

UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

Metabolômica para análise de cultivares, quimiotaxonomia e triagem de inibidores de lactato desidrogenase-5 (LDH-5) em espécies de Asteraceae

Felipe Antunes dos Santos

Ribeirão Preto

2019

Felipe Antunes dos Santos

Metabolômica para análise de cultivares, quimiotaxonomia e triagem de inibidores de lactato desidrogenase-5 (LDH-5) em espécies de Asteraceae

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(a): Prof. Dr. Fernando Batista da Costa

Ribeirão Preto

2019

LOMBADA

<p>SANTOS, F.A.</p>	<p>Metabólica para análise de cultivares, quimiotaquonomia e triagem de inibidores de lactato desidrogenase-5 (LDH-5) em espécies de Asteraceae</p>		<p>DOUTORADO FCFRPUSP 2019</p>
--------------------------------	--	--	---

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.

Santos, Felipe Antunes

Metabolômica para análise de cultivares, quimiotaonomia e triagem de inibidores de lactato desidrogenase-5 (LDH-5) em espécies de Asteraceae. Ribeirão Preto, 2019, 178 p., 30 cm.

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

Orientador: Da Costa, Fernando Batista

1. Metabolômica. 2. Quimiometria. 3. *Decision Tree*. 4. Asteraceae. 5. Quimiotaonomia. 6. Germinação. 7. Cultivo. 8. Lactato desidrogenase-5 (LDH-5). 9. Sazonalidade

FOLHA DE APROVAÇÃO

Felipe Antunes dos Santos

Metabolômica para análise de cultivares, quimiotaxonomia e triagem de inibidores de lactato desidrogenase-5 (LDH-5) em espécies de Asteraceae

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: _____

Em tempos sombrios, com constantes cortes orçamentários nas áreas da ciência, tecnologia e inovação brasileiras, é mais que um nobre desafio realizar pesquisa de qualidade que possa contribuir com o desenvolvimento nacional...

Dedico aos pós-graduandos do Brasil, em especial aos bolsistas CAPES, CNPq e não bolsistas...

Agradecimentos

Primeiramente, agradeço a Deus, por estar em todas as coisas, emoções, sentimentos, natureza, que me fazem bem e me deram forças para prosseguir com este trabalho.

Agradeço à minha família, em especial a meus pais, José Antonio R. dos Santos e Sebastiana Lopes A. R. dos Santos, e ao meu irmão, Fabrício A. dos Santos. Vocês sempre me enxergam muito melhor do que sou. São o meu porto seguro. Desejam sempre o melhor para mim. Me auxiliaram diversas vezes com apoio material, compreensão e carinho. Agradeço também à Rose, não somente pelo carinho, compreensão e apoio, mas por também ter me ajudado laboralmente no início deste meu trabalho, peneirando a terra para o cultivo.

Agradeço ao meu orientador, Prof. Dr. Fernando Batista da Costa, primeiramente por ter aceitado me orientar por todos estes anos. Agradeço pela paciência, ensinamentos, conselhos, apoio e amizade.

Agradeço de maneira geral a todos do laboratório de farmacognosia e do grupo Asterbiochem, pela torcida, apoio, desejos de boa sorte... Em especial à Debora e Ingrid, por serem colegas desde a graduação, ao Lucas e Anny, pelo auxílio com os dados de LC-MS das espécies de Asteraceae, ao Gari, pelo auxílio com o *speed-vac*, ao Victor, pelo auxílio com o liofilizador, ao Jolindo, pelos momentos que se dispôs a me ajudar, à Rosana, pelas dicas com ensaios e ajuda no processamento de dados do LC-MS, à Marina, pela valiosa ajuda nos ensaios de inibição enzimática, ao Luizão, pela aquisição de dados do LC-MS, à Gelly, pelo carinho, apoio e dedicação, ao Mario, pelo acessoramento na farmacognosia, à Prof. Simone, por ter me permitido realizar os estudos de germinação em seu laboratório e me permitido usar a estufa do horto medicinal, ao Anderson, aluno da Profa Luciane do lab. de bioquímica, pelas ajudas iniciais com os ensaios, ao Jony, também aluno da Profa Luciane, por se dispor diversas vezes em me ajudar e me dar dicas nos ensaios, ao Emerson, do lab. do Prof. Sergio Akira, por ter me dado suporte ao usar a leitora de placas para os ensaios. Agradeço à Profa Luciane e ao Prof. Sergio Akira por me permitirem usar a leitora de placas para realizar meus experimentos com a enzima LDH. Agradeço ao valioso apoio que o Prof. Cid, da UFPR, deu à coleta de sementes e folhas de *A. trichophylla*.

Agradeço à Faculdade de Ciências Farmacêuticas de Ribeirão Preto – USP, por todo apoio e suporte ao desenvolvimento deste trabalho.

Agradeço à FAPESP (processos 2010/51454-3 e 2014/26866-7) pelo financiamento das análises no LC-MS e coletas de material vegetal.

Agradeço ao CNPq (processo 141229/2015-2) pela bolsa concedida.

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

“Não deve haver limites para o esforço humano. Somos todos diferentes. Por pior do que a vida possa parecer, sempre há algo que podemos fazer em que podemos obter sucesso. Enquanto houver vida, haverá esperança.”

Stephen Hawking

RESUMO

SANTOS, F. A. **Metabolômica para análise de cultivares, quimiotaonomia e triagem de inibidores de lactato desidrogenase-5 (LDH-5) em espécies de Asteraceae** 2019. 178 f. Tese (Doutorado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2019.

Metabolômica pode ser definida como o estudo de todos os metabólitos de um dado sistema biológico. Tal estudo possibilita abordagens holísticas e integradas de diversos ramos da ciência. Deste modo, problemas em delimitações taxonômicas, análise da influência ambiental sobre plantas, e triagem de metabólitos bioativos constituem exemplos de áreas onde a metabolômica pode atuar. A família Asteraceae, com aproximadamente 25.000 espécies, tem sido alvo de estudos taxonômicos por décadas. Assim sendo, a abordagem metabolômica aplicada à quimiotaonomia pode auxiliar em tais classificações, ou revelar novas descobertas. Por outro lado, discute-se que os metabólitos secundários sofrem muita influência de diversos fatores ambientais. Deste modo, este trabalho investigou também a influência da sazonalidade, luminosidade, adubação e ritmo circadiano no metaboloma de cultivares de *Aldama robusta* e *A. trichophylla* (Asteraceae). O cultivo de *A. robusta* foi também estudado quanto à sua germinação e desenvolvimento morfológico. A ampla variabilidade química em Asteraceae também a coloca como uma família com potencial para a triagem de metabólitos bioativos. Um alvo bioquímico relevante é a enzima lactato desidrogenase-5 (LDH-5), a qual esta relacionada com o aumento da sobrevivência de células cancerígenas. Deste modo, o estudo metabolômico em espécies de Asteraceae pode possibilitar a descoberta de substâncias ativas contra determinado alvo. O metaboloma de todas as espécies deste trabalho foi adquirido por UHPLC-ESI-UV/DAD-MS, com analisador Orbitrap™. Para a quimiotaonomia e estudos de inibição da LDH-5, foram estudadas 250 espécies de Asteraceae, provenientes de diversos gêneros de 16 tribos. Para os estudos da influência de determinados fatores ambientais no metaboloma de *A. robusta* e *A. trichophylla*, foram realizadas coletas em diferentes ritmos circadianos, estações do ano e diferentes cultivares com base na adubação do solo e luminosidade. Os dados químicos foram analisados por diversas técnicas quimiométricas e de *data mining* (mineração de dados), tais como PCA, HCAbp, OPLS-DA e *decision tree*. Deste modo, a análise por OPLS-DA com espécies de Asteraceae foi capaz de separar algumas tribos e apontar metabólitos discriminantes em cada uma delas, não obstante a PCA e HCAbp não tenham revelado agrupamento. Do mesmo modo, a OPLS-DA e a *decision tree* foram capazes de apontar alguns prováveis metabólitos correlacionados com extratos considerados “ativos” contra a LDH-5, bem como realizar a predição de extratos provenientes de outras espécies de Asteraceae que não haviam ainda sido ensaiados. Já os estudos com a germinação e o desenvolvimento dos cultivares pode revelar condições mais adequadas para o cultivo de *A. robusta* e *A. trichophylla*. Além disso, a metabolômica de tais cultivares também foi capaz de apontar acúmulo de metabólitos a depender da estação climática. Isso sugere uma necessidade maior em se padronizar a data da coleta do material vegetal para estudos quimiotaonômicos. Em breve, os cultivares terão suas atividades preditas frente à LDH-5, assim como serão isolados e ensaiados os metabólitos que foram considerados ativos nas espécies de Asteraceae.

Palavras-chave: Metabolômica, Quimiometria, *Decision Tree*, Asteraceae, Quimiotaonomia, Germinação, Cultivo, Lactato Desidrogenase-5 (LDH-5), Estacional.

ABSTRACT

SANTOS, F. A. **Metabolomics for cultivar analysis, chemotaxonomy and screening of lactate dehydrogenase-5 (LDH-5) inhibitors in Asteraceae species**. 2019. 178 f. Thesis (Doctoral). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2019.

Metabolomics can be defined as the study of all the metabolites of a given biological system. Such a study enables holistic and integrated approaches from several fields of science. Thus, problems in taxonomic delimitations, analysis of environmental influence on plants, and screening of bioactive metabolites are examples of areas where metabolomics can aid. The Asteraceae family, with approximately 25,000 species, has been the subject of taxonomic studies for decades. Thus, the metabolomics approach applied to chemotaxonomy may aid in such classifications, or reveal new findings. On the other hand, it is argued that secondary metabolites are influenced by many environmental factors. In this way, this work also investigated the influence of seasonality, luminosity, fertilization and circadian rhythm on the metabolome of cultivars of *Aldama robusta* and *A. trichophylla* (Asteraceae). The cultivation of *A. robusta* was also studied for its germination and morphological development. The wide chemical variability in Asteraceae also places it as a family with potential for the screening of bioactive metabolites. A relevant biochemical target is the enzyme lactate dehydrogenase-5 (LDH-5), which is related to the increased survival of cancer cells. In this way, the metabolomic study in Asteraceae species may allow the discovery of active compounds against LDH-5. The metabolome of all species of this work was acquired by UHPLC-ESI-UV/DAD-MS, with Orbitrap™ analyzer. For the chemotaxonomy and LDH-5 inhibition studies, 250 species of Asteraceae, from several genera of 16 tribes, were studied. For the studies of the influence of certain environmental factors on the metabolome of *A. robusta* and *A. trichophylla*, collections were made at different circadian rhythms, seasons of the year and different cultivars based on soil fertilization and luminosity. The chemical data were analyzed by several chemometric and data mining techniques, such as PCA, HCAbp, OPLS-DA and decision tree. Thus, OPLS-DA analysis with Asteraceae species was able to separate some tribes and point out discriminant metabolites in each of them, although PCA and HCAbp did not reveal clustering. Likewise, OPLS-DA and the decision tree were able to point out some probable metabolites correlated with extracts considered "active" against LDH-5, as well as to predict extracts from other Asteraceae species that had not yet been assayed. As for the germination and development studies of cultivars, more adequate conditions for the cultivation of *A. robusta* and *A. trichophylla* were described. In addition, the metabolomics of such cultivars were also able to indicate accumulation of metabolites depending on the climatic season. This suggests a greater need to standardize the date of collection of plant material for chemotaxonomic studies. Soon, the cultivars will have their activities predicted against the LDH-5, as well as the metabolites that were considered active in the species of Asteraceae will be isolated and tested.

Keywords: Metabolomics, Chemometrics, Decision Tree, Asteraceae, Chemotaxonomy, Germination, Cultivation, Lactate Dehydrogenase-5 (LDH-5), Seasonal.

LISTA DE FIGURAS

Figura 1. Comparação de diferentes softwares de processamento de dados (Extraído de Castillo et al., 2011).....	05
Figura 2. Árvore filogenética de Asteraceae construída com base em dados de DNA e inferência Bayesiana. Figura extraída de Panero e colaboradores (PANERO et al., 2014).....	10
Figura 3. Diagrama da família Asteraceae construído com base em caracteres morfológicos (BREMER et al., 1996). Figura oriunda de Calabria e coautores (CALABRIA et al., 2006).....	11
Figura 4. Algumas das principais classes químicas de metabólitos secundários em Asteraceae.....	12
Figura 5. Exemplo da reprodutibilidade da análise destacando padrões internos.....	28
Figura 6. Inspeção das réplicas de ID 14, com os respectivos padrões internos, cafeína e daidzeína.....	28
Figura 7. Inspeção das réplicas de ID 159, com os respectivos padrões internos, cafeína e daidzeína.....	29
Figura 8. Inspeção de réplicas de extratos provenientes de espécies das tribos Eupatorieae, Vernonieae, Astereae e Millerieae.....	30
Figura 9. PCA (4 componentes, 0,29) com espécies de Asteraceae das tribos Heliantheae, Eupatorieae, Vernonieae, Astereae, Barnadesieae e Millerieae.....	32
Figura 10. Gráfico de erro padrão por valor-p AU.....	33
Figura 11. HCAbp realizado com algumas espécies das tribos Barnadesieae, Eupatorieae, Astereae, Millerieae, Vernonieae e Heliantheae.....	34
Figura 12. OPLS-DA.1 (3 + 7) ($R^2 = 0,90$ e $Q^2 = 0,37$) mostrando um agrupamento das espécies de Asteraceae com base no metaboloma das tribos Barnadesieae, Heliantheae, Eupatorieae e Vernonieae.....	36
Figura 13. OPLS-DA.2 (4 + 3) ($R^2 = 0,93$ e $Q^2 = 0,78$) mostrando um agrupamento das espécies de Asteraceae com base no metaboloma das tribos Astereae, Millerieae, Eupatorieae e Vernonieae.....	38

Figura 14. OPLS-DA.3 (4 + 3) ($R^2 = 0,92$ e $Q^2 = 0,79$) mostrando um agrupamento das espécies de Asteraceae com base no metaboloma das tribos Astereae, Milleriae, Heliantheae e Vernonieae.....	48
Figura 15. Otto Henrich Warburg, quem descreveu o “efeito Warburg” das células cancerígenas. Figura retirada da página Wikipedia.....	67
Figura 16. Obtenção de energia pela via da fosforilação oxidativa ou glicolítica. Esquema retirado de Granchi et al., 2011.....	67
Figura 17. OPLS-DA (1 + 7) ($R^2 = 0,99$ e $Q^2 = 0,77$) para extratos foliares de espécies de Asteraceae ensaiados contra LDH-5. Foram considerados ativas (active) as amostras que inibiram mais de 45% da atividade enzimática, as outras amostras foram consideradas inativas (no activity).....	78
Figura 18. <i>Decision tree</i> realizada no Weka com base no metaboloma de 75% das espécies (grupo treino) de Asteraceae ensaiadas contra a enzima LDH-5.....	94
Figura 19. <i>Decision tree</i> com base no metaboloma de 75% das espécies (grupo treino) de Asteraceae ensaiadas contra a enzima LDH-5 após aplicar o filtro supervisionado “attribute selection” do Weka às variáveis.....	94
Figura 20. Esquema mostrando as dimensões dos canteiros sem e com sombrite 35% construídos. Figura de minha autoria.....	109
Figura 21. Canteiro construído com sombrite 35%. À esquerda plantas adubadas, à direita, plantas sem adubo. Figura de minha autoria.....	117
Figura 22. Canteiro construído sem sombrite. À esquerda plantas adubadas, à direita, plantas sem adubo. Figura de minha autoria.....	118
Figura 23. Sobreposição das amostras correspondentes ao pool 1. A substância que elui em ap. 15 min. corresponde ao padrão interno crisina.....	131
Figura 24. Sobreposição das amostras correspondentes ao pool 1. A substância que elui em ap. 15 min. corresponde ao padrão interno crisina.....	132
Figura 25. TIC das triplicatas de branco injetadas antes e durante as análises. Não é possível visualizar nenhuma substância de concentração significativa que possa comprometer a análise.....	132
Figura 26. Controle de qualidade metabolômico inspecionado pelo agrupamento de pools.....	133
Figura 27. PCA com pools e todas as amostras provenientes das coletas dos cultivos. Para visualizar o código dos cultivares, checar Apêndice B.....	134

Figura 28. OPLS-DA mostrando diferença entre as coletas das 11 horas realizadas com N ₂ líquido (nitrogen) versus secamento por estufa (air oven). OPLS-DA (1 + 6): R ² = 0,96, Q ² = 0,79.....	135
Figura 29. PCA (R ² = 0,50) mostrando tendência de agrupamento das amostras coletadas no inverno (Winter) e primavera (Spring) contra as amostras coletadas no verão (Summer) e outono (Autumn).....	139
Figura 30. OPLS-DA (3 + 10) (R ² = 0,95 e Q ² = 0,69) mostrando clara distinção entre amostras coletadas no inverno (winter) e primavera (spring) em relação às coletas do verão (summer) e outono (autumn).....	140
Figura 31. OPLS-DA (2 + 10) (R ² = 0,94 e Q ² = 0,78) com amostras coletadas no verão (summer) e outono (autumn) consideradas um único grupo indistinto.....	141
Figura 32. TIC dos padrões para curva de calibração.....	145
Figura 33. Curva de calibração para 9 substâncias com base nas suas concentrações e intensidade de ionização registrada pela área do TIC.....	146

LISTA DE TABELAS E QUADROS

Tabela 1. Classificação das espécies das tribos Heliantheae, Eupatorieae, Vernonieae e Barnadesieae, segundo o modelo OPLS-DA-1 (Figura 12), construído a partir do metaboloma de tribos de Asteraceae, utilizado para encontrar variáveis discriminantes de Barnadesieae.....	37
Tabela 2. Identificação das variáveis discriminantes da tribo Barnadesieae, segundo modelo OPLS-DA-1.....	39
Tabela 3. Classificação das espécies das tribos Eupatorieae, Vernonieae, Astereae e Milleriae, segundo o modelo OPLS-DA-2 (Figura 13), construído a partir do metaboloma de tribos de Asteraceae, utilizado para encontrar variáveis discriminantes de Eupatorieae.....	43
Tabela 4. Identificação das variáveis discriminantes da tribo Eupatorieae, segundo modelo OPLS-DA-2.....	44
Tabela 5. Classificação das espécies das tribos Heliantheae, Vernonieae, Astereae e Milleriae, segundo o modelo OPLS-DA-3 (Figura 14), construído a partir do metaboloma de tribos de Asteraceae, utilizado para encontrar variáveis discriminantes de Heliantheae, Vernonieae, Astereae e Milleriae.....	49
Tabela 6. Identificação das variáveis discriminantes da tribo Heliantheae, Vernonieae, Astereae e Milleriae, segundo modelo OPLS-DA-3.....	50
Tabela 7. Espécies investigadas no ensaio de inibição da enzima LDH-5 e seus respectivos códigos.....	72
Tabela 8. Validação externa (27 extratos) ($P^2 = 81,48\%$) da OPLS-DA dos extratos foliares de espécies de Asteraceae avaliados contra a inibição da LDH-5.....	76
Tabela 9. Classificação dos extratos foliares de espécies de Asteraceae que foram retirados da análise (35 extratos) para a otimização do modelo de OPLS-DA para a inibição da atividade de LDH-5.....	79
Tabela 10. Identificação das variáveis discriminantes para os extratos foliares das espécies de Asteraceae classificados como “sem atividade” (no activity) e “ativo” (active), segundo modelo OPLS-DA.....	88
Tabela 11. Valores de classificação, validação e erro para os modelos 1 (todas as amostras) e 2 (75% das amostras – grupo treino) analisados por <i>Decision Tree</i> para extratos foliares de espécies de Asteraceae avaliados contra LDH-5.....	93

Tabela 12. Identificação das variáveis discriminantes para os extratos foliares das espécies de Asteraceae classificados como “sem atividade” (no activity) e “ativo” (active), segundo modelo <i>Decision Tree</i>	96
Tabela 13. A atividade fotoblástica das sementes de <i>A. robusta</i> e <i>A. trichophylla</i> foi analisada comparando-se a germinação sob exposição a 12 horas de luz versus a ausência total de luz (escuro).....	112
Tabela 14. Resultados do monitoramento da germinação de <i>A. robusta</i> e <i>A. trichophylla</i> expostos ao GA ₃	113
Tabela 15. Resultados do monitoramento da germinação de sementes de <i>T. diversifolia</i> expostas ao GA ₃	115
Tabela 16. Resultados do monitoramento da germinação de sementes de <i>T. diversifolia</i> expostas a aquecimento a 80 °C por 30 s e 3 min.....	116
Tabela 17. Valores de médias para os parâmetros de desenvolvimento dos cultivares de <i>A. robusta</i>	119
Tabela 18. Identificação das variáveis discriminantes da análise comparativa de coletas seguidas pela secagem de planta com N ₂ líquido e liofilização versus coletas seguidas da secagem em estufa de ar-circulante (40 °C).....	136
Tabela 19. Identificação das variáveis discriminantes de cada estação de coleta.....	143
Quadro 1. Área da curva de cada substância, seus respectivos tempos de retenção e [M+H] ⁺ em relação à diluição seriada em µg/ml.....	146
Quadro 2. Concentração estimada de substâncias químicas em <i>A. robusta</i> com base em curva de calibração.....	148

LISTA DE ABREVIATURAS E SIGLAS

3D	tridimensional
%G	porcentagem de sementes germinadas
°C	graus celsius
AsterDB	banco de dados do grupo de pesquisa Asterbiochem
ATP	trifosfato de adenosina
CH ₃ CN	acetonitrila
COX	ciclooxigenase
CV	cross validation
ESI	ionização por electrospray
EtOH	etanol
EUA	Estados Unidos da América
FID	ionização em chama
FTMS	espectrometro de massas por transformada de Fourier
g	grama
<i>g</i>	gravidade
GA ₃	ácido giberélico
GSI	índice de velocidade de germinação
h	horas
H ₂ O	água
há	hectare
HCA	hierarchical cluster analysis
HCAbp	hierarchical cluster analysis por reamostragem bootstrap
HCD	higher energy collisional dissociation
HIF	fator induzido por hipóxia
HRMS	espectrometro de massas de alta resolução
ID	identificação/número de identificação
kV	kilo volt
LC-MS	cromatografia líquida acoplada a espectrometria de massas
LDH	lactato desidrogenase
LOX	lipoxigenase

LST	lactonas sesquiterpênicas
m	metros
m/z	relação massa-carga (designação para massas protonadas ou desprotonadas neste trabalho)
MAE	mean absolute error
MeOH	metanol
mg	miligrama
MGT	tempo médio de germinação
min	minutos
ml	mililitro
MS	espectrometria de massas (espectrômetro de massas)
n.i.	não identificado
NADH	dinucleotídeo de nicotinamida e adenina
NCE	energia de coalisão normalizada
NCI	Instituto Nacional de Câncer
neg	modo de ionização negativo
OMS	Organização Mundial da Saúde
ONU	Organização das Nações Unidas
OPLS-DA	orthogonal partial least squares - discriminant analysis
PCA	principal component analysis
PLS	partial least squares
pos	modo de ionização positivo
ppm	parte por milhão
RMN	ressonância magnética nuclear
RMSEcv	root mean squared error of cross validation
RT	tempo de retenção
s	segundos
SD	desvio padrão
std	substância padrão
TIC	cromatograma de íons totais
UHPLC-UV/DAD-ESI-HRMS	cromatografia líquida de ultra performance acoplada a detector de UV/DAD e espectrometria de massas de alta resolução e ionização por ESI
UV/DAD	ultravioleta com arranjo de diodos

VC

coeficiente de variação

VIP

Variable importance in projection

WHO

Organização Mundial da Saúde

LISTA DE SÍMBOLOS

®	marca registrada
μ	micro
<i>E</i>	<i>entgegen</i> (trans)
<i>i</i>	linha da matriz
<i>j</i>	coluna da matriz
TM	trademark
α	alfa
β	beta
Σ	somatório

SUMÁRIO

INTRODUÇÃO GERAL	1
1 – METABOLÔMICA.....	1
2 – TÉCNICAS QUIMIOMÉTRICAS.....	5
3 – FAMÍLIA ASTERACEAE.....	9
OBJETIVOS GERAIS	15
CAPÍTULO 1: ABORDAGEM METABOLÔMICA PARA QUIMIOTAXONOMIA DE ASTERACEAE 16	
LISTA DE FIGURAS DO CAPÍTULO 1	16
LISTA DE TABELAS DO CAPÍTULO 1	17
LISTA DE ABREVIATURAS, SIGLAS E SIMBOLOS DO CAPÍTULO 1.....	17
RESUMO DO CAPÍTULO 1	18
SUMMARY OF CHAPTER 1	19
1 – INTRODUÇÃO.....	20
2 – OBJETIVOS.....	23
3 – MATERIAL E MÉTODOS.....	24
4 – RESULTADOS DE DISCUSSÃO.....	27
5 - CONCLUSÃO DO CAPÍTULO 1	61
CAPÍTULO 2: ESTUDOS DE INIBIÇÃO DE LDH-5 E TRIAGEM DE METABÓLITOS ATIVOS COM BASE NA ABORDAGEM METABOLÔMICA DE ESPÉCIES DE ASTERACEAE.....	62
LISTA DE FIGURAS DO CAPÍTULO 2	62
LISTA DE TABELAS DO CAPÍTULO 2	63
LISTA DE ABREVIATURAS, SIGLAS E SIMBOLOS DO CAPÍTULO 2.....	63
RESUMO DO CAPÍTULO 2	64
SUMMARY OF CHAPTER 2	65
1 – INTRODUÇÃO.....	66
2 – OBJETIVOS.....	68
3 – MATERIAL E MÉTODOS.....	69
4 – RESULTADOS DE DISCUSSÃO.....	70
5 - CONCLUSÃO DO CAPÍTULO 2	99
CAPÍTULO 3: ESTUDOS DE GERMINAÇÃO E CULTIVO DE <i>Aldama robusta</i> E <i>A. trichophylla</i> COMPARADOS COM <i>Tithonia diversifolia</i> (ASTERACEAE).....	100
LISTA DE FIGURAS DO CAPÍTULO 3	100
LISTA DE TABELAS DO CAPÍTULO 3	100
LISTA DE ABREVIATURAS, SIGLAS E SIMBOLOS DO CAPÍTULO 3.....	100
RESUMO DO CAPÍTULO 3.....	101
SUMMARY OF CHAPTER 3	102
1 – INTRODUÇÃO.....	103
2 – OBJETIVOS.....	105
3 – MATERIAL E MÉTODOS.....	106
4 – RESULTADOS E DISCUSSÃO.....	111
5 – CONCLUSÃO DO CAPÍTULO 3	120
CAPÍTULO 4: METABOLÔMICA DO CULTIVO DE <i>Aldama robusta</i> E <i>Aldama trichophylla</i> (ASTERACEAE) PARA ANÁLISE DA INFLUÊNCIA DO SOLO, LUMINOSA, CIRCADIANA E ESTACIONAL.....	122
LISTA DE FIGURAS DO CAPÍTULO 4.....	122
LISTA DE TABELAS E QUADROS DO CAPÍTULO 4.....	123
LISTA DE ABREVIATURAS, SIGLAS E SIMBOLOS DO CAPÍTULO 4.....	123
RESUMO DO CAPÍTULO 4.....	124
SUMMARY OF CHAPTER 4	125
1 – INTRODUÇÃO.....	126
2 – OBJETIVOS.....	128
3 – MATERIAL E MÉTODOS.....	129
4 – RESULTADOS E DISCUSSÃO.....	131
5 – CONCLUSÃO DO CAPÍTULO 4	148
DISCUSSÃO GERAL	149
CONCLUSÕES.....	151
REFERÊNCIAS.....	152
APÊNDICES	161

INTRODUÇÃO GERAL

1 – METABOLÔMICA

Metabolômica é uma ciência emergente, que visa o estudo abrangente e integrado de uma ampla gama de metabólitos de baixo peso molecular, incluindo abordagens de triagem global e análise quantitativa e/ou qualitativa de todo o metaboloma presente em um dado sistema biológico, como uma célula específica, tecido ou organismo, ou o conjunto deles, seja oriundo de plantas, microrganismos, animais ou organismos marinhos (VILLAS-BÔAS, 2005; DE VOS et al., 2007; KRASTANOV, 2010). A metabolômica é conduzida por meio de técnicas de detecção e identificação de metabólitos de acordo com suas características físico-químicas. Há diversas formas de se analisar o metaboloma, mas as principais técnicas hoje empregadas são as hífenadas, que combinam dois ou mais equipamentos para se detectar e identificar os metabólitos. Como técnicas de separação podemos citar a Cromatografia Gasosa (GC), Cromatografia Líquida de Alta Performance (HPLC) e Cromatografia Líquida de Ultra Performance (UHPLC), por meio das quais pode-se obter o tempo de eluição de cada metabólito e assim tentar identificá-lo. Como técnicas de detecção, podemos citar a Espectrometria de Massas (MS), Ressonância Magnética Nuclear (RMN), Ultravioleta (UV), UV com de Arranjo de Diodos (UV/DAD) e Ionização em Chama (FID).

Cada equipamento possui vantagens e desvantagens, os quais irão nortear a escolha da técnica hífenada a ser empregada na análise do metaboloma. Por exemplo, no que diz respeito às técnicas cromatográficas descritas anteriormente, todas tem um mesmo princípio de funcionamento: as substâncias são conduzidas por uma fase móvel e, ao interagir com uma fase estacionária, são mais ou menos retidas e então eluídas; deste modo, cada substância terá um tempo de eluição, relacionado com suas descrições físico-químicas (tamanho molecular, polaridade, solubilidade, etc). Entretanto, a GC tem seu uso restrito a metabólitos que não são termolábeis, uma vez que ela requer altas temperaturas para volatilizá-los e separá-los através da fase estacionária. Na GC, a fase móvel é um gás inerte, e a fase estacionária é um líquido ou polímero sobre um sólido inerte. Nas HPLC e UHPLC, por outro lado, a fase móvel é um líquido, e a fase estacionária é composta por partículas microscópicas (geralmente sílica igual ou menor que 5 µm) que podem ser porosas (exemplo, ACE C18®, ACE), ou envoltas por uma película porosa (exemplo, Kinetex® – Phenomenex), ou ainda são compostas por polímeros, como as colunas monolíticas (exemplo, Onyx monolithic C8– Phenomenex). As HPLC e UHPLC utilizam bombas que possibilitam gradientes de alta pressão e, deste modo, a utilização

de fases estacionárias com partículas de menor tamanho ou porosidade. São justamente estas partículas de menor tamanho e/ou porosidade que possibilitam uma melhor separação dos metabólitos, uma vez que há maior interação físico-química entre a substância e a fase estacionária. As designações “C8” e “C18” expressas neste parágrafo informam que a sílica da fase estacionária é tratada com grupos alquilas que podem ser C_8H_{17} ou $C_{18}H_{37}$. Este tratamento torna a fase estacionária apolar (cromatografia de fase reversa), o que permite o uso de fases móveis polares de baixo custo (como água) e uma melhor separação de substâncias de polaridade intermediária, como metabólitos secundários.

Já no que se diz respeito aos métodos de detecção, sua escolha também depende de vantagens e desvantagens, bem como das finalidades. Por exemplo, o detector UV/DAD permite analisar um cromatograma em diferentes comprimentos de onda. O UV/DAD é importante, uma vez que permite inferir a classe química do metabólito que estamos visualizando. A espectroscopia por RMN é capaz de determinar com precisão a estrutura química das substâncias. O fenômeno de RMN ocorre quando núcleos alinhados a um campo magnético aplicado são induzidos a absorver energia e a mudar a orientação do “giro” (*spin*) em relação ao campo aplicado (PAVIA et al., 2010). Na RMN são estudados principalmente a ressonância dos átomos de hidrogênio e carbono. A RMN é útil pois cada núcleo de 1H (ou ^{13}C) da estrutura química possuirá frequência diferente, a depender de outros átomos próximos. Por exemplo, a densidade eletrônica de valência varia de um próton para o outro na estrutura da substância. Tais elétrons criam um campo magnético que se opõe ao efeito do campo magnético aplicado, “protegendo” (*blindagem diamagnética* ou *anisotropia*) tais núcleos (PAVIA et al., 2010). Tendo em vista que o “ambiente químico” (quantidade de elétrons próximos aos núcleos de 1H ou ^{13}C) é diferente dentro das estruturas químicas, é possível gerar uma impressão digital do metabólito, baseado em sua estrutura química. Entretanto, a técnica de RMN tem a desvantagem de, em geral, ter baixo limite de detecção/quantificação e baixa sensibilidade quando comparada às técnicas de MS. Por outro lado, por meio da MS é possível detectar e identificar uma substância com base em sua capacidade de ionização e peso molecular, bem como nos fragmentos produzidos quando é exposta a energias elevadas de ionização, ou gases de colisão. Dentre as técnicas de ionização, podemos descrever: ionização por elétrons (EI), ionização química à pressão atmosférica (APCI), ionização por *electrospray* (ESI) e ionização por dessorção a laser assistida por matriz (MALDI). A ESI foi a técnica utilizada nos experimentos deste trabalho. Na ESI, um potencial elétrico é aplicado ao capilar onde há fase móvel contendo as substâncias que estão eluindo do cromatógrafo, formando um aerossol (cone

de Taylor). Esse processo resulta na formação de gotas com superfícies carregadas. Um gás inerte secante aquecido é utilizado para evaporar o solvente, aumentando a repulsão eletrostática entre as cargas formais nas gotículas, promovendo a explosão *coulombica*, liberando os íons. A fonte de ESI pode oxidar ou reduzir a substância, gerando íons moleculares ($M^{+\circ}$ ou $M^{-\circ}$), mas também pode formar moléculas protonadas ($[M + H]^+$) ou desprotonadas ($[M + H]^-$), bem como coordenação com íons metálicos Na^+ , Cl^- , K^+ e NH_4^+ , por meio de reações ácido-base (CROTTI et al., 2006). A extensão da ionização das substâncias é um processo que depende não apenas da voltagem aplicada ao capilar, mas também dos grupamentos químicos das substâncias. Assim, substâncias que apresentam grupamentos básicos, como aminas, amidas e ésteres, são mais facilmente protonadas e, portanto, analisadas (visualizadas/detectadas) no modo positivo. Por outro lado, substâncias com grupos ácidos, como os fenóis e ácidos carboxílicos, são mais facilmente desprotonados e, portanto, analisadas no modo negativo (CROTTI et al., 2006). A técnica de ESI é considerada uma ionização “suave”, uma vez que ela gera pouca fragmentação. Sendo assim, ela é muito útil em metabolômica, quando acoplada a cromatógrafos e espectrômetros de alta resolução, como CG, HPLC ou UHPLC. A técnica torna-se ainda mais eficaz em metabolômica quando há espectrômetros de massas em *tandem* (em sequência, MS/MS). Por meio desta técnica, é possível selecionar e promover a fragmentação de íons específicos (de interesse). Do mesmo modo, há diversos analisadores que podem ser utilizados em MS. O espectrômetro utilizado neste trabalho foi o de alta resolução com transformada de Fourier (HRFT), com analisador do tipo Orbitrap. O orbitrap opera aprisionando o íon em uma armadilha de Kingdon (*Kingdon trap*), que consiste em um eletrodo externo em forma de barril e um eletrodo interno fusiforme. Assim, por meio de um campo eletrostático, o íon fica aprisionado e oscilando em órbita ao redor do eletrodo fusiforme. A frequência das oscilações harmônicas do íon independe de sua velocidade, mas é inversamente proporcional à relação massa carga m/z (VESSECCHI et al., 2011), segundo a equação:

$$\omega = \sqrt{k/(m/z)}$$

onde ω é a frequência angular e k é força constante do potencial. Os dados químicos (metaboloma) deste projeto foram adquiridos por UHPLC-UV/DAD-ESI-MS.

Tendo em vista os diferentes equipamentos e técnicas para se realizar a metabolômica, esta vem sendo subdividida em demais subáreas, cada qual com sua finalidade e propósito. De maneira geral, metabolômica pode ser subdividida em *targeted* e *untargeted*. Essas definições são mais conceituais do que analíticas. *Targeted* refere-se à metabolômica direcionada, ou seja, análises do metaboloma que objetivam a detecção/identificação, e mesmo a quantificação de metabólitos específicos, ao passo que *untargeted* seria a detecção/identificação de tantos metabólitos quanto for possível. Assim, desde a maneira como se coleta uma amostra, se extrai seus metabólitos, e se adquire os dados químicos, pode influenciar em se classificar uma análise como *targeted* ou *untargeted*.

Independente da abordagem de escolha e, embora tenha sido comentado no primeiro parágrafo da introdução deste capítulo que “há diversas formas de se realizar a metabolômica”, a sociedade científica de metabolômica, *Metabolomics Society*, vem realizando esforços para se padronizar as condições mínimas e aceitáveis para estudos metabolômicos (SUMNER et al., 2007, SALEK et al., 2013; SCHRIMPE-RUTLEDGE et al., 2016). Novamente, as técnicas, em geral, podem variar de acordo com os objetivos da análise, mas é de suma importância que elas sejam minuciosamente descritas a fim de se garantir a confiabilidade do experimento, bem como sua reprodutibilidade (SUMNER et al., 2007). Por exemplo, no caso da metabolômica de plantas, não há um procedimento único sobre como se deve processar o material coletado que se vai analisar (com N₂ líquido ou com gelo seco, por exemplo), ou como se deve realizar a extração das amostras com solvente (com ou sem subsequente etapa de *clean up*, por exemplo) (SUMNER et al., 2007), mas é necessário especificar tais condições, padronizar a técnica e evitar o máximo de variações externas à análise (DE VOS et al., 2007, SUMNER et al., 2007, NAZ et al., 2014). Não obstante, a metabolômica sempre gera uma grande quantidade de dados químicos, os quais necessitam de subsequentes passos para serem analisados. Assim, os dados brutos provenientes das técnicas de LC-MS podem ser pré-tratados por *softwares* de processamento como o Metalign (LOMMEN, 2009; LOMMEN e KOOLS, 2012) e MZmine (KATAJAMMA et al., 2006; PLUSKAL et al., 2010), dentre outros, sendo que cada qual possui vantagens e desvantagens, como apresentado na Figura 1 (CASTILLO et al., 2011). Após o processamento, são geradas matrizes de dados que são normalizados pelas mais diferentes técnicas estatísticas, como *Pareto scaling*, *level scaling*, *log transformation*, etc (VAN DEN BERG et al., 2006). Após o pré-tratamento e a normalização, os dados do metaboloma podem ser analisados por diferentes técnicas quimiométricas (item 2, Capítulo 1).

Data preprocessing steps covered by the proteomic and metabolomic software tools discussed in this paper.									
Software	Filtering	Peak detection	Deisotoping	Alignment	Gap filling	Visualization	Identification	Normalization	Batch mode
apLCMS	✓	✓	-	✓	✓	-	-	-	✓
Decon2LS	✓	✓	✓	-	-	✓	-	-	-
Madison Metabolomics Consortium Database	-	-	-	-	-	-	✓	-	-
MAVEN	✓	✓	-	✓	-	✓	✓	-	-
MetAlign	✓	✓	-	✓	-	-	-	-	✓
MetaboAnalyst	-	✓	-	-	✓	-	-	✓	-
MSight	-	-	-	-	-	✓	-	-	-
MZedDB	-	-	-	-	-	-	✓	-	-
MZmine 2	✓	✓	✓	✓	✓	✓	✓	✓	✓
OpenMS/TOPP	✓	✓	✓	✓	-	✓	✓	-	✓
Sirius	-	-	✓	-	-	✓	✓	-	-
SuperHirn	-	✓	-	✓	✓	-	-	✓	-
XCMS	✓	✓	-	✓	✓	✓	✓	-	-

Figura 1 - Comparação de diferentes softwares de processamento de dados (Extraído de Castillo *et al.*, 2011).

Por fim, as diferentes abordagens metabolômicas podem ser utilizadas para diversas finalidades, em estudos envolvendo conceitos de ecologia química (ecoquímica ou ecofisiologia), quimiotaxonomia, triagem de extratos para a busca de substâncias ativas, controle de qualidade de amostras complexas, biogeografia, dentre outros. De maneira geral, em estudos de ecologia a metabolômica pode ser útil para identificar substâncias altamente correlacionadas com sazonalidade, ritmo circadiano, condições ambientais, interações ecoquímicas entre os seres vivos (como alelopatia), dentre outros (GOBBO-NETO e LOPES, 2007; SAMPAIO *et al.*, 2016; PADILLA-GONZÁLEZ *et al.*, 2017). Já estudos de quimiotaxonomia (e por vezes de quimiosistemática) buscam agrupar e caracterizar *taxa* (plural de *taxon*) de acordo com seus respectivos metabólitos biossintetizados. Tais estudos podem ser úteis na sistemática de plantas (PADILLA-GONZÁLEZ *et al.*, 2016; GALLON *et al.*, 2018) ou na triagem de rotas biossintéticas específicas de certos *taxa*, estudos de distribuição biogeográfica ou mesmo no controle de qualidade de amostras biológicas, como por exemplo, plantas morfológicamente similares, mas quimicamente distintas (PADILLA-GONZÁLEZ *et al.*, 2017). A metabolômica para triagem de substâncias biologicamente ativas é útil para apontar substâncias discriminantes e altamente correlacionadas com determinada atividade biológica em amostras complexas (CHAGAS-PAULA *et al.*, 2015; SANTOS *et al.*, 2018).

2 – TÉCNICAS QUIMIOMÉTRICAS

A enorme quantidade de dados gerados em metabolômica pode ser analisada pelas mais diversas técnicas quimiométricas, a fim de se extrair informações relevantes. Os dados obtidos são usualmente multidimensionais e, neste caso, são chamados de “multivariados”. Os diferentes métodos matemáticos e estatísticos aplicados aos dados podem ser a Análise das Componentes Principais (*Principal Component Analysis – PCA*), Análise de Agrupamento

Hierárquico (*Hierarchical Cluster Analysis* – HCA), Mínimos Quadrados Parciais (*Partial Least Squares* – PLS) e sua variante ortogonal (OPLS), sendo que estas duas últimas podem ser utilizadas para Análises Discriminantes (*Discriminant Analysis* – DA) (WOLD et al., 2001; TRYGG e WOLD, 2002), compondo as conhecidas técnicas PLS-DA e OPLS-DA. Assim sendo, a HCA e PCA são análises multivariadas não supervisionadas, pois elas são úteis para o reconhecimento de padrões, tendências e agrupamentos não direcionados das amostras. Já o PLS-DA é considerado uma análise multivariada supervisionada, pois o agrupamento observado se correlacionará com certa resposta previamente informada à análise, as quais podem ser numérica ou descritiva, como o valor de dada bioatividade ou os *clusters* formados por gêneros ou origem biogeográfica.

Melhor explicando, tanto a PCA quanto o PLS-DA são técnicas de redução de dimensionalidade (RENCHER, 2002), onde um conjunto de amostras com variáveis multidimensionais são reduzidas no que se chama de “componentes principais”, no caso da PCA, e “variáveis latentes”, no caso do PLS-DA. Ambas técnicas projetam em um hiperplano bidimensional o que seria impossível de se observar em um espaço multidimensional. A PCA sempre traça a primeira componente na direção de maior variância encontrada entre as variáveis do sistema. Uma segunda componente, ortogonal à primeira, explica uma quantidade de variância menor que a primeira componente, mas maior que a terceira componente, à qual é ortogonal às duas primeiras. Quanto maior a complexidade de um conjunto de dados multidimensionais, tão mais difícil é de se explicar com apenas duas componentes a maior variância encontrada, sendo, portanto, necessários mais componentes. O PLS funciona da mesma forma, e o termo componentes principais é utilizado às vezes para designar as variáveis latentes. Entretanto, a diferença é que as variáveis latentes sempre apontam na direção de maior variância sobre uma determinada informação (a resposta em Y) fornecida ao sistema. Assim sendo, o PLS-DA correlaciona a variância do espaço multidimensional (a matriz X) com uma resposta informada (matriz Y) (WOLD et al., 2001). Já o OPLS-DA é considerado uma modificação do PLS-DA, onde é realizada uma separação das variações da matriz X em duas partes, uma linearmente relacionada à matriz Y, chamada de componentes preditivas, e outra que é ortogonal a Y, chamada de não preditiva com a resposta Y e usualmente importante para detectar variações internas dentro dos grupos, mas não correlacionadas com a resposta (TRYGG e WOLD, 2002). Isso gera uma maior interpretabilidade, visto que as variações ortogonais a Y nos informam as diferenças dentro de X, mas que não se correlacionam com Y. Todas estas técnicas de redução dimensional do espaço multidimensional geram um novo

gráfico (projeção do hiperplano bidimensional) de amostras, chamado agora de *scores*, os quais se relacionam diretamente com o novo gráfico gerado para as variáveis, chamado agora de *loadings*.

Já a HCA é um método de “*Cluster Analysis*” propriamente dito. Em outras palavras, a HCA necessariamente agrupa por similaridade as amostras do estudo, procurando por padrões e tendências dentro dos agrupamentos (*clusters*). Ela é representada na forma de uma “árvore” onde os *clusters* se formam ao longo de uma distância. Quanto mais próximo a 0, mais similar. Uma distância comum bastante utilizada é a Euclidiana, a qual se trata da distância entre dois vetores, um x (x_1, x_2, \dots, x_p) e outro y (y_1, y_2, \dots, y_p), definido como:

$$d(x, y) = \sqrt{x - y'(x - y)} = \sqrt{\sum_{j=1}^p (x_j - y_j)(x_j - y_j)}$$

Há muitas formas de se fazer uma HCA, sendo as mais utilizadas o *Single Linkage*, o qual se utiliza da distância entre as amostras mais próximas para construir os *clusters*, e o método de Ward (*Ward's Method*) (WARD, 1963), o qual usa a variância total dentro das amostras para definir os *clusters*. O resultado é que o método de Ward minimiza a somatória da distância dentro dos *clusters* (RENCHER, 2002; LAROSE, 2006). A escolha de qual método se deve utilizar é empírica e depende das necessidades ou do que se pretende informar. Entretanto, a HCA não informa a confiabilidade dos *clusters* que se formam e, portanto, os resultados devem ser interpretados com cautela. Todavia, a HCA por reamostragem (*bootstrapping*), ou HCAbp, pode ser utilizada se obter não somente os agrupamentos hierárquicos, mas também seus graus de incertezas, fornecendo estimativas da probabilidade do *cluster*. A HCAbp pode ser realizada por meio do pacote *pvclust* no *software* R (The R Project website, <http://www.r-project.org/>) (SUZUKI, R. e SHIMODARA, H., 2006).

Por fim, é necessário introduzir aqui os conceitos de *Data Mining* (Mineração de Dados) e como ela tem sido utilizada pela quimiometria recentemente (BUYDENS et al., 1999; MUTIHAC, L. e MUTIHAC R., 2008) para a classificação, predição de amostras e descoberta de variáveis latentes. *Data Mining* trata da resolução de problemas presentes em bancos de dados, explorando-os por meio de técnicas computacionais, visando o reconhecimento de padrões e/ou tendências e classificações. Tais padrões nos permitem realizar predições em novos dados. Atualmente, muitas das técnicas desenvolvidas em *Data Mining* podem ser

utilizadas pela quimiometria, tais como Análise Bayesiana (*NaiveBayes*), Redes Neurais Artificiais e *Decision Tree* (Árvores de Decisão). Na representação gráfica da *Decision Tree*, os “nós” da árvore representam um determinado atributo (uma variável) que foi testado por condições específicas (concentração, por exemplo) e apontado como uma variável latente, discriminante, de uma classe (WITTEN, 2011). Há vários algoritmos que calculam uma *decision tree* e não convém explicá-los aqui, embora os mais utilizados sejam o C4.5, desenvolvido por Ross Quinlan (QUINLAN, 1993), e o J.48 (BHARGAVA et al., 2013), implementação do C4.5 no *software* gratuito Weka (HALL et al., 2009).

Durante as técnicas quimiométricas, são apontados pelos modelos diversos picos do cromatograma (como o cromatograma de íons totais – TIC) relevantes ou discriminantes de amostras específicas. A etapa final da metabolômica é, portanto, identificar (e por vezes até mesmo quantificar) tais picos. Para tanto, faz-se uso de diversas técnicas de identificação, tal como o isolamento por meio de técnicas cromatográficas da substância correspondente ao pico de interesse e sua identificação por meio de técnicas espectroscópicas e espectrométricas usuais. Entretanto, muitas vezes tais picos são de baixa intensidade (portanto podendo ser pouco concentrado na amostra) e, além disso, às vezes a resolução cromatográfica não permite um fácil isolamento e identificação da substância. Deste modo, a identificação necessita ser realizada sem o prévio isolamento do pico, mas no próprio extrato, o que é chamado de “desrepliação” (do inglês, *dereplication*) (HUBERT et al., 2017). A desrepliação é enormemente facilitada pela presença de substâncias previamente isoladas ou adquiridas comercialmente (biblioteca de substâncias padrão), quando se pode injetá-las no sistema LC-UV-MS/MS e confirmar a substância que se quer identificar por meio de similaridades (correlação/correspondência/*match*) com o tempo de retenção, classe química, massa, e o perfil de ligações intramoleculares. Para tanto, é importante que se tenha as informações físico-químicas (banco de dados) de tal biblioteca. Ademais, se o usuário dispuser de um conjunto de substâncias químicas estruturalmente diversificadas e que foram analisadas pelo mesmo gradiente de eluição, é possível gerar modelos de tempos de retenção cromatográfico com base na estrutura química e prever o tempo de retenção de outras substâncias que não fazem parte da biblioteca de padrões do laboratório (EUGSTER et al., 2014; OLIVEIRA et al., 2015). Além disso, se o laboratório dispuser de espectros MS/MS ou MSⁿ em seu banco de dados, pode-se também construir modelos de predição de fragmentação de substâncias (inclusive de outros bancos de dados, de outros laboratórios) e aumentar a quantidade picos identificados (YANG et al., 2013; ALLARD et al., 2016; WANG et al., 2016). A identificação/desrepliação de

substâncias químicas em metabolômica também segue parâmetros e níveis de identificação estabelecidos pela *Metabolomics Society* (SUMNER et al., 2007; SALEK et al., 2013). A equipe AsterBioChem (www.asterbiochem.org/), juto à qual este trabalho foi realizado, possui o banco de dados AsterDB, que pode ser acessado por meio do *link* “AsterDB” no site ou do catálogo da ZINC (<http://zinc15.docking.org/catalogs/asternp/>), banco de dados gratuito com milhões de estruturas químicas de substâncias.

3 – FAMÍLIA ASTERACEAE

A família Asteraceae (Compositae) compreende um dos maiores grupos de plantas das Angiospermas, com aproximadamente 25.000 espécies organizadas em 1.600 gêneros (VERDI et al., 2005; FUNK et al., 2009), 44 tribos e 13 subfamílias (Figura 2) (PANERO et al., 2014), representando 10% de todas as Angiospermas (PAVARINI et al., 2018). Esta família de plantas apresenta ampla distribuição geográfica em todos os continentes, sendo mais presente em climas tropicais e subtropicais (PAVARINI et al., 2018). A classificação mais recente (PANERO et al., 2014) das subfamílias de Asteraceae a subdivide em 13 subfamílias (Figura 2), sendo descrita a mais nova subfamília Faminanthoideae, um grupo “irmão” das subfamílias Multisoideae-Asteroideae, as quais correspondem a mais de 99% das Asteraceae (PANERO et al., 2014).

A taxonomia de Asteraceae se inicia em 1816, com o botânico Henri Cassini agrupando as espécies em 19 tribos (CASSINI, 1816). Desde então, a família passou por diversos reagrupamentos e classificações. Destas, talvez ainda a mais conhecida seja a de Bremer (BREMER, 1996), o qual, baseado em critérios morfológicos, subdividiu a família em 4 subfamílias (Barnadesioideae, Carduoideae, Cichorioideae e Asteroideae) e 14 tribos (Figura 3), sendo a subfamília Barnadesioideae, com uma única tribo, considerada um clado basal dentro da família. O *taxon* Mutiseae foi considerado indefinido (Figura 3).

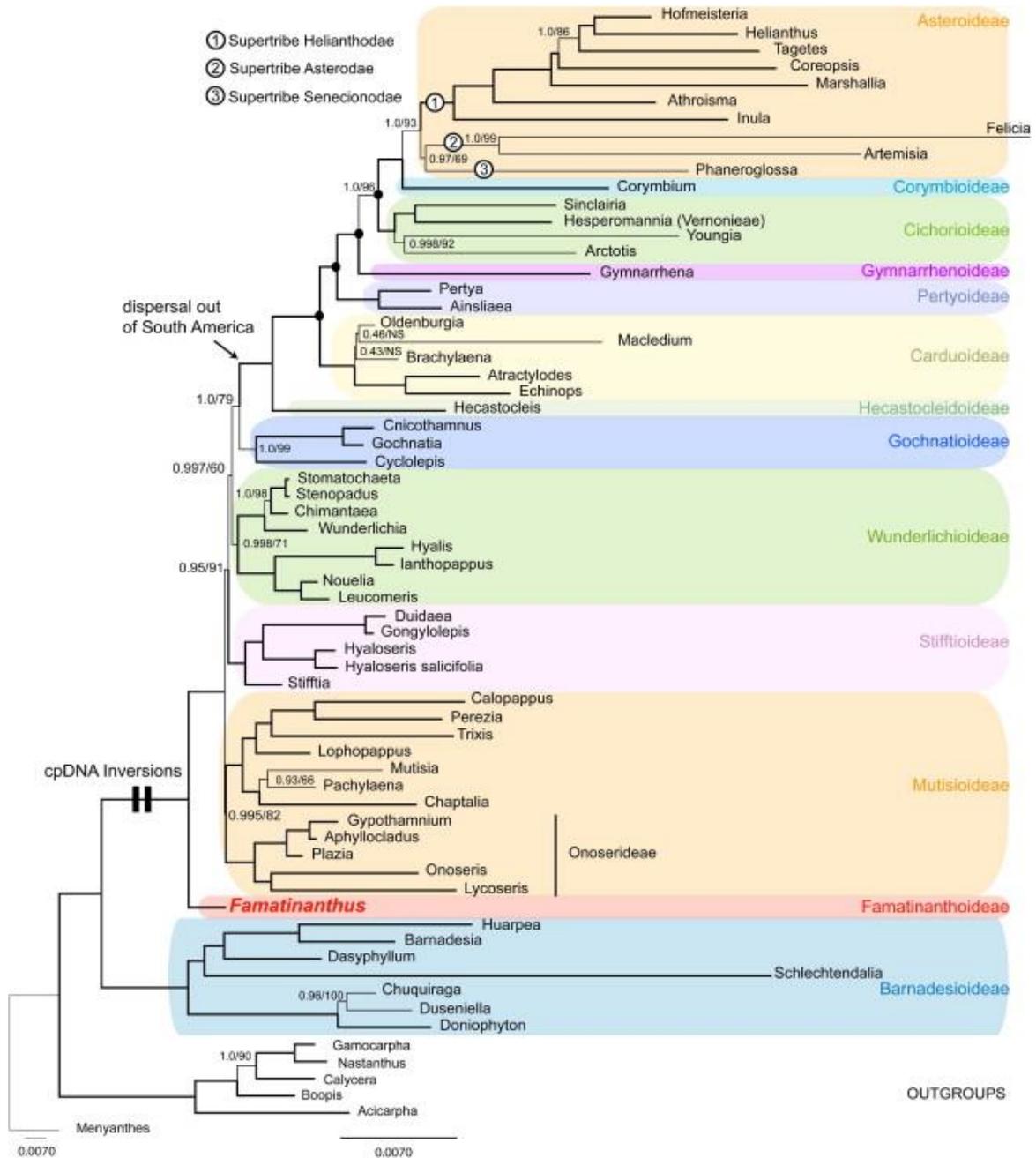


Figura 2 - Árvore filogenética de Asteraceae construída com base em dados de DNA e inferência Bayesiana. Destacam-se as 13 subfamílias. Figura extraída de Panero e colaboradores (PANERO et al., 2014).

Uma característica marcante de Asteraceae é a presença de capítulos (*capitula* plural em latim), que é uma inflorescência especializada que pode conter centenas de pequenas flores em um receptáculo discoide ou arredondado contendo brácteas. Muitos membros de Asteraceae possuem dois tipos de flores: flores do disco, localizadas na porção central do capítulo, e flores do raio, localizadas na periferia. O ovário das Asteraceae é ínfero (se encontra dentro do

receptáculo) e apresenta um único óvulo. A família é frequentemente conhecida como “família do girassol” (*Helianthus annuus* L.), espécie da tribo Heliantheae, subfamília Asteroideae.

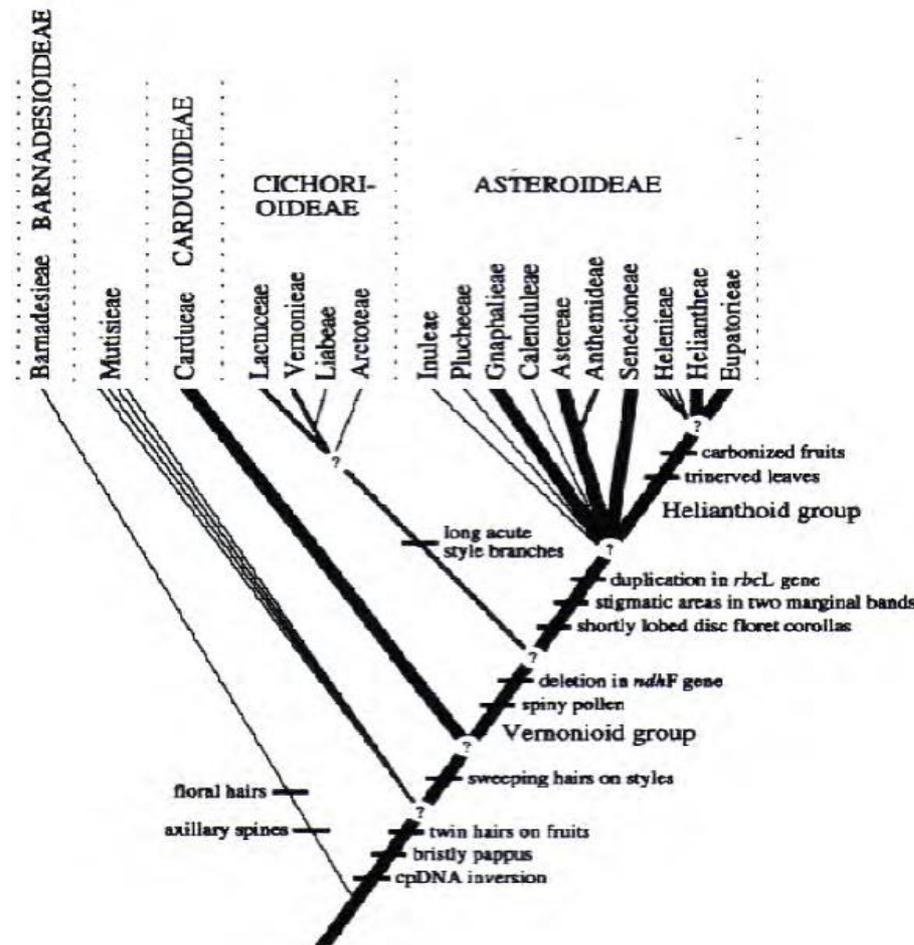


Figura 3 - Diagrama da família Asteraceae construído com base em caracteres morfológicos (BREMER et al., 1996). Evidenciam-se as 4 subfamílias (Barnadesioideae, Carduoideae, Cichorioideae e Asteroideae, e suas respectivas tribos). Figura oriunda de Calabria e coautores (CALABRIA et al., 2006).

Do ponto de vista químico, a família Asteraceae é caracterizada pela presença de lactonas sesquiterpênicas (LST) e sesquiterpenos furanos (SEAMAN, 1982; ALVES et al., 2018, in: SÜLSEN e MARTINO, 2018). Deste modo, elas poderiam ser estudadas computacionalmente na classificação das espécies de Asteraceae (DA COSTA et al., 2005; SCOTTI et al., 2012) (Figura 4). Além destes compostos, a família também possui outras classes de metabólitos secundários, tais como os diterpenos (labdanos, cauranos, pimaranos, clerodanos), triterpenos, derivados do ácido quínico (ácidos gálico e quínico, por exemplo), flavonoides (flavonois, flavonas, flavanonas, flavanonois, isoflavonas, flavan-3-ois, chalconas, antocianinas, auronas e isoflavonas), derivados do ácido *trans*-cinâmico e poliacetilenos

(Figura 4) (ZDERO e BOHLMANN, 1990; BOHM e STUESSY, 2001). Compostos voláteis das classes dos monoterpênos, sesquiterpênos e fenilpropanóides também estão presentes em óleos essenciais. Por outro lado, não é comum encontrar na família alcalóides, iridóides, estilbenos, policetídeos e taninos.

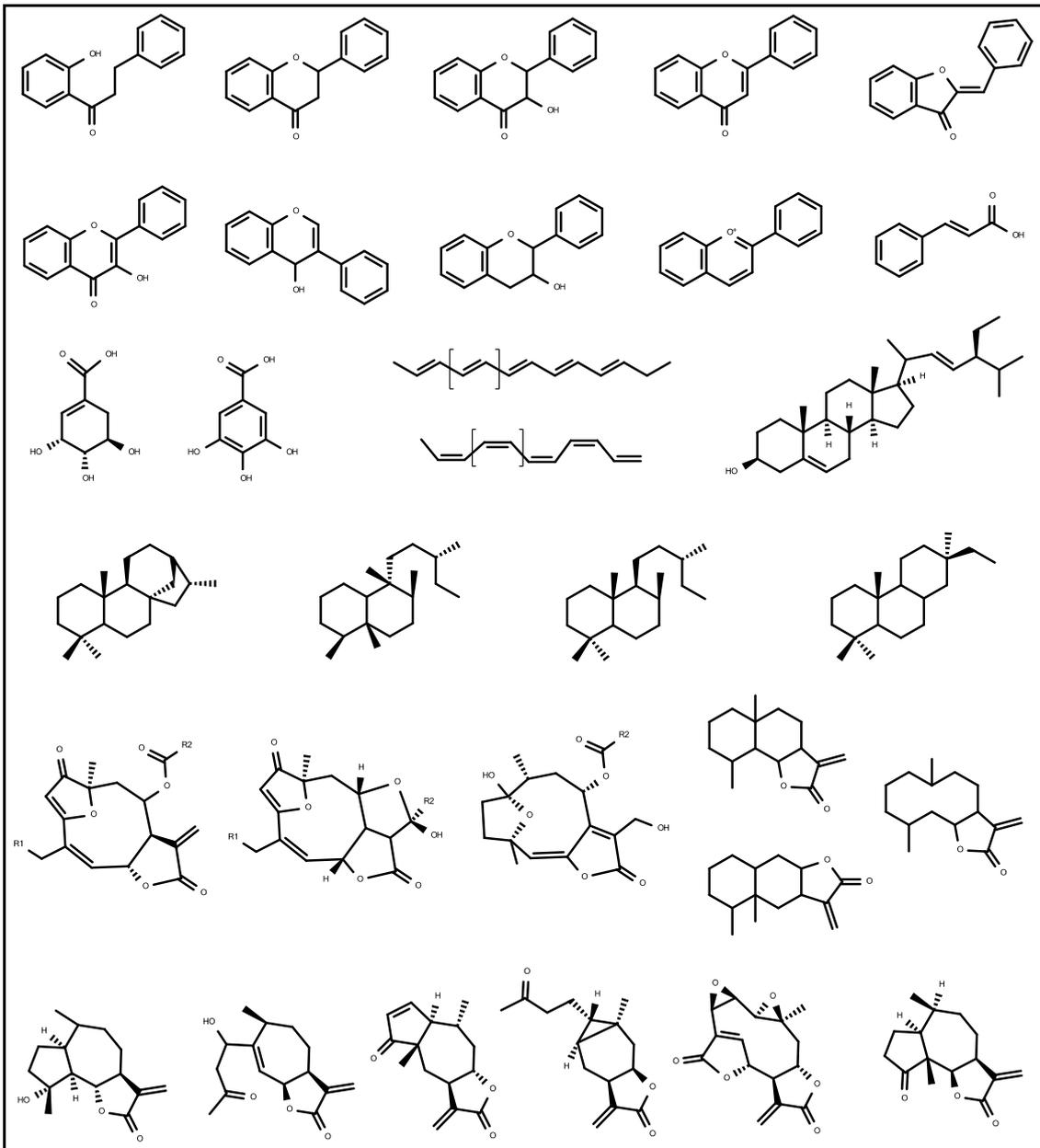


Figura 4 – Algumas das principais classes químicas de metabólitos secundários em Asteraceae. Da esquerda para a direita e de cima para baixo: flavonoides chacona, flavanona, flavanonol, flavona, aurona, flavonol, isoflavona, flavan-3-ol, antocianina; ácido trans-cinâmico; ácido quínico, ácido gálico; poliacetilenos; triterpeno (estigmasterol); diterpenos caurano, clerodano, labdano, pimarano; LSTs furanoheliangolido, eremantolido, hirsutinolido, eudesmanolido, eremofilanolido, germacrolido, guaianolido, xantanolido, helenanolido, carabrone, mikanolido e ambrosanolido.

Neste trabalho, foi realizado um estudo quimiotaxonômico cerca de 250 espécies de Asteraceae, distribuídas em 16 tribos (ver Apêndice A), sendo que algumas delas são de grande importância do ponto de vista fitoquímico e, portanto, pode-se comentar com mais detalhes suas constituições químicas. O gênero *Baccharis* L. (tribo Astereae, subfamília Asteroideae), por exemplo, foi amplamente estudado, tendo sido descrito diversas atividades biológicas para seus constituintes químicos, dentre as quais efeito antioxidante, antibacteriano, antiprotozoário, antiviral, antifúngico, anti-inflamatório, antimutagênico, dentre outras (CAMPOS et al., 2016). Pertence a este gênero, inclusive, a planta medicinal, de amplo uso etnofarmacológico, carqueja (*B. trimera*), conhecida por melhorar problemas gastrointestinais. De maneira geral, as principais classes químicas do gênero são: flavonoides (flavonois, flavonas, flavanonas, flavanonois, glicosilados ou não), derivados do ácido *trans*-cinâmico (derivados cumarínicos e ácidos cafeoilquínicos, principalmente), fenilpropanoides (ácido ferúlico, por exemplo), monoterpenos, sesquiterpenos, diterpenos (clerodanos, labdanos, cauranos, por exemplo), triterpenos e esteróis. Outro gênero de grande importância, *Mikania* Willd. (tribo Eupatorieae, subfamília Asteroideae), merece ser sucintamente comentado, tendo em vista inclusive a importância etnofarmacológica e econômica de duas de suas espécies, *M. laevigata* e *M. glomerta*, popularmente conhecidas como “guaco”. Ambas espécies foram oficializadas como plantas medicinais (Formulário de Fitoterápicos da Farmacopeia Brasileira, 1ª edição). O xarope de guaco possui o marcador químico cumarina (1,2-benzopirona, derivada do ácido *trans*-cinâmico) (NAPIMOGA e YASTUDA, 2010) e é amplamente utilizado pelo seu efeito antitussígeno. Outras classes químicas do gênero são: diterpenos (labdanos, pimaranos e cauranos), derivados do ácido *trans*-cinâmico (ácidos cafeoilquínicos), flavonoides, triterpenos, esteróis e LST, em especial as do tipo mikanolido (RUFATTO et al., 2012). Outros gêneros importantes em Asteraceae são a *Tithonia* Desf. ex Juss. e *Helianthus* L. (ambos da tribo Heliantheae e subfamília Asteroideae), bem representados pelas suas espécies *T. diversifolia* e *H. annuus*. Em ambas *T. diversifolia* e *H. annuus* foi relatada a presença de derivados cumarínicos, flavonois, flavonas, diterpenos caurenóicos, LST dos tipos guaianolido, eudesmanolido, germacrolido, heliangolido and furanoheliagolidos, triterpenos e esteróis (CHAGAS-PAULA et al., 2012; MARSNI et al., 2015). As atividades biológicas relatadas para *T. diversifolia* (extratos brutos ou substâncias isoladas) são diversas, dentre as quais efeito antimalárico, antiviral, antimicrobiano, antiespasmódico, anti-inflamatório e citotóxico (CHAGAS-PAULA et al., 2012). *H. annuus* é mais estudado por causa do óleo fixo extraído

de suas sementes, tendo grande importância agrônômica e econômica. Ainda assim, foram relatadas algumas atividades para seus metabólitos secundários, tais como efeitos herbicida, antimicrobiano, antifúngico e regulares de crescimento (MARSNI et al., 2015). Outro gênero de constituição química muito semelhante à de *Helianthus* e *Tithonia* é *Aldama* La Llave (outrora *Viguiera* Kunth). De fato, não apenas são relativamente semelhantes em suas constituições químicas, como também pertencem à mesma tribo e subtribo (Helianthinae). Para os constituintes químicos e extratos brutos deste gênero, foram relatadas diversas atividades biológicas, como inseticida (MARQUINA et al., 2001), antimicrobiana (CANALES et al., 2008; PORTO et al., 2009), anti-tripanosoma (NOGUEIRA et al., 2016), anti-inflamatória (SHORR et al., 2002; ARAKAWA et al., 2008; CHAGAS-PAULA et al., 2015), muitas das quais principalmente relacionadas à presença de LST budleína A e ao diterpeno do tipo caurano ácido *ent-caur-16-en-19-oico*. Por fim, outros dois gêneros muito importantes em Asteraceae serão brevemente descritos aqui, *Vernonia* Schreb. e *Lychnophora* Mart., ambos da tribo Vernoniaceae (subfamília Cichorioideae), mas respectivamente das subtribos Vernoniinae e Lychnophorinae. Uma planta medicinal de *Vernonia* bem conhecida é a assa-peixe (*V. polyanthes*), sendo utilizada para o tratamento de gripe e como expectorante. De maneira geral, em ambas as subtribos foram relatados os mesmos constituintes químicos, como fenilpropanoides, flavonoides (principalmente flavonois, favanonas, flavonas e chalconas), derivados do ácido *trans*-cinâmico, derivados do ácido quínico, sesquiterpenos, diterpenos (pouco comum) e LST germacrolidos, eudesmanolido e guaianolido. A diferença sutil se encontra nas LST do tipo furanoheliangolido, com os tipos eremantolido e goyazensolido como marcadores de Lychnophorinae (KELES et al., 2010), e hirsutinolido como marcador de Vernoniinae (BOHLMAN et al., 1981; KOS et al., 2006). Para a subtribo Lychnophorinae, são descritas na literatura atividades antioxidante, antineoplásica, genotóxica, citotóxica, tripanomicida, antimicrobiana e anti-inflamatória (KELES et al., 2010).

Pelo o que foi exposto neste item do capítulo 1 deste trabalho, fica evidente a grande importância taxonômica, farmacológica e econômica de muitas espécies de Asteraceae. Entretanto, muitas das espécies destes gêneros estão sendo extintas, tendo em vista as atividades econômicas predatórias no campo, a inexistência de estudos de cultivo, a negligência e o descaso das autoridades políticas. De maneira geral, a população não vê fins práticos para se preservar tais espécies, sendo algumas inclusive consideradas como ervas daninhas e pragas, a exemplo da *T. diversifolia*. Deste modo, cabe à comunidade científica demonstrar os ganhos econômicos e ambientais que se pode conseguir por meio da preservação e até mesmo do cultivo

de tais espécies. Para tanto, investigações a respeito das atividades biológicas dos metabólitos de tais gêneros, seja em termos farmacológicos ou ecoquímicos, são de grande importância para que se justifiquem suas preservações.

OBJETIVOS GERAIS

- 1) Realizar estudos quimiotaxonômicos de Asteraceae, com base no metaboloma de mais de 250 espécies de 16 tribos diferentes (capítulo 1);
- 2) Realizar estudos de inibição de LDH-5 com extratos de espécies de Asteraceae, bem como a triagem de metabólitos correlacionados com extratos considerados ativos (capítulo 2);
- 3) Estudar melhores condições de germinação e cultivo de *A. robusta*, *A. trichophylla* e *T. diversifolia* (capítulo 3);
- 4) Realizar estudos da influência de fatores ambientais no metaboloma de cultivares de *Aldama* com base em abordagens metabolômicas (capítulo 4).

CAPÍTULO 1: ABORDAGEM METABOLÔMICA PARA QUIMIOTAXONOMIA DE ASTERACEAE

LISTA DE FIGURAS DO CAPÍTULO 1

Figura 5. Exemplo da reprodutibilidade da análise destacando padrões internos.....	28
Figura 6. Inspeção das réplicas de ID 14, com os respectivos padrões internos, cafeína e daidzeína.....	28
Figura 7. Inspeção das réplicas de ID 159, com os respectivos padrões internos, cafeína e daidzeína.....	29
Figura 8. Inspeção de réplicas de extratos provenientes de espécies das tribos Eupatorieae, Vernonieae, Astereae e Millerieae.....	30
Figura 9. PCA (4 componentes, 0,29) com espécies de Asteraceae das tribos Heliantheae, Eupatorieae, Vernonieae, Astereae, Barnadesieae e Milleriae.....	32
Figura 10. Gráfico de erro padrão por valor-p AU.....	33
Figura 11. HCAbp realizado com algumas espécies das tribos Barnadesieae, Eupatorieae, Astereae, Millerieae, Vernonieae e Heliantheae.....	34
Figura 12. OPLS-DA.1 (3 + 7) ($R^2 = 0,90$ e $Q^2 = 0,37$) mostrando um agrupamento das espécies de Asteraceae com base no metaboloma das tribos Barnadesieae, Heliantheae, Eupatorieae e Vernonieae.....	36
Figura 13. OPLS-DA.2 (4 + 3) ($R^2 = 0,93$ e $Q^2 = 0,78$) mostrando um agrupamento das espécies de Asteraceae com base no metaboloma das tribos Astereae, Milleriae, Eupatorieae e Vernonieae.....	38
Figura 14. OPLS-DA.3 (4 + 3) ($R^2 = 0,92$ e $Q^2 = 0,79$) mostrando um agrupamento das espécies de Asteraceae com base no metaboloma das tribos Astereae, Milleriae, Heliantheae e Vernonieae.....	48

LISTA DE TABELAS DO CAPÍTULO 1

Tabela 1. Classificação das espécies das tribos Heliantheae, Eupatorieae, Vernonieae e Barnadesieae, segundo o modelo OPLS-DA-1 (Figura 12), construído a partir do metaboloma de tribos de Asteraceae, utilizado para encontrar variáveis discriminantes de Barnadesieae.....	37
Tabela 2. Identificação das variáveis discriminantes da tribo Barnadesieae, segundo modelo OPLS-DA-1.....	39
Tabela 3. Classificação das espécies das tribos Eupatorieae, Vernonieae, Astereae e Millerieae, segundo o modelo OPLS-DA-2 (Figura 13), construído a partir do metaboloma de tribos de Asteraceae, utilizado para encontrar variáveis discriminantes de Eupatorieae.....	43
Tabela 4. Identificação das variáveis discriminantes da tribo Eupatorieae, segundo modelo OPLS-DA-2.....	44
Tabela 5. Classificação das espécies das tribos Heliantheae, Vernonieae, Astereae e Millerieae, segundo o modelo OPLS-DA-3 (Figura 14), construído a partir do metaboloma de tribos de Asteraceae, utilizado para encontrar variáveis discriminantes de Heliantheae, Vernonieae, Astereae e Millerieae.....	49
Tabela 6. Identificação das variáveis discriminantes da tribo Heliantheae, Vernonieae, Astereae e Millerieae, segundo modelo OPLS-DA-3.....	50

LISTA DE ABREVIATURAS, SIGLAS E SIMBOLOS DO CAPÍTULO 1

AsterDB	banco de dados do grupo de pesquisa
CH ₃ CN	acetonitrila
ESI	ionização por electrospray
EtOH	etanol
g	grama
g	gravidade
H ₂ O	água
HCAbp	hierarchical cluster analysis por reamostragem bootstrap
m/z	relação massa-carga (designação para massas protonadas ou desprotonadas neste trabalho)
mg	miligrama
min	minutos
ml	mililitros
OPLS-DA	orthogonal partial least squares - discriminant analysis
PCA	principal component analysis
RMSEcv	root mean squared error of cross validation
RT	tempo de retenção
UHPLC-UV/DAD-ESI-HRMS	cromatografia líquida de ultra performance acoplada a detector de UV/DAD e espectrometria de massas de alta resolução e ionização por ESI
UV/DAD	ultravioleta com arranjo de diodos

RESUMO DO CAPÍTULO 1

No capítulo 1 foi analisada a metabolômica aplicada à quimiotaxonomia de algumas espécies de Asteraceae. O objetivo foi verificar se seria possível agrupar as espécies em tribos, com base em caracteres químicos provenientes do metaboloma e de acordo com a classificação de Bremer (BREMER, 1996) ou Panero (PANERO et al., 2014). Para tanto, foi extraído o metaboloma de mais de 250 espécies, provenientes de 16 tribos. A aquisição dos dados químicos se deu por UHPLC-UV/DAD-ESI-HRMS, em cromatografia de fase reversa. Os dados foram analisados por PCA, HCAbp e OPLS-DA. Pelas técnicas não supervisionadas não foi possível correlacionar o agrupamento quimiotaxonômico com as classificações de Bremer (BREMER, 1996) ou Panero (PANERO et al., 2014). Deste modo, mais investigações devem ser realizadas antes de se concluir inequivocadamente que os caracteres químicos não são bons marcadores químicos em Asteraceae ao nível da classificação de tribos. Por outro lado, diversas análises de OPLS-DA permitiram separar as tribos de Barnadesieae, Heliantheae, Vernonieae, Milleriae, Astereae e Eupatorieae, com base na diferença de seus constituintes químicos. Com base nisso, picos discriminantes de cada tribo foram indicados e desreplicados com base em bancos de massa e tempo de retenção de substâncias de Asteraceae. Tais substâncias podem ser consideradas marcadores quimiotaxonômicos das tribos, estando correlacionadas com a adaptação ao meio em que vivem suas respectivas espécies, ou com a diferenciação evolutiva entre as tribos.

SUMMARY OF CHAPTER 1

In chapter 1 the metabolomics applied to the chemotaxonomy of some Asteraceae species were analyzed. The objective was to verify if it was possible to group the species into tribes, based on chemical characters from the metabolome and according to Bremer's classification (Bremer 1996) or Panero (PANERO et al., 2014). For that, the metabolome of more than 250 species from 16 tribes was extracted. The acquisition of the chemical data was by UHPLC-UV/DAD-ESI-HRMS, in reverse phase chromatography. Data were analyzed by PCA, HCABp and OPLS-DA. By the unsupervised techniques it was not possible to correlate the chemotaxonomic grouping with the classifications of Bremer (BREMER, 1996) or Panero (PANERO et al., 2014). Thus, further investigations must be conducted before it is unequivocally concluded that chemical characters are not good chemical markers in Asteraceae at the level of tribes classification. On the other hand, several analyzes of OPLS-DA allowed to separate the tribes of Barnadesieae, Heliantheae, Vernoniae, Milleriae, Astereae and Eupatorieae, based on the difference of their chemical constituents. Based on this, discriminant peaks of each tribe were indicated and dereplicated based on mass banks and retention time of compounds from Asteraceae. These substances can be considered chemotaxonomic markers of the tribes, being correlated with the adaptation to the environment in which their respective species live, or with the evolutionary differentiation between the tribes.

1 – INTRODUÇÃO

1.1 – Quimiotaxonomia

A taxonomia é um campo das ciências biológicas que determina a organização das espécies, classificando-as e delimitando-as em relação umas às outras em posições hierárquicas (BEAMAN, 1986), com base em dados morfológicos, genéticos, químicos, dentre outros. Analogamente, quimiotaxonomia vegetal pode ser descrita como um método de classificação de plantas baseado na similaridade de sua constituição química (HEGNAUER, 1967; BEAMAN, 1986; WINK, 2003; REYNOLDS, 2007). No entanto, depois da obra “Origem das Espécies”, de Darwin, publicada em 1859, a relação evolutiva tornou-se a base da classificação de espécies (DARWIN, 1859; BEAMAN, 1986). Nesse sentido, a sistemática, uma das áreas da biologia, está relacionada à classificação das espécies de acordo com suas relações evolutivas (filogenia) (BEAMAN, 1986). Deste modo, a sistemática é um estudo mais amplo que taxonomia, englobando-a. Por assim dizer, a confusão entre os termos “quimiotaxonomia” e “quimiosistemática” decorre da confusão que se faz entre os conceitos de taxonomia e sistemática. Neste trabalho, portanto, considerar-se-á como sendo mais adequado usar o termo “quimiotaxonomia” em vez de “quimiosistemática” para estudos que não utilizam o conceito de “ancestral comum”, que é o único critério de classificação da sistemática.

Assim sendo, um estudo quimiotaxonômico pretende, primeiramente, verificar se existe algum agrupamento dos seres vivos em função de seus constituintes químicos, compreendendo a biodiversidade entre diferentes *taxa* e suas relações com o meio ambiente. Após isso, a quimiotaxonomia pode ser útil em auxiliar estudos de sistemática, quando ela aponta metabólitos específicos (marcadores taxonômicos) que podem estar relacionados com a filogenia. Neste último caso, estaríamos nos referindo à quimiosistemática.

A quimiotaxonomia também está relacionada à “taxonomia numérica” (fenética), a qual se ocupa em classificar os seres vivos de acordo com suas similaridades fenotípicas, mesmo que o processo não reflita na ancestralidade comum. Assim, quanto maiores as similaridades fenotípicas, como os metabólitos, mais próximas estarão as espécies dentro de um *cluster* e, portanto, maior será o grau de parentesco entre elas. Ainda assim, a taxonomia numérica afirma que as mudanças fenotípicas adquiridas ao longo do tempo por um organismo poderiam ser usadas para fins filogenéticos (SNEATH e SOKAL, 1973; SOKAL e SNEATH, 1966). Na taxonomia numérica, uma quantidade maior de dados é desejada para melhores resultados. Nesse sentido, as abordagens metabolômicas são úteis para a quimiotaxonomia.

Mais precisamente em relação à quimiotaxonomia em Asteraceae, já em 1982 Seaman (SEAMAN, 1982) relatou que as LST poderiam ser utilizadas como marcadores taxonômicos, desde que fossem utilizadas com cautela e sob corretas abordagens analíticas (ALVES et al., 2018, in: SÜLSEN e MARTINO, 2018). De fato, Seaman advertiu que as LST até então relatadas apresentavam descontínua distribuição entre os grupos, e que dados completos de todos os *taxa* incluídos em um estudo seriam necessários para os estudos em sistemática (SEAMAN, 1982; ALVES et al., 2018, in: SÜLSEN e MARTINO, 2018). Em outras palavras, a metabolômica poderia solucionar tais problemas e auxiliar a sistemática, tendo em vista sua abordagem mais completa.

Ainda em relação às LST, Spring e colaboradores propuseram alterações taxonômicas em *Helianthus* (SPRING et al., 1989, 1990, 1991). Pode-se também notar diferenças entre as principais cadeias carbônicas de LST quando são comparados gêneros ou tribos de Asteraceae. Assim, as LST dos tipos germacrolido, heliangolido e furanoheliangolido com oxigênio no C-8 sempre com orientação em β são marcadores em Heliantheae. Por outro lado, a tribo Vernonieae possui furanoheliangolidos com oxigênio com orientação em α , como os eremantolidos e goyazensolidos. Pode-se notar também que a subtribo Vernoniinae (Vernonieae) possui os biomarcadores hirsutinolido (LST do tipo uranoheliangolido) e glaucolido (LST do tipo germacrolido). Por outro lado, o gênero *Mikania* possui como marcador taxonômico as LST do tipo mikanolido, como comentado no terceiro item da Introdução Geral (Figura 4).

Como foi descrito no terceiro item da Introdução Geral (Figura 4), há outros metabólitos que são frequentemente documentados em Asteraceae, ao passo que outros não o são. Por exemplo, além das LST, as flavonas e flavonóis também podem ser descritos como marcadores em Asteraceae. Os padrões de substituição mais frequentes para as flavonas são a trioxigenação 5,7,4' (tipo apigenina) e a tetraoxigenação 5,7,3',4' (tipo luteolina). Para os flavonóis, a tetraoxigenação 3,5,7,4' (tipo canferol) e pentaoxigenação 3,5,7,3',4' (tipo de quercetina) são mais comuns (BOHM e STUESSY, 2001; GUO et al., 2017).

Em 1987, Emerenciano e colaboradores descreveram que os níveis de oxigenação e especialização do esqueleto de LST e flavonoides poderiam estar sugerindo evolução entre as tribos de Asteraceae, uma vez que se distribuíam discriminantemente entre elas (EMERENCIANO et al., 1987). Além disso, Emerenciano et al. (2006) construiu um modelo com PLS empregando um banco de dados de metabólitos secundários de Asteraceae coletados da literatura, visando avaliar os níveis de oxidação de acordo com a classificação taxonômica

(EMERENCIANO et al., 2006). Seus resultados mostraram uma interdependência entre os estados de oxidação de cada classe de metabólito secundário em níveis de tribo e subtribo de Asteraceae (EMERENCIANO et al., 2006). Em 2001, um estudo similar realizado com base em PCA demonstrou correlação entre metabólitos secundários e tribos de Asteraceae (ALVARENGA et al., 2001). Um outro estudo de filogenia foi realizado com base em 389 caracteres químicos de ca. 7.000 metabólitos descritos na literatura para 3.159 espécies de Asteraceae (CALABRIA et al., 2006). Esse estudo concluiu que a falta de dados químicos da literatura, bem como sua inconsistência, estaria comprometendo os resultados sobre a filogenia das tribos de Asteraceae (CALABRIA et al., 2006), o que certamente pode ser superado por abordagens metabolômicas. Em 2005, Da Costa et al. utilizou 144 estruturas químicas em formato 3D para classificar 7 subtribos dentro das tribos Eupatorieae, Heliantheae e Vernonieae por meio de redes neurais artificiais (Da Costa et al., 2005). A correlação entre dados químicos e a classificação taxonômica desse estudo indica que redes neurais podem ser utilizadas no auxílio em classificações taxonômicas de Asteraceae (DA COSTA et al., 2005). Utilizando a mesma técnica, Correia e seus colaboradores concluíram que as redes neurais são uma eficaz técnica em se classificar tribos de Asteroideae, bem como em classificar subtribos de Heliantheae e gêneros de Helianthinae, com base em diferenças estruturais de seus metabólitos (CORREIA et al., 2008 e 2012).

Conforme comentado no parágrafo anterior, a utilização de técnicas de quimiometria e *data mining* aplicada a dados químicos coletados da literatura certamente é útil para a realização de classificações quimiotaxonômicas e podem auxiliar na compreensão das relações das plantas com o meio-ambiente, ou mesmo auxiliar em problemas de delimitações taxonômicas, em estudos de quimiosistemática. Entretanto, podemos também notar que a falta de dados na literatura torna-se uma limitação das técnicas. De fato, não é porque uma substância não foi descrita para determinado táxon, que ela não foi biossintetizada. O isolamento e caracterização de substâncias químicas depende drasticamente das técnicas de extração e fracionamento empregadas, bem como dos objetivos do isolamento. Por outro lado, o metaboloma engloba uma maior quantidade de constituintes químicos produzidos em um tecido/órgão de estudo. Deste modo, uma vez que a metabolômica seja padronizada para todos os indivíduos do estudo, ela pode trazer melhores resultados quimiotaxonômicos. Deste modo, Martucci et al. (2015) pode agrupar em quatro grupos distintos dez espécies de *Vernonia* com base em seu metaboloma adquirido por LC-MS (MARTUCCI et al., 2015). Um outro trabalho ainda mais amplo, com dados de LC-MS de 79 espécies das subtribos Vernoniinae,

Lepidaploinae e Rolandrinae, também foi capaz de classificar as espécies em quatro grupos distintos e, por meio de análises supervisionadas, como OPLS-DA, apontar seus metabólitos discriminantes (GALLON et al., 2018). Além disso, 62 espécies de plantas de seis famílias diferentes (Asteraceae, Rosaceae, Fabaceae, Fagaceae, Betulaceae e Aceraceae) também foram agrupadas por análises não supervisionadas, como PCA e HCA, e distinguidas de acordo com seu metaboloma (LEE et al., 2015). Na ocasião, Lee e seus colaboradores também aplicaram análises baseadas em PLS-DA e OPLS-DA a fim de separarem as espécies com base na taxonomia e atividade biológica (LEE et al., 2015). Outro trabalho recente utilizou metabolômica direcionada para a identificação de flavonoides e, de acordo com os dados, distinção de quatro espécies e duas subespécies de *Chrysanthemum* L. (Asteraceae) (UEHARA et al., 2017). Um outro trabalho utilizou metabolômica baseada na espectroscopia de RMN uni e bidimensional para agrupar onze espécies de *Ilex* L. (Aquifoliaceae) através de várias abordagens quimiométricas (KIM et al., 2010). Já por meio de metabolômica baseada em GC-MS, um estudo agrupou 18 espécies de *Ferula* L. (Apiace) em quatro grupos principais (KANANI et al., 2011), caracterizando-os quimicamente. Por fim, um outro trabalho, também baseado em GC-MS, pode agrupar e caracterizar quimicamente trinta e uma espécies da subtribo Espeletiinae (Asteraceae) (PADILLA-GONZÁLEZ et al., 2016).

Deste modo, considerando-se o potencial da quimiotaxonomia assistida por metabolômica, os gêneros (e grupos infragenéricos) de Asteraceae estudados durante este trabalho poderiam ser caracterizados quimicamente, apontando seus metabólitos discriminantes e biomarcadores. Uma vez encontrados os metabólitos discriminantes, estes poderiam ser correlacionados com a filogenia dos gêneros (GALLON et al., 2018) em estudos quimiosistemáticos, bem como evidenciar possíveis interações entre o habitat e a química do gênero, em estudos ecoquímicos.

2 – OBJETIVOS

Os objetivos deste capítulo foram:

- 1) realizar a quimiotaxonomia de Asteraceae e observar se o agrupamento das espécies se corrobora com a taxonomia das tribos proposta por BREMER et al., 1996 ou de PANERO et al., 2014;

2) apontar metabólitos discriminantes ao nível da classificação de tribos das espécies de Asteraceae investigadas neste estudo por meio de OPLS-DA com dados do metaboloma;

3 – MATERIAL E MÉTODOS

3.1 – Material vegetal

O material vegetal utilizado neste capítulo são folhas coletadas ou adquiridas comercialmente de ca. 250 espécies de Asteraceae, classificadas em 16 tribos (Apêndice A), dentre as quais constam espécies dos gêneros *Aldama*, *Baccharis*, *Espeletia* Mutis ex Bonpl., *Helianthus*, *Lychnophora*, *Mikania*, *Tithonia*, *Vernonia*, dentre outras. Neste trabalho, algumas espécies foram analisadas em replicata, avaliando-se a reprodutibilidade da intensidade dos picos e os tempos de retenção cromatográficos.

3.2 – Extração e aquisição dos dados químicos

Para a extração, foram pesadas 10 mg de droga vegetal seca previamente moídas em moinho analítico (IKA[®] A11, Sigma Aldrich, Alemanha) e extraídas com 500 µL de EtOH-H₂O (4:1). Após o respectivo tempo de extração, a mistura foi centrifugada a 10.000 g por 15 min (Eppendorf 5810R, Alemanha). Após esta etapa, foi realizado o *clean-up* com *n*-hexano grau HPLC (Macron Chemicals, EUA). O sobrenadante contendo o extrato foi coletado e colocado em novos microtubos (Eppendorf, Hamburgo, Alemanha) previamente pesados. Tal extrato foi secado em centrífuga a vácuo a 40 °C (SpeedVac SAVANT SPD 2010, Thermo Electron Corporation, EUA) e ressuspendido em CH₃CN-H₂O (1:1) a uma concentração de 1 mg/ml. Por fim, o extrato foi filtrado em membranas Mini-Uniprep[™] PTFE de 0.2 µm (Whatman, EUA) e 4 µL foram injetados em UHPLC-UV/DAD-ESI-HRMS, com analisador Exactive Plus Orbitrap[™] (Thermo Scientific, EUA). Três padrões internos (cafeína, daidzeína e ác. 3- α -acetóxi-copálico) também foram injetados juntamente com cada uma das réplicas de todas as amostras, a fim de se validar o método.

Foi utilizada uma coluna cromatográfica de fase reversa (XB-C-18, *coreshell*, 150 x 2.1 mm, 1.7 µm de diâmetro de partícula, Kinetex[®], Phenomenex, EUA) para o UHPLC (Accela 1250, Thermo Scientific, Alemanha), protegida por uma pré-coluna equivalente, operando em forno de coluna a 30 °C. A análise foi realizada através de um sistema de eluição por gradiente: 0-15 min (5-50% de solvente B), 15-20 (50-50% B), 20-30 (50-100% B), 30-35 (100-100% B), 35-40 (100-5% B), sendo o solvente A (H₂O deionizada (MilliQ^R Merck Millipore Corporation, EUA) e solvente B (CH₃CN (grau LC-MS, Macron Fine Chemicals, EUA), ambos com 0.1%

de ácido fórmico grau HPLC (Sigma-Aldrich, Alemanha). A vazão foi estabelecida em 400 $\mu\text{L}/\text{min}$.

Para a análise de espectrometria de massas, a voltagem do *spray* foi selecionada em 3,6 kV para o modo positivo e 3,2 kV para o modo negativo (*heater temperature* = 300 °C; *capilar temperature* = 320 °C). O intervalo de varredura foi selecionado de 100,0 a 1500,0 *m/z*. A resolução foi de 70.000 para ambos modos de ionização. A fragmentação foi realizada em cela HCD (*higher-energy collisional dissociation*) com NCE (*normalized collision energy*) de 35%. Para todos os calibrantes (solução de calibrantes Thermo calmix, Sigma-Aldrich) observou-se acurácia de 0,1 ppm (modo positivo) e 0,2 ppm (modo negativo).

3.3 – Processamento dos dados do LC-MS

Os dados brutos adquiridos do LC-MS em formato .raw foram transformados em formato .mzXML por linha de comando usando-se a função MSconvert do *software* ProteoWizard (KESSNER et al., 2008). Após isso, o processamento dos dados foi conduzido no *software* MZmine 2.23 (PLUSKAL e al., 2010) sequencialmente por: *Mass Detection*, *FTMS Filter Gaussian*, *Chromatogram builder*, *Smoothing*, *Deconvolution*, *Isotopic peaks grouper*, *Filtering*, *Alignment*, *Gap filling* e *Duplicate Filter*. Os parâmetros específicos estão no Apêndice C. Os valores de $[M+H]^+$ e $[M-H]^-$ dos picos detectados foram confrontados com os valores de massa exata de aproximadamente 2.000 substâncias descritas na literatura provenientes de diferentes gêneros da família Asteraceae (AsterDB - asterbiochem.org/asterdb). Também foram utilizados os valores de massa e tempo de retenção do banco de dados *in-house* da biblioteca de substâncias (cerca de 200 substâncias) para a identificação dos picos detectados.

3.4 – Análises quimiométricas

Os dados químicos de todos os extratos após serem processados foram exportados em planilha (.csv), organizados de modo a colocar as variáveis nas colunas e amostras nas linhas, e as áreas sob os picos detectados foram normalizadas para posterior análise multivariada e *decision tree*. Para as análises de quimiotaxonomia os dados foram normalizados por *log transformation* e por *Pareto*, de acordo com os seguintes cálculos:

Log Transformation: $t_{ij} = \log_{10}(x_{ij})$, seguida por $y_{ij} = t_{ij} - \bar{t}_i$, onde \log é o logaritmo na base 10 da área do pico detectado, e \bar{x} é a média das áreas dos picos.

Pareto: $y_{ij} = x_{ij} - \bar{x}_i / \sqrt{SD_{ij}}$, onde \bar{x} é a média da área dos picos detectados e SD é o desvio padrão.

A PCA e OPLS-DA foram realizadas no *software* SIMCA v.13.0.3 (Umetrics – Suécia). Para apenas as análises de triagem de inibidores de LDH-5, os dados foram também analisados por *decision tree* pelo *software* Weka 3.6 (Universidade de Waikato, Nova Zelândia).

A avaliação dos resultados da PCA e OPLS-DA foi realizada por meio da inspeção dos valores de coeficiente de correlação múltipla (R_x^2 e R_y^2), capacidade preditiva (Q^2 , *leave-one-out 7-fold cross validation*) e raiz quadrada do erro médio da validação cruzada (RMSEcv).

Valores de R^2 superiores a 0.6 e de Q^2 superiores a 0.3 são desejáveis em metabolômica, considerando-se presença de *overfitting* quando o valor de $R^2 - Q^2$ for superior a 0.3. Valores de RMSEcv próximos a 0 são desejáveis. Além disso, os resultados foram também avaliados a um nível de significância de 95% (valor $p < 0,005$).

A HCA por reamostragem (*bootstrapping*), ou HCAbp, foi realizada para se avaliar os dados de quimiotaonomia. Nesta análise, os dados foram normalizados por *Pareto*. A finalidade foi obter não somente os agrupamentos hierárquicos, mas também seus graus de incertezas, fornecendo estimativas da probabilidade do *cluster*. A HCAbp pode ser realizada por meio do pacote *pvclust* no *software R v.3.4.3 (The R Project for Statistical Computing, <http://www.r-project.org>, SUZUKI, R. e SHIMODARA, H., 2006)*. Para acessar a confiabilidade do *cluster* foram realizadas 10.000 replicatas *bootstrap* ($nboot = 10.000$).

4 – RESULTADOS DE DISCUSSÃO

4.1 – Resultados da quimiotaonomia de *Asteraceae*

Primeiramente foi realizada a checagem da reprodutibilidade do tempo de retenção e da intensidade dos picos das moléculas protonadas dos padrões internos que foram adicionados aos extratos para fins de validação. Tal inspeção foi realizada escolhendo-se os extratos de ID 10, 50, 100, 150 e 200 no *software MZmine* (Figura 5). Constatou-se que o tempo de retenção dos padrões internos se alterou muito pouco, o que gera confiabilidade em relação ao método cromatográfico utilizado (Figura 5). Por outro lado, a intensidade dos picos, a qual reflete a ionização e concentração do analito no extrato, variou um pouco (Figura 5). À primeira vista, pode parecer que há muita variação na intensidade dos picos, mas se for levado em consideração a escala dos picos, todos estão acima de 10^9 .

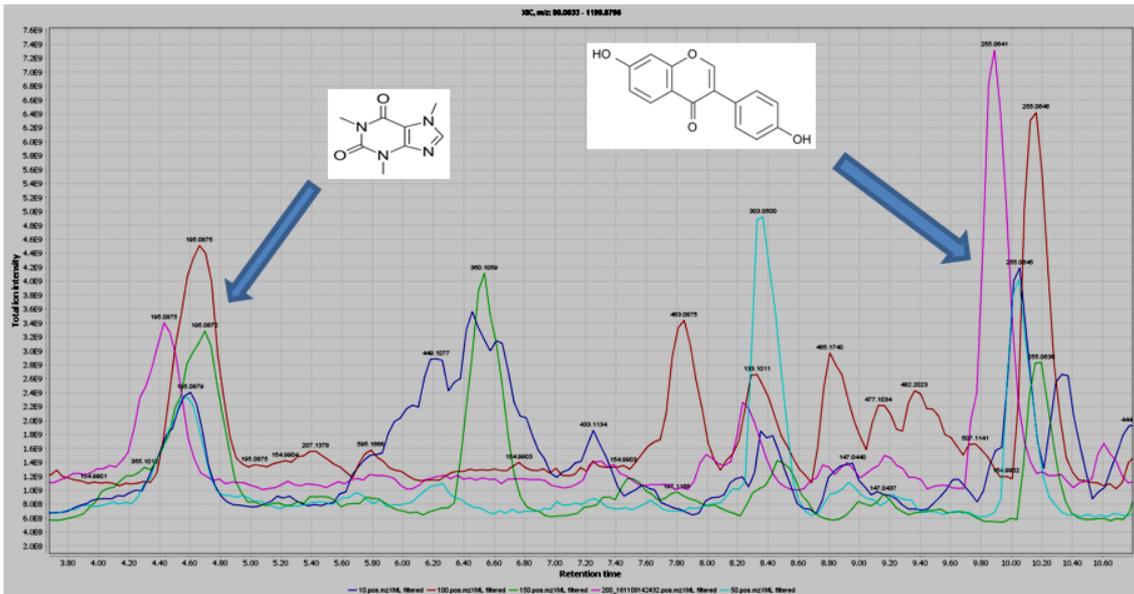


Figura 5 – Exemplo da reprodutibilidade da análise destacando padrões internos (cafeína, RT = 4,39 e $[M+H]^+ = 195$; daidzeína, RT = 9,78 e $[M+H]^+ = 253$) nos extratos de ID 10, 50, 100, 150 e 200, analisados no *software* MZmine.

Também foi avaliada reprodutibilidade do tempo de retenção e da intensidade dos picos das moléculas protonadas (ou desprotonadas) do estudo por meio da inspeção de algumas das réplicas (Figuras 6 e 7). Assim, tendo em vista os extratos de ID 14 (Figura 6) e 159 (Figura 7), por exemplo, pode-se notar alta reprodutibilidade nos tempos de retenção e intensidade das moléculas protonadas, com total sobreposição dos picos, inclusive dos padrões internos, cafeína e daidzeína, que eluem em aproximadamente 4,39 e 9,78 min, respectivamente.

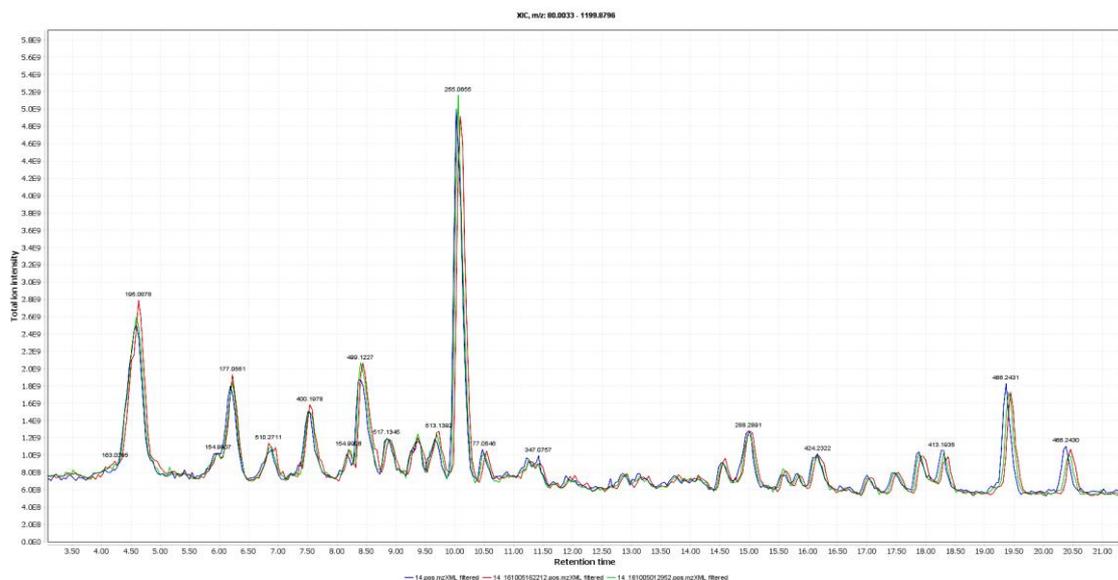


Figura 6 - Inspeção das réplicas de ID 14, com os respectivos padrões internos, cafeína e daidzeína, em 4,39 e 9,78 min., respectivamente.

Em tais figuras, nota-se também uma boa resolução cromatográfica. De fato, a variação do tempo de retenção para tais amostras e sistema cromatográficos escolhidos foi menor que 5% (dados não disponibilizados). Foi concluído que o método “se mostrou robusto, preciso, acurado, linear e seletivo” (ROSA, 2018, dados não publicados). Do mesmo modo, o coeficiente de variação (VC), ou desvio padrão relativo, do tempo de retenção dos padrões internos foi menor que 1%. O VC é uma medida de dispersão de variáveis, sendo definido pela razão do desvio padrão (σ) pela média (μ), de acordo com a fórmula:

$$VC = \sigma / \mu$$

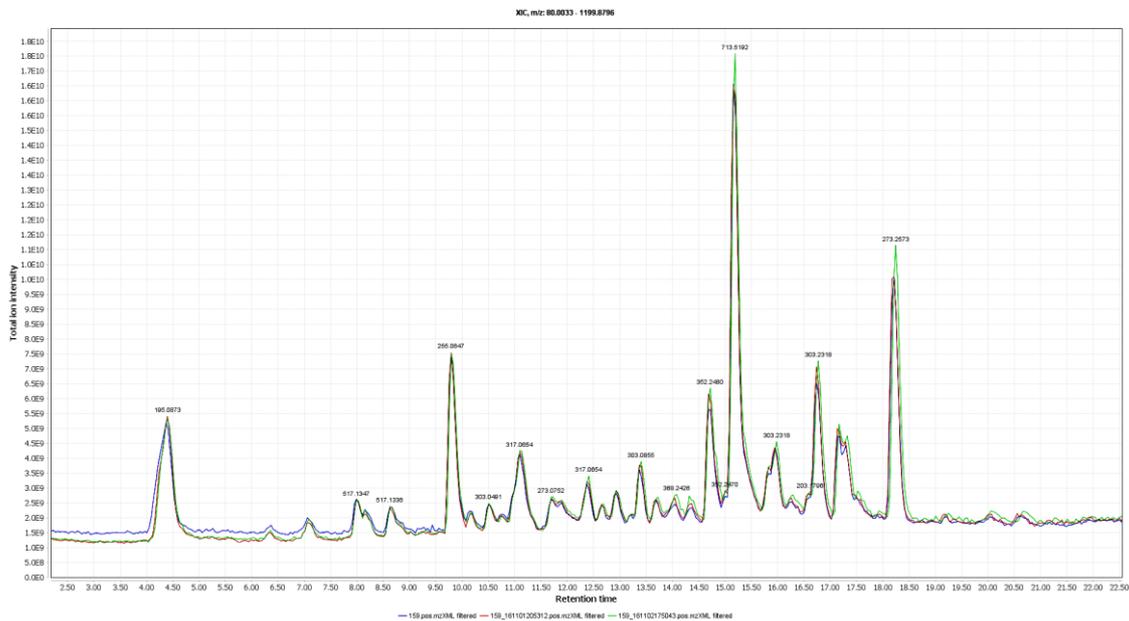


Figura 7 - Inspeção das réplicas de ID 159, com os respectivos padrões internos, cafeína e daidzeína, em 4,39 e 9,78 min., respectivamente.

Os dados químicos já processados e normalizados foram também inspecionados em PCA. De fato, seria esperado que em uma análise multivariada do metaboloma extraído no extratos de diversas espécies de Asteraceae conseguisse agrupar não somente as réplicas, mas também os *taxa* mais próximos. Assim sendo, a análise de PCA representada pela Figura 10 pode, de fato, revelar ligeira semelhança entre réplicas de extratos provenientes das tribos Eupatorieae, Vernonieae, Astereae e Millerieae, embora não tenha sido observado agrupamento entre tais tribos, nem mesmo entre amostras do gênero *Espeletia* (IDs 9, 33 e 40) (Figura 8).

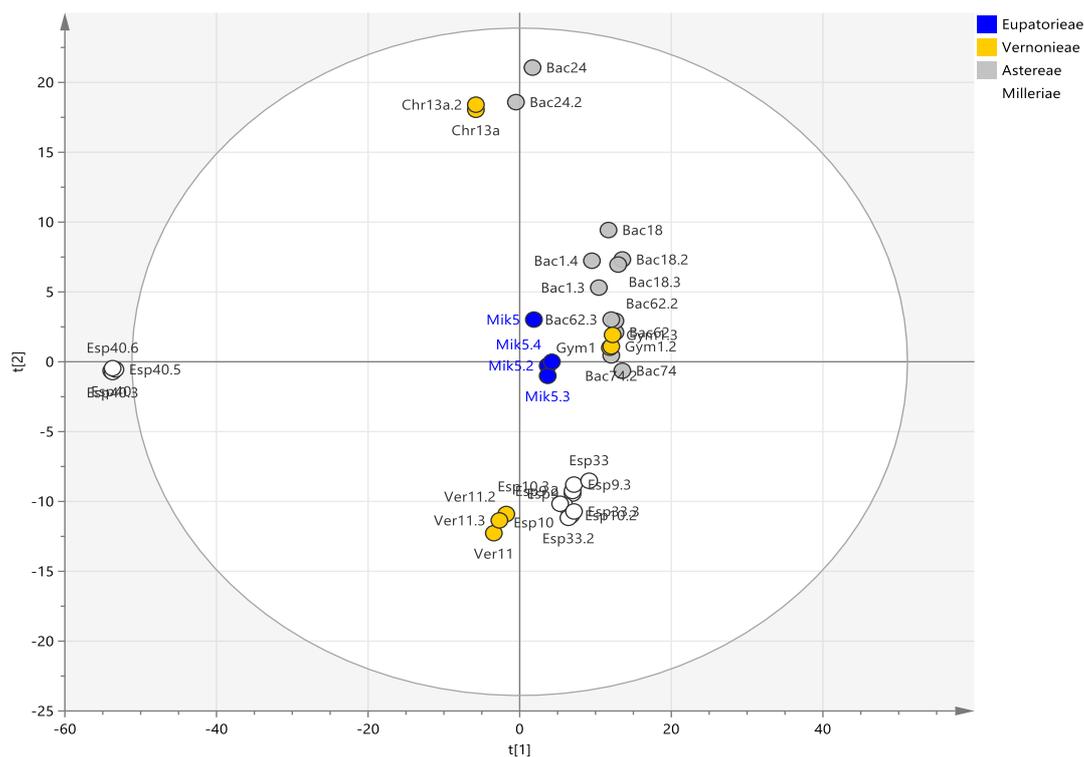


Figura 8 - Inspeção de réplicas de extratos provenientes de espécies das tribos Eupatorieae, Vernonieae, Astereae e Milleriae. A identificação dos códigos e seus respectivos IDs pode ser consultada no Apêndice A.

Assim, tendo em vista a sobreposição de picos referentes a padrões internos e às réplicas, e à ligeira semelhança entre as réplicas, quando analisadas do ponto de vista metabolômico por PCA, uma maior análise foi realizada com todas as 248 espécies de 16 tribos de Asteraceae deste trabalho, presentes no Apêndice A. Para tal análise, a matriz de dados contendo 2.046 picos provenientes do modo negativo e 2.066 provenientes do modo positivo foi analisada por PCA, HCAbp e OPLS-DA. O intuito seria visualizar agrupamento entre espécies da mesmo tribo de acordo com seus constituintes químicos. Deste modo, por meio da PCA não foi possível visualizar qualquer agrupamento entre as espécies de modo que se pudesse realizar comparação com dados da literatura (PCA não mostrada), sejam aqueles provenientes de filogenia (BREMER, 1996; PANERO et al., 2014), ou aqueles com base em caracteres químicos (CALABRIA et al., 2006). Além disso, também foi realizado HCAbp com todas as mencionadas espécies e, do mesmo modo, não foi possível visualizar clara distinção entre as tribos (HCA não mostrada). Os prováveis motivos para tais resultados podem ser diversos, tal como origem do material vegetal (não padronização da coleta, por exemplo), extração química das folhas, falhas analíticas no sistema LC-MS e mau processamento dos dados brutos. Por exemplo, assim, tal como será melhor abordado no capítulo 4, os metabólitos secundários sofrem influência do meio ambiente (GOBBO-NETO e LOPES, 2007), tal como

ritmo circadiano, luminosidade, disponibilidade hídrica, sazonalidade, dentre outras. Decorre disso que, caso não tenha ocorrido padronização na coleta, pode ter sido gerado uma alteração metabólica não relacionada com a taxonomia, mas com o meio ambiente. Além disso, nenhuma destas espécies foi coletada e prontamente congeladas em N₂ líquido para cessar seu metabolismo, tal como protocolado pela literatura (DE VOS et al., 2007), o que pode ter gerado maior quantidade de produtos de degradação metabólicos, os quais não se relacionam necessariamente com a classificação taxonômica (MOREIRA et al., 2018). De fato, foi constatado que ácidos cafeoilquínicos, derivados biossinteticamente do ácido *p*-cumárico, se acumulam nas folhas após a coleta, caso não sejam congeladas (MOREIRA et al., 2018). Ademais, a extração foi realizada e seguida por *cleanup*, o que pode ter descartado metabólitos importantes para quimiotaxonomia. Entretanto, as causas do não agrupamento observado entre as tribos são hipotéticas e serão melhor investigadas futuramente.

Persistindo com as análises não supervisionadas e com a quimiotaxonomia, a análise dos dados foi realizada, mas com uma menor quantidade de tribos, selecionando-se aquelas mais representativas deste trabalho: Eupatorieae, Astereae, Millerieae, Vernonieae, Heliantheae e Barnadesieae, táxon basal em Asteraceae. A PCA realizada com tais tribos mostra que houve certo agrupamento entre as tribos Vernonieae, Astereae e Millerieae (Figura 9), com algumas exceções. Ainda assim, do ponto de vista metabólico, há pouca tendência de agrupamento entre estas tribos analisadas, bem como pouca separação entre as tribos.

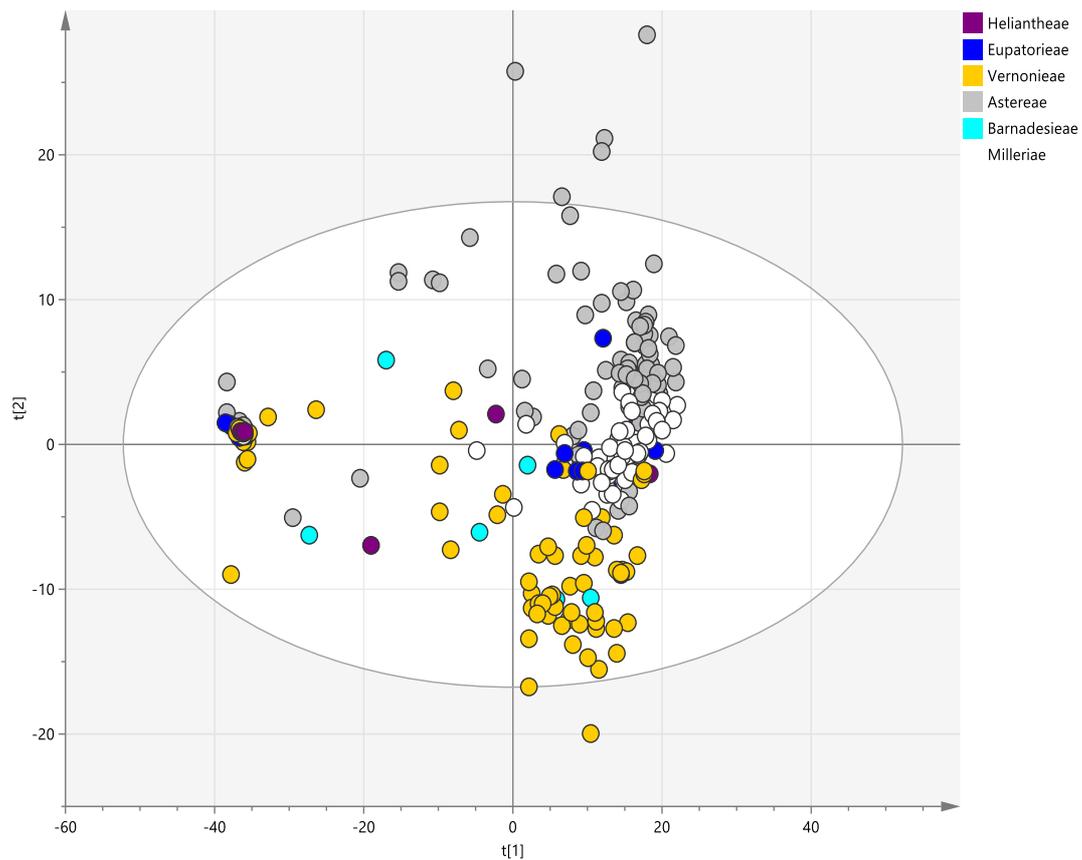


Figura 9 – PCA (4 componentes, 0,29) com espécies de Asteraceae das tribos Heliantheae, Eupatorieae, Vernoniaeae, Astereae, Barnadesieae e Milleriae. Percebe-se que, do ponto de vista metabólico, há pouca tendência de agrupamento entre estas tribos analisadas. Ainda assim, Vernoniaeae, Milleriae e Astereae tendem a se localizar em clusters, embora não totalmente separados.

Foi também realizada diversas HCABp pelo método de Ward e distância Euclidiana com tais tribos com o intuito de visualizar agrupamento entre as espécies de acordo com a classificação taxonômica de Asteraceae ao nível de tribos. Na ocasião, foi esta análise foi constantemente inspecionada em termos de amostras que apresentavam alto valor de desvio padrão em termos de valor-p AU (*approximately unbiased*) e foram removidas da análise. Assim, uma HCABp mais simples e com menores erros era gerada subsequentemente em relação à anterior. Ainda assim, a última HCABp realizada revela pouco padrão de agrupamento e separação entre as tribos (Figura 11). Além disso, pode-se notar que a confiabilidade dos agrupamentos é baixa, revelados pelos valores-p AU e BP (*bootstrap probability*) menores que 60 ou 0 (Figura 11), o que ainda indica amostras com alto teor de desvio padrão em termos de valor-p AU (Figura 10). Na Figura 10 destaca-se em retângulo os clusters com confiabilidade maior que 95% para o conjunto de dados analisados.

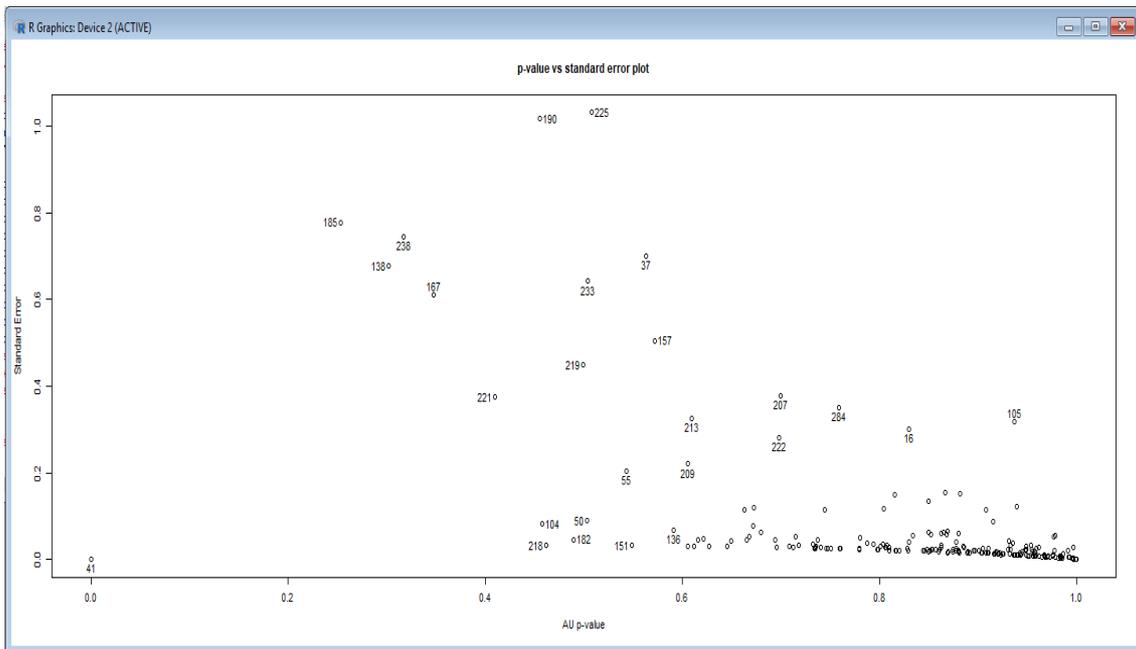


Figura 10 - Gráfico de erro padrão por valor-p AU. Os pequenos círculos são as amostras dispersas em relação ao erro e valor-p AU. Os números próximos aos círculos indicam amostras que foram deletadas, a fim de se verificar se ocorria melhora na HCAbp.

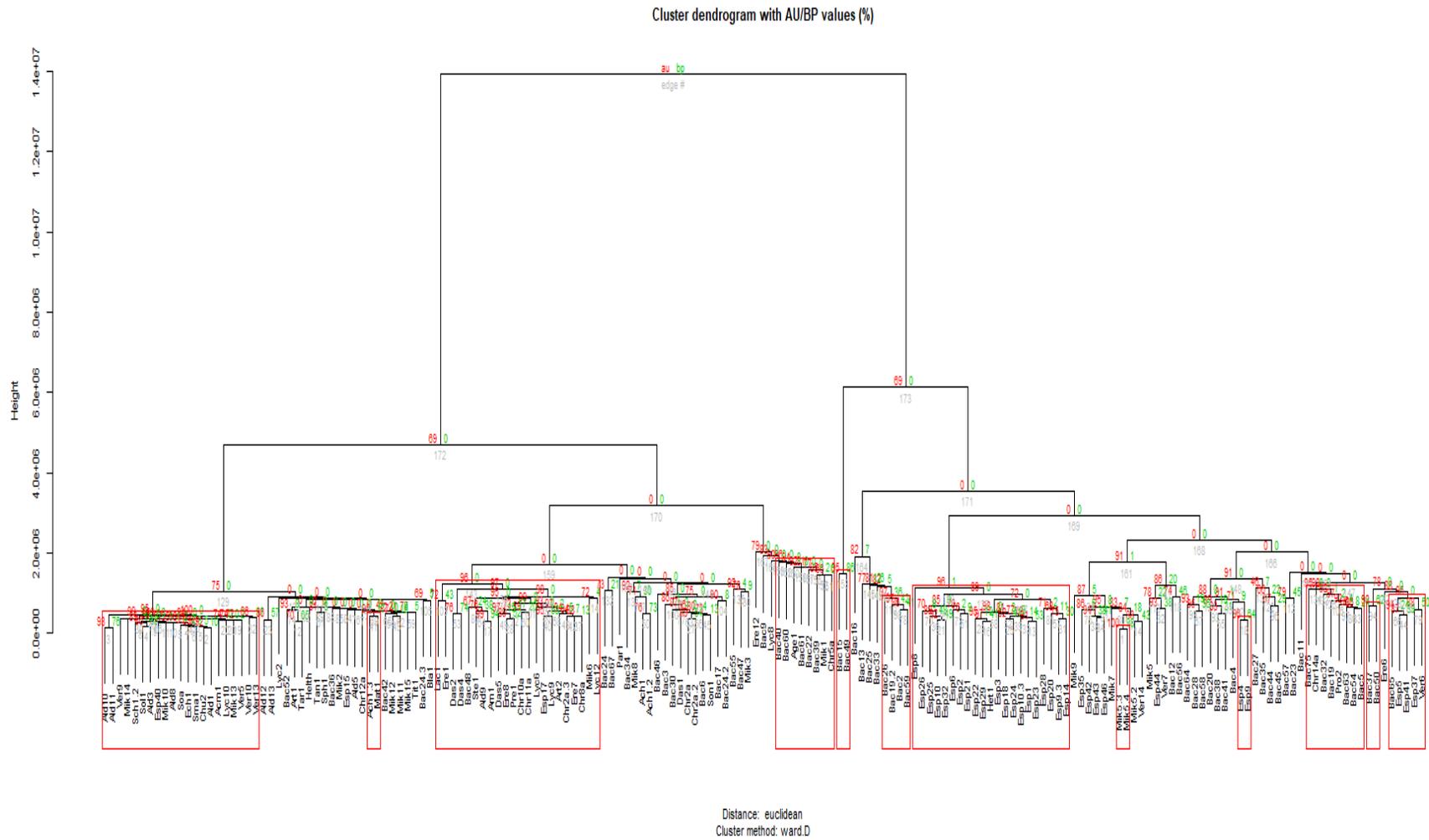


Figura 11 – HCAbp realizado com algumas espécies das tribos Barnadesieae, Eupatorieae, Astereae, Millerieae, Vernoniae e Heliantheae. Os retângulos indicam clusters com confiabilidade acima de 95%. Ainda assim, observa-se pouco agrupamento e separação das tribos de Asteraceae.

Pela HCAbp da Figura 13 não foi possível observar distinção entre as tribos, com base em seus caracteres químicos e com base nas classificações propostas por Bremer (BREMER, 1996) ou Panero e colaboradores (PANERO et al., 2014). Calabria e colaboradores também indicaram problemas no agrupamento filogenético quando construído com 389 caracteres químicos de cerca de 7.000 metabólitos descritos na literatura para 3.159 espécies de Asteraceae (CALABRIA et al., 2006). Ainda assim, mais estudos deverão ser realizados, a fim de se obter melhores conclusões a respeito da comparação da quimiotaxonomia obtida por técnicas não supervisionadas e as classificações filogenéticas de Asteraceae. Devemos também levar em consideração efeitos de fatores ambientais sobre o metaboloma, o que pode causar equívocos na quimiotaxonomia (PADILLA-GONZÁLEZ et al., 2017). Assim sendo, alguns fatores ambientais foram melhor estudados no capítulo 4, desta tese, utilizando-se como modelo *A. robusta* e *A. trichophylla* (Asteraceae).

Posteriormente às análises quimiotaxonômicas realizadas por técnicas não supervisionadas, foi realizado OPLS-DA a fim de se encontrar metabólitos discriminantes de cada tribo, ou seja, mais correlacionados com cada tribo de Asteraceae. De fato, o não agrupamento das tribos visualizado em PCA e HCAbp não significa que não existam diferenças metabólicas entre as tribos, mas apenas que as técnicas não supervisionadas não encontram padrão de agrupamento entre todas as variáveis de maneira geral e naturalmente. As técnicas supervisionadas, por outro lado, apontam variáveis latentes, ou seja, aquelas que são importantes para a discriminação de grupos, mas que não são facilmente visualizadas no conjunto das variáveis. Deste modo, fica muito clara a utilidade do OPLS-DA na análise de dados de metaboloma. Com base nisso, foram realizadas sucessivas OPLS-DA com as tribos mais representativas deste trabalho, a fim de se encontrar metabólitos discriminantes para cada uma delas. Sempre em alguma OPLS-DA foi necessário excluir alguma tribo a fim de tornar o modelo mais simples e facilitar a separação entre as tribos pela técnica empregada. Dito isso, foi possível realizar uma OPLS-DA (OPLS-DA.1) para se apontar as variáveis discriminantes correlacionadas com a tribo Barnadesieae (Figura 12).

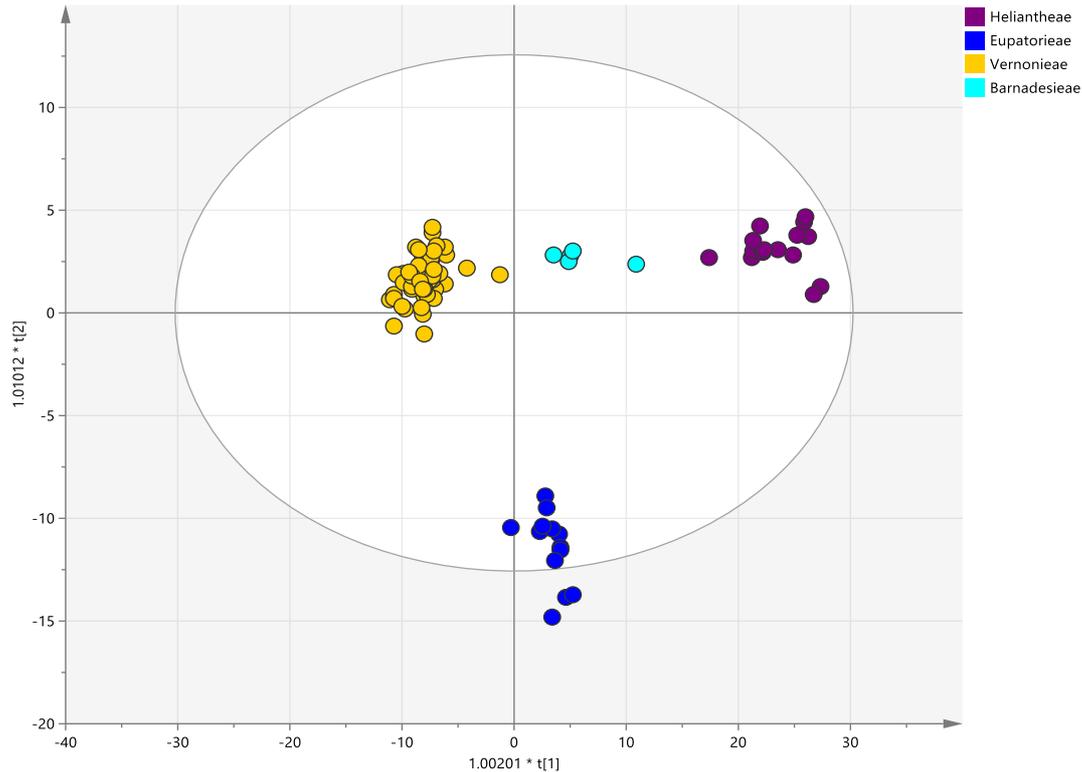


Figura 12- OPLS-DA.1 (3 + 7) ($R^2 = 0,90$ e $Q^2 = 0,37$) mostrando um agrupamento das espécies de Asteraceae com base no metaboloma das tribos Barnadesieae, Heliantheae, Eupatorieae e Vernoniaeae. Por meio desta análise é possível identificar metabólitos discriminantes de cada tribo, embora o modelo não esteja totalmente livre de *overfitting* (Q^2 relativamente baixo em relação ao R^2). Outras análises de OPLS-DA foram realizadas, por meio das quais se conseguiu maiores valores de Q^2 (Figura 13 e 14), mas sem a tribo Barnadesieae. RMSEcv < 0,29, Valor-p = $7,98 \times 10^{-10}$.

A OPLS-DA.1 (Figura 12) apresentou boa separação entre as tribos, mas um valor relativamente baixo de Q^2 , o que pode sugerir *overfitting* na análise. Ainda assim, os valores de RMSEcv e valor-p foram baixos, o que indicam baixo desvio padrão das amostras e baixa probabilidade de que a separação tenha sido ocasionada pelo acaso. Com base nesta OPLS-DA.1 foi também realizada uma classificação das espécies a fim de se verificar se elas seriam corretamente classificadas de acordo com suas tribos (Tabela 1). Por meio da OPLS-DA.1 foram classificadas corretamente, de acordo com o metaboloma, 78,3% das espécies, um valor relativamente bom, embora ainda tenha sido visualizado maiores erros justamente para Barnadesieae. Ainda assim, este foi o melhor modelo que se conseguiu construir para tal tribo e, por meio dele, gerar suas variáveis discriminantes (Tabela 2).

Tabela 1 - Classificação das espécies das tribos Heliantheae, Eupatorieae, Vernonieae e Barnadesieae, segundo o modelo OPLS-DA-1 (Figura 12), construído a partir do metaboloma de tribos de Asteraceae, utilizado para encontrar variáveis discriminantes de Barnadesieae. Destaca-se em negrito a quantidade de espécies classificadas corretamente.

Tribos	Espécies	Correto	Classificação			
			Heliantheae	Eupatorieae	Vernonieae	Barnadesieae
Heliantheae	15	100%	15	0	0	0
Eupatorieae	20	65%	6	13	1	0
Vernonieae	78	79,49%	15	1	62	0
Barnadesieae	7	57,14%	1	0	2	4
Total	120	78,33%	37	14	65	4

As variáveis discriminantes da tribo Barnadesieae foram identificadas com base no banco AsterDB (ver o item 3.3 de Material e Métodos, pág. 22) e são, provavelmente, principalmente pertencentes à classe dos diterpenos, embora também tenham sido sugeridos policetídeos, uma catequina e uma LST do tipo furanoheliangolido (Tabela 2). Esta última pouco provável, uma vez que, até o presente momento, não foi relatada tal classe na tribo Barnadesieae.

A segunda OPLS-DA.2 realizada foi necessária para encontrar variáveis discriminantes da tribo Eupatorieae (Figura 13). Neste modelo, tanto Heliantheae quanto Barnadesieae tiveram de ser excluídas a fim de se otimizar o modelo. Foi possível visualizar uma clara distinção entre as tribos (alto valor de R^2) e boa capacidade preditiva em termos de validação cruzada (alto valor de Q^2), o que não indica *overfitting*. Além disso, os baixos valores para o erro da análise (RMSEcv) e a probabilidade de significância (valor-p) também indicam bons resultados da análise (Figura 13).

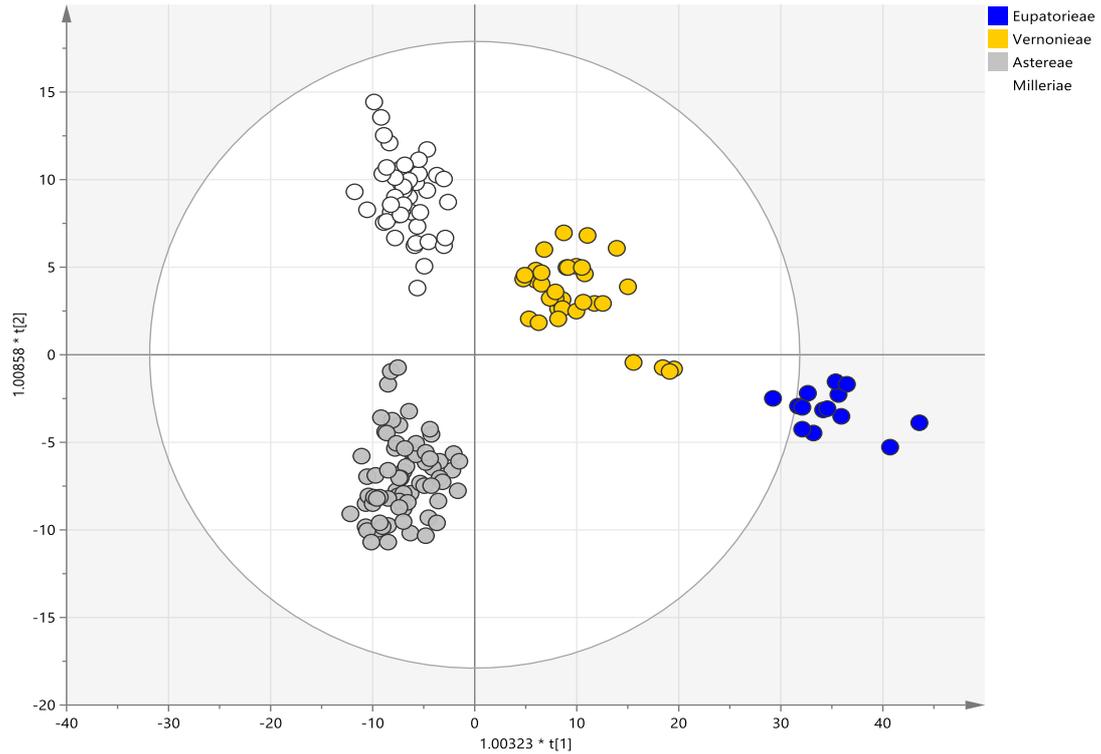


Figura 13 - OPLS-DA.2 (4 + 3) ($R^2 = 0,93$ e $Q^2 = 0,78$) mostrando um agrupamento das espécies de Asteraceae com base no metaboloma das tribos Astereae, Milleriae, Eupatorieae e Vernonieae. A análise evidencia clara separação entre os grupos (alto valor de R^2) e alta capacidade preditiva (alto valor da validação cruzada, Q^2). Os representantes da tribo Eupatorieae desta análise se distinguem claramente das demais tribos, razão pela qual se encontram fora do limite do Hotelling T2, com intervalo de confiança de 95%. RMSEcv < 0,21, Valor-p = 0,00.

Tabela 2 - Identificação das variáveis discriminantes da tribo Barnadesieae, segundo modelo OPLS-DA-1. O sobrescrito “std” se refere a “standard” (substâncias do banco de padrões do AsterBioChem, previamente isolados ou comprados), os quais foram injetados no mesmo método cromatográfico a fim de se obter valores de tempo de retenção e perfil de fragmentação. Deste modo, as desreplicações que apontam para substâncias pertencentes ao banco de padrões são mais confiáveis, uma vez que não somente houve correlação com os valores de *m/z*, mas também com o tempo de retenção (RT). Var ID = número de identificação e respectivo modo de ionização em que se detectou o pico (pos = positivo; neg = negativo). As estruturas podem ser visualizadas em <http://zinc15.docking.org/catalogs/asternp/substances/>, em “Lookup”, ou em <http://143.107.203.160:8080/>, em “Compound Search”, inserindo-se seu nome usual ou código InChIKey. N.i. = pico não identificado.

Var ID	RT	<i>m/z</i>	Identificação discriminantes de Barnadesieae (nome usual ou código InChIKey)	Classe
183 pos	4,68	122,081	n.i.	
1110 pos	14,06	369,227	n.i.	
634 pos	10,21	365,159	CDLFKJLFYKQREE-DVGHFYJSSA-N	LST-furanoheliangolido
21 pos	1,21	182,985	n.i.	
352 pos	6,83	713,155	n.i.	
33 pos	1,29	162,076	n.i.	
431 pos	8,04	116,071	n.i.	
1708 pos	20,96	285,185	UANHYLJGDUVWED-ZHACJKMWNA-N; CNXRXTXGTZPXPV-VTMWINSUNA-N	Poliacetilenos
243 pos	5,63	195,088	n.i.	
1390 pos	16,05	321,242	XJSSRXITQUJZLO-IVJAHPGJSA-N; CASWWQVDGSTYMZ-UDVNOLRYNA-N	Diterpeno e Poliacetileno

			PMUKMASMUOTCMA-FIXWDHIASA-N	Diterpeno e Poliacetileno
1541 pos	17,64	503,279	n.i.	
610 pos	9,99	401,156	n.i.	
1283 pos	15,30	287,236	kaurenal; JCAVDWHQNFTFBW-XRNRSJMDSA-N; IXIINGQFXAAIPH-XRNRSJMDSA-N;	Diterpenos
			JCAVDWHQNFTFBW-DAUOMPHXSA-N; MUDDKIQZZPCIQA-FGSPNWDHSA-N	Diterpenos
1874 pos	25,86	722,503	n.i.	
717 pos	10,79	286,201	n.i.	
1647 pos	19,34	259,115	n.i.	
1957 pos	27,19	467,315	n.i.	
944 pos	12,60	333,132	IJCWCJRLHJAVFD-UHFFFAOYNA-N	Catequina
1075 pos	13,73	305,247	n.i.	
1308 pos	15,48	275,162	n.i.	

1360 pos	15,83	331,19	PKAIECBWQZFYRP-SZUIKGGKSA-N; AYCMMSGIYHDUSS-MPIDJNAQSA-N	Derivado do ácido cinâmico e diterpeno
75 pos	1,98	113,06	n.i.	
1786 pos	23,62	447,253	n.i.	
1010 pos	13,10	289,141	n.i.	
962 pos	12,74	303,231	ent-kaurenoic acid; pimaric acid; abietic acid	Diterpenos
1969 pos	27,31	507,272	n.i.	

Assim, por meio da OPLS-DA.2 foi possível indicar com confiabilidade bons discriminantes para a tribo Eupatorieae (Tabela 4). Os representantes da tribo Eupatorieae desta análise se distinguem claramente das demais tribos, razão pela qual se encontram fora do limite do Hotelling T2. Além disso, houve um bom acerto da classificação das espécies em suas respectivas tribos (Tabela 3), com 84% de acerto total, dos quais 70% de acerto para Eupatorieae.

Tabela 3 - Classificação das espécies das tribos Eupatorieae, Vernonieae, Astereae e Milleriae, segundo o modelo OPLS-DA-2 (Figura 13), construído a partir do metaboloma de tribos de Asteraceae, utilizado para encontrar variáveis discriminantes de Eupatorieae.

			Classificação			
Tribos	Espécies	Correto	Eupatorieae	Vernonieae	Astereae	Milleriae
Eupatorieae	20	70%	14	3	3	0
Vernonieae	78	89,74%	3	70	5	0
Astereae	94	82,98%	0	16	78	0
Milleriae	58	82,76%	1	8	1	48
Total	250	84%	18	97	87	48

Já os picos apontados como discriminantes de Eupatorieae (Tabela 4) foram identificados principalmente em sesquiterpenos, flavonoides glicosilados, LST dos tipos guaianolido, eudesmanolido, *seco*-eudesmanolido e heliangolido, além de diterpenos, principalmente com base no AsterDB. Os flavonoides glicosilados isorhamnetina-3-glucosídeo e quercitrina-*O*-rhamnosídeo (Tabela 4), foram identificados com base em dados de tempo de retenção e $[M+H]^+$ de padrões externos da biblioteca de substâncias do AsterBioChem. Deste modo, tais identificações têm bons níveis de confiabilidade (SUMNER et al., 2007; SALEK et al., 2013).

Por fim, foi realizada mais um OPLS-DA.3 para se apontar substâncias discriminantes de Vernonieae, Milleriae, Astereae e Heliantheae (Figura 14). Neste modelo, foram excluídas as tribos Barnadesieae e Eupatorieae, a fim de se otimizar o modelo. Também foi possível visualizar clara separação entre as tribos e boa capacidade preditiva, sendo os valores de R^2 , Q^2 , RMSEcv e valor-p muito semelhantes aos do modelo anterior. Entretanto, foi escolhida a OPLS-DA.3 para se apontar os metabólitos discriminantes de Vernonieae, Milleriae e

Astereae, pois seu valor de Q^2 e RMSEcv foram ligeiramente melhores que aqueles visualizados no modelo OPLS-DA.2.

Tabela 4 - Identificação das variáveis discriminantes da tribo Eupatorieae, segundo modelo OPLS-DA-2. O sobrescrito “std” se refere a “standard” (substâncias do banco de padrões do AsterBioChem, previamente isolados ou comprados), os quais foram injetados no mesmo método cromatográfico a fim de se obter valores de tempo de retenção e perfil de fragmentação. Deste modo, as desreplicações que apontam para substâncias pertencentes ao banco de padrões são mais confiáveis, uma vez que não somente houve correlação com os valores de *m/z*, mas também com o tempo de retenção (RT). Var ID = número de identificação e respectivo modo de ionização em que se detectou o pico (pos = positivo; neg = negativo). As estruturas podem ser visualizadas em <http://zinc15.docking.org/catalogs/asternp/substances/>, em “Lookup”, ou em <http://143.107.203.160:8080/>, em “Compound Search”, inserindo-se seu nome usual ou código InChIKey. N.i. = pico não identificado.

Var ID	RT	<i>m/z</i>	Identificação discriminantes Eupatorieae (nome usual ou código InChIKey)	Classe
1782 pos	23,61	284,295	n.i.	
390 pos	7,55	495,113	n.i.	
30 pos	1,25	641,17	n.i.	
1103 pos	14,01	127,112	n.i.	
1995 pos	28,09	509,403	n.i.	
1922 pos	26,63	605,421	n.i.	
1985 pos	27,69	301,216	(-)-dehydroabiatic acid; RJIPNPHMQGDUBW-RFGKEDTNSA-N e isômeros	Diterpenos
560 pos	9,42	449,107	quercitrin-O-rhamnosideo ^{std} ; cynaroside; luteolin 7-O-glucoside; astragalin; quercitrin	Flavonoides-glicosilados
			OVSQVDMCBVZWGM-QSOFNFLRSA-N e isômeros; VNTMXJLNIJFLIF-QNDFHXLGSA-N e isômeros	Flavonoides-glicosilados
1501 pos	17,11	273,111	n.i.	

398 pos	7,61	231,101	n.i.	
543 pos	9,22	457,206	n.i.	
1887 pos	26,09	239,149	n.i.	
1733 pos	21,58	149,132	n.i.	
515 pos	8,92	241,704	n.i.	
1655 pos	19,58	361,234	AQBQBBLJTDSVLC-RVXFQSDISA-N	Diterpeno
450 pos	8,26	479,118	isorhamnetin-3-glucoside ^{std}	Flavonóide-glicosilado
1541 pos	17,64	503,279	n.i.	
1736 pos	21,63	191,179	n.i.	
1968 pos	27,30	421,352	n.i.	
1747 pos	22,22	189,164	n.i.	
846 pos	11,94	301,216	(-)-dehydroabietic acid; RJIPNPHMQGDUBW-RFGKEDTNSA-N e isômeros	Diterpenos
940 pos	12,58	273,221	n.i.	

840 pos	11,87	147,044	Coumarin	Derivado do ácido cinâmico
654 pos	10,39	279,122	artecanin e canin; ANTYQIKMKBUESI-NVULAGHYSA-N; DLMKPBDNRSJGBM-DBMMHUENSA-N	LST guaianolido, acíclica,
			WFYBOEGLIBIGPA-ZPBKDSQCSA-N; RTOGTHJTQOMSQZ-QMPIGLIWSA-N e estereoisômeros;	eudesmanolido, heliangolido
			OURDJVKQQOPENX-OTOORWMLSA-N	
1548 pos	17,77	445,256	n.i.	
1488 pos	17,01	279,159	YVXVQMWUXKMTIT-VBCXXGINSAN; RCKXDYPRQOIBKT- CUOATXAZSAN	LST <i>seco</i> -eudesmanolido
338 pos	6,73	203,143	n.i.	
1506 pos	17,16	255,21	n.i.	
1778 pos	23,39	205,195	elemene, cadinene, bourbonene, germacrene, bisbolene, humulene, isocomene, isocaryophyllene, cubebene, farnesene,	Sesquiterpenos
			gamma-curcumene, selinene, bulnesene, zingiberene, bergamotene, himachalene, caryophyllene, santalene,	Sesquiterpenos
			gurjene, valencene, biciclogermacrene, amorpho-4,11-diene e estereoisômeros	Sesquiterpenos
311 pos	6,51	209,153	n.i.	

1963 pos	27,25	344,228	n.i.	
1610 pos	18,54	329,174	SIVSTSIMZBDXBG-CJKOKKRSA-N e estereoisomeros	LST guianolido
1510 pos	17,22	389,268	AGFZWPUDRZOSPR-KYKUXKSPSA-N e estereoisomeros	Diterpeno
1479 pos	16,88	253,195	n.i.	
1721 pos	21,21	337,104	n.i.	

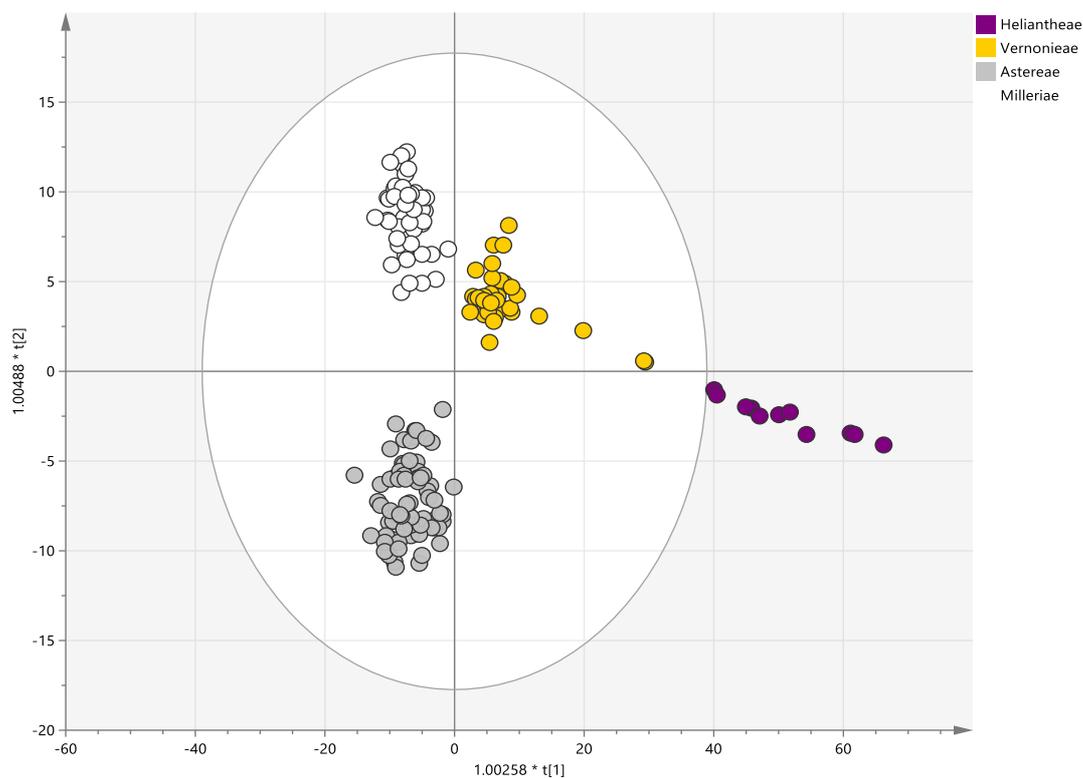


Figura 14 - OPLS-DA.3 (4 + 3) ($R^2 = 0,92$ e $Q^2 = 0,79$) mostrando um agrupamento das espécies de Asteraceae com base no metaboloma das tribos Astereae, Milleriae, Heliantheae e Vernoniae. A análise evidencia clara separação entre os grupos (alto valor de R^2) e alta capacidade preditiva (alto valor da validação cruzada, Q^2). Os representantes da tribo Heliantheae desta análise se distinguem claramente das demais tribos, razão pela qual se encontram fora do limite do Hotelling T2, com intervalo de confiança de 95%. RMSEcv < 0,20, Valor-p = 0,00.

Com base na OPLS-DA.3 também foi realizada uma classificação das espécies com base em suas tribos (Tabela 5), o que revelou acerto de 80% ou mais para todas as tribos, e acertou de 83,7% para todas as espécies. Deste modo, foi gerada uma lista de substâncias discriminantes para todas estas tribos (Tabela 6).

Tabela 5 – Classificação das espécies das tribos Heliantheae, Vernoniae, Astereae e Milleriae, segundo o modelo OPLS-DA-3 (Figura 14), construído a partir do metaboloma de tribos de Asteraceae, utilizado para encontrar variáveis discriminantes de Heliantheae, Vernoniae, Astereae e Milleriae. Destaca-se em negrito a quantidade de espécies classificadas corretamente.

			Classificação			
Tribos	Espécies	Correto	Heliantheae	Vernoniae	Astereae	Milleriae
Heliantheae	15	80,0%	12	0	3	0
Vernoniae	78	84,6%	5	66	7	0
Astereae	94	84,0%	9	6	79	0
Milleriae	58	82,8%	1	7	2	48
Total	245	83,7%	27	79	91	48

Tabela 6 - Identificação das variáveis discriminantes da tribo Heliantheae, Vernoniae, Astereae e Millerieae, segundo modelo OPLS-DA-3. O sobrescrito “std” se refere a “standard” (substâncias do banco de padrões do AsterBioChem, previamente isolados ou comprados), os quais foram injetados no mesmo método cromatográfico a fim de se obter valores de tempo de retenção e perfil de fragmentação. Deste modo, as desreplicações que apontam para substâncias pertencentes ao banco de padrões são mais confiáveis, uma vez que não somente houve correlação com os valores de *m/z*, mas também com o tempo de retenção (RT). Var ID = número de identificação e respectivo modo de ionização em que se detectou o pico (pos = positivo; neg = negativo). As estruturas podem ser visualizadas em <http://zinc15.docking.org/catalogs/asternp/substances/>, em “Lookup”, ou em <http://143.107.203.160:8080/>, em “Compound Search”, inserindo-se seu nome usual ou código InChIKey. N.i. = pico não identificado.

Var ID	RT	<i>m/z</i>	Identificação de Heliantheae (nome usual ou código InChIKey)	Classe
38 pos	1,34	186,994	n.i.	
1975 pos	27,51	307,193	n.i.	
318 pos	6,55	119,09	n.i.	
192 pos	4,78	207,138	2-phenylethyl 3-methylbutanoate	
1874 pos	25,86	722,503	n.i.	
396 pos	7,59	191,179	n.i.	
1286 pos	15,32	317,102	eriodictyol 7,3'-dimethyl ether ^{std} ; persicogenin	Flavonoides
1922 pos	26,63	605,421	n.i.	
588 pos	9,75	193,049	isoscopoletin; scopoletin	Derivados do ácido cinâmico

1543 pos	17,70	431,24	n.i.	
338 pos	6,73	203,143	n.i.	
1995 pos	28,09	509,403	n.i.	
679 pos	10,63	303,05	quercetin ^{std} ; morin ^{std} ; 6-hydroxyluteolin	Flavonoides
Var ID	RT	m/z	Identificação de Vernoniae (nome usual ou código InChIKey)	Classe
2056 pos	29,67	409,383	n.i.	
653 pos	10,37	379,138	diacetilpiptocarfol (hirsutinolideo) ^{std}	LST furanoheliangolido
2063 pos	29,89	122,081	n.i.	
492 pos	8,72	447,092	n.i.	
516 pos	8,93	381,154	n.i.	
1960 pos	27,21	487,359	n.i.	
525 pos	9,01	477,102	n.i.	

714 pos	10,78	625,155	n.i.	
601 pos	9,89	613,144	n.i.	
1063 pos	13,59	287,055	morin [M-OH] ^{std} ; luteolin, quercetin, scutellarein, kaempferol, isoscutellarein	Flavonoides
2047 pos	29,25	243,121	n.i.	
88 pos	2,23	319,186	ASZVUCPOPSJEJZ-OCVRCQPFSAN e estereoisomeros	LST germacranolido
			ZPHXKYPRCCFEEM-CKFGLLJZSAN e estereoisomeros	LST eudesmanolido
496 pos	8,74	351,143	n.i.	
560 pos	9,42	449,107	quercitrin-O-rhamnoside ^{std} ; cynaroside; quercitrin, triofolin, astragalin, luteolin 7-O-glucoside	Flavonoides glicosilados
984 pos	12,95	377,159	16a-(1',2'-epoxy-1'-methylprop)-eremantholide ^{std} ; KOMKVTZGCJDXPN-RYZYPSNUSAN, eupatundin,	LST eremantolido
			ILVSKWHMMYOGQO-IWJZJBDRSAN; WLFVPPSXKYSJGO-WJVSDBCXSAN,	LST germacranolido,
			HHMGKVCAMWORU-BVLUOJBOSAN, JWUUXOHPRFBRM-NVNZQMSSAN	eudesmanolido, guaianolido
			AFRRMOVQZXLLAY-HZACFNRPSAN e estereoisomeros	

1485 pos	16,95	419,17	UKUKZUAKMMJYFZ-NVGRPZBPSA-N, euparotin acetate, JIEVJXRUYDZFTD-CVQTXJQTSA-N e estereoismeros	LST guaianolido e germacranolido
880 pos	12,24	215,106	Ichthyothereol	Poliino
888 pos	12,29	301,07	chrysoeriol ^{std} ; hipidulin ^{std} ; diosmetin ^{std} ; rhamnazin ^{std} ; geraldol; sorbifolin; luteolin 7-methyl ether	Flavonóides
1952 pos	27,13	443,333	n.i.	
1948 pos	27,04	355,281	n.i.	
782 pos	11,40	125,096	n.i.	
1984 pos	27,67	439,381	n.i.	
587 pos	9,75	415,135	n.i.	
721 pos	10,81	523,287	n.i.	
1430 pos	16,38	423,201	ALKJEOWDCFQCTJ-DGRDYZOHSAN; UVRIFAYGSSDVER-QGYAONMESA-N	LST germacranolido
			KYJZJPDRJIRSOL-SEXCCMIOSA-N e estereoismeros	
2010 pos	28,25	371,257	n.i.	

108 pos	3,06	357,094	n.i.	
1078 pos	13,75	301,07	chrysoeriol ^{std} ; hipidulin ^{std} ; diosmetin ^{std} ; rhamnazin ^{std} ; geraldol; sorbifolin; luteolin 7-methyl ether	Flavonóides
Var ID	RT	m/z	Identificação de Astereae (nome usual ou código InChIKey)	Classe
167 pos	4,52	251,167	IKUJXDOLQGOQFB-MODKSPOCSA-N; arbusculin A; LEXBBZCFWJNTGC-FOENDFONSA-N;	Sesquiterpeno, LST germacranolido,
			JWJOVYZAXACWBK-XYFSKKJPSA-N; KVWUOMNJFDNWQM-RMIQQSQVSA-N; MTIXBBDLFRVGOQ-WQFFSLPPSA-N;	xantanolido, <i>seco</i> -eudesmanolido,
			FEYQLIJGWGWQNT-OVIZQLOJSA-N; FEYQLIJGWGWQNT-ZGILNNDFA-N; KQEADOUJYBGFC-AIHGEPKZSA-N;	eudesmanolido
			UEOBEAJSRJXKEQ-XPCJQDJLSA-N e estereoisomeros	
1545 pos	17,71	305,247	RXXOJGKRNHHLDL-WVMNWWHYSA-N; GMXGAPWNZFBSDY-NQOFJFAASA-N; GWTHBPSLWRXRSF-KVPLUYHFSA-N;	Diterpeno
			SSGGJEQQCQEEEL-JTBPYRKJSA-N; KARUSPOBGJZEMI-INTGFUSFSA-N; OJSKJQFODPKTBT-XUFBONQXSA-N;	
			NOMCEPDYZFZNLH-JGFIGGHZSA-N; IQQKVAVYUIDKR-UQEPQNNOSA-N e estereoisomeros	

1708 pos	20,96	285,185	UANHYLJGDUVWED-ZHACJKMUNA-N; CNXRXTXGTZXPV-VTMWINSUNA-N	Poliacetilenos
512 pos	8,91	277,216	n.i.	
1173 pos	14,50	345,096	nevadensin; pedunculin; 2-(3,4-dimethoxyphenyl)-3,5-dihydroxy-7-methoxychromen-4-one; eupatorin	Flavonóides
130 pos	3,64	177,055	n.i.	
691 pos	10,70	347,076	SWRUBNXINAMYJO-UHFFFAOYNA-N; dihydroquercetin 3-acetate	Flavonóides
866 pos	12,11	361,092	acerosin; sudachitin	Flavonóides
542 pos	9,20	287,091	sakuranetin; isosakuranetin	Flavonóides
1817 pos	24,40	285,221	n.i.	
535 pos	9,13	273,221	n.i.	
1722 pos	21,22	217,195	n.i.	
1531 pos	17,43	285,221	n.i.	
1794 pos	23,82	305,247	RXXOJGKRNHLDL-WVMNWWHYSA-N; GMXGAPWNZFBSY-NQOFJFAASA-N; SSGGJEQQCEEEL-JTBPYRKJSA-N;	Diterpenos

			KARUSPOBGJZEMI-INTGFUSFSA-N; IQQKVAVYUIDKR-UQEPQNNOSA-N; QFNNNZPXWWTDRE-IHIZMXTDSA-N	Diterpenos
1604 pos	18,44	449,286	PUQJNULVCTWMIY-RHFDULIZSA-N	Diterpeno
973 pos	12,87	283,17	n.i.	
1907 pos	26,39	271,241	n.i.	
1654 pos	19,56	305,247	n.i.	
1844 pos	25,23	337,237	QFOKXRAQWYQLQM-VIRHNSBINA-N; ACCLBNYEFRAUCK-BIDMWUBCSA-N	Polieno
163 pos	4,42	215,081	n.i.	
1284 pos	15,31	323,258	VAKACABDOYKFHJ-MAPHFDCTSA-N; OCUPFQFFVVEGAS-HZTYIQCYSA-N; JRMZVZSBORMZSD-RUSOHIBLSA-N;	Diterpenos
			VAKACABDOYKFHJ-JLPZLSSXSA-N; JRMZVZSBORMZSD-QMBFBYQJSA-N	Diterpenos
1854 pos	25,40	287,237	kaurenal; IXIINGQFXAAIPH-XRNRSJMDSA-N; JCAVDWHQNFTFBW- DAUOMPHXSA-N	Diterpenos
			MUDDKIQZZPCIQA-FGSPNWDHSA-N	Diterpenos
1140 pos	14,22	217,105	n.i.	

836 pos	11,83	331,081	tricin; 2-(3,4-dihydroxyphenyl)-5-hydroxy-3,6-dimethoxychromen-4-one; quercetin 3,4'-dimethyl ether; jaceosidin; desmethoxycentaureidin; cirsiolol	Flavonóides
1437 pos	16,43	285,221	n.i.	
370 pos	7,28	277,07	n.i.	
Var ID	RT	m/z	Identificação de Millerieae (nome usual ou código InChIKey)	Classe
602 pos	9,90	710,133		
1721 pos	21,21	337,104		
1242 pos	15,06	619,096		
1171 pos	14,49	589,097		
1205 pos	14,82	617,092		
816 pos	11,67	355,066		
1076 pos	13,74	573,103		
1468 pos	16,75	617,092		

636 pos	10,24	629,092		
1169 pos	14,48	603,076		
645 pos	10,30	579,134		
1478 pos	16,85	917,119		
1435 pos	16,41	389,159	KTIPQHCMSXAIBC-XGHQBKJUSA-N	
352 pos	6,83	713,155		
907 pos	12,43	261,112	VCMZMMIHEOEXSR-GEEBZKHNSA-N; IQFFYN SHANWLIR-FJKOOJNLSA-N; SPZKJZJNPQNEKT-WRXSAAJRSA-N	LST germacranolido, guaianolido
292 pos	6,28	329,065		
1734 pos	21,59	253,195		
1479 pos	16,88	253,195		
1319 pos	15,57	397,125		
1321 pos	15,59	459,162		
1411 pos	16,20	349,127		

1810 pos	24,16	345,169	deoxyelephantopin; LLFFQXOEXCHHKX-ZVMMMAAASA-N; FKDIIXZIKCNXAT-SLFMQDIPSA-N;	LST germacranolido, guaianolido
			DWYSOHJKGRHGTF-SJBKCFBNSA-N; JEAWGNYMYMHNGM-LSLHFRIISA-N; FKDIIXZIKCNXAT-USVHXFPUSA-N	
			JEAWGNYMYMHNGM-NYVXQQPPSA-N e estereoisomeros	
1168 pos	14,45	345,132	15-desoxy-goyazensolide ^{std} ; MSTZNVVCBOEAGA-ZZTPXLKCSA-N; JMUOPRSXUVOHFE-MCYOVBASSA-N	LST furanoheliangolido, guaianolido,
			RYBHZNMPMHOBAR-MFRWKCHVSA-N e estereoisomeros	
895 pos	12,34	375,107	hymenoxin; quercetin 3-methyl ether	Flavonóides
1186 pos	14,66	317,065	nepetin; isorhamnetin; tamarixetin; pedaltin; onopordin; 6-methoxykaempferol	Flavonóides
1436 pos	16,43	289,107		

De acordo com a OPLS-DA.3 e a base de dados AsterDB, poucas variáveis puderam ser identificadas para a tribo Heliantheae, embora com grande confiabilidade dois flavonoides, sendo eles o eriodictiol 7,3'-dimetil eter e a quercetina, ou seu estereoisômero morina, todos presentes em na biblioteca de substâncias puras do AsterBioChem (Tabela 6). Em Vernonieae, também foram identificadas com bom nível de identificação metabolômica (SUMNER et al., 2007; SALEK et al., 2013) outros flavonoides, como o quercitrina-*O*-rhamnosídeo e crisoeriol, ou seus estereoisômeros, e também uma LST do tipo eremantolido, 16 α -(1',2'-epoxy-1'-metilprop)-eremantolideo, e outra LST do tipo hirsutinolido, diacetilpiptocarfol (Tabela 6). É importante destacar que, tal como foi abordado na introdução deste capítulo, as LST do tipo eremantolido e hirsutinolido são bons marcadores taxonômicos de Vernonieae (KELES et al., 2010; BOHMAN et al., 1981; KOS et al., 2006), fato que se corrobora com as análises e validam o emprego de OPLS-DA em apontar discriminantes taxonômicos.

Na tribo Milleriae foi identificado o metabólito 15-desoxi-goyazensolideo, da classe das LST do tipo furanoheliangolido, de acordo com o banco de padrões (Tabela 6), e o restante dos outros supostos metabólitos (*putative metabolites*) seriam pertencentes a outras LST dos tipos germacrolido ou guaianolido. Além disso, é importante destacar que o 15-desoxi-goyazensolideo é marcador do gênero *Lychnophora* e *Eremanthus*, ambos da tribo Vernonieae, mas não foi apontado como discriminante em tal tribo. É importante destacar que o OPLS-DA é uma análise estatística multivariada que emprega validação cruzada. Sendo assim, probabilisticamente, o 15-desoxi-goyazensolideo seria mais discriminante de Milleriae do que de Vernonieae e, além disso, poderia ser considerado um marcador taxonômico em comum para as duas tribos, o que se corrobora com a proximidade de Vernonieae (subfamília Cichorioideae) com a subfamília Asteroideae (PANERO et al, 2014, Figura 2), à qual pertence a tribo Milleriae. Quanto à tribo Astereae, todos os metabólitos foram identificados somente com base nos valores de [M+H]⁺ e [M-H]⁻ das substâncias presentes no banco AsterDB. Pode-se destacar para esta tribo principalmente metabólitos da classe dos diterpenos e flavonoides, com alguns casos de LST e um provável caso de poliacetileno e polieno.

5 - CONCLUSÃO DO CAPÍTULO 1

Pelas técnicas não supervisionadas não foi possível correlacionar o agrupamento quimiotaxonômico com as classificações de Bremer (BREMER, 1996) ou Panero (PANERO et al., 2014). Deste modo, mais investigações devem ser realizadas antes de se concluir inequivocadamente que os caracteres químicos não são bons marcadores químicos em Asteraceae ao nível da classificação de tribos. Por outro lado, diversas análises de OPLS-DA permitiram separar as tribos de Barnadesieae, Heliantheae, Vernonieae, Milleriae, Astereae e Eupatorieae, com base na diferença de seus constituintes químicos. Com base nisso, picos discriminantes de cada tribo foram indicados e desreplicados com base em bancos de massa e tempo de retenção de substâncias de Asteraceae. Tais substâncias podem ser consideradas marcadores quimiotaxonômicos das tribos, estando correlacionadas com a adaptação ao meio em que vivem suas respectivas espécies, ou com a diferenciação evolutiva entre as tribos.

CAPÍTULO 2: ESTUDOS DE INIBIÇÃO DE LDH-5 E TRIAGEM DE METABÓLITOS ATIVOS COM BASE NA ABORDAGEM METABOLÔMICA DE ESPÉCIES DE ASTERACEAE

LISTA DE FIGURAS DO CAPÍTULO 2

- Figura 15. Otto Henrich Warburg, quem descreveu o “efeito Warburg” das células cancerígenas. Figura retirada da página Wikipedia.....67
- Figura 16. Obtenção de energia pela via da fosforilação oxidativa ou glicolítica. Esquema retirado de Granchi et al., 2011.....67
- Figura 17. OPLS-DA (1 + 7) ($R^2 = 0,99$ e $Q^2 = 0,77$) para extratos foliares de espécies de Asteraceae ensaiados contra LDH-5. Foram considerados ativas (active) as amostras que inibiram mais de 45% da atividade enzimática, as outras amostras foram consideradas inativas (no activity).....78
- Figura 18. *Decision tree* realizada no Weka com base no metaboloma de 75% das espécies (grupo treino) de Asteraceae ensaiadas contra a enzima LDH-5.....94
- Figura 19. *Decision tree* com base no metaboloma de 75% das espécies (grupo treino) de Asteraceae ensaiadas contra a enzima LDH-5 após aplicar o filtro supervisionado “attribute selection” do Weka às variáveis.....94

LISTA DE TABELAS DO CAPÍTULO 2

Tabela 7. Espécies investigadas no ensaio de inibição da enzima LDH-5 e seus respectivos códigos.....	72
Tabela 8. Validação externa (27 extratos) ($P^2 = 81,48\%$) da OPLS-DA dos extratos foliares de espécies de Asteraceae avaliados contra a inibição da LDH-5.....	76
Tabela 9. Classificação dos extratos foliares de espécies de Asteraceae que foram retirados da análise (35 extratos) para a otimização do modelo de OPLS-DA para a inibição da atividade de LDH-5.....	79
Tabela 10. Identificação das variáveis discriminantes para os extratos foliares das espécies de Asteraceae classificados como “sem atividade” (no activity) e “ativo” (active), segundo modelo OPLS-DA.....	88
Tabela 11. Valores de classificação, validação e erro para os modelos 1 (todas as amostras) e 2 (75% das amostras – grupo treino) analisados por <i>Decision Tree</i> para extratos foliares de espécies de Asteraceae avaliados contra LDH-5.....	93
Tabela 12. Identificação das variáveis discriminantes para os extratos foliares das espécies de Asteraceae classificados como “sem atividade” (no activity) e “ativo” (active), segundo modelo <i>Decision Tree</i>	96

LISTA DE ABREVIATURAS, SIGLAS E SIMBOLOS DO CAPÍTULO 2

ATP	trifosfato de adenosina
COX	ciclooxigenase
HIF	fator induzido por hipóxia
LDH-5	lactato desidrogenase-5
LOX	lipoxigenase
NADH	dinucleotídeo de nicotinamida e adenina
OMS	Organização Mundial da Saúde
ONU	Organização das Nações Unidas
WHO	Organização Mundial da Saúde

RESUMO DO CAPÍTULO 2

No capítulo 2 foi analisada a metabolômica aplicada à triagem de inibidores de lactato desidrogenase-5 (LDH-5) de extratos provenientes de cerca de 80 espécies de Asteraceae. A enzima LDH-5 está superexpressada em diversos cânceres e relacionada com as suas sobrevivências. Deste modo, os objetivos foram: analisar se os extratos seriam capazes de inibir a enzima e, após esse ensaio, realizar análises metabolômicas com o intuito de se descobrir substâncias discriminantes de extratos ativos, além de realizar a predição de outros extratos que não foram ensaiados. Nos ensaios, a atividade enzimática foi mensurada de acordo com a oxidação de NADH em NADH⁺, após a conversão de piruvato em lactato (GRANCHI et al., 2011). A metabolômica foi realizada por meio da extração do metaboloma de mais de 250 espécies de Asteraceae, com aquisição dos dados químicos via UHPLC-UV/DAD-ESI-HRMS, em cromatografia de fase reversa. Os dados do metaboloma foram analisados por OPLS-DA e *decision tree*. Deste modo, a análise de inibição de LDH-5 revelou que 11, dos 80 extratos ensaiados, foram capazes de diminuir a atividade enzimática em mais de 45%, ou mesmo quase 100% (completa inibição). O modelo de OPLS-DA construído permitiu uma clara separação dos extratos ativos, dos não ativos, com base em suas constituições químicas. Os valores de validação desta OPLS-DA ($Q^2 = 0,77$; $P^2 = 0,81$) indicam um modelo robusto e com alta capacidade preditiva. Com base nisso, os outros extratos não ensaiados tiveram suas atividades preditas. Além disso, foram propostas algumas substâncias discriminantes dos extratos considerados ativos. As análises por *decision tree* igualmente tiveram bons valores de validação ($Q^2 = 0,72$; $P^2 = 0,56$, sem aplicação do algoritmo “*attribute selection*”), o que também permitiu indicar prováveis substâncias químicas discriminantes dos extratos ativos. Tais substâncias ativas serão isoladas futuramente e terão suas atividades confirmadas.

SUMMARY OF CHAPTER 2

In Chapter 2 the metabolomics applied to the screening of lactate dehydrogenase-5 (LDH-5) inhibitors from extracts from about 80 Asteraceae species were analyzed. The LDH-5 enzyme is overexpressed in several cancers and related to its survival. Thus, the objectives were: to analyze if the extracts would be able to inhibit the enzyme and, after this test, carry out metabolomic analyzes with the purpose of discovering discriminant substances of active extracts, in addition to predicting other extracts that were not tested. In the assays, the enzymatic activity was measured according to the NADH oxidation in NADH +, after the conversion of pyruvate to lactate (GRANCHI et al., 2011). The metabolomics were performed by extracting the metabolome of more than 250 Asteraceae species, with the acquisition of chemical data via UHPLC-UV/DAD-ESI-HRMS, in reverse phase chromatography. The metabolome data were analyzed by OPLS-DA and decision tree. Thus, LDH-5 inhibition analysis revealed that 11 out of the 80 extracts tested were able to decrease enzyme activity by more than 45%, or even almost 100% (complete inhibition). The built-in OPLS-DA model allowed a clear separation of active and non-active extracts based on their chemical constitutions. The validation values of this OPLS-DA ($Q^2 = 0.77$; $P^2 = 0.81$) indicate a robust model with high predictive capacity. Based on this, the other untested extracts had their activities predicted. In addition, some discriminant substances of the extracts considered active have been proposed. The decision tree analyzes also had good validation values ($Q^2 = 0.72$, $P^2 = 0.56$, without application of the "attribute selection" algorithm), which also allowed to indicate probable discriminant chemical substances of the active extracts. Such active substances will be isolated in the future and their activities confirmed.

1 – INTRODUÇÃO

1.1 – A enzima lactato desidrogenase-5 (LDH-5) como alvo anticancerígeno

Atualmente, o câncer é responsável por uma em cada seis mortes no mundo. Mais de 14 milhões de pessoas desenvolvem câncer todos os anos, e esse número deve subir para mais de 21 milhões de pessoas em 2030 (ONU Brasil, 2017). Além disso, o Instituto Nacional de Câncer (NCI) revelou em 2017 que as mortes por câncer aumentaram 31% no Brasil em 15 anos, de acordo com a OMS (NCI, 2017). Em 2017 a Organização Mundial da Saúde (OMS) informou que a cada ano 8,8 milhões de pessoas morrem de câncer no mundo, sendo esta a segunda doença mais mortal (WHO, 2017).

Câncer, também conhecido como neoplasia maligna, é o nome dado a um conjunto de mais de 100 doenças que têm em comum o crescimento desordenado de células, com potencial para invadir e se espalhar para outros tecidos e órgãos do corpo, fenômeno conhecido como metástase. Para que uma célula sadia se transforme em uma célula cancerígena, os genes que regulam o crescimento e a diferenciação celular têm de sofrer mutação. Nesse sentido, várias rotas bioquímicas estão envolvidas no funcionamento das células cancerígenas.

No entanto, a maioria dos cânceres possui uma característica metabólica em comum, conhecida como “efeito Warburg” (WARBURG et al., 1926; WARBURG et al., 1956), em homenagem ao fisiologista e médico alemão Otto Heinrich Warburg (Figura 15), laureado com o Nobel de medicina em 1931 por suas descobertas. O efeito Warburg consiste em uma observação experimental de que as células cancerígenas utilizam a via glicolítica para a obtenção de energia, preferencialmente à via da fosforilação oxidativa (Figura 16). Na via glicolítica, moléculas de piruvato são convertidas em lactato pela enzima lactato desidrogenase (LDH), com a produção de ATP para energia celular (WARBURG et al., 1956; GOGVADZE et al., 2010; WARD e THOMPSON, 2012) (Figura 16). Sabe-se também que a perda de função do gene supressor tumoral p53, ou sua baixa expressão, pode aumentar a glicólise, além de impedir a apoptose celular em condições de estresse metabólico (WARD e THOMPSON, 2012).



Figura 15 - Otto Heinrich Warburg, quem descreveu o “efeito Warburg” das células cancerígenas. Figura retirada da página Wikipedia (en.wikipedia.org/wiki/Otto_Heinrich_Warburg).

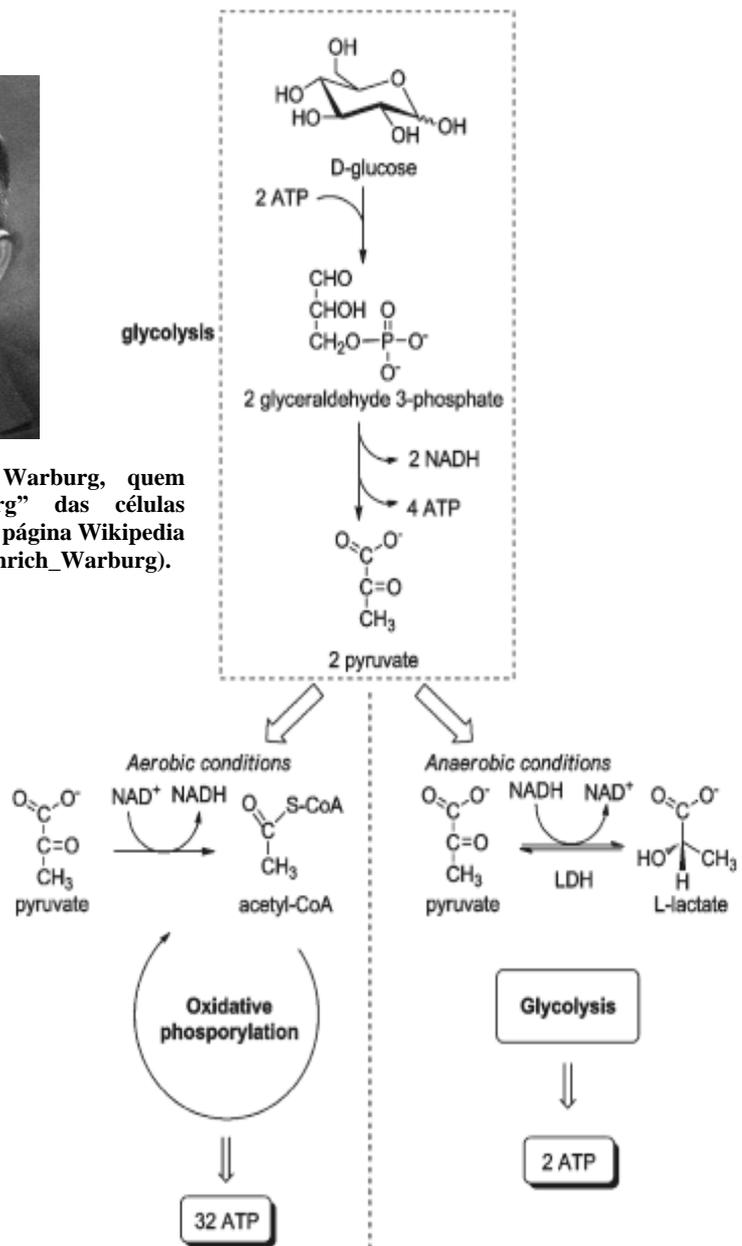


Figura 16 - Obtenção de energia pela via da fosforilação oxidativa ou glicolítica. Esquema retirado de Granchi et al., 2011.

Para que as células cancerígenas forneçam ATP pela via glicolítica deve ocorrer indução do heterodímero “fator induzido por hipóxia-1” (HIF-1) (WANG e SEMENZA, 1993), o qual estimula a expressão do gene LDH-A. Em humanos, a LDH é uma enzima tetramérica que pode existir em cinco isoformas (*hLDH1-5*). Existem apenas dois tipos de subunidades: LDH-A e LDH-B. As cinco isoformas são constituídas pelas seguintes combinações destas duas subunidades: LDH-1 = LDH-B₄; LDH-2 = LDH-AB₃; LDH-3 = LDH-A₂B₂; LDH-4 = LDH-A₃B; LDH-5 = LDH-A₄. Destas isoformas, é sabido que apenas a LDH-5 está superexpressada em muitos carcinomas (GRANCHI et al., 2011; AUGOFF et al., 2015), sendo, portanto, um importante alvo farmacológico. Deste modo, trabalhos recentes têm avaliado a inibição da LDH-5, ou de sua subunidade, frente às substâncias químicas isoladas (GRANCHI et al., 2011) e extratos de plantas (DEIAB et al., 2013). Além disso, um recente trabalho demonstrou que a inibição da LDH-A foi capaz de reverter alterações induzidas pela inflamação no câncer de cólon (MANERB et al., 2013), bem como reduzir a tumorigênese em modelos de ratos com tumor (XIE et al., 2014).

Diversas espécies da família Asteraceae foram investigadas do ponto de vista biológico frente a diversos alvos, muitos dos quais farmacológicos, tal como descrito no terceiro item da Introdução Geral. Muitas de suas substâncias e extratos brutos foram avaliados por meio de triagens que revelaram promissores efeitos farmacológicos. Entretanto, apesar desses promissores resultados, ainda há poucas substâncias investigadas quanto à sua propriedade de inibição da LDH-5, ou de sua subunidade A. Portanto, tendo em vista a ampla variedade química dos gêneros aqui investigados, tal como descrito no terceiro item da Introdução Geral, bem como da necessidade pela busca de mais fármacos anticancerígenos, faz sentido realizar uma triagem de inibição das LDH-5, a fim de ser verificar o potencial inibitório das espécies de Asteraceae deste projeto, bem como de seus metabólitos isolados.

2 – OBJETIVOS

Os objetivos deste capítulo foram:

- 1) avaliar o potencial de extratos de algumas espécies de Asteraceae em inibir a enzima LDH-5;
- 2) realizar a triagem de inibidores de LDH-5 por meio do metaboloma das espécies de Asteraceae e análises por OPLS-DA e *decision tree*;

3) realizar um estudo de predição de inibição de LDH-5 com base em OPLS-DA e *decision tree* para extratos de outras espécies de Asteraceae que não foram ensaiadas.

3 – MATERIAL E MÉTODOS

3.1 – *Material vegetal*

O material vegetal utilizado foi o mesmo do capítulo 1.

3.2 – *Extração e aquisição dos dados químicos*

A extração e aquisição dos dados foram realizadas de acordo com o capítulo 1.

3.3 – *Análise de inibição de LDH-5*

A atividade de inibição de LDH-5 (Lee Biosolution, Inc., EUA) foi determinada a 37 °C por meio da mensuração espectrométrica em 340 nm da oxidação do NADH (Sigma-Aldrich, Inc., EUA) como função do tempo. A cinética enzimática foi mensurada em 5 min, com tempo de agitação de 20 s (GRANCHI et al., 2011). O piruvato (Sigma-Aldrich, Inc., EUA) foi utilizado como substrato natural da enzima. Deste modo, quando a enzima converte piruvato em lactato, ocorre a oxidação de NADH para NAD⁺. As concentrações da análise foram: 0,00125u/μl (quantidade final no poço de 0,005 u) de LDH-5 por 0,1 M de tampão fosfato em pH 7.4; 25 μM de NADH; 1 mM de piruvato; 800 μg de cada extrato solubilizados em 1 ml de solução MeOH 10% (GRANCHI et al., 2011, adaptado). Todas as análises foram realizadas em quintuplicata e o branco (piruvato, NADH e solução de extrato, sem enzima) em triplicata. Também foi realizada uma análise em triplicata da inibição da enzima provocada pelo MeOH, sendo esta solução considerada o controle negativo da reação. Deste modo, a atividade dos extratos foi comparada com a solução de MeOH 10%, sendo considerados ativos aqueles que foram capazes de reduzir a atividade enzimática em mais de 45% (porcentagem de inibição relativa, GRANCHI et al., 2011, adaptado).

3.4 – *Processamento dos dados do LC-MS*

Os dados brutos adquiridos do LC-MS em formato .raw foram transformados de acordo com o capítulo 1.

3.5 – *Análises quimiométricas e data mining com Decision Tree*

Os dados químicos de todos os extratos após serem processados foram exportados em planilha (.csv), organizados de modo a colocar as variáveis nas colunas e amostras nas linhas, e as áreas sob os picos detectados foram normalizadas para posterior análise multivariada e

Decision Tree. Para as análises de triagem de inibidores de LDH-5, foram normalizados por *Pareto*, de acordo com o seguinte cálculo:

$$Pareto: y_{ij} = x_{ij} - \bar{x}_i / \sqrt{SD_{ij}}, \text{ onde } \bar{x} \text{ é a média da área dos picos detectados e } SD \text{ é o}$$

desvio padrão.

A PCA e OPLS-DA foram realizadas no *software* SIMCA v.13.0.3 (Umetrics – Suécia) e analisadas de acordo com o capítulo 1.

Os parâmetros escolhidos para rodar a *decision tree* foram escolhidos “*by default*”. Os resultados foram avaliados em termos de quantidade de amostras classificadas corretamente empregando métricas de estatística, tais como “*mean absolute error*” (MAE) e “*root mean squared error*” (RMSE), valores de validação cruzada (Q^2) e validação externa (P^2) (25% das amostras escolhidas aleatoriamente). Num segundo momento, e a fim de se produzir modelos mais simples e com melhores resultados, foi também utilizado o algoritmo “*attribute selection*”, presente no *software* Weka, o qual se trata de um filtro supervisionado que seleciona automaticamente os atributos (variáveis) mais correlacionados com o modelo.

4 – RESULTADOS DE DISCUSSÃO

4.1 – Resultados de triagem de inibidores de LDH-5

Para que fosse possível construir um modelo com base em metabolômica e estatística multivariada para a identificação de variáveis correlacionadas com a inibição de LDH-5, foi realizado o ensaio de alguns extratos de espécies de Asteraceae, de diferentes tribos e gêneros, a fim de garantir variabilidade química e, portanto, variabilidade de resposta de inibição. A enzima LDH-5 está superexpressada em diversos tecidos cancerígenos e sua inibição tem sido relatada pela literatura como um potencial alvo farmacológico (GRANCCHI et al., 2011). As amostras selecionadas para tal análise são as mesmas das do item anterior (quimiotaxonomia), as quais passaram por *cleanup*, podendo ser observadas no Apêndice A. O intuito de se realizar ensaio com extratos que passaram por *cleanup* seria eliminar o excesso de material graxo e pigmentos do extrato e, deste modo, falsos positivos na análise, como consta em protocolos de triagem de extratos. Deste modo, a análise de inibição de LDH-5 revelou que 11, dos 80 extratos foliares de espécies diferentes de Asteraceae, foram capazes de diminuir a atividade enzimática em mais de 45%, ou mesmo quase 100% (completa inibição) (Tabela 7). Contudo, pode-se

notar que os extratos foliares da maioria das espécies apresentaram uma atividade muito baixa, ou mesmo insignificante, quando comparada com o controle. É interessante também observar que as únicas duas espécies de *Lychnophora* (*L. salicifolia* e *L. ericoides*) utilizadas neste estudo apresentaram atividade próxima e acima de 50% de inibição, podendo ser considerado um gênero promissor para este alvo bioquímico (Tabela 7). Entretanto, de maneira geral, não parece que o extrato foliar de um gênero apresente mais atividade que outro neste ensaio. De maneira geral, tais resultados são semelhantes ao que se apresenta na literatura. Por exemplo, Deiab e colaboradores verificaram que 115 extratos alcoólicos de plantas, de um total de 905, a uma concentração de 500 µg/ml, foram capazes de inibir 50% da enzima LDH-5 (DEIAB et al., 2013). Do mesmo modo, em outro estudo relacionado com a inibição de alvos enzimáticos por extratos, Chagas-Paula e colaboradores verificaram que 13 de um total de 57 extratos hidroalcoólicos provenientes de espécies de Asteraceae foram capazes de inibir as enzimas inflamatórias ciclooxigenase-1 (COX-1) e lipoxigenase-5 (5-LOX) (CHAGAS-PAULA et al., 2015). Destas 13 espécies, nenhuma se encontra entre as 11 mais promissoras descritas na tabela 7. Pelo contrário, espécies como *A. robusta* e *A. trichophylla* (escolhidas para estudos de cultivo presentes neste trabalho, descritos no capítulo 3) apresentaram baixa capacidade de inibição frente a LDH-5.

Tabela 7– Espécies investigadas no ensaio de inibição da enzima LDH-5 e seus respectivos códigos. A coluna “Inibição” se refere ao quanto o extrato foi capaz de diminuir a atividade enzimática em comparação com o controle (enzima atuando sem inibidores). Com base nos valores de inibição foram criados dois modelos, o primeiro considerou como de “baixa atividade” (low activity) os extratos que diminuíram a atividade enzimática de 1% a 45%. O segundo modelo considerou como “sem atividade” (no activity) todos os extratos que não foram capazes de diminuir em 45% a atividade enzimática. Os dois modelos consideraram como “ativos” (active) os extratos que foram capazes de diminuir em mais de 45% a atividade enzimática. Destacam-se em negrito as espécies cultivadas neste projeto: *A. robusta* e *A. trichophylla*, ambas sem atividade de inibição. O modelo 2 apresentou melhores resultados em termos de R^2 e Q^2 para a busca de variáveis (possíveis metabólitos) discriminantes dos extratos ativos.

Espécie	Código Ensaio	Código mzXML	Inibição	Modelo 1	Modelo 2
<i>Matricaria chamomila</i>	Extrato 54	41	100%	active	active
<i>Schlechtendalia luzulaefolia</i>	Extrato 64	47	93%	active	active
<i>Pluchea quitoc</i>	Extrato 61	46	90%	active	active
<i>Baccharis dracunculifolia</i>	Extrato 28	74	82%	active	active
<i>Lychnophora salicifolia</i>	Extrato 53	39	71%	active	active
<i>Lychnophora ericoides</i>	Extrato 52	128	59%	active	active
<i>Proteopsis argentea</i>	Extrato 63	145	59%	active	active
<i>Gymnanthemum amygdalinum</i>	Extrato 50	230	58%	active	active
<i>Baccharis sessifolia</i>	Extrato 25	156	49%	active	active
<i>Aldama discolor</i>	Extrato 17	61	49%	active	active
<i>Chuiraga spinosa</i>	Extrato 13 A	26	48%	active	active
<i>Dasyphyllum spinensis</i> - caule	Extrato 34	193	31%	low activity	no activity
<i>Lactuca sativa</i>	Extrato 51	37	29%	low activity	no activity
<i>Mikania cordifolia</i>	Extrato 56	42	27%	low activity	no activity

<i>Vernonia platensis</i>	Extrato 68	150	26%	low activity	no activity
<i>Eremanthus veadeiroensis</i>	Extrato 47	115	26%	low activity	no activity
<i>Mikania fulva</i>	Extrato 55	240	24%	low activity	no activity
<i>Arnica montana</i>	Extrato 21	11	24%	low activity	no activity
<i>Baccharis augustifolia</i>	Extrato 26	64	23%	low activity	no activity
<i>Achillea millefolium</i>	Extrato 1 A	1	20%	low activity	no activity
<i>Baccharis trimera</i>	Extrato 24	91	20%	low activity	no activity
<i>Paralychnophora atkinsiae</i>	Extrato 60	44	19%	low activity	no activity
<i>Espeletia cayetana</i>	Extrato 49	117	18%	low activity	no activity
<i>Stevia rebaudiana</i>	Extrato 65	53	17%	low activity	no activity
<i>Vernonia polyanthes</i>	Extrato 69	59	15%	low activity	no activity
<i>Eremanthus uniflorus</i>	Extrato 46	114	15%	low activity	no activity
<i>Aldama gardneri</i>	Extrato 18	8	14%	low activity	no activity
<i>Espeletia arbelaezii</i>	Extrato 48	194	13%	low activity	no activity
<i>Artemisia annua</i>	Extrato 23	13	12%	low activity	no activity
<i>Baccharis crispa</i>	Extrato 27	72	12%	low activity	no activity
<i>Dasyphyllum brasiliensis</i>	Extrato 33	29	11%	low activity	no activity
<i>Prestelia eriopus</i>	Extrato 62	144	11%	low activity	no activity
<i>Aldama robusta</i>	Extrato 2 A	9	11%	low activity	no activity
<i>Aldama nudibasilaris</i>	Extrato 3 A	3AC	11%	low activity	no activity

<i>Tithonia diversifolia</i>	Extrato 4 A	148	11%	low activity	no activity
<i>Dasyphyllum tomentosum</i>	Extrato 36	106	11%	low activity	no activity
<i>Eremanthus goyazensis</i>	Extrato 44	34	11%	low activity	no activity
<i>Artemisia absinthium</i>	Extrato 22	12	11%	low activity	no activity
<i>Taraxacum officinale</i>	Extrato 67	55	10%	low activity	no activity
<i>Echinacea purpurea</i>	Extrato 41	30	9%	low activity	no activity
<i>Tanacetum parthenium</i>	Extrato 66	54	9%	low activity	no activity
<i>Dasyphyllum varians</i>	Extrato 35	107	8%	low activity	no activity
<i>Calendula officinalis</i>	Extrato 37	192	8%	low activity	no activity
<i>Eremanthus pabstii</i>	Extrato 45	113	8%	low activity	no activity
<i>Aldama filifolia</i>	Extrato 19	7	6%	low activity	no activity
<i>Mikania laevigata</i>	Extrato 5 A	136	1%	low activity	no activity
<i>Chresta pycnocephala</i>	Extrato 38	24	1%	low activity	no activity
<i>Aldama pilosa</i>	Extrato 9 A	9AC	0%	no activity	no activity
<i>Emilia sonchifolia</i>	Extrato 42	31	0%	no activity	no activity
<i>Aldamda grandiflora</i>	Extrato 8 A	8AC	0%	no activity	no activity
<i>Baccharis lychnophora</i>	Extrato 29	80	0%	no activity	no activity
<i>Bidens pilosa</i>	Extrato 31	17	0%	no activity	no activity
<i>Vernonanthura brasiliiana</i>	Extrato 72	245	-1%	no activity	no activity
<i>Solidago chilensis</i>	Extrato 70	244	-2%	no activity	no activity

<i>Smallanthus sonchifolius</i>	Extrato 6 A	48	-2%	no activity	no activity
<i>Chresta sphaerocephala</i>	Extrato 39	104	-5%	no activity	no activity
<i>Ageratum conyzoides</i>	Extrato 20	4	-5%	no activity	no activity
<i>Aldama trichophylla</i>	Extrato 7 A	7AC	-6%	no activity	no activity
<i>Solidago microglossa</i>	Extrato 11 A	50	-6%	no activity	no activity
<i>Vernonia condensata</i>	Extrato 71	58	-15%	no activity	no activity
<i>Minasia scapigera</i>	Extrato 59	141	-17%	no activity	no activity
<i>Helianthus annuus</i>	Extrato 10 A	10AC	-19%	no activity	no activity
<i>Mikania glomerata</i>	Extrato 58	134	-19%	no activity	no activity
<i>Calea cuneifolia</i>	Extrato 12 A	19	-23%	no activity	no activity
<i>Aldama bracteata</i>	Extrato 16	16AC	-25%	no activity	no activity
<i>Cichorium intybus</i>	Extrato 40	27	-26%	no activity	no activity
<i>Blanchetia heterotricha</i>	Extrato 32	18	-31%	no activity	no activity

Tendo em vista os resultados apresentados na tabela 7, foram criados dois modelos por OPLS-DA para se apontar as variáveis correlacionadas com inibição de LDH-5. No primeiro modelo os extratos foram divididos em três classes: ativas (*active*), baixa atividade (*low activity*) e sem atividade (*no activity*) (Tabela 7). Entretanto, tal modelo não revelou bons valores de Q^2 , além de possuir um elevado valor de RMSEcv (dados não mostrados). Por outro lado, o segundo modelo, o qual considera como ativas apenas as amostras que conseguiram reduzir em mais de 50% atividade enzimática, foi melhor, além de mais simples. Por meio deste modelo podemos visualizar o gráfico dos *scores* (amostras) correlacionadas ou não com a atividade de inibição da LDH-5 (Figura 17). Esta OPLS-DA foi validada não somente pelos valores de Q^2 (validação cruzada), mas também por validação externa (25% das amostras ativas e 25% das amostras negativas da tabela 7, escolhidas aleatoriamente). A validação externa (P^2) foi de 81,48% (Tabela 8).

Tabela 8 – Validação externa (27 extratos) ($P^2 = 81,48\%$) da OPLS-DA dos extratos foliares de espécies de Asteraceae avaliados contra a inibição da LDH-5. Destaca-se em negrito a quantidade de espécies classificadas corretamente. Pela validação, todos os três extratos considerados como “ativo” (*active* = capaz de reduzir a atividade enzimática em mais de 45%) foram classificados como “sem atividade” (*no active*), de acordo com a análise criada pela OPLS-DA e seus respectivos constituintes químicos.

			Predição	
	Extratos	Correto	no activity	active
no activity	24	91,67%	22	2
Active	3	0%	3	0
Total	27	81,48%	25	2
			Predição	
ID Código	Classe	Espécie	no activity	active
Age1	no activity	<i>Ageratum conyzoides</i>	0,997	0,003
Ald9	no activity	<i>Aldama filifolia</i>	0,938	0,062
Ald8	no activity	<i>Aldama gardneri</i>	1,030	-0,030
Ech1	no activity	<i>Echinacea purpurea</i>	0,968	0,032
Ere7	no activity	<i>Eremanthus goyazensis</i>	0,974	0,026
Sma1	no activity	<i>Smallanthus sonchifolius</i>	0,994	0,006
Ver2	no activity	<i>Vernonia polyanthes</i>	0,920	0,080
Ald7	Active	<i>Aldama discolor</i>	0,901	0,099
Bac10	no activity	<i>Baccharis angustifolia</i>	-0,044	1,044
Bac10.2	no activity	<i>Baccharis angustifolia</i>	0,905	0,095

Bac10.3	no activity	<i>Baccharis angustifolia</i>	0,690	0,310
Bac10.4	no activity	<i>Baccharis angustifolia</i>	0,699	0,301
Bac22	no activity	<i>Baccharis crispa</i>	0,611	0,389
Bac40	no activity	<i>Baccharis lychnophora</i>	0,998	0,002
Bac1	no activity	<i>Baccharis trimera</i>	1,001	-0,001
Bac1.2	no activity	<i>Baccharis trimera</i>	-0,053	1,053
Bac1.3	no activity	<i>Baccharis trimera</i>	1,001	-0,001
Bac1.4	no activity	<i>Baccharis trimera</i>	0,707	0,293
Esp8	no activity	<i>Espeletia cayetana</i>	0,823	0,177
Lyc2	Active	<i>Lychnophora ericoides</i>	0,952	0,048
Pro1	Active	<i>Proteopsis argentea</i>	1,001	-0,001
Tit1	no activity	<i>Tithonia diversifolia</i>	1,036	-0,036
Ver1	no activity	<i>Vernonia platensis</i>	1,079	-0,079
Mik11	no activity	<i>Mikania fulva</i>	1,005	-0,005
Ald11	no activity	<i>Aldama trichophylla</i>	0,988	0,012
Ald13	no activity	<i>Aldama. pilosa</i>	0,928	0,072
Helth	no activity	<i>Helianthus annus</i>	1,004	-0,004

Pela validação, todos os três extratos considerados como “ativo” (*active* = capazes de reduzir a atividade enzimática em mais de 45%) foram classificados como “sem atividade” (*no active*), de acordo com a análise criada pela OPLS-DA e seus respectivos constituintes químicos. A OPLS-DA conseguiu separar claramente as amostras consideradas ativas das consideradas não ativas ($R^2 = 0,99$) com altos valores de validação cruzada ($Q^2 = 0,77$) (Figura 17).

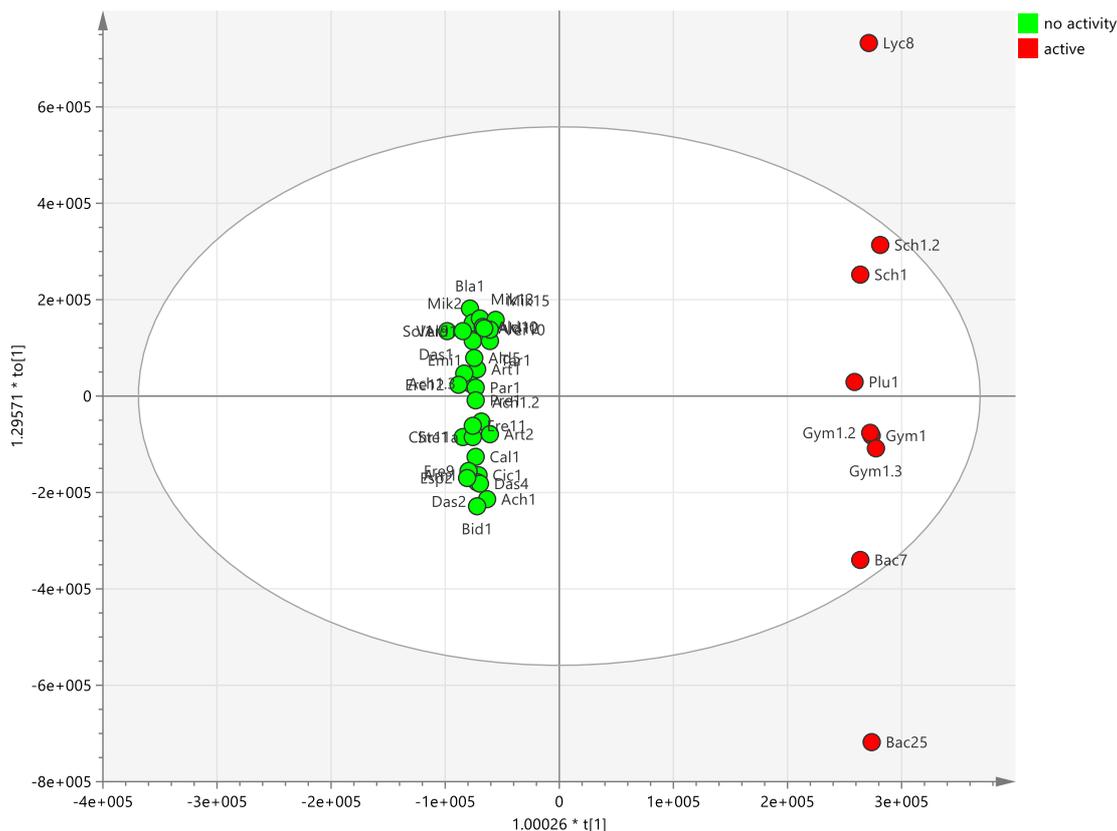


Figura 17 - OPLS-DA (1 + 7) ($R^2 = 0,99$ e $Q^2 = 0,77$) para extratos foliares de espécies de Asteraceae ensaiados contra LDH-5. Foram consideradas ativas (active) as amostras que inibiram mais de 45% da atividade enzimática, as outras amostras foram consideradas inativas (no activity). Pode ser observado pelo modelo que houve clara separação entre os grupos (alto valor de R^2) com base na inibição enzimática e no metaboloma das espécies, sem a presença de overfitting (alto valor de Q^2). Esta análise apresentou também um alto valor de validação externa ($P^2 = 0,81$, Tabela 8), baixo valor de RMSEcv (0,20) e Valor-p ($5,16 \times 10^{-5}$), o que permitiu gerar variáveis discriminantes (Tabela 9) com confiabilidade.

Esta OPLS-DA, e seus respectivos valores de validação, permitiu a realizar a predição da atividade dos outros extratos que não haviam sido ensaiados nas atividades de inibição de LDH-5 deste experimento, com base em seus metabolomas adquiridos por LC-MS. Assim, os resultados da tabela 9 mostram que foi possível prever a atividade da maioria dos extratos (Tabela 9). A tabela 9 contém também a classificação dos extratos que foram ensaiados, mas que tiveram de ser retirados do modelo para fins de otimização dos da OPLS-DA, o que resulta em mais amostras classificadas corretamente e melhor capacidade preditiva. Assim, dos cinco extratos classificados como ativos, de acordo com o resultado do ensaio, todos foram classificados como “sem atividade”, de acordo com suas constituições químicas (Tabela 9). Do mesmo modo, a maioria dos extratos não ensaiados foi classificada como “sem atividade”. Resumidamente, todas as espécies de *Aldama*, *Chresta* e *Eremanthus* foram classificadas como “sem atividade”.

Tabela 9 – Classificação dos extratos foliares de espécies de Asteraceae que foram retirados da análise (35 extratos) para a otimização do modelo de OPLS-DA para a inibição da atividade de LDH-5. E predição da atividade dos extratos foliares das espécies que não foram avaliadas contra LDH-5. Destaca-se em negrito as classificações consideradas “incorretas”, ou seja, aquelas que não se correlacionam com a informação previamente estabelecida e de acordo com o resultado do ensaio. Os números com asterisco indicam classificação/predição duvidosa.

			Classificação	
	Membros	Correto	no activity	active
no activity	30	90%	27	3
active	5	0%	5	0
No class	215		166	49
Total	250	77,14%	198	52
			Predição	
ID Código	Classe	Espécie	no activity	active
Acy1		<i>Achyrocline satureioides</i>	0,85	0,15
Acm1		<i>Acmella oleracea</i>	1,02	-0,02
Age1	no activity	<i>Ageratum conyzoides</i>	1,00	0,00
Ald3		<i>Aldama arenaria</i>	1,02	-0,02
Ald9	no activity	<i>Aldama filifolia</i>	0,94	0,06
Ald8	no activity	<i>Aldama gardneri</i>	1,03	-0,03
Ant1		<i>Anteremanthus hatschbachii</i>	0,89	0,11
Bac8		<i>Baccharis albilanosa</i>	1,11	-0,11
Bac8.2		<i>Baccharis albilanosa</i>	0,96	0,04
Bac8.3		<i>Baccharis albilanosa</i>	1,04	-0,04
Bac29		<i>Baccharis glaziovii</i>	0,91	0,09
Bac48		<i>Baccharis organensis</i>	0,85	0,15
Chr3c		<i>Chresta exsucca</i>	0,99	0,01
Chr5a		<i>Chresta hatschbachii</i>	0,95	0,05
Chr6c		<i>Chresta martii</i>	1,01	-0,01
Chr13a		<i>Chresta plantaginifolia</i>	1,01	-0,01
Chr13a.2		<i>Chresta plantaginifolia</i>	1,02	-0,02
Chr7b	no activity	<i>Chresta pycnocephala</i>	0,93	0,07
Cho1		<i>Chronopappus bifrons</i>	0,95	0,05
Chu2	active	<i>Chuiraga spinosa</i>	0,93	0,07
Cyn1		<i>Cynara scolymus</i>	1,04	-0,04

Ech1	no activity	<i>Echinacea purpurea</i>	0,97	0,03
Ere1		<i>Eremanthus argenteus</i>	1,14	-0,14
Ere4		<i>Eremanthus capitatus</i>	1,06	-0,06
Ere7	no activity	<i>Eremanthus goyazensis</i>	0,97	0,03
Ere10		<i>Eremanthus polycephalus</i>	0,98	0,02
Het2		<i>Heterocoma gracilis</i>	0,90	0,10
Lac1	no activity	<i>Lactuca sativa</i>	0,80	0,20
Lyc6		<i>Lychnophora mellobarretoii</i>	0,99	0,01
Lyc9		<i>Lychnophora santosii</i>	0,83	0,17
Mat1	active	<i>Matricaria chamomilla</i>	1,03	-0,03
Mik10		<i>Mikania sessilifolia</i>	0,94	0,06
Pip1		<i>Piptolepis monticola</i>	1,06	-0,06
Sma1	no activity	<i>Smallanthus sonchifolius</i>	0,99	0,01
Soa		<i>Soaresia velutina</i>	1,03	-0,03
Son1		<i>Sonchus oleraceus</i>	1,03	-0,03
Sph1		<i>Sphagneticola trilobata</i>	0,72	0,28
Tan1	no activity	<i>Tanacetum parthenium</i>	0,77	0,23
Tri1		<i>Tridax procumbens</i>	0,89	0,11
Ver8		<i>Vernonanthura beyrichii</i>	0,95	0,05
Ver2	no activity	<i>Vernonia polyanthes</i>	0,92	0,08
Ald7	active	<i>Aldama discolor</i>	0,90	0,10
Arc1		<i>Arctium lappa</i>	0,58*	0,42*
Bac9		<i>Baccharis alpina</i>	1,07	-0,07
Bac10	no activity	<i>Baccharis angustifolia</i>	-0,04	1,04
Bac10.2	no activity	<i>Baccharis angustifolia</i>	0,90	0,10
Bac10.3	no activity	<i>Baccharis angustifolia</i>	0,69	0,31
Bac10.4	no activity	<i>Baccharis angustifolia</i>	0,70	0,30
Bac11		<i>Baccharis aphylla</i>	0,39*	0,61*
Bac13		<i>Baccharis bicolor</i>	0,08	0,92
Bac15		<i>Baccharis boliviensis</i>	0,76	0,24
Bac16		<i>Baccharis buchtienii</i>	0,58*	0,42*
Bac17		<i>Baccharis caespitosa</i>	0,40*	0,60*

Bac19		<i>Baccharis cenicienta</i>	0,31	0,69
Bac19.2		<i>Baccharis cenicienta</i>	0,52*	0,48*
Bac21		<i>Baccharis ciliata</i>	0,41*	0,59*
Bac22	no activity	<i>Baccharis crispa</i>	0,61*	0,39*
Bac24		<i>Baccharis densiflora</i>	0,56*	0,44*
Bac24.2		<i>Baccharis densiflora</i>	0,77	0,23
Bac24.3		<i>Baccharis densiflora</i>	0,92	0,08
Bac27		<i>Baccharis floripa</i>	0,43*	0,57*
Bac31		<i>Baccharis grisebachii</i>	0,81	0,19
Bac32		<i>Baccharis halimifolia</i>	0,08	0,92
Bac33		<i>Baccharis incarum</i>	0,39*	0,61*
Bac35		<i>Baccharis jelskii</i>	0,10	0,90
Bac40	no activity	<i>Baccharis lychnophora</i>	1,00	0,00
Bac47		<i>Baccharis nitida</i>	0,91	0,09
Bac49		<i>Baccharis papillosa</i>	0,83	0,17
Bac51		<i>Baccharis papillosa longipedicellata</i>	0,26	0,74
Bac54		<i>Baccharis pilularis</i>	0,47*	0,53*
Bac55		<i>Baccharis plummerae</i>	0,98	0,02
Bac57		<i>Baccharis quitensis</i>	0,49*	0,51*
Bac60		<i>Baccharis rodriguezii</i>	0,82	0,18
Bac65		<i>Baccharis sergiloides</i>	0,45*	0,55*
Bac67		<i>Baccharis stylosa</i>	0,71	0,29
Bac70		<i>Baccharis tola</i>	0,99	0,01
Bac1	no activity	<i>Baccharis trimera</i>	1,00	0,00
Bac1.2	no activity	<i>Baccharis trimera</i>	-0,05	1,05
Bac1.3	no activity	<i>Baccharis trimera</i>	1,00	0,00
Bac1.4	no activity	<i>Baccharis trimera</i>	0,71	0,29
Bac72		<i>Baccharis trinervis</i>	0,31	0,69
Bac74		<i>Baccharis umbellata</i>	0,28	0,72
Bac74.2		<i>Baccharis umbellata</i>	0,21	0,79
Bac75		<i>Baccharis urvilleana</i>	0,46*	0,54*
Bac77		<i>Baccharis wrightii</i>	0,50*	0,50*

Chr1a		<i>Chresta angustifolia</i>	0,90	0,10
Chr2a		<i>Chresta curumbensis</i>	0,99	0,01
Chr2a.2		<i>Chresta curumbensis</i>	1,02	-0,02
Chr2a.3		<i>Chresta curumbensis</i>	0,92	0,08
Chr4a		<i>Chresta harleyi</i>	1,01	-0,01
Chr14a		<i>Chresta phlycifolia</i>	0,57*	0,43*
Chr12a		<i>Chresta scapigera</i>	1,01	-0,01
Chr8a		<i>Chresta scapigera</i>	0,92	0,08
Chr9a		<i>Chresta souzae</i>	0,93	0,07
Chr10a		<i>Chresta speciosa</i>	1,01	-0,01
Das5		<i>Dasyphyllum tomentosum</i>	1,11	-0,11
Ere2		<i>Eremanthus auriculatus</i>	1,02	-0,02
Ere3		<i>Eremanthus brevifolius</i>	1,03	-0,03
Ere5		<i>Eremanthus elaeagnus</i>	0,97	0,03
Ere6		<i>Eremanthus erythropappus</i>	0,75	0,25
Ere8		<i>Eremanthus incanus</i>	1,12	-0,12
Esp5		<i>Espeletia barclayana</i>	0,58*	0,42*
Esp8	no activity	<i>Espeletia cayetana</i>	0,82	0,18
Esp12		<i>Espeletia curialensis</i>	0,67	0,33
Esp15		<i>Espeletia estanislan</i>	1,02	-0,02
Esp16		<i>Espeletia frontinoensis</i>	0,76	0,24
Esp17		<i>Espeletia garciae</i>	0,70	0,30
Esp26		<i>Espeletia manguana</i>	0,85	0,15
Esp28		<i>Espeletia murilloi</i>	0,70	0,30
Esp40		<i>Espeletia schultzii</i>	1,03	-0,03
Esp40.2		<i>Espeletia schultzii</i>	0,64*	0,36*
Esp40.3		<i>Espeletia schultzii</i>	1,03	-0,03
Esp40.4		<i>Espeletia schultzii</i>	0,50*	0,50*
Esp40.5		<i>Espeletia schultzii</i>	1,02	-0,02
Esp40.6		<i>Espeletia schultzii</i>	1,02	-0,02
Esp41		<i>Espeletia semiglobulata</i>	0,68	0,32
Gor1		<i>Gorceixia decurrens</i>	1,04	-0,04

Het1		<i>Heterocoma ekmaniana</i>	0,61*	0,39*
Lyc2	active	<i>Lychnophora ericoides</i>	0,95	0,05
Lyc5		<i>Lychnophora markgravii</i>	0,91	0,09
Lyc7		<i>Lychnophora passerina</i>	1,03	-0,03
Lyc10		<i>Lychnophora tomentosa</i>	1,04	-0,04
Mik8		<i>Mikania aff. burchelli</i>	0,90	0,10
Mik12		<i>Mikania angularis</i>	0,87	0,13
Mik14		<i>Mikania hirsutissima</i>	0,88	0,12
Mik6		<i>Mikania psilostachya</i>	0,83	0,17
Mik3		<i>Mikania</i>	1,05	-0,05
Mik4		<i>Mikania thapsoides</i>	1,04	-0,04
Mik1		<i>Mikania triangularis</i>	1,00	0,00
Min1	no activity	<i>Minasia scapigera</i>	1,00	0,00
Par2		<i>Paralychnophora bicolor</i>	1,05	-0,05
Par.2		<i>Paralychnophora bicolor</i>	0,80	0,20
Par.3		<i>Paralychnophora bicolor</i>	1,07	-0,07
Pip2		<i>Piptolepis schultzi</i>	0,69	0,31
Pro1	active	<i>Proteopsis argentea</i>	1,00	0,00
Pro2		<i>Proteopsis hermogenesii</i>	0,57*	0,43*
Ver4		<i>Solidago chilensis</i>	1,02	-0,02
Tit1	no activity	<i>Tithonia diversifolia</i>	1,04	-0,04
Ver12		<i>Vernonanthura brasiliensis</i>	0,92	0,08
Ver1	no activity	<i>Vernonia platensis</i>	1,08	-0,08
Bac2		<i>Baccharis aff altimontana</i>	0,43*	0,57*
Bac3		<i>Baccharis aff boliviensis</i>	1,10	-0,10
Bac4		<i>Baccharis aff curitybensis</i>	0,32	0,68
Bac5		<i>Baccharis aff leucopappa</i>	0,55*	0,45*
Bac6		<i>Baccharis aff malmei</i>	0,99	0,01
Bac12		<i>Baccharis arguta</i>	0,36*	0,64*
Bac14		<i>Baccharis bifronsa</i>	0,43*	0,57*
Bac18		<i>Baccharis calliprinos</i>	0,50*	0,50*
Bac18.2		<i>Baccharis calliprinos</i>	0,45*	0,55*

Bac18.3		<i>Baccharis calliprinos</i>	0,53*	0,47*
Bac20		<i>Baccharis chionolaenoides</i>	0,13	0,87
Bac23		<i>Baccharis debilis</i>	0,53*	0,47*
Bac26		<i>Baccharis dubia</i>	0,20	0,80
Bac28		<i>Baccharis genistelloides</i>	0,56*	0,44*
Bac30		<i>Baccharis glomerulifolia</i>	0,79	0,21
Bac34		<i>Baccharis inexpectata</i>	0,66	0,34
Bac36		<i>Baccharis kurtziana</i>	0,96	0,04
Bac37		<i>Baccharis latifolia</i>	0,48*	0,52*
Bac38		<i>Baccharis linearifolia</i>	0,22	0,78
Bac39		<i>Baccharis longipetiolata</i>	0,90	0,10
Bac41		<i>Baccharis macrophylla</i>	0,22	0,78
Bac42		<i>Baccharis marchesii</i>	1,02	-0,02
Bac43		<i>Baccharis molina</i>	0,48*	0,52*
Bac44		<i>Baccharis multifolia</i>	0,35	0,65
Bac45		<i>Baccharis napaea</i>	-0,02	1,02
Bac46		<i>Baccharis neglecta</i>	0,92	0,08
Bac50		<i>Baccharis papillosa chaparensisv</i>	0,30	0,70
Bac52		<i>Baccharis pentlandii</i>	0,98	0,02
Bac62		<i>Baccharis papillosa ssp integrifolia</i>	0,44*	0,56*
Bac62.2		<i>Baccharis papillosa ssp integrifolia</i>	0,43*	0,57*
Bac62.3		<i>Baccharis papillosa ssp integrifolia</i>	0,51*	0,49*
Bac56		<i>Baccharis polycephala</i>	0,53*	0,47*
Bac58		<i>Baccharis reticularia</i>	0,64*	0,36*
Bac59		<i>Baccharis retusa</i>	0,26	0,74
Bac61		<i>Baccharis salicifolia</i>	0,93	0,07
Bac53		<i>Baccharis salicina</i>	0,32	0,68
Bac63		<i>Baccharis saliens</i>	0,62*	0,38*
Bac64		<i>Baccharis sarothroides</i>	0,43*	0,57*
Bac66		<i>Baccharis singularis</i>	0,93	0,07
Bac68		<i>Baccharis subalata</i>	0,39*	0,61*

Bac69		<i>Baccharis tarchonanthoides</i>	0,61*	0,39*
Bac71		<i>Baccharis tola frimbriata</i>	0,62*	0,38*
Bac73		<i>Baccharis trineura</i>	0,61*	0,39*
Bac76		<i>Baccharis vargasii</i>	0,56*	0,44*
Cae1	no activity	<i>Calendula officinalis</i>	0,76	0,24
Esp1		<i>Espeletia annemariana</i>	0,79	0,21
Esp3		<i>Espeletia ariana</i>	0,88	0,12
Esp4		<i>Espeletia aristeguietana</i>	0,49*	0,51*
Esp6		<i>Espeletia brachyaxiantha</i>	0,82	0,18
Esp7		<i>Espeletia brassicoidea</i>	0,83	0,17
Esp9		<i>Espeletia cleefii</i>	0,59*	0,41*
Esp9.2		<i>Espeletia cleefii</i>	0,89	0,11
Esp9.3		<i>Espeletia cleefii</i>	0,80	0,20
Esp10		<i>Espeletia congestiflora</i>	0,96	0,04
Esp10.2		<i>Espeletia congestiflora</i>	0,68	0,32
Esp10.3		<i>Espeletia congestiflora</i>	0,82	0,18
Esp11		<i>Espeletia congestiflora</i>	0,67	0,33
Esp13		<i>Espeletia delicatissima</i>	0,75	0,25
Esp14		<i>Espeletia episcopalis</i>	0,71	0,29
Esp18		<i>Espeletia grandiflora</i>	0,82	0,18
Esp19		<i>Espeletia incana</i>	0,87	0,13
Esp20		<i>Espeletia jajoensis</i>	0,73	0,27
Esp21		<i>Espeletia jaramilloi</i>	0,97	0,03
Esp22		<i>Espeletia jaramilloi</i>	0,58*	0,42*
Esp23		<i>Espeletia killipii</i>	0,67	0,33
Esp24		<i>Espeletia litocrassa</i>	0,85	0,15
Esp25		<i>Espeletia lopezii</i>	1,08	-0,08
Esp27		<i>Espeletia murilloi</i>	0,65	0,35
Esp29		<i>Espeletia nana</i>	0,65*	0,35*
Esp30		<i>Espeletia nemekenei</i>	0,61*	0,39*
Esp31		<i>Espeletia paipana</i>	0,83	0,17
Esp32		<i>Espeletia paipana</i>	0,87	0,13

Esp33		<i>Espeletia pescana</i>	0,96	0,04
Esp33.2		<i>Espeletia pescana</i>	0,94	0,06
Esp33.3		<i>Espeletia pescana</i>	0,85	0,15
Esp34		<i>Espeletia pisbana</i>	0,69	0,31
Esp35		<i>Espeletia pleiochasia</i>	0,45*	0,55*
Esp36		<i>Espeletia rositae</i>	0,91	0,09
Esp37		<i>Espeletia saboyensis</i>	0,82	0,18
Esp38		<i>Espeletia schultesiana</i>	1,02	-0,02
Esp39		<i>Espeletia schultzii</i>	0,46*	0,54*
Esp42		<i>Espeletia standleyana</i>	0,63*	0,37*
Esp43		<i>Espeletia steyermarkii</i>	0,67	0,33
Esp44		<i>Espeletia summapacis</i>	0,60*	0,40*
Esp45		<i>Espeletia tenorae</i>	0,94	0,06
Esp46		<i>Espeletia uribei</i>	0,60*	0,40*
Lyc1		<i>Lychnophora crispa</i>	1,00	0,00
Lyc4		<i>Lychnophora gardneri</i>	0,42*	0,58*
Lyc12		<i>Lychnophora triflora</i>	0,89	0,11
Lyc13		<i>Lychnophora grannogolensis</i>	0,55*	0,45*
Lyc14		<i>Lychnophorella leucodrendon</i>	0,80	0,20
Mik5		<i>Mikania firmula</i>	0,27	0,73
Mik5.2		<i>Mikania firmula</i>	0,45*	0,55*
Mik5.3		<i>Mikania firmula</i>	0,54*	0,46*
Mik5.4		<i>Mikania firmula</i>	0,52*	0,48*
Mik7		<i>Mikania guaco</i>	0,52*	0,48*
Mik9		<i>Mikania cuzcoensis</i>	0,40*	0,60*
Mik11	no activity	<i>Mikania fulva</i>	1,01	-0,01
Ver3		<i>Vernonia</i> 168 (?)	0,50*	0,50*
Ver5		<i>Vernonia</i> 170	1,00	0,00
Ver6		<i>Vernonia</i> 171	0,65	0,35
Ver7	no activity	<i>Vernonia</i> 172 (?)	0,43*	0,57*
Ver11		<i>Vernonanthura brasiliiana</i>	0,57*	0,43*
Ver11.2		<i>Vernonanthura brasiliiana</i>	0,48*	0,52*

Ver11.3		<i>Vernonanthura brasiliiana</i>	0,47*	0,53*
Ver13		<i>Vernonanthura brasiliiana-2</i>	0,91	0,09
Ver14		<i>Vernonia</i> 166 - Não identificada	0,52*	0,48*
Ald11	no activity	<i>Aldama trichophylla</i>	0,99	0,01
Ald13	no activity	<i>Aldama pilosa</i>	0,93	0,07
Helth	no activity	<i>Helianthus annus</i>	1,00	0,00

Com base no modelo de OPLS-DA construído foi possível também selecionar variáveis correlacionadas não somente com as amostras consideradas “ativas”, mas também com aquelas consideradas “sem atividade” (Tabela 10).

Tabela 10 – Identificação das variáveis discriminantes para os extratos foliares das espécies de Asteraceae classificados como “sem atividade” (no activity) e “ativo” (active), segundo modelo OPLS-DA. O sobrescrito “std” se refere a “standard” (substâncias do banco de padrões do AsterBioChem, previamente isolados ou comprados), os quais foram injetados no mesmo método cromatográfico a fim de se obter valores de tempo de retenção e perfil de fragmentação. Deste modo, as desreplicações que apontam para substâncias pertencentes ao banco de padrões são mais confiáveis, uma vez que não somente houve correlação com os valores de *m/z*, mas também com o tempo de retenção (RT). Var ID = número de identificação e respectivo modo de ionização em que se detectou o pico (pos = positivo; neg = negativo). As estruturas podem ser visualizadas em <http://zinc15.docking.org/catalogs/asternp/substances/>, em “Lookup”, ou em <http://143.107.203.160:8080/>, em “Compound Search”, inserindo-se seu nome usual ou código InChIKey. N.i. = pico não identificado. As variáveis marcadas com asterisco (97 pos, 338 neg, 249 neg e 30 neg) também foram apontadas pela Decision Tree.

Variáveis discriminantes para as amostras "no activity" fornecidas pelo OPLS-DA				
ID	RT	m/z	Identificação (nome usual ou código InChIKey)	Classes
73 pos	11,96	255,065	n.i.	
97 pos*	27,87	279,159	YVXVQMWUXKMTIT-VBCXXGINSAN; RCKXDYPRQOIBKT-CUOATXAZSAN	LST- <i>seco</i> -eudesmanolido
118 pos	25,90	171,993	n.i.	
Variáveis discriminantes para as amostras "active" fornecidas pelo OPLS-DA				
ID	RT	m/z	Identificação (nome usual ou código InChIKey)	Classes
134 pos	28,99	195,088	n.i.	
338 neg*	7,49	443,062	n.i.	

350 neg	2,88	283,046	n.i.	
249 neg*	8,02	471,056	n.i.	
30 neg*	2,07	317,055	n.i.	
136 neg	7,44	165,054	p-coumaric acid ^{std} ; methyl anisate	Derivados do ácido cinâmico e quínico
9 pos	5,63	195,088	n.i.	
196 pos	1,34	344,133	n.i.	
948 neg	7,64	112,987	n.i.	
15 pos	11,56	110,020	n.i.	
1261 neg	5,41	325,057	n.i.	
166 pos	12,11	361,092	acerosin; sudachitin	Flavonóides
214 neg	5,31	349,060	n.i.	
108 pos	27,69	301,216	(-)-dehydroabietic acid; NHWJVANFEVHGEL-ZNLXNUBKSA-N; ZPJLCCRNYWMMRT-GFBDQPOWSA-N	Diterpenos

532 neg	4,25	651,192	n.i.	
1768 neg	7,64	179,034	5-acetyl-2-hydroxybenzoic acid; 3,4-dihydroxycinnamic acid; caffeic acid	Derivados do ácido cinâmico e quínico
673 neg	4,28	388,088	n.i.	
293 pos	10,31	147,044	coumarin ^{std}	Derivados do ácido cinâmico
285 neg	9,21	507,114	n.i.	
17 pos	9,82	110,020	n.i.	
56 pos	12,29	301,070	chrysoeriol ^{std} ; hipidulin ^{std} ; diosmetin ^{std} ; rhamnazin ^{std} ; geraldol; sorbifolin; luteolin 7-methyl ether	Flavonóides
879 neg	5,44	390,104	n.i.	
86 neg	3,18	365,134	n.i.	
1226 pos	4,27	349,088	n.i.	
236 neg	26,34	459,275	n.i.	

235 pos	20,55	277,216	n.i.	
338 pos	18,29	233,153	lychnophoic acid ^{std} ; costunolide; isoalantolactone; IFASGTOWHLMHEZ-QVHKTLOISA-N	Sequiterpeno e LST eudesmanolido, <i>seco</i> -eudesmanolido e germacrolido

Como variáveis correlacionadas com as amostras sem atividade foram apontados três picos, todos detectados no modo de ionização positivo. Um deles pode ser uma LST do tipo *seco-eudesmanolido*, de acordo com o banco AsterDB. Este mesmo pico também foi identificado pela *Decision Tree* (Figura 18) como sendo uma variável latente correlacionada com amostras sem atividade. Variáveis correlacionadas com amostras sem atividade poderiam significar substâncias que atuam no sítio ativo da enzima, mas não inibem a conversão de piruvato em lactato, com conseqüente oxidação de NADH em NAD⁺. Quanto aos picos correlacionados com amostras ativas apontadas pelo OPLS-DA, alguns tiveram seus valores de [M+H]⁺ e [M-H]⁻ correlacionados não somente com o banco AsterDB, mas também com o banco de dados proveniente da biblioteca de substâncias padrão. A maioria dos picos identificados pertence à classe dos flavonoides e derivados do ácido *trans*-cinâmico e quínico. O ácido *p*-cumárico, cumarina, hispidulina (e estereoisômeros) e o sesquiterpeno ácido liquinofólico, são exemplos de estruturas que tiveram não apenas correlação da massa, mas também do tempo de retenção, de acordo com dados de padrões externos à análise.

É importante analisar os dados com cautela, ao considerar que tais substâncias, apontadas como correlacionadas com extratos ativos, certamente possuem atividade de inibição de LDH-5. Por exemplo, Baell, após realizar diversos ensaios de inibição com metabólitos secundários, revelou que muitas substâncias químicas eram frequentemente apontadas pelos ensaios como sendo ativas, embora apresentem, de fato, respostas falso positivas (BAELL, 2016). Foi verificado que, mesmo em diversas condições de ensaio, alguns metabólitos apresentavam atividade positiva indiscriminadamente (BAELL, 2016). Uma característica estrutural destes compostos é a presença de grupo ou catecol ou pirogalol, muito presentes em flavonoides e derivados do ácido *trans*-cinâmico, por exemplo. Assim, a substância hispidulina (e seus enantiômeros) deveria ser analisada com maior cautela. De fato, mais análises seriam relevantes a fim de se validar a atividade de tais substâncias, como o ensaio delas isoladas, o que será realizado em breve.

Além de OPLS-DA, também foram realizadas análises via *decision tree*, tendo em vista que se trata de outro algoritmo e, portanto, poderia apontar outras substâncias, ou corroborar aquelas já apontadas pelo OPLS-DA.

Destemo modo, alguns modelos foram construídos por *decision tree*, o que pode ser sumarizado na tabela 11.

Tabela 11 – Valores de classificação, validação e erro para os modelos 1 (todas as amostras) e 2 (75% das amostras – grupo treino) analisados por *Decision Tree* para extratos foliares de espécies de Asteraceae avaliados contra LDH-5. Cross. Valid. = Validação cruzada do conjunto de dados. Cross. Valid. Filtro = Validação cruzada aplicada ao conjunto de dados após a aplicação de filtro sobre as variáveis. Valid. Ext. = Validação externa. Filtro = “Attribute Selection” do Weka (utiliza um algoritmo que seleciona os atributos/variáveis mais correlacionados com o modelo).

Modelo 1 (Todas amostras)	Todas amostras	Cross. Valid.	Cross. Valid. Filtro	
Amostras classificadas corretamente:	71 (92,20%)	57 (74,03%)	65 (84,41%)	
Erro absoluto médio:	0,13	0,27	0,19	
RMSE:	0,25	0,47	0,39	
Modelo 2 (75% Treino)	Valid. Ext.	Cross. Valid.	Valid. Ext. Filtro	Cross. Valid. Filtro
Amostras classificadas corretamente:	15 (55,56%)	36 (72,00%)	16 (59,26%)	42 (84,00%)
Erro absoluto médio:	0,45	0,30	0,43	0,19
RMSE:	0,64	0,51	0,62	0,39

Na tabela 11 constam os resultados de dois modelos construídos na *decision tree*: “Modelo 1”, com todas as amostras, e “Modelo 2”, com a mesma partição de dados realizada para validar o modelo de OPLS-DA. Ambos os modelos foram utilizados para se verificar a quantidade de amostras classificadas corretamente e para apontar variáveis correlacionadas com extratos dito ativos, ou sem atividade. Por meio da validação externa ($P^2 = 55,6\%$) e cruzada ($Q^2 = 72,0\%$) realizada pela *Decision Tree*, pode-se notar que a quantidade de amostras corretamente classificadas/preditas é menor que aquela visualizada pelo OPLS-DA (Tabela 11). De fato, a OPLS-DA foi constantemente otimizada e variáveis não correlacionadas com o modelo ($VIP < 0,4$) foram deletadas, assim como amostras com altos erros de predição excluídas. Assim sendo, na *Decision Tree* também foi realizado um experimento onde se aplicou o algoritmo “*attribute selection*” do Weka, o qual trata-se de um filtro supervisionado que seleciona apenas variáveis correlacionadas com o modelo. Após este filtro, cada modelo gerou um novo conjunto de variáveis e os novos valores de P^2 e Q^2 foram melhores: 59,3% e 84,0%, para o modelo 2, respectivamente (Tabela 11). Ainda assim, é importante mencionar que as variáveis latentes apontadas após a utilização deste filtro não são substancialmente diferentes daquelas que foram apontadas antes, mesmo porque o filtro basicamente refaz o modelo quase que exclusivamente com apenas variáveis latentes, por isso os valores de predição

melhoram. Podemos visualizar estas variáveis correlacionadas com amostras classificadas como “ativas” e “sem atividade” na figura da *decision tree* antes do filtro (Figura 18), e depois do filtro (Figura 19).

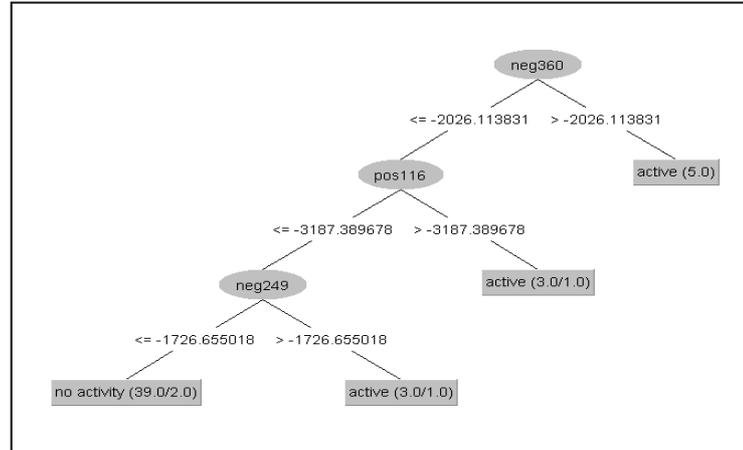


Figura 18 - *Decision tree* realizada no Weka com base no metaboloma de 75% das espécies (grupo treino) de Asteraceae ensaiadas contra a enzima LDH-5. Para esta análise, considerou-se como ativo (active) extratos que diminuíram mais de 45% da atividade enzimática. Do contrario, os extratos foram considerados inativos (no activity). Nesta análise, considera-se metaboloma a área dos picos do TIC dos extratos transformados por Pareto. Neg e pos referem-se ao modo de ionização que foi detectado o pico (neg = negativo; pos = positivo). As respectivas numerações que precedem o modo de ionização se referem ao número de identificação do pico. Assim, se o pico nº 360 detectado no modo negativo tiver área maior que -2026,1138, o extrato será ativo. Se sua área for menor que tal valor, o extrato será inativo, salvo se o extrato tiver o pico nº 116 detectado no modo positivo. A *decision tree* informa graficamente as variáveis (picos) mais importantes para a classificação dos extratos.

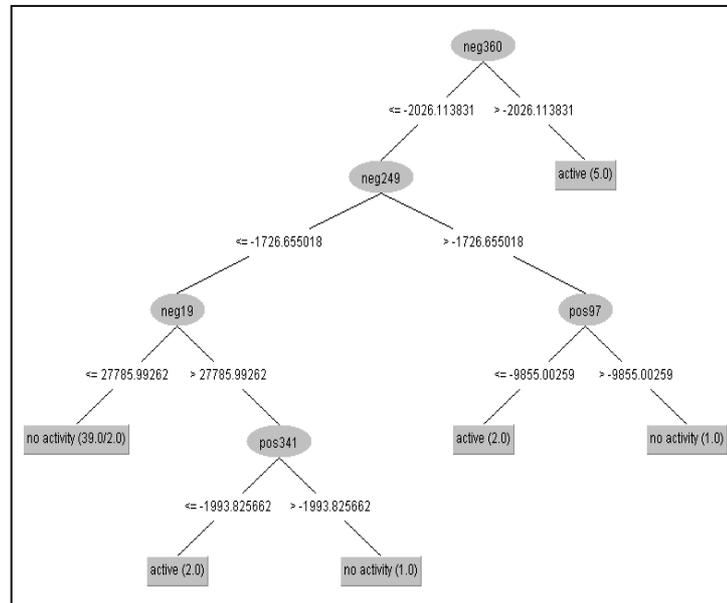


Figura 19 - *Decision tree* com base no metaboloma de 75% das espécies (grupo treino) de Asteraceae ensaiadas contra a enzima LDH-5 após aplicar o filtro supervisionado “attribute selection” do Weka às variáveis. Esse filtro utiliza um algoritmo que seleciona os atributos (variáveis/picos) mais correlacionados com o modelo. O filtro selecionou os picos pos97, pos341, pos361, pos491, pos511, pos884, neg19, neg30, neg249, neg360 para a construção do novo modelo. Para esta análise, considerou-se como ativo (active) extratos que diminuíram mais de 45% da atividade enzimática. Do contrario, os extratos foram considerados inativos (no activity). Nesta análise, considera-se metaboloma a área dos picos do TIC dos extratos transformados por Pareto. Neg e pos referem-se ao modo de ionização que foi detectado o pico (neg = negativo; pos = positivo). As respectivas numerações que precedem o modo de ionização se referem ao número de identificação do pico. Assim, se o pico nº 360 detectado no modo negativo tiver área maior que -2026,1138, o extrato será ativo. Se sua área for menor que tal valor, o extrato será inativo, salvo se o extrato tiver o pico nº 249 detectado no modo negativo e, ao mesmo tempo, não tiver o pico nº 97 detectado no modo positivo. A *decision tree* informa graficamente as variáveis (picos) mais importantes para a classificação dos extratos.

Assim, pela *decision tree* (Figura 18 e 19), pode-se compreender que as variáveis de ID 360 neg, 249 neg, 116 pos, 19 neg classificam o extrato como “ativo”. Isso significa que, se tais variáveis forem substâncias e estiverem em um extrato, ele provavelmente será ativo contra LDH-5. Do mesmo modo, tais substâncias teriam atividade quando isoladas. Já as substâncias de ID 97 pos e 341 pos, classificam o extrato como “sem atividade” e, portanto, atuam propiciando a não conversão de piruvato em lactato, com consequente não oxidação de NADH em NAD⁺. A literatura traz poucas aplicações práticas de *decision tree* na classificação de amostras de plantas em relação à atividade biológica. Por exemplo, Chagas-Paula e colaboradores ao utilizarem *decision tree* em seus experimentos com dados de metabolômica baseada em LC-MS e atividade de inibição de COX-1 e 5-LOX, propuseram que determinados extratos de espécies de Asteraceae teriam atividade dual contra tais enzimas, caso contivessem determinadas substâncias (CHAGAS-PAULA et al., 2015). Igualmente, Santos e colaboradores, ao partirem de dados de metabolômica baseada em GC-MS de plantas e ensaios contra *Streptococcus mutans*, conseguiram apontar substâncias altamente correlacionadas com a inibição da bactéria (SANTOS et al., 2018). Nessa ocasião, tais resultados foram validados ao se confirmar por concentração inibitória mínima (MIC) que a substância ácido palmítico possuía atividade antibacteriana (MIC = 31.25 µg/ml).

A tabela 12 descreve a identificação dos picos apontados pelo modelo 1 e 2 de *decision tree* construídos. Pode-se notar que alguns picos também foram apontados pela OPLS-DA, tais como os de ID 97 pos, 226 pos, 30 neg, 249 neg e 338 neg. Destes, apenas o 226 pos teve sua massa correlacionada com LSTs do banco AsterDB. Outras substâncias foram aqui identificadas. Por exemplo, uma cumarina, cuja identificação é pouco provável estar correta, uma vez que o tempo de retenção não se correlaciona com nosso banco de dados, e flavonoides glicosilados, os quais deveriam ser tomados com cautela, tendo em vista a presença de grupo ou catecol ou pirogalol em suas estruturas, o que poderia sugerir falso positivo nas substâncias (BAELL et al., 2016). Ainda assim, o flavonol santina, cujo tempo de retenção e [M+H]⁺ se correlacionou com o banco de dados provenientes da biblioteca de substâncias padrões, pode ser uma boa substância para se investigar o real potencial de inibição da LDH-5, tendo em vista que hidroxila na posição 6 do anel A se encontra metilada e, portanto, não há grupamento pirogalol livre na estrutura da substância.

Tabela 12 - Identificação das variáveis discriminantes para os extratos foliares das espécies de Asteraceae classificados como “sem atividade” (no activity) e “ativo” (active), segundo modelo *Decision Tree*. Pela *Decision Tree* foram realizados dois modelos, no primeiro (Modelo 1) foram utilizadas todas as amostras avaliadas contras LDH-5 para se encontrar as variáveis discriminantes; no segundo modelo (Modelo 2), foi utilizado apenas o grupo treino. O sobrescrito “std” se refere a “standard” (substâncias do banco de padrões do AsterBioChem, previamente isolados ou comprados), os quais foram injetados no mesmo método cromatográfico a fim de se obter valores de tempo de retenção e perfil de fragmentação. Deste modo, as desreplicações que apontam para substâncias pertencentes ao banco de padrões são mais confiáveis, uma vez que não somente houve correlação com os valores de *m/z*, mas também com o tempo de retenção (RT). Var ID = número de identificação e respectivo modo de ionização em que se detectou o pico (pos = positivo; neg = negativo). As estruturas podem ser visualizadas em <http://zinc15.docking.org/catalogs/asternp/substances/>, em “Lookup”, ou em <http://143.107.203.160:8080/>, em “Compound Search”, inserindo-se seu nome usual ou código InChIKey. N.i. = pico não identificado. As variáveis marcadas com um asterisco (97 pos, 338 neg, 249 neg e 30 neg) também foram apontadas pela OPLS-DA. Já a variável marcada com dois asteriscos (226 pos) foi apontada pela *Decision Tree* como discriminante tanto das amostras sem atividade, se em altas concentrações, como das amostras com atividade, quando em baixas concentrações.

Variáveis discriminantes fornecidas pela Decision Tree para as amostras "no activity" (Modelo 1 = Todas amostras)

ID	RT	m/z	Identificação (nome usual ou código InChIKey)	Classes
226 pos**	11,22	345,169	FKDIIXZIKCNXAT-SLFMQDIPSA-N; DWYSOHJKGRHGTF-SJBKCFBNSA-N;	LST germacranolido, guaianolido, helenanolido e estereoisomeros
			JEAWGNYMYMHNGM-NYVXQQPPSA-N; IRVAQGHUUDIVGEA-UZYNNVTJSA-N;	
			FTPHYXGWMIZVMP-RZYABJHASA-N	

Variáveis discriminantes fornecidas pela Decision Tree para as amostras "active" (Modelo 1 = Todas amostras)

ID	RT	m/z	Identificação (nome usual ou código InChIKey)	Classes
226 pos**	11,22	345,169	FKDIIXZIKCNXAT-SLFMQDIPSA-N; DWYSOHJKGRHGTF-SJBKCFBNSA-N;	LST germacranolido, guaianolido, helenanolido e estereoisomeros
			JEAWGNYMYMHNGM-NYVXQQPPSA-N; IRVAQGHUUDIVGEA-UZYNNVTJSA-N;	

			FTPHYXGWMIZVMP-RZYABJHASA-N	
48 pos	1,33	162,112	n.i.	
249 neg*	8,02	471,056	n.i.	
30 neg*	2,07	317,055	n.i.	
54 neg	8,21	1031,24 4	n.i.	
709 pos	13,56	147,044	Coumarin	Derivado do ácido cinâmico
19 neg	5,82	593,151	kaempferol 3-O-rutinoside; vicianin-2	Flavonóides-glicosilados
338 neg*	7,49	443,06	n.i.	

Variáveis discriminantes fornecidas pela Decision Tree para as amostras "no activity" (Modelo 2 = 75% das amostras = Treino)

ID	RT	m/z	Identificação (nome usual ou código InChIKey)	Classes
97 pos*	27,87	279,159	YVXVQMWUXKMTIT-VBCXXGINSAN; RCKXDYPRQOIBKT-CUOATXAZSAN	LST- <i>seco</i> -eudesmanolido
341 pos	29,89	122,081	n.i.	

Variáveis discriminantes fornecidas pela Decision Tree para as amostras "active" (Modelo 2 = 75% das amostras = Treino)

ID	RT	m/z	Identificação (nome usual ou código InChIKey)	Classes
116 pos	16,98	345,097	santin ^{std} ; nevadensin; pedunculin; eupatorin	Flavonóides-glicosilados
360 neg	4,35	327,075	n.i.	
249 neg	8,02	471,056	n.i.	
19 neg	5,82	593,151	kaempferol 3-O-rutinoside; vicenin-2	Flavonóides-glicosilados

5 - CONCLUSÃO DO CAPÍTULO 2

A análise de inibição de LDH-5 revelou que 11, dos 80 extratos ensaiados, foram capazes de diminuir a atividade enzimática em mais de 45%, ou mesmo quase 100% (completa inibição). O modelo de OPLS-DA construído permitiu uma clara separação dos extratos ativos, dos não ativos, com base em suas constituições químicas. Os valores de validação desta OPLS-DA ($Q^2 = 0,77$; $P^2 = 0,81$) indicam um modelo robusto e com alta capacidade preditiva. Com base nisso, os outros extratos não ensaiados tiveram suas atividades preditas. Além disso, foram propostas algumas substâncias discriminantes dos extratos considerados ativos. As análises por *decision tree* igualmente tiveram bons valores de validação ($Q^2 = 0,72$; $P^2 = 0,56$, sem aplicação do algoritmo “*attribute selection*”), o que também permitiu indicar prováveis substâncias químicas discriminantes dos extratos ativos. Tais substâncias ativas serão isoladas futuramente e terão suas atividades confirmadas.

CAPÍTULO 3: ESTUDOS DE GERMINAÇÃO E CULTIVO DE *Aldama robusta* E *A. trichophylla* COMPARADOS COM *Tithonia diversifolia* (ASTERACEAE)

LISTA DE FIGURAS DO CAPÍTULO 3

Figura 20. Esquema mostrando as dimensões dos canteiros sem e com sombrite 35% construídos. Figura de minha autoria.....	109
Figura 21. Canteiro construído com sombrite 35%. À esquerda plantas adubadas, à direita, plantas sem adubo. Figura de minha autoria.....	117
Figura 22. Canteiro construído sem sombrite. À esquerda plantas adubadas, à direita, plantas sem adubo. Figura de minha autoria.....	118

LISTA DE TABELAS DO CAPÍTULO 3

Tabela 13. A atividade fotoblástica das sementes de <i>A. robusta</i> e <i>A. trichophylla</i> foi analisada comparando-se a germinação sob exposição a 12 horas de luz versus a ausência total de luz (escuro).....	112
Tabela 14. Resultados do monitoramento da germinação de <i>A. robusta</i> e <i>A. trichophylla</i> expostos ao GA ₃	113
Tabela 15. Resultados do monitoramento da germinação de sementes de <i>T. diversifolia</i> expostas ao GA ₃	115
Tabela 16. Resultados do monitoramento da germinação de sementes de <i>T. diversifolia</i> expostas a aquecimento a 80 °C por 30 s e 3 min.....	116
Tabela 17. Valores de médias para os parâmetros de desenvolvimento dos cultivares de <i>A. robusta</i>	119

LISTA DE ABREVIATURAS, SIGLAS E SIMBOLOS DO CAPÍTULO 3

%G	porcentagem de germinação
GA ₃	ácido giberélico
GSI	índice de velocidade de germinação
há	hectare
MGT	tempo médio de germinação
ppm	parte por milhão

RESUMO DO CAPÍTULO 3

O capítulo 1 apontou uma necessidade em se avaliar a influência que fatores ambientais teriam em alterar o metaboloma de espécies de Asteraceae, a ponto de causar equívocos nas interpretações quimiotaxonômicas. Para tanto, antes de se realizar tal análise (capítulo 4), com base em uma abordagem metabolômica, seria necessário cultivar algumas espécies e parametrizá-las. Espécies com significativo potencial farmacológico foram escolhidas para o cultivo: *A. robusta*, *A. trichophylla* e *T. diversifolia*. Entretanto, o gênero *Aldama* não se propaga vegetativamente (BOMBO et al., 2015) e não pode ser facilmente cultivado por estaquias. Assim, a germinação de sementes deste gênero é obrigatoriamente o passo inicial do desenvolvimento de novos espécimes. Por outro lado, as sementes de *T. diversifolia* apresentam dormência (AGBOOLA et al., 2005). Deste modo, foi necessário realizar um estudo prévio quanto à germinação e desenvolvimento morfológico de tais espécies. A germinação das sementes de *Aldama* foi analisada frente à exposição à luz e exposição ao ácido giberélico (GA₃). As sementes de *T. diversifolia* foram analisadas quanto à exposição ao GA₃ e aquecimento em água. Os parâmetros analisados foram: porcentagem de germinação (%G), índice de velocidade de germinação (GSI) e tempo médio de germinação (MGT). Após a germinação as plântulas seriam analisadas quanto aos seus desenvolvimentos morfológicos frente à adubação e luminosidade, aferindo-se área foliar, altura do ramo principal, quantidade de ramos secundários, quantidade de folhas do ramo principal e diâmetro do ramo principal. Foi verificado que apenas *A. trichophylla* possui atividade fotoblástica. Observou-se também que todas as espécies respondem ao aumento da concentração de GA₃, obtendo-se maiores valores de %G, GSI e MGT, embora pouco significativo em *T. diversifolia*. Estas, responderam melhor ao aquecimento em água (80 °C por 3 min). As plântulas de *T. diversifolia* morreram precocemente, e *A. trichophylla* apresentou um crescimento lento e frágil de seus ramos. Deste modo, o desenvolvimento morfológico foi realizado apenas para *A. robusta*, onde se constatou que a adubação de espécies sombreadas foi capaz de gerar plantas com maior número de folhas, maior diâmetro do ramo principal e maior área foliar. Por outro lado, a adubação em canteiro sem sombrite foi capaz de produzir apenas maior número de folhas do ramo principal. Os resultados aqui descritos contribuem para um maior número de germinação de sementes, bem como melhor desenvolvimento morfológico de *Aldama*, os quais poderão ser comparados com outras análises de outras espécies.

SUMMARY OF CHAPTER 3

Chapter 1 pointed to a need to evaluate the influence that environmental factors would have on altering the metabolome of Asteraceae species, causing mistakes in chemotaxonomic interpretations. Therefore, before carrying out such an analysis, based on a metabolomic approach (chapter 4), it would be necessary to cultivate some species to standardize them. Species with significant pharmacological potential were chosen for the cultivation: *A. robusta*, *A. trichophylla* and *T. diversifolia*. However, the genus *Aldama* does not propagate vegetatively (BOMBO et al., 2015) and can not be easily cultivated by plant cutting. Thus, the germination of seeds of this genus is obligatorily the initial step of the development of new specimens. On the other hand, the seeds of *T. diversifolia* present dormancy (AGBOOLA et al., 2005). Therefore, it was necessary to carry out a previous study on the germination and morphological development of such species. The germination of *Aldama* seeds was analyzed by exposure to light and exposure to gibberellic acid (GA₃). The seeds of *T. diversifolia* were analyzed for exposure to GA₃ and heating in water. The parameters analyzed were: percentage of germination (%G), germination speed index (GSI) and mean germination time (MGT). After germination the seedlings would be analyzed for their morphological development in relation to fertilization and luminosity, and the leaf area, height of the main branch, number of secondary branches, number of leaves of the main branch and diameter of the main branch. It was verified that only *A. trichophylla* has photoblastic activity. It was also observed that all species respond to the increase of the concentration of GA₃, obtaining higher values of %G, GSI and MGT, although little significant in *T. diversifolia*. These responded better to heating in water (80 °C for 3 min). Seedlings of *T. diversifolia* died early, and *A. trichophylla* showed a slow and fragile growth of their branches. Thus, the morphological development was performed only for *A. robusta*, where it was found that the fertilization of shaded species was able to generate plants with greater number of leaves, larger diameter of the main branch and greater leaf area. On the other hand, the fertilization without shade was able to produce only a greater number of leaves of the main branch. The results described here contribute to a greater number of seed germination as well as better morphological development of *Aldama*, which can be compared with other analyzes of other species.

1 – INTRODUÇÃO

O gênero *Aldama* possui 118 espécies, acomodando vários *taxa* infragênicos de *Viguiera* Kunth sensu Blake (Asteraceae) e três outros gêneros, *Alvordia* Brandege (Asteraceae), *Rhysolepis* Blake (Asteraceae) e *Stuessya* Turner & Davis (Asteraceae) (SCHILLING e PANERO, 2011). Existem espécies de *Aldama* do sudoeste da América do Norte e do México até a América do Sul, mas o grupo mais representativo de *Aldama* são as espécies de *Viguiera* da América do Sul. Por outro lado, o gênero *Tithonia* compreende 11 espécies e *T. diversifolia* é a espécie mais representativa e estudada do gênero (CHAGAS-PAULA et al., 2012). Esta espécie é encontrada em regiões tropicais e equatoriais do mundo, com exemplares nos continentes da América Central, América do Sul, África e Ásia.

O gênero *Viguiera*, atualmente *Aldama*, foi extensivamente estudado quanto ao seu potencial farmacológico. Várias substâncias foram isoladas e tiveram suas atividades biológicas documentadas. Entre as substâncias avaliadas, a LST budleína A foi a mais estudada, tendo sido relatada atividade antimicrobiana (PORTO et al., 2009), citotóxica (DELGADO et al., 1982), atividades anti-inflamatória e analgésica (VALERIO et al., 2007), tripanocida (NOGUEIRA et al., 2016) e inibição da interação célula leucocitária-endotelial e TNF- α , IL-8, MCP-1, nitritos e liberação CAM (NICOLETE et al., 2009) e inibição da atividade do NF- κ B (ZARPELON et al., 2017). Quanto à *T. diversifolia*, mais de 150 compostos foram isolados, principalmente as LST tagitininas A, C e F (CHAGAS-PAULA et al., 2012). A tagitinina C é um biomarcador da espécie e, possivelmente, tem atividades antiplasmodial (GOFFIN et al., 2002), antifeedant (AMBROSIO et al., 2008) e leishmanicida (TOLEDO et al., 2014). Outras seis LST isoladas de *T. diversifolia* apresentaram atividade leishmanicida (TOLEDO et al., 2014). No entanto, a atividade biológica de *T. diversifolia* não está relacionada apenas à presença de LST. O ácido 5-*O*-(*E*)-cafeoilquínico, derivado *trans*-cinâmico, apresentou melhor atividade anti-inflamatória oral e tópica do que a tagitinina C e até indometacina (CHAGAS-PAULA et al., 2011). A maioria das atividades biológicas publicadas para *T. diversifolia* está relacionada aos usos etnofarmacológicos de extratos, como propriedades antimalárica, anti-inflamatória e analgésica, anticancerígena, antidiabética e antioxidante (AJAO e MOTEETEE, 2017). É importante mencionar que *Aldama* e *Tithonia* possuem alguns metabólitos em comum, como as LST rotundina, leptocarpina, 15-hidroileptocarpina, diversifolina, orizabina, tirofundifolina A, os diterpenos ácidos *ent*-kaur-16-en-19-oico, grandiflorico e *ent*-17-hidroxi-15-en-19-oico, os flavonoides himenoxina, hispidulina, luteolina, e os esteróis estigmasterol e β -sitosterol, o que pode sugerir atividades biológicas similares por seus extratos brutos. Recentemente,

extratos hidroalcoólicos brutos de folhas de *A. robusta*, *A. trichophylla* e *T. diversifolia* apresentaram atividade inibitória das enzimas inflamatórias ciclooxigenase e lipoxigenase (CHAGAS-PAULA et al., 2015).

Tendo em vista o potencial farmacológico de *A. robusta*, *A. trichophylla* e *T. diversifolia*, foi planejado cultivá-los para posteriores estudos. No entanto, o gênero *Aldama* não se propaga vegetativamente (BOMBO et al., 2015) e não pode ser facilmente cultivado por estaquias. Assim, a germinação de sementes deste gênero é obrigatoriamente o passo inicial do desenvolvimento de novos espécimes. Estudos anteriores avaliaram as temperaturas ideais para a germinação de cinco espécies de *Aldama*, e para *A. robusta* e *A. trichophylla*, e foi estabelecida em 25 °C (BOMBO et al., 2015). Ainda assim, de acordo com tal estudo, houve apenas 54% de germinação para as espécies de *A. trichophylla* nas melhores condições experimentais. Quanto à *T. diversifolia*, sabe-se que, embora suas plantas apresentem rápida propagação, suas sementes apresentam dormência, gerando também baixos números de germinação (AGBOOLA et al., 2005). Uma maneira de contornar problemas com baixas quantidades de sementes germinadas é usar hormônios que podem aumentar seu rendimento. Um hormônio amplamente estudado na germinação de sementes é o ácido giberélico (GA₃). Possivelmente, o GA₃ aumenta a síntese de proteína de semente (MCDONALD e KHAN, 1983), estimula o crescimento do embrião e induzindo a produção de hidrolases para enfraquecer as estruturas ao redor do embrião (HOOLEY, 1994). Além disso, o GA₃ promove um aumento na síntese de α -amilase que hidrolisa as reservas de amido e aumenta os níveis de glicose necessários para o desenvolvimento de energia embrionária (KANEKO et al., 2002).

Considerando a necessidade de preservar plantas de grande importância ecológica, química e farmacológica, estudos são necessários para melhorar suas condições de cultivo. A germinação de sementes é um passo essencial no cultivo de muitas espécies. Tal estudo visa fornecer melhores estratégias para a germinação e preservação das espécies destes gêneros e outras da família Asteraceae. Além disso, após a germinação as plântulas podem ser cultivadas e estudadas mais a fundo, em termos da influência de fatores externos em sua anatomia e metaboloma.

No capítulo anterior foi discutido que a quimiotaxonomia pode auxiliar não somente na caracterização química de grupos de plantas, bem como em auxiliar na sistemática das espécies envolvidas no estudo. Entretanto, também foi discutido que os metabólitos secundários podem não constituir bons marcadores taxonômicos, uma vez que eles sofrem forte influência

do meio ambiente (PADILLA-GONZÁLEZ et al., 2017; ALVES et al., 2018, in: SÜLSEN e MARTINO, 2018). Deste modo, foi idealizado cultivar as sementes germinadas e estudá-las do ponto de vista metabólico frente a diferentes parâmetros ambientais (capítulo 4). Neste capítulo, por outro lado, dar-se-á enfoque às possíveis diferenças fisiológicas causadas por parâmetros como adubação e sombreamento. De fato, não há trabalho na literatura que demonstra a influência de tais parâmetros no crescimento e quantidade de folhas em *A. robusta*, e, pelo o que exposto no terceiro item da Introdução Geral, tais estudos podem contribuir com a conservação de espécie, especialmente aquelas do Cerrado e ameaçadas de extinção. Não obstante, um trabalho demonstrou a influência que diferentes fotoperíodos causam no crescimento e floração de *A. robusta* (RUGGIERO et al., 1997). Um outro trabalho com praticamente a mesma abordagem avaliou também o crescimento e a floração de *A. discolor* frente a diferentes fotoperíodos (ISEJIMA et al., 1991). Desde então, poucos trabalhos foram conduzidos quanto ao cultivo de espécies de *Aldama*, tal como será abordado neste capítulo.

2 – OBJETIVOS

Os objetivos deste capítulo são:

- 1) analisar a influência de fatores exógenos, como a luz, o GA₃ e o aquecimento da água, sobre a germinação de *A. robusta*, *A. trichophylla* e *T. diversifolia*;
- 2) analisar a influência da adubação e sombreamento em parâmetros anatômicos de *A. robusta*, tais como altura, quantidade de folhas e ramos, e área foliar.

3 – MATERIAL E MÉTODOS

3.1 – Coleta de sementes

Capitula de *A. robusta* (*V. robusta*) e *A. trichophylla* (*V. trichophylla*) foram coletadas em Altinópolis (SP) (21°01'49.9" S, 47°18'25.4" W) em junho de 2015, e em Palmeira (PR) (25°22'53.40" S; 49°48'21.00" W), em abril de 2015, respectivamente. *Capitula* de *T. diversifolia* foram coletadas em junho de 2015 em Ribeirão Preto (SP) (21°11'7.16" S; 47°51'43.48" W). Todas as coletas de *capitula* foram realizadas durante o período de frutificação. Os *capitula* foram armazenados em sacos de papel a (aproximadamente 18 °C em média) até o momento da utilização.

3.2 – Configurações experimentais

A germinação teve início em agosto de 2015. As sementes foram retiradas manualmente dos *capitula*, posteriormente lavadas com hipoclorito de sódio a 1% para fins de esterilização (BARBOSA et al., 1999) e, em seguida, enxaguadas com água da torneira. Em seguida, vinte sementes por placa foram distribuídas em duas lâminas de papel filtro, hidratadas com 13 mL de água destilada em caixas plásticas de gerbox previamente esterilizadas (BOMBO et al., 2015). Os experimentos foram conduzidos em relação ao fotoblastismo de sementes de *A. robusta* e *A. trichophylla*: 12 h de fotoperíodo diário *versus* ausência de luz.

Além disso, outros experimentos foram conduzidos com sementes de *A. robusta*, *A. trichophylla* e *T. diversifolia*, a fim de verificar a influência da concentração (5, 40, 320 ppm) e do tempo de exposição (1, 12, 24 h) ao ácido giberélico (GA₃ - 90% de pureza, Sigma Aldrich) (FARIMAN et al. 2011 e KIRMIZI et al., 2011). Para isso, as sementes foram submersas em GA₃ usando tubos Falcon. Sementes sem exposição ao GA₃ (experimento de atividade foblástica) foram consideradas como grupo controle. Para as sementes de *T. diversifolia*, também foi verificada a influência do tempo de exposição (30 s *versus* 3 min) ao banho em água a 80 ° C (AGBOOLA et al., 2006) para sementes sem exposição ao GA₃ (grupo controle) e para os grupos que proporcionaram níveis mais altos de germinação à exposição ao GA₃.

Os experimentos foram conduzidos em germinadores (Marconi MA 402) com temperatura constante de 25 °C fornecida por lâmpada fluorescente (BOMBO et al., 2015). Os experimentos foram realizados em triplicata, para exposição ao GA₃, e em quintuplicata, para

atividade fotoblástica (apenas *A. trichophylla* e *A. robusta*) e exposição ao calor (apenas *T. diversifolia*).

A germinação foi contada a partir do segundo dia após a incubação. Após o terceiro dia de incubação, a contagem foi realizada a cada dois dias. O último dia de contagem de sementes germinadas variou de acordo com a espécie e o experimento observado: 15 dias para *A. trichophylla*, 17 para *A. robusta*, 13 para *T. diversifolia* exposta a GA₃, 13 para *T. diversifolia* exposta a aquecimento por 30 s (sem exposição a GA₃), 15 dias para *T. diversifolia* exposta ao aquecimento por 3 min (sem exposição ao GA₃), e 17 min para *T. diversifolia* exposta ao aquecimento por 3 min (com exposição ao GA₃). Considerou-se como o último dia de contagem enquanto a germinação foi observada.

3.3 – Monitoramento da germinação

A monitoração da germinação foi observada de acordo com a porcentagem final de sementes germinadas (%G), índice de velocidade de germinação (GSI) (MAGUIRE, 1962) e tempo médio de germinação (MGT) (DANTAS et al., 2001), com base nos seguintes cálculos:

$$\%G = gf/ni$$

onde gf é o número de sementes germinadas até o último dia de observação e ni é a quantidade inicial de sementes por placa em cada experimento, e

$$GSI = \sum_{i=j}^k gi/ti \quad e \quad MGT = \sum_{i=j}^k giti / \sum_{i=j}^k gi$$

onde gi é o número de sementes germinadas no respectivo dia ti, j é a primeira observação (segundo dia de germinação) e k é o último dia de germinação.

Os valores brutos de GSI e MGT foram normalizados (N) de acordo com a quantidade inicial de sementes por placa de cada experimento, de acordo com o seguinte cálculo:

$$N = x/ni$$

onde x representa os valores experimentais brutos obtidos dos cálculos GSI e MGT e ni é a quantidade inicial de sementes por placa em cada experimento. Esta normalização é importante para comparações adicionais entre os resultados presentes neste estudo ou com os

resultados de outros estudos, uma vez que o número de sementes germinadas também é determinado pelo número de sementes iniciais.

Para a análise estatística, os dados normalizados (N) e %G foram transformados (T) de acordo com (DANTAS et al., 2001; BOMBO et al., 2015, ambos adaptados):

$$T = \arcsen\sqrt{A}$$

onde A expressa os valores obtidos de N (para a análise de GSI e MGT) e %G.

A análise de variância foi utilizada para os resultados, e as médias foram comparadas por meio do teste t e teste de Tukey ($p < 0,05$). A maioria das análises foi realizada com o *software* Microsoft Excel (2013), e o teste de Tukey no *software* PAST 3.15 (Universidade de Oslo, Noruega).

3.4 – Cultivo

Após a germinação das sementes de *A. robusta*, *A. trichophylla* e *T. diversifolia* suas plântulas, ainda na caixa gerbox, foram gentilmente enxaguadas e, posteriormente, cultivadas em bandejas de germinação de 5 x 5 x 8 cm contendo substrato de plantio e mantidas em estufa de plantas com sombrite 50% e irrigação automática por 20 min e três vezes ao dia. Posteriormente, quando as *A. robusta* atingiram aproximadamente 5 cm elas foram transpostas para vasos de 3,8 L contendo terra peneirada proveniente da FCFRP-USP e mantidas na mesma estufa. Após tal espécie atingir pouco mais que 40 cm de altura elas foram divididas em dois grupos: adubadas e não adubadas. Também foram divididas em grupos e adubadas as *T. diversifolia* e *A. trichophylla*, mas a primeira não resistiu a infestação de pragas locais, e a segunda teve um lento crescimento, mesmo na presença de adubação, e não foi monitorada morfologicamente (ver resultados deste capítulo). Deste modo, o cultivo de *A. trichophylla* foi utilizado apenas para estudos de metabolômica aplicada a cultivo (capítulo 4).

Após as *A. robusta* atingirem cerca de 70 cm, elas foram divididas em grupos de 10 e transplantadas para novos locais fora da estufa (10 plantas/tratamento): um canteiro com 35% de sombrite e outro canteiro sem sombrite, cada qual com uma parte que seria novamente adubada. As plantas de *A. robusta* foram transplantadas para os canteiros em julho de 2016, e *A. trichophylla* foi transplantada em janeiro de 2017. Os canteiros construídos, assim como a estufa 50%, se encontram no horto medicinal da FCFRP-USP, um local de área verde de

aproximadamente 10.000 m² contendo diversas plantas de espécies de gêneros e famílias diferentes. Para cada canteiro foi roçada uma área de aproximadamente 21 m² e seu solo foi arejado de modo que as plantas pudessem se acondicionar em dois grupos distintos (adubadas e não adubadas) separados um do outro por cerca de 80 cm (Figura 20). O plantio foi realizado dentro de uma faixa de terra dentro do canteiro, de modo que se fez um perímetro de cerca de 50 cm de largura ao redor das plantas (Figura 20). Dentro do mesmo tratamento, cada planta ficou distante aproximadamente 30 cm uma da outra (Figura 20). Para a construção do canteiro com sombrite 35%, 8 vigas de madeira com mais de 4,5 m foram colocadas no respectivo local de modo que se construísse um viveiro de cerca de 4 m de altura. Para uma maior durabilidade do viveiro, toda a madeira foi vernizada. Além disso, visando o controle de ervas daninhas dentro dos canteiros, no perímetro ao redor dos tratamentos foi colocada manta permeável à água e pedras brita por cima da mesma.

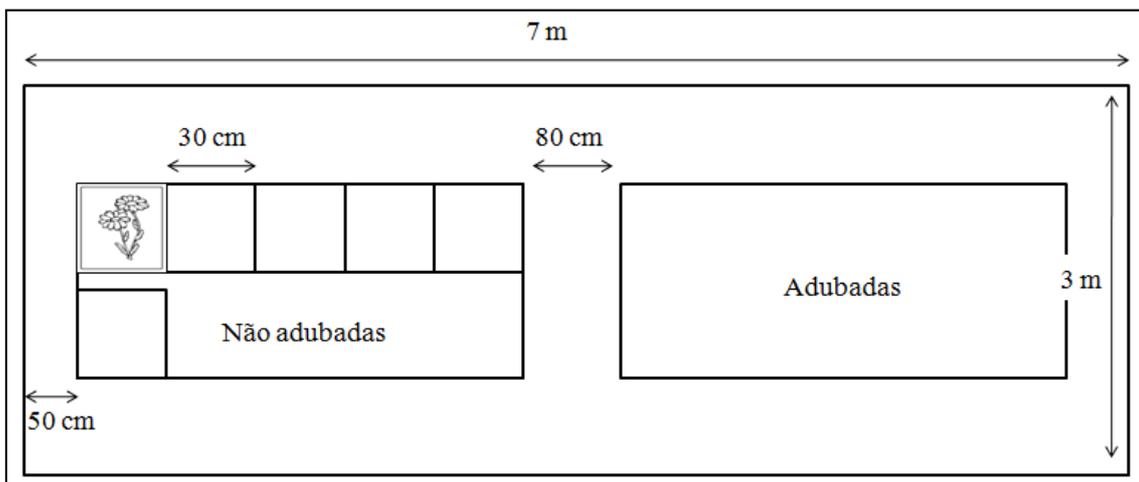


Figura 20 - Esquema mostrando as dimensões dos canteiros sem e com sombrite 35% construídos. Figura de minha autoria.

As plantas dos canteiros foram irrigadas duas vezes ao dia por um período de dois meses até que se verificou uma melhor adaptação aos novos locais. Após estes dois meses, as irrigações diárias foram diminuindo até não se fazer mais irrigação (duração de adaptação de ap. 3 meses). O intuito de não se irrigar mais as plantas é poder fazer a comparação com as plantas selvagens coletadas em seus respectivos *habitat*.

As adubações foram realizadas três vezes, primeiramente em março de 2016, quando todas as plantas ainda se encontravam dentro da estufa com 50% de sombrite, e depois em setembro de 2016 e em março de 2017. Pesou-se 25 g de adubo comercial (1 e ½ colher de sopa cheia, ou cerca de 2.700 kg/ha) (Forth, São Paulo, Brasil) contendo N (13%), P (5%), K (13%),

Ca (1%), B (0,04%), Mn (0,08%), Mg (1%), Cu (0,05%), Mo (0,005%), S (5%), Fe (0,2%), Zn (0,15%) foi espalhado cuidadosamente ao redor das plantas nas duas datas de adubação.

3.5 – Análise morfológica dos cultivares de *A. robusta*

Os cultivares de *T. diversifolia* e *A. trichophylla* não foram avaliadas em termos morfológicos. *T. diversifolia* não resistiu à infestação de insetos dentro da estufa (ver resultados deste capítulo) e *A. trichophylla* apresentou um lento crescimento, mesmo quando exposta à adubação, sendo utilizada apenas para estudos de metabolômica aplicada a cultivo (ver capítulo 4).

Deste modo, o desenvolvimento dos cultivares de *A. robusta* foi analisado sob os seguintes parâmetros: altura do maior ramo principal (em m), quantidade de folhas do maior ramo principal, quantidade de ramos principais, quantidade de ramos secundários, área foliar (em cm²), diâmetro do maior ramo principal (em cm). As comparações foram realizadas aos pares: plantas adubadas *versus* (*vs*) não adubados, sob uma mesma luminosidade; sombrite 35% *vs* sem sombrite, sob uma mesma condição de adubação. Tais parâmetros foram aferidos em meados de dezembro (2016) e janeiro (2017). Para verificar a presença de algum *outlier* em algum determinado parâmetro aferido foi utilizado o seguinte critério:

$L_{sup} < outlier < L_{inf}$, onde L_{sup} = limite superior e L_{inf} = limite inferior, dados pelas fórmulas:

$L_{sup} = \bar{x} + 1.5 \times IQR$ e $L_{inf} = \bar{x} - 1.5 \times IQR$, onde \bar{x} é a média dos valores aferidos e $IQR = \text{Quartil } 3 - \text{Quartil } 1$.

Todos os valores relacionados ao desenvolvimento da planta foram adquiridos em quintuplicata.

Em todos os experimentos, para se comparar a equivalência entre as variâncias foi utilizado o teste F, para testar a diferença das médias foi aplicado o teste t e ANOVA e teste de Tukey para comparação múltipla entre as médias ($p < 0.05$). A maioria das análises foi conduzida via Microsoft Excel e o teste de Tukey no *software* PAST 3.15 (Universidade de Oslo, Noruega).

4 – RESULTADOS E DISCUSSÃO

4.1 – Resultados dos estudos de germinação

A análise da atividade fotoblástica de sementes de *A. robusta* e *A. trichophylla* revelou que apenas sementes de *A. trichophylla* são positivas ao fotoblastismo. Quando exposto a 12 h de luz por dia, houve um total de 51% de sementes de *A. trichophylla* germinadas contra 34% de sementes germinadas no escuro (Tabela 13). Valores de GSI para *A. trichophylla* também mostram que as sementes germinam mais rapidamente quando expostas à luz por 12 h (Tabela 13). Por outro lado, não houve diferença estatisticamente significativa entre os valores de %G, GSI e MGT para as sementes de *A. robusta* germinadas sob 12 h de luz ou no escuro (Tabela 13). Outros autores relataram que a luz é um estímulo de germinação comum para muitas espécies de Asteraceae (BASKIN e BASKIN, 1996). É importante notar que a quantidade de sementes germinadas de *A. trichophylla* (51% sob luz) foi muito próxima do número de sementes germinadas no experimento Bombo e seus colaboradores (54%) (BOMBO et al. 2015), com a diferença de que tais autores utilizaram 8 h de luz, ao invés de 12 h. Por outro lado, a quantidade de sementes de *A. robusta* germinada neste estudo (61%, sob luz) é significativamente menor do que a quantidade que Bombo e seus colaboradores obtiveram (95%). Essa comparação pode indicar que a quantidade ideal de exposição à luz seja, de fato, de 8 h por dia, e não 12 h, como usado neste estudo. No entanto, a fim de manter a padronização das condições iniciais, todas as análises subsequentes foram realizadas com exposição a 12 h de luz por dia.

Tabela 13 - A atividade fotoblástica das sementes de *A. robusta* e *A. trichophylla* foi analisada comparando-se a germinação sob exposição a 12 horas de luz *versus* a ausência total de luz (escuro). Valores de $p < 0,05$ foram considerados estatisticamente significantes. Nota-se que a luz influenciou positivamente a germinação de sementes de *A. trichophylla*, mas não influenciou a germinação de *A. robusta*.

Species	Germination Monitoring	12 h light	dark	p-value
<i>A. trichophylla</i>	%G	(51%) 0.796	(34%) 0.622	0.007
	GSI	0.305	0.220	0.010
	MGT	0.820	0.797	0.126
<i>A. robusta</i>	%G	(61%) 0.897	(58%) 0.866	0.466
	GSI	0.286	0.320	0.135
	MGT	0.881	0.868	0.199

Sementes de *A. robusta* e *A. trichophylla* responderam bem ao aumento da concentração e exposição ao GA₃ (Tabela 14). No entanto, é interessante notar que os valores de %G de *A. trichophylla* observados para o controle e as sementes expostas por 24 h a uma concentração de 5 ppm de GA₃ não são estatisticamente diferentes (Tabela 14). O mesmo pode ser observado para os valores do GSI. Apenas o tratamento que expôs as sementes por 24 h a 40 ppm de GA₃ (Tabela 14) é estatisticamente diferente do controle. De fato, pode-se observar na tabela 14 que o tratamento que fornece maior %G e GSI para sementes de *A. trichophylla* é aquele que expõe as sementes a 40 ppm de GA₃ por 24 h. Quanto a *A. robusta*, os valores de %G do controle diferem significativamente dos tratamentos de 12 h de exposição para 40 ou 320 ppm de GA₃, ou de 24 h de exposição para 5 ppm de GA₃ (Tabela 14). No entanto, os valores de GSI do controle diferem significativamente de 12 h de exposição para qualquer concentração de GA₃ testada, ou de 1 h de exposição a 320 ppm de GA₃ (Tabela 14). Os resultados de *A. robusta* %G e GSI de *A. robusta* descritos na tabela 14 mostram que as sementes de *A. robusta* atingem maior %G (98,33%) e GSI (0,779) na maior concentração de GA₃ e expostas no maior tempo. O mesmo não ocorreu com *A. trichophylla*, que parece começar a sofrer com toxicidade quando exposto por 24 h a 320 ppm de GA₃.

Tabela 14 - Resultados do monitoramento da germinação de *A. robusta* e *A. trichophylla* expostos ao GA₃. Letras iguais na mesma coluna indicam que as diferenças entre médias, considerando o desvio padrão, não são estatisticamente significantes ($p < 0,05$).

Specie	Monitoring	%G	GSI
<i>A. trichophylla</i>	Control	(51.00%) 0.796 a	0.305 a
	1h - 5 ppm	(51.67%) 0.802 a	0.279 a
	1h - 40 ppm	(53.33%) 0.819 a	0.298 a
	1h - 320 ppm	(53.33%) 0.819 a	0.286 a
	12h - 5 ppm	(53.33%) 0.819 a	0.295 a
	12h - 40 ppm	(65.00%) 0.938 a,b	0.317 a
	12h - 320 ppm	(53.33%) 0.819 a	0.290 a
	24h - 5 ppm	(66.67%) 0.956 a,b	0.360 a,b
	24h - 40 ppm	(88.33%) 1.230 c	0.430 b
	24h - 320 ppm	(75.00%) 1.049 b,c	0.345 a,b
<i>A. robusta</i>	Control	(61.00%) 0.897 a	0.286 a
	1h - 5 ppm	(61.67%) 0.904 a	0.446 a,b
	1h - 40 ppm	(65.00%) 0.938 a,b	0.482 a,b,c
	1h - 320 ppm	(65.00%) 0.938 a,b	0.548 b,c,d
	12h - 5 ppm	(71.67%) 1.012 a,b,c	0.607 b,c,d,e
	12h - 40 ppm	(80.00%) 1.109 b,c	0.617 b,c,d,e
	12h - 320 ppm	(93.33%) 1.313 d	0.686 d,e
	24h - 5 ppm	(80.00%) 1.109 b,c	0.626 b,c,d,e
	24h - 40 ppm	(86.67%) 1.198 c,d*	0.675 c,d,e
	24 - 320 ppm	(98.33%) 1.496 d*	0.779 e

Outra observação interessante dos resultados da tabela 14 é que tanto *A. robusta* quanto *A. trichophylla* parecem responder melhor ao tempo de exposição ao GA₃ do que à sua concentração, uma vez que não foi observada diferença significativa entre o controle e as sementes expostas por apenas 1 h a 320 ppm de GA₃, com exceção do GSI observado em *A. robusta*. A observação de que o aumento da concentração de GA₃ influencia positivamente a porcentagem de sementes germinadas está de acordo com a literatura. Por exemplo, Fariman e seus colaboradores mostraram um aumento significativo na porcentagem de sementes de

Echinacea purpurea L. (Asteraceae) após o tratamento com concentrações acima de 100 ppm de GA₃ por 2 h (FARIMAN et al., 2011). Por outro lado, Kirmizi e seus colaboradores também observaram um aumento na porcentagem de sementes germinadas de *Tripleurospermum pichleri*, *Cirsium leucopsis* e *Senecio olympicus* (Asteraceae) tratadas com 100 ppm de GA₃ por 24 h (KIRMIZI et al., 2011).

As sementes de *T. diversifolia* também foram germinadas a 25 °C com 12 h de luz. Embora seja possível observar que as sementes de *T. diversifolia* respondem ao tempo de exposição e às concentrações de GA₃, os valores de %G e GSI para as sementes de *T. diversifolia* foram muito baixos, mesmo quando expostos por 24 h a 320 ppm de GA₃ (Tabela 15). Os baixos valores podem estar relacionados à temperatura e à luz, fatores que não foram previamente avaliados para *T. diversifolia*. Com base na literatura, pode-se inferir que se as sementes de *T. diversifolia* tivessem sido germinadas no escuro e em temperaturas acima de 28 °C, teria sido possível chegar a aproximadamente 30% das sementes germinadas, como nos resultados obtidos por Agboola e seus colaboradores (AGBOOLA et al., 2006). Outro fator que pode estar relacionado à baixa concentração de sementes germinadas é o curto período entre a coleta e a germinação. Em outras palavras, o aumento no período de armazenamento das sementes coletadas até a data de sua germinação parece estar relacionado à quebra de dormência de sementes de *T. diversifolia* (AGBOOLA et al., 2006). No entanto, há uma relação clara entre a concentração de GA₃ e o tempo de exposição com o aumento de sementes de *T. diversifolia* germinadas, fato também observado para *T. rotundifolia* (PATEL et al., 2014) (Tabela 15).

Tabela 15 - Resultados do monitoramento da germinação de sementes de *T. diversifolia* expostas ao GA₃. Letras iguais na mesma coluna indicam que as diferenças entre médias, considerando o desvio padrão, não são estatisticamente significantes ($p < 0,05$). Letras com asteriscos indicam que são estatisticamente diferentes. Apesar da baixa quantidade de sementes germinadas, o aumento de %G e GSI é evidente à medida que a concentração e o tempo de exposição ao GA₃ aumentam. Os valores de TMG não foram estatisticamente diferentes (valor de p : 0,8186).

Monitoring	%G	GSI
Control	(6.67%) 0.215 a	0.022 a
1h - 5 ppm	(6.67%) 0.215 a	0.056 a,b
1h - 40 ppm	(8.33%) 0.290 a,b	0.059 a,b
1h - 320 ppm	(16.67%) 0.420 a,b,c	0.115 a,b,c
12h - 5 ppm	(11.67%) 0.347 a,b,c*,d	0.109 a,b,c
12h - 40 ppm	(16.67%) 0.416 a,b,c,d	0.132 b,c
12h - 320 ppm	(20.00%) 0.462 a,b,c,d	0.143 b,c,d
24h - 5 ppm	(18.33%) 0.442 a,b,c,d	0.135 b,c
24h - 40 ppm	(26.67%) 0.541 b,c,d	0.165 c,d
24 - 320 ppm	(36.67%) 0.650 c*	0.233 d

As sementes de *T. diversifolia* responderam positivamente ao aquecimento em água a 80 °C por 30 s ou 3 min (Tabela 16), obtendo maiores %G e GSI quando comparados ao controle (sem GA₃ e sem aquecimento, Tabela 15). De fato, parece que as sementes de *T. diversifolia* respondem melhor ao aquecimento em água do que a exposição e concentração de GA₃, uma vez que maiores %G e MGT foram obtidos quando as sementes foram expostas à água a 80 °C por 3 min do que quando expostas 24 h a 320 ppm de GA₃ (Tabela 16). Neste caso, o GSI não foi estatisticamente diferente. No entanto, houve um efeito somatório para os valores de %G, GSI e MGT quando o GA₃ foi combinado com o aquecimento da água (Tabela 16). De fato, foi possível germinar 73% das sementes de *T. diversifolia* quando elas foram expostas por 24 h a 320 ppm de GA₃ e subsequentemente submersas em água a 80 °C por 3 min (Tabela 16). A correlação positiva observada entre a dormência de sementes de *T. diversifolia* submersas em água quente corrobora dados da literatura (AGBOOLA et al., 2006), onde foi possível germinar quase 90% das sementes de *T. diversifolia* quando aquecidas em água a 80 °C durante 3 min.

Tabela 16 - Resultados do monitoramento da germinação de sementes de *T. diversifolia* expostas a aquecimento a 80 °C por 30 s e 3 min. Letras iguais na mesma coluna indicam que as diferenças entre médias, considerando o desvio padrão, não são estatisticamente significantes ($p < 0,05$). Letras com asteriscos indicam que são estatisticamente diferentes. Pode-se notar que o aquecimento por 3 min tem uma influência maior do que o GA₃ na germinação de *T. diversifolia*, embora ambos os fatores combinados tenham resultado em um maior nível de germinação (73%).

Monitoring	%G	GSI	MGT
Control 24h - 320 ppm	(36.67%) 0.650 a	0.233 a	0.776 a
80 °C - 30 s	(11.25%) 0.339	0.105	0.865 b
80 °C - 3 min	(51.25%) 0.798	0.255 a	0.863 b
80 °C - 30 s – GA ₃	(37.50%) 0.659 a	0.208 a	0.821 a,b*
80 °C - 3 min – GA ₃	(73.00%) 1.026	0.314	0.908 b*

4.2 – Resultados do cultivo de *A. robusta*

A espécie *T. diversifolia* não resistiu às doenças geradas principalmente por uma constante infestação de mosca branca (*Bemisia argentifolii*) dentro da estufa com 50% de sombrite e irrigação automática (três vezes ao dia e por 20-30 min). Dezenas desta mosca incessantemente se alojavam na parte abaxial das folhas de todas as espécies cultivadas. Entretanto, parecia haver uma predileção pela *T. diversifolia*. Na tentativa de exterminar a praga, foram realizadas diversas aspersões de fumo, óleo de neem e até mesmo organoclorado. O intervalo entre uma aplicação e outra nunca foi realizado dentro de um intervalo de tempo inferior a três semanas, visando não intoxicar as plantas. Entretanto, a população de *B. argentifolli* sempre diminuiu, mas nunca era exterminada de fato. Parece que os fatores que influenciaram para as re-infestações foram a umidade do local, a presença de pedras brita (as moscas se alojavam entre os vãos úmidos das pedras) e a baixa luminosidade. Além disso, por duas vezes foi necessário aplicar solução de CuSO₄ visando conter infestação por fungos.

No que diz respeito ao acúmulo de mosquitos na parte abaxial das folhas de *T. diversifolia*, descrita no parágrafo anterior, é sabido que as LST se acumulam em tricomas glandulares justamente na parte abaxial das folhas. O efeito tóxico e deterrente de LST tem sido estudado frente aos insetos. Assim, Ambrósio e colaboradores relataram que extratos foliares de *T. diversifolia* ricos em LST possuem atividade inseticida contra lagarta da espécie *Chlosyne lacinia* (Lepidoptera) (AMBRÓSIO et al., 2008). Por outro lado, o trabalho de Martucci e Gobbo-Neto sugere que a larva *C. lacinia* pode ter apresentado tolerância à *T. diversifolia* e

apontam outro metabólito como tendo atividade deterrente de insetos, o ácido 3-*O-E*-cafeoilquínico, da classe dos derivados dos ácidos *trans*-cinâmicos, e não LST (MARTUCCI e GOBBO-NETO, 2016). Assim sendo, tendo em vista dados da literatura e a infestação de mosquitos na parte abaxial das folhas de *T. diversifolia* de nosso cultivo, consideramos que mais pesquisas sejam importantes a fim de se investigar melhor os metabólitos com capacidade inseticida em plantas.

A intensa umidade e presença de insetos na estufa motivou a construção de canteiros e de um viveiro com sombrite 35% sem irrigação automática. Deste modo, os canteiros construídos (com e sem sombrite 35%) (Figuras 21 e 22) passaram a abrigar apenas as *A. robusta*. As *A. trichophylla* foram transplantadas para canteiros sem sombrite apenas em janeiro de 2017, portanto, mais tardiamente que *A. robusta* (julho de 2016). O motivo foi que *A. trichophylla* estava apresentando um lento crescimento e seus ramos mostravam-se mais frágeis que as espécies selvagens do Paraná, mesmo aquelas que receberam adubação. Deste modo, tal espécie foi apenas avaliada em termos de metabolômica de cultivo (capítulo 4).

Os ramos de *A. robusta* também apresentaram pouca rigidez, ainda mais tendo em vista muita presença de ventanias no local. Deste modo, todas as *Aldama* tiveram que ter seu ramo principal amarrado a um suporte afixado ao solo.



Figura 21 - Canteiro construído com sombrite 35%. À esquerda plantas adubadas, à direita, plantas sem adubo. Figura de minha autoria.



Figura 22 - Canteiro construído sem sombrite. À esquerda plantas adubadas, à direita, plantas sem adubo. Figura de minha autoria.

Para aos cultivares de *A. robusta*, não foi verificado *outlier* em qualquer aferição dos parâmetros estudados. As comparações foram realizadas aos pares: plantas adubadas *versus* não adubadas, sob uma mesma luminosidade; sombrite 35% *versus* sem sombrite, sob uma mesma condição de adubação. Os valores das médias dos parâmetros analisados para cada tratamento encontram-se na tabela 17. Foi constatado que o parâmetro “altura do ramo principal” obteve diferença significativa apenas quando se comparou plantas adubadas provenientes de cultivos com e sem sombrite. O “número de folhas do maior ramo principal” parece ser o parâmetro mais importante, dando diferença entre todas as comparações, menos quando se compara plantas não adubadas sob diferentes luminosidades. Quando se compara plantas sem adubação, mas cultivadas em diferentes condições de luminosidade, verifica-se que somente houve diferença significativa no “número de ramos secundários”. Este parâmetro também foi significativo quando comparado plantas adubadas e não adubadas sem sombrite. Por outro lado, quando se compara plantas com adubação, mas cultivadas em diferentes condições de luminosidade, verifica-se que houve diferença na “altura do maior ramo principal” e seu “número de folhas”. O “diâmetro do ramo principal” e a “área do ramo principal” foram parâmetros analisados cujas médias obtidas foram estatisticamente significativas para plantas adubadas e não adubadas cultivadas sob sombrite 35%. Não houve diferença significativa apresentada pela média do “número de ramos principais” quando analisado qualquer tratamento.

Tabela 17 - Valores de médias para os parâmetros de desenvolvimento dos cultivares de *A. robusta*. Letras iguais na mesma linha indicam diferenças estatisticamente significativas ($p < 0.05$).

Parâmetros	Sombrite 35%		Sem sombrite	
	não adubada	adubada	não adubada	adubada
Altura do ramo principal (m)	3.8	4.14 c	3.66	3.84 c
n° de folhas do ramo principal	112 a	154 a,c	140 b	198 b,c
Área foliar (cm ²)	45.4 a	64.6 a	54.8	63.8
n° de ramos principais	3.6	4.6	3	4.6
n° de ramos secundários	8.4 c	8.2	1 b,c	6 b
diâmetro do ramo principal (cm)	1.18 a	1.48 a	1.4	1.56

Pelos resultados expostos na tabela 17, pode-se resumir que a adubação de espécies dentro do viveiro com sombrite foi capaz de gerar plantas com maior número de folhas, maior diâmetro do ramo principal e maior área foliar. Por outro lado, a adubação em canteiro sem sombrite foi capaz de produzir apenas maior número de folhas do ramo principal. A floração foi um parâmetro que não foi aferido neste estudo, mesmo porque a floração de *A. robusta* ocorre em meados de abril e maio e, tal como comentado em Material e Métodos (item 3.5), a coleta de dados morfológicos foi realizada em dezembro (2016) e janeiro (2017). Posteriormente à coleta destes dados, verificou-se que algumas plantas dentro do canteiro com sombrite começavam a florescer em março. De fato, a floração das plantas de dentro do viveiro aconteceu antes que as de fora. Aqui pode-se comparar estes resultados com aqueles observados por Ruggiero e colaboradores, quando eles afirmam que *A. robusta* floresce de novembro a abril (RUGGIERO et al., 1997). Em nossas observações de campo, até o presente momento não foi constatada florada de *A. robusta* antes de março. Nosso grupo de pesquisas frequentemente realizou coletas de *A. arenaria* em Itirapina (SP) em meados de fevereiro e março, quando é observada a florada de tal espécie. É importante notar que Ruggiero e colaboradores relatam que coletaram *A. robusta* em Itirapina (RUGGIERO et al., 1997). Também é interessante notar a semelhança anatômica entre tais espécies, tal como destacado pela taxonomista M. Magenta (MAGENTA, 2006) em seu trabalho de doutorado “*Viguiera* Kunth (Asteraceae, Heliantheae) na América do Sul e sistemática das espécies do Brasil”. Ademais, nossa observação *a posteriori* (dados não demonstrados) de que houve floração mais cedo e mais volumosa dentro

do viveiro, do que fora, está parcialmente de acordo com a análise de Ruggiero e coautores, os quais afirmam que há *A. robusta* floresce mais cedo, mas em menor volume, quando cultivada sob menor exposição à luz (RUGGIERO et al., 1997). A floração mais cedo também foi verificada para *A. discolor*, quando exposta a menor exposição fotodiária (ISEJIMA et al., 1991). Outra comparação que se pode traçar com dados da literatura, é a altura do ramo principal atingida por *A. robusta* neste experimento. Neste caso, contando mais de um ano de cultivo, as plantas atingiram em média 4 m de altura, valor ligeiramente maior que o observado em campo, para as plantas selvagens. No experimento de Ruggiero e coautores, *A. robusta* após 285 dias de cultivo (cerca de 9 meses) alcançou cerca de 1 m e já apresentando florada (RUGGIERO et al., 1997). Durante praticamente o mesmo período, a *A. robusta* cultivada neste trabalho apresentou altura similar quando ainda se encontrava em vaso e dentro da estufa com irrigação automática, passando a crescer rapidamente quando foi cultivada em canteiro. De fato, o experimento de Ruggiero também foi conduzido em vasos, o que pode sugerir que esta condição estivesse impedindo a planta de atingir maiores altura (RUGGIERO et al., 1997). O trabalho de Ruggiero não informa como foi feito o aporte de água às plantas, mas fotoperíodos mais curtos (ou sombrite) somadas a uma maior umidade podem influenciar em floração mais cedo em *A. robusta*. Por fim, a única constatação feita para *A. trichophylla* (dados não informados), é que seus espécies cultivados neste trabalho nunca atingiram 1,5 m, ao passo que as plantas selvagens alcançam cerca de 3 m de altura. Solos foram coletados nos respectivos locais de coleta de sementes e poderão ser analisados futuramente, tendo em vista que o solo do local de coleta de sementes de *A. trichophylla* apresentou-se mais arenoso e pobre em nutrientes que o solo utilizado nos cultivos deste trabalho.

5 – CONCLUSÃO DO CAPÍTULO 3

Foi verificado que apenas *A. trichophylla* possui atividade fotoblastica. Observou-se também que todas as espécies respondem ao aumento da concentração de GA₃, obtendo-se maiores valores de %G, GSI e MGT, embora pouco significativo em *T. diversifolia*. Estas, responderam melhor ao aquecimento em água (80 °C por 3 min). As plântulas de *T. diversifolia* morreram precocemente, e *A. trichophylla* apresentou um crescimento lento e frágil de seus ramos. Em *A. robusta*, verificou-se que a adubação de espécies sombreadas foi capaz de gerar plantas com maior número de folhas, maior diâmetro do ramo principal e maior área foliar. Por outro lado, a adubação em canteiro sem sombrite foi capaz de produzir apenas maior número

de folhas do ramo principal. Os resultados aqui descritos contribuem para um maior número de germinação de sementes, bem como melhor desenvolvimento morfológico de *Aldama*, os quais poderão ser comparados com outras análises de outras espécies.

CAPÍTULO 4: METABOLÔMICA DO CULTIVO DE *Aldama robusta* E *Aldama trichophylla* (ASTERACEAE) PARA ANÁLISE DA INFLUÊNCIA DO SOLO, LUMINOSA, CIRCADIANA E ESTACIONAL

LISTA DE FIGURAS DO CAPÍTULO 4

Figura 23. Sobreposição das amostras correspondentes ao pool 1. A substância que elui em ap. 15 min. corresponde ao padrão interno crisina.....	131
Figura 24. Sobreposição das amostras correspondentes ao pool 1. A substância que elui em ap. 15 min. corresponde ao padrão interno crisina.....	132
Figura 25. TIC das triplicatas de branco injetadas antes e durante as análises. Não é possível visualizar nenhuma substância de concentração significativa que possa comprometer a análise.....	132
Figura 26. Controle de qualidade metabolômico inspecionado pelo agrupamento de pools.....	133
Figura 27. PCA com pools e todas as amostras provenientes das coletas dos cultivos. Para visualizar o código dos cultivares, checar Apêndice B.....	134
Figura 28. OPLS-DA mostrando diferença entre as coletas das 11 horas realizadas com N ₂ líquido (nitrogen) versus secamento por estufa (air oven). OPLS-DA (1 + 6): R ² = 0,96, Q ² = 0,79.....	135
Figura 29. PCA (R ² = 0,50) mostrando tendência de agrupamento das amostras coletadas no inverno (Winter) e primavera (Spring) contra as amostras coletadas no verão (Summer) e outono (Autumn).....	139
Figura 30. OPLS-DA (3 + 10) (R ² = 0,95 e Q ² = 0,69) mostrando clara distinção entre amostras coletadas no inverno (winter) e primavera (spring) em relação às coletas do verão (summer) e outono (autumn).....	140
Figura 31. OPLS-DA (2 + 10) (R ² = 0,94 e Q ² = 0,78) com amostras coletadas no verão (summer) e outono (autumn) consideradas um único grupo indistinto.....	141
Figura 32. TIC dos padrões para curva de calibração.....	145
Figura 33. Curva de calibração para 9 substâncias com base nas suas concentrações e intensidade de ionização registrada pela área do TIC.....	146

LISTA DE TABELAS E QUADROS DO CAPÍTULO 4

Tabela 18. Identificação das variáveis discriminantes da análise comparativa de coletas seguidas pela secagem de planta com N ₂ líquido e liofilização versus coletas seguidas da secagem em estufa de ar-circulante (40 °C).....	136
Tabela 19. Identificação das variáveis discriminantes de cada estação de coleta.....	143
Quadro 1. Área da curva de cada substância, seus respectivos tempos de retenção e [M+H] ⁺ em relação à diluição seriada em µg/ml.....	146
Quadro 2. Concentração estimada de substâncias químicas em <i>A. robuta</i> com base em curva de calibração.....	148

LISTA DE ABREVIATURAS, SIGLAS E SIMBOLOS DO CAPÍTULO 4

As mesmas do capítulo 1, além de

TIC

Cromatograma de íons totais

RESUMO DO CAPÍTULO 4

No capítulo 4 foi analisado a influências de alguns fatores ambientais (GOBBO-NETO e LOPES, 2007) no metaboloma de plantas, tomando-se como modelo experimental espécies de *A. robusta* e *A. trichophylla* cultivadas. A ideia inicial seria não somente verificar a alteração metabólica ocasionada por tais fatores, mas também analisar se isso poderia ocasionar equívocos em análise quimiotaconômicas. Assim, por meio dos resultados seria indicado quais são os fatores ambientais que mais podem alterar o metaboloma e, portanto, a classificação quimiotaconômica. Além disso, os resultados poderiam indicar marcadores químicos ambientais, relacionados ou não com fatores adversos, como o estresse hídrico em plantas. Para tanto, a análise metabolômica foi conduzida com base em dados químicos, adquiridos em UHPLC-ESI-HRMS, provenientes de diversas coletas relacionadas com o ritmo circadiano, estação do ano, adubação do solo, luminosidade e rápida interrupção do metabolismo (com congelamento em N₂ líquido). Além disso, foi realizado uma curva de calibração com base na intensidade do pico de 9 substâncias de diferentes classes químicas registrado no TIC, a fim de se estimar a concentração dos metabólitos desreplcados. Os dados foram analisados por PCA e OPLS-DA. De maneira geral, tanto por PCA quanto OPLS-DA pode-se observar que as diferenças metabolômicas provenientes de fatores como o ritmo circadiano, adubação do solo e luminosidade são aleatórias, não sendo possível agrupar as amostras. Por outro lado, foi possível verificar diferenças metabólicas em folhas coletadas que não foram rapidamente congeladas, mas foram secadas em estufa. Com base nisso, alguns metabólitos correlacionados com amostras congeladas em N₂ líquidos foram propostos e corroborados pela literatura (MOREIRA et al., 2018). Além disso, também foi possível notar clara influência que as estações do ano tem no metabolismo das plantas, com agrupamento de amostras coletadas em estações mais úmidas (verão e outono), separadas das estações mais secas (inverno e primavera). Do ponto de vista estacional, não há diferença entre coletas realizadas no verão e outono. Com base nas análises de OPLS-DA, foi possível indicar alguns picos discriminantes de tais estações e propor alguns metabólitos que se acumulam no inverno, tendo em vista, provavelmente, o estresse hídrico. A curva de calibração construída também permitiu estimar a concentração de alguns metabólitos que se acumulam em cada estação. Os resultados deste capítulo nos permite concluir que o fator ambiental que mais pode causar equívocos na quimiotaconomia é a data da coleta ao longo do ano, bem como o processamento do material vegetal (congelado ou secado em estufa). Além disso, foi apontado alguns metabólitos que podem ser considerados como marcadores estacionais, ou provenientes de degradação pelo não congelamento das amostras. Estes ultimo resultados permite melhor controle de qualidade de amostras vegetais e estudos quimiotaconômicos.

SUMMARY OF CHAPTER 4

In chapter 4 the influences of some environmental factors (GOBBO-NETO and LOPES, 2007) on the metabolome of plants were analyzed, taking as an experimental model species of *A. robusta* and *A. trichophylla* cultivated. The initial idea was not only to verify the metabolic alteration caused by such factors, but also to analyze if this could cause mistakes in chemotaxonomic analysis. Thus, the results would indicate which are the environmental factors that can most alter the metabolome and, therefore, chemotaxonomic classification. In addition, the results could indicate environmental chemical markers, related or not with adverse factors, such as water stress in plants. For this, the metabolomic analysis was conducted based on chemical data, acquired in UHPLC-ESI-HRMS, from various collections related to circadian rhythm, season of the year, soil fertilization, luminosity and rapid interruption of metabolism (with freezing in N₂ liquid). In addition, a calibration curve was performed based on the intensity of the peak of 9 substances of different chemical classes registered in the TIC, in order to estimate the concentration of the dereplicated metabolites. The data were analyzed by PCA and OPLS-DA. In general, according both PCA and OPLS-DA can be observed that metabolomic differences from factors such as circadian rhythm, soil fertilization and luminosity are aleatory, and it is not possible to group the samples. On the other hand, it was possible to verify metabolic differences in collected leaves that were not rapidly frozen, but were dried in an air-oven. Based on this, some metabolites correlated with samples frozen in liquid N₂ were proposed and corroborated by the literature (MOREIRA et al., 2018). In addition, it was also possible to observe a clear influence of the seasons on plant metabolism, with the collection of samples collected in wetter seasons (summer and autumn), separated from the drier seasons (winter and spring). From the seasonal point of view, there is no difference between collections performed in the summer and autumn. Based on the OPLS-DA analyzes, it was possible to indicate some discriminant peaks of these seasons and to propose some metabolites that accumulate in the winter, probably due to the lack of water. The constructed calibration curve also allowed to estimate the concentration of some metabolites that accumulate in each season. The results of this chapter allow us to conclude that the environmental factor that may cause misunderstandings in chemotaxonomy is the date of collection throughout the year, as well as the processing of the plant material (frozen or dried in the air-oven). In addition, it was pointed out some metabolites that can be considered as seasonal markers, or from degradation by not freezing the samples. The latter results allows better quality control of plant samples and chemotaxonomic studies.

1 – INTRODUÇÃO

As análises em metabolômica também podem ser utilizadas para investigar os efeitos de fatores ambientais no metabolismo de plantas com base em uma abordagem holística. Isso pode ser trazer novas descobertas em estudos relacionados à ecoquímica, adaptabilidade de plantas, quimiotaxonomia e quimiossistemática. Especificamente em relação a estas duas últimas, frequentemente se comenta que os metabólitos secundários não são bons marcadores taxonômicos, uma vez que quali e quantitativamente suas presenças sofrem mais influência do meio ambiente do que os marcadores genéticos, ditos “conservados”. Em relação a isso, a metabolômica ambiental aplicada a cultivares poderia apontar melhores formas de padronização dos estudos quimiotaxonomia/quimiossistemática, validando tais estudos. Por exemplo, uma vez determinado o padrão de biossíntese (ou acúmulo) de metabólitos secundários frente a diversas condições ambientais, como tipo de solo, umidade, temperatura, ciclo circadiano, luminosidade e sazonalidade, as análises quimiotaxonômicas e quimiossistemáticas poderiam ser padronizadas precisamente em relação a tal resultado. Deste modo, tendo as condições ambientais padronizadas para todos os indivíduos do estudo, as análises quimiotaxonômicas e quimiossistemáticas seriam realizadas com maior confiabilidade.

A literatura descreve diversos fatores ambientais que podem alterar o metaboloma das plantas (GOBBO-NETO e Lopes, 2007), tais como idade da planta (ou do material vegetal), sazonalidade, ritmo circadiano, disponibilidade hídrica, tipo de solo, luminosidade, altitude (o que reflete em temperaturas mais baixas, menor concentração de O₂ e maior taxa de radiação ultravioleta), poluição atmosférica, relação com outros seres vivos (GOBBO-NETO e Lopes, 2007; HARMER et al., 2009; ROZEMA et al., 1997). Por exemplo, flavonoides glicosilados como apigenina 7,4’ di-*O*-glucuronídeo e luteolina 7,3’ di-*O*-glucuronídeo poderiam ter seu acúmulo alterado pela influência da intensidade ou quantidade de radiação ultravioleta (GOBBO-NETO e LOPES, 2007). De fato, a biossíntese de flavonoides em plantas tem sido frequentemente relatada na literatura como um metabólito que desempenha papel protetor contra os danos da radiação UV (AGATI et al., 2012), além de outros metabólitos, como taninos e ligninas (ROZEMA et al., 1997). Tais metabólitos tem em comum grupamentos fenólicos e, portanto, poderiam desempenhar função protetora em plantas contra estresses oxidativo. Tendo em vista diversos fatores ambientais que podem alterar o metaboloma de plantas, a metabolômica ambiental tem sido uma abordagem inovadora em encontrar sutis diferenças metabólicas causadas por alterações ambientais, especialmente aquelas causadas por estresses (LANKADURAI et al., 2013). Por exemplo, a herbivoria seria capaz de aumentar o acúmulo

de ácido quínico em folhas de *Solanum lycopersicum* (tomate) e diminuir o ácido *trans*-3-cafeoilquínico (LANKADURAI et al., 2013).

Pode-se também citar exemplos de como a fenologia altera o metaboloma de plantas, tal como demonstrado por um estudo metabolômico de LC-MS de *Hypericum organifolium* (Hypericaceae) (ÇIRAK et al., 2007). Da mesma forma, os estudos metabolômicos também desempenham um papel importante na compreensão da relação das plantas com outros seres vivos, como um estudo metabolômico realizado por GC-MS e LC-MS que demonstrou que interações entre a via do ácido jasmônico e salicílico modulam o metaboloma da planta e afetam herbívoros de diferentes tipos de alimentação (SCHWEIGER et al., 2014). Outro estudo demonstra a influência da alelopatia e a sazonalidade no metabolismo vegetal, como demonstrado por análises metabolômicas baseada na RMN para a discriminação e caracterização química de espécies de plantas mediterrâneas (SCOGNAMIGLIO et al., 2015; SCOGNAMIGLIO et al, 2014). A metabolômica baseada em RMN também foi utilizada para avaliar a relação entre perfil metabólico e análise elementar de extratos de folhas de *Erica multiflora* (Ericaceae) de acordo com a sazonalidade, crescimento de plantas e condições de estresse moderado de seca e aquecimento (RIVAS-UBACH et al., 2012). Por fim, um estudo focado não apenas na sazonalidade, mas também na origem geográfica e influência de fatores abióticos ambientais, foi capaz de discriminar cultivares de *T. diversifolia* (Asteraceae) com base em dados concatenados de RMN e LC-MS (SAMPAIO et al, 2016). Os resultados mostraram que um padrão sazonal foi observado para os açúcares, LST e fenólicos na folha e caules, que podem ser correlacionados à quantidade de chuva e às mudanças de temperatura. Além disso, a distribuição dos metabólitos nas inflorescências e raízes foi afetada principalmente pela variação de alguns macro e micronutrientes do solo, como Ca, Mg, P, K e Cu.

Deste modo, visando a uma maior compreensão da influência de fatores ambientais no metaboloma de plantas, foram escolhidos como modelo experimental cultivares de *A. robusta* e *A. trichophylla* (tribo Heliantheae, Asteraceae) para uma análise metabolômica ambiental, com a finalidade de encontrar padrões de agrupamento químico de acordo com fatores externos, como ritmo circadiano, tipo de solo (adubação), sazonalidade (estações do ano) e luminosidade. O intuito é auxiliar na compreensão da quimiotaxonomia de *Aldama*, ou mesmo de Asteraceae, padronizando condições ambientais de coleta. Além disso, foi avaliada a influência da secagem de plantas coletas em N₂ líquido, ou em estufa de ar-circulante, uma vez que não há trabalhos na literatura que tenham comparado tais abordagens, embora alguns tenham protocolado que

análises metabolômicas sejam realizadas em N₂ líquido (DE VOS et al., 2007), e outros tenham indicado o acúmulo de metabólitos em plantas coletadas e não congeladas (MOREIRA et al., 2018). O intuito de se utilizar N₂ líquido seria realizar o “*quenching metabolic*”, ou a parada do metabolismo por congelamento do material vegetal. Isso evitaria achados equivocados, identificando como metabólitos discriminantes produtos de degradação, por exemplo (DE VOS et al., 2007). Além disso, o N₂ líquido é útil para se diminuir o ruído em análises metabolômicas (DE VOS et al., 2007), problema este facilmente solucionado pelo processamento dos dados brutos (no MZmine, por exemplo).

2 – OBJETIVOS

Os objetivos deste capítulo são:

- 1) verificar se o agrupamento dos cultivares de *A. robusta* e *A. trichophylla* obedece mais a fatores intrínsecos (genótipo), do que fatores extrínsecos (meio ambiente);
- 2) verificar por análises não supervisionadas se os fatores ambientais analisados são capazes de agrupar os cultivares;
- 3) apontar por meio de OPLS-DA os metabólitos mais correlacionados com algum parâmetro ambiental analisado.

3 – MATERIAL E MÉTODOS

3.1 – *Material vegetal*

O material vegetal dos cultivares de *A. robusta* e *A. trichophylla* foi a coleta de suas folhas. O código de cada coleta se encontra no Apêndice B.

Três folhas de cinco espécies foram coletadas aleatoriamente de cada experimento, sendo que cada experimento possui 10 espécimes.

3.2 – *Parâmetros ambientais e de coleta analisados*

Foi analisada a coleta de folhas e sua subsequente secagem em estufa de ar-circulante a 40 °C *versus* congelamento em N₂ líquido seguido de liofilização.

A coleta foi também realizada em ritmo circadiano (7, 11, 15, 20, 24 h), estacional (verão, outono, inverno, primavera; sempre no segundo mês do início da estação), em plantas não sombreadas e sombreadas (dentro de casa de vegetação com sombrite 35%) e em plantas não adubadas e adubadas.

A comparação de secagem em estufa ou N₂ líquido foi apenas realizada para as coletas das 11 h, e não foi realizada esta comparação nas coletas do outono. A espécie *A. trichophylla* não foi coletada no verão, uma vez que ela ainda estava em crescimento e adaptação ao canteiro de cultivo. A espécie *A. trichophylla* também só foi coletada às 11 h, com exceção do outono, quando foi coletada em ritmo circadiano. Maiores descrições podem ser observadas no Apêndice 2.

3.3 – *Extração e aquisição dos dados químicos*

A droga vegetal foi pulverizada em moinho analítico (IKA[®] A11, Sigma Aldrich) quando proveniente da estufa a 40 °C ou em gral com pistilo com auxílio de N₂ líquido. As folhas congeladas e pulverizadas com N₂ líquido foram posteriormente secadas em liofilizador. Após a pulverização, foram extraídas na proporção de 20 mg de droga vegetal seca por 1.000 µL de EtOH:H₂O (4:1) por 10 min, posteriormente vortexados e centrifugados a 10.000 g por 15 min (Eppendorf 5810R, Alemanha). O sobrenadante contendo o extrato foi coletado e diluído na proporção de 1:3, segundo padronizações prévias com outros extratos que visaram verificar o rendimento médio da extração. Por fim, o extrato foi filtrado em membranas Mini-Uniprep[™] PTFE de 0.2 µm (Whatman, EUA) e 4 µL foram injetados em UHPLC -ESI-HRMS, com analisador Exactive Plus Orbitrap[™] (Thermo Scientific, Alemanha).

Para o controle de qualidade da extração e da análise em LC-MS (monitoramento do tempo de retenção e ionização), foram realizados 2 *pools*, cada qual com uma mistura de folhas de espécies diferentes. Esses *pools* foram extraídos nas mesmas condições três vezes e re-injetados três vezes. Também foi utilizado o flavonoide crisina como controle interno da análise, estando a uma concentração de 10 µg/ml na amostra. Triplicata de brancos foi utilizada no início e no meio da eluição, a fim de se monitorar a coluna cromatográfica e a presença de contaminantes.

Os *pools* foram constituídos por: *pool 1* = *Achillea millefolium*, *A. robusta*, *A. nudibasilaris*, *T. diversifolia*, *Mikania laevigata*, *Smallanthus sonchifolius*, *Calea* sp e *Chiquiraga spinosa*; *pool 2* = *A. trichophylla*, *A. grandiflora*, *A. pilosa*, *H. annuus*, *Solidago microglosa*, *Calea* sp e *Chiquiraga spinosa*.

Para o monitoramento da concentração de metabólitos conhecidos e majoritários nos cultivares, foi construída uma curva de calibração com 10 substâncias químicas de diferentes classes, nas concentrações de 60; 30; 15; 7,5; 3,75; 1,875; e 0,937 µg/mL. As substâncias selecionadas foram (B) budleína A, (P) partenolideo, (G) 2-ona-8β-metacrilóiloxi-guaia-3,10(14),11(13)-trien-6α,12-olideo (guaianolido), (E) 15-hidroxieremantolido, (K) canferol, (R) rutina, (A) apigenina, (C) catequina, (Q) ácido 5-*O-E*-cafeoilquínico e (H) 7-hidroxycumarina. Em parênteses, os respectivos códigos de cada substância.

Os métodos cromatográficos e a aquisição de dados químicos são os mesmos do capítulo 1.

3.4 – Processamento dos dados químicos

O processamento de dados químicos são os mesmos do capítulo 1.

3.5– Análises quimiométricas

Os dados químicos de todos os extratos após serem processados foram exportados em planilha (.csv), organizados de modo a colocar as variáveis nas colunas e amostras nas linhas, e as áreas sob os picos detectados foram normalizadas por Pareto para posterior análise multivariada. Normalização por Pareto:

Pareto: $y_{ij} = x_{ij} - \bar{x}_i / \sqrt{SD_{ij}}$, onde \bar{x} é a média da área dos picos detectados e SD é o

desvio padrão.

A PCA e OPLS-DA foram realizadas no *software* SIMCA v.13.0.3 (Umetrics – Suécia). A avaliação dos resultados foi realizada por meio da inspeção dos valores de coeficiente de correlação múltipla (R_x^2 e R_y^2), capacidade preditiva (Q^2 , *leave-one-out 7-fold cross validation*) e raiz quadrada do erro médio da validação cruzada (RMSEcv). Valores de R^2 superiores a 0.6 e de Q^2 superiores a 0.3 são desejáveis em metabolômica, considerando-se presença de *overfitting* quando os valores de $R^2 - Q^2$ forem superiores a 0.3. Valores de RMSEcv próximos a 0 são desejáveis. Além disso, os resultados foram também avaliados a um nível de significância de 95% (valor $p < 0,005$).

4 – RESULTADOS E DISCUSSÃO

O controle de qualidade dos dados químicos para a investigação dos efeitos ambientais sobre o metaboloma dos cultivares de *A. robusta* e *A. trichophylla* foi realizado por meio de *pools* (ver Material e Métodos, pág. X), bem como da flavona crisina adicionada às amostras a uma concentração de 10 $\mu\text{g/mL}$. Deste modo, antes que se fizesse a análise metabolômica dos cultivares, foi realizada a inspeção dos tempos de retenção e ionização das substâncias dos *pools* (Figuras 23 e 24).

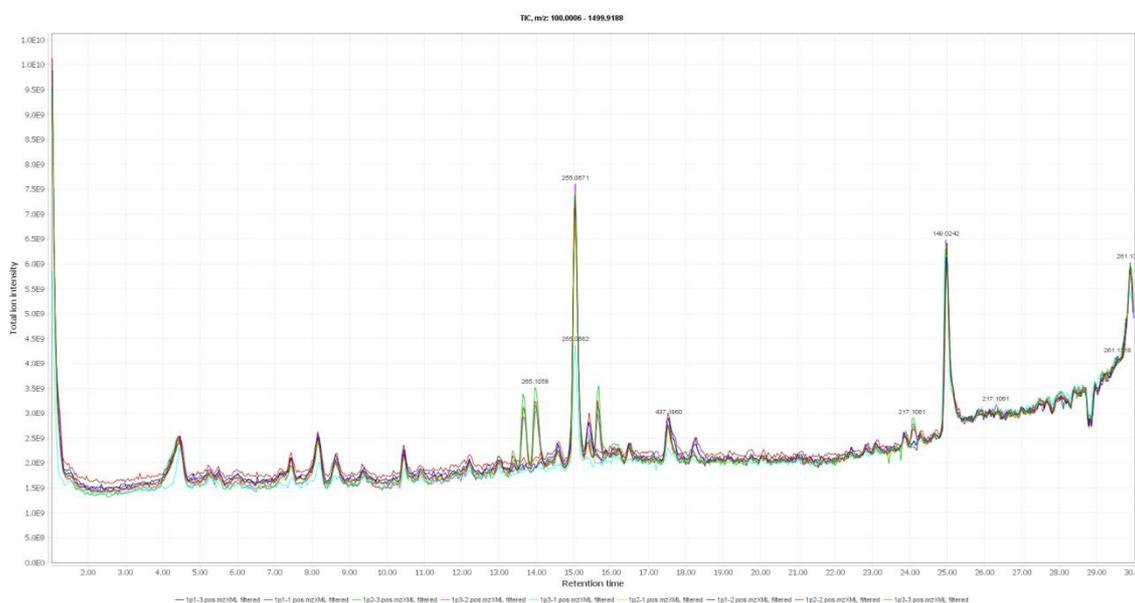


Figura 23 - Sobreposição das amostras correspondentes ao *pool* 1. A substância que elui em ap. 15 min. corresponde ao padrão interno crisina.

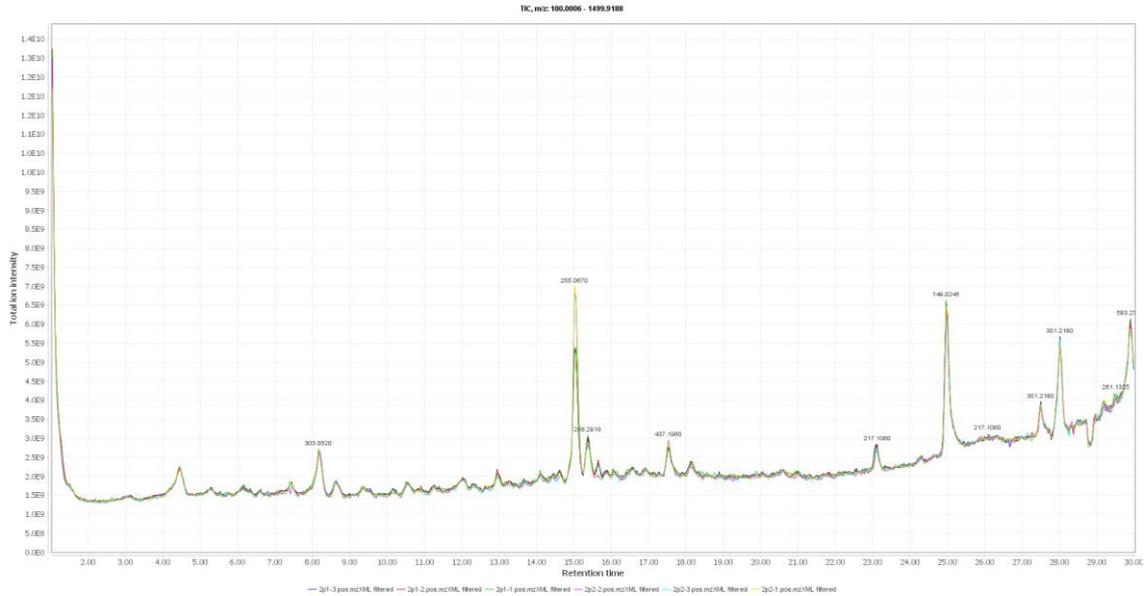


Figura 24 - Sobreposição das amostras correspondentes ao *pool* 1. A substância que elui em ap. 15 min. corresponde ao padrão interno crisina. O *pool* correspondente à terceira injeção (2p3 apresentou problemas de análise e não foi apresentado neste cromatograma).

Além disso, os cromatograma de íons totais (TIC) da triplicata de brancos injetada antes, e durante, as análises não demonstra nenhuma substância contaminante do solvente em concentração significativa que possa comprometer a análise, exceto aquela que elui por volta de 25 min (Figura 25).



Figura 25 – TIC das triplicatas de branco injetadas antes e durante as análises. Não é possível visualizar nenhuma substância de concentração significativa que possa comprometer a análise. O pico que aparece em 15,10 min de m/z 255,0660 [M + H]⁺ corresponde ao padrão interno crisina. O pico que sai em 25 min é um contaminante do solvente

Os controle de qualidade também foram realizados em termos metabolômicos com a inspeção dos *pools* por PCA (Figura 26). Por meio desta figura, pode-se notar que há agrupamento dos *pools* quando comparado com o metaboloma de amostras proveniente de extratos de apenas uma espécie. Essa observação confirma que as substâncias do nosso sistema LC-MS estão tendo boa reprodutibilidade em termos de tempo de retenção e ionização em ESI, tendo em vista que os *pools* deste experimento são réplicas de extração (triplicata) e aquisição de dados LC-MS (triplicata).

Assim, tendo em vista a qualidade dos dados obtidos e a confiabilidade do método, as análises de metabolômica aplicadas sobre a influência de fatores ambientais ao cultivo de *A. robusta* e *A. trichophylla* foram analisadas. Com base nisso, foi obtida uma PCA (Figura 27) com todas as amostras provenientes de todas as coletas de todos os cultivares deste estudo (para visualizar o código dos cultivares, checar Apêndice B. Para visualizar os parâmetros ambientais analisados em cada coleta, checar Material e Métodos deste capítulo e Apêndice B).

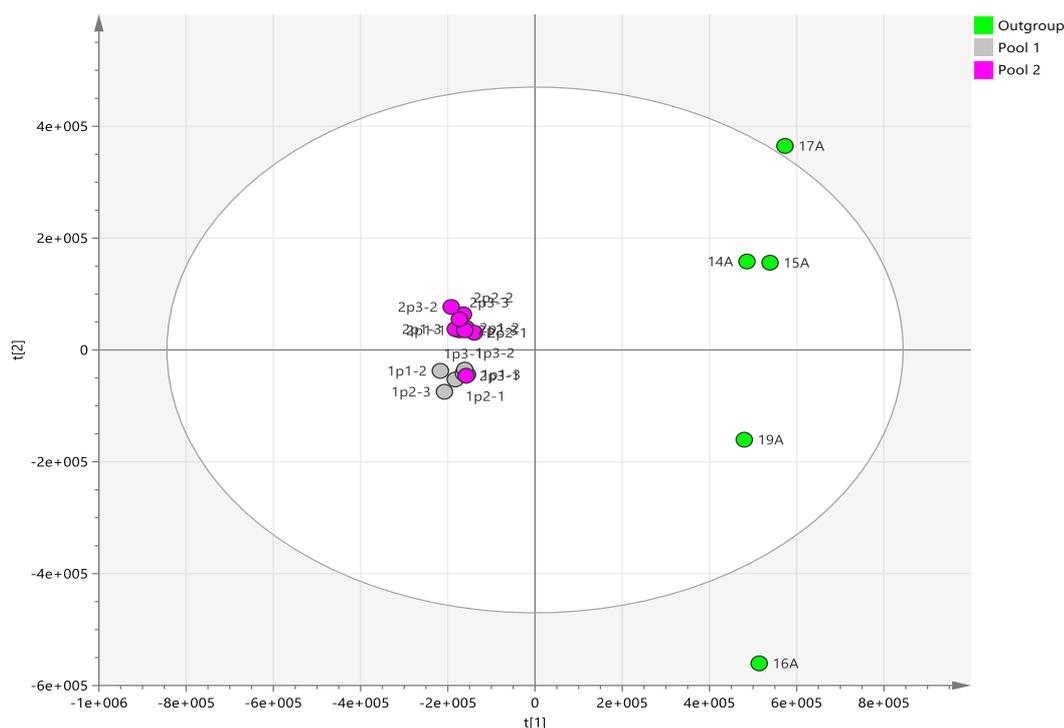


Figura 26 - Controle de qualidade metabolômico inspecionado pelo agrupamento de *pools*.

Pela inspeção da PCA apresentada na Figura 27 pode-se notar que mesmo com mais amostras acrescentadas ao sistema, os *pools* ainda permanecem agrupados. Também pode-se notar que a PCA parece mostrar uma tendência de agrupamento entre amostras provenientes de

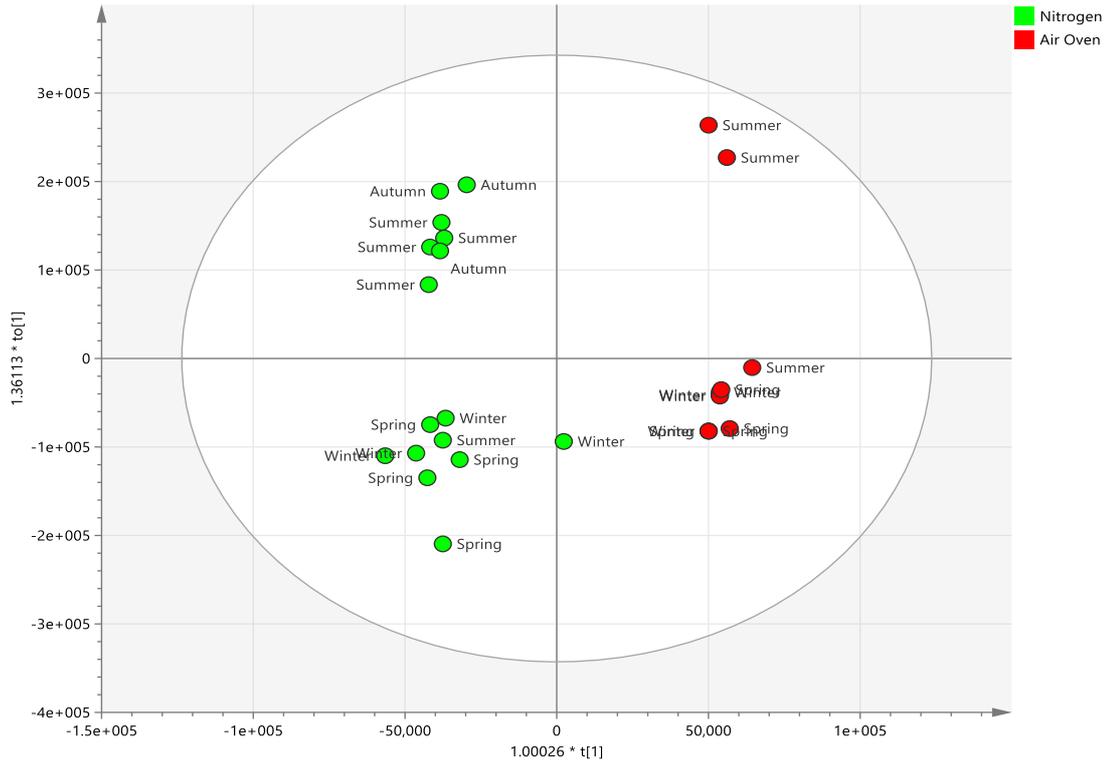


Figura 28 - OPLS-DA mostrando diferença entre as coletas das 11 horas realizadas com N₂ líquido (nitrogen) versus secamento por estufa (air oven). OPLS-DA (1 + 6): R² = 0,96, Q² = 0,79. Por meio desta análise já foi possível visualizar separação das amostras de acordo com a estação de coleta. Valor-p = 0,0048 e RMSEcv = 0,22.

Ainda com tais dados foi possível notar que para as coletas realizadas com N₂ líquido houve separação entre as amostras em dois grupos distintos, relacionados com coletas realizadas no verão e outono e aquelas realizadas na primavera e inverno (Figura 28). Esta análise revelou bons indicadores analíticos, tais como alto valor de Q², e baixos valores de RMSEcv e valor-p, o que sugere ausência de *overfitting* e alta significância na análise (Figura 28). Deste modo, foi possível apontar metabólitos discriminantes correlacionados com coletas realizadas com N₂ líquido e aquelas realizadas com subsequente secamento em estufa (Tabela 18).

Tabela 18 – Identificação das variáveis discriminantes da análise comparativa de coletas seguidas pela secagem de planta com N₂ líquido e liofilização versus coletas seguidas da secagem em estufa de ar-circulante (40 °C). O sobrescrito “std” se refere a “standard” (substâncias do banco de padrões do AsterBioChem, previamente isolados ou comprados), os quais foram injetados no mesmo método cromatográfico a fim de se obter valores de tempo de retenção e perfil de fragmentação. Deste modo, as desreplicações que apontam para substâncias pertencentes ao banco de padrões são mais confiáveis, uma vez que não somente houve correlação com os valores de *m/z*, mas também com o tempo de retenção (RT). Var ID = número de identificação e respectivo modo de ionização em que se detectou o pico (pos = positivo; neg = negativo). As estruturas podem ser visualizadas em <http://zinc15.docking.org/catalogs/asternp/substances/>, em “Lookup”, ou em <http://143.107.203.160:8080/>, em “Compound Search”, inserindo-se seu nome usual ou código InChIKey.

Var ID	RT	<i>m/z</i>	Identificação de variáveis discriminantes para os extratos de coletas secados em N ₂ líquido e liofilizador (nome usual ou código InChIKey)	Classe
5 pos	1,05	144,102	n.i.	
213 pos	14,03	274,274	n.i.	
177 pos	11,08	245,064	n.i.	
26 pos	1,71	113,965	n.i.	
31 pos	1,99	355,106	chlorogenic acid; neochlorogenic acid (pouco prováveis, tempo de retenção baixo)	Derivados do ácido cinâmico
105 pos	6,84	591,212	n.i.	
158 pos	9,27	481,21	n.i.	
254 pos	18,23	217,107	n.i.	
32 pos	2,04	102,035	n.i.	
76 pos	5,351	435,189	n.i.	
1 pos	1,00	118,087	L-valine ^{std}	Aminoácido
253 pos	18,04	316,325	n.i.	
237 pos	16,10	173,08	n.i.	
251 pos	17,61	255,068	n.i.	
86 pos	5,74	163,041	4-hydroxycoumarin; umbelliferone; p-coumaric acid ^{std}	Derivados do ácido cinâmico

20 pos	1,41	236,152	n.i.	
190 pos	11,98	331,085	tricin; 2-(3,4-dihydroxyphenyl)-5-hydroxy-3,6-dimethoxychromen-4-one;	Flavonóides
			quercetin 3,4'-dimethyl ether; desmethoxycentaureidin; cirsiliol; jaceosidin;	Flavonóides
			(2R,3R)-5,7-dihydroxy-2-(4-hydroxyphenyl)-4-oxo-2,3-dihydro-1-benzopyran-3-yl acetate	Flavonóides
95 pos	6,26	177,057	n.i.	
66 pos	4,64	135,046	n.i.	
168 pos	10,44	163,041	4-hydroxycoumarin; umbelliferone	Derivados do ácido cinâmico
93 pos	6,10	423,131	n.i.	
19 pos	1,35	162,114	n.i.	
Var ID	RT	m/z	Identificação de variáveis discriminantes para os extratos de coletas secadas em estufa (nome usual ou código InChIKey)	Classe
24 pos	1,66	182,987	n.i.	
34 pos	2,32	154,992	n.i.	

Todavia, foram poucas as variáveis estatisticamente confiáveis que puderam ser geradas para os extratos provenientes de coletas secadas em estufa. Já para as variáveis apontadas para coletas realizadas com N₂ líquido, foi possível identificar alguns picos com base no banco de AsterDB e biblioteca de substâncias puras. Assim, foram identificados flavonoides e derivados do ácido *trans*-cinâmico, dentre os quais, o ácido *p*-cumárico teve seu tempo de retenção encontrado no banco de dados de substâncias puras. Interessante observar que tal substância apontada como discriminante para plantas coletadas com N₂ líquido e liofilizadas, em vez de secadas a 40 °C, vai ao encontro com o que foi constatado por Moreira e coautores (MOREIRA et al., 2018). Nesse trabalho os autores concluíram que ácidos cafeoilquínicos, derivados biossinteticamente do ácido *p*-cumárico, se acumulam nas folhas após a coleta, caso não sejam congeladas (MOREIRA et al., 2018). Isso indica a necessidade em se proceder estudos de metabolômica ambiental, ou quimiotaixonomia, com N₂ líquido a fim de se evitar falsas conclusões, ao se apontar metabólitos que estão mais correlacionados com o estresse pós-coleta, do que o metabolismo da planta propriamente dito. A abordagem metabolômica deste experimento foi capaz de apontar um metabólito correlacionado com o *quenching metabolite* e corroborada pela literatura.

As demais análises deste capítulo não evidenciaram qualquer agrupamento entre as coletas com base no ritmo circadiano, na adubação e no sombreamento (sombrite 35%) (dados não mostrados). Deste modo, pode-se inferir que, neste estudo, não há correlação entre o metaboloma e tais variações ambientais. Ou seja, tais parâmetros não são capazes de alterar significativamente o metabolismo vegetal a ponto de produzir diferenças entre as amostras. Dito isso, a coleta para fins quimiotaixonômicos poderia ser realizada em qualquer horário do dia, por exemplo, que não seria capaz de afetar profundamente os agrupamentos visualizados em uma PCA ou HCAbp. Por outro lado, o processamento vegetal após a coleta (parágrafo anterior) e a sazonalidade (próximos parágrafos) foram capazes de formar grupos que se distinguiram com base na constituição química. Quanto à sazonalidade, pode se perceber agrupamento das espécies de *A. robusta* de acordo com a estação em que foram coletadas. Deste ponto em diante as análises foram realizadas somente com *A. robusta* coletadas e congeladas em N₂ líquido a fim de facilitar o entendimento do modelo. Assim, a PCA demonstra clara separação entre coletas realizadas no outono e verão, daquelas realizadas na primavera e inverno (Figura 29).

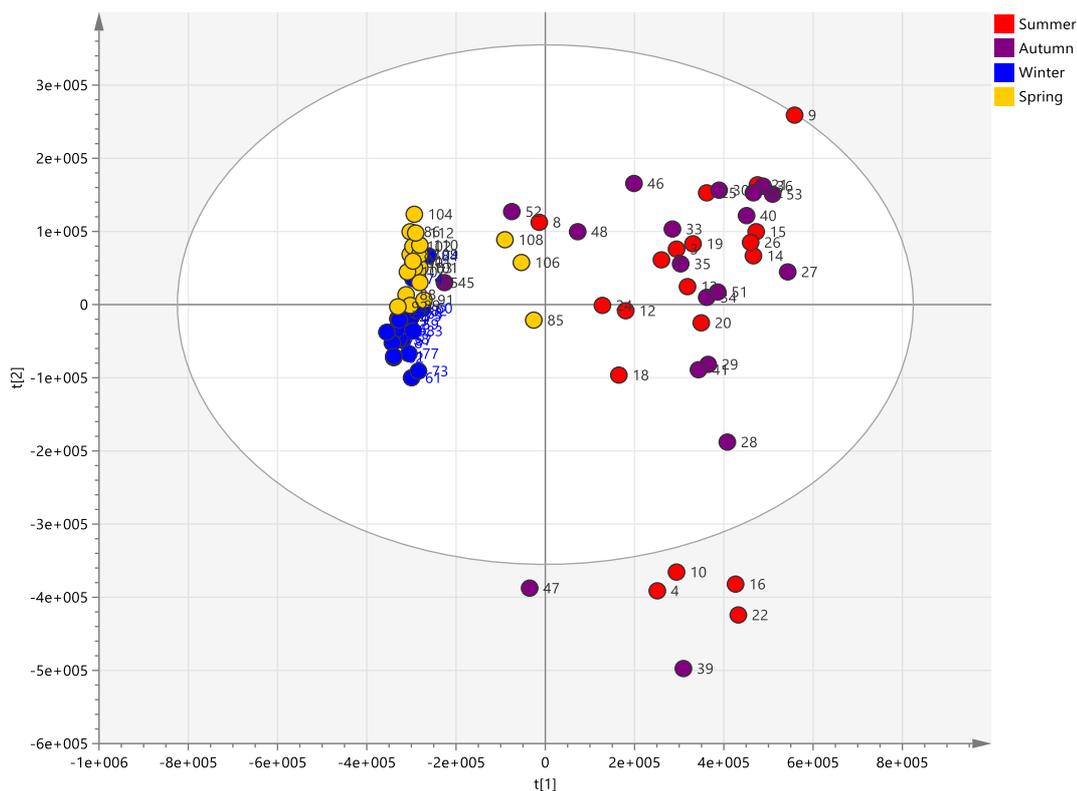


Figura 29 - PCA ($R^2 = 0,50$) mostrando tendência de agrupamento das amostras coletadas no inverno (Winter) e primavera (Spring) contra as amostras coletadas no verão (Summer) e outono (Autumn). Pela PCA podemos perceber que as amostras das estações secas ficam mais agrupadas, o que revela menor diversidade de metabólitos.

Pela PCA pode-se perceber que as amostras das estações mais secas (inverno e primavera) ficam mais agrupadas, o que revela menor diversidade de metabólitos. Além disso, pode-se notar também que quase ocorre uma separação entre as amostras do inverno e primavera, ao passo que aquelas provenientes do verão e outono são mais homogêneas, com alterações maiores, porém aleatórias em seu metabolismo. Dando continuidade à investigação de tais dados, foi realizada uma OPLS-DA. Com uma certa surpresa, ao observar o gráfico das amostras do OPLS-DA (Figura 30), observa-se que houve maior agrupamento entre as amostras provenientes das coletas realizadas no verão e outono, e uma clara separação entre aquelas realizadas na primavera e inverno. Isso revela que os agrupamentos observados na análise da PCA (Figura 29) não estão totalmente correlacionados com as estações. Do ponto de vista estacional, não houve diferença entre coletas realizadas no verão ou outono e o agrupamento das coletas do inverno e primavera se deve mais à pouca variabilidade química, do que a fatores climáticos. Contudo, mesmo com pouca variabilidade química, o OPLS-DA apontou clara diferença entre as duas estações (Figura 30). Tais constatações levaram-nos a considerar para a próxima OPLS-DA as coletas de outono e verão como sendo uma classe única (Figura 31).

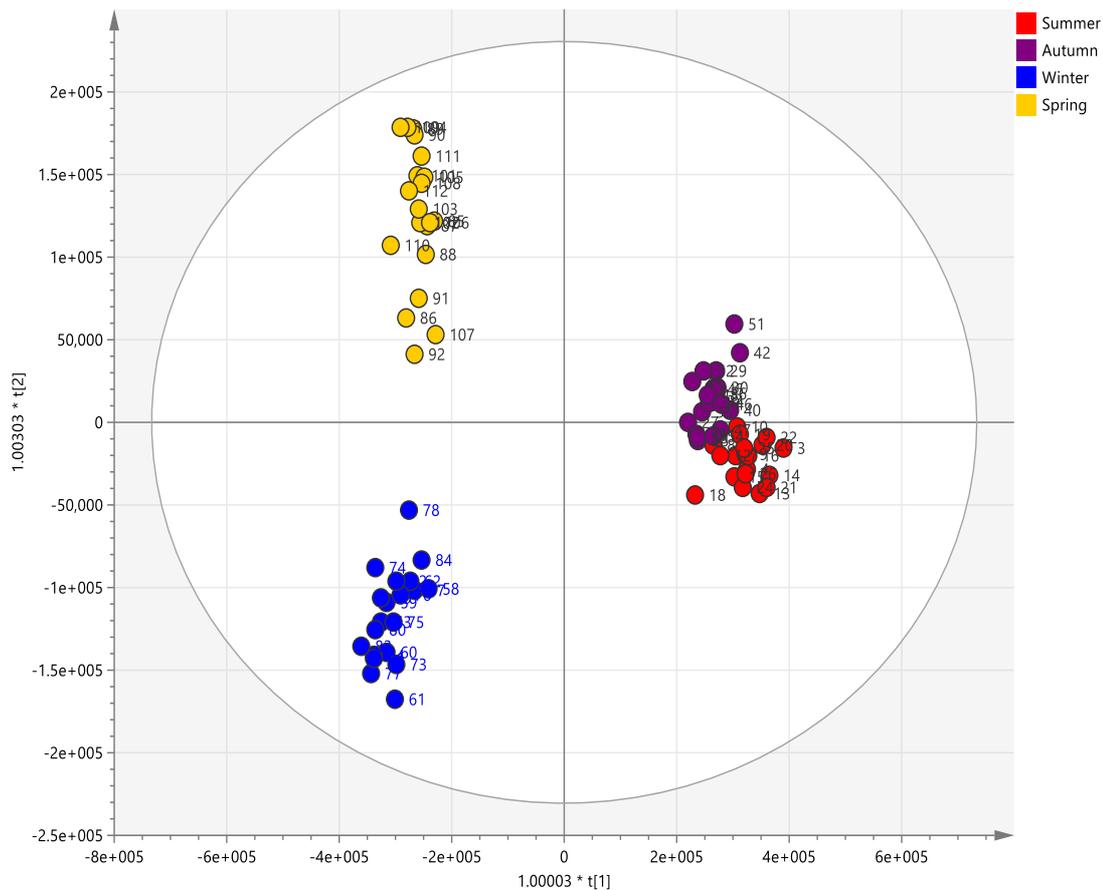


Figura 30 - OPLS-DA (3 + 10) ($R^2 = 0,95$ e $Q^2 = 0,69$) mostrando clara distinção entre amostras coletadas no inverno (winter) e primavera (spring) em relação às coletas do verão (summer) e outono (autumn). Entretanto, percebe-se que, quando o metaboloma é analisado do ponto de vista estacional, há pouca diferença entre as coletas realizadas no verão e outono, ao passo que as coletas realizadas na primavera e inverno se distinguiram entre si, formando outros dois grupos distintos (eixo ortogonal). Valor-p = $2,06 \times 10^{-13}$ e RMSEcv < 0,27.

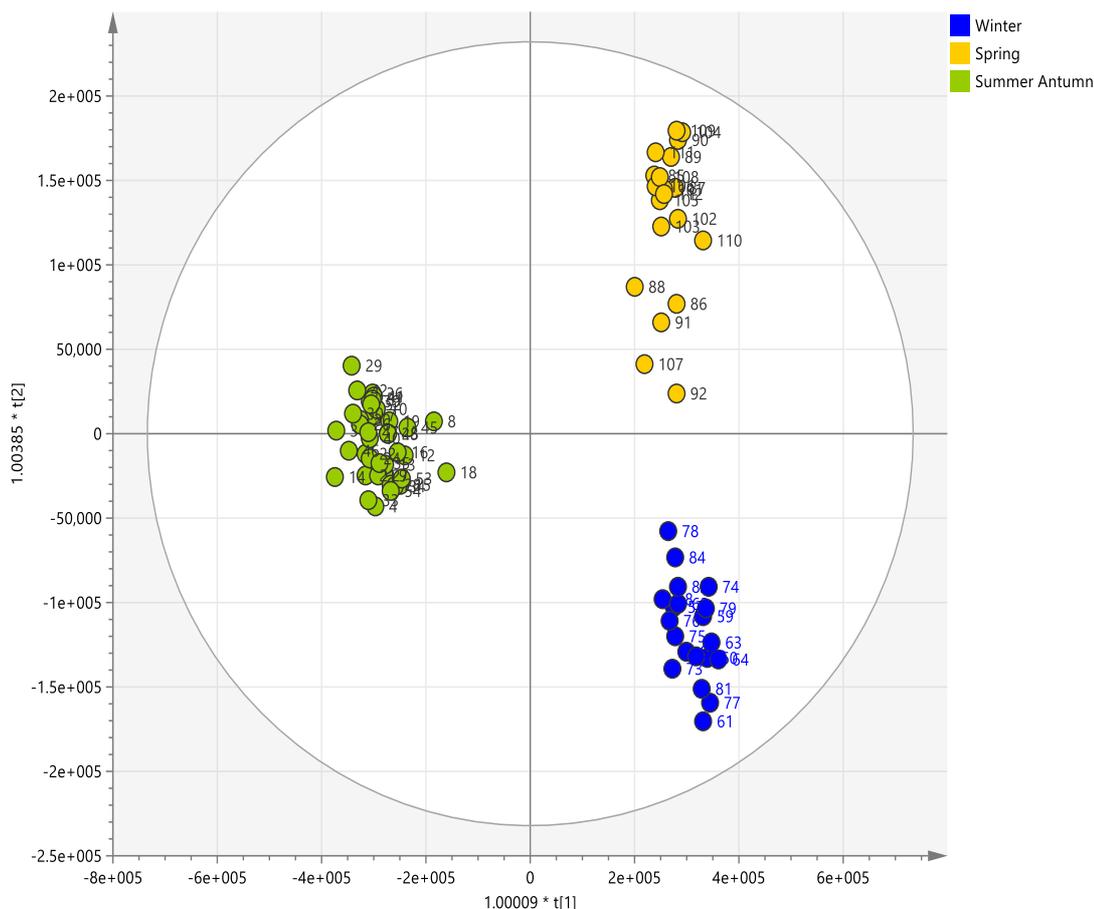


Figura 31 - OPLS-DA (2 + 10) ($R^2 = 0,94$ e $Q^2 = 0,78$) com amostras coletadas no verão (summer) e outono (antumn) consideradas um único grupo indistinto. Percebe-se a distinção deste grupo das coeltas realizadas no inverno (winter) e primavera (spring). Por meio deste gráfico de OPLS-DA foi gerado uma lista de variáveis discriminantes (picos, que podem ser metabólitos) de cada classe. Valor-p = $4,45 \times 10^{-23}$ e $RMSE_{cv} < 0,22$.

Este último modelo gerado permitiu clara separação entre as classes ($R^2 = 0,94$), com alto valor de validação cruzada ($Q^2 = 0,78$) e baixo valor de $RMSE_{cv}$ e valor-p, o que sugere que o agrupamento observado não se deve ao acaso e possui significância estatística. Deste modo, foi possível gerar uma lista de variáveis discriminantes e correlacionadas com as coletas realizadas em cada estação (Tabela 19). Assim, observa-se na tabela 19 que quase todos os metabólitos identificados como discriminantes de folhas coletadas no inverno são da classe dos sesquiterpenos e derivados do ácido *trans*-cinâmico, com exceção do aldeído protocatecuico, um derivado do ácido quínico. Alguns dos metabólitos aqui identificados têm bons níveis de identificação metabolômico (SUMNER et al., 2007; SALEK et al., 2013), tendo sido encontrado não somente sua $[M+H]^+$ ou $[M-H]^-$ nos bancos de dados, mas também o tempo de retenção, tais como os ácidos 3-O-cafeoil-2-(2-butil)-2-hidroxiбутanodioico o e o 1,5-O-dicafeoilquinico (e estereoisômeros). Entretanto, não foi possível identificar qualquer pico para

as folhas coletadas no inverno (variáveis discriminantes não mostradas), e apenas foi identificada uma substância, da classe das LST do tipo ambrosanolido, para uma única variável discriminante das amostras provenientes de coletas realizadas no verão e outono (variáveis discriminantes não mostradas). Ainda assim, pode-se considerar que a correlação das variáveis desreplicadas como derivados do ácidos *trans*-cinâmicos, em especial os ácidos clorogênicos, e quínico com a estação de inverno se corrobora com informações da literatura, onde trabalhos demonstram concentrações aumentadas de ácido cafeico e clorogênico em plantas de *Nicotiana tabacum* e *Mahonia repens* cultivadas sob baixas temperaturas (GOBBO-NETO e LOPES, 2007). O ácido cafeico foi também encontrado em concentrações maiores em amostras de *Baccharis dentata* coletadas no inverno, quando comparadas com outras estações (SARTOR et al., 2013). Tais resultados podem sugerir marcadores químicos ambientais de Asteraceae, possivelmente relacionados com o estresse hídrico ocasionado em períodos de inverno (ou pós-inverno). Deste modo, para resultados mais robustos em quimiotaxonomia, as datas das coletas deveriam ser padronizadas e realizadas preferencialmente foram de tais períodos.

Tabela 19 – Identificação das variáveis discriminantes de cada estação de coleta. O sobrescrito “std” se refere a “standard” (substâncias do banco de padrões do AsterBioChem, previamente isolados ou comprados), os quais foram injetados no mesmo método cromatográfico a fim de se obter valores de tempo de retenção e perfil de fragmentação. Deste modo, as desreplicações que apontam para substâncias pertencentes ao banco de padrões são mais confiáveis, uma vez que não somente houve correlação com os valores de *m/z*, mas também com o tempo de retenção (RT). Var ID = número de identificação e respectivo modo de ionização em que se detectou o pico (pos = positivo; neg = negativo). As estruturas podem ser visualizadas em <http://zinc15.docking.org/catalogs/asternp/substances/>, em “Lookup”, ou em <http://143.107.203.160:8080/>, em “Compound Search”, inserindo-se seu nome usual ou código InChIKey. N.i. = pico não identificado.

Var ID	RT	m/z	Identificação Inverno (nome usual ou código InChIKey)	Classe
278 pos	21,29	203,181	(1R,4S)-1,6-dimethyl-4-(propan-2-yl)-1,2,3,4-tetrahydronaphthalene	Sesquiterpeno
			(1S,4S)-1,6-dimethyl-4-(propan-2-yl)-1,2,3,4-tetrahydronaphthalene	Sesquiterpeno
			curcumene	Sesquiterpeno
288 pos	23,32	203,180	(1R,4S)-1,6-dimethyl-4-(propan-2-yl)-1,2,3,4-tetrahydronaphthalene	Sesquiterpeno
			(1S,4S)-1,6-dimethyl-4-(propan-2-yl)-1,2,3,4-tetrahydronaphthalene	Sesquiterpeno
			curcumene	Sesquiterpeno
269 neg	8,65	190,928	n.i.	
203 pos	13,00	247,098	6-hydroxy-7-[(3-methylbut-2-en-1-yl)oxy]chromen-2-one	Derivado do ácido cinâmico
124 neg	4,92	177,019	n.i.	
5 pos	1,05	144,102	n.i.	

90 neg	4,00	137,024	protocatechuic aldehyde	Derivado do ác. quínico
308 neg	10,04	195,030	n.i.	
65 neg	2,49	335,892	n.i.	
148 pos	8,28	163,039	4-hydroxycoumarin; umbelliferone	Derivado do ácido cinâmico
252 neg	8,15	351,073	3-O-caffeoyl-2-(2-propyl)-2-hydroxybutanedioic acid ^{std}	Derivado do ácido cinâmico
63 pos	4,47	163,040	4-hydroxycoumarin; umbelliferone	Derivado do ácido cinâmico
297 neg	9,55	365,089	3-O-caffeoyl-2-(2-butyl)-2-hydroxybutanedioic acid ^{std}	Derivado do ácido cinâmico
140 pos	8,15	517,137	4,5-di-O-E-caffeoylquinic acid ^{std} ; 3,4-di-O-E-caffeoylquinic acid ^{std} ;	Derivado do ácido cinâmico
			1,5-O-dicaffeoylquinic acid ^{std} ; 3,5-O-dicaffeoylquinic acid ^{std} ;	Derivado do ácido cinâmico
			4,5-O-dicaffeoylquinic acid ^{std}	Derivado do ácido cinâmico

Por fim, para o monitoramento da concentração de metabólitos conhecidos e majoritários nos cultivares, foi construída uma curva de calibração com 10 substâncias químicas de diferentes classes, nas concentrações de 60, 30, 15, 7,5, 3,75, 1,875, 0,937 $\mu\text{g/ml}$. As substâncias escolhidas foram: (B) budleína A, (P) partenolideo, (G) 2-ona-8 β -metacrililoxi-guaia-3,10(14),11(13)-trien-6 α ,12-olideo (guaianolido), (E) 15-hidroxi-remantolido, (K) canferol, (R) rutina, (A) apigenina, (C) catequina, (Q) ácido 5-*O-E*-cafeoilquínico e (H) 7-hidroxycumarina. Em parênteses, os respectivos códigos de cada substância. A curva de calibração de tais substâncias, primeiramente, revela que a intensidade do pico observado no cromatograma de íons totais (TIC) se correlaciona não somente com a substância, mas também com a concentração da substância (Figura 32 e Quadro 1). A substância catequina não ioniza no modo positivo e, portanto, não foi analisada. Isso revela que toda a abordagem metabolômica utilizada em todas as análises desta tese que correlaciona concentração de pico com atividade biológica é válida. Além disso, podemos observar boa resolução cromatográfica até mesmo para substâncias que eluem em tempos de retenção próximos (Figura 32).

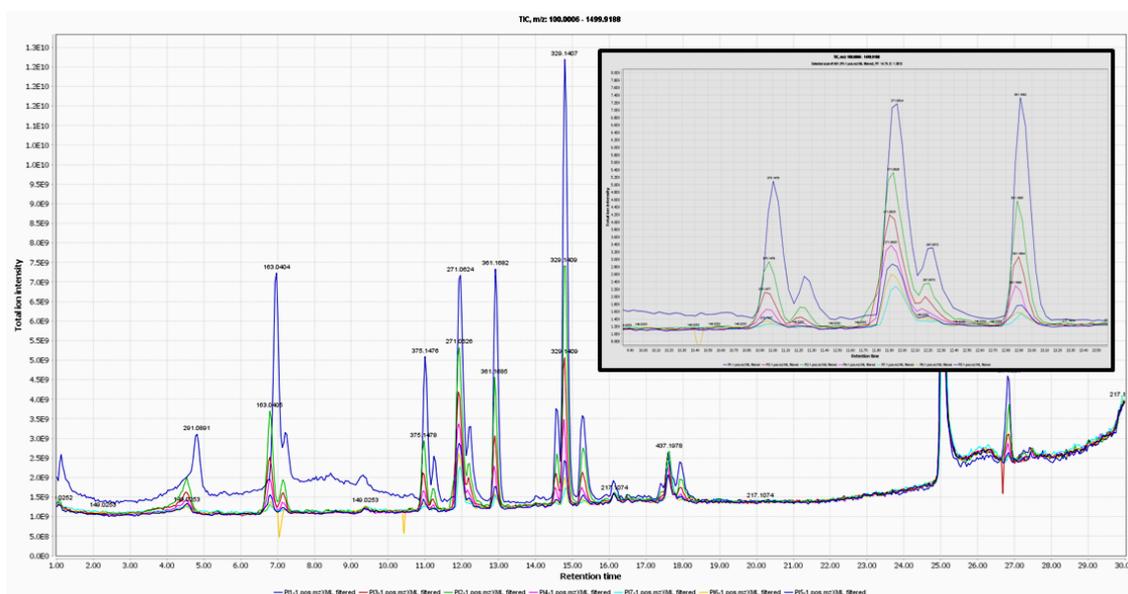


Figura 32 - TIC dos padrões para curva de calibração. Em destaque o zoom do cromatograma dos picos que eluem em de 10 min. a 13 min., aproximadamente. A ordem em que as substâncias aparecem é: Q = 4,87 min., H = 6,91 min., R = 7,09 min., B = 11,12 min., A = 11,86 min., K = 12,5 min., E = 12,83 min., P = 14,17 min., G = 15,00 min. Para consultar os códigos, checar material e métodos deste capítulo. Podemos notar pelo zoom em destaque que a intensidade do pico sofreu proporcionalmente em relação à concentração de substâncias. A substância catequina não ionizou no modo positivo.

	[M+H] ⁺	355.102	163.039	611.160	375.144	271.060	287.055	379.174	231.138	329.138
	RT	4.87	6.91	7.09	11.12	11.86	12.15	12.83	14.17	15.00
Diluição	Diluição (ug/ml)	Q	H	R	B	A	K	E	P	G
1	60.0	1.40E+08	9.10E+08	1.40E+08	4.90E+08	1.20E+09	3.40E+08	5.80E+08	1.90E+08	1.00E+09
2	30.0	6.20E+07	5.00E+08	8.80E+07	2.40E+08	8.40E+08	1.90E+08	3.30E+08	9.80E+07	5.70E+08
3	15	4.70E+07	2.80E+08	5.20E+07	1.40E+08	6.10E+08	1.20E+08	1.90E+08	5.60E+07	3.50E+08
4	7.5	3.50E+07	1.70E+08	2.90E+07	8.70E+07	4.50E+08	7.20E+07	1.10E+08	3.10E+07	2.10E+08
5	3.75	3.00E+07	9.60E+07	2.00E+07	4.40E+07	3.40E+08	4.40E+07	5.60E+07	2.00E+07	1.10E+08
6	1.875	0.00E+00	5.50E+07	0.00E+00	2.50E+07	2.70E+08	2.80E+07	3.10E+07	0.00E+00	6.50E+07
7	0.9375	0.00E+00	3.20E+07	0.00E+00	0.00E+00	2.10E+08	0.00E+00	0.00E+00	0.00E+00	4.00E+07
	Código subst.	Q	H	R	B	A	K	E	P	G

Quadro 1 – Área da curva de cada substância, seus respectivos tempos de retenção e [M+H]⁺ em relação à diluição seriada em µg/ml. Para consultar os códigos, checar material e métodos deste capítulo. Em destaque (retângulos pintados diferente) o limite de detecção de cada substância relacionado à concentração. Assim, na maioria das substâncias não foi possível sua detecção a uma concentração abaixo de 2 µg/ml.

Por meio dos dados de intensidade de pico da [M+H]⁺ das nove substâncias (a catequina não ionizou no modo positivo) para a curva de calibração e suas respectivas concentrações foi possível notar que a resposta é linear, podendo-se gerar gráficos e equações para todas as substâncias, as quais permitem estimar a concentração da substância no extrato, ou mesmo na folha (Figura 33).

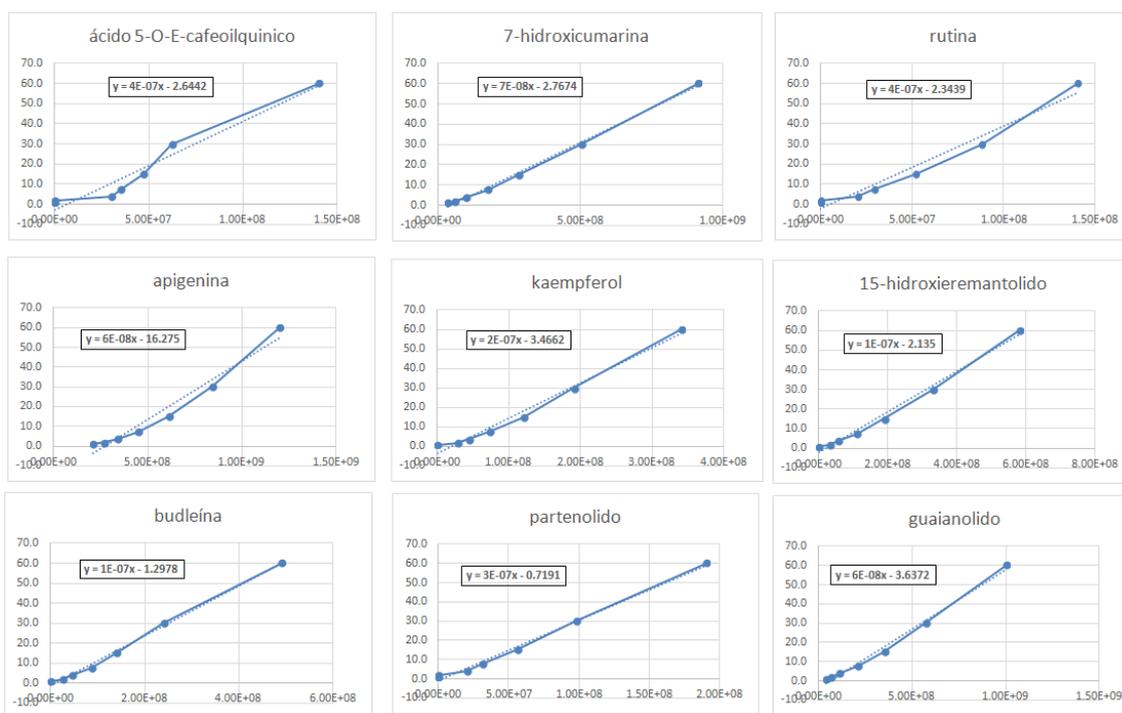


Figura 33 - Curva de calibração para 9 substâncias com base nas suas concentrações e intensidade de ionização registrada pela área do TIC.

Por meio das equações geradas para a curva de calibração das substâncias e presentes na Figura 33, a concentração de algumas substâncias puderam ser estimadas com base na inspeção da matriz de dados provenientes do processamento no MZmine, comparando-se a concentração de algumas substâncias nas diferentes estações de coleta (Quadro 2). Assim, pode-se perceber que no inverno a flavona santina, a 7-hidroxycumarina (ou ácido *p*-cumarico) e o ácido 5-*O*-*E*-cafeoilquínico estariam mais acumulados do que as demais estações. Do mesmo modo o verão estaria acumulando a substância 4-hidroxycumarina e a quercetina-3-*O*-glucoronideo nas folhas de *A. robusta*, ao passo que no outono haveria acúmulo de 3-*O*-cafeoil-2-*C*-metil-D-eritrano-1,4-lactona. A literatura também nos informa que a quercetina foi detectada em maiores concentrações em no verão, em amostras de *B. dentata* (SARTOR et al., 2013), o que corrobora a constatação de acúmulo de quercetina-3-*O*-glucoronideo nas folhas coletadas no verão em nosso estudo. Deste modo, a abordagem metabolômica aqui apresentada pode indicar diversos metabólitos correlacionados com fatores climáticos, como estação, muitos dos quais validados por bons níveis de identificação (SUMNER et al., 2007; SALEK et al., 2013), bem como por dados da literatura.

Porcentagem de substância encontradas em <i>A. robusta</i> em cada estação						
RT	4.46	5.73	7.01	7.26	7.52	15.57
compound	5-O-caffeoylquinic acid	4-hydroxycoumarin	7-hydroxycoumarin ou p-coumaric acid	3-O-Caffeoyl-2-C-methyl-D-erythrono-1,4-lactone	Quercetin-3-O-glucoronide	santin
[M+H] ⁺	355.105	163.041	163.039	325.095	301.037	345.097
Verão	0%	65%	0%	13%	56%	0%
Outono	0%	31%	0%	87%	44%	0%
Inverno	100%	3%	100%	0%	0%	100%
Primavera	0%	1%	0%	0%	0%	0%
Concentração em ug/ml por cultivar de <i>A. robusta</i>						
RT	4.46	5.73	7.01	7.26	7.52	15.57
compound	5-O-caffeoylquinic acid	4-hydroxycoumarin	7-hydroxycoumarin ou p-coumaric acid	3-O-Caffeoyl-2-C-methyl-D-erythrono-1,4-lactone	Quercetin-3-O-glucoronide	santin
[M+H] ⁺	355.105	163.041	163.039	325.095	301.037	345.097
Verão		7.50			1.51	
Outono		2.17		11.38	0.67	
Inverno	1.92		0.17			21.53
Primavera						
Concentração em mg por folha de <i>A. robusta</i> , supondo rendimento de 20 mg (droga vegetal) = 5 mg de rendimento. Supondo peso de 15 g/folha						
RT	4.46	5.73	7.01	7.26	7.52	15.57
compound	5-O-caffeoylquinic acid	4-hydroxycoumarin	7-hydroxycoumarin ou p-coumaric acid	3-O-Caffeoyl-2-C-methyl-D-erythrono-1,4-lactone	Quercetin-3-O-glucoronide	santin
[M+H] ⁺	355.105	163.041	163.039	325.095	301.037	345.097
Verão	0.00	5.63	0.00	0.00	1.13	0.00
Outono	0.00	1.63	0.00	8.53	0.50	0.00
Inverno	1.44	0.00	0.13	0.00	0.00	16.15
Primavera	0.00	0.00	0.00	0.00	0.00	0.00

Quadro 2 – Concentração estimada de substâncias químicas em *A. robusta* com base em curva de calibração.

5 – CONCLUSÃO DO CAPÍTULO 4

Os resultados deste capítulo nos permite concluir que o fator ambiental que mais pode causar equívocos na quimiotaxonomia é a sazonalidade, bem como o processamento do material vegetal (congelado ou secado em estufa). Além disso, foi apontado alguns metabólitos que podem ser considerados como marcadores estacionais, ou provenientes de degradação pelo não congelamento das amostras. Estes ultimo resultados permite melhor controle de qualidade de amostras vegetais e estudos quimiotaxonômicos.

DISCUSSÃO GERAL

Com base na abordagem metabolômica, não foi possível agrupar por PCA e HCAbp as espécies de acordo com as respectivas classificações propostas por Bremer (BREMER, 1996) e Panero (PANERO et al., 2014). Tais dados merecem ser melhor investigados a fim de se corroborar se de fato os caracteres químicos (ou a abordagem metabolômica) não poderiam ser considerados como bons marcadores de Asteraceae ao nível da classificação em tribos. Além disso, deve-se levar em consideração que fatores ambientais podem influenciar o metabolismo vegetal (GOBBO-NETO e LOPES, 2007) e, portanto, causar equívocos nas interpretações quimiotaxonômicas, assunto analisado no capítulo 4 do presente trabalho. Por outro lado, por meio da OPLS-DA foi possível diferenciar tribos de Asteraceae com base na composição química das espécies. Essa diferenciação possibilitou apontar prováveis metabólitos discriminantes das tribos Barnadesieae, Astereae, Millerieae, Vernonieae, Heliantheae e Eupatorieae, os quais podem estar relacionados com as adaptações de cada tribo em seus respectivos *habitats*, ou com a diferenciação evolutiva entre elas.

As análises de inibição de LDH-5 com extratos foliares de espécies de Asteraceae também apontaram que algumas espécies podem ser promissoras em fornecer metabólitos ativos contra tal alvo bioquímico. Assim, modelos com base em metabolômica e técnicas quimiométricas, como OPLS-DA, e *data mining*, como a *decision tree*, possibilitaram não somente classificar extratos de Asteraceae que ainda não haviam sido analisados frente a LDH-5, bem como apontar prováveis substâncias químicas altamente correlacionadas com os extratos classificados como “ativos” nos ensaios. Tais substâncias serão isoladas futuramente e terão suas atividades confirmadas. Deste modo, a metabolômica assistida por técnicas quimiométricas e *data mining* possibilitam rápida classificação de amostras ainda não ensaiadas e a identificação de metabólitos que podem ser ativos e, portanto, serem submetidos à fitoquímica direcionada, o que possibilita uma redução no tempo e custo das análises.

Os resultados de germinação (capítulo 3) evidenciaram que apenas as sementes de *A. trichophylla* possuem atividade fotoblástica, enquanto *A. robusta* não possui, germinando preferencialmente quando expostas a 12 h de luz. Além disso, a concentração e o tempo de exposição de GA₃ influenciaram positivamente a germinação de *A. robusta*, *A. trichophylla* e *T. diversifolia*. Para *A. trichophylla* e *A. robusta*, a %G e o GSI foram maiores quando as sementes foram expostas a 40 e 320 ppm de GA₃, ambas por 24 h, respectivamente. A germinação das sementes de *T. diversifolia* também foi maior em 320 ppm de GA₃ por 24 h.

No entanto, *T. diversifolia* parece ser mais sensível ao aquecimento em água a 80 °C (51,2% de germinação) do que a exposição a GA₃ (36,7% de germinação). Tais resultados se corroboram com a literatura (BOMBO et al., 2015; AGBOOLA et al., 2005) e permitem melhores estratégias para aumentar a germinação das sementes e os subsequentes cultivos. Quanto a análise de adubação e sombreamento sobre o desenvolvimento de *A. robusta*, foi verificado que, de maneira geral, a adubação e sombreamento foram capazes de gerar plantas com maior número de folhas, maior diâmetro do ramo principal e maior área foliar. Por outro lado, a adubação em canteiro sem sombrite foi capaz de produzir apenas maior número de folhas do ramo principal. Foi observado também que a floração e a altura de *A. robusta* ocorreram semelhantemente às plantas selvagens, mas mais lentamente do que quando comparado com a literatura (RUGGIERO et al., 1997). De acordo com a literatura (RUGGIERO et al., 1997), a baixa estatura da planta pode estar relacionado com o cultivo em vasos, e a floração precoce com a presença de adubação e sombreamento.

Por fim, a metabolômica aplicada ao estudo de influências ambientais, tais como ritmo circadiano, estação do ano, luminosidade e solo, sobre cultivares de *A. robusta*, foi capaz de agrupar os espécimes com base nas coletas realizadas em diferentes estações. Isso sugere que as datas das coletas devem ser padronizadas a fim de se evitar equívocos em estudos quimiotaxonômicos. Por outro lado, não houve agrupamento com base no ritmo circadiano, luminosidade e adubação do solo. De maneira geral, foi verificada tanto por PCA quanto por OPLS-DA uma dispersão aleatória entre as amostras, o que sugere pouca influência de tais fatores ambientais no metaboloma de *A. robusta* e, possivelmente, em estudos quimiotaxonômicos. Por outro lado, a OPLS-DA foi capaz de apontar metabólitos discriminantes das estações de inverno, primavera e outono conjuntamente com verão, os quais podem ser considerados marcadores ambientais de Asteraceae. Tais resultados encontram respaldo na literatura (SARTOR et al., 2013). Os resultados mostram que também foi possível separar coletas de folhas secadas a 40 °C das folhas congeladas e secadas em liofilizador, apontando seus prováveis metabólitos discriminantes. Esses metabólitos também são confirmados pela literatura (MOREIRA et al., 2018). Esse resultado também sugere que se deve padronizar a coletas em N₂ líquido, especialmente quando se objetiva um maior controle de qualidade em plantas, ou quando se analisa fatores ambientais nas plantas e sua quimiotaxonomia.

CONCLUSÕES

O presente trabalho objetivou utilizar a metabolômica e técnicas quimiométricas para auxiliar na compreensão da quimiotaxonomia de Asteraceae, ao nível da classificação de tribos, realizar a triagem em extratos de metabólitos ativos na inibição de LDH-5, e analisar a influência de determinados fatores ambientais no metaboloma de plantas. De maneira geral, podemos concluir que as abordagens metabolômicas apresentadas neste trabalho foram capazes de:

- a) apontar metabólitos discriminantes de cada tribos, os quais podem ser considerados marcadores quimiotaxonômicos (capítulo 1);
- b) prever a atividade de extratos de outras espécies de Asteraceae ainda não ensaiados, bem como apontar prováveis substâncias correlacionadas com extratos ativos (capítulo 2), o que possibilita menor tempo e custo de análise, assim como o isolamento direcionado dos metabólitos;
- c) apontar marcadores químicos reacionados com estações do ano, o que permite um melhor controle de qualidade de amostras vegetais, bem como advertir para uma padronização das datas das coletas e secagem de folhas que antecedem os estudos quimiotaxonômicos (capítulo 3).

Além disso, o presente trabalho também traz informações de condições ideais de germinação e influência da adubação e sombreamento sobre a morfologia de *A. robusta* (e *A. trichophylla* e *T. diversifolia*, para o caso da germinação), o que permite um melhor delineamento experimental que vise a obtenção de uma maior quantidade de cultivares, bem como de um maior número de folhas e maiores ramos por planta.

REFERÊNCIAS

- Agati, G., Azzarello, E., Pollastri, S., Tattini, M. Flavonoids as antioxidants in plants: location and functional significance. *Plant science*. 2012, 196, 67-76.
- Agboola, D. A., Idowu, W. F., Kadiri, M. Seed germination and seedling growth of the Mexican sunflower *Tithonia diversifolia* (Compositae) in Nigeria, Africa. *Revista de biología tropical*, 2006. 395-402.
- Ajao, A. A., Moteetee, A. N. *Tithonia diversifolia* (Hemsl) A. Gray.(Asteraceae: Heliantheae), an invasive plant of significant ethnopharmacological importance: A review. *South African Journal of Botany*. 2017, 113, 396-403.
- Alcantara, G. B., Honda, N. K., Ferreira, M. M. C., Ferreira, A. Chemometric analysis applied in ¹H HR-MAS NMR and FT-IR data for chemotaxonomic distinction of intact lichen samples. *Analytica Chimica Acta*, 2007, 595, 3–8
- Allard, P. M., Péresse, T., Bisson, J., Gindro, K., Marcourt, L., Pham, V. C., ... Wolfender, J. L. Integration of molecular networking and in-silico MS/MS fragmentation for natural products dereplication. *Analytical chemistry*. 2016, 88, 3317-3323.
- Alvarenga, S. A. V., Ferreira, M. J. P., Emerenciano, V. P., Cabrol-Bass, D. Chemosystematic studies of natural compounds isolated from Asteraceae: characterization of tribes by principal component analysis. *Chemometric and Intelligent Laboratory Systems*. 2001, 56, 27-37.
- Ambrósio, S. R., Arakawa, N. S., Esperandim, V. R., de Albuquerque, S., Da Costa, F. B. Trypanocidal activity of pimarane diterpenes from *Viguiera arenaria* (Asteraceae). *Phytotherapy Research: An International Journal Devoted to Pharmacological and Toxicological Evaluation of Natural Product Derivatives*. 2008, 22, 1413-1415.
- Arakawa, N. S., Schorr, K., Ambrosio, S. R., Merfort, I., Da Costa, F. B. Further sesquiterpene lactones from *Viguiera robusta* and the potential anti-inflammatory activity of a heliangolide: inhibition of human neutrophil elastase release. *Zeitschrift für Naturforschung C*. 2008, 63, 533-538.
- Augoff, K., Hryniewicz-Jankowska, A., Tabola, R. Lactate dehydrogenase 5: An old friend and a new hope in the war on cancer. *Cancer Letters*. 2015, 358, 1-7.
- Barbosa, A. R., Yamamoto, K., Valio, I. F. M. Effect of light and temperature on germination and early growth of *Vochysia tucanorum* Mart., Vochysiaceae, in cerrado and forest soil under different radiation levels. *Revista Brasileira de Botânica*. 1999, 22, 275-280,.
- Baskin, C. C., Baskin, J. M. Role of temperature and light in the germination ecology of buried seeds of weedy species of disturbed forests. II. *Erechtites hieracifolia*. *Canadian Journal of Botany*. 1996, 74, 2002-2005.
- Beaman, J. H. *Plant Taxonomy*. *Clin. Dermatol.*, 1986, 4, 23-30.
- Bhargava, N., Sharma, G., Bhargava, R., Mathuria, M. Decision tree analysis on J48 algorithm for data mining. 2013, 3, 1114-1119.
- Blaženović, I., Kind, T., Sa, M. R., Ji, J., Vaniya, A., Wancewicz, B., ... , Showalter, M. R. Structure annotation of all mass spectra in untargeted metabolomics. *Analytical chemistry*. 2019, 91, 2155-2162.
- Bohlmann, F., Zdero, C., King, R. M., Robinson, H. Hirsutinolides and other sesquiterpene lactones from *Vernonia* species. *Phytochemistry*. 1982, 21, 695-699.

Bohm, B. A., Stuessy, T. F. Flavonoids of the sunflower family (Asteraceae). Springer Science Business Media. 2001

Bombo, A. B., Oliveira, T. S. D., Appezzato-da-Glória, B., Novembre, A. D. D. L. C. Seed germination of Brazilian *Aldama* species (Asteraceae). Journal of Seed Science, (AHEAD), 2015.

Bremer, K. Major clades and grades of the Asteraceae. In: Hind DN, Beebtj HJ (es) Compositae: systematic. Proceedings of the international Compositae conference, Kew. Hind DJN (Editor-in-Chief), v 1, Royal Botanic Gardens, Kew, 1994, 1-7.

Buydens, L.M.C., Reijmers, T.H., Beckers, M.L.M., Wehrens, R. Molecular data-mining: a challenger for chemometrics. Chemometrics and Intelligent Laboratory Systems. 1999, 49, 121-133.

Calabria, L. M., Emerenciano, V. P., Ferreira, M. J., SCotti, M. T., Mabry, T. J. A phylogenetic analysis of tribes of the Asteraceae based on phytochemical data. Natural Product Communications. 2007, 2, 277-285.

Calabria, L. M., Emerenciano, V. P., Ferreira, M. J. P. et al. A phylogenetic analysis of tribes of the Asteraceae based on phytochemical data. Natural Products Communications. 2007, 2, 277-285.

Campos, F., R., Bressan, J., Godoy Jasinski, V. C., Zuccolotto, T., da Silva, L. E., Bonancio Cerqueira, L. *Baccharis* (Asteraceae): chemical constituents and biological activities. Chemistry biodiversity. 2016, 13, 1-17.

Canales, M. et al. Antimicrobial Activity of the Extracts and Essential Oil of *Viguiera dentata*. Pharmaceutical Biology. 2008, 46, 719–723.

Chagas-Paula, D. A., de Oliveira, R. B., da Silva, V. C., Gobbo-Neto, L., Gasparoto, T. H., Campanelli, A. P., ... Da Costa, F. B. Chlorogenic acids from *Tithonia diversifolia* demonstrate better anti-inflammatory effect than indomethacin and its sesquiterpene lactones. Journal of Ethnopharmacology. 2011, 136, 355-362.

Chagas-Paula, D. A., Oliveira, R. B., Rocha, B. A., Da Costa, F. B. Ethnobotany, chemistry, and biological activities of the genus *Tithonia* (Asteraceae). Chemistry biodiversity. 2012, 9, 210-235.

Chagas-Paula, D. A., Oliveira, T. B., Zhang, T., Edrada-Ebel, R., Da Costa, F. B. Prediction of anti-inflammatory plants and discovery of their biomarkers by machine learning algorithms and metabolomic studies. Planta medica. 2015, 81, 450-458.

Chagas-Paula, D., Zhang, T., Da Costa, F., Edrada-Ebel, R. A metabolomic approach to target compounds from the Asteraceae family for dual COX and LOX inhibition. Metabolites. 2015, 5, 404-430.

Çirak, C., Radušienė, J., Ivanauskas, L., Janulis, V. Variation of bioactive secondary metabolites in *Hypericum origanifolium* during its phenological cycle. Acta Physiol. Plant. 2007, 29, 197–203.

Correia, M. V., Fokou, H. H., Marcelo, J. et al., Self-organizing maps a good tool for classification of subfamily Asteroideae. Journal of Medicinal Plants Research. 2012, 6, 1207-1218.

Correia, M. V., Scotti, M. T., Ferreira, M. J. P., et al. Self-organizing maps as a new tool for classification of plants at lower hierarchical levels. Natural Products Communications. 2008, 3, 1723-1730.

Da Costa, F. B., Terflloth, L., Gasteiger, J. Sesquiterpene lactone-based classification of three Asteraceae tribes: a study based on self-organizing neural networks applied to chemo systematics. Phytochemistry. 2005, 66, 345-353.

Dantas, B. F., Alves, E., Aragão, C. A., Tofanelli, B. D., Corrêa, M. R., Rodrigues, J. D., ... Nakagawa, J. Germinação de sementes de capim-marmelada (*Brachiaria plantaginea* (link) hitchc.) tratadas com ácido giberélico. *Revista Brasileira de Sementes*. 2001, 23, 27-34.

Darwin, C. *On the Origin of Species by Means of Natural Selection* (J. Murray, London). 1859.

De Toledo, J., Ambrósio, S., Borges, C., Manfrim, V., Cerri, D., Cruz, A., Da Costa, F. In vitro leishmanicidal activities of sesquiterpene lactones from *Tithonia diversifolia* against *Leishmania braziliensis* promastigotes and amastigotes. *Molecules*. 2014, 19, 6070-6079.

De Vos, R. C., Moco, S., Lommen, A., Keurentjes, J. J., Bino, R. J., & Hall, R. D. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. *Nature Protocols*, 2007, 2, 778-791.

Deiab, S., Mazzi, E., Messeha, S., Mack, N., & Soliman, K. F. A. High-throughput screening to identify plant derived human LDH-A inhibitors. *European journal of medicinal plants*. 2013, 3, 603.

Delgado, G., de Vivar, A. R., Herz, W. Sesquiterpene lactones from *Viguiera* species. *Phytochemistry*. 1982, 21, 1305-1308.

El Marsni, Z., Torres, A., Varela, R. M., Molinillo, J. M., Casas, L., Mantell, C., ... Macias, F. A. Isolation of bioactive compounds from sunflower leaves (*Helianthus annuus* L.) extracted with supercritical carbon dioxide. *Journal of agricultural and food chemistry*. 2015, 63, 6410-6421.

Emerenciano, V. D. P., Ferreira, Z. S., Kaplan, M. A. C., Gottlieb, O. R. A chemosystematic analysis of tribes of Asteraceae involving sesquiterpene lactones and flavonoids. *Phytochemistry*. 1987, 26, 3103-3115.

Emerenciano, V. P., Cabrol-Bass, D., Ferreira, M. J. P. et al. Chemical evolution in the Asteraceae. The oxidation-reduction mechanism and production of secondary metabolites. *Nat Prod Comm*. 2006, 1, 495-507.

Scotti, M. T., Emerenciano, V., Ferreira, M. J. P. et al. Self-organizing maps of molecular descriptors for sesquiterpene lactones and their application to the chemotaxonomy of the Asteraceae family. *Molecules*. 2012, 17, 468-4702.

Eugster, P.J., Boccard, J., Debrus, B., Bréant, L., Wolfender, J-L., Martel, S., Carrupt, P-A. Retention time prediction for dereplication of natural products (C_xH_yO_z) in LC-MS metabolite profiling. *Phytochemistry*. 2014, 108, 196-207.

Fariman, Z. K., Azizi, M., Noori, S. Seed germination and dormancy breaking techniques for *Echinacea purpurea* L. *Journal of Biology Environmental Science* 2011, 5, 7-10.

Funk, V. A., Susanna, A., Stuessy, T. F., Bayer, R. J. Systematics, evolution, and biogeography of Compositae. *International Association for Plant Taxonomy*. Vienna, 2009.

Gallon, M. E., Monge, M., Casoti, R., Da Costa, F. B., Semir, J., Gobbo-Neto, L. Metabolomic analysis applied to chemosystematics and evolution of megadiverse Brazilian *Vernoniaeae* (Asteraceae). *Phytochemistry*. 2018, 150, 93-105.

Gasteiger J. Engel, T. (Eds.), *Chemoinformatics*. WILEY-VCH Verlag GmbH Co. KGaA, Weinheim, 2003.

Gobbo-Neto, L., Lopes, N. P. Plantas medicinais: fatores de influência no conteúdo de metabólitos secundários. *Química nova*. 2007, 30, 374.

Goffin, E., Ziemons, E., De Mol, P., de Madureira, M. D. C., Martins, A. P., da Cunha, A. P., ... Frederich, M. In vitro antiplasmodial activity of *Tithonia diversifolia* and identification of its main active constituent: tagitinin C. *Planta medica*. 2002, 68, 543-545.

Granchi, C., Roy, S., Giacomelli, C., Macchia, M., Tuccinardi, T., Martinelli, A., ... & Funel, N. Discovery of N-hydroxyindole-based inhibitors of human lactate dehydrogenase isoform A (LDH-A) as starvation agents against cancer cells. *Journal of medicinal chemistry*. 2011, 54, 1599-1612.

Guo, S., Ge, Y., Jom, K. N. A review of phytochemistry, metabolite changes, and medicinal uses of the common sunflower seed and sprouts (*Helianthus annuus* L.). *Chemistry Central Journal*. 2017, 11, 95.

Hall, M., Frak, E., Holmes, G., Pfahinger, B., Reutemann, P., Witten, I.H. The weka data mining software: An update. 2009, 11, 10-18.

Harmer, S. L. The circadian system in higher plants. *Annual Review of Plant Biology*. 2009, 60, 357-377.

Hegnauer, R. Chemical characters in plant taxonomy: some possibilities and limitations. *Pure and Applied Chemistry*. 1967, 14, 173-188.

Hooley, R. Gibberellins: perception, transduction and responses. In *Signals and Signal Transduction Pathways in Plants*. 1994, 293-319. Springer, Dordrecht.

Hubert, J., Nuzillard, J. M., & Renault, J. H. Dereplication strategies in natural product research: How many tools and methodologies behind the same concept? *Phytochemistry Reviews*. 2017, 16: 55-95.

INCA (2017) Mortes por câncer aumentaram 31% no Brasil em 15 anos. Disponível em: <http://www2.inca.gov.br/wps/wcm/connect/agencianoticias/site/home/noticias/2017/mortes_por_cancer_aumentaram_31_no_brasil_em_15_anos_diz_oms>. Acessado em 20/fevereiro/2019.

Isejima, E. M., Figueiredo-Rrieiro, R. C. L., Zaidan, L. B. P. Fructan composition in adventitious tuberous roots of *Viguiera discolor* Baker (Asteraceae) as influenced by daylength. *New Phytologist*. 1991, 119, 149-154.

Panero, J. L., Freire, S. E., Espinar, L. A., Crozier, B. S., Crozier, G. E., Barboza, J. J. C. Resolution of deep nodes yields an improved backbone phylogeny and a new basal lineage to study early evolution of Asteraceae. *Molecular Phylogenetics and Evolution*. 201, 80, 43-53

Kanani, M. R., Rahiminejad, M. R., Sonboli, A., Mozaffarian, V., Osaloo, S. K., Ebrahimi, S. N. Chemotaxonomic significance of the essential oils of 18 *Ferula* species (Apiaceae) from Iran. *Chemistry & Biodiversity*, 2011, 8, 503-517.

Kaneko, M., Itoh, H., Ueguchi-Tanaka, M., Ashikari, M., Matsuoka, M. The α -amylase induction in endosperm during rice seed germination is caused by gibberellin synthesized in epithelium. *Plant Physiology*. 2002, 128, 1264-1270.

Katajamaa, M., Miettinen, J., Oresic, M. MZmine: toolbox for processing and visualization of mass spectrometry based molecular profile data. *Bioinformatics*. 2006, 22, 634-636.

Keles, L. C., Melo, N. I. D., Aguiar, G. D. P., Wakabayashi, K. A. L., Carvalho, C. E. D., Cunha, W. R., ... Lopes, N. P. Lychnophorinae (Asteraceae): a survey of its chemical constituents and biological activities. *Quimica Nova*. 2010, 33, 2245-2260.

Kessner, D., Chambers, M., Burke, R., Agus, D., Mallick, P. ProteoWizard: open source software for rapid proteomics tools development. *Bioinformatics*. 2008, 24, 2534-2536.

Kim, H.K., Khan, S., Wilson, E.G., Kricun, S.D.P., Meissner, A., Goral, S., Deelder, A.M., Choi, Y.H., Verpoorte, R. Metabolic classification of South American *Ilex* species by NMR-based metabolomics. *Phytochemistry*. 2010, 71, 773–784.

Kirmizi, S., Guleryuz, G., Arslan, H. Germination responses to GA3 and short-time chilling of three endemic species: *Tripleurospermum picheleri*, *Cirsium leucopis* and *Senecio olympicus* (Asteraceae). *Plant Species Biology*. 2011, 26, 51-57.

Kos, O., Castro, V., Murillo, R., Poveda, L., Merfort, I. Ent-kaurane glycosides and sesquiterpene lactones of the hirsutinolide type from *Vernonia triflosculosa*. *Phytochemistry*. 2006, 67, 62-69.

Krastanov A. Metabolomics – The state of art. *Biotechnol. Biotechnol. Eq.* 2010, 1, 1537-1543.

Lankadurai, B. P., Nagato, E. G., Simpson, M. J. Environmental metabolomics: an emerging approach to study organism responses to environmental stressors. *Environmental Reviews*. 2013, 21, 180-205.

Larose, D. T. (2006). *Data Mining – Methods and Models*. Central Connecticut State University.

Lee, S., Oh, D. G., Lee, S., Kim, G., Lee, J., Son, Y., ... Lee, C. Chemotaxonomic metabolite profiling of 62 indigenous plant species and its correlation with bioactivities. *Molecules*. 2015, 20, 19719-19734.

Lommen, A. Kools, H.J. MetAlign 3.0: performance enhancement by efficient use of advances in computer hardware. *Metabolomics*. 2012, 8, 719-726.

Lommen, A. MetAlign: Interface-Driven, versatile metabolomics tool for hyphenated full-scan mass spectrometry data preprocessing. *Analytical Chemistry*. 2009, 81, 3079-3086.

Magenta, M.A.G. *Viguiera* Kunth (Asteraceae , Heliantheae) na América do Sul e sistemática das espécies do Brasil. Tese de Doutorado. 339. Universidade de São Paulo, 2006.

Maguire, J. D. Speed of Germination—Aid In Selection And Evaluation for Seedling Emergence And Vigor 1. *Crop science*. 1962, 2, 176-177.

Manerba, M., Di Ianni, L., Govoni, M., Roberti, M., Recanatini, M., Di Stefano, G. Lactate dehydrogenase inhibitors can reverse inflammation induced changes in colon cancer cells. *European Journal of Pharmaceutical Sciences*. 2017, 96, 37-44.

Marquina, S., Maldonado, N., Garduño-Ramírez, M. L., Aranda, E., Villarreal, M. L., Navarro, V., ... Alvarez, L. Bioactive oleanolic acid saponins and other constituents from the roots of *Viguiera decurrens*. *Phytochemistry*. 2001, 56, 93-97.

Martucci, M. E. P., Gobbo-Neto, L. Differential secondary metabolite accumulation and performance of *Chlosyne lacinia* fed with *Tithonia diversifolia* or *Vernonia polyanthes*. *Biochemical systematics and ecology*. 2016, 68, 156-162.

Martucci, M. E. P., De Vos, R. C., Carollo, C. A., Gobbo-Neto, L. Metabolomics as a potential chemotaxonomical tool: application in the genus *Vernonia* Schreb. *PLoS One*. 2014, 9, e93149.

McDonald, M. B., Khan, A. A. Acid Scarification and protein synthesis during Seed Germination 1. *Agronomy journal*. 1983, 75, 111-114.

Moreira, E. A., Pilon, A. C., Andrade, L. E., Lopes, N. P. New Perspectives on Chlorogenic acid accumulation in harvested leaf tissue: impact on traditional medicine Preparations. *ACS Omega*. 2018, 3, 18380-18386.

Mutihac, L. Mutihac, R. Mining in chemometrics. *Analytica Chimica ACTA*. 2008, 612, 1-18.

Napimoga, M. H., Yatsuda, R. Scientific evidence for *Mikania laevigata* and *Mikania glomerata* as a pharmacological tool. *Journal of Pharmacy and Pharmacology*. 2010, 62, 809-820.

Naz, S., Vallejo, M., García, A., Barbas, C. Method validation strategies involved in non-targeted metabolomics. *Journal of Chromatography A*. 2014, 1353, 99-105.

Nicolete, R. et al. Budlein A from *Viguiera robusta* inhibits leukocyte-endothelial cell interactions, adhesion molecule expression and inflammatory mediators release. *Phytomedicine*. 2009, 16, 904–915.

Nogueira, M., Da Costa, F., Brun, R., Kaiser, M., Schmidt, T. ent-pimarane and ent-kaurane diterpenes from *Aldama discolor* (Asteraceae) and their antiprotozoal activity. *Molecules*. 2016, 21, 1237.

Oliveira, A. B., Oliveira, T. S., Glória, B. A., Novembre, A. D. L. C. Seed germination of Brazilian *Aldama* species (Asteraceae). *Journal of Seed Science*. 2015, 37, 185-191.

Oliveira, T. B., Gobbo-Neto, L., Schmidt, T. J., Da Costa, F. B. Study of chromatography retention of natural terpenoids by cheminformatic tools. *Journal of Chemical Information and Modeling*. 2015, 55, 26-38, 2015.

ONU BRASIL (2017) OMS: câncer mata 8,8 milhões de pessoas anualmente no mundo. Disponível em: < <https://nacoesunidas.org/oms-cancer-mata-88-milhoes-de-pessoas-anualmente-no-mundo/>>. Acessado em 20/fevereiro/2019.

Padilla-González, G. F., Diazgranados, M., Da Costa, F. B. Biogeography shaped the metabolome of the genus *Espeletia*: a phytochemical perspective on an Andean adaptive radiation. *Scientific reports*. 2017, 7, 8835.

Padilla-González, G.F., Aldana, J.A., Da Costa, F.B. Chemical characterization of two morphologically related *Espeletia* (Asteraceae) species and chemometric analysis based on essential oil components. *Revista Brasileira de Farmacognosia*. 2016, 1–7.

Patel, R. G., Mankad, A. U. Effect of gibberellins on seed germination of *Tithonia rotundifolia* Blake. *International Journal of Innovative Research in Science, Engineering and Technology*. 2014, 3, 10680-10684.

Pavarini, D., Bauermeister, A., Semir, J., Monge, M., Lopes, J., Pepporine, N. P. (2018) Phenolic Compounds from the Brazilian genus *Lychnophora* Mart. (Asteraceae). Chapter in ACS Symposium Series · January 2018

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, 11, 395–406.

Porto, T. S., Furtado, N. A., Heleno, V. C., Martins, C. H., Da Costa, F. B., Severiano, M. E., ... Ambrosio, S. R. Antimicrobial ent-pimarane diterpenes from *Viguiera arenaria* against Gram-positive bacteria. *Fitoterapia*. 2009, 80, 432-436.

Quinlan, J. R. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers. 1993.

Rencher, A. C. (2002). *Methods of Multivariate Analysis*. Bringham Young University, 2nd Ed.

- Reynolds, T. The evolution of chemosystematics. *Phytochemistry*. 2007, 68, 22-24.
- Rivas-ubach, A., Sardans, J., Pérez-trujillo, M., Estiarte, M., Peñuelas, J. Strong relationship between elemental stoichiometry and metabolome in plants. *Proceedings of the National Academy of Sciences*. 2012, 109, 4181–4186.
- Rosa, A. L. Aplicação de ferramentas computacionais e analíticas na construção de bibliotecas de produtos naturais da família Asteraceae. Tese de Doutorado. Universidade de São Paulo, 2018.
- Rozema, J., Van de Staaij, J., Björn, L. O., Caldwell, M. UV-B as an environmental factor in plant life: stress and regulation. *Trends in Ecology and Evolution*. 1997, 12, 22–28.
- Rufatto, L. C., Gower, A., Schwambach, J., Moura, S. Genus *Mikania*: chemical composition and phytotherapeutical activity. *Revista Brasileira de Farmacognosia*. 2012, 22, 1384-1403.
- Ruggiero, P. G., Zaidan, L. B. Estudo de desenvolvimento de *Viguiera robusta* Gardn., uma Asteraceae do cerrado. *Brazilian Journal of Botany*. 1997, 20, 01-09.
- Salek, R. M., Steinbeck, C., Viant, M. R., Goodacre, R., Dunn, W. B. The role of reporting standards for metabolite annotation and identification in metabolomic studies. *GigaScience*. 2013, 2, 13.
- Santos, F. A. D., Sousa, I. P., Furtado, N. A., Costa, F. B. D. Combined OPLS-DA and decision tree as a strategy to identify antimicrobial biomarkers of volatile oils analyzed by gas chromatography–mass spectrometry. *Revista Brasileira de Farmacognosia*. 2018, 28, 647-653.
- Sartor, T., Xavier, V. B., Falcão, M. A., Mondin, C. A., Dos Santos, M. A., Cassel, E., ... Santarém, E. R. Seasonal changes in phenolic compounds and in the biological activities of *Baccharis dentata* (Vell.) GM Barroso. *Industrial crops and products*. 2013, 51, 355-359.
- Schilling, E. E. Panero, J. L. A revised classification of subtribe Helianthinae (Asteraceae: Heliantheae) II. Derived lineages. *Botanical Journal of the Linnean Society*. 2011, 167, 311–331.
- Schorr, K., García-Piñeres, A.J., Siedle, B., Merfort, I., Da Costa, F.B. Guaianolides from *Viguiera gardneri* inhibit the transcription factor NF-kappaB. *Phytochemistry*. 2002, 60, 733–40.
- Schweiger, R., Heise, A., Persicke, M., Müller, C. Interactions between the jasmonic and salicylic acid pathway modulate the plant metabolome and affect herbivores of different feeding types. *Plant Cell and Environment*. 2014, 37, 1574–1585.
- Scognamiglio, M., D’Abrosca, B., Esposito, A., Fiorentino, A. Chemical composition and seasonality of aromatic Mediterranean plant species by NMR-based metabolomics. *Journal of Analytical Methods in Chemistry*. 2015, 1–9.
- Scognamiglio, M., Fiumano, V., D’Abrosca, B., Esposito, A., Choi, Y. H., Verpoorte, R., Fiorentino, A. Chemical interactions between plants in Mediterranean vegetation: The influence of selected plant extracts on *Aegilops geniculata* metabolome. *Phytochemistry*. 2014, 106, 69–85.
- Seaman, F. C. Sesquiterpene lactones as taxonomic characters in the Asteraceae. *The Botanical Review*. 1982. 48, 121-595.
- Sneath, P. H., Sokal, R. R. Numerical taxonomy. The principles and practice of numerical classification. 1973.
- Sokal, R. R., Sneath, P. H. Principles of numerical taxonomy. Principles of numerical taxonomy. 1963
- Spring, O. Schilling, E.E. The sesquiterpene lactone chemistry of *Helianthus* sect. *Atrorubentes* (Asteraceae: Heliantheae). *Biochemical Systematics and Ecology*. 1991, 19, 59–79.

Spring, O. Schillingt, E.E. Chemosystematic investigation of the annual species of *Helianthus* (Asteraceae)*. *Biochemical Systematics and Ecology*. 1989, 17, 519–528.

Spring, O. Schillingt, E.E. The sesquiterpene lactone chemistry of *Helianthus* Sect. *Ciliares* (Asteraceae: Heliantheae)*. *Biochemical Systematics and Ecology*. 1990, 18, 139–143.

Sumner, L.W., Amberg, A., Barrett, D., Beale, M.H., Beger, R., Daykin, C.A., Fan, T.W., Fiehn, O., Goodacre, R., Griffin, J.L., Hankemeier, T., Hardy, N., Harnly, J., Higashi, R. Proposed minimum reporting standards for chemical analysis. *Metabolomics*. 2007, 3, 211–221.

Suzuki, R., Shimodaira, H. Pvcust: an R package for assessing the uncertainty in hierarchical clustering. *Bioinformatics*. 2006, 22, 1540–1542.

Trygg, J., Wold, S. Orthogonal projections to latent structures (O-PLS). *Journal of Chemometrics*. 2002, 16, 119–128.

Uehara, A., Nakata, M., Uchida, A., Iwashina, T. Chemotaxonomic consideration of flavonoids from the leaves of *Chrysanthemum arcticum* subsp. *arcticum* and *yezoense*, and related species. *Biochemical Systematics and Ecology*. 2017, 73, 11–15.

Valério, D. A., Cunha, T. M., Arakawa, N. S., Lemos, H. P., Da Costa, F. B., Parada, C. A., ... Verri Jr, W. A. Anti-inflammatory and analgesic effects of the sesquiterpene lactone budlein A in mice: inhibition of cytokine production-dependent mechanism. *European journal of pharmacology*. 2007, 562, 155–163.

Van den Berg, R. A., Hoefsloot, H. C., Westerhuis, J. A., Smilde, A. K., Van der Werf, M. J. Centering, scaling, and transformations: improving the biological information content of metabolomics data. *BMC genomics*. 2006, 7, 142.

Verdi, L.G., Brighente, I.M.C., Pizzolatti, M.G. Gênero *Baccharis* (Asteraceae): aspectos químicos, econômicos e biológicos. *Química Nova*. 2005, 28, 85–94.

Villas-Bôas, S.G., Rasmussen, S., Lane, G.A. Metabolomics or metabolite profiles? Trends in biotechnology. 2005, 23, 385–386.

Wang, M., Carver, J. J., Phelan, V. V., Sanchez, L. M., Garg, N., Peng, Y., ... Porto, C. Sharing and community curation of mass spectrometry data with GNPS. *Nature biotechnology*. 2016, 34, 828.

Wang, G. L., Semenza, G. L. General involvement of hypoxia-inducible factor 1 in transcriptional response to hypoxia. *Proceedings of the National Academy of Sciences*. 1993, 90, 4304–4308.

Warburg, O., *Über den Stoffwechsel der Tumore*. 1926. Springer, Berlin. Translated: *The metabolism of tumors*, 1930. Arnold Constable, London.

Warburg, O. On the origin of cancer cells. *Science*. 1956, 123: 309–314.

Ward, P. S., Thompson, C. B. Metabolic reprogramming: a cancer hallmark even warburg did not anticipate. *Cancer cell*. 2012, 21, 297–308.

WHO – World Health Organization. Cancer. Disponível em < <http://www.who.int/news-room/fact-sheets/detail/cancer>>. Acessado em 19/fevereiro/2019.

Wink, M. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. *Phytochemistry*. 2003, 64, 3–19.

Witten, I. H., Eibe, F., Hall, M. A. (2011). *Data Mining – practical machine learning tools and techniques*. 3rd Ed., Morgan Kaufmann Publishers.

Wold, S., Sjöström, M., Eriksson, L. Partial least squares projections to latent structures (PLS) in chemistry. *Encyclopedia of computational chemistry*. 2002, 3.

Xie, H., Hanai, J. I., Ren, J. G., Kats, L., Burgess, K., Bhargava, P., ... Wang, X. Targeting lactate dehydrogenase-a inhibits tumorigenesis and tumor progression in mouse models of lung cancer and impacts tumor-initiating cells. *Cell metabolism*. 2014, 19, 795-809.

Yang, J. Y., Sanchez, L. M., Rath, C. M., Liu, X., Boudreau, P. D., Bruns, N., ... & Wong, W. R. Molecular networking as a dereplication strategy. *Journal of natural products*. 2013, 76, 1686-1699.

Ward Jr, J. H. Hierarchical grouping to optimize an objective function. *Journal of the American statistical association*, 1963, 58, 236-244.

Zarpon, A. C., Fattori, V., Souto, F. O., Pinto, L. G., Pinho-Ribeiro, F. A., Ruiz-Miyazawa, K. W., ... Casagrande, R. The sesquiterpene lactone, budlein A, Inhibits antigen-induced arthritis in mice: Role of NF- κ B and Cytokines. *Inflammation*. 2017, 40, 2020-2032.

Zdero, C., Bohlmann, F. Systematics and evolution within the Compositae, seen with the eyes of a chemist. *Plant Systematics and evolution*, 1990, 171, 1-14.

APÊNDICES

Apêndice A – Espécies da família Asteraceae, suas respectivas tribos, códigos/números de identificação e réplicas de análise.

Código 1 mzXML	Código 2 mzXML	Tribo	Nome da espécie	Código 1 mzXML	Código 2 mzXML	Tribo	Nome da espécie
1	Ach1	Anthemideae	Achillea millefolium	125	Esp41	Millerieae	Espeletia semiglobulata
1,2	Ach1.2	Anthemideae	Achillea millefolium	126	Gor1	Vernonieae	Gorceixia decurrens
1,3	Ach1.3	Anthemideae	Achillea millefolium	127	Het1	Vernonieae	Heterocoma ekmaniana
2	Acy1	Gnaphalieae	Achyrocline satureioides	128	Lyc2	Vernonieae	Lychnophora ericoides
3	Acm1	Heliantheae	Acmella oleracea	129	Lyc5	Vernonieae	Lychnophora markgravii
4	Age1	Eupatorieae	Ageratum conyzoides	130	Lyc7	Vernonieae	Lychnophora passerina
5	Ald3	Heliantheae	Aldama arenaria	131	Lyc10	Vernonieae	Lychnophora tomentosa
7	Ald9	Heliantheae	Aldama filifolia	132	Mik8	Eupatorieae	Mikania aff. burchelli
8	Ald8	Heliantheae	Aldama gardneri	133	Mik12	Eupatorieae	Mikania angularis
9	Ald1	Heliantheae	Aldama robusta	134	Mik13	Eupatorieae	Mikania glomerata
10	Ant1	Vernonieae	Anteremanthus hatschbachii	135	Mik14	Eupatorieae	Mikania hirsutissima
11	Arn1	Heliantheae	Arnica montana	136	Mik15	Eupatorieae	Mikania laevigata
12	Art1	Anthemideae	Artemisia absinthium	137	Mik6	Eupatorieae	Mikania psilostachya
13	Art2	Anthemideae	Artemisia annua	138	Mik3	Eupatorieae	Mikania sp
14	Bac8	Astereae	Baccharis albilanosa	139	Mik4	Eupatorieae	Mikania thapsoides
14,2	Bac8.2	Astereae	Baccharis albilanosa	140	Mik1	Eupatorieae	Mikania triangularis
14,3	Bac8.3	Astereae	Baccharis albilanosa	141	Min1	Vernonieae	Minasia scapigera
15	Bac29	Astereae	Baccharis glaziovii	142	Par2	Vernonieae	Paralychnophora bicolor
16	Bac48	Astereae	Baccharis organensis	142,2	Par.2	Vernonieae	Paralychnophora bicolor
17	Bid1	Coreopsidae	Bidens pilosa	142,3	Par.3	Vernonieae	Paralychnophora bicolor
18	Bla1	Vernonieae	Blanchetia heterotricha	143	Pip2	Vernonieae	Piptolepis schultiziana
19	Cal1	Neurolaeneae	Calea cuneifolia	144	Pre1	Vernonieae	Prestelia eriopus
20	Chr3c	Vernonieae	Chresta exsucca	145	Pro1	Vernonieae	Proteopsis argentea
21	Chr5a	Vernonieae	Chresta hatschbachii	146	Pro2	Vernonieae	Proteopsis hermogenesii
22	Chr6c	Vernonieae	Chresta martii	147	Ver4	Vernonieae	Solidago chilensis
23	Chr13a	Vernonieae	Chresta plantaginifolia	148	Tit1	Heliantheae	Tithonia diversifolia
23,2	Chr13a.2	Vernonieae	Chresta plantaginifolia	149	Ver12	Vernonieae	Vernonanthura brasiliiana
24	Chr7b	Vernonieae	Chresta pycnocephala	150	Ver1	Vernonieae	Vernonia platensis

25	Cho1	Vernonieae	Chronopappus bifrons	151	Bac2	Astereae	Baccharis aff altimontana
26	Chu2	Barnadesieae	Chuquiraga spinosa Less.	152	Bac3	Astereae	Baccharis aff boliviensis
27	Cic1	Cichorieae	Cichorium intybus L.	153	Bac4	Astereae	Baccharis aff curitybensis
28	Cyn1	Cynareae	Cynara scolymus L.	154	Bac5	Astereae	Baccharis aff leucopappa
29	Das1	Barnadesieae	Dasyphyllum brasiliensis	155	Bac6	Astereae	Baccharis aff malmei
30	Ech1	Heliantheae	Echinacea purpurea	156	Bac7	Astereae	Baccharis aff sessilifolia
31	Emi1	Senecioneae	Emilia sonchifolia	157	Bac12	Astereae	Baccharis arguta
32	Ere1	Vernonieae	Eremanthus argenteus	158	Bac14	Astereae	Baccharis bifronsa
33	Ere4	Vernonieae	Eremanthus capitatus	159	Bac18	Astereae	Baccharis calliprinos
34	Ere7	Vernonieae	Eremanthus goyazensis	159,2	Bac18.2	Astereae	Baccharis calliprinos
35	Ere10	Vernonieae	Eremanthus polycephalus	159,3	Bac18.3	Astereae	Baccharis calliprinos
36	Het2	Vernonieae	Heterocoma gracilis	160	Bac20	Astereae	Baccharis chionolaenoides
37	Lac1	Cichorieae	Lactuca sativa	161	Bac23	Astereae	Baccharis debilis
38	Lyc6	Vernonieae	Lychnophora mellobarretoii	162	Bac26	Astereae	Baccharis dubia
39	Lyc8	Vernonieae	Lychnophora salicifolia	163	Bac28	Astereae	Baccharis genistelloides
40	Lyc9	Vernonieae	Lychnophora santosii	164	Bac30	Astereae	Baccharis glomerulifolia
41	Mat1	Anthemideae	Matricaria chamomilla	165	Bac34	Astereae	Baccharis inexpectata
42	Mik2	Eupatorieae	Mikania cordifolia	166	Bac36	Astereae	Baccharis kurtziana
43	Mik10	Eupatorieae	Mikania sessilifolia	167	Bac37	Astereae	Baccharis latifolia
44	Par1	Vernonieae	Paralychnophora atkinsiae	168	Bac38	Astereae	Baccharis linearifolia
45	Pip1	Vernonieae	Piptolepis monticola	169	Bac39	Astereae	Baccharis longipetiolata
46	Plu1	Inuleae	Pluchea quitoc	170	Bac41	Astereae	Baccharis macrophylla
47	Sch1	Barnadesieae	Schlechtendalia luzulaefolia	171	Bac42	Astereae	Baccharis marchesii
47,2	Sch1.2	Barnadesieae	Schlechtendalia luzulaefolia	172	Bac43	Astereae	Baccharis molina
48	Sma1	Vernonieae	Smallanthus sonchifolius	173	Bac44	Astereae	Baccharis multifolia
49	Soa	Vernonieae	Soaresia velutina	174	Bac45	Astereae	Baccharis napaea
50	Sol1	Astereae	Solidago microglossa	175	Bac46	Astereae	Baccharis neglecta
51	Son1	Cichorieae	Sonchus oleraceus	176	Bac50	Astereae	Baccharis papillosa chaparensisv
52	Sph1	Heliantheae	Sphagneticola trilobata	177	Bac52	Astereae	Baccharis pentlandii
53	Ste1	Eupatorieae	Stevia rebaudiana	178	Bac62	Astereae	Baccharis papillosa ssp integrifolia
54	Tan1	Anthemideae	Tanacetum parthenium	178,2	Bac62.2	Astereae	Baccharis papillosa ssp integrifolia
55	Tar1	Cichorieae	Taraxacum officinale	178,3	Bac62.3	Astereae	Baccharis papillosa ssp integrifolia
56	Tri1	Millerieae	Tridax procumbens	179	Bac56	Astereae	Baccharis polycephala
57	Ver8	Vernonieae	Vernonanthura beyrichii	180	Bac58	Astereae	Baccharis reticularia
58	Ver9	Vernonieae	Vernonia condensata	181	Bac59	Astereae	Baccharis retusa

59	Ver2	Vernonieae	Vernonia polyanthes	182	Bac61	Astereae	Baccharis salicifolia
60,2	Ald5	Heliantheae	Aldama bracteata	183	Bac53	Astereae	Baccharis salicina
61,2	Ald7	Heliantheae	Aldama discolor	184	Bac63	Astereae	Baccharis saliens
62,2	Arc1	Cynareae	Arctium lappa	185	Bac64	Astereae	Baccharis sarothroides
63	Bac9	Astereae	Baccharis alpina	186	Bac66	Astereae	Baccharis singularis
64	Bac10	Astereae	Baccharis angustifolia	187	Bac68	Astereae	Baccharis subalata
64,2	Bac10.2	Astereae	Baccharis angustifolia	188	Bac69	Astereae	Baccharis tarchonanthoides
64,3	Bac10.3	Astereae	Baccharis angustifolia	189	Bac71	Astereae	Baccharis tola frimbriata
64,4	Bac10.4	Astereae	Baccharis angustifolia	190	Bac73	Astereae	Baccharis trineura
65	Bac11	Astereae	Baccharis aphylla	191	Bac76	Astereae	Baccharis vargasii
66	Bac13	Astereae	Baccharis bicolor	192	Cae1	Calenduleae	Calendula officinalis
67	Bac15	Astereae	Baccharis boliviensis	194	Esp2	Millerieae	Espeletia arbelaezii
68	Bac16	Astereae	Baccharis buchtienii	195	Esp1	Millerieae	Espeletia annemariana
69	Bac17	Astereae	Baccharis caespitosa	196	Esp3	Millerieae	Espeletia ariana
70	Bac19	Astereae	Baccharis cenicienta	197	Esp4	Millerieae	Espeletia aristeguietana
70,2	Bac19.2	Astereae	Baccharis cenicienta	198	Esp6	Millerieae	Espeletia brachyaxiantha
71	Bac21	Astereae	Baccharis ciliata	199	Esp7	Millerieae	Espeletia brassicoidea
72	Bac22	Astereae	Baccharis crispa	200	Esp9	Millerieae	Espeletia cleefii
73	Bac24	Astereae	Baccharis densiflora	200,2	Esp9.2	Millerieae	Espeletia cleefii
73,2	Bac24.2	Astereae	Baccharis densiflora	200,3	Esp9.3	Millerieae	Espeletia cleefii
73,3	Bac24.3	Astereae	Baccharis densiflora	201	Esp10	Millerieae	Espeletia congestiflora
74	Bac25	Astereae	Baccharis dracunculifolia	201,2	Esp10.2	Millerieae	Espeletia congestiflora
75,2	Bac27	Astereae	Baccharis floripa	201,3	Esp10.3	Millerieae	Espeletia congestiflora
76,2	Bac31	Astereae	Baccharis grisebachii	202	Esp11	Millerieae	Espeletia congestiflora
77	Bac32	Astereae	Baccharis halimifolia	203	Esp13	Millerieae	Espeletia delicatissima
78,2	Bac33	Astereae	Baccharis incarum	204	Esp14	Millerieae	Espeletia episcopalis
79	Bac35	Astereae	Baccharis jelskii	205	Esp18	Millerieae	Espeletia grandiflora
80,2	Bac40	Astereae	Baccharis lychnophora	206	Esp19	Millerieae	Espeletia incana
81	Bac47	Astereae	Baccharis nitida	207	Esp20	Millerieae	Espeletia jajoensis
82	Bac49	Astereae	Baccharis papillosa	208	Esp21	Millerieae	Espeletia jaramilloi × E. boyacensis
83,3	Bac51	Astereae	Baccharis papillosa	209	Esp22	Millerieae	Espeletia jaramilloi
84	Bac54	Astereae	Baccharis pilularis	210	Esp23	Millerieae	Espeletia killipii
85	Bac55	Astereae	Baccharis plummerae	211	Esp24	Millerieae	Espeletia litocrassa
86	Bac57	Astereae	Baccharis quitensis	212	Esp25	Millerieae	Espeletia lopezii
87	Bac60	Astereae	Baccharis rodriguezii	213	Esp27	Millerieae	Espeletia murilloi

88	Bac65	Astereae	Baccharis sergiloides	214	Esp29	Millerieae	Espeletia nana
89	Bac67	Astereae	Baccharis stylosa	215	Esp30	Millerieae	Espeletia nemekenei
90	Bac70	Astereae	Baccharis tola	216	Esp31	Millerieae	Espeletia paipana × E. murilloi
91	Bac1	Astereae	Baccharis trimera	217	Esp32	Millerieae	Espeletia paipana
91,2	Bac1.2	Astereae	Baccharis trimera	218	Esp33	Millerieae	Espeletia pescana
91,3	Bac1.3	Astereae	Baccharis trimera	218,2	Esp33.2	Millerieae	Espeletia pescana
91,4	Bac1.4	Astereae	Baccharis trimera	218,3	Esp33.3	Millerieae	Espeletia pescana
92	Bac72	Astereae	Baccharis trinervis	219	Esp34	Millerieae	Espeletia pisbana
93	Bac74	Astereae	Baccharis umbellata	220	Esp35	Millerieae	Espeletia pleiochasia
93,2	Bac74.2	Astereae	Baccharis umbellata	221	Esp36	Millerieae	Espeletia rositae
94	Bac75	Astereae	Baccharis urvilleana	222	Esp37	Millerieae	Espeletia saboyensis
95	Bac77	Astereae	Baccharis wrightii	223	Esp38	Millerieae	Espeletia schultesiana
96	Chr1a	Vernonieae	Chresta angustifolia	224	Esp39	Millerieae	Espeletia schultzii × E. jajoensis
97	Chr2a	Vernonieae	Chresta curumbensis	225	Esp42	Millerieae	Espeletia standleyana
97,2	Chr2a.2	Vernonieae	Chresta curumbensis	226	Esp43	Millerieae	Espeletia steyermarkii
97,3	Chr2a.3	Vernonieae	Chresta curumbensis	227	Esp44	Millerieae	Espeletia summapacis
98,2	Chr4a	Vernonieae	Chresta harleyi	228	Esp45	Millerieae	Espeletia tenorae × E. schultzii
99	Chr14a	Vernonieae	Chresta phyllicifolia	229	Esp46	Millerieae	Espeletia uribei
100	Chr12a	Vernonieae	Chresta scapigera	230	Gym1	Vernonieae	Gymnanthemum amygdalinum
101	Chr8a	Vernonieae	Chresta scapigera x C. pycnocephala	230,2	Gym1.2	Vernonieae	Gymnanthemum amygdalinum
102	Chr9a	Vernonieae	Chresta souzae	230,3	Gym1.3	Vernonieae	Gymnanthemum amygdalinum
103	Chr10a	Vernonieae	Chresta speciosa	231	Lyc1	Vernonieae	Lychnophora crispa
104	Chr11a	Vernonieae	Chresta sphaerocephala	232	Lyc4	Vernonieae	Lychnophora gardneri
105	Das2	Barnadesieae	Dasyphyllum spinensis	234	Lyc12	Vernonieae	Lychnophora triflora
106	Das5	Barnadesieae	Dasyphyllum tomentosum	235	Lyc13	Vernonieae	Lychnophora grannogolensis
107	Das4	Barnadesieae	Dasyphyllum varians	236	Lyc14	Vernonieae	Lychnophorella leucodrendon
108	Ere2	Vernonieae	Eremanthus auriculatus	237,2	Mik5	Eupatorieae	Mikania firmula
109	Ere3	Vernonieae	Eremanthus brevifolius	237,3	Mik5.2	Eupatorieae	Mikania firmula
110	Ere5	Vernonieae	Eremanthus elaeagnus	237,4	Mik5.3	Eupatorieae	Mikania firmula
111	Ere6	Vernonieae	Eremanthus erythropappus	237,5	Mik5.4	Eupatorieae	Mikania firmula
112	Ere8	Vernonieae	Eremanthus incanus	238	Mik7	Eupatorieae	Mikania guaco
113	Ere9	Vernonieae	Eremanthus pabstii	239	Mik9	Eupatorieae	Mikania cuzcoensis
114	Ere11	Vernonieae	Eremanthus uniflorus	240	Mik11	Eupatorieae	Mikania fulva
115	Ere12	Vernonieae	Eremanthus veadeiroensis	241	Ver3	Vernonieae	Vernonia 1 - 168 - Solidago chilensis

116	Esp5	Millerieae	Espeletia barclayana	242	Ver5	Vernonieae	Vernonia 3 - 170
117	Esp8	Millerieae	Espeletia cayetana	243	Ver6	Vernonieae	Vernonia 4 - 171
118	Esp12	Millerieae	Espeletia curialensis	244	Ver7	Vernonieae	Vernonia 5 - 172 - Solidago chilensis
119	Esp15	Millerieae	Espeletia estanislana	245	Ver10	Vernonieae	Vernonia Coleta 2 ponto 1 - 162
120	Esp16	Millerieae	Espeletia frontinoensis	246	Ver11	Vernonieae	Vernonia Coleta 2 ponto 2 - 163
121	Esp17	Millerieae	Espeletia garciae	246,2	Ver11.2	Vernonieae	Vernonia Coleta 2 ponto 2 - 163
122	Esp26	Millerieae	Espeletia manguana	246,3	Ver11.3	Vernonieae	Vernonia Coleta 2 ponto 2 - 163
123	Esp28	Millerieae	Espeletia murilloi	247	Ver13	Vernonieae	Vernonia Coleta 2 ponto 4.1 - 165
124	Esp40	Millerieae	Espeletia schultzii	248	Ver14	Vernonieae	Vernonia Coleta 2 ponto 4.2 - 166
124,2	Esp40.2	Millerieae	Espeletia schultzii	3AC	Ald10	Heliantheae	A. nudibasilaris
124,3	Esp40.3	Millerieae	Espeletia schultzii	7AC	Ald11	Heliantheae	A. trichophylla
124,4	Esp40.4	Millerieae	Espeletia schultzii	8AC	Ald12	Heliantheae	A. grandiflora
124,5	Esp40.5	Millerieae	Espeletia schultzii	9AC	Ald13	Heliantheae	A. pilosa
124,6	Esp40.6	Millerieae	Espeletia schultzii	10AC	Helth	Heliantheae	Helianthus annuus

Apêndice B – Cultivo de código do cultivo de *A. robusta* e *A. trichophylla* (Asteraceae), e suas respectivas espécies coletadas em seus habitats naturais (“selvagem”). Na tabela constam outras espécies do gênero *Aldama* La Llave, bem como *T. diversifolia* e *H. annuus* (Asteraceae). Season = estação da coleta; Drying = secagem da planta coletada; Circadian = horário da coleta; Soil and Luminosity = condições do cultivo (com/sem sombrite e com/sem adubação); Specie = nome da espécie no estudo. Flowering period = período de florada; Air Oven = secagem em estufa de ar-circulante (40 °C); Nitrogen = secagem em N₂ líquido (congelamento seguido por liofilização); hours = hora da coleta; with/without = com/sem; shad = sombra (cultivo sob sombrite); fert = fertilizante (cultivo com adubação). Summer = verão; Autumn = outono; Spring = primavera; Winter = inverno. Pool = Mistura de drogas vegetais para controle de qualidade em metabólica. Pool 1 = *Achillea millefolium*, *A. robusta*, *A. nudibasilaris*, *T. diversifolia*, *Mikania laevigata*, *Smilax sonchifolius*, *Calea* sp, *Chiquiraga spinosa*. Pool 2 = *A. trichophylla*, *A. grandiflora*, *A. pilosa*, *H. annuus*, *Solidago microglosa*, *Calea* sp, *Chiquiraga spinosa*.

ID Código	Season	Drying	Circadian	Soil and Luminosity	Specie
1	Flowering period	Air Oven			<i>Aldama robusta</i> selvagem
2	Flowering period	Air Oven			<i>Aldama trichophylla</i> selvagem
3	Summer	Nitrogen	7 hours	with shad and without fert	<i>A. robusta</i> 7 h, c/ sombra, s/ adubo
4	Summer	Nitrogen	11 hours	with shad and without fert	<i>A. robusta</i> 11 h, c/ sombra, s/ adubo
5	Summer	Air Oven	11 hours	with shad and without fert	<i>A. robusta</i> 11 h, c/ sombra, s/ adubo (seca)
6	Summer	Nitrogen	15 hours	with shad and without fert	<i>A. robusta</i> 15 h, c/ sombra, s/ adubo
7	Summer	Nitrogen	20 hours	with shad and without fert	<i>A. robusta</i> 20 h, c/ sombra, s/ adubo
8	Summer	Nitrogen	20 hours	with shad and without fert	<i>A. robusta</i> 24 h, c/ sombra, s/ adubo
9	Summer	Nitrogen	7 hours	with shad and with fert	<i>A. robusta</i> 7 h, c/ sombra, c/ adubo

10	Summer	Nitrogen	11 hours	with shad and with fert	A. robusta 11 h, c/ sombra, c/ adubo
11	Summer	Air Oven	11 hours	with shad and with fert	A. robusta 11 h, c/ sombra, c/ adubo (seca)
12	Summer	Nitrogen	15 hours	with shad and with fert	A. robusta 15 h, c/ sombra, c/ adubo
13	Summer	Nitrogen	20 hours	with shad and with fert	A. robusta 20 h, c/ sombra, c/ adubo
14	Summer	Nitrogen	24 hours	with shad and with fert	A. robusta 24 h, c/ sombra, c/ adubo
15	Summer	Nitrogen	7 hours	without shad and without fert	A. robusta 7 h, s/ sombra, s/ adubo
16	Summer	Nitrogen	11 hours	without shad and without fert	A. robusta 11 h, s/ sombra, s/ adubo
17	Summer	Nitrogen	11 hours	without shad and without fert	A. robusta 11 h, s/ sombra, s/ adubo (seca)
18	Summer	Nitrogen	15 hours	without shad and without fert	A. robusta 15 h, s/ sombra, s/ adubo
19	Summer	Nitrogen	20 hours	without shad and without fert	A. robusta 20 h, s/ sombra, s/ adubo
20	Summer	Nitrogen	24 hours	without shad and without fert	A. robusta 24 h, s/ sombra, s/ adubo
21	Summer	Nitrogen	7 hours	without shad and with fert	A. robusta 7 h, s/ sombra, c/ adubo
22	Summer	Nitrogen	11 hours	without shad and with fert	A. robusta 11 h, s/ sombra, c/ adubo

23	Summer	Air Oven	11 hours	without shad and with fert	A. robusta 11 h, s/ sombra, c/ adubo (seca)
24	Summer	Nitrogen	15 hours	without shad and with fert	A. robusta 15 h, s/ sombra, c/ adubo
25	Summer	Nitrogen	20 hours	without shad and with fert	A. robusta 20 h, s/ sombra, c/ adubo
26	Summer	Nitrogen	24 hours	without shad and with fert	A. robusta 24 h, s/ sombra, c/ adubo
27	Autumn	Nitrogen	7 hours	with shad and without fert	A. robusta 7 h, c/ sombra, s/ adubo
28	Autumn	Nitrogen	7 hours	with shad and with fert	A. robusta 7 h, c/ sombra, c/ adubo
29	Autumn	Nitrogen	7 hours	without shad and without fert	A. robusta 7 h, s/ sombra, s/ adubo
30	Autumn	Nitrogen	7 hours	without shad and with fert	A. robusta 7 h, s/ sombra, c/ adubo
31	Autumn	Nitrogen	7 hours	without shad and without fert	A. trichophylla 7 h s/ adubo
32	Autumn	Nitrogen	7 hours	without shad and with fert	A. trichophylla 7 h c/ adubo
33	Autumn	Nitrogen	11 hours	with shad and without fert	A. robusta 11 h, c/ sombra, s/ adubo
34	Autumn	Nitrogen	11 hours	with shad and with fert	A. robusta 11 h, c/ sombra, c/ adubo
35	Autumn	Nitrogen	11 hours	without shad and without fert	A. robusta 11 h, s/ sombra, s/ adubo

36	Autumn	Nitrogen	11 hours	without shad and with fert	A. robusta 11 h, s/ sombra, c/ adubo
37	Autumn	Nitrogen	11 hours	without shad and without fert	A. trichophylla 11 h s/ adubo
38	Autumn	Nitrogen	11 hours	without shad and with fert	A. trichophylla 11 h c/ adubo
39	Autumn	Nitrogen	15 hours	with shad and without fert	A. robusta 15 h, c/ sombra, s/ adubo
40	Autumn	Nitrogen	15 hours	with shad and with fert	A. robusta 15 h, c/ sombra, c/ adubo
41	Autumn	Nitrogen	15 hours	without shad and without fert	A. robusta 15 h, s/ sombra, s/ adubo
42	Autumn	Nitrogen	15 hours	without shad and with fert	A. robusta 15 h, s/ sombra, c/ adubo
43	Autumn	Nitrogen	15 hours	without shad and without fert	A. trichophylla 15 h s/ adubo
44	Autumn	Nitrogen	15 hours	without shad and with fert	A. trichophylla 15 h c/ adubo
45	Autumn	Nitrogen	20 hours	with shad and without fert	A. robusta 20 h, c/ sombra, s/ adubo
46	Autumn	Nitrogen	20 hours	with shad and with fert	A. robusta 20 h, c/ sombra, c/ adubo
47	Autumn	Nitrogen	20 hours	without shad and without fert	A. robusta 20 h, s/ sombra, s/ adubo
48	Autumn	Nitrogen	20 hours	without shad and with fert	A. robusta 20 h, s/ sombra, c/ adubo

49	Autumn	Nitrogen	20 hours	without shad and without fert	A. trichophylla 20 h s/ adubo
50	Autumn	Nitrogen	20 hours	without shad and with fert	A. trichophylla 20 h c/ adubo
51	Autumn	Nitrogen	24 hours	with shad and without fert	A. robusta 24 h, c/ sombra, s/ adubo
52	Autumn	Nitrogen	24 hours	with shad and with fert	A. robusta 24 h, c/ sombra, c/ adubo
53	Autumn	Nitrogen	24 hours	without shad and without fert	A. robusta 24 h, s/ sombra, s/ adubo
54	Autumn	Nitrogen	24 hours	without shad and with fert	A. robusta 24 h, s/ sombra, c/ adubo
55	Autumn	Nitrogen	24 hours	without shad and without fert	A. trichophylla 24 h s/ adubo
56	Autumn	Nitrogen	24 hours	without shad and with fert	A. trichophylla 24 h c/ adubo
57	Winter	Nitrogen	7 hours	with shad and without fert	A. robusta 7 h, c/ sombra, s/ adubo
58	Winter	Nitrogen	7 hours	with shad and with fert	A. robusta 7 h, c/ sombra, c/ adubo
59	Winter	Nitrogen	7 hours	without shad and without fert	A. robusta 7 h, s/ sombra, s/ adubo
60	Winter	Nitrogen	7 hours	without shad and with fert	A. robusta 7 h, s/ sombra, c/ adubo
61	Winter	Nitrogen	11 hours	with shad and without fert	A. robusta 11 h, c/ sombra, s/ adubo

62	Winter	Nitrogen	11 hours	with shad and with fert	A. robusta 11 h, c/ sombra, c/ adubo
63	Winter	Nitrogen	11 hours	without shad and without fert	A. robusta 11 h, s/ sombra, s/ adubo
64	Winter	Nitrogen	11 hours	without shad and with fert	A. robusta 11 h, s/ sombra, c/ adubo
65	Winter	Air Oven	11 hours	with shad and without fert	A. robusta 11 h, c/ sombra, s/ adubo (seca)
66	Winter	Air Oven	11 hours	with shad and with fert	A. robusta 11 h, c/ sombra, c/ adubo (seca)
67	Winter	Air Oven	11 hours	without shad and without fert	A. robusta 11 h, s/ sombra, s/ adubo (seca)
68	Winter	Air Oven	11 hours	without shad and with fert	A. robusta 11 h, s/ sombra, c/ adubo (seca)
69	Winter	Nitrogen	11 hours	without shad and without fert	A. trichophylla 11 h s/adubo
70	Winter	Nitrogen	11 hours	without shad and with fert	A. trichophylla 11h c/adubo
71	Winter	Air Oven	11 hours	without shad and without fert	A. trichophylla 11 h s/adubo (seca)
72	Winter	Air Oven	11 hours	without shad and with fert	A. trichophylla 11h c/adubo (seca)
73	Winter	Nitrogen	15 hours	with shad and without fert	A. robusta 15 h, c/ sombra, s/ adubo
74	Winter	Nitrogen	15 hours	with shad and with fert	A. robusta 15 h, c/ sombra, c/ adubo

75	Winter	Nitrogen	15 hours	without shad and without fert	A. robusta 15 h, s/ sombra, s/ adubo
76	Winter	Nitrogen	15 hours	without shad and with fert	A. robusta 15 h, s/ sombra, c/ adubo
77	Winter	Nitrogen	20 hours	with shad and without fert	A. robusta 20 h, c/ sombra, s/ adubo
78	Winter	Nitrogen	20 hours	with shad and with fert	A. robusta 20 h, c/ sombra, c/ adubo
79	Winter	Nitrogen	20 hours	without shad and without fert	A. robusta 20 h, s/ sombra, s/ adubo
80	Winter	Nitrogen	20 hours	without shad and with fert	A. robusta 20 h, s/ sombra, c/ adubo
81	Winter	Nitrogen	24 hours	with shad and without fert	A. robusta 24 h, c/ sombra, s/ adubo
82	Winter	Nitrogen	24 hours	with shad and with fert	A. robusta 24 h, c/ sombra, c/ adubo
83	Winter	Nitrogen	24 hours	without shad and without fert	A. robusta 24 h, s/ sombra, s/ adubo
84	Winter	Nitrogen	24 hours	without shad and with fert	A. robusta 24 h, s/ sombra, c/ adubo
85	Spring	Nitrogen	7 hours	with shad and without fert	A. robusta 7 h, c/ sombra, s/ adubo
86	Spring	Nitrogen	7 hours	with shad and with fert	A. robusta 7 h, c/ sombra, c/ adubo
87	Spring	Nitrogen	7 hours	without shad and without fert	A. robusta 7 h, s/ sombra, s/ adubo

88	Spring	Nitrogen	7 hours	without shad and with fert	A. robusta 7 h, s/ sombra, c/ adubo
89	Spring	Nitrogen	11 hours	with shad and without fert	A. robusta 11 h, c/ sombra, s/ adubo
90	Spring	Nitrogen	11 hours	with shad and with fert	A. robusta 11 h, c/ sombra, c/ adubo
91	Spring	Nitrogen	11 hours	without shad and without fert	A. robusta 11 h, s/ sombra, s/ adubo
92	Spring	Nitrogen	11 hours	without shad and with fert	A. robusta 11 h, s/ sombra, c/ adubo
93	Spring	Air Oven	11 hours	with shad and without fert	A. robusta 11 h, c/ sombra, s/ adubo (seca)
94	Spring	Air Oven	11 hours	with shad and with fert	A. robusta 11 h, c/ sombra, c/ adubo (seca)
95	Spring	Air Oven	11 hours	without shad and without fert	A. robusta 11 h, s/ sombra, s/ adubo (seca)
96	Spring	Air Oven	11 hours	without shad and with fert	A. robusta 11 h, s/ sombra, c/ adubo (seca)
97	Spring	Nitrogen	11 hours	without shad and without fert	A. trichophylla 11 h s/adubo
98	Spring	Nitrogen	11 hours	without shad and with fert	A. trichophylla 11h c/adubo
99	Spring	Air Oven	11 hours	without shad and without fert	A. trichophylla 11 h s/adubo (seca)
100	Spring	Air Oven	11 hours	without shad and with fert	A. trichophylla 11h c/adubo (seca)

101	Spring	Nitrogen	15 hours	with shad and without fert	A. robusta 15 h, c/ sombra, s/ adubo
102	Spring	Nitrogen	15 hours	with shad and with fert	A. robusta 15 h, c/ sombra, c/ adubo
103	Spring	Nitrogen	15 hours	without shad and without fert	A. robusta 15 h, s/ sombra, s/ adubo
104	Spring	Nitrogen	15 hours	without shad and with fert	A. robusta 15 h, s/ sombra, c/ adubo
105	Spring	Nitrogen	20 hours	with shad and without fert	A. robusta 20 h, c/ sombra, s/ adubo
106	Spring	Nitrogen	20 hours	with shad and with fert	A. robusta 20 h, c/ sombra, c/ adubo
107	Spring	Nitrogen	20 hours	without shad and without fert	A. robusta 20 h, s/ sombra, s/ adubo
108	Spring	Nitrogen	20 hours	without shad and with fert	A. robusta 20 h, s/ sombra, c/ adubo
109	Spring	Nitrogen	24 hours	with shad and without fert	A. robusta 24 h, c/ sombra, s/ adubo
110	Spring	Nitrogen	24 hours	with shad and with fert	A. robusta 24 h, c/ sombra, c/ adubo
111	Spring	Nitrogen	24 hours	without shad and without fert	A. robusta 24 h, s/ sombra, s/ adubo
112	Spring	Nitrogen	24 hours	without shad and with fert	A. robusta 24 h, s/ sombra, c/ adubo
2A	Flowering period	Air Oven			Aldama robusta selvagem 2

3A	Flowering period	Air Oven			<i>A. nudibasilaris</i>
4A	Flowering period	Air Oven			<i>Tithonia diversifolia</i>
7A	Flowering period	Air Oven			<i>Aldama trichophylla selvagem 2</i>
8A	Flowering period	Air Oven			<i>A. grandiflora</i>
9A	Flowering period	Air Oven			<i>A. pilosa</i>
10A	Flowering period	Air Oven			<i>Helianthus annuus</i>
14A	Flowering period	Air Oven			<i>A. arenaria</i>
15A	Flowering period	Air Oven			<i>A. bakerina</i>
16A	Flowering period	Air Oven			<i>A. bracteata</i>
17A	Flowering period	Air Oven			<i>A. discolor</i>
18A	Flowering period	Air Oven			<i>A. gardneri</i>
19A	Flowering period	Air Oven			<i>A. filifolia</i>
1p1-1	Pool 1	Nitrogen	Pool 1	Pool 1	pool 1 - extração 1 - injeção 1

1p1-2	Pool 1	Nitrogen	Pool 1	Pool 1	pool 1 - extração 1 - injeção 2
1p1-3	Pool 1	Nitrogen	Pool 1	Pool 1	pool 1 - extração 1 - injeção 3
1p2-1	Pool 1	Nitrogen	Pool 1	Pool 1	pool 1 - extração 2 - injeção 1
1p2-2	Pool 1	Nitrogen	Pool 1	Pool 1	pool 1 - extração 2 - injeção 2
1p2-3	Pool 1	Nitrogen	Pool 1	Pool 1	pool 1 - extração 2 - injeção 3
1p3-1	Pool 1	Nitrogen	Pool 1	Pool 1	pool 1 - extração 3 - injeção 1
1p3-2	Pool 1	Nitrogen	Pool 1	Pool 1	pool 1 - extração 3 - injeção 2
1p3-3	Pool 1	Nitrogen	Pool 1	Pool 1	pool 1 - extração 3 - injeção 3
2p1-1	Pool 2	Nitrogen	Pool 2	Pool 2	pool 2 - extração 1 - injeção 1
2p1-2	Pool 2	Nitrogen	Pool 2	Pool 2	pool 2 - extração 1 - injeção 2
2p1-3	Pool 2	Nitrogen	Pool 2	Pool 2	pool 2 - extração 1 - injeção 3
2p2-1	Pool 2	Nitrogen	Pool 2	Pool 2	pool 2 - extração 2 - injeção 1
2p2-2	Pool 2	Nitrogen	Pool 2	Pool 2	pool 2 - extração 2 - injeção 2

2p2-3	Pool 2	Nitrogen	Pool 2	Pool 2	pool 2 - extração 2 - injeção 3
2p3-1	Pool 2	Nitrogen	Pool 2	Pool 2	pool 3 - extração 3 - injeção 1
2p3-2	Pool 2	Nitrogen	Pool 2	Pool 2	pool 3 - extração 3 - injeção 2
2p3-3	Pool 2	Nitrogen	Pool 2	Pool 2	pool 3 - extração 3 - injeção 3

Apêndice C – Parâmetros utilizados no processamentos de dados no MZmine (Pluskal et al., 2010).

1. Raw data import

2. Raw data filtering:

Raw data filtering > Scan by filtering

Filter: Savitzky-Golay filter. Number of data points: 5

3. Mass detection:

Peak detection > Mass detection. MS level 1. Set filters: + ou – (dependendo do modo de ionização).

Mass detector: Exact mass. Noise level: 1.0E6

4. FTMS shoulder peaks filter:

Mass resolution: 70.000

Peak model function: Lorentzian extended

5. Chromatogram builder:

Peak detection > Chromatogram builder

Min time spam (min): 0.2

Min height: 5E6

m/z tolerance: 0.001 m/z or 2 ppm

6. Smoothing:

Peak smoothing: 5

7. Chromatogram deconvolution:

Peak recognition: Local minimum search:

Chromatographic threshold: 20%

Search minimum in RT range: 0.20

Minimum relative height: 10%

Minimum absolute height: 4E6

Min ratio of peak top/edge: 2

Peak duration range: 0.25 to 2.00

8. Deisotope:

Isotopic peaks grouper:

m/z tolerance: 0.001

RT tolerance: 0.2

Maximum charge: 2

Representative isotope: most intense

9. Alignment > Ransac aligner algorithm

m/z tolerance: 0.001 ou 2 ppm

RT tolerance: 0.8

RT after correction: 0.5

Ransac interections: 0

Min. number of points: 50% /// Threshold value: 10

10. Gap filling:

Intensity tolerance: 30%

m/z tolerance: 0.001 ou 2 ppm

RT tolerance: 0.7

RT correction

11. Filtering: Duplicate peaks filter

m/z tolerance: 0.001 ou 2 ppm

RT tolerance: 0.7

12. Identification:

Complex search (select H+, if positive mode)/Adduct search (modo positivo: Na+, K+; modo negativo: Cl-, COO-):

m/z tolerance: 0.001 ou 1 ppm

RT tolerance: 0.1; Mass adduct peak height: 50%

Custom database search: AsterDB ou standar compounds

m/z tolerance: 0.01 ou 5 ppm; RT tolerance: 0.3