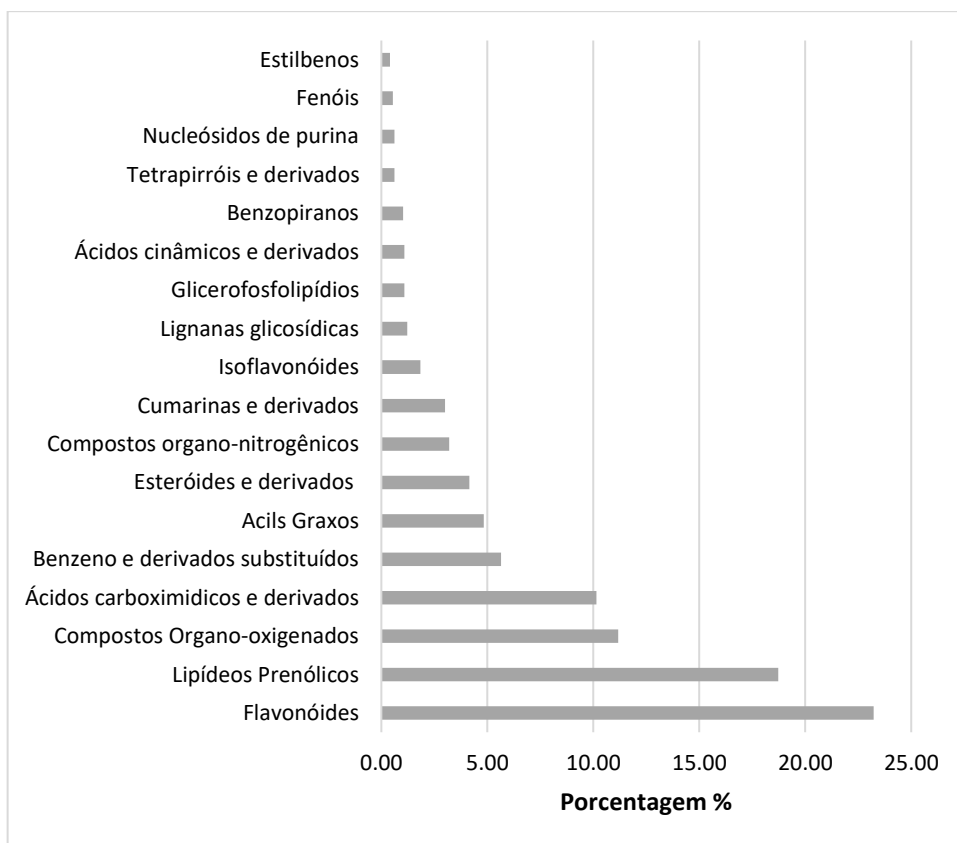


Gráfico 5. Dezoito classes mais abundantes dentre os compostos anotados na rede molecular de compostos de média-alta polaridade advindos dos extratos de Orchidaceae e analisados por LC-MS.

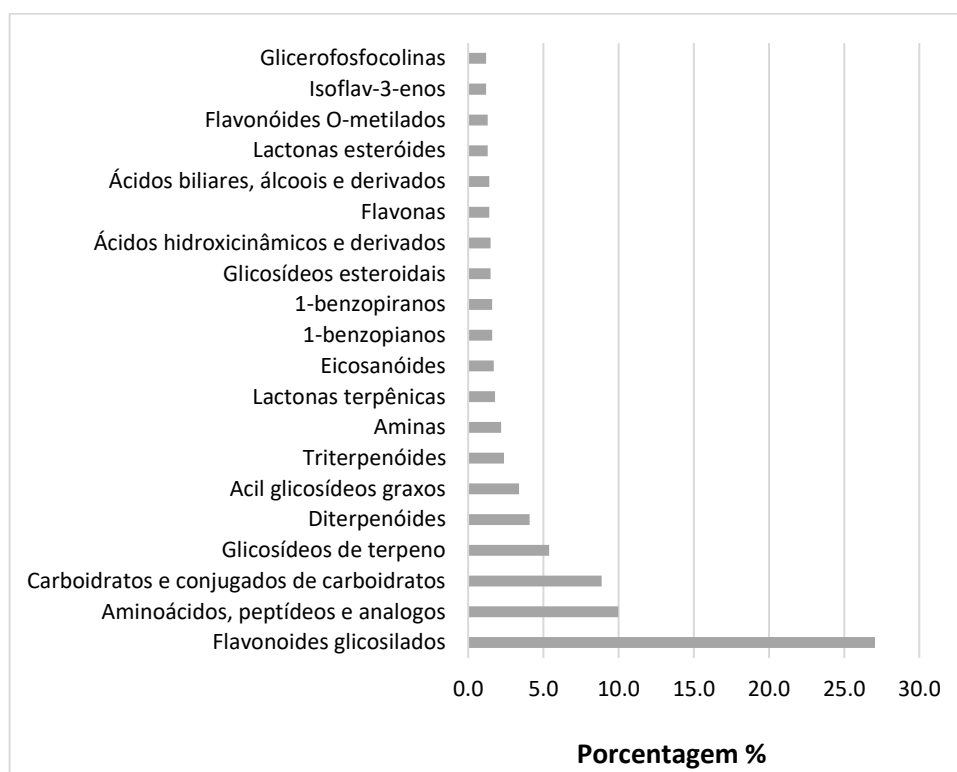


Gráfico 6. Dez subclasses mais abundantes dentre os compostos anotados na rede molecular de compostos de média-alta polaridade advindos dos extratos de Orchidaceae e analisados por LC-MS.

Dentre as classes anotadas, cinco se destacam por estarem presentes em quase todas as amostras: os flavonoides, lipídeos prenólicos (terpenos), esteroides e derivados e compostos organo-nitrogenados. Esses também estão entre as classes mais abundantes dentre os compostos anotados. Além disso, outras classes interessantes para a atividade biológica também foram anotadas, como cumarinas, derivados de ácido cinâmico e determinadas classes de alcaloides, como os derivados da vinca, emetina e alcaloides esteroidais.

Algumas das classes e subclasses anotadas para os compostos de média-alta polaridades também foram observadas nas análises dos compostos de média-baixa polaridade, como os terpenos e seus derivados, estilbenos, antraquinonas, chalconas e esteroides, em quantidades menores e em alguns casos glicosilados, como os terpe

A tabela 4 mostra a distribuição entre as espécies de algumas subclasses interessantes no estudo do potencial químico e biológico de Orchidaceae.

Tabela 4. Distribuição de algumas subclasses presentes nas amostras analisadas e de interesse farmacológico, de acordo com os compostos anotados nas redes moleculares advindos dos extratos de Orchidaceae analisados por LC-MS

Subclasse (ClassyFire)	%	Subclasse (ClassyFire)	%
Aminoácidos, peptídeos e análogos	100	Acil Graxos glicosídicos	38
Triterpenos	99	ésteres de ácido 2-benzimidazolilcarbâmico	35.6
Flavonóides O-metilados	98	alcalóides esteroidais	1.5
Glicosídeos esteróides	98	Glicosídeos esteroidais	9.8
Aminas	98	ácido hidroxicinâmico e derivados	9.1
Diterpenos	98	Flavonas	6.8
Clorinas	95	Flavonóides O-metilados	3.8
Carboidratos e conjugados	80	Isoflav-3-enos	9.8
Sesquiterpenos	77	Sesquiterpenos	18.9
Terpenos glicosídicos	76	Chalconas e dihidrochalconas	3.0
Ácidos carboximídicos	73	Monoterpenos	6.8
Flavanos	72	Isoflavonóides O-glicosilados	2.3
Flavonóides glicosilados	60	éster de ácido cinâmico	2.3
Éteres	57	Biflavonóides e poliflavonóides	1.5
Ácidos indolil carboxílicos e derivados	54	Cumarinas glicosídicas	1.5
lactonas esteroidais	50	Estilbenos glicosilados	2.3
lactonas terpênicas	42	Alcalóides esteroidais	1.5
Álcoois graxos	41	Quinolina carboxamidas	3.8

%. Porcentagem

Mais uma vez esta abordagem metabolômica permitiu a anotação de diferentes classes químicas que revelam as espécies brasileiras de Orchidaceae como uma fonte rica em diversidade química e de compostos de interesse farmacológico. Abaixo estão algumas dessas classes, assim como na figura 2.

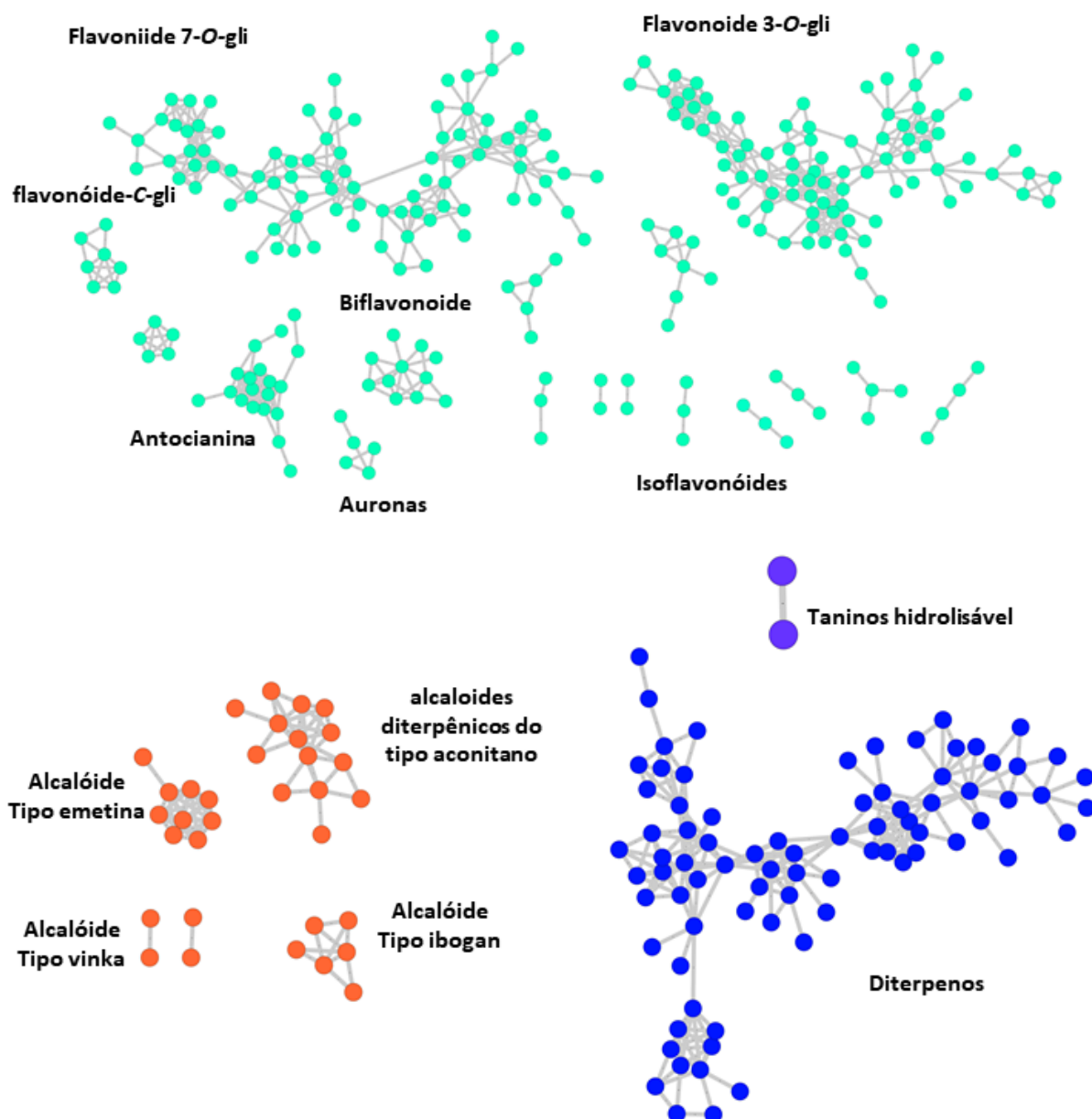


Figura 2. Algumas Famílias moleculares, relacionadas a atividades biológicas, da rede molecular advinda dos extratos de Orchidaceae analisados por LC-MS, coloridas de acordo com as classes químicas. O tamanho do cluster (nodo) está relacionado ao número de espectros.

- *Flavonóides*

Os flavonoides são uma classe importante de polifenóis, sintetizados a partir da via dos fenilpropanóides, além de ser um dos grupos fenólicos mais diversificados e importantes dentre os produtos naturais, e uma das mais abundantes classes que ocorrem no reino vegetal (WANG, 2021).

As funções atribuídas a eles são extensas, como a proteção contra danos causados por estresse bióticos e abióticos, incluindo a alimentação por insetos,

infecção por fungos, vírus e bactérias, e a irradiação UV e visível. Podem atuar como sinalizadores na atração de polinizadores, antioxidante, inibidores de enzimas, agentes alelopáticos e controle da ação de hormônios vegetais e participando do metabolismo da auxina (LIU et al., 2021). Além disso, são um dos principais responsáveis pelos pigmentos das plantas, através da presença das antocianinas (pigmentos vermelhos, laranja, azuis e roxos), chalconas e auronas (pigmentos amarelos), flavonóis e flavonas (pigmentos brancos e amarelo-claros), que conferem às plantas a grande variedade de e auxílio na atração de polinizadores (LIU et al., 2021)

Neste estudo foram anotados 341 compostos classificados como flavonoides, dentre eles as subclasses dos biflavonóides e poliflavonóides, flavanos, flavonas, flavanonas, flavonóis, flavonoides glicosilados (O-gli, como 7-O-gli, 3-O-gli; C-gli, cpmp 8-C-gli), flavonoides metilados (2'-O-Met, 7-O-Met, 4'-O-Met, 3'-O-Met, 6-O-Met), antocianinas (3-O-gli, 5-O-gli), chalconas, dihidrochalconas, auronas, isoflavonóides (isoflavonas, isoflav-3-ene, isoflav-3-enones, pterocarpanos, 2-arilbenzofurano, benzopiranos, isoflav-3-ene, isoflavonas, isoflav-3-enones e isoflavonoides O-gli, pterocarpanos. Os flavonoides estão presentes em 100% das amostras analisadas.

Os flavonoides são conhecidos na família Orchidaceae. William (1979) reportou em seu estudo de 142 espécies de 75 gêneros diferentes que as flavonas C-glicosídeos (presentes em 53% das espécies) foi encontrado como o componente mais comum entre todas as estudadas. Agora já são conhecidos mais de 23 flavonoides extraídos somente de folhas do gênero *Dendrobium* (Wang, 2021).

Esses também possuem importância farmacológica, atribuindo a alguns compostos dessa classe a ação antitumoral, anti-inflamatória, antioxidante, antiviral, antidiabética, imunomodulatória, hepatoprotetiva, entre outros (POLIMATI et al., 2022). Na ação contra o câncer, flavonoides podem modificar a produção hormonal e prevenir o crescimento de células cancerígenas. Além disso, há a hipótese de que polifenóis em extratos de orquídeas inibem as enzimas metabolizadoras de xenobióticos, alterando a ativação metabólica de potenciais carcinogênicos, inibindo as células cancerígenas (ŚLIWIŃSKI et al., 2022)

Os flavonóides, como as fitoalexinas ou antioxidantes, têm a função de eliminação de espécies reativas de oxigênio (ROS), auxiliando no tratamentos de

doenças crônicas inflamatórias, como artrite reumatoide e lúpus eritematoso(Liu et.al., 2021). Além disso, seu potencial antioxidante auxilia no retardo do envelhecimento do sistema nervoso, órgãos imunológicos, sistema reprodutivo, fígado e pele, e contribui para a prevenção da osteoporose, doenças cardiovasculares, doença de Alzheimer e câncer de mama (Liu et.al., 2021).

Flavonas, como apigenina, crisina, acacetina e gardenina-S, demonstraram possui ação antiperoxidase e, além disso, a negleteína, luteolina e apigenina mostraram atividade hepatoprotetora significativa em hepatócitos e ratos estimulados por IL-1, inibindo a produção de óxido nítrico em comparação com o padrão (Polimati et.al., 2022).

As antocianinas são os pigmentos mais comum entre as plantas, responsável pelas cores vermelho, roxo e azul. Além disso, possuem diversas funções ecológicas, fisiológicas e bioquímicas, como ação antimicrobiana e antifúngica, servir de sumidouro de carboidratos, ação antioxidante pela eliminação de espécies reativas de oxigênio (ROS) e quelação de metais (Hughes et.al, 2021).

As auronas fazem parte da família dos polifenóis, e de acordo com Mazziotti et.al. (2022), elas têm sido frequentemente negligenciadas pelos pesquisadores em comparação com os outros membros da superfamília dos flavonoides. Mas agora, devido ao seu potencial efetivo nos usos terapêuticos, ele começou a ser revelado. Portanto, eles representam uma classe de compostos naturais que vale a pena aprofundar para fornecer compostos bioativos. Em virtude de sua natureza poli-hidroxilada, são capazes de extinguir as espécies reativas de oxigênio (ERO), apresentando propriedades antioxidantes.

- *Taninos Hidrolisáveis*

Os taninos são compostos estrategicamente produzidos para a defesa contra agentes nocivos e a herbivoria, devido à sua capacidade de complexar proteínas de forma irreversível, sabor adstringente e deterrente. Dentre os taninos hidrolisáveis podemos encontrar galotaninos e elagitaninos (Maugeri et.al., 2022). Os taninos hidrolisáveis foram relatados como possuidores de várias atividades biológicas para a saúde humana, como a ação antioxidante, anti-inflamatória, anticâncer, antimicrobiana, antidiabética, atividade neuroprotetora, dentre outras.

- *Alcaloides*

Os alcaloides são classes estudadas profundamente na química de Orchidaceae, pois são utilizados por décadas na medicinal popular da China, principalmente utilizados os alcalóides advindos do gênero *Dendrobium*.

Entre os alcalóides, compostos com esqueletos diferentes foram relatados em Orchidaceae: sesquiterpenóides, indolizidinas, pirrolidinas, ftalidas e derivados de imidazol. Um exemplo são os alcalóides do tipo dendrobine: dendrobine, dendramina, nobilonina, dendroxina e 6-hidroxidendroxina (Sut et.al, 2017).

Neste estudo foram anotados 39 compostos classificados como alcalóides, dentre eles foram anotados os alcaloides esteroidais, alcaloides diterpênicos do tipo aconitano, alcaloides do tipo emetina, Ibogan e vinka, e 9b-azoaphenales.

Os alcaloides advindos de *Dendrobium* são compostos ativos com ação anti-inflamatória, antitumoral e antiviral, além de serem reguladores da produção de lipídeos pelo fígado, glicogênese e a proteção contra a hiperglicemia.

4.3. Padrões de Similaridade Química

A análise das redes moleculares utilizando o Cytoscape® para visualização das famílias moleculares e clusters permitiu observar que alguns estavam correlacionados a um táxon específico, conforme mostrado na Figura 3 e 4. Algumas famílias moleculares estão presentes em todas as subfamílias de Orchidaceae, indicando uma semelhança entre elas, mas houve outras famílias moleculares e alguns grupos de clusters que estão presentes em uma determinada subfamília ou tribo (Fig.2.A,B). Para avaliar se existem de fato padrões de similaridade ou dissimilaridade química entre espécies a nível de subfamília e tribo, foram realizadas análises estatísticas multivariadas não supervisionadas de PCA e HCA.

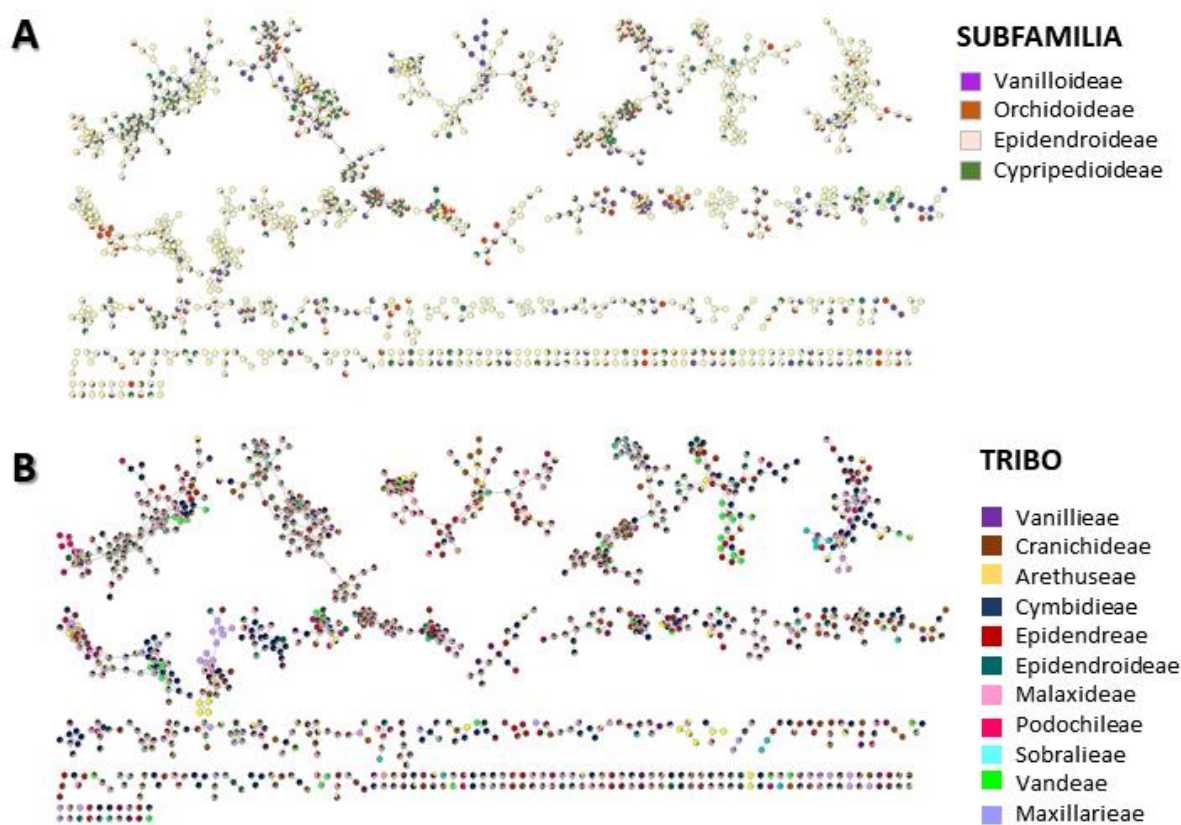


Figure 3. Rede molecular do metaboloma apolar (GC-MS) de espécies de Orchidaceae, coloridas de acordo com a subfamília (A) e tribo (B) observando os padrões de correlação com a classificação taxonômica e filogenética, segundo Givnish et al. (2015 e 2016) e Chase, et al. (2015).

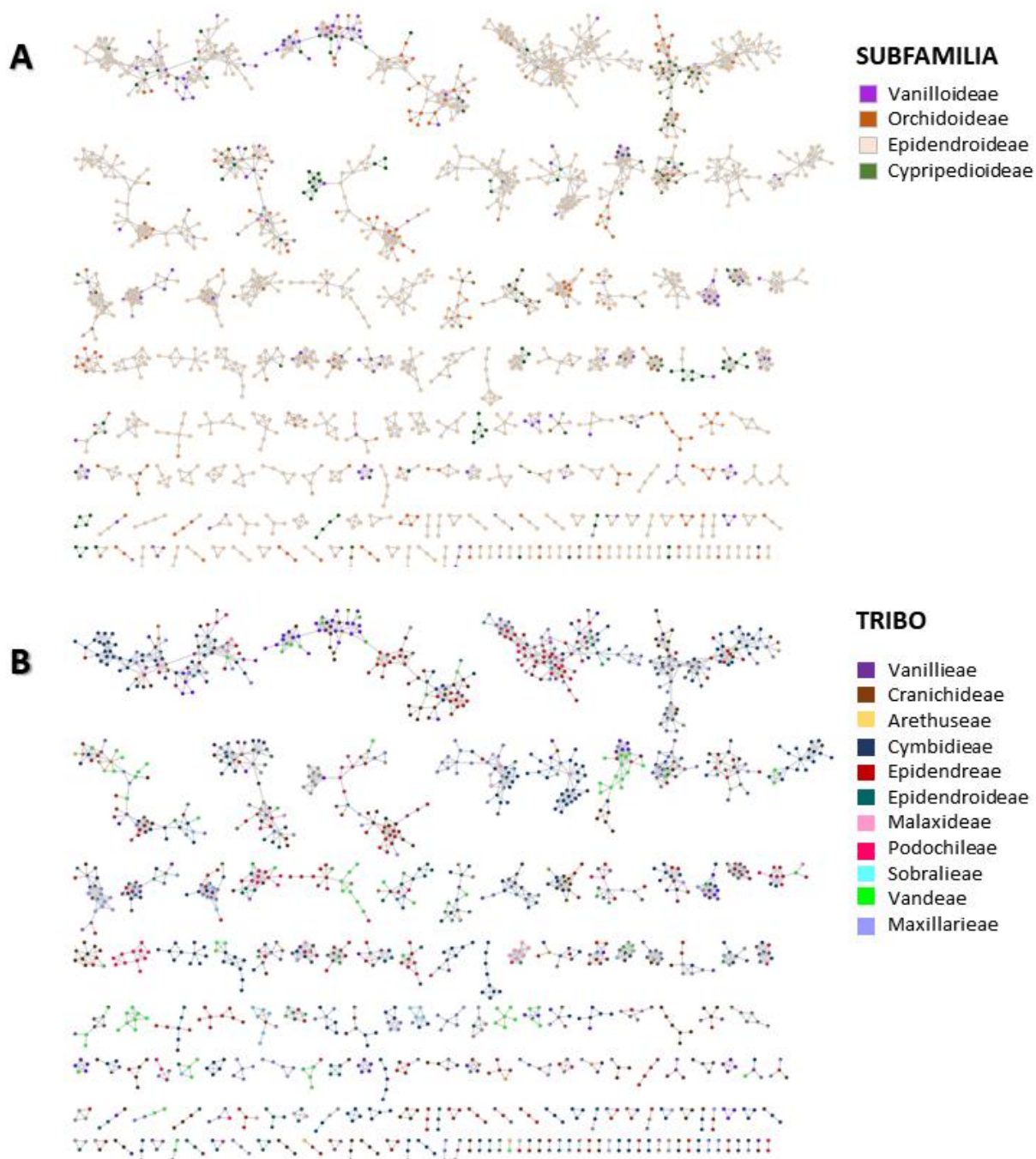


Figure 4. Rede molecular do metaboloma semi-polar (LC-MS, modo positivo) de espécies de Orchidaceae, coloridas de acordo com a subfamília (A) e tribo (B) observando os padrões de correlação com a classificação taxonômica e filogenética, segundo Givnish et al. (2015 e 2016) e Chase, et al. (2015).

No PCA, não foram observados padrões de agrupamento definidos em relação às subfamílias ou tribo de Orchidaceae (Fig. 5 e 6), mas foi observada uma tendência de dissimilaridade entre as duas maiores tribos da subfamília Epidendroideae, as Cymbidieae e Epidendreae tribos (Fig. 6). Essa tendência é observada nas análises utilizando tanto os dados do metaboloma apolar e semi-polar separadamente (A e B) quanto juntos (C). Quando utilizado os dados com conjunto (C), foi observada uma melhor distribuição das espécies dentro do respectivo táxon. A análise das tendências de similaridade realizadas em diferentes perspectivas químicas é importante, pois há a possibilidade de a similaridade ou dissimilaridade estar presente predominantemente em um conjunto de metabólitos do que em outro.

Através dos HCAs foi possível obter uma visão geral da formação de alguns agrupamentos entre subfamílias e tribos, especialmente a dissimilaridade observada no PCA entre as espécies das tribos Cymbidieae (azul) e Epidendreae (vermelho) (Fig. 7e 8), mas não houve a formação estruturada de um padrão de agrupamentos de acordo com cada táxon.

Estudos de similaridade química também são uma grande ferramenta para afinar a seleção de espécies e amostras para estudos de atividade. Utilizando análises como HCA, é possível selecionar somente as espécies mais próximas hierarquicamente de uma que já possui sua atividade conhecida. Um exemplo neste estudo seria avaliar as espécies que possuem maior similaridade química com espécies do gênero *Dendrobium*, utilizadas como plantas medicinais na medicina popular chinesa e detentora dos alcalóides do tipo dendrobina, além de possuir produtos comercializados com o pó da planta, extrato ou fração alcaloídica (Dan Yang *et. al.*, 2018).

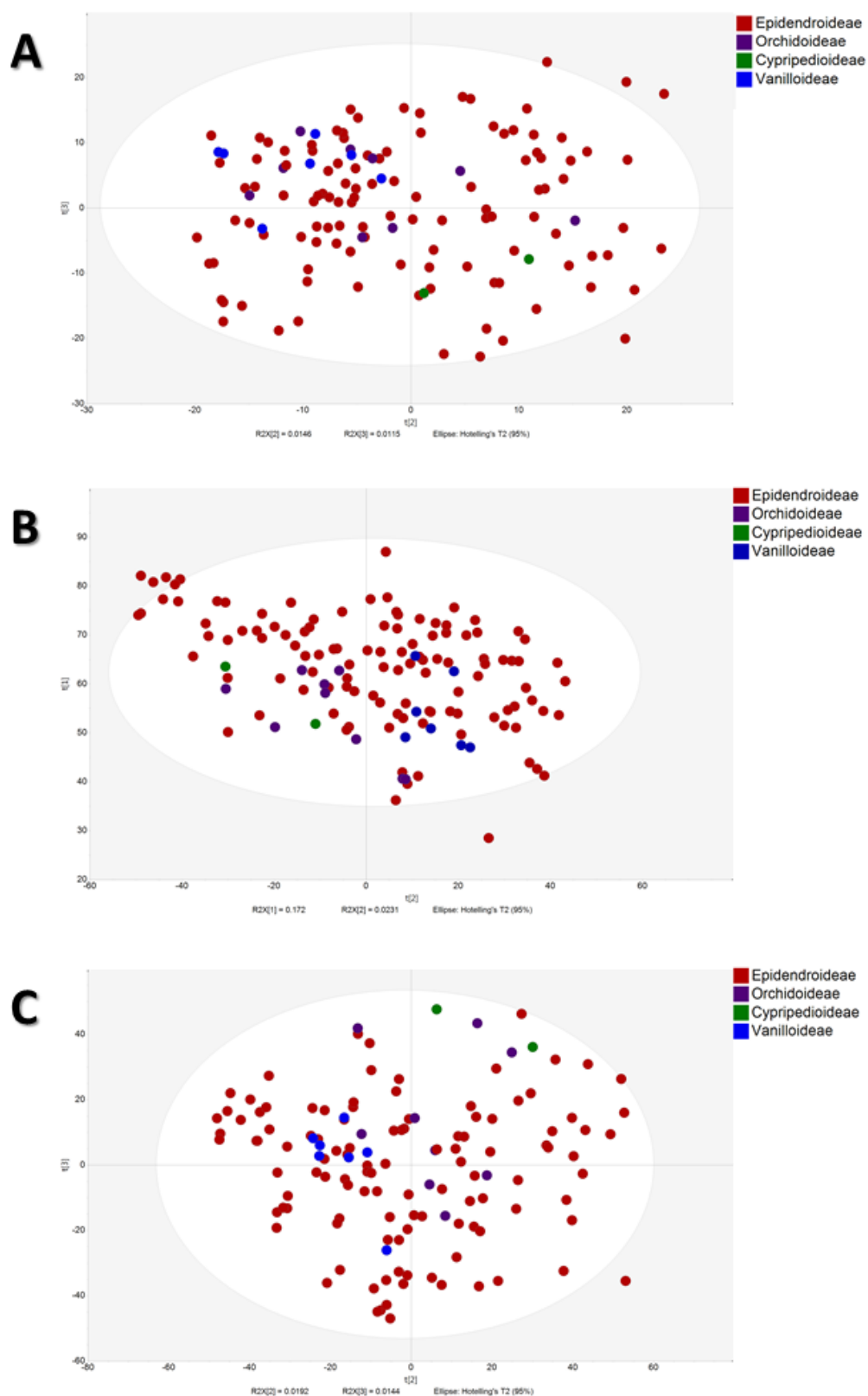


Figura 5. Análises de Componentes Principais (PCA) do Metaboloma apolar **(A)**, semi-polar **(B)** e da combinação dos dois metabolomas **(C)** das espécies de Orchidaceae, observando tendências de agrupamento de acordo com a subfamília. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015).

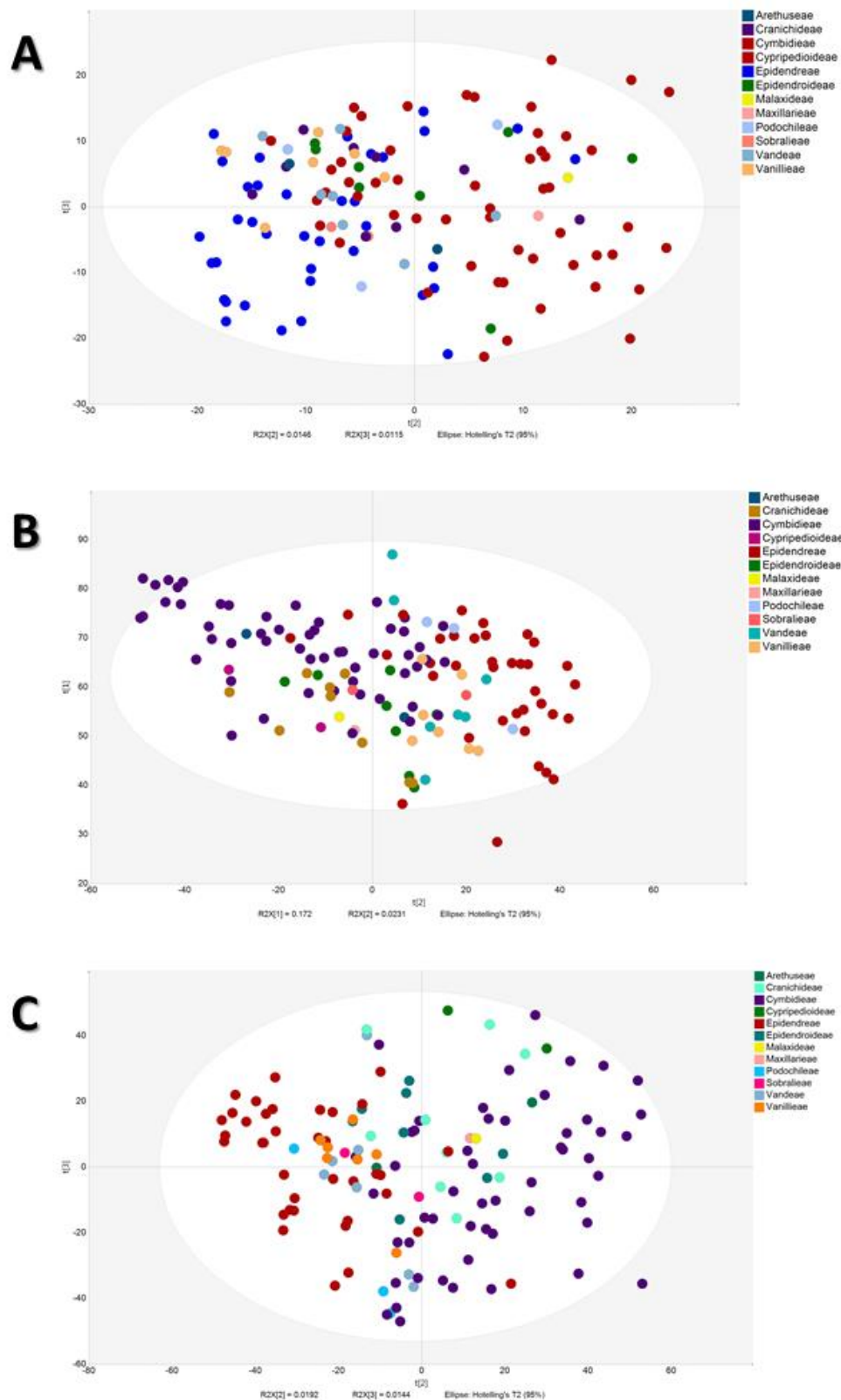


Figura 6. Análises de Componentes Principais (PCA) do Metaboloma apolar **(A)**, semi-polar **(B)** e da combinação dos dois metabolomas **(C)** das espécies de Orchidaceae, observando tendências de agrupamento de acordo com a tribo. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015).

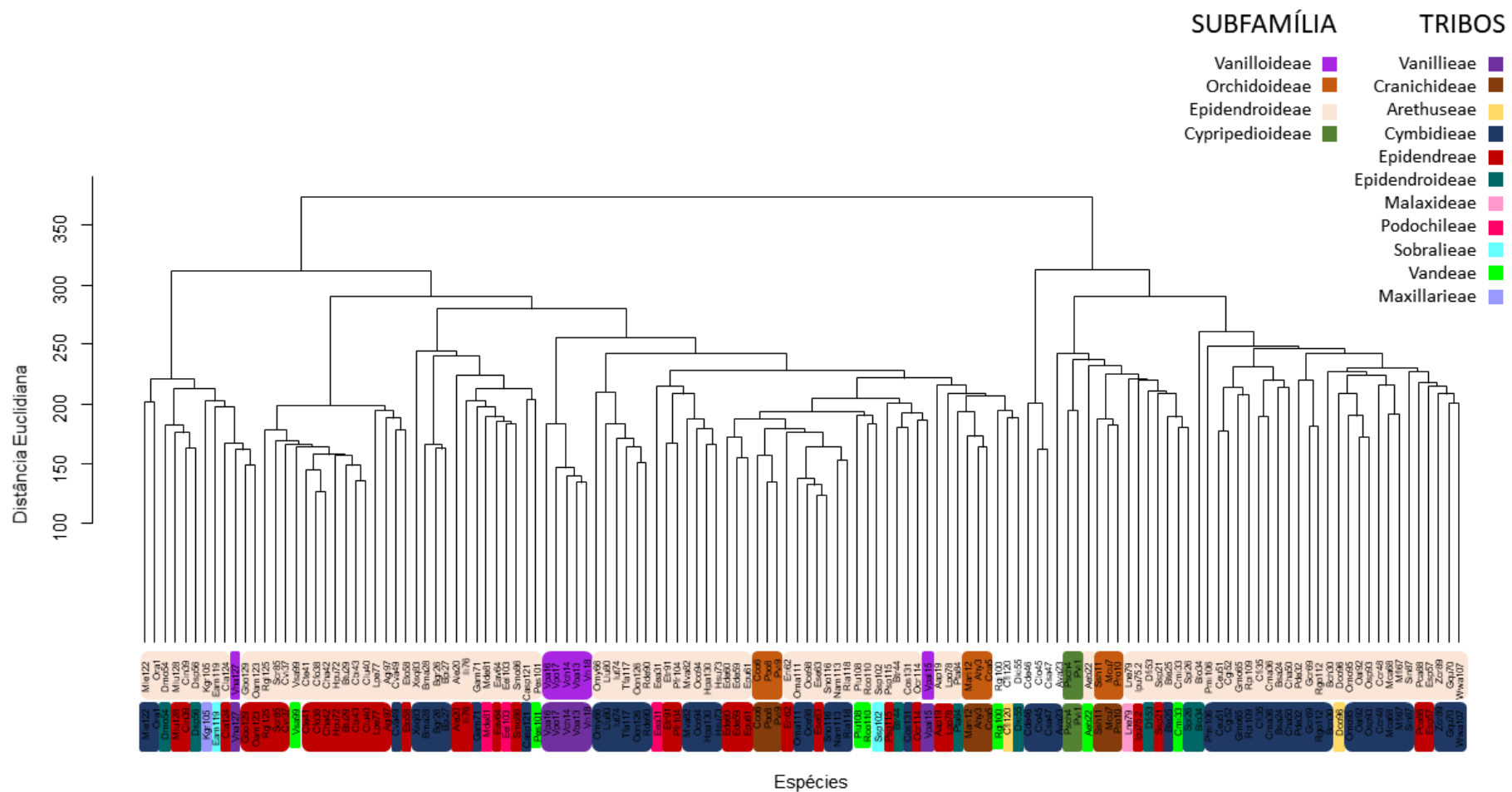


Figura 7. Análise de Agrupamentos Hierárquicos (HCA) do metaboloma semi-polar das espécies de Orchidaceae, adquirido por LC-MS no modo de ionização positivo e negativo, observando tendências de agrupamento de acordo com a subfamília e tribo. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015).

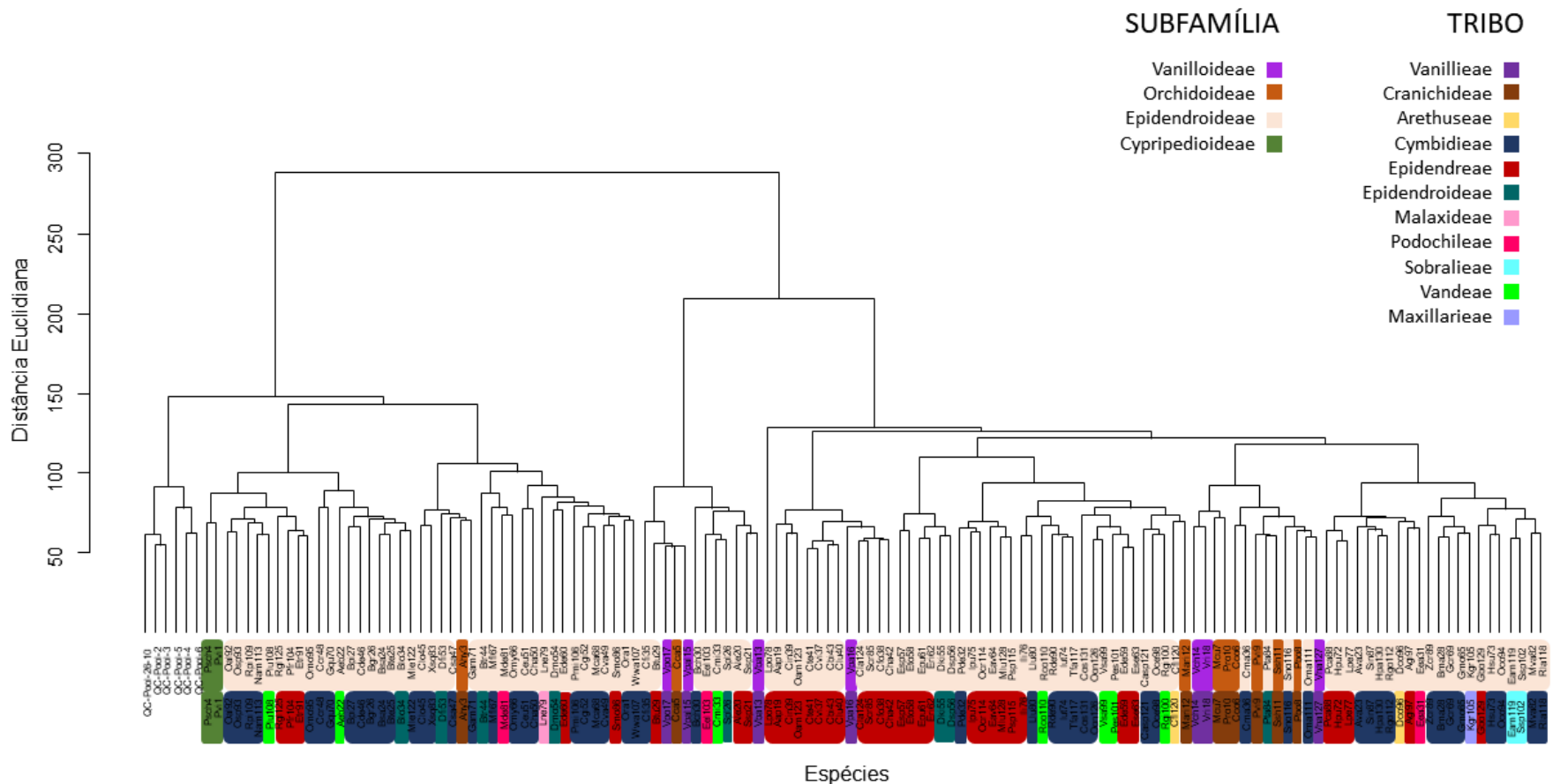


Figura 8. Análise de Agrupamentos Hierárquicos (HCA) do metaboloma apolar das espécies de Orchidaceae, adquirido por GC-MS, observando tendências de agrupamento de acordo com a subfamília e tribo. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015).

Devido à tendência observada de alguns agrupamentos à nível de subfamília e tribo, especificamente entre as tribos Cymbidieae e Epidendreae, foram realizadas análises de enriquecimento de classe com o objetivo de observar quais classes de compostos, dentre os anotados, estavam enriquecidas em determinado táxon, em relação aos outros.

Dentre as subfamílias, a figura 9 mostra uma comparação entre as subfamílias Orchidoideae, Vanilloideae e Epidendroideae, demonstrando a diferença entre elas perante algumas das classes anotadas. A subfamília Cypripedioideae não participou desta análise por não possuir número de amostras suficiente (mínimo de 3 amostras). A escala de calor demonstra que a classe está enriquecida quanto mais intensa a cor vermelha, ou que possui níveis similares dentro das subfamílias quanto mais esbranquiçada, ou que possui déficit dessa classe quanto mais intensa a cor azul.

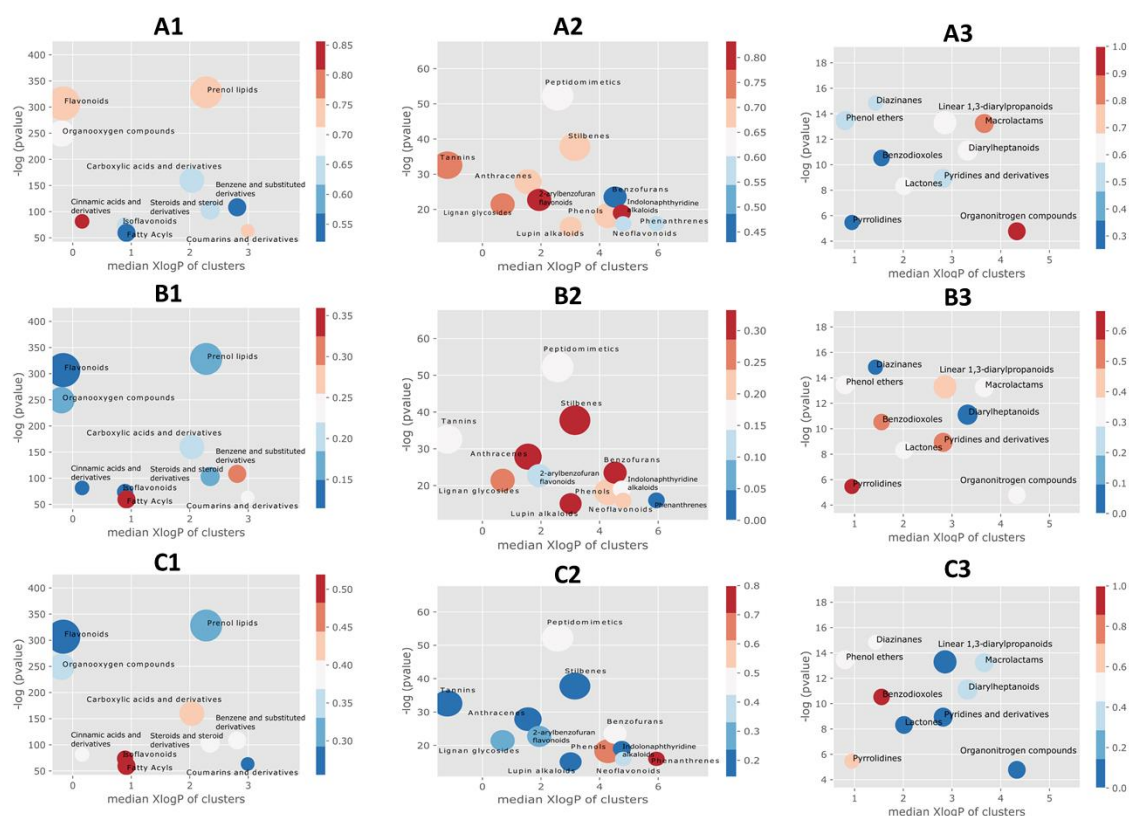


Figura 9. Análise de Enriquecimento de Classes presentes nas subfamílias das espécies de Orchidaceae estudadas, destacando as 10 classes dentre as subfamílias. **A.** Epidendroideae. **B.** Orchidoideae. **C.** Vanilloideae.. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015).

O enriquecimento de classe à nível de subfamília permitiu observar que a família Epidendroideae possui enriquecidas, em relação às outras subfamílias, as classes dos terpenos, flavonoides, ácido cinâmico e derivados, taninos, alcalóides indolonaftiridinos e compostos organonitrogenados. Já a Orchidoideae possui enriquecidas as classes dos benzenos e derivados substituídos, estilbenos, antracenos, benzofurandos e alcalóides lupínicos. As Vanilloideae possuem enriquecidas as classes dos Acil graxos, isoflavonóides, fenantrenos e benzodioxoles.

O enriquecimento de classes é uma ferramenta muito interessante para auxiliar no afinilamento da (i) seleção das espécies a serem utilizadas em estudos de atividade biológica, ecologia química, entre outros, selecionando s que possui os compostos de interesse (estudo direcionado), (ii) seleção da atividade biológica a ser estudada, de acordo com a classe de metabólitos enriquecidas e seu histórico de atividade, e (ii) direcionar estudos de quimiotaxonomia à níveis de classe química. A seguir está a tabela das classes enriquecidas de acordo com cada tribos de Orchidaceae.

A figura 10 mostra uma consideração mais detalhada sobre às tribos Cymbidieae e Epidendreae em relação às algumas classes metabólicas enriquecidas, em comparação com todas as tribos estudadas. Através do enriquecimento de classe é possível perceber o quão diferentes essas tribos são, mesmo que o enriquecimento seja referente à todas as tribos do estudo. Cymbidieae possui enriquecidas, em relação à Epidendreae, as classes dos esteroides e derivados, piridinas e derivados, compostos organo-nitrogenados, diazinas e antracenos. A tribo Epidendreae possui enriquecidas, comparando com a tribo Cymbidieae, as subclasses bezenos e derivados substituídos, compostos peptidomiméticos, pirrolodinas, alcalóides do tipo indolonaftiridino e neoflavonóides .

As espécies fornecidas para este estudo são cultivadas no orquidário LBMBP-USP, em casa de vegetação com intensidade de luz, irrigação, umidade e contato com polinizadores e predadores controlados. Assim, todas foram cultivadas nas mesmas condições e expostas aos mesmos fatores bióticos e abióticos, fazendo com que a resposta a esses fatores seja a mesma para todas, e dependendo do tipo rotas metabólicas e maquinarias enzimáticas expressas por cada genótipo.

Através do estudo dos padrões de correlação observados entre compostos e classes taxonômicas e filogenéticas, foi possível confirmar que há tendências de

agrupamentos relacionados à similaridade química que seguem a filogenia e taxonomia mais recentes (Givnish et al., 2015 e 2016; Chase, et al., 2015), chegando em níveis de agrupamentos de acordo com a subfamília e tribo. Assim, este estudo mostra que é possível realizar estudos quimiotaconômicos e aprofundamentos em filogenia das espécies através do estudo do metaboloma. Para fins de aprofundamento em estudos quimiotaconômicos, o desenvolvimento um estudo utilizando uma amostragem maior de cada subtribo, com réplicas de indivíduos e de populações diferentes, seria muito interessante.

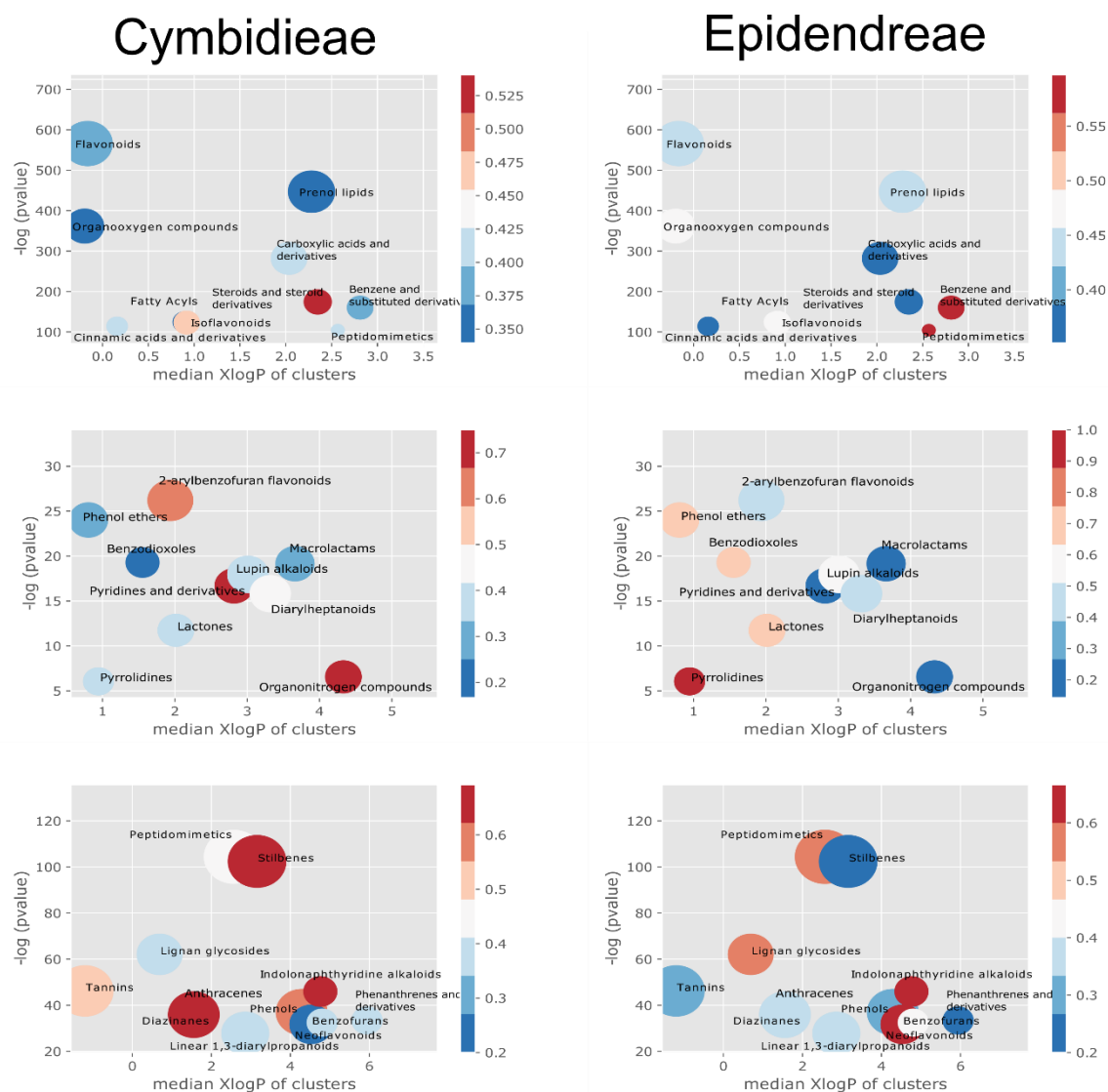


Figura 10. Análise de Enriquecimento de Classes presentes nas tribos Cymbidieae e Epidendreae, avaliando a diferença entre elas à nível de classe. Classificação taxonômica e filogenética de acordo com Givnish et al. (2015 e 2016) e Chase, et al. (2015).

5. CONCLUSÕES

O uso de ferramentas metabolômicas e bioinformáticas tornou mais ágil o acesso ao conhecimento sobre os metabólitos secundários de mais de 100 espécies de Orchidaceae brasileiras, sendo este o primeiro estudo que fornece uma visão geral dos compostos das orquídeas brasileiras e a base para futuros estudos.

Neste estudo foram anotados mais de 1700 compostos, dentre compostos de média-baixa e média-alta polaridades. A grande maioria das classes observadas neste estudo são inéditas para espécies de orquídeas brasileiras, e algumas classes são inéditas para a família Orchidaceae. Dentre as anotações obtidas, algumas classes possuem um histórico muito interessante sobre seu potencial ativo e utilização farmacológica.

Assim, além da grande diversidade ecológica, da estrutura floral e do uso ornamental desta família, existe uma grande diversidade química entre as espécies brasileiras que nos traz a possibilidade de potenciais alvos na busca de compostos ativos.

Além disso, as redes moleculares mostraram algumas correlações entre clusters e famílias moleculares e certos táxons da família, confirmados por análises de HCA e PCA, especialmente a dissimilaridade entre as tribos Epidendreae e Cymbideae, duas das maiores tribos de Orchidaceae brasileiras. Análises de enriquecimento de classe foram realizadas, possibilitando a compreensão de diferenças entre subfamílias e as entre as duas tribos citadas.

CAPÍTULO 3

1. INTRODUÇÃO

O uso medicinal de espécies de Orquidaceae datam de séculos, principalmente na Medicina Tradicional Chinesa e na Medicina Ayurvédica, além da utilização dessas na medicina popular do Sudeste Asiático, Japão, Europa, África e Américas (MUSHAROF HOSSAIN, 2011).

O estudo etnofarmacológico realizado por Hajdu e Hohmann (2012) na comunidade indígena Porvenir, na amazônia boliviana, cita a utilização dos frutos de espécies do gênero *Vanilla* para furúnculos (inflamação de folículos capilares), diarreia, dores de estômago e de garganta. Boscolo e Valle (2008) observaram que orquídeas como *Cyrtopodium paranaense* Schltr. (“Sumaré”) e *Vanilla sp.* (baunilha) são utilizadas na cicatrização de feridas, sendo o pseudobulbo e as folhas as partes utilizadas na medicina popular. As orquídeas são amplamente utilizadas na medicina tradicional chinesa, como a espécie *Gastrodia elata*, *Bletilla striata* e diferentes espécies do gênero *Dendrobium*, sendo cultivadas e comercializadas na China como parte de uma grande indústria da medicina herbal na fitoterapia do país (Bulpitt *et al.*, 2007). O estudo de Zheng *et al.* (2013) na Ilha Hainan (China) mostrou a utilização de *Tropidia curculigoides* Lind. para diabetes a partir da ingestão do decocto.

Os pseudobulbos de *Catasetum Macroglossum* são popularmente utilizados como anti-inflamatório e antirreumático nas florestas e terras médias do Equador. Em Ramos *et al.* (2012), o potencial anti-inflamatório da espécie foi confirmado através de ensaios de indução de edema de pata por carragenina, em comparação com controles indometacina e salina. A atividade, de acordo com estudos fitoquímicos, pode ser devido à presença de um glucomanano e vestígios de fenantreno e estilbeno, tal como em outras espécies de *Catasetum*. Já o *Cyrtopodium punctatum* é utilizado na medicina tradicional no Brasil para tratamento de inflamações, abscessos e queimaduras, mostrando propriedades curativas e anti-hemorrágicas. Parente *et al.* (2014) demonstrou que um dos componentes do extrato dessa espécie, o polissacarídeo *andersonii* glucomanano, interfere no mecanismo ulcerogênico, mostrando propriedade citoprotetora, e possui potencial anti-inflamatório, justificando o uso na medicina tradicional.

Um exemplo de gênero altamente estudado é o *Dendrobium*, cujas espécies são muito utilizadas na medicina popular chinesa, sendo caracterizado pela presença

de sesquiterpenos, alcalóides, compostos fenólicos, polissacarídeos bioativos, e mais de 19 tipos de alcalóides denominados dendrobina (Dan Yang *et al.*, 2018). Os estudos de Dan Yang *et al.* (2018) identificaram quatro novos alcalóides (finflaynes A-D) e quatro novos compostos fenólicos (dendrofindlaphenoil A-C e 6"-de-O-methildendrofindlafenol A) no extrato etanólico (Etanol/Água 95:5) do caule de *Dendrobium findlayanum*, além de outras substâncias como dendrobina, 2-hidroxi-dendrobina e 6"-de-O-methildendrofindlaphenol A. Os compostos fenólicos isolados demonstraram atividade citotóxica contra linhagens celulares de câncer, diminuíram a produção de NO e demonstraram a propriedade de eliminação de radicais livres.

Mesmo com todo o potencial ativo e diversidade química da família Orchidaceae sendo apresentado por estudos químicos, etnofarmacológicos e estudos de atividade biológica, as orquídeas brasileiras são pouco estudadas, o que desperta interesse em seus metabólitos e potencial biológico.

A família Orquidaceae possui, devido ao seu uso medicinal e curiosas interações ecológicas, estudos químicos que apontara diversas classes e subclasses com grande potencial antioxidante, anti-inflamatória e de diminuição de espécies reativas de oxigênio (ROS), como os flavonoides, xalconas, auronas, taninos hidrolisáveis, dentre outros (CHAN *et al.*, 2016; HUGHES *et al.*, 2021; LIU *et al.*, 2021; SEMWAL *et al.*, 2021). Assim, se tornam excelentes alvos de estudos para doenças autoimunes que possui como uma linha de tratamento eliminar a produção de ERO e, se possível, diminuir a função do neutrófilo. Uma dessas doenças é o lúpus eritematoso (LES).

O lúpus eritematoso sistêmico (LES) é uma doença autoimune multifatorial que acomete múltiplos órgãos e ocorre, principalmente, em mulheres. O desenvolvimento do LES é atribuído a rupturas na imunidade adaptativa, desencadeadas por fatores genéticos predisponentes e vários fatores ambientais, que juntos contribuem para a perda de tolerância aos auto antígenos. A doença começa com uma fase pré-clínica caracterizada pela produção desses auto anticorpos e prossegue com uma fase autoimune onde os sinais clínicos específicos da doença se tornam aparentes (TOLLER-KAWAHISA *et al.*, 2015).

A progressão do LES é caracterizada por dano tecidual promovendo a deposição de imunocomplexos e ativação do sistema complemento, com posterior

recrutamento de neutrófilos. Assim, os neutrófilos englobam e digerem os imunocomplexos, o que resultam na produção de espécies de oxigênio (ROS) e liberação de moléculas citotóxicas.

Além do recrutamento exacerbado de neutrófilos para destruir os imunocomplexos resultando em um aumento de ERO no tecido, os neutrófilos de pacientes com LES exibem alterações marcantes em suas funções-resposta, por exemplo produção desequilibrada de ERO, quimiotaxia e fagocitose que podem contribuir para a fisiopatologia da doença e a lesões em vários órgãos, em particularmente no rim, sistema vascular e pele. A responsividade alterada dos neutrófilos, associada à depuração ineficiente de ERO pelo sistema de defesa antioxidante, agrava o dano tecidual oxidativo em alguns pacientes com LES (TOLLER-KAWAHISA et al., 2015).

A análise metabolômica, a qual inclui a metabolômica de plantas (*Plant Metabolomics*), é uma ferramenta inovadora e valiosa para a busca de metabólitos a partir de extratos complexos de plantas, oferecendo oportunidades para a descoberta de novos fármacos a partir de produtos naturais (CHAGAS-PAULA et al., 2015; YULIANA et al., 2011). É uma abordagem rápida, eficaz e otimizada na investigação qualitativa e quantitativa de metabólitos, a qual permitir a correlação dos metabólitos com propriedades atribuídas às amostras (PILON et al., 2020).

A identificação de metabólitos secundários em estudos de produtos naturais é um processo manual, trabalhoso, complexo e demorado, que demanda tempo e experiência do espectroscópio, para ser utilizado em estudos que envolvem grandes conjuntos de dados (Big data) e análise de grandes quantidades de dados. A análise metabolômica utiliza a desrepliação como ferramenta de anotação molecular, uma caracterização rápida de compostos a fim de anotar aqueles já relatados na literatura, evitando assim o isolamento de produtos naturais conhecidos (PILON et al., 2020; CARNEVALE et al., 2016).

2. OBJETIVOS

Este capítulo teve como objetivos principais avaliar os extratos hidrometanólicos de 128 espécies de Orchidaceae, dentre elas brasileiras e exóticas, quanto a sua atividade antioxidante, citotóxica e moduladora da produção de espécies

reativas de oxigênio por neutrófilos humanos e, através de análises metabolômicas e bioinformáticas, propor as classes de metabólitos secundários e substâncias responsáveis pela atividade.

Os objetivos específicos foram:

- a. Coleta das espécies brasileiras e exóticas de Orchidaceae;
- b. Obtenção dos extratos de média-alta polaridade;
- c. Avaliar o potencial antioxidante via ensaios com DPPH;
- d. Avaliar a citotoxicidade dos extratos;
- e. Avaliar o potencial modulador dos extratos quanto a produção de espécies reativas de oxigênio por neutrófilos humanos;
- f. Análise dos extratos por UFLC-DAD-ESI-MS/MS (micrOTOF-QII);
- g. Processamento das impressões digitais metabólicas;
- h. Construção das redes moleculares pela plataforma GNPS, utilizando as ferramentas NAP, MS2LDA e MolNetEnhancer;
- i. Análises de mineração de dados e correlação *in silico* entre impressões digitais metabólicas e as atividades, visando propor classes e substâncias bioativas.

3. PROCEDIMENTOS EXPERIMENTAIS

3.1. Material Vegetal

As 130 espécies de Orchidaceae foram fornecidas pelo Prof. Emerson Ricardo Pasarin, coordenador e responsável pelo orquidário do Laboratório de Biologia Molecular e Biosistemática de Plantas (LBMBP-USP) da Faculdade de Filosofia, Ciências e Letras de Ribeirão Preto (FFCLRP-USP).

Foram utilizadas as partes aéreas (folhas e pseudobulbos), que se apresentaram saudáveis e completamente desenvolvidas. As espécies foram coletadas e transferidas para tubos *falcon* de 50 mL, mergulhadas em nitrogênio líquido e liofilizadas a -90°C até eliminação total de água. Após esse processo, as amostras foram trituradas e o tamanho de partículas padronizado em tamiz de malha granulométrica 45 mesh. Finalmente o material vegetal foi armazenado em congelador -26 °C em frascos de vidro vedados.

552. Processo de Extração

As amostras foram preparadas utilizando 10,0 mg de droga vegetal pulverizada e extraídas com 1,0 mL de solução MeOH:H₂O (7:3, v/v) em banho de ultrassom por

10 minutos, seguida de centrifugação à 15000 rpm (26 °C) por 10 minutos, filtração em filtro de membrana PTFE (0,22 µm) e secagem em centrífuga à vácuo à temperatura ambiente, com posterior pesagem para cálculo de rendimento em massa. Os extratos secos foram acondicionados em freezer à -26°C até o momento da análise.

3.2. Avaliação do Potencial Bioativo dos Extratos

Os ensaios para a avaliação da atividade antioxidante, atividade citotóxica e do efeito dos extratos sobre a produção de ERO pelos neutrófilos humanos foram realizados no Laboratório de Imunologia e Citologia dos Fluidos Biológicos em colaboração com a Profa. Dra. Cleni M. Marzocchi Machado do Departamento de Análises Clínicas, Toxicológicas e Bromatológicas (Faculdade de Ciências Farmacêuticas de Ribeirão Preto, USP), pelas alunas de doutorado Camila Capel Godinho e Nathália Cristina Canicoba.

3.2.1. Obtenção de Sangue

Este protocolo de pesquisa foi aprovado pelo Comitês de Ética em Pesquisa da FCFRP-USP de acordo com a resolução 466/12 do Conselho Nacional de Saúde, Brasil (Anexos, Figura 1).

Os participantes da pesquisa foram recrutados na comunidade do campus USP Ribeirão Preto, como estudantes de graduação, pós-graduação, funcionários e docentes. A faixa etária determinada é entre 18 e 50 anos, sendo dez o número de voluntários. O número de participantes foi calculado baseado em estudos anteriores do grupo de pesquisa sob orientação de estatístico.

Todos estes participantes foram convidados pessoalmente pelas pós-graduandas e pesquisadoras responsáveis, Camila Capel Godinho e Nathália Cristina Canicoba, farmacêutica e biomédica, respectivamente, e doutorandas nos programas de Ciências Farmacêuticas e Biociências e Biotecnologia, da Faculdade de Ciências Farmacêuticas de Ribeirão Preto (FCFRP-USP). Elas foram responsáveis pelo recrutamento dos participantes, apresentação do Termo de Esclarecimento Livre e Esclarecido (TCLE), colheita e processamento das amostras biológicas doadas pelos participantes.

Foram colhidos 20mL (2 colheres de sopa) de sangue venoso periférico, para

a obtenção dos neutrófilos humanos. O sangue é colhido com sistema de coleta a vácuo e estéril, e em tubos sem anticoagulante. O critério de exclusão geral é tabagismo, uso de medicamentos controlados, gravidez e o uso de anti-inflamatórios há menos de sete dias.

As amostras de sangue foram processadas no dia da colheita e os neutrófilos obtidos utilizados no mesmo dia. Não houve armazenamento de material biológico de nenhuma das amostras colhidas. Ao final do dia que ocorreu a colheita do material, o material remanescente foi descartado. Não houve armazenamento de qualquer material biológico durante o período da pesquisa e tão pouco após o término dela.

3.2.2. Purificação e Obtenção dos Neutrófilos

Os neutrófilos foram obtidos a partir das amostras de sangue total, anticoagulado, por método de diferença de densidade segundo metodologia descrita por Lucisano e Mantovani (1984).

As amostras de sangue total de cada doador foram misturadas ao anticoagulante Alsever e submetidas à centrifugação a 664xg por 10 minutos em temperatura ambiente (centrífuga Eppendorf, modelo 5810R). O plasma foi desprezado juntamente com a camada de células mononucleares. Os neutrófilos foram separados das outras células pela adição de solução de gelatina 2,5% em NaCl 0,15 M em quantidade equivalente a duas vezes o volume de células e as amostras foram mantidas em banho-maria a 37°C por 15 minutos. Ao sobrenadante, o qual continha os neutrófilos, foi adicionado NaCl 0,15 M e as amostras foram centrifugadas a 526xg por 10 minutos em temperatura ambiente. As células foram lavadas com NH₄Cl 0,83% pH 7,2 para promover a lise dos eritrócitos remanescentes. A mistura foi mantida em banho-maria por 5 minutos a 37°C e, então, centrifugada a 562xg por 10 minutos em temperatura ambiente. O precipitado foi lavado com NaCl 0,15 M e submetido a 562xg por 10 minutos. As células foram ressuspensas em 1 mL de solução de Hanks pH 7,2 contendo gelatina 0,1%.

Para calcular a concentração de neutrófilos obtidos na purificação, as células foram diluídas em líquido de Turk e os neutrófilos foram contados em câmara de Neubauer. A viabilidade celular foi analisada pela diluição de uma parte da suspensão de neutrófilos em solução de Azul de Tripán 0,4% e os resultados foram sempre maiores que 95% de neutrófilos de pureza e de viabilidade.

3.2.3. Avaliação da Atividade Antioxidante

Para avaliar se os extratos de orquídeas apresentavam atividade antioxidante, foi realizado o teste com a substância 2,2-difenil-1-picril-hidrazil (DPPH), de acordo com Oliveira *et al.* (2014) Fernandes *et al.* (2014). Em uma cubeta foram adicionados 800µL do diluente (uma mistura de solução de tampão de acetato 0,1mol/L e etanol) e 200µL de uma solução de DPPH 0,5mol/L. A mistura foi lida em espectrofotômetro no comprimento de onda de 510nm. Em seguida, foram adicionados 5µL do extrato, a mistura foi homogeneizada e incubada em temperatura ambiente por 5 minutos no espectrofotômetro, em seguida feita a leitura no mesmo comprimento de onda. Ao final foi calculada a porcentagem final de DPPH reduzido utilizando-se a seguinte fórmula:

$$DPPH_{reduzido}(\%) = \frac{Absorvância_{inicial} - Absorvância_{final}}{Absorvância_{inicial}} \times 100$$

3.2.4. Avaliação da citotoxicidade dos extratos sobre os neutrófilos

Os neutrófilos tratados com os diferentes extratos de orquídeas foram incubados por 30min em banho-maria a 37°C e, posteriormente centrifugados a 970 xg por 10min a 4°C. O sobrenadante foi descartado e ao *pellet* será adicionado 100µL Hanks/gelatina 0,1% pH 7,2 e 100µL de azul de Tripán 0,4%. Para a avaliação da citotoxicidade, 200 neutrófilos humanos foram contados em microscopia ótica, sendo determinado o número de células vivas e mortas.

3.2.5. Avaliação do efeito dos extratos sobre a produção de ERO por neutrófilos

A produção das espécies reativas de oxigênio (ERO) foi medida por quimioluminescência (QL) dependente de luminol (Sigma-Aldrich, St Louis, MO, EUA, nº catálogo A8511) na concentração de 10⁻⁴M, baseando-se nos ensaios descritos por Alves *et al.* (2003). Os neutrófilos (1x10⁶células/mL em 200µL de reação) em Hanks pH 7,2 contendo gelatina 0,1% foram tratados com os extratos das orquídeas em diferentes concentrações (200, 100, 50, 25 e 12,5µg/mL). Essa mistura foi incubada em um luminômetro (EG&G Berthold, AutoLumat Plus LB 953, Bad Wildbad, Alemanha) por 30 minutos a 37°C, seguida pela ativação dos neutrófilos por PMA na concentração de 10⁻⁷M. A QL foi acompanhada por 20 minutos a 37°C e registrada em contagem de fótons por minuto (cpm). Os controles de QL espontânea consistiram

em incubar as células e luminol na ausência dos estímulos e das substâncias. Os resultados apresentados nos gráficos foram expressos como a área integrada sob o perfil de QL registrado.

3.3. Análises dos extratos por UFLC-DAD-ESI-MS/MS (microTOF-QII)

Os extratos de média-alta polaridade foram analisados em um sistema UFLC-DAD-ESI-MS/MS (microTOF-QII) (Bruker Daltonics), utilizando uma coluna C18 (Kinetex™, Phenomenex) de tamanho de partícula de 2,7 µm (150 x 4,6 mm). O detector de UV (DAD) foi programado na faixa de comprimentos de onda entre 100 e 900 nm e a temperatura do forno foi de 35 °C. Foi utilizado CH₃CN-HCO₂H (0,1%) (v/v) (B) e H₂O-HCO₂H (0,1%) (v/v) (A) como fase móvel, a uma vazão de 1,0 mL.min⁻¹, volume de injeção 25 µL e o seguinte gradiente de eluição: 0 – 2 min, 3% B; 2 – 30 min, 3 - 100% B; 30 – 34 min, 100% B; 34 – 37 min, 100 - 3% B, 37 – 40 min, 3% B, sendo o calibrante analisado no intervalo de 37 – 40 min.

As análises por espectrometria de massas foram realizadas nos modos de ionização positivo e negativo, separadamente, utilizando a taxa de espectros (*spectra rate*) de 2,0Hz e faixa de massa de 50 a 1300 *m/z*. Os parâmetros da fonte de ionização foram: voltagem do *End Plate Offset* de 450V; voltagem do capilar, 3,0 kV (modo negativo) e 3,6 kV (modo positivo); nebulizador a 5,0 bar; vazão do gás de secagem de 10,5 L/min; temperatura do gás de secagem, 220 °C.

Os parâmetros para a fragmentação foram Auto MS/MS com a seleção dos íons precursores em ciclos de 3,0 s, excluindo a seleção após detecção de 2 espectros do mesmo íon ou após 0,60min, reconsiderando sua seleção se a intensidade do íon precursor for o dobro da anterior. Para as fragmentações dos íons precursores, foi utilizada uma rampa de energia de colisão de 20 a 50 eV.

De modo a obter massas de alta resolução, o instrumento e os experimentos foram calibrados a partir de uma solução de padrão interno de Na-TFA (400 mg/100 mL) ao final de cada análise. Os dados foram analisados usando o software DataAnalysis 4.3 (Bruker, Daltonik). A cada dez amostras, um branco contendo por solução extratora e um QC-Pool foi analisado. O QC-Pool é uma mistral realizada para controle de qualidade das análises realizadas, contendo 10,0 µl de cada amostra a ser analisada.

As amostras foram preparadas solubilizando 1,0 mg de extratos secos com 1,0 mL de solução MeOH:H₂O (7:3, v/v), foram filtradas com filtro de seringa com membrana filtrante de PTFE (0,22 µm), acondicionada em *vials* de vidro de 2,0 mL e, em seguida, analisadas.

3.4. Pré-processamento dos Dados

Os dados adquiridos nas análises por LC-MS/MS no modo de ionização positivo e negativo foram convertidos, separadamente, em arquivos .mzxml utilizando o software livre ProteoWizard - MSconvert (v.3, Proteowizard Software Foundation, Palo Alto, CA, EUA) (Chambers et.al. , 2012). Posteriormente, foram processados utilizando o software *MzMineTM* (v. 2.51, BMC Bioinformatics, Reino Unido) (Pluskal et.al., 2010)., com o intuito de obter cromatogramas alinhados, deconvoluídos, com mínima influência de ruído e isótopos, e que representem ao máximo os extratos analisados. Os parâmetros utilizados no processamento dos dados adquiridos no modo positivo e negativo estão nas Tabelas 1 e 2, respectivamente.

Após o processamento, foram exportados 3 arquivos de dados: arquivo .mgf, com as informações dos espectros de MS², um arquivo .csv, com a matriz de dados contendo os íons precursores selecionados para fragmentação (MS²), as espécies em que aparecem, as áreas dos picos, razões massa-carga (*m/z*) e tempos de retenção; e outro arquivo .csv com uma matriz de dados contendo os íons precursores totais (selecionados ou não para a fragmentação) (MS¹) com os mesmo descritores do arquivo contendo MS².

Tabela 1. Parâmetros utilizados no software *MzMine*TM para o processamento dos dados adquiridos no modo de ionização positivo por análise em UFLC-DAD-ESI-MS/MS (microTOF-QII).

Raw data methods/ Peak detection	<p>Mass detection: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS1 <u>Mass Detector:</u> Centroid - Noise level= 1.0E3</p> <p>Mass detection: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS2 <u>Mass Detector:</u> Centroid - Noise level= 1.0E1</p> <p>Chromatogram builder: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS1 Min time span (min)= 0.01; Min height= 3.0E3; <i>m/z</i> tolerance= 20,0 ppm</p>
Peak list methods	<p>Chromatogram deconvolution: <u>Baseline cut-off:</u> Min peak height = 3.0E3; peak duration= 0.01 – 4.0; baseline level= 3.0E3 <i>m/z</i> range for MS2 scan pairing (Da): 0.01; RT range for MS2 scan pairing (min): 0.2</p> <p>Isotopic peaks grouper: <i>m/z</i> tolerance= 20 ppm; Retention time tolerance= 0,2 min; Maximum charge= 3; Representative isotope= most intense.</p> <p>Alignment: <u>Join aligner:</u> <i>m/z</i> tolerance= 20 ppm; Weight for <i>m/z</i>= 75; Retention time tolerance= 0.4 min; weight for RT= 25</p> <p>Filtering: <u>Peak list rows filter:</u> Keep only peaks with MS2 scan (GNPS): OK Reset the peak number ID: OK</p>

Legenda: min, minutos; *m / z*, massa/carga; ppm, partes por milhão.

Tabela 2. Parâmetros utilizados no software *MzMine*TM para o processamento dos dados adquiridos no modo de ionização negativo por análise em UFLC-DAD-ESI-MS/MS (micrOTOF-QII).

Raw data methods/ Peak detection	<p>Mass detection: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS1 <u>Mass Detector:</u> Centroid - Noise level= 1.0E3</p> <p>Mass detection: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS2 <u>Mass Detector:</u> Centroid - Noise level= 1.0E1</p> <p>ADAP Chromatogram builder: <u>Scans:</u> Retention Time: 0-33 min; MS level: MS1 Min group size in # of scan= 5 ; Group intensity threshold= 1.0E3; Min highest intensity= 1.5E3; <i>m/z</i> tolerance= 20,0 ppm</p>
Peak list methods	<p>Chromatogram deconvolution: <u>Wavelets (ADAP):</u> S/N threshold= 10; S/N estimator= Intensity window S/N; min feature hight= 2.0E3; coefficient/area threshold= 10; Peak duration range= 0.01 – 4.0; RT wavelet range= 0.001-0.1 <i>m/z</i> center calculation: MEDIAN; <i>m/z</i> range for MS2 scan paring (Da): 0.01; RT range for MS2 scan paring (min): 0.2</p> <p>Isotopic peaks grouper: <i>m/z</i> tolerance= 20 ppm; Retention time tolerance= 0,2 min; Maximum charge= 3; Representative isotope= most intense.</p> <p>Alignment: <u>Join aligner:</u> <i>m/z</i> tolerance= 20 ppm; Weight for <i>m/z</i>= 75; Retention time tolerance= 0.2 min; weight for RT= 25</p> <p>Filtering: Peak list rows filter: Keep only peaks with MS2 scan (GNPS): OK Reset the peak number ID: OK</p>

Legenda: min, minutos; *m / z*, massa/carga; ppm, partes por milhão.

3.5. Análises de Correlação in silico

As análises estatísticas multivariadas supervisionadas foram realizadas para com o intuito de propor as substâncias envolvidas na ação antioxidante, citotóxica e moduladora da produção de ERO por neutrófilos humanos. As Análises Discriminantes por Mínimos Quadrados Parciais (*Partial Least Square – Discriminant Analysis – PLS-DA*) e as Análises Discriminantes por Mínimos Quadrados Parciais Ortogonais (*Orthogonal Partial Least Square – Discriminant Analysis – OPLS-DA*) foram realizadas utilizando o software SIMCA 13.0.3.0[©] (Umetrics) utilizando a transformação por Log em todas as matrizes.

3.6. Anotações via Redes Moleculares

Os arquivos .mgf e .csv (MS²) adquiridos após processamento dos dados das análises por LC-MS/MS no modo de ionização positivo e negativo foram enviados para

a plataforma online do GNPS (<https://gnps.ucsd.edu>, Wang et al., 2016) pelo do software WinSCP (<https://winscp.net/eng/index.php>), para a construção das redes moleculares utilizando a ferramenta “*Feature-Based Molecular Networking (FBMN)*” (Nothias et al., 2019). Foram construídas uma rede para cada modo de ionização.

A formação da rede considerou os valores de tolerância de massa de íons precursores e de íons fragmento de, respectivamente, 0,08 Da e 0,02 Da. Na construção dos agrupamentos (*cluster*), o valor mínimo de cosseno de similaridade para que fosse formada uma borda (*edge*) e conectasse dois espectros consenso (nodos) foi de 0.7, com no mínimo 3 íons fragmento compatível no espectro. As bordas foram mantidas entre dois nodos somente se cada nodo está entre os 10 mais similares (*Network TopK*) e o tamanho máximo de uma família molecular (*cluster*) foi definido como 100 (*Maximum Connected Component Size*). A máxima diferença de massa entre os precursores foi de 500Da.

Os espectros da rede foram pesquisados nas bibliotecas espectrais do GNPS (Wang et al., 2016; Horai et al., 2010) considerando compatíveis os espectros que possuíam cosseno de similaridade maiores que 0,7 e pelo menos 3 íons fragmento em comum. A visualização das redes moleculares foi realizada no Cytoscape 3.2 (Shannon et al., 2003), o qual permitiu a visualização na forma de uma rede de agrupamentos de nodos (íons precursores) e arestas (valores de cosseno de similaridade entre análogos). Cada agrupamento de nodos corresponde a uma família molecular (*cluster*).

4. Resultados e Discussão

4.1. Atividade Antioxidante dos extratos brutos

O teste utilizando DPPH (2,2-difenil-1-picril-hidrazil) permite avaliar a capacidade redutora dos extratos de orquídeas, avaliando seus metabólitos secundários com capacidade antioxidante. Ao final do ensaio, é possível calcular a porcentagem final de DPPH reduzido através da leitura espectrofotométrica.

As amostras foram consideradas potenciais antioxidantes em valores acima de 70% de DPPH reduzido. Foi selecionado esse valor pois buscamos compostos que tenham alta capacidade de redução dos reagentes de oxigênio produzidos pelos neutrófilos humanos em uma resposta autoimune, como ocorre no LES. Dentre as 129 amostras analisadas, 27 promoveram a redução de DPPH acima de 70% do DPPH e

66 amostras, acima de 50%. Abaixo está a tabela com os resultados das análises por DPPH (Tabela 1).

Tabela 3. Resultados das análises de ação antioxidante por DPPH de extratos hidrometanólicos de espécies brasileiras de Orchidaceae. *Maior que 70% de inibição.

ID	DPPH reduzido (%)	ID	DPPH reduzido (%)	ID	DPPH reduzido (%)	ID	DPPH reduzido (%)
Aap19	20.60	Cmi33*	83.35	Iut74	60.73	Pro10	58.85
Aeb22	30.60	Csa47	13.93	Kgr105	53.82	Psch4	20.83
Agj97	31.53	Cte41	-78.88	Llu80	27.63	Psp115	69.55
Ahy3	16.81	Cva49	34.02	Lne79	28.98	Pta84	30.76
Ale20	32.86	Cvi37	67.75	Lpe77	46.66	Pvi2*	77.44
Ava23	53.10	Dco96	63.74	Lpo78	29.06	Pvi9	-5.40
Bch30	9.88	Dfi53	41.35	Mari12	36.92	Rco110	42.75
Bgr26	64.57	Dki55	67.11	Mca68	51.81	Rde90*	69.90
Blo34	54.05	Dmo54	57.47	Mcu7	52.18	Rgi100	15.57
Bma28 *	72.58	Dsp56	47.60	Mde81	75.31	Rgl125	25.34
Bpi27	38.31	Eam119*	77.49	Mfl67*	72.89	Rgo112	55.30
Bte25	29.79	Eav64*	86.08	Mle122*	86.30	Rla118	24.73
Btr44 *	87.24	Ede59	66.61	Mlu128	61.31	Rpi109	35.07
Btu29	58.54	Ede60	49.58	Mva82	36.24	Scr85	43.56
Casp121 *	81.62	Eea31	47.73	Nam113	68.35	Sin87*	75.35
Cbi43	6.40	Eel103*	80.53	Oal92	58.20	Smo86	65.17
Cca5	46.03	Epu61*	82.63	Oam123	43.28	Sno116	12.40
Cco6	65.70	Eri62	48.29	Oce98	66.90	Spl26*	72.88
Ccr48	67.31	Ese63	14.87	Oco94	39.19	Ssh11**	72.54
Cde46	51.95	Esp57*	77.94	Ocr114	64.22	Ssp102	88.98
Ceu51	29.05	Etr91	66.78	Oma111	30.47	Ssp21	13.70
Cfi35	63.20	Eto58	41.00	Omo95*	70.71	Tfa117	55.51
Cfl120	44.63	Gam71*	81.38	Omy66	27.34	Vba13	23.86
Cfo38	30.63	Gbo129	56.58	Oon126	53.51	Vch14*	72.04
Cgi52	24.03	Gcr69	25.91	Ora1	13.14	Vha127	18.45
Cha42	36.74	Gmo65*	79.20	Osp93	65.42	Vpa16	45.20
Cha50	7.56	Gqu70	63.72	Pbo8	7.13	Vpal15	49.59
Cin39	27.11	Hpa130	15.96	Pca88	58.78	Vpo17	12.07
Cla124	59.39	Hpu72*	79.76	Pes101	90.27	Vri18	8.70
Clo45	41.47	Hsu73	25.27	Pfr104	61.23	Vsa99*	81.44
Clu40*	71.64	Ili768*	92.88	Plu108	39.95	Wwa107*	77.50
Cma36	30.50	Ipu75	64.17	Pmi106	54.03	Xsq83	49.72
						Zcr89	68.74

4.2. Citotoxicidade dos extratos brutos

Os ensaios de citotoxicidade em células de neutrófilo são um importante passo para selecionar quais amostras, e suas respectivas doses, devem ser estudadas nas próximas etapas. Assim, foram consideradas aptas para as próximas etapas das análises em neutrófilos humanos os extratos que obtiveram no mínimo de 80% de viabilidade celular.

As análises de viabilidade celular em neutrófilos humanos foram realizadas em cinco concentrações diferentes: 200 $\mu\text{g}\cdot\text{ml}^{-1}$, 100 $\mu\text{g}\cdot\text{ml}^{-1}$, 50 $\mu\text{g}\cdot\text{ml}^{-1}$, 25 $\mu\text{g}\cdot\text{ml}^{-1}$, 12.5 $\mu\text{g}\cdot\text{ml}^{-1}$. Assim, para as 129 amostras, foram realizadas 647 análises. Dentre as espécies estudadas, somente uma foi tóxica em todas as concentrações possíveis, *Cycnoches haagii* Barb.Rodr. (Cha50), havendo uma diminuição linear da viabilidade celular com o aumento da concentração do extrato.

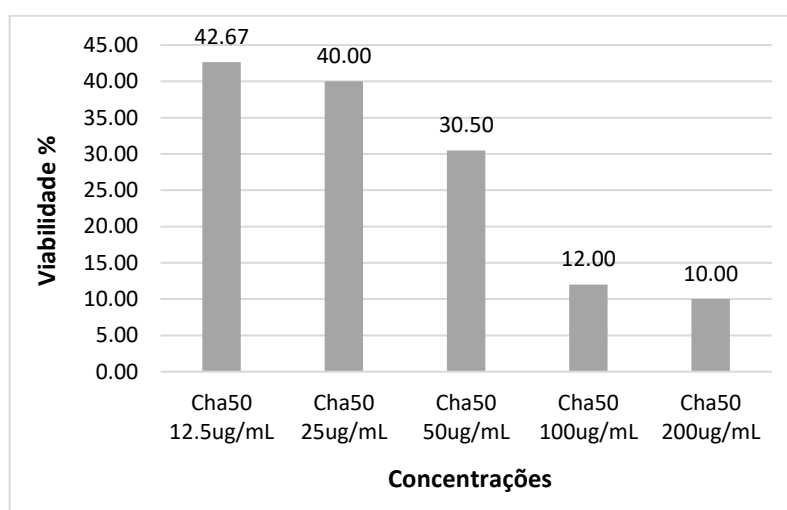


Gráfico 1. Gráfico da viabilidade celular em neutrófilos humanos na presença do extrato hidrometanólico de *Cycnoches haagii* Barb.Rodr. (Cha50) em cinco concentrações diferentes, demonstrando o seu grande potencial citotóxico.

Somente 53 doses entre 28 amostras não foram selecionadas para as próximas etapas, sendo que eram, em maioria, doses em altas concentrações (200 $\mu\text{g}\cdot\text{ml}^{-1}$ e 100 $\mu\text{g}\cdot\text{ml}^{-1}$). Em 6 espécies (Brg26, Cha50, Eam119, Eri62, Ipu75, Pmi106) foi observada citotoxicidade utilizando concentração de 50 $\mu\text{g}\cdot\text{ml}^{-1}$, e duas delas (Cha50 e Eri62) apresentaram toxicidade a 25 $\mu\text{g}\cdot\text{ml}^{-1}$.

4.3. Efeitos sobre a Produção de ERO por Neutrófilos dos extratos brutos

A produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos foi avaliada em 128 amostras e, dentre essas amostras, 590 doses entre $200\mu\text{g}\cdot\text{ml}^{-1}$, $100\mu\text{g}\cdot\text{ml}^{-1}$, $50\mu\text{g}\cdot\text{ml}^{-1}$, $25\mu\text{g}\cdot\text{ml}^{-1}$, $12.5\mu\text{g}\cdot\text{ml}^{-1}$. Os resultados estão demonstrados nos gráficos a seguir. Os gráficos foram separados por dias de experimento, para os cálculos estatísticos estarem relacionados ao controle utilizado no dia. A amostra controle equivale a produção basal de ERO em um neutrófilo saudável.

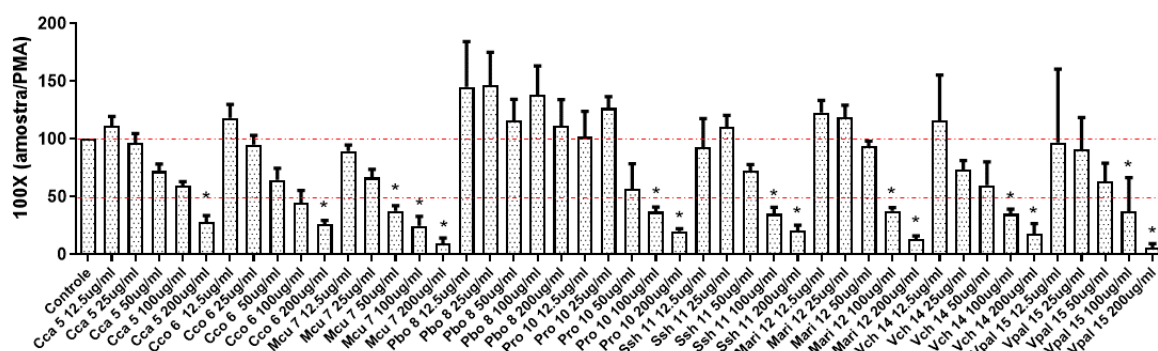


Gráfico 2. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies Cca 5, Cco 6, Mari 12, Mcu 7, Pbo 8, Pro 10, Ssh 11, Vch 14 e Vpal 15 de Orchidaceae.

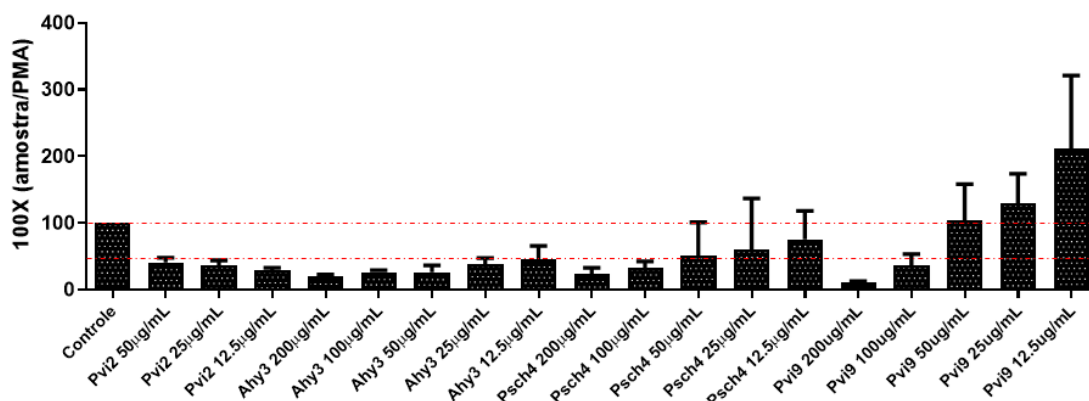


Gráfico 3. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies Ahy3, Psch4, Pvi2 e Pvi9 de Orchidaceae.

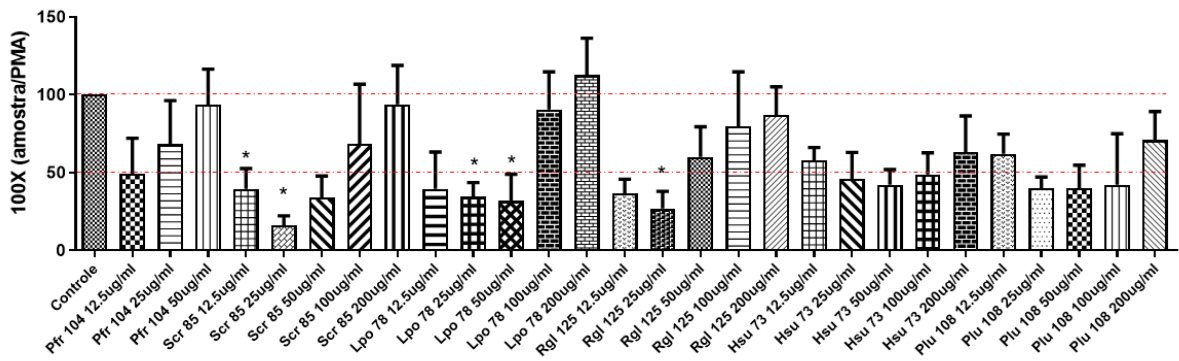


Gráfico 4. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies Hsu 73, Lpo 78, Pfr 104, Plu 108, Rgl 125 e Scr 85 de Orchidaceae.

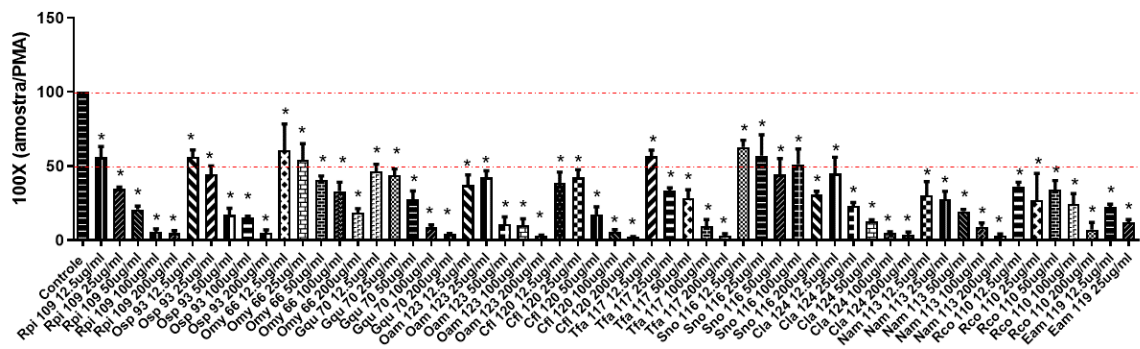


Gráfico 5. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies Cfl 120, Cia 124, Eam 119, Gqu 70, Nam 113, Oam 123, Omy 66, Osp 93, Rco 110, Rpi 109, Sno 116 e Tfa 117 de Orchidaceae.

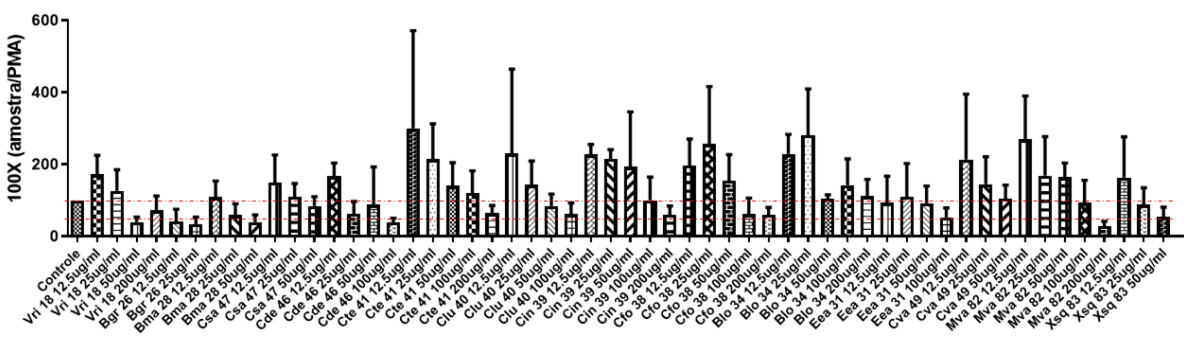


Gráfico 6. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies Bgr 26, Blo 34, Bma 28, Cde 46, Cfo 38, Cin 39, Clu 40, Csa 47, Cte 41, Cva 49, Eea 31, Mva 82, Vri 18 e Xsq 83 de Orchidaceae.

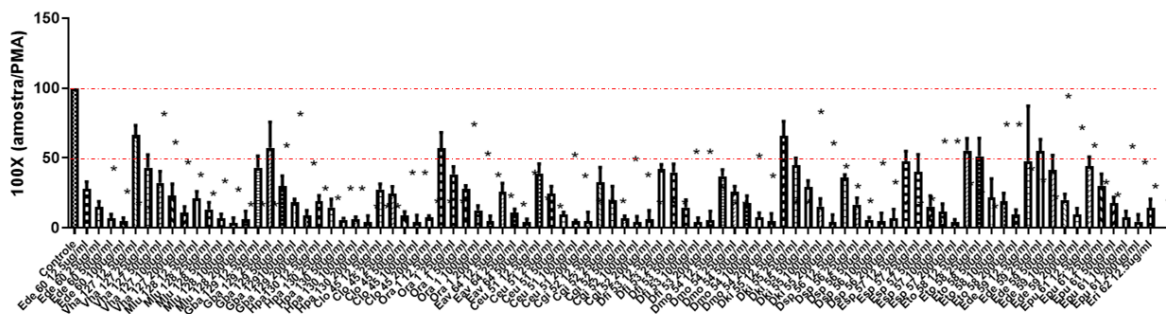


Gráfico 7. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies *Ceu 51*, *Cgi 52*, *Clo 45*, *Dfi 53*, *Dki 55*, *Dmo 54*, *Dsp 56*, *Eav 64*, *Ede 59*, *Ede 60*, *Epu 61*, *Eri 62*, *Esp 57*, *Eto 58*, *Gba 129*, *Hpa130*, *Mlu 128*, *Ora 1* e *Vha 127* de Orchidaceae.

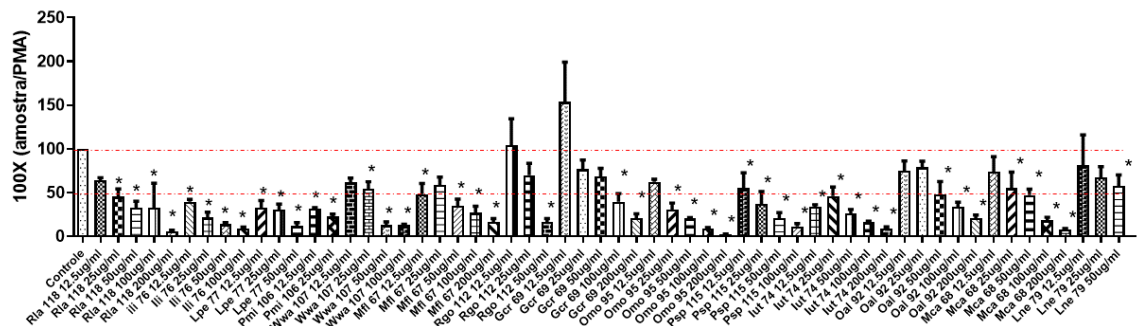


Gráfico 8. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies *Gcr 69*, *Ili 76*, *lut 74*, *Lne 79*, *Lpe 77*, *Mca 68*, *Mfl 67*, *Oal 92*, *Omo 95*, *Pmi 106*, *Psp 115*, *Rgo 112*, *Rla 118* e *Wwa 107* de Orchidaceae.

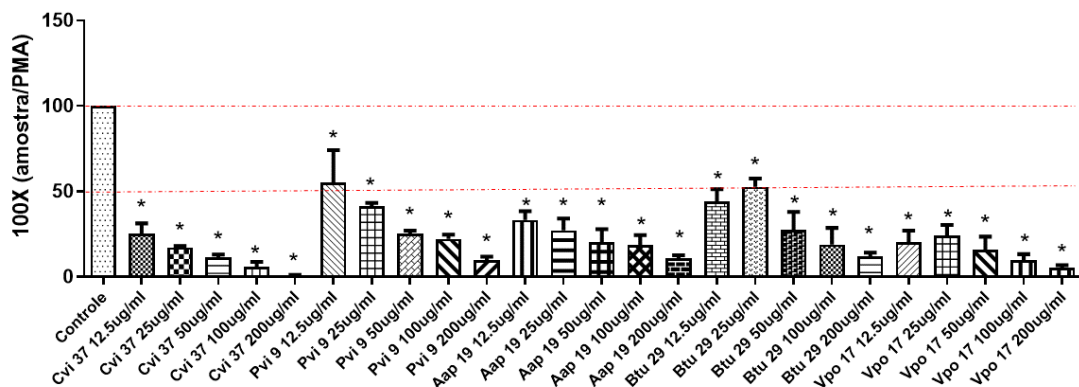


Gráfico 9. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies *Aap 19*, *Btu 29*, *Cvi 37* e *Pvi 9* de Orchidaceae.

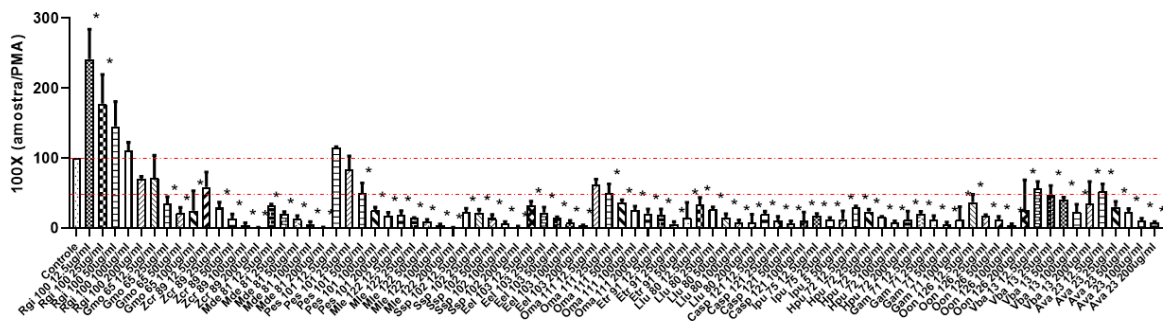


Gráfico 10. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies Ava 23, Casp 121, Eel 103, Etr 91, Gam 71, Gmo 65, Hpu 72, Ipu 75, Llu 80, Mde 81, Mle 122, Oma 111, Oon 126, Pes 101, Rgi 100, Ssp 102 Vba 13 e Zcr 89 de Orchidaceae.

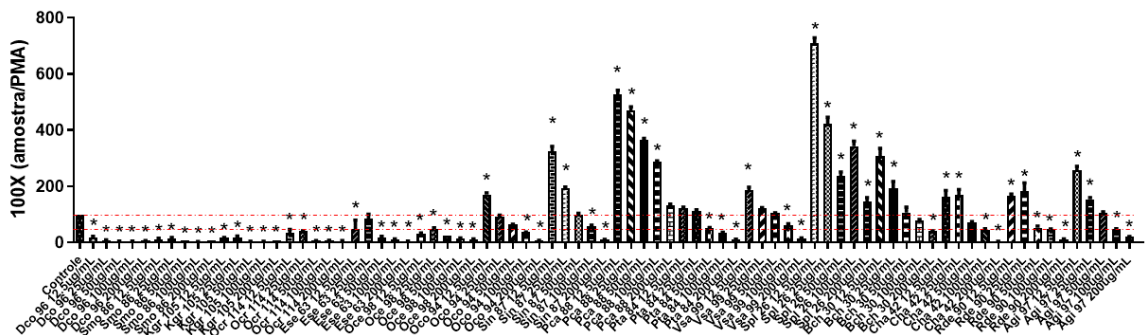


Gráfico 11. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies Agi 97, Bch 30, Cha 42, Dco 96, Ese 63, Kgr 105, Oce 98, Oco 94, Ocr 11,4 Pca 88, Pta 84, Rde 90, Sin 87, Smo 86, Spl 26 e Vsa 99 de Orchidaceae.

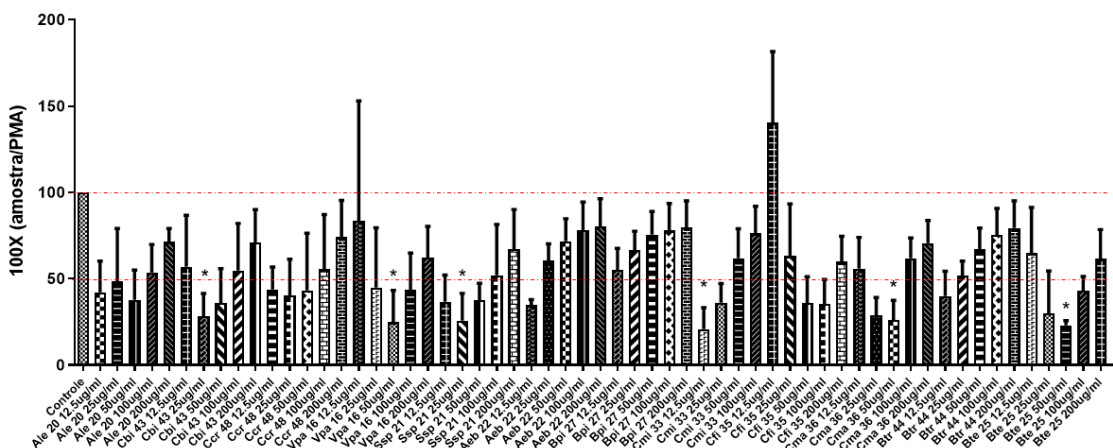


Gráfico 12. Gráfico da produção das espécies reativas de oxigênio (ERO) por neutrófilos humanos em contato com os extratos das espécies Aeb 22, Ale 20, Bpi 27, Bte 25, Btr 44, Cbi 43, Ccr 48, Cfi 35, Cma 36, Cmi 33, Ssp 21 e Vpa 16 de Orchidaceae.

O espaço entre as linhas tracejadas vermelhas indica o intervalo entre 100% e 50% de produção de ERO em relação ao controle. Esse é o intervalo resposta que se considerou uma amostra/dose ativa (A). Acima de 100% considerou-se amostras inativas (B), já que não auxiliaram na diminuição da produção de ERO. E abaixo de 50% considerou-se inibitório, colocando-os na categoria B pois abaixo desse valor a diminuição da produção de ERO pode estar relacionada a inibição da função do neutrófilo em si, o que não é buscado neste estudo.

Considerando ativos (A) os extratos que tiveram pelo menos uma dose analisada com resposta no intervalo entre 100-50% e os outros B (Inativos ou Inibitórios), obtivemos 75 amostras que diminuíram a produção de ERO por neutrófilos humanos, seja por inibição de rotas de produção diretamente no neutrófilo ou por interação direta com o reativo de oxigênio.

4.4. Biomarcadores do Potencial Ativo de Orchidaceae Brasileiras

4.4.1. Determinação dos Biomarcadores para a Atividade Antioxidante

Em uma família tão rica em compostos de potencial antioxidante, possuir 51% das amostras com ação redutora acima de 50% faz jus ao seu potencial. Os extratos brutos são misturas complexas de metabólitos secundários em diferentes concentrações, e a interação entre eles em uma reação química de redução de DPPH é algo inesperado, principalmente para estas as espécies estudadas. O potencial antioxidante conseguiria avaliar mais precisamente após processos de partição e seleção dos metabólitos, mas seria um processo manual laborioso e demorado, principalmente para 128 espécies de Orchidaceae.

Com o intuito de propor as classes e compostos biomarcadores da atividade antioxidante, foram realizadas análises estatísticas multivariadas supervisionadas por PLS-DA (*Partial Least Square – Discriminant Analysis*), utilizando os extratos com valores acima de 70% classificados como “Ativo” e os restantes como “Inativo”. Nesta análise de correlação *in silico*, foram correlacionadas as impressões digitais metabólicas (modos de ionização positivos e negativo) com as propriedades antioxidantes citadas acima. A figura 1 mostra o gráfico de scores da análise por PLS-DA, no qual é possível observar uma boa separação entre os dois grupos.

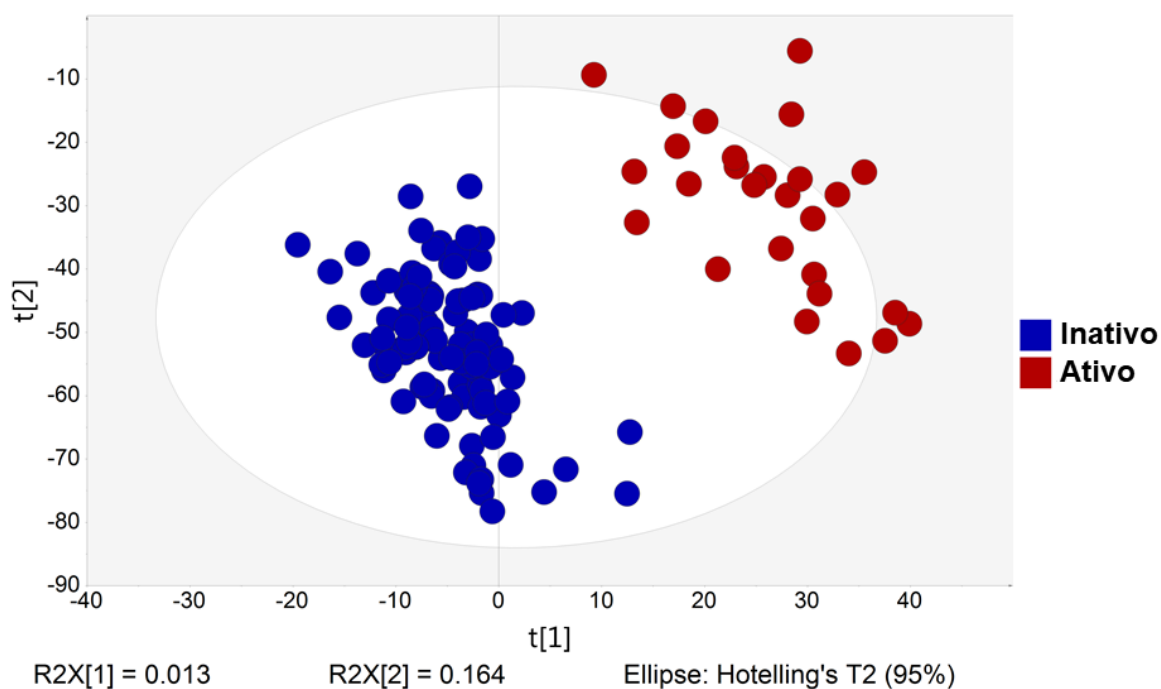


Figura 1. Gráfico de dispersão de Score gerados pela análise estatística multivariada supervisionada PLS-DA (*Partial Least Square – Discriminant Analysis*) correlacionando as impressões digitais metabólicas com as propriedades antioxidantes dos extratos analisados de Orchidaceae brasileiras. R^2Y . 0,95. Q^2 .0,68.

A partir das análises por PLS-DA obteve-se o gráfico de VIP (*Variable Influence on Projection*), que fornece as variáveis mais importantes para a projeção dos grupos, e os gráficos de coeficientes relacionados ao Y (propriedade antioxidante). Utilizando esses gráficos, foram selecionados os metabólitos relacionados à ação antioxidante (grupo ativo), considerando os valores de $VIP \geq 1,0$ e de coeficientes relacionado ao Y positivos.

Dentre as variáveis indicadas pelo modelo para o grupo ativo, 21 compostos foram anotados pela rede molecular (Anexos, tabela 2), e dentre eles as classes: flavonóides-7-O-glicosilados, ácidos cumáricos e derivados, cumarinas glicosiladas, cromonas, ácidos hidrocicinâmicos, xantonas, sesquiterpenóides, triterpeno, Acilglicosídeos graxos de mono e dissacarídeos, e até ácidos fenilpiruvicos e derivados. Algumas dessas classes foram discutidas no capítulo 2 desta tese, as quais já demonstraram seu poder antioxidante em estudos anteriores. É importante frisar que a anotação molecular nos traz o conhecimento à nível de classes e subclasses moleculares. Para confirmar a nível de moléculas e estruturas, outras análises devem ser realizadas.

4.4.2. Determinação dos Biomarcadores da Ação Moduladora da Produção de ERO por Neutrófilos Humanos

Nas análises do efeito dos extratos na produção de ERO por neutrófilos humanos, a resposta positiva (59% das amostras) foi além da resposta da atividade antioxidante (21% das amostras). O efeito dos extratos na produção de ERO pelo neutrófilos não depende exclusivamente do potencial antioxidante do composto, que atuaria diretamente na espécie em si, mas ele pode atuar em outras rotas relacionadas à produção de ERO que o neutrófilo, inibindo receptores, impedindo a liberação, etc.

Um exemplo é a influência na via NADPH oxidase, a qual é responsável pela produção de componentes do complemento, ERO e quimiocinas como CCL2/MCP-1 e CXCL8/ IL-8 (Pereira-Crott et. Al, 2020).

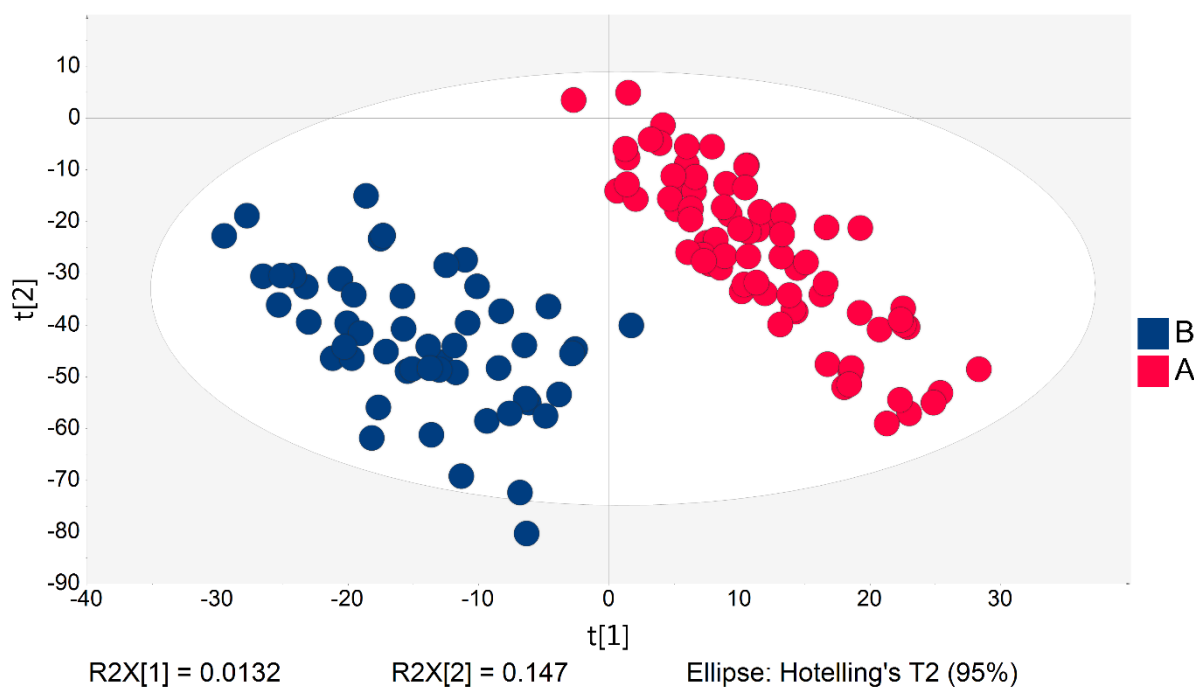


Figura 2. Gráfico de dispersão de Score gerados pela análise estatística multivariada supervisionada PLS-DA (*Partial Least Square – Discriminant Analysis*) correlacionando as impressões digitais metabólicas com as propriedades moduladoras da produção de ERO por neutrófilos humanos, dos extratos analisados de Orchidaceae brasileiras. R^2Y . 0,98. Q^2 .0,73.

Com o intuito de propor as classes e compostos biomarcadores do efeito modulador da produção de ERO por neutrófilos humanos, foram realizadas análises estatísticas multivariadas supervisionadas por PLS-DA (*Partial Least Square – Discriminant Analysis*), considerando (A) os extratos que tiveram pelo menos uma

dose analisada com resposta no intervalo entre 100-50% e **(B)** os que tiveram valores acima de 100% (Inativos) ou abaixo de 50% (Inibitórios).

Nesta análise de correlação *in silico*, foram correlacionadas as impressões digitais metabólicas (modos de ionização positivos e negativo) com as propriedades citadas acima. A figura 2 mostra o gráfico de scores da análise por PLS-DA, no qual é possível observar uma boa separação entre os dois grupos.

Dentre as variáveis indicadas pelo modelo para o grupo ativo, 35 compostos foram anotados pela rede molecular (Anexos, tabela 3), e dentre eles as classes: 2'-Hidroxicalconas, flavanonas 8-preniladas, peptídeos cíclicos, flavonoides -7-O-glicosilados, lignanas glicosiladas, acilglicosídeos graxos de mono e dissacarídeos, quinolinas carboxamidas, diterpenos, sesquiterpenos, dentre outros. Assim como no modelo anterior, algumas dessas classes foram discutidas no capítulo 2 desta tese, as quais já demonstraram seu poder antioxidante em estudos anteriores.

5. CONCLUSÕES

Este estudo permitiu evidenciar o grande potencial ativo das espécies brasileiras de Orchidaceae, de diferentes tribos e gêneros, quanto à atividade antioxidante e moduladora da produção de ERO por neutrófilos humanos. Dentre as espécies estudadas, 27 apresentaram atividade antioxidante, 75 apresentaram ação modulatória na produção de ERO por neutrófilos humanos e 14 delas apresentaram as duas atividades, havendo uma grande chance de as moléculas responsáveis pela ação antioxidante também atuarem como moduladoras da produção de ROS.

As análises metabolômicas, quimiométricas e ferramentas de anotação tornaram mais rápido o processo de acesso ao metaboloma e ao conhecimento biológico contido dele, possibilitando a indicação das classes dos flavonoides, auronas, chalconas, alcaloides, ácido chiquímicos como atuantes na atividade oxidante.

Os resultados das atividades foram promissores, demonstrando 50% de espécies com potencial modulados e antioxidante, e que apresentam baixos níveis de toxicidade do extrato bruto. São promissores pois não apresentaram citotoxicidade, sugerindo estas como potenciais plantas para serem selecionadas para estudos mais avançados sobre moduladores e doenças relacionadas.

CONCLUSÕES GERAIS e PERSPECTIVAS

Este é o primeiro estudo que avalia mais de 100 espécies brasileiras das 4 subfamílias de Orchidaceae, contribuindo com o conhecimento farmacológico dessas que foram por tanto tempo negligenciadas na ciência brasileira. E isso só foi possível devido à estudos químicos, *in vitro* e análises metabolômicas, que consomem menor quantidade de amostra, tempo e trabalho para trazer um alicerce para futuros estudos.

A análise metabolômica tornou possível o acesso ao conhecimento sobre os metabólitos secundários de mais de 100 espécies de Orchidaceae brasileiras, sendo que a grande maioria das classes observadas neste estudo são inéditas para espécies de orquídeas brasileiras, e algumas classes são inéditas para a família Orchidaceae.

Dentre as anotações obtidas, algumas classes possuem um histórico muito interessante sobre seu potencial ativo e utilização farmacológica. Algumas delas foram propostas como detentoras das atividades antioxidantes e moduladoras da produção de ERO por neutrófilos humanos. Neste estudo foi demonstrado o grande potencial biológico que as orquídeas brasileiras possuem, antioxidante e modulador.

Através deste trabalho, o trazemos como guia para próximos estudos químicos e farmacológicos, através da seleções de espécies mais potentes e classes de metabólitos de interesse para a hipótese a ser estudada.

REFERÊNCIAS GERAIS

Aksenov AA, Laponogov I, Zhang Z, Doran SL, Belluomo I, Veselkov D, Bittremieux W, Nothias LF, Nothias-Esposito M, Maloney KN, Misra BB. Auto-deconvolution and molecular networking of gas chromatography–mass spectrometry data. *Nature Biotechnology*. 2021 Feb;39(2):169-73.

Beale, D.J., Pinu, F.R., Kouremenos, K.A. et al. Review of recent developments in GC–MS approaches to metabolomics-based research. *Metabolomics* 14, 152 (2018). <https://doi.org/10.1007/s11306-018-1449-2>

Beale, D.J., Pinu, F.R., Kouremenos, K.A. et al. Review of recent developments in GC–MS approaches to metabolomics-based research. **Metabolomics**, v. 14, p.152, 2018.

Beale, D.J., Pinu, F.R., Kouremenos, K.A. et al. Review of recent developments in GC–MS approaches to metabolomics-based research. *Metabolomics* 14, 152 (2018). <https://doi.org/10.1007/s11306-018-1449-2>

Brunetti AE , Carnevale Neto F , Vera MC , Taboada C , Pavarini DP , Bauermeister A , Lopes NP . An integrative omics perspective for the analysis of chemical signals in

ecological interactions. *Chem Soc Rev.* 2018 Mar 5;47(5):1574-1591. doi: 10.1039/c7cs00368d. PMID: 29114668.

Carnevale Neto F, Pilon AC, Selegato DM, Freire RT, Gu H, Raftery D, Lopes NP and Castro-Gamboa I (2016) Dereplication of Natural Products Using GC-TOF Mass Spectrometry: Improved Metabolite Identification by Spectral Deconvolution Ratio Analysis. *Front. Mol. Biosci.* 3:59. doi: 10.3389/fmolb.2016.00059

Chambers, M., Maclean, B., Burke, R. *et al.* A cross-platform toolkit for mass spectrometry and proteomics. *Nat Biotechnol* **30**, 918–920 (2012). <https://doi.org/10.1038/nbt.2377>

Creydt, M; Arndt, M.; Hudzik, D.; Fischer, M. Plant Metabolomics: Evaluation of Different Extraction Parameters for Nontargeted UPLC-ESI-QTOF-Mass Spectrometry at the Example of White Asparagus officinalis. *Journal of Agricultural and Food Chemistry* **2018** 66 (48), 12876-12887. DOI: 10.1021/acs.jafc.8b06037

Da Silva, R. R.; Wang, M.; Nothias, L-F.; Van Der Hoof, J. J. J.; Caraballo Rodríguez, A. M.; Fox, E.; Balunas, M. J.; Klassen, J. L.; Lopes, N. P.; Dorrestein, P. C. *Plos Comput. Biol.* V. 14, E100608, 2018

Da Silva, R. R.; Wang, M.; Nothias, L-F.; Van Der Hoof, J. J. J.; Caraballo Rodríguez, A. M.; Fox, E.; Balunas, M. J.; Klassen, J. L.; Lopes, N. P.; Dorrestein, P. C. *PLOS Comput. Biol.* V. 14, e100608, 2018

De Vos, R. C. *et al.* Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Natural Protocols*, v. 2, p. 778–791, 2007.

Ernst M, Silva DB, Silva RR, Vêncio RZ, Lopes NP. Mass spectrometry in plant metabolomics strategies: from analytical platforms to data acquisition and processing. *Nat Prod Rep.* 2014;31(6):784-806. doi:10.1039/c3np70086k

Ernst M, Silva DB, Silva RR, Vêncio RZ, Lopes NP. Mass spectrometry in plant metabolomics strategies: from analytical platforms to data acquisition and processing. *Nat Prod Rep.* 2014;31(6):784-806. doi:10.1039/c3np70086k

Fiehn O. Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling. *Curr Protoc Mol Biol.* 2016;114:30.4.1-30.4.32. Published 2016 Apr 1. doi:10.1002/0471142727.mb3004s114

Fiehn O. Metabolomics by Gas Chromatography-Mass Spectrometry: Combined Targeted and Untargeted Profiling. *Curr Protoc Mol Biol.* 2016;114:30.4.1-30.4.32. Published 2016 Apr 1. doi:10.1002/0471142727.mb3004s114

VarQuest: Gurevich, A. *et al.* Increased diversity of peptidic natural products revealed by modification-tolerant database search of mass spectra. *Nature Microbiology* 3, 319-327 (2018), <https://dx.doi.org/10.1038/s41564-017-0094-2>.

For the DEREPLICATOR: Mohimani, H. *et al.* Dereplication of microbial metabolites through database search of mass spectra, *Nature Communications* 9, 4035 (2018), <https://dx.doi.org/10.1038/s41467-018-06082-8>.

Frank, A. M. *et al.* Clustering millions of tandem mass spectra. *J. Proteome Res.* **7**, 113–122 (2008).

Jia Liu, Yang Liu, Yu Wang, Ann Abozeid, Yuan-Gang Zu, Zhong-Hua Tang. The integration of GC–MS and LC–MS to assay the metabolomics profiling in *Panax*

ginseng and *Panax quinquefolius* reveals a tissue and species-specific connectivity of primary metabolites and ginsenosides accumulation. *Journal of Pharmaceutical and Biomedical Analysis*. v. 135, p. 176–185, 2017.

Kanani H, Chrysanthopoulos PK, Klapa MI (2008) Standardizing GC-MS metabolomics. *J Chromatogr B* 871(2):191–201. <https://doi.org/10.1016/j.jchromb.2008.04.049>

Kanani H, Chrysanthopoulos PK, Klapa MI (2008) Standardizing GC-MS metabolomics. *J Chromatogr B* 871(2):191–201. <https://doi.org/10.1016/j.jchromb.2008.04.049>

Kanani, H., Chrysanthopoulos, P.K., Klapa, M.I. Standardizing GC-MS metabolomics. *Journal of Chromatography B*, v. 871(2), p.191–201, 2008.

Katajamaa M. et al, MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics* 22, 634-636 (2006), <https://doi.org/10.1186/1471-2105-11-395>.

Kind T, Wohlgemuth G, Lee DY, et al. FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. ***Analytical Chemistry***. V.81(24), p.10038-10048, 2009.

Kind T, Wohlgemuth G, Lee DY, et al. FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal Chem*. 2009;81(24):10038-10048. doi:10.1021/ac9019522

Kind T, Wohlgemuth G, Lee DY, et al. FiehnLib: mass spectral and retention index libraries for metabolomics based on quadrupole and time-of-flight gas chromatography/mass spectrometry. *Anal Chem*. 2009;81(24):10038-10048. doi:10.1021/ac9019522

Marília Elias Gallon, Olakunle Adeboye Jaiyesimi, Leonardo Gobbo-Neto. LC-UV-HRMS dereplication of secondary metabolites from Brazilian *Vernoniae* (Asteraceae) species supported through in-house database, *Biochemical Systematics and Ecology*, Volume 78, 2018, Pages 5-16, ISSN 0305-1978, <https://doi.org/10.1016/j.bse.2018.03.002>.

Martucci, M.E.P., Loeuille, B., Pirani, J.R., Gobbo-Neto, L. Comprehensive untargeted metabolomics of *Lychnophorinae* subtribe (Asteraceae: *Vernoniae*) in a phylogenetic context. *PLOS ONE*, v. 13, n°1, 2018.

Mushtaq, M.Y., Choi, Y.H., Verpoorte, R. and Wilson, E.G. (2014), Extraction for Metabolomics: Access to The Metabolome. *Phytochem. Anal.*, 25: 291-306. <https://doi.org/10.1002/pca.2505>

Myers, O.D. et al, One Step Forward for Reducing False Positive and False Negative Compound Identifications from Mass Spectrometry Metabolomics Data: New Algorithms for Constructing Extracted Ion Chromatograms and Detecting Chromatographic Peaks [<https://pubs.acs.org/doi/abs/10.1021/acs.analchem.7b00947>]. *Anal. Chem.* 89, 17, 866-8703 (2017)

Nothias LF et al. Feature-based Molecular Networking in the GNPS Analysis Environment. *bioRxiv* 812404 (2019). <https://doi.org/10.1101/812404>.

Nothias, L.F., Petras, D., Schmid, R. *et al.* Feature-based molecular networking in the GNPS analysis environment. *Nat Methods* **17**, 905–908 (2020). <https://doi.org/10.1038/s41592-020-0933-6>

Papadimitropoulos M.E.P., Vasilopoulou C.G., Maga-Nteve C., Klapa M.I. (2018) Untargeted GC-MS Metabolomics. In: Theodoridis G., Gika H., Wilson I. (eds) *Metabolic Profiling. Methods in Molecular Biology*, vol 1738. Humana Press, New York, NY.

Papadimitropoulos M.E.P., Vasilopoulou C.G., Maga-Nteve C., Klapa M.I. (2018) Untargeted GC-MS Metabolomics. In: Theodoridis G., Gika H., Wilson I. (eds) **Metabolic Profiling. Methods in Molecular Biology**, vol 1738. Humana Press, New York, NY

Papadimitropoulos M.E.P., Vasilopoulou C.G., Maga-Nteve C., Klapa M.I. (2018) Untargeted GC-MS Metabolomics. In: Theodoridis G., Gika H., Wilson I. (eds) *Metabolic Profiling. Methods in Molecular Biology*, vol 1738. Humana Press, New York, NY

Pilon, Alan C., Selegato, Denise M., Fernandes, Richard P., Bueno, Paula C. P., Pinho, Danielle R., Carnevale Neto, Fausto, Freire, Rafael T., Castro-Gamboa, Ian, Bolzani, Vanderlan S., & Lopes, Norberto P.. (2020). METABOLÔMICA DE PLANTAS: MÉTODOS E DESAFIOS. *Química Nova*, *43*(3), 329-354. Epub June 01, 2020. <https://doi.org/10.21577/0100-4042.20170499>

Pluskal T *et al.* MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics* **11**, 395 (2010), <https://doi.org/10.1093/bioinformatics/btk039>.

Pluskal T, Castillo S, Villar-Briones A, Oresic M. MZmine 2: modular framework for processing, visualizing, and analyzing mass spectrometry-based molecular profile data. *BMC Bioinformatics*. 2010 Jul 23;11:395. doi: 10.1186/1471-2105-11-395. PMID: 20650010; PMCID: PMC2918584.

Salem, M. A., Perez de Souza, L., Serag, A., Fernie, A. R., Farag, M. A., Ezzat, S. M., & Alseekh, S. (2020). Metabolomics in the Context of Plant Natural Products Research: From Sample Preparation to Metabolite Analysis. *Metabolites*, *10*(1), 37. <https://doi.org/10.3390/metabo10010037>

Schripsema J. Application of NMR in plant metabolomics: techniques, problems and prospects. *Phytochem Anal.* 2010 Jan-Feb;21(1):14-21. doi: 10.1002/pca.1185. PMID: 19904731.

Sedio, B.E., Boya, C.A., Echeverri, J.C.R. A protocol for high-throughput, untargeted forest community metabolomics using mass spectrometry molecular networks. **Applications in Plant Sciences**. v.6(3), ed.1033, 2018.

SUMNER, L. W.; MENDES, P.; DIXON, R. A. Plant metabolomics: large-scale phytochemistry in the functional genomics era. **Phytochemistry**, v. 62, p. 817–836, 2003.

Valli, M.; Russo, H. M.; Pilon, A. C.; Pinto, M. E. F.; Dias, N. B.; Freire, R. T.; Castro-Gamboa, I.; Bolzani V. S.; *Phys. Sci. Rev.* (2019), doi:10.1515/psr-2018-0167

Valli, M.; Russo, H. M.; Pilon, A. C.; Pinto, M. E. F.; Dias, N. B.; Freire, R. T.; Castro-Gamboa, I.; Bolzani V. S.; *Phys. Sci. Rev.* (2019), doi:10.1515/psr-2018-0167

Villas-Boas, S., Mas, S., Åkesson, J.S., Nielsen, J. Mass spectrometry in metabolome analysis. **Mass Spectrometry Reviews**, v. 24, p. 613-646, 2005.

Wang M et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nature Biotechnology* 34.8 (2016): 828-837. <https://doi.org/10.1038/nbt.3597>.

Wang, M., Carver, J., Phelan, V. *et al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. *Nat Biotechnol* **34**, 828–837 (2016). <https://doi.org/10.1038/nbt.3597>

Wang, M.X., Carver, J.J., Phelan, V.V., Sanchez, L.M., Garg, N., Peng, Y. *et.al.* Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. **Nature Biotechnology**, v.34, p. 828-837, 2016.

Wang, Xiaochuan Sun, Liwang Liu. Metabolomic analysis with GC-MS to reveal potential metabolites and biological pathways involved in Pb & Cd stress response of radish roots. **Scientific Reports**. v.5, p. 18296, 2015.

Watrous J.; Roach. P.; Alexandrov, T.; Heath B. S.; Yang, J. Y.; Kersten, R. D.; Van Der Voort, M.; Pogliano, K.; Gross, H.; Raaijmakers, J. M.; Moore, B.S.; Laskin, J.; Bandeira, N.; Dorrestein, P.C. *Proc. Natl. Acad. Sci. U S A*. v. 109, p. E1743-52, 2012.

Watrous J.; Roach. P.; Alexandrov, T.; Heath B. S.; Yang, J. Y.; Kersten, R. D.; Van Der Voort, M.; Pogliano, K.; Gross, H.; Raaijmakers, J. M.; Moore, B.S.; Laskin, J.; Bandeira, N.; Dorrestein, P.C. *Proc. Natl. Acad. Sci. U S A*. v. 109, p. E1743-52, 2012.

AL., Alan C. Pilon et. METABOLÔMICA DE PLANTAS: MÉTODOS E DESAFIOS. **Quim.Nova**, [S. l.], v. 43, n. 3, p. 329–354, 2020.

AUBERON, Florence; OLATUNJI, Opeyemi Joshua; RAMINOSON, Diamondra; MULLER, Christian D.; SOENGAS, Beatriz; BONTÉ, Frédéric; LOBSTEIN, Annelise. Isolation of novel stilbenoids from the roots of *Cyrtopodium paniculatum* (Orchidaceae). **Fitoterapia**, [S. l.], v. 116, p. 99–105, 2017. DOI: 10.1016/j.fitote.2016.11.015. Disponível em: <http://dx.doi.org/10.1016/j.fitote.2016.11.015>.

BOLYEN, Evan et al. Reproducible, interactive, scalable and extensible microbiome data science using QIIME 2. **Nature Biotechnology**, [S. l.], v. 37, n. 8, p. 852–857, 2019. DOI: 10.1038/s41587-019-0209-9.

CHAGAS-PAULA, Daniela A.; ZHANG, Tong; DA COSTA, Fernando B.; EDRADA-EBEL, Ru Angelie. A metabolomic approach to target compounds from the asteraceae family for dual COX and LOX inhibition. **Metabolites**, [S. l.], v. 5, n. 3, p. 404–430, 2015. DOI: 10.3390/metabo5030404.

CHAN, Weng Keong; TAN, Loh Teng Hern; CHAN, Kok Gan; LEE, Learn Han; GOH, Bey Hing. **Nerolidol: A sesquiterpene alcohol with multi-faceted pharmacological and biological activities**. [s.l: s.n.]. v. 21 DOI: 10.3390/molecules21050529.

CHANDRA DE, Lakshman; PATHAK, Promila; RAO, A. N.; RAJEEVAN, P. K. 12 Medicinal and Aromatic Orchids. **Commercial Orchids**, [S. l.], 2014. DOI: 10.2478/9783110426403.12.

DA SILVA, Ricardo R. et al. Propagating annotations of molecular networks using in silico fragmentation. **PLoS Computational Biology**, [S. l.], v. 14, n. 4, p. 1–26, 2018. DOI: 10.1371/journal.pcbi.1006089.

DAHMÉN, Jan; LEANDER, Kurt. A new triterpene glucoside from *Cymbidium giganteum*. **Phytochemistry**, [S. l.], v. 17, n. 11, p. 1975–1978, 1978. DOI: 10.1016/S0031-9422(00)88745-2.

DEMARQUE, Daniel P.; CROTTI, Antonio E. M.; VESSECCHI, Ricardo; LOPES, João L. C.; LOPES, Norberto P. Fragmentation reactions using electrospray ionization mass spectrometry: An important tool for the structural elucidation and characterization of synthetic and natural products. **Natural Product Reports**, [S. l.], v. 33, n. 3, p. 432–455, 2016. DOI: 10.1039/c5np00073d.

DUECKER, Fenja L.; REUSS, Franziska; HERETSCH, Philipp. Rearranged ergostane-type natural products: Chemistry, biology, and medicinal aspects. **Organic and Biomolecular Chemistry**, [S. l.], v. 17, n. 7, p. 1624–1633, 2019. DOI: 10.1039/c8ob02325e.

ERNST, Madeleine et al. Molnetenhancer: Enhanced molecular networks by integrating metabolome mining and annotation tools. **Metabolites**, [S. l.], v. 9, n. 7, 2019. DOI: 10.3390/metabo9070144.

FEUNANG, Yannick Djoumbou et al. ClassyFire: automated chemical classification with a comprehensive, computable taxonomy. **Journal of Cheminformatics**, [S. l.], v. 8, n. 1, p. 1–20, 2016. DOI: 10.1186/s13321-016-0174-y.

GONZÁLEZ-COLOMA, Azucena; LÓPEZ-BALBOA, Carmen; SANTANA, Omar; REINA, Matías; FRAGA, Braulio M. Triterpene-based plant defenses. **Phytochemistry Reviews**, [S. l.], v. 10, n. 2, p. 245–260, 2011. DOI: 10.1007/s11101-010-9187-8.

HABER, Ariela I.; SIMS, James W.; MESCHER, Mark C.; DE MORAES, Consuelo M.; CARR, David E. A key floral scent component (β -trans-bergamotene) drives pollinator preferences independently of pollen rewards in seep monkeyflower. **Functional Ecology**, [S. l.], v. 33, n. 2, p. 218–228, 2019. DOI: 10.1111/1365-2435.13246.

HUANG, Hsin; KUO, Yi Wei; CHUANG, Yu Chen; YANG, Ya Ping; HUANG, Li Min; JENG, Mei Fen; CHEN, Wen Huei; CHEN, Hong Hwa. Terpene Synthase-b and Terpene Synthase-e/f Genes Produce Monoterpenes for Phalaenopsis bellina Floral Scent. **Frontiers in Plant Science**, [S. l.], v. 12, n. July, 2021. DOI: 10.3389/fpls.2021.700958.

HUGHES, Nicole M.; CONNORS, Michaela K.; GRACE, Mary H.; LILA, Mary Ann; WILLANS, Brooke N.; WOMMACK, Andrew J. The same anthocyanins served four different ways: Insights into anthocyanin structure-function relationships from the wintergreen orchid, *Tipularia discolor*. **Plant Science**, [S. l.], v. 303, n. November 2020, p. 110793, 2021. DOI: 10.1016/j.plantsci.2020.110793. Disponível em: <https://doi.org/10.1016/j.plantsci.2020.110793>.

ISHMURATOV, G. Yu; VYDRINA, V. A.; GALKINA, Yu A.; YAKOVLEVA, M. P.; ISHMURATOVA, N. M.; TOLSTIKOV, A. G. Natural Seven-Membered Terpene Lactones: Synthesis and Biological Activity. **Chemistry of Natural Compounds**, [S. l.], v. 51, n. 6, p. 1011–1034, 2015. DOI: 10.1007/s10600-015-1483-5.

JIA, Min; CHEN, Ling; XIN, Hai Liang; ZHENG, Cheng Jian; RAHMAN, Khalid; HAN, Ting; QIN, Lu Ping. A friendly relationship between endophytic fungi and medicinal plants: A systematic review. **Frontiers in Microbiology**, [S. l.], v. 7, n. JUN, p. 1–14, 2016. DOI: 10.3389/fmicb.2016.00906.

KRÓL, Sylwia Katarzyna; KIEŁBUS, Michał; RIVERO-MÜLLER, Adolfo; STEPULAK, Andrzej. Comprehensive review on betulin as a potent anticancer agent. **BioMed Research International**, [S. l.], v. 2015, 2015. DOI: 10.1155/2015/584189.

LAM, Yau; NG, Tzi Bun; YAO, Ren Ming; SHI, Jun; XU, Kai; SZE, Stephen Cho Wing; ZHANG, Kalin Yanbo. Evaluation of chemical constituents and important mechanism of pharmacological biology in *Dendrobium* plants. **Evidence-based Complementary and Alternative Medicine**, [S. l.], v. 2015, 2015. DOI: 10.1155/2015/841752.

LIU, Li; CHENG, Xuexiang; ZHAO, Weiwei; WANG, Yunhao; DONG, Xiang; CHEN, Lili; ZHANG, Dangquan; PENG, Wanxi. Systematic characterization of volatile organic components and pyrolyzates from *Camellia oleifera* seed cake for developing high value-

added products. **Arabian Journal of Chemistry**, [S. l.], v. 11, n. 6, p. 802–814, 2018. DOI: 10.1016/j.arabjc.2017.12.031. Disponível em: <https://doi.org/10.1016/j.arabjc.2017.12.031>.

LIU, Weixin; FENG, Yi; YU, Suhang; FAN, Zhengqi; LI, Xinlei; LI, Jiyuan; YIN, Hengfu. The flavonoid biosynthesis network in plants. **International Journal of Molecular Sciences**, [S. l.], v. 22, n. 23, p. 1–18, 2021. DOI: 10.3390/ijms222312824.

LIU, Xiao Jing; CHUANG, Yao Nung; CHIOU, Chung Yi; CHIN, Dan Chu; SHEN, Fu Quan; YEH, Kai Wun. Methylation effect on chalcone synthase gene expression determines anthocyanin pigmentation in floral tissues of two *Oncidium* orchid cultivars. **Planta**, [S. l.], v. 236, n. 2, p. 401–409, 2012. DOI: 10.1007/s00425-012-1616-z.

MAJUMDER, P. L.; BANDYOPADHYAY, Saswati. Stilbenoids and sesquiterpene derivatives of the orchids *Gastrochilum calcoelaria* and *Dendrobium amoenum*: Application of 2D NMR spectroscopy in structural elucidation of complex natural products. **Journal of the Indian Chemical Society**, [S. l.], v. 87, n. 2, p. 221–234, 2010. DOI: 10.1002/chin.201051205.

MÜLLER, Marianne; BUCHBAUER, Gerhard. Essential oil components as pheromones. A review. **Flavour and Fragrance Journal**, [S. l.], v. 26, n. 6, p. 357–377, 2011. DOI: 10.1002/ffj.2055.

MUSHAROF HOSSAIN, Mohammad. Therapeutic orchids: Traditional uses and recent advances - An overview. **Fitoterapia**, [S. l.], v. 82, n. 2, p. 102–140, 2011. DOI: 10.1016/j.fitote.2010.09.007. Disponível em: <http://dx.doi.org/10.1016/j.fitote.2010.09.007>.

NETO, Fausto Carnevale; PILON, Alan C.; SELEGATO, Denise M.; FREIRE, Rafael T.; GU, Haiwei; RAFTERY, Daniel; LOPES, Norberto P.; CASTRO-GAMBOA, Ian. Dereplication of natural products using GC-TOF mass spectrometry: Improved metabolite identification by spectral deconvolution ratio analysis. **Frontiers in Molecular Biosciences**, [S. l.], v. 3, n. SEP, p. 1–13, 2016. DOI: 10.3389/fmolb.2016.00059.

NOTHIAS, Louis Felix et al. Feature-based molecular networking in the GNPS analysis environment. **bioRxiv**, [S. l.], n. October, 2019. DOI: 10.1101/812404.

NUNES, Carlos E. P.; GERLACH, Günter; BANDEIRA, Karen D. O.; GOBBO-NETO, Leonardo; PANSARIN, Emerson R.; SAZIMA, Marlies. Two orchids, one scent? Floral volatiles of *Catasetum cernuum* and *Gongora bufonia* suggest convergent evolution to a unique pollination niche. **Flora: Morphology, Distribution, Functional Ecology of Plants**, [S. l.], v. 232, p. 207–216, 2017. DOI: 10.1016/j.flora.2016.11.016. Disponível em: <http://dx.doi.org/10.1016/j.flora.2016.11.016>.

POLIMATI, Haritha; PRAGADA, Rajeswara Rao; THUAN, Nguyen Huy; TATIPAMULA, Vinay Bharadwaj. Hepatoprotective potential of bioflavonoids. **Studies in Natural Products Chemistry**, [S. l.], v. 72, p. 259–285, 2022. DOI: 10.1016/B978-0-12-823944-5.00014-4.

RAMYA, Mummadireddy; AN, Hye Ryun; BAEK, Yun Su; REDDY, Kondreddy Eswar; PARK, Pue Hee. Orchid floral volatiles: Biosynthesis genes and transcriptional regulations. **Scientia Horticulturae**, [S. l.], v. 235, n. March, p. 62–69, 2018. DOI: 10.1016/j.scienta.2017.12.049. Disponível em: <https://doi.org/10.1016/j.scienta.2017.12.049>.

RIVIÈRE, Céline; PAWLUS, Alison D.; MÉRILLON, Jean Michel. Natural stilbenoids: Distribution in the plant kingdom and chemotaxonomic interest in Vitaceae. **Natural Product Reports**, [S. l.], v. 29, n. 11, p. 1317–1333, 2012. DOI: 10.1039/c2np20049j.

ROY, Ram V.; SUMAN, Suman; DAS, Trinath P.; LUEVANO, Joe E.; DAMODARAN, Chendil. Withaferin a, a steroidal lactone from *withania somnifera*, induces mitotic catastrophe and growth arrest in prostate cancer cells. **Journal of Natural Products**, [S. l.], v. 76, n. 10, p. 1909–1915, 2013. DOI: 10.1021/np400441f.

SEMWAL, Ruchi Badoni; SEMWAL, Deepak Kumar; COMBRINCK, Sandra; VILJOEN, Alvaro. Emodin - A natural anthraquinone derivative with diverse pharmacological activities.

Phytochemistry, [S. l.], v. 190, n. July, p. 112854, 2021. DOI: 10.1016/j.phytochem.2021.112854. Disponível em: <https://doi.org/10.1016/j.phytochem.2021.112854>.

SHANNON, Paul; MARKIEL, Andrew; OWEN OZIER, NITIN S. BALIGA, JONATHAN T. WANG, Daniel Ramage; AMIN, Nada; BENNO SCHWIKOWSKI, And Trey Ideker. Cytoscape: a software environment for integrated models of biomolecular interaction networks. **Genome Research**, [S. l.], v. 13, n. 11, p. 504, 2003. DOI: 10.1101/gr.1239303.metabolite. Disponível em: <http://ci.nii.ac.jp/naid/110001910481/>.

SHEN, Tao; WANG, Xiao Ning; LOU, Hong Xiang. Natural stilbenes: An overview. **Natural Product Reports**, [S. l.], v. 26, n. 7, p. 916–935, 2009. DOI: 10.1039/b905960a.

SHIAO, Young Ji; CHEN, Wan Ping; LIN, Yun Lian. New polyphenols and triterpene from the pseudobulbs of pleione formosana. **Journal of the Chinese Chemical Society**, [S. l.], v. 56, n. 4, p. 828–833, 2009. DOI: 10.1002/jccs.200900122.

ŚLIWIŃSKI, Tomasz; KOWALCZYK, Tomasz; SITAREK, Przemysław; KOLANOWSKA, Marta. Orchidaceae-Derived Anticancer Agents: A Review. **Cancers**, [S. l.], v. 14, n. 3, p. 1–31, 2022. DOI: 10.3390/cancers14030754.

SUT, Stefania; MAGGI, Filippo; DALL'ACQUA, Stefano. Bioactive Secondary Metabolites from Orchids (Orchidaceae). **Chemistry and Biodiversity**, [S. l.], v. 14, n. 11, 2017. DOI: 10.1002/cbdv.201700172.

TEKA, Tekleab; LELE, Zhang; XIAOYAN, Ge; LI, Yanjie; LIFENG, Han; XIAOHUI, Yan. Stilbenes: Source plants, chemistry, biosynthesis, pharmacology, application and problems related to their clinical Application-A comprehensive review. **Phytochemistry**, [S. l.], v. 197, n. August 2021, p. 113128, 2022. DOI: 10.1016/j.phytochem.2022.113128. Disponível em: <https://doi.org/10.1016/j.phytochem.2022.113128>.

TOLLER-KAWAHISA, Juliana Escher; CANICOBA, Nathália Cristina; VENANCIO, Vinicius Paula; KAWAHISA, Rogério; ANTUNES, Lusânia Maria Gregg; CUNHA, Thiago Mattar; MARZOCCHI-MACHADO, Cleni Mara. Systemic lupus erythematosus onset in lupus-prone B6.MRL/lpr mice is influenced by weight gain and is preceded by an increase in neutrophil oxidative burst activity. **Free Radical Biology and Medicine**, [S. l.], v. 86, p. 362–373, 2015. DOI: 10.1016/j.freeradbiomed.2015.06.005. Disponível em: <http://dx.doi.org/10.1016/j.freeradbiomed.2015.06.005>.

VAN DER HOOFT, Justin Johan Jozias; WANDY, Joe; BARRETT, Michael P.; BURGESS, Karl E. V.; ROGERS, Simon. Topic modeling for untargeted substructure exploration in metabolomics. **Proceedings of the National Academy of Sciences of the United States of America**, [S. l.], v. 113, n. 48, p. 13738–13743, 2016. DOI: 10.1073/pnas.1608041113.

WANG, Mingxun et al. Sharing and community curation of mass spectrometry data with Global Natural Products Social Molecular Networking. **Nature Biotechnology**, [S. l.], v. 34, n. 8, p. 828–837, 2016. DOI: 10.1038/nbt.3597.

WANG, Yue Hu. Traditional uses, chemical constituents, pharmacological activities, and toxicological effects of Dendrobium leaves: A review. **Journal of Ethnopharmacology**, [S. l.], v. 270, n. January, p. 113851, 2021. DOI: 10.1016/j.jep.2021.113851. Disponível em: <https://doi.org/10.1016/j.jep.2021.113851>.

XIAO, Kai; ZHANG, Hong Jun; XUAN, Li Jiang; ZHANG, Juan; XU, Ya Ming; BAI, Dong Lu. Stilbenoids: Chemistry and bioactivities. **Studies in Natural Products Chemistry**, [S. l.], v. 34, n. C, p. 453–646, 2008. DOI: 10.1016/S1572-5995(08)80032-4. Disponível em: [http://dx.doi.org/10.1016/S1572-5995\(08\)80032-4](http://dx.doi.org/10.1016/S1572-5995(08)80032-4).

YE, Qinghua; QIN, Guowei; ZHAO, Weimin. Immunomodulatory sesquiterpene glycosides from Dendrobium nobile. **Phytochemistry**, [S. l.], v. 61, n. 8, p. 885–890, 2002. DOI: 10.1016/S0031-9422(02)00484-3.

YULIANA, Nancy Dewi; KHATIB, Alfi; CHOI, Young Hae; VERPOORTE, Robert. Metabolomics for bioactivity assessment of natural products. **Phytotherapy Research**, [S. l.], v. 25, n. 2, p. 157–169, 2011. DOI: 10.1002/ptr.3258.

ZHANG, Cong et al. Sesquiterpene amino ether and cytotoxic phenols from *Dendrobium wardianum* Warner. **Fitoterapia**, [S. l.], v. 122, n. August, p. 76–79, 2017. DOI: 10.1016/j.fitote.2017.08.015. Disponível em: <http://dx.doi.org/10.1016/j.fitote.2017.08.015>.

ZHAO, Weimin; YE, Qinghua; DAI, Jingqiu; ZHU, Jieping. Sesquiterpenes from *Dendrobium moniliforme*. [S. l.], 2003.

ZHUANG, Chunlin; ZHANG, Wen; SHENG, Chunquan; ZHANG, Wannian; XING, Chengguo; MIAO, Zhenyuan. Chalcone: A Privileged Structure in Medicinal Chemistry. **Chemical Reviews**, [S. l.], v. 117, n. 12, p. 7762–7810, 2017. DOI: 10.1021/acs.chemrev.7b00020.

Anexos

Tabela 1. Espécies brasileiras e exóticas de Orchidaceae utilizadas no estudo da metabolômica de espécies brasileiras de Orchidaceae.

Código	Espécies	País	Voucher	Subfamília	Tribo	Subtribo
Aap19	<i>Acianthera aphthosa</i> (Lindl.) Pridgeon & M.W. Chase	Brasil	LBMBP 258	Epidendroideae	Epidendreae	Pleurothallidinae
Aeb22	<i>Angraecum eburneum</i> Bory	Madagascar	LBMBP 627	Epidendroideae	Vandeae	Angraecinae
Agi97	<i>Arpophyllum giganteum</i> Hartw. ex Lindl.	México	LBMBP 246	Epidendroideae	Epidendreae	Laeliinae
Ale20	<i>Acianthera leptotifolia</i> (Barb. Rodr.) Pridgeon & M.W. Chase	Brasil	LBMBP 244	Epidendroideae	Epidendreae	Pleurothallidinae
Ava23	<i>Aspasia variegata</i> Lindl.	Brasil	LBMBP 654	Epidendroideae	Cymbidieae	Oncidiinae
Bch30	<i>Brassia chloroleuca</i> Barb. Rodr	Brasil	LBMBP 655	Epidendroideae	Cymbidieae	Oncidiinae
Bgr26	<i>Brasiliorchis gracilis</i> (Lodd.) R.B. Singer, S. Koehler & Carnevali	Brasil	LBMBP 674	Epidendroideae	Cymbidieae	Maxillariinae
Blo34	<i>Bulbophyllum lobbii</i> Lindl.	Indonésia	LBMBP 42	Epidendroideae	-	Dendrobiinae
Bma28	<i>Brasiliorchis marginata</i> (Lindl.) R.B. Singer, S. Koehler & Carnevali	Brasil	LBMBP 310	Epidendroideae	Cymbidieae	Maxillariinae
Bpi27	<i>Brasiliorchis picta</i> (Hook.) R.B. Singer, S. Koehler & Carnevali	Brasil	LBMBP 309	Epidendroideae	Cymbidieae	Maxillariinae
Bsa24	<i>Baptistonia sarcodes</i> (Lindl.) Chiron & V.P. Castro	Brasil	LBMBP 1074	Epidendroideae	Cymbidieae	Oncidiinae
Bte25	<i>Bifrenaria tetragona</i> (Lindl.) Schltr.	Brasil	LBMBP 1247	Epidendroideae	Cymbidieae	Maxillariinae
Btr44	<i>Bulbophyllum tripetalum</i> Lindl.	Brasil	LBMBP 1181	Epidendroideae	-	Dendrobiinae
Btu29	<i>Brassavola tuberculata</i> Hook.	Brasil	LBMBP 1235	Epidendroideae	Epidendreae	Laeliinae
Casp121	<i>Camaridium ochroleucum</i> Lindl.	Brasil	LBMBP s.n.	Epidendroideae	Cymbidieae	Maxillariinae
Cbi43	<i>Caularthron bicornutum</i> (Hook.) Raf.	Brasil	LBMBP 525	Epidendroideae	Epidendreae	Laeliinae
Ccr48	<i>Gomesa crispa</i> (Lindl.) Klotzsch ex Rchb.f.	Brasil	LBMBP s.n.	Epidendroideae	Cymbidieae	Oncidiinae
Cde46	<i>Cirrhaea dependens</i> (Lodd.) Loudon	Brasil	LBMBP 268	Epidendroideae	Cymbidieae	Stanhopeinae
Ceu51	<i>Cyrtopodium eugenii</i> Rchb.f. & Warm.	Brasil	LBMBP 354	Epidendroideae	Cymbidieae	Catasetinae
Cfi35	<i>Catasetum fimbriatum</i> (E. Morren) Lindl. & Paxton	Brasil	LBMBP 304	Epidendroideae	Cymbidieae	Catasetinae
Cfl120	<i>Coelogyne flaccida</i> Lindl.	Nepal	LBMBP 70	Epidendroideae	Arethuseae	Coelogyneinae
Cfo38	<i>Cattleya forbesii</i> Lindl.	Brasil	LBMBP 255	Epidendroideae	Epidendreae	Laeliinae
Cgi52	<i>Cyrtopodium gigas</i> Vell.	Brasil	LBMBP s.n.	Epidendroideae	Cymbidieae	Catasetinae
Cha42	<i>Cattleya harrisoniana</i> Bateman ex Lindl.	Brasil	LBMBP 419	Epidendroideae	Epidendreae	Laeliinae
Cha50	<i>Cycnoches haagii</i> Barb. Rodr.	Brasil	LBMBP 927	Epidendroideae	Cymbidieae	Catasetinae
Cin39	<i>Cattleya intermedia</i> Graham	Brasil	LBMBP 270	Epidendroideae	Epidendreae	Laeliinae

Cla124	<i>Cattleya labiata</i> Lindl.	Brasil	LBMBP 367	Epidendroideae	Epidendreae	Laeliinae
Clo45	<i>Cirrhaea longiracemosa</i> Hoehne	Brasil	LBMBP 219	Epidendroideae	Cymbidieae	Stanhopeinae
Clu40	<i>Cattleya lueddemanniana</i> Rchb.f.	Brasil	LBMBP 632	Epidendroideae	Epidendreae	Laeliinae
Cma36	<i>Catasetum mattosianum</i> Bicalho	Brasil	LBMBP 305	Epidendroideae	Cymbidieae	Catasetinae
Cmi33	<i>Campylocentrum micranthum</i> (Lindl.) Rolfe	Brasil	LBMBP 460	Epidendroideae	Vandaeae	Angraecinae
Cos131	<i>Catasetum osculatum</i> Lacerda & V.P.Castro	Brasil	LBMBP 920	Epidendroideae	Cymbidieae	Catasetinae
Csa47	<i>Cirrhaea saccata</i> Lindl.	Brasil	LBMBP 182	Epidendroideae	Cymbidieae	Stanhopeinae
Cte41	<i>Cattleya tenuis</i> Campacci & Vedov.	Brasil	LBMBP 369	Epidendroideae	Epidendreae	Laeliinae
Cva49	<i>Coppensia varicosa</i> (Lindl.) Campacci	Brasil	LBMBP 505	Epidendroideae	Cymbidieae	Oncidiinae
Cvi37	<i>Cattleya violacea</i> (Kunth) Rolfe	Brasil	LBMBP 276	Epidendroideae	Epidendreae	Laeliinae
Dco96	<i>Dendrochilum cobbianum</i> Rchb.f.	Filipinas	LBMBP 44	Epidendroideae	Arethuseae	Coelogyninae
Dfi53	<i>Dendrobium fimbriatum</i> Hook.	Himalaia	LBMBP 254	Epidendroideae	-	Dendrobiinae
Dki55	<i>Dendrobium kingianum</i> Bidwill ex Lindl.	Austrália	LBMBP 1225	Epidendroideae	-	Dendrobiinae
Dmo54	<i>Dendrobium moschatum</i> (Buch.-Ham.) Sw.	Tailândia	LBMBP 15	Epidendroideae	-	Dendrobiinae
Dsp56	<i>Dendrobium speciosum</i> Sm.	Austrália	LBMBP 308	Epidendroideae	-	Dendrobiinae
Eam119	<i>Elleanthus amethystinoides</i> Garay	Equador	LBMBP s.n.	Epidendroideae	Sobralieae	-
Eav64	<i>Epidendrum avicula</i> Lindl.	Brasil	LBMBP 344	Epidendroideae	Epidendreae	Laeliinae
Ede59	<i>Epidendrum denticulatum</i> Barb. Rodr.	Brasil	LBMBP 580	Epidendroideae	Epidendreae	Laeliinae
Ede60	<i>Epidendrum densiflorum</i> Hook.	Brasil	LBMBP s.n.	Epidendroideae	Epidendreae	Laeliinae
Eea31	<i>Eria hyacinthoides</i> (Blume) Lindl.	Malásia	LBMBP s.n.	Epidendroideae	Podochileae	Eriinae
Eel103	<i>Eria elata</i> Hook.f.	Malásia	LBMBP s.n.	Epidendroideae	Podochileae	Eriinae
Epu61	<i>Epidendrum puniceoluteum</i> F.Pinheiros & F.Barros	Brasil	LBMBP 596	Epidendroideae	Epidendreae	Laeliinae
Eri62	<i>Epidendrum rigidum</i> Jacq.	Brasil	LBMBP 851	Epidendroideae	Epidendreae	Laeliinae
Ese63	<i>Epidendrum secundum</i> Jacq.	Brasil	LBMBP 357	Epidendroideae	Epidendreae	Laeliinae
Esp57	<i>Encyclia spiritusanctensis</i> L.C. Menezes	Brasil	LBMBP 322	Epidendroideae	Epidendreae	Laeliinae
Eto58	<i>Encyclia tocantinenses</i> V.P.Castro & Campacci	Brasil	LBMBP 556	Epidendroideae	Epidendreae	Laeliinae
Etr91	<i>Epidendrum tridactylum</i> Lindl.	Brasil	LBMBP 340	Epidendroideae	Epidendreae	Laeliinae
Gam71	<i>Grobya amherstiae</i> Lindl.	Brasil	LBMBP 605	Epidendroideae	Cymbidieae	Catasetinae

Gbo129	<i>Guarianthe bowringiana</i> (O'Brien) Dressler & W. E. Higgins	Ecuador	LBMBP 233	Epidendroideae	Epidendreae	Laeliinae
Gcr69	<i>Gomesa crispa</i> (Lindl.) Klotzsch ex Rchb.f.	Brasil	LBMBP 1208	Epidendroideae	Cymbidieae	Oncidiinae
Gmo65	<i>Galeandra montana</i> Barb. Rodr.	Brasil	LBMBP s.n.	Epidendroideae	Cymbidieae	Catasetinae
Gqu70	<i>Gongora quinquenervis</i> Ruiz & Pav.	Brasil	LBMBP 176	Epidendroideae	Cymbidieae	Stanhopeinae
Hpa130	<i>Christensonella pachyphylla</i> . (Schltr. ex Hoehne) Szlach	Brasil	LBMBP s.n.	Epidendroideae	Cymbidieae	Maxillariinae
Hpu72	<i>Hadrolaelia purpurata</i> (Lindl. & Paxton) Chiron & V.P.Castro	Brasil	LBMBP 64	Epidendroideae	Epidendreae	Laeliinae
Hsu73	<i>Heterotaxis superflua</i> (Rchb.f.) F.Barros	Brasil	LBMBP s.n.	Epidendroideae	Cymbidieae	Maxillariinae
Ili76	<i>Isochilus linearis</i> (Jacq.) R.Br.	Brasil	LBMBP s.n.	Epidendroideae	Epidendreae	Ponerinae
Ipu75	<i>Isabelia pulchella</i> (Kraenzl.) Senghas & Teusch.	Brasil	LBMBP 662	Epidendroideae	Epidendreae	Laeliinae
Iut74	<i>Ionopsis utricularioides</i> (Sw.) Lindl.	Brasil	LBMBP 455	Epidendroideae	Cymbidieae	Oncidiinae
Kgr105	<i>Koellensteinia graminea</i> (Lindl.) Rchb.f.	Brasil	LBMBP s.n.	Epidendroideae	Maxillarieae	Zygopetalinae
Llu80	<i>Lockhartia lunifera</i> (Lindl.) Rchb.f.	Brasil	LBMBP 786	Epidendroideae	Cymbidieae	Oncidiinae
Lne79	<i>Liparis nervosa</i> (Thunb.) Lindl.	Brasil	LBMBP 736	Epidendroideae	Malaxideae	-
Lpe77	<i>Laelia perrinii</i> (Lindl.) Bateman	Brasil	LBMBP 1064	Epidendroideae	Epidendreae	Laeliinae
Lpo78	<i>Leptotes pohlitinocoi</i> V.P.Castro & Chiron	Brasil	LBMBP 672	Epidendroideae	Epidendreae	Laeliinae
Mca68	<i>Miltonia candida</i> Lindl.	Brasil	LBMBP 796	Epidendroideae	Cymbidieae	Oncidiinae
Mde81	<i>Mediocalcar decoratum</i> Schuit.	Brasil	LBMBP 1145	Epidendroideae	Podochileae	-
Mfl67	<i>Miltonia flavescens</i> Lindl.	Brasil	LBMBP 329	Epidendroideae	Cymbidieae	Oncidiinae
Mle122	<i>Maxillaria leucaimata</i> Barb.Rodr.	Brasil	LBMBP 527	Epidendroideae	Cymbidieae	Maxillariinae
Mlu128	<i>Microlaelia lundii</i> (Rchb.f. & Warm.) Chiron & V.P.Castro	Brasil	LBMBP s.n.	Epidendroideae	Epidendreae	Laeliinae
Mva82	<i>Maxillaria valenzuelana</i> (A.Rich.) Nash.	Brasil	LBMBP	Epidendroideae	Cymbidieae	dendro
Nam113	<i>Notylia amesii</i> L.B. Sm. & S.K. Harris	Brasil	LBMBP	Epidendroideae	Cymbidieae	Oncidiinae
Oal92	<i>Oncidium altissimum</i> (Jacq.) Sw.	Brasil	LBMBP 651	Epidendroideae	Cymbidieae	Oncidiinae
Oam123	<i>Orleanesia amazonia</i> Barb.Rodr.	Brasil	LBMBP 1239	Epidendroideae	Epidendreae	Laeliinae
Oce98	<i>Oncidium cebolleta</i> (Jacq.) Sw.	Brasil	LBMBP 638	Epidendroideae	Cymbidieae	Oncidiinae
Oco94	<i>Ornithidium coccineum</i> (Jacq.) Salisb. ex R. Br.	Brasil	LBMBP 490	Epidendroideae	Cymbidieae	Maxillariinae
Ocr114	<i>Octomeria crassifolia</i> Lindl.	Brasil	LBMBP 273	Epidendroideae	Epidendreae	Pleurothallidinae
Oma111	<i>Oeceoclades maculata</i> (Lindl.) Lindl	Brasil	LBMBP 482	Epidendroideae	Cymbidieae	Eulophiinae

Omo95	<i>Oncidium montanum</i> Barb. Rodr.	Brasil	LBMBP 884	Epidendroideae	Cymbidieae	Oncidiinae
Omy66	<i>Ornithocephalus myrticola</i> Lindl.	Brasil	LBMBP 384	Epidendroideae	Cymbidieae	Oncidiinae
Oon126	<i>Oncidium onustum</i> Lindl.	Brasil	LBMBP 477	Epidendroideae	Cymbidieae	Oncidiinae
Ora1	<i>Ornithophora radicans</i> (Rchb. f.) Garay & Pabst	Brasil	LBMBP 395	Epidendroideae	Cymbidieae	Oncidiinae
Osp93	<i>Oncidium sphacelatum</i> Lindl.	Brasil	LBMBP 618	Epidendroideae	Cymbidieae	Oncidiinae
Pca88	<i>Pseudolaelia canaanensis</i> (Ruschi) F. Barros	Brasil	LBMBP 426	Epidendroideae	Epidendreae	Laeliinae
Pde32	<i>Phymatidium delicatulum</i> Lindl.	Brasil	LBMBP 746	Epidendroideae	Cymbidieae	Oncidiinae
Pes101	<i>Polystachya estrellensis</i> Rchb.f.	Brasil	LBMBP 294	Epidendroideae	Vandaeae	Polystachyineae
Pfr104	<i>Prosthechea fragrans</i> (Sw.) W.E.Higgins	Brasil	LBMBP 526	Epidendroideae	Epidendreae	Laeliinae
Plu108	<i>Phalaenopsis lueddemanniana</i> Rchb.f.	Filipinas	LBMBP s.n.	Epidendroideae	Vandaeae	Aeridinae
Pmi106	<i>Paradisanthus micranthus</i> (Barb.Rodr.) Schltr.	Brasil	LBMBP	Epidendroideae	Cymbidieae	Zygopetalinae
Psp115	<i>Pleurothallis</i> sp.	Brasil	LBMBP 1005	Epidendroideae	Epidendreae	Pleurothallidinae
Pta84	<i>Phaius tankervilleae</i> (L'Her.) Blume	Brasil	LBMBP s.n.	Epidendroideae	-	Collabinae
Rco110	<i>Renanthera coccinea</i> Lour.	Brasil	LBMBP 871	Epidendroideae	Vandaeae	Aeridina
Rde90	<i>Rodriguezia decora</i> (Lem.) Rchb.f.	Brasil	LBMBP 779	Epidendroideae	Cymbidieae	Oncidiinae
Rgi100	<i>Rhynchostylis gigantea</i> (Lindl.) Ridl.	Brasil	LBMBP 1131	Epidendroideae	Vandaeae	Aeridinae
Rgl125	<i>Rhynchoaelia glauca</i> (Lindl.) Schltr.	Brasil	LBMBP 510	Epidendroideae	Epidendreae	Laeliinae
Rgo112	<i>Rodrigueziella gomezoides</i> Berman	Brasil	LBMBP 448	Epidendroideae	Cymbidieae	Oncidiinae
Rla118	<i>Rodriguezia lanceolata</i> Ruiz & Pav.	Brasil	LBMBP 443	Epidendroideae	Cymbidieae	Oncidiinae
Rpi109	<i>Rudolfiella picta</i> (Schltr.) Hoehne	Brasil	LBMBP 621	Epidendroideae	Cymbidieae	Maxillariinae
Scr85	<i>Schomburgkia crispa</i> Lindl.	Brasil	LBMBP	Epidendroideae	Epidendreae	Laeliinae
Sin87	<i>Stanhopea insignis</i> Frost ex W. Hook.	Brasil	LBMBP 229	Epidendroideae	Cymbidieae	Stanhopeinae
Smo86	<i>Scaphyglottis modesta</i> (Rchb.f.) Schltr.	Brasil	LBMBP 1220	Epidendroideae	Epidendreae	Laeliinae
Sno116	<i>Scuticaria novaesi</i> Barros & Cath.	Brasil	LBMBP 1137	Epidendroideae	Cymbidieae	Maxillariinae
Spl26	<i>Spathoglottis plicata</i> Blume	Filipinas	LBMBP 481	Epidendroideae	-	Collabinae
Ssp102	<i>Sobralia</i> sp.	Brasil	LBMBP 335	Epidendroideae	Sobralieae	-
Ssp21	<i>Stelis</i> sp.	Brasil	LBMBP 1072	Epidendroideae	Epidendreae	Pleurothallidinae
Tfa117	<i>Trizeuxis falcata</i> Lindl.	Brasil	LBMBP 392	Epidendroideae	Cymbidieae	Oncidiinae

Vsa99	<i>Vanda sanderiana</i> (Rchb.f.) Rchb.f.	Filipinas	LBMBP 890	Epidendroideae	Vandaeae	Aeridinae
Wwa107	<i>Warrea warreana</i> Lodd. ex Lindl.) C.Schweinf.	Brasil	LBMBP 1205	Epidendroideae	Cymbidieae	Zygopetalinae
Xsq83	<i>Xylobium squalens</i> (Lindl.) Lindl.	Brasil	LBMBP 256	Epidendroideae	Cymbidieae	Maxillariinae
Zcr89	<i>Zygopetalum crinitum</i> Lodd.	Brasil	LBMBP 21	Epidendroideae	Cymbidieae	Zygopetalinae
Ahy3	<i>Aspidogyne hylibates</i> (Rchb.f.) Garay	Brasil	LBMBP	Orchidoideae	Cranichideae	Goodyerinae
Cca5	<i>Cranichis candida</i> (Barb.Rodr.) Cogn.	Brasil	LBMBP 732	Orchidoideae	Cranichideae	Cranichidinae
Cco6	<i>Cyclopogon congestus</i> (Vell.) Hoehne	Brasil	LBMBP 743	Orchidoideae	Cranichideae	Spiranthinae
Mari12	<i>Microchilus arietinus</i> (Rchb.f. & Warm.) Ormerod	Brasil	LBMBP s.n.	Orchidoideae	Cranichideae	Goodyerinae
Mcu7	<i>Mesadenella cuspidata</i> (Lindl.) Garay	Brasil	LBMBP 745	Orchidoideae	Cranichideae	Spiranthinae
Pbo8	<i>Pelexia bonariensis</i> (Lindl.) Schltr.	Brasil	LBMBP 730	Orchidoideae	Cranichideae	Spiranthinae
Pro10	<i>Pteroglossa EROeoalba</i> (Rchb.f.) Salazar & M.W. Chase	Brasil	LBMBP 699	Orchidoideae	Cranichideae	Spiranthinae
Pvi9	<i>Pelexia vinosa</i> A. C. Ferreira, M. I. S. Lima & Pansarin	Brasil	LBMBP 712	Orchidoideae	Cranichideae	Spiranthinae
Ssh11	<i>Sarcoglottis schwackei</i> (Cogn.) Schltr.	Brasil	LBMBP 739	Orchidoideae	Cranichideae	Spiranthinae
Vba13	<i>Vanilla bahiana</i> Hoehne	Brasil	LBMBP 689	Vanilloideae	Vanillieae	-
Vch14	<i>Vanilla chamissonis</i> Klotzsch	Brasil	LBMBP 685	Vanilloideae	Vanillieae	-
Vha127	<i>Vanilla hartii</i> Rolfe	Peru	LBMBP 684	Vanilloideae	Vanillieae	-
Vpa16	<i>Vanilla paulista</i> (sp. nova)	Brasil	LBMBP 693	Vanilloideae	Vanillieae	-
Vpal15	<i>Vanilla palmarium</i> (Salzm. ex Lindl.) Lindl.	Brasil	LBMBP 697	Vanilloideae	Vanillieae	-
Vpo17	<i>Vanilla pompona</i> Schiede	Brasil	LBMBP 1007	Vanilloideae	Vanillieae	-
Vri18	<i>Vanilla ribeiroi</i> Hoehne	Brasil	LBMBP 687	Vanilloideae	Vanillieae	-
Psch4	<i>Phragmipedium schroederiae</i> (Lindl.) Rolfe	Brasil	LBMBP 239	Cypripedioideae	-	-
Pvi2	<i>Phragmipedium vittatum</i> (Vell.) Rolfe	Brasil	LBMBP 422	Cypripedioideae	-	-

Tabela 2. Compostos biomarcadores da propriedade antioxidantes dos extratos analisados de Orchidaceae brasileiras, obtidos nas análise estatística multivariada supervisionada PLS-DA (*Partial Least Square – Discriminant Analysis*) e anotados nas redes moleculares produzidas na plataforma GNPS.

Compostos	m/z	Tempo de Retenção	CF_Dparent	CF_subclass	CF_classe	CF_superclass
14-hydroxy-14-(hydroxymethyl)-5,9-dimethyltetracyclo[11.2.1.0.?.0?.?]hexadecane-5-carboxylic acid	335.2188	18.7931	Carboximidic acids	Carboximidic acids	Prenol lipids	Lipids and lipid-like molecules
3-Formyl-6-isopropylchromone	217.116	30.6176	Chromones	O-methylated flavonoids	Flavonoids	Phenylpropanoids and polyketides
Ferulic acid ethyl ester	177.0572	8.7163	Coumaric acids and derivatives	Hydroxycinnamic acids and derivatives	Cinnamic acids and derivatives	Phenylpropanoids and polyketides
3-Hydroxy-4-methoxycinnamic acid from NIST14	177.0572	12.4427	Coumaric acids and derivatives	nan	Benzene and substituted derivatives	Phenylpropanoids and polyketides
NCGC00385000-01_C15H18O8_beta-D-Glucopyranose, 1-O-[(2E)-3-(2-hydroxyphenyl)-1-oxo-2-propen-1-yl]-	309.0973	12.5725	Coumarin glycosides	Hydroxycinnamic acids and derivatives	Organooxygen compounds	Phenylpropanoids and polyketides
NCGC00385799-01_C28H36O10_2-Hydroxy-13,20-dimethoxy-4,7,17,22,22-pentamethyl-11-oxo-5,10,21,23-tetraoxahexacyclo[18.2.1.0~1,17~.0~4,16~.0~6,14~.0~8,12~]tricosan-6(14),7,12-trien-18-yl acetate	555.2387	8.4002	Coumarin glycosides	Hydroxycinnamic acids and derivatives	Organooxygen compounds	Phenylpropanoids and polyketides
NCGC00384905-01_C24H40O12_(2R,3E)-4-[(1S,6R)-1-Hydroxy-2,2,6-trimethyl-4-oxocyclohexyl]-3-buten-2-yl 6-O-beta-D-xylopyranosyl-beta-D-glucopyranoside	543.2397	8.661	Fatty acyl glycosides of mono- and disaccharides	Fatty acyl glycosides	Fatty Acyls	Lipids and lipid-like molecules
3',4',5,7-tetrahydroxyflavone	287.0546	10.4632	Flavonoid-7-O-glycosides	Flavonoid glycosides	Flavonoids	Phenylpropanoids and polyketides
5,9-dimethyltetracyclo[11.2.1.0.?.0?.?]hexadecane-5,14-dicarboxylic acid	333.2022	18.6159	Flavonoid-7-O-glycosides	Flavonoid glycosides	Flavonoids	Phenylpropanoids and polyketides
7-[(2S,3R,4S,5S,6R)-6-[[[(2S,3R,4R)-3,4-dihydroxy-4-(hydroxymethyl)oxolan-2-yl]oxymethyl]-3,4,5-trihydroxyoxan-2-yl]oxy-5-hydroxy-3-(4-hydroxyphenyl)-6-methoxychromen-4-one	595.1657	11.1601	Flavonoid-7-O-glycosides	Flavonoid glycosides	Flavonoids	Phenylpropanoids and polyketides
Hexanoside of (iso)vitexin	595.1657	10.8359	Flavonoid-7-O-glycosides	Flavonoid glycosides	Flavonoids	Phenylpropanoids and polyketides
Spectral Match to Peonidin 3-galactoside cation from NIST14	463.1234	11.6488	Flavonoid-7-O-glycosides	Flavonoid glycosides	Flavonoids	Phenylpropanoids and polyketides

p-Coumaric acid	165.0555	7.7234	Hydroxycinnamic acids	Amino acids, peptides, and analogues	Carboxylic acids and derivatives	Organic acids and derivatives
N-(Phenylacetyl)-L-phenylalanine	166.0876	4.1519	Phenylalanine and derivatives	Amino acids, peptides, and analogues	Carboxylic acids and derivatives	Organic acids and derivatives
trans-2-Hydroxycinnamic acid	165.0555	8.7153	Phenylpyruvic acid derivatives	Phenylpyruvic acid derivatives	Benzene and substituted derivatives	
14-(hydroxymethyl)-5,9-dimethyl-15-oxapentacyclo[11.3.1.0,?.0?,?.0?,?]heptadecane-5-carboxylic acid	333.2022	19.0934	Sesquiterpenoids	Sesquiterpenoids	Prenol lipids	Lipids and lipid-like molecules
3',4'-Dihydroxy-.beta.-naphthoflavone	305.1577	16.3955	Sesquiterpenoids	nan	Prenol lipids	Lipids and lipid-like molecules
8,8-dimethyl-2-phenylpyrano[2,3-f]chromen-4-one	305.1577	18.1553	Sesquiterpenoids	nan	Prenol lipids	Lipids and lipid-like molecules
Euparin	217.116	31.4298	TRITERPENO	nan	nan	nan
NCGC00385936-01!4-hydroxy-3-(3-methylbut-2-enyl)benzoic acid	207.1107	15.0979	Xanthenes	1-benzopyrans	Benzopyrans	Organoheterocyclic compounds
Tricin methyl ether	345.0965	14.5436	Xanthenes	nan	Flavonoids	Organoheterocyclic compounds

Tabela 3. Compostos biomarcadores da propriedade moduladora da produção de ERO por neutrófilos humanos dos extratos analisados de Orchidaceae brasileiras, obtidos nas análise estatística multivariada supervisionada PLS-DA (*Partial Least Square – Discriminant Analysis*) e anotados nas redes moleculares produzidas na plataforma GNPS.

Compound_Name	m/z	Tempo de Retenção	CF_Dparent	CF_class	CF_subclass	CF_superclass
Spectral Match to Alpinetin from NIST14	271.096	14.5528	2'-Hydroxychalcones	Linear 1,3-diarylpropanoids	Chalcones and dihydrochalcones	Phenylpropanoids and polyketides
Isobavachin - 10eV	323.109	5.7848	8-prenylated flavanones	Flavonoids	Flavans	Phenylpropanoids and polyketides
Isobavachin - 10eV	323.11	6.6818	8-prenylated flavanones	Flavonoids	Flavans	Phenylpropanoids and polyketides
Oleoyl 3-carbacyclic phosphatidic acid	439.317	31.9736	Cyclic peptides	Prenol lipids	nan	Lipids and lipid-like molecules
Microsporin B	537.303	20.1173	Cyclic peptides	Prenol lipids	nan	Lipids and lipid-like molecules
NCGC00179861-02_C15H26O2_1,8-Azulenediol, 1,2,3,3a,4,7,8,8a-octahydro-3a,6-dimethyl-1-(1-methylethyl)-, (1R,3aR,8S,8aS)-	221.201	14.1294	Diterpenoids	Prenol lipids	nan	Lipids and lipid-like molecules
NCGC00168965-02_C21H32O11_beta-D-Glucopyranoside, 3-(4-hydroxyphenyl)-1-methylpropyl 6-O-[(2S,3R,4R)-tetrahydro-3,4-dihydroxy-4-(hydroxymethyl)-2-furanyl]-	483.219	11.3401	Fatty acyl glycosides of mono- and disaccharides	Fatty Acyls	Fatty acyl glycosides	Lipids and lipid-like molecules
NCGC00381276-01_C22H34O14_Hexitol, 2,6-anhydro-, 5-[4-(beta-D-glucopyranosyloxy)-2-methylenebutanoate] 1-(4-hydroxy-2-methylenebutanoate)	545.199	13.6206	Fatty acyl glycosides of mono- and disaccharides	Fatty Acyls	Fatty acyl glycosides	Lipids and lipid-like molecules
NCGC00180129-02!5-hydroxy-2-(4-hydroxyphenyl)-8-[3,4,5-trihydroxy-6-(hydroxymethyl)oxan-2-yl]-7-(3,4,5-trihydroxy-6-methyloxan-2-yl)oxychromen-4-one	579.169	10.2665	Flavonoid-7-O-glycosides	Flavonoids	Flavonoid glycosides	Phenylpropanoids and polyketides
NCGC00170012-03!5-hydroxy-3-[(2S,3R,4S,5S)-4-hydroxy-5-(hydroxymethyl)-3-[(2S,3R,4R,5R,6S)-3,4,5-trihydroxy-6-methyloxan-2-yl]oxyoxolan-2-yl]oxy-2-(4-hydroxyphenyl)-7-methoxychromen-4-one	579.17	12.157	Flavonoid-7-O-glycosides	Flavonoids	Flavonoid glycosides	Phenylpropanoids and polyketides
BUDDLEOFLAVONOLOSIDE	593.186	12.528	Flavonoid-7-O-glycosides	Flavonoids	Flavonoid glycosides	Phenylpropanoids and polyketides

NCGC00169522-02_C27H34O11_4-[[{(3R,4R)-4-(3,4-Dimethoxybenzyl)-2-oxotetrahydro-3-furanyl]methyl]-2-methoxyphenyl beta-D-glucopyranoside	557.256	9.1965	Lignan glycosides	Lignan glycosides	nan	Lignans, neolignans and related compounds
(+)-Fluprostenol	457.134	6.4837	no matches	no matches	no matches	no matches
1-(9Z-Octadecenoyl)-sn-glycero-3-phosphocholine	522.353	22.615	no matches	no matches	no matches	no matches
5,9-dimethyltetracyclo[11.2.1.0,?.0?,?]hexadecane-5,14-dicarboxylic acid	333.202	21.1089	no matches	no matches	no matches	no matches
Strychnine	335.217	18.6441	no matches	no matches	no matches	no matches
14-hydroxy-14-(hydroxymethyl)-5,9-dimethyltetracyclo[11.2.1.0,?.0?,?]hexadecane-5-carboxylic acid	335.2173	20.4796	no matches	no matches	no matches	no matches
Cortodoxone	347.22	15.1967	no matches	no matches	no matches	no matches
Estrone sulfate	349.09	7.3922	no matches	no matches	no matches	no matches
Rosmarinic acid	359.131	2.408	no matches	no matches	no matches	no matches
4'-Benzyloxy-5,7-dimethoxyflavone	389.175	7.247	no matches	no matches	no matches	no matches
MeO-Glu(NMe2)-Val-hexanoate	408.186	3.4575	no matches	no matches	no matches	no matches
Aminopterin	441.169	7.7807	no matches	no matches	no matches	no matches
Cyclohexylaminopropanesulfonic acid	443.156	15.7446	no matches	no matches	no matches	no matches
NCGC00385099-01_C19H28O11_beta-D-Glucopyranoside, 2-(4-hydroxyphenyl)ethyl 2-O-[(2S,3R,4R)-tetrahydro-3,4-dihydroxy-4-(hydroxymethyl)-2-furanyl]-	455.117	1.4929	no matches	no matches	no matches	no matches
15-formyl-16-hydroxy-1,2,14,17,17-pentamethyl-8-(prop-1-en-2-yl)pentacyclo[11.7.0.0,?.0?,?.0?,?]icosane-5-carboxylic acid	469.326	28.9328	no matches	no matches	no matches	no matches
Dexamethasone phosphate	471.145	6.2908	no matches	no matches	no matches	no matches
Steviolbioside	479.299	14.1011	no matches	no matches	no matches	no matches
NCGC00380271-01_C24H40O11_2-Cyclohexen-1-one, 3,5,5-trimethyl-4-[3-[(6-O-beta-D-xylopyranosyl-beta-D-glucopyranosyl)oxy]butyl]-	527.246	11.8942	no matches	no matches	no matches	no matches
NCGC00347731-02_C26H32O12_beta-D-Glucopyranoside, 2-methoxy-4-[(1S,3aS,4R,6aR)-tetrahydro-3a-hydroxy-4-(4-hydroxy-3-methoxyphenyl)-1H,3H-furo[3,4-c]furan-1-yl]phenyl	559.178	12.451	no matches	no matches	no matches	no matches

Spectral Match to Neodiosmin from NIST14	609.181	11.9022	no matches	no matches	no matches	no matches
trans-2-Hydroxycinnamic acid	165.054	10.937	Phenylpyruvic acid derivatives	Benzene and substituted derivatives	Phenylpyruvic acid derivatives	Organoheterocyclic compounds
3-Quinolinecarboxamide, 1,4-dihydro-8-methoxy-4-oxo-1-pentyl-N-tricyclo[3.3.1.1 ^{3,7}]dec-1-yl-	445.204	11.4503	Quinoline-3-carboxamides	Quinolines and derivatives	Quinoline carboxamides	Organoheterocyclic compounds
FARNESOL	261.219	18.3206	Sesquiterpenoids	Prenol lipids	nan	Lipids and lipid-like molecules
19(R)-Hydroxyprostaglandin F2.alpha.	393.286	28.4799	Sesquiterpenoids	Prenol lipids	nan	Lipids and lipid-like molecules



UNIVERSIDADE DE SÃO PAULO
Faculdade de Ciências Farmacêuticas de Ribeirão Preto
Comitê de Ética em Pesquisa

Of. CEP/FCFRP nº. 007/2020
kms

Ribeirão Preto, 22 de maio de 2020.

À pesquisadora
Camila Capel Godinho
FCFRP/USP

Prezada Pesquisadora,

Informamos que o protocolo de pesquisa intitulado “**Estudos Metabolômicos de Espécies Brasileiras de Orchidaceae**” (CAAE: 31758720.3.0000.5403 – Protocolo CEP/FCFRP nº 542), foi aprovado pelo Comitê de Ética em Pesquisa da FCFRP em sua 196ª reunião ordinária realizada em 22/04/2020, conforme Parecer Consubstanciado nº 4.043.526.

Lembramos que, de acordo com a Resolução 466/2012, item IV.5, letra d, o TCLE deverá “ser elaborado em duas vias, rubricadas em todas as suas páginas e assinadas, ao seu término, pelo convidado a participar da pesquisa, ou por seu representante legal, assim como pelo pesquisador responsável, ou pela(s) pessoa(s) por ele delegada(s), devendo as páginas de assinaturas estar na mesma folha. Em ambas as vias deverão constar o endereço e contato telefônico ou outro, dos responsáveis pela pesquisa e do CEP local”.

Informamos que deverá ser encaminhado ao CEP o relatório final da pesquisa em formulário próprio deste Comitê, bem como comunicada qualquer alteração, intercorrência ou interrupção do mesmo, tais como eventos adversos e eventuais modificações no protocolo ou nos membros da equipe, através da interposição de emenda na Plataforma Brasil.

Atenciosamente,

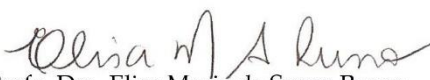

Profa. Dra. Elisa Maria de Sousa Russo
Vice-Coordenadora do CEP/FCFRP

Figura 1. Termo de aprovação do Comitê de Ética em Pesquisa da FCFRP para utilização de células humanas (neutrófilos) neste estudo.

