UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

**Quimiometria aplicada à metabolômica de *Aldama* La Llave:
Contribuições quimiotaxonômicas e fitoquímica direcionada
baseada em inibição de cicloxigease-1 e 5-lipoxigease**

Felipe Antunes dos Santos

Ribeirão Preto

2014

UNIVERSIDADE DE SÃO PAULO

FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO

**Quimiometria aplicada à metabolômica de *Aldama* La Llave:
Contribuições quimiotaxonômicas e fitoquímica direcionada
baseada em inibição de cicloxigease-1 e 5-lipoxigease**

Dissertação de Mestrado apresentada ao
Programa de Pós-Graduação em Ciências
Farmacêuticas para obtenção do Título de
Mestre em Ciências

Área de Concentração: Produtos Naturais e
Sintéticos

Orientado(a): Felipe Antunes dos Santos

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

Ribeirão Preto

2014

RESUMO

SANTOS, F. A. **Quimiometria aplicada à metabolômica de *Aldama* La Llave: Contribuições quimiotaxonômicas e fitoquímica direcionada baseada em inibição de ciclooxigease-1 e 5-lipoxigease.** 2014. 143f. Dissertação (Mestrado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2014.

As espécies do gênero *Viguiera* Kunth, recentemente transferidas para *Aldama* La Llave (Asteraceae, Heliantheae), ainda possuem problemas de delimitação taxonômica. Tal gênero possui 35 espécies distribuídas por todo o território brasileiro, principalmente no cerrado. Análises químicas têm demonstrado seu potencial em auxiliar na resolução de problemas em vários níveis taxonômicos com base em grupos químicos de metabólitos secundários. Além disso, análises fitoquímicas tem revelado que determinadas substâncias de *Aldama* possuem potencial para serem estudadas tendo vista a investigação de mecanismos moleculares anti-inflamatórios. Desta forma, foi proposto neste trabalho realizar a abordagem da metabolômica não-direcionada auxiliada por quimiometria, visando-se fornecer dados químicos tanto para contribuições quimiotaxonômicas quanto para encontrar substâncias bioativas. Por meio de tal abordagem, análises multivariadas não supervisionadas (PCA e HCA), bem como supervisionadas (OPLS-DA), com dados provenientes de UHPLC-DAD-Orbitrap, demonstraram que o gênero *Aldama* é quimioconsistente e, deste modo, determinadas substâncias químicas discriminantes foram sugeridas. Além disso, espécies que apresentam problemas taxonômicos tiveram os seus posicionamentos infragenéricos explicado do ponto de vista quimiotaxonômico. Quanto às análises para encontrar substâncias bioativas (fitoquímica direcionada), os alvos inflamatórios pesquisados foram as enzimas pró-inflamatórias COX-1 e 5-LOX, com as quais se realizou ensaios *in vitro* verificando-se a atividade positiva de extratos de espécies de *Aldama*. *A. trichophylla* foi a espécie selecionada para realizar a fitoquímica direcionada, uma vez que apresentou inibição dupla contra as enzimas e por se tratar de uma espécie muito pouco estudada do ponto de vista fitoquímico. Os melhores modelos de inibição de COX-1 e 5-LOX criados com dados químicos de *Aldama* foram selecionados a partir de diversas combinações de algoritmos, *softwares* de processamento (*MZmine*, *MetAlign* e *MSClust*) e técnicas-hifenadas, tais como UHPLC-Orbitrap e HPLC-TOF. Deste modo, determinados picos de íons foram apontados pela quimiometria como sendo discriminantes para a atividade de inibição dupla. Tais picos foram desreplicados com base em seus perfis de UV e padrões de fragmentação via HCD-Orbitrap e Ion-Trap. Deste modo, foi possível desreplicar três prováveis substâncias químicas inibidoras de COX-1 e 5-LOX: kaempferol-3-*O*-glucuronideo, quercetina-3-*O*-metil-7-glucuronideo e kaempferol-3-*O*-(6"-malonil-glucosideo). Por fim, foi realizada a fitoquímica direcionada. O fracionamento cromatográfico permitiu isolar tais substâncias e uma análise preliminar de RMN ¹H foi realizada com o intuito de realizar a identificação estrutural. Tais substâncias terão a suas atividades confirmadas na inibição das duas enzimas em um futuro próximo.

Palavras-chave: *Aldama* La Llave, Quimiotaxonomia, Fitoquímica direcionada, Quimiometria, Metabolômica não-direcionada

ABSTRACT

SANTOS, F. A. **Chemometrics applied to metabolomics of *Aldama* La Llave: chemotaxonomic contributions and targeted phytochemistry based on cyclooxygenase-1 and 5-lipoxygenase inhibition.** 2014. 143f. Dissertation (Master). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2014.

The species of the genus *Viguiera* Kunth, recently transferred to the genus *Aldama* La Llave (Asteraceae, Heliantheae), still show problems of taxonomic boundaries. This genus has 35 species distributed throughout the Brazilian territory, mostly in the cerrado. Chemical analyses has demonstrated its potential to help in the solution of problems at various taxonomic levels based on chemical groups of secondary metabolites. Furthermore, phytochemical analyses have shown that certain compounds from *Aldama* have the potential to be studied aiming to investigate anti-inflammatory molecular mechanisms. Thus, in this work it was proposed to carry out the untargeted metabolomic approach aided by chemometrics, aiming to provide chemical data either for chemotaxonomic contributions and to find bioactive compounds. Through this approach, both unsupervised (PCA and HCA) and supervised multivariate analysis (OPLS-DA) combined with data from UHPLC-DAD-Orbitrap analyses showed that the genus *Aldama* is “chemoconsistent” and thus certain discriminant compounds were suggested. Additionally, some species with taxonomic problems had their infrageneric positioning explained from the chemotaxonomic point of view. With regards to the analyses carried out to find bioactive compounds (targeted phytochemistry), the inflammatory targets investigated in this study were the COX-1 and LOX-5 pro-inflammatory enzymes with which *in vitro* assays were made and positive activity of extracts from species of *Aldama* was observed. *A. trichophylla* was the selected species for the targeted phytochemistry because it showed dual inhibition activity against both enzymes and it is still little investigated from the chemical point of view. The best models for inhibition of COX-1 and 5-LOX obtained with chemical data of *Aldama* were selected by means of various combinations of algorithms, processing software (MZmine, MetAlign and MScLust) and hyphenated techniques, such as HPLC-TOF and UHPLC-Orbitrap. Thus, certain peak ions were appointed by chemometrics as discriminant for the dual inhibition activity. These peaks were dereplicated based on their UV profiles and fragmentation patterns via HCD-Orbitrap and Ion-Trap. Thus, it was possible to dereplicate three probable chemical inhibitors of COX-1 and 5-LOX: kaempferol-3-*O*-glucuronide, quercetin-3-*O*-methyl-7-glucuronide and kaempferol-3-*O*-(6"-malonyl-glucoside). Finally, the targeted phytochemistry was carried out. The chromatographic fractionation allowed us to isolate these three compounds and a preliminary ¹H NMR analysis was performed in order to conclude their structural identification. In the near future, these substances will have their activity evaluated by the inhibition of the two enzymes.

Keywords: *Aldama* La Llave, Chemotaxonomy, Targeted Phytochemistry, Chemometrics, Untargeted Metabolomics

Introdução

A família Asteraceae (Compositae) é uma das mais numerosas dentre as Angiospermas, tendo cerca de 1.100 gêneros e 25.000 espécies (Funk *et al.*, 2009). De acordo com Verdi e seus colaboradores (2005), por volta de “98% dos gêneros são constituídos de plantas de pequeno porte, sendo encontradas em todos os habitats, principalmente nas regiões tropicais da América do Sul”. O gênero *Viguiera* Kunth (tribo Heliantheae, subtribo Helianthinae) possui 35 espécies brasileiras distribuídas em diversas regiões do território nacional, principalmente no cerrado (Magenta e Pirani, 2014; Blake, 1918). Tal gênero possui 180 espécies e apresenta problemas de delimitação taxonômica e têm sido amplamente estudado por botânicos, taxonomistas e fitoquímicos. Pesquisas recentes sugeriram uma posição derivada do gênero dentro da subtribo Helianthinae e uma origem recente dentro do bioma cerrado, bem como a necessidade de se realizar uma reavaliação taxonômica das espécies brasileiras (Schilling *et al.*, 2000 e 2012).

Assim sendo, em 2006 foi realizada uma revisão taxonômica das espécies de *Viguiera* que ocorrem no Brasil (Magenta, 2006), verificando-se extensivamente a morfoanatomia do grupo, concluindo-se que havia a necessidade de se realizar estudos moleculares (com material genético) e fitoquímicos integrados às características morfoanatômicas do gênero. Por exemplo, as espécies *V. anchusifolia* e *V. megapotamica* apresentam morfologia externa muitas vezes semelhante, o que dificulta a sua delimitação. Além disso, foi observado que além das espécies *V. robusta* e *V. arenaria* serem facilmente confundidas em herbários e, muitas vezes, só distinguidas pelas características do grão de pólen, as espécies *V. discolor*, *V. squalida* e *V. bakerina* podem ser a mesma. Há outros problemas de delimitação que foram destacados por outros autores, tais como a difícil delimitação entre *V. linearifolia*, *V. filifolia* e *V. trichophylla*, por exemplo, sendo que as duas primeiras foram confundidas entre si por Blake (1884) e Blake (1918), ao passo que *V. trichophylla* foi considerada coespecífica de *V. linearifolia* por Robinson e Moore (2004). Não obstante todos estes problemas observados, recentemente os resultados obtidos das análises de ITS (região do espaçador ribossomal interno) com sítios de restrição de cpDNA e de combinações entre as análises de ITS e ETS (região do espaçador ribossomal externo), o grupo parafilético *Viguiera* foi reduzido a apenas uma espécie, *Viguiera dentata*, sendo, portanto, considerado agora um grupo monoespecífico, transferindo-se todas as outras espécies para o gênero *Aldama* (Schilling e Panero, 2002 e 2011). Entretanto, há de se levar em consideração que poucas

espécies brasileiras foram incluídas neste estudo. Ainda assim, Magenta considerou em seu último trabalho publicado o nome *Aldama* para as espécies brasileiras, destacando ainda novidades infraespecíficas e que havia um equívoco cometido por Blake ao diferenciar *V. bakeriana* de *V. subdentata*, sendo que ambas tratam-se da mesma espécie (Magenta, 2014). Portanto, entrando em perfeita consonância com os estudos morfoanatômicos e filogenéticos aqui apresentados, nesta dissertação o gênero *Viguiera* Kunth será de agora em diante designado como *Aldama* La Llave.

As últimas pesquisas têm demonstrado os avanços em se caracterizar e traçar uma delimitação taxonômica consistente para o gênero *Aldama*. Sendo assim, os caracteres químicos podem dar a sua contribuição às classificações taxonômicas naquilo que se denomina quimiotaxonomia. Em tempo, cabe aqui explicar *a priori* algumas definições conceituais antes de entrar nas contribuições químicas para a taxonomia de plantas. Primeiramente, é de suma importância mencionar que quimiotaxonomia não pode ser confundida com quimiosistemática. A primeira diz respeito à contribuição química para classificações taxonômicas, sendo que táxon significa uma unidade taxonômica, que por sua vez indica um determinado nível de sistema de classificação científica. Assim sendo, taxonomia é o estudo que descreve os taxa. Já quimiosistemática se trata da mesma utilização dos caracteres químicos para classificações taxonômicas, porém, integrados a outros estudos, como os filogenéticos, ambientais, sazonais, circadianos, ecoquímicos, etc. Fica claro que quimiotaxonomia é um estudo mais restrito, ao passo que quimiosistemática é um estudo mais amplo. Apenas para deixar mais claras as definições, pode-se mencionar aqui que filogenia é o estudo das relações evolutivas entre os seres vivos.

Deste modo, alguns químicos de produtos naturais têm utilizado a química em classificações taxonômicas, uma vez que têm se demonstrado que determinadas substâncias químicas do metabolismo secundário de plantas são específicas de determinadas famílias, tribos, gêneros, ou até mesmo de espécies. Por exemplo, para o nosso caso particular, pode-se citar os trabalhos de Seaman (1982) e de Zdero e Bohlman (1990) que evidenciaram que a família Asteraceae possui uma tendência em biossintetizar lactonas sesquiterpênicas (LST). Observaram também que a tribo Heliantheae tem uma tendência em biossintetizar diterpenos do tipo *ent*-caurano e LST do tipo furanoeliangolido com configuração β no carbono 8 (C-8), diferentemente das espécies de Vernonieae, cuja configuração no C-8 é sempre α . Pode-se visualizar estas estruturas químicas na Figura 1.

Figura 1 – Diterpeno do tipo *ent*-caurano e LST do tipo furanoeliangolido orientada em β no C-8, comumente encontradas em Heliantheae, e orientada em α no C-8, típicas de Vernoniaeae. Os grupos R, R₁ e R₂ podem ser átomos de hidrogênio ou hidroxila e R₄ são grupos acila.

Também poderíamos notar que as LST do tipo furanoeliangolido 1-oxo-2-en-3,10-epoxieliangolido (budleína A – Figura não mostrada) não ocorrem nos gêneros *Tithonia* e *Helianthus* (Da Costa, 2008), os quais são filogeneticamente muito próximos a *Aldama*, sendo que esta é tida como um ancestral comum destes gêneros, segundo Blake (1918). Ainda assim, muitos outros metabólitos são compartilhados por estes gêneros e pode-se perceber que, especificamente em relação às LST, as classes químicas possuem esqueletos mais simples, tais como aqueles representados pelos níveis I e II da Figura 2 (Seaman, 1982). Com base nisso, pode-se também traçar vários níveis evolutivos para outras classes de metabólitos, tal como o exemplo das LST (Figura 2).

Figure 2 - Níveis de especiação de esqueleto carbônico de LST.

De fato, de maneira geral pode-se observar que níveis taxonômicos elevados de Asteraceae, tais como subfamílias e tribos, podem ser classificados de acordo com o tipo de esqueleto carbônico das estruturas químicas, ao passo que níveis mais baixos, como gêneros, espécies e variedades, podem ser classificados quimiotaxonomicamente de acordo com padrões de substituições, tais como grupos funcionais, estereoquímica e isomeria de posição, e níveis de oxidação (Da Costa, 2008).

Por falar em níveis de oxidação, cabe aqui mencionar a “teoria redox” da evolução química proposta por Gottlieb (1992) que descreve a relação diretamente proporcional entre os níveis de oxidação dos metabólitos secundários e o grau evolutivo das plantas a que eles pertencem. Esta teoria acredita ser o oxigênio um elemento que dá grau de especiação às estruturas químicas, possibilitando que elas participem de inúmeras reações enzimáticas, por

exemplo. Por fim, esta teoria postula que o elevado grau de especiação de metabólitos secundários confere não somente às plantas, mas também aos insetos, uma maior biodiversidade.

O referido artigo de Otto Gotlieb (1992) traz a influência de conceitos da “Taxonomia Numérica” de Sokal (1973), assim definida no livro “*Systematics, Evolution, and Biogeography of Compositae*”, de Funk (2009):

““*Taxonomia numérica mede o grau de alteração no estado de oxidação de “micromoléculas” (metabólitos secundários), indicando uma tendência no processo evolucionário que pode ser expressado como índices numéricos*””

Em relação à quimiotaxonomia, este livro traz a ideia geral de que a especialização que ocorre em estruturas químicas por meio de um crescente grau de oxidação, bem como a simplificação das cadeias carbônicas, acompanha a evolução das plantas. Além disso, sabe-se que muitas evoluções estão inter-correlacionadas entre diversos seres vivos, como plantas e insetos, e, deste modo, a quimiotaxonomia poderia ser explicada do ponto de vista prático, visto que ela pode evidenciar o relacionamento causa-consequência entre os metabólitos e o meio ambiente.

Mais precisamente em relação à caracterização química do gênero *Viguiera*, a grande maioria das substâncias isoladas tem sido as lactonas sesquiterpênicas (LST) da classe dos heliangolidos e seus derivados furanoeliangolidos e epoxieliangolidos, típicos do gênero (Da Costa, 2008). Demais caracterizações químicas de várias espécies do gênero foram relatadas, observando-se a ocorrência de LST dos tipos guaianolido e germacrolido (Schorr *et al.*, 2002), diterpenos, como os ácidos *ent*-caurenóico e *ent*-pimaradienóico, bem como os seus derivados (Da Costa *et al.*, 1996, Ambrosio *et al.*, 2004), e ainda flavonoides, poucas cumarinas e um novo derivado do mioinositol, fato único no gênero (Spring *et al.*, 2001). Por outro lado, é rara a ocorrência de triterpenos, sendo que há poucos estudos com óleos essenciais e seus mono- e sesquiterpenos. Portanto, de tal maneira, a caracterização química das plantas é útil para o auxílio da taxonomia e filogenia, tal como pode ser observado em trabalhos realizados anteriormente com o gênero *Helianthus*, onde os autores propuseram inclusive alterações taxonômicas e de parentesco (Spring *et al.*, 1989, 1990, 1991).

Além das contribuições quimiotaxonômicas, vale mencionar que muitos dos metabólitos secundários descritos acima são bioativos e, portanto, factíveis de serem

estudados do ponto de vista terapêutico (Verpoort, 1998; Ngo, 2013) e toxicológico (Schmidt, 199). Mais precisamente em relação ao potencial farmacológico do gênero *Aldama*, há diversos estudos que demonstram a importância deste gênero. Como exemplo, há um estudo muito interessante que avalia a atividade tripanomicida de diterpenos derivados do ácido *ent*-caurenóico (Da Costa *et al.*, 1996). Outro trabalho revelou a ação de diterpenos derivados da classe dos *ent*-pimaranos na ação inibidora da contratilidade da carótida de ratos (Ambrósio *et al.*, 2002; Tirapelli *et al.*, 2004). Outros trabalhos demonstraram a atividade antimicrobial do óleo essencial de *V. dentata* (Caneles *et al.*, 2008), bem como a atividade anticancerígena em ratos, desempenhada por LST isoladas de *A. sylvatica* (Taylor *et al.*, 2008).

Por fim, há artigos que demonstraram a importância dos metabólitos secundários em processos anti-inflamatórios (Gobbo-Neto *et al.*, 2005; Schorr *et al.*, 2005), alguns específicos ao gênero *Aldama* (Schorr *et al.*, 2002; Siedle *et al.*, 2004; Nicolette *et al.*, 2009). Recentemente também têm sido muito explorados os compostos fenólicos derivados dos ácidos cafeoilquínico e clorogênico, no que diz respeito à ação anti-inflamatória (Chagas-Paula *et al.*, 2011).

Com relação ao estudo das propriedades anti-inflamatórias de *Aldama* é importante demonstrar alguns dos processos bioquímicos que desencadeiam a inflamação. Sabe-se que há duas enzimas muito importantes envolvidas nos processos inflamatórios: as ciclooxigenases (COXs) e as lipoxigenases (LOXs), responsáveis por oxidar o ácido araquidônico, resultando na produção de prostanoídes (prostaglandinas, prostaciclina e tromboxanos) na via COX-dependente, ou leucotrienos, na via LOX-dependente (Rang e Dale, 2004). As prostaglandinas contribuem para os sintomas da inflamação aguda, tais como aumento da permeabilidade vascular, vasodilatação, formação de edema, produção de dor e febre. Já os leucotrienos são potentes quimiotáticos e agentes broncoconstritores, e desta forma estimulam a infiltração de leucócitos, contribuindo para a ulceração por oclusão de micro vasos, bem como provocando crise de asma e conversão da inflamação aguda para crônica (Rang e Dale, 2004; Parente, 2001). Os anti-inflamatórios não esteroidais (AINEs) atuais são capazes de inibir as COXs e não são capazes de inibir as LOXs (Parente, 2001; Rang e Dale, 2004). Assim, a via de metabolização do ácido araquidônico é deslocada para ser metabolizado apenas pelas LOXs, resultando maior formação de leucotrienos, aumentando os efeitos adversos dos AINEs (Rang e Dale, 2004; Parente, 2001; Fiorucci *et al.*, 2001). Então, recentemente destacou-se a importância de procurar substâncias que sejam

capazes de inibir tanto as COXs quanto as LOXs (Parente, 2001; Fiorucci *et al.*, 2001; Charlier, 2003).

Deste modo, cabe aqui explicar o conceito de metabolômica e como ela pode nos auxiliar tanto na obtenção de dados químicos para a quimiotaxonomia como para realizar estudos de bioatividade. Metabolômica, por definição, seria a detecção de todos os metabólitos de um dado organismo, assim como genômica é o estudo do genoma completo de um dado organismo, e proteômica é o estudo de todas as proteínas e suas isoformas. O primeiro conceito de metaboloma introduzido (Oliver *et al.*, 1998) menciona a necessidade de se fazer conexão entre metabolômica e genômica, observando-se as alterações de concentração do metabólito como um resultado da deleção ou expressão de algum gene. Logo em seguida (Fiehn *et al.*, 2001 e 2002) a metabolômica é descrita como sendo “uma abrangente e quantitativa análise de todos os metabólitos”, sendo que tal abordagem revelaria o metaboloma de um sistema biológico e nos daria uma maior compreensão da genômica funcional. Já Goodacre e seus colaboradores (Goodacre *et al.*, 2004) pela primeira vez listaram diferentes abordagens de metabolômica, embora concordem que ela deva ser quantitativa e complementar à transcriptômica e genômica. As diferentes abordagens definidas foram: *metabolite target analysis* (análise direcionada ao metabólito), *metabolite profiling* (perfil metabólito), *metabolomics* (metabolômica), *metabolic fingerprint* (impressão digital metabólica), *metabolic profiling* (perfil químico metabólico) e *metabonomics* (metabonomica). Porém, muitos destes conceitos já se alteraram desde 2004 e outros caíram em desuso. Por exemplo, a metabonomica, a qual fora colocada como sendo uma medida das alterações nos níveis dos metabólitos causados por perturbações causadas por drogas e toxinas, não tem sido mais recomendada. Metabolômica é colocada como uma ciência e não mais como uma abordagem, visto que ainda é impraticável a detecção e quantificação de todos os metabólitos. Assim sendo, os conceitos mais modernos do que seria metabolômica as descrevem como uma nova área da ciência, em vez de uma abordagem analítica (Villas-Bôas *et al.*, 2005). Definem ainda que metabolômica é a caracterização do fenótipo metabólico (metaboloma) sob específicas condições (estágios de desenvolvimento do organismo, condições ambientais e genéticas, etc.) e a correlação ao seu correspondente genótipo (Villas-Bôas *et al.*, 2005 e 2006; Summer *et al.*, 2003; Schwab *et al.*, 2003). Este fenótipo metabólico pode ser caracterizado sob as diferentes abordagens, onde a técnica do *metabolite profiling* (perfil metabólico) pode ser direcionada ou não. Nesta abordagem, objetiva-se a detecção dos metabólitos por um conjunto de técnicas analíticas específicas

como, por exemplo, técnicas hifenadas baseadas em cromatografia e espectroscopia (Villas-Bôas *et al.*, 2005 e 2006).

A metabolômica não-direcionada (*untargeted metabolomics*), que seria a mesma coisa que perfil metabólico não-direcionado, é uma técnica que visa à obtenção dos dados químicos de todo e qualquer metabólito, sem especificidade, por exemplo primário ou secundário. Metabolômica direcionada é quando, já de antemão, se propõe detectar determinada classe específica de um metabólito, ou uma substância química (ou derivados) em específico. Assim sendo, a extração fitoquímica de todo e qualquer metabólito pode ser considerada uma metabolômica não-direcionada, pois, embora o objetivo seja a detecção de metabólitos secundários, não se direciona a técnica para uma determinada classe química, como flavonoides, por exemplo.

Por fim, ainda no que diz respeito à metabolômica, é necessário descrever aqui conceitos que trazem uma relação direta entre as diferentes abordagens da metabolômica e a identificação e quantificação dos metabólitos, que por sua vez são inversamente proporcionais ao número de metabólitos detectados (Krastanov, 2010). Assim sendo, quanto mais conhecimento se tem sobre os tipos de metabólitos detectados, inclusive realizando a sua quantificação, mais a abordagem se aproxima de uma análise direcionada, ao passo que, quanto menos isso ocorre, mais a análise passa a ser classificada como de perfil metabólico ou até mesmo de impressão digital metabólica (Figura 3), a qual diz respeito apenas à aquisição dos espectros e cromatogramas dos perfis químicos. Estes conceitos não entram em conflito com o que fora exposto por Villas-Bôas, apenas descrevem melhor em termos quantitativos e qualitativos as abordagens metabolômicas.

Figure 3 - Tipos de análises metabolômicas (Retirado de Krastanov, 2010; pág 1538).

Entretanto, devido à grande quantidade de dados químicos gerados pela abordagem da metabolômica não-direcionada, torna-se demasiado complexa e demorada a análise dos mesmos. Para tanto, é necessário a utilização de análise estatística multivariada supervisionada e não supervisionada, as quais estão inseridas na quimiometria, ramo da ciência que trata da aplicação de métodos estatísticos e matemáticos aos dados químicos, extraindo-se informações químicas de tais dados (Gasteiger e Engel, 2003). Os dados químicos são usualmente multidimensionais e, neste caso, são chamados de “multivariados”. Os diferentes métodos matemáticos e estatísticos aplicados aos dados químicos 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), 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 segue uma tendência (um direcionamento) previamente informada à análise, por exemplo, as amostras se agrupam de acordo com uma específica informação fornecida à análise, a qual pode ser a bioatividade ou *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, ou seja, é justamente nesta direção que apenas uma componente principal teria maior precisão em explicar tudo aquilo que somente seria possível se observássemos o espaço multidimensional. Entretanto, embora a primeira componente sempre explique a maior variância encontrada, sempre é necessário mais de uma para explicar um espaço multidimensional. Assim sendo, uma segunda componente, ortogonal à primeira, explica uma maior quantidade de variância, menor que a primeira componente, mas maior que a terceira componente, à qual é ortogonal às duas primeiras. Quando 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, onde é realizada uma separação das variações da matriz X em duas partes, uma linearmente relacionada à matriz Y (componentes preditivas) e outra que é ortogonal à Y (não preditiva com a resposta Y) (Trygg *et al.*, 2002; Verron *et al.*, 2004; Manual SIMCA 13.0.3). 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 trata-se 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*), 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.

A quimiometria não se ocupa apenas da análise dos dados que visa gerar informação, mas também do processamento e tratamento dos mesmos, os quais são etapas anteriores às análises e cruciais na geração de adequados dados químicos. Por exemplo, os dados provenientes de cromatogramas e espectros geralmente necessitam de suavização da linha de base, alinhamento, normalização, entre outras técnicas de tratamento, visando a adequação dos dados químicos à posterior análise e geração de informação. Desta forma, para os dados químicos provenientes do LC-MS há diversos *softwares* de processamento de dados, dentre os quais daremos destaque ao *MZmine* (Katajamma *et al.*, 2006; Pluskal *et al.*, 2010) e *MetAlign* (Lommen, 2009; Lommen e Kools, 2012), entre outros, sendo que cada qual traz sua vantagem e desvantagem, como apresentado na Figura 4 (Castillo *et al.*, 2011).

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	✓	✓	-	✓	✓	✓	✓	-	-

Figure 4 - Comparação de diferentes softwares de processamento de dados (Extraído de Castillo *et al.*, 2011).

Os *softwares* *MZmine* e *MetAlign* são gratuitos, sendo que este último pode ainda ser utilizado em conjunto com o *software* *MSClust* (Tikunov *et al.*, 2012), o qual traz algoritmos capazes de reduzir a quantidade de sinais de um espectrômetro de massas, oriundos de fragmentos moleculares, adutos e isótopos. Assim sendo, este *software* traz a promessa de agrupar os sinais em supostos metabólitos íntegros, podendo-se fazer identificação *a posteriori*.

Por fim, é necessário introduzir aqui os conceitos de *Data Mining* (Mineração de Dados) e de como ela tem sido utilizada pela quimiometria recentemente (Buydens *et al.*, 1999; Mutihac, L. e Mutihac R., 2008), visto que a *Decision Tree*, uma de suas técnicas, é utilizada neste trabalho. Basicamente, os significados de *Data Mining* não diferem muito dos de quimiometria. *Data Mining* trata da resolução de problemas presentes em bancos de dados (os quais podem ser químicos), explorando-os por meio de técnicas computacionais, visando o reconhecimento de padrões e/ou tendências e classificações. Tais padrões nos permitem realizar previsões 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 Trees* (Árvores de Decisão). *Decision Tree* é uma representação gráfica proveniente de “*divide-and-conquer*” (dividir e conquistar), um algoritmo que trabalha subdividindo problemas maiores em problemas menores que possam ser resolvidos (Quinlan, 1986 e 1987). Em uma *Decision Tree*, os “nós” da árvore representam um determinado atributo (uma variável) a ser testado por condições específicas (concentração, por exemplo). Tais testes fornecem as “folhas” da árvore, as quais contêm a classificação do teste (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).

Por fim, pode-se destacar aqui trabalhos (Alvarenga *et al.*, 2001; Emerenciano *et al.*, 2006) que se utilizam da quimiometria para explicar a quimiotaxonomia de Asteraceae, descrevendo muito apropriadamente os metabólitos mais específicos de cada subtribo e gênero, sendo que o segundo trabalho traz aplicações diretas sobre a teoria redox de Gottlieb, mencionada anteriormente. Entretanto, estes trabalhos foram realizados com dados químicos provenientes da literatura e, portanto, se limitam apenas aos metabólitos isolados e identificados pela fitoquímica clássica. Aqui, nesta Dissertação, menciona-se a aplicação de

quimiometria aos dados químicos provenientes de uma abordagem metabolômica não-direcionada.

Portanto, a quimiometria aplicada à metabolômica não-direcionada auxilia nos estudos quimiotaxonômicos do gênero, fundamentando-se no fato de que algumas classes de substâncias são específicas de alguns táxons (Sharma, 2007), possibilitando-nos fazer a desreplicação de substâncias discriminantes de determinados gêneros ou grupos infragenéricos. Além disso, esta combinação de técnicas nos permite direcionar a fitoquímica para metabólitos que seriam potencialmente ativos frente a um determinado alvo molecular ou enzimático, sendo que neste trabalho os alvos referem-se às enzimas anti-inflamatórias COX-1 e 5-LOX.

Conclusão

Neste trabalho foram abordados dois assuntos distintos: quimiotaonomia e inibição das enzimas pró-inflamatórias COX-1 e 5-LOX, os quais foram analisados com dados químicos de *Aldama* La Llave provenientes da abordagem metabolômica não-direcionada e tratados por meio de análises quimiométricas. Deste modo, esta conclusão pode-se subdividir, do ponto de vista lógico, em dois grandes tópicos, cada qual tratando-se de um assunto diferente.

Primeiramente, quanto às contribuições quimiotaonômicas, tendo em vista a metodologia empregada, foi possível verificar que o gênero *Aldama* La Llave é quimicamente consistente, formando *clusters* bem definidos em relação aos demais gêneros analisados. Tal fato foi mais observado por meio da HCA que se utilizou de dados provenientes dos modos de ionização positivo e negativo concomitantemente. A HCA pode explicar alguns agrupamentos infragenéricos para determinadas espécies que apresentam problemas de classificação taxonômica com base em dados de morfoanatomia. Futuramente, esta HCA poderá ser comparada com os dados de filogenia baseada em dados moleculares. Além disso, pouca correlação pôde-se observar entre os dados químicos e a classificação ambiental de diversas regiões do Brasil, bem como a classificação de Blake (1918), por meio de todas as PCA. Entretanto, diversos modelos de predição com base em OPLS-DA foram apresentados nesta Dissertação, revelando o grande potencial desta técnica em predizer amostras desconhecidas ou externas ao grupo de estudo. Deste modo, foi gerada uma lista de picos de íons, detectados pelo UHPLC-Orbitrap, como sendo discriminantes de *Aldama* La Llave. Tais picos foram correlacionados com um banco de massas criado para substâncias químicas específicas do gênero *Aldama* e outra proveniente de Asteraceae, por meio dos quais foi possível encontrar 37 prováveis substâncias químicas. Contudo, tais substâncias químicas não puderam ser desreplicadas com base no perfil de UV e padrão de fragmentação, visto que muitas das quais ou estavam em baixa intensidade no espectro, ou possuíam um confuso padrão de fragmentação visualizado pelo espectro de MS-MS. Foi comentado que tal dificuldade da análise de fragmentação possa se dever ao fato do Orbitrap Exactive Plus não conter um analisador Quadrupolo, o qual poderia selecionar íons específicos para a desreplicação, ou ainda o fato do UHPLC não estar conseguindo separar bem os picos de acordo com a análise empregada. Ainda assim, em um futuro próximo, tais substâncias terão

seus tempos de retenção informados por meio de modelos experimentais de predição o que, certamente, favorecerá uma desreplicação mais inequívoca destas substâncias.

Quanto às análises realizadas para predição de prováveis substâncias químicas responsáveis pela inibição dupla das enzimas pro-inflamatórias COX-1 e 5-LOX, primeiramente foi listado que apenas *A. trichophylla* e *A. robusta* apresentaram a inibição dupla desejada. A primeira espécie foi escolhida para a fitoquímica direcionada, uma vez que ela é menos estudada do ponto de vista fitoquímico. Para o intuito de se saber quais seriam os prováveis picos responsáveis pela inibição dupla e, presentes em *A. trichophylla*, uma série de modelos experimentais com base nos diferentes analisadores TOF e Orbitrap foram criados empregando-se diferentes combinações de algoritmos do *software* *MZmine*, bem como os *softwares* *MetAlign* e *MSClust*. Tais comparações são inéditas na literatura para o propósito aqui abordado. Assim sendo, dentro do material e das técnicas utilizadas, foi constatado que os algoritmos do *MZmine* tinham pouca influência na melhora do modelo experimental, assim como os *softwares* *MetAlign* e *MSClust*. Entretanto, de uma maneira geral, os dados químicos provenientes do modo de ionização negativo pareceram um pouco melhores na criação dos modelos de inibição enzimática. Ainda assim, foram escolhidos seis melhores modelos (três provenientes de dados do TOF e três do Orbitrap) para serem otimizados por meio da deleção de variáveis não importantes na projeção ($VIP < 1$). Tais modelos também foram rodados no *software* Weka, por meio do qual foi realizada a *Decision Tree*, informando quais seriam os picos mais importantes na classificação do modelo. Com base nisso, os modelos otimizados e a *Decision Tree* forneceram uma série de prováveis substâncias responsáveis pela inibição dupla das enzimas COX-1 e 5-LOX, resultando em apenas três substâncias desreplicadas com base no perfil de UV e padrões de fragmentação visualizados por meio de análises realizadas no HCD-Orbitrap e Ion-Trap: kaempferol-3-*O*-glucuronideo, quercetina-3-*O*-metil-7-glucuronideo e kaempferol-3-*O*-(6"-malonil-glucosideo). Estas três substâncias foram isoladas de frações em coluna de *Sephadex LH-20* da fração AcOEt de *A. trichophylla* e uma prévia análise pelo RMN 1H foi realizada, indicando que, possivelmente, foi obtido sucesso no isolamento. Por fim, a análise de elucidação estrutural será melhor concluída e as substâncias poderão ter o seu potencial de inibição testado por meio das análises *in vitro* com COX-1 e 5-LOX.

Referências

- 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*, v. 56, p. 27-37, 2001.
- Alvares, C.A.; Stape, J.L.; Sentelhas, P.C.; de Moraes, G.; Leonardo, J.; Sparovek, G. Köppen's climate classification map for Brazil. *Meteorologische Zeitschrift*, v. 22, p. 711-728, 2013.
- Ambrosio, S.R.; Schorr, K.; Da Costa, F.B. Terpenoids of *Viguiera arenaria* (Asteraceae). *Biochemical Systematics and Ecology*, v. 32, p. 221–224, 2004.
- Ambrosio, S.R.; Tirapelli, C.R.; Bonaventura, D.; Oliveira, A.M.; Da Costa, F.B. Pimarane diterpene from *Viguiera arenaria* (Asteraceae) inhibit rat carotid contraction. *Fitoterapia*, v. 73, p. 484–489, 2002.
- Bhargava, N.; Sharma, G.; Bhargava, R.; Mathuria, M. Decision tree analysis on J48 algorithm for data mining, v. 3, p. 1114-1119, 2013.
- BLAKE, S.F. A revision of the genus *Viguiera*. *Contributions of the Gray Herbarium*, v. 54, p. 1-205, 1918.
- Blaker, J.G. Compositae IV: Helianthoideae. In C.F.P Martius & A.W. Eichler (eds) *Flora brasiliensis Monacchii*. Typografia Regia, v. 6, p. 135-268, 1884.
- 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*, v. 49, p. 121-133, 1999.
- Canales, M. *et al.* Antimicrobial Activity of the Extracts and Essential Oil of *Viguiera dentata*. *Pharmaceutical Biology*, v. 46, p. 719–723, 2008.

Castillo, S.; Gopalacharyulu, P.; Yetukuri, L.; Orešič, M. Algorithms and tools for the preprocessing of LC–MS metabolomics data. *Chemometrics and Intelligent Laboratory Systems*, v. 108, p. 23–32, 2011.

Cazarolli, L.H. *et al.* Flavonoids: prospective drug candidates. *Mini-Reviews in Medicinal Chemistry*, v. 8, p. 1429-1440, 2008.

Chagas-Paula, D. A. *et al.* Chlorogenic acids from *Tithonia diversifolia* demonstrate better anti-inflammatory effect than indomethacin and its sesquiterpene lactones. *Journal of ethnopharmacology*, v. 136, p. 355–362, 2011.

Charlier, C. & Michaux, C. Dual inhibition of cyclooxygenase-2 (COX-2) and 5-lipoxygenase (5-LOX) as a new strategy to provide safer non-steroidal anti-inflammatory drugs. *European Journal of Medicinal Chemistry*, v. 38, p. 645–659, 2003.

ChemAxon – Cheminformatic Platforms and Desktop. Disponível em: www.chemaxon.com.

Clifford, M.N.; Johnston, S.K.; Knight, S.; Kuhnert, N. Hierarchical scheme for LC-MSⁿ identification of chlorogenic acids. *Journal Agricultural and Food Chemistry*, v. 51, p. 2900-2911.

Clifford, M.N.; Knight, S.; Kuhnert, N. Discriminating between the six isomers of dicaffeoylquinic acid by LC-MSⁿ. *Journal of Agricultural and Food Chemistry*, v. 53, p. 3821-3832.

Da Costa, F. B.; Albuquerque, S.; Vichnewski, W. Diterpenes and synthetic derivatives from *Viguiera aspillioides* with trypanomicidal activity. *Planta medica*, v. 62, p. 557–559, 1996.

Da Costa, F.B. *Asteraceae: Química e Potencial Farmacológico*. 244. Texto sistematizado para livre docência. Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2008.

Dictionary of Natural Products. Disponível em: <<http://dnp.chemnetbase.com/intro/>>.

Domon, B. & Costello, C. A systematic nomenclature for carbohydrate fragmentation in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate Journal*, v. 5, p. 397-409, 1988.

Emerenciano, V.P.; Cabrol-Bass, D.; Ferreira, M.J.P.; Alvarenga, S.A.V.; Brant, A.J.C.; Scotti, M.T.; Barbosa, K.O. Chemical evolution in the Asteraceae. The oxidation-reduction mechanism and production of secondary metabolites. *Natural Product Communications*, v. 1, p. 495-507, 2006.

Equipe Científica AsterbioBioChem. Disponível em: <www.asterbiochem.org>.

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*, v. 108, p. 196-207, 2014.

Fiehn, O. Combining genomics, metabolome analysis, and biochemical modelling to understand metabolic networks. *Comparative and functional genomics*, v. 2, p. 155–168, 2001.

Fiehn, O. Metabolomics-the link between genotypes and phenotypes. *Plant molecular biology*, v. 48, p. 155–71, 2002.

Fiorucci, S.; Meli, R.; Bucci, M.; Cirino, G. Dual inhibitors of cyclooxygenase and 5-lipoxygenase. A new avenue in anti-inflammatory therapy? *Biochemical pharmacology*, v. 62, p. 1433–1438, 2001.

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

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

Gobbo-Neto, L. *et al.* Evaluation of the anti-inflammatory and antioxidant activities of di-C-glucoflavones from *Lychnophora ericoides* (Asteraceae). *Planta medica*, v. 71, p. 3–6, 2005.

Goodacre, R.; Vaidyanathan, S.; Dunn, W.B.; Harrigan, G.G.; Kell, D.B. Trend in Biotechnology, v. 22, p. 245-252. 2004.

Gottlieb, O.R. Biodiversidade: Uma teoria molecular. *Química Nova*, v. 15, p. 167-172, 1992.

Gottlieb, O.R. & Borin, M.R.M.B. Químico-Biologia quantitativa: Um novo paradigma? Química Nova, v. 35, p. 2105-2114, 2012.

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

Henchoz, Y.; Guillaume, D.; Martel, S.; Rudaz, S.; Veuthey, J-L.; Carrupt, P-A. Fast log P determination by ultra-high-pressure-liquid chromatography coupled with UV and mass spectrometry detections. Analytical and Bioanalytical Chemistry, v. 394, p. 1919-1930, 2009.

Holčapek, M.; Jirásko, R.; Lísa, M. Recent developments in liquid chromatography-mass spectrometry and related techniques. Journal of chromatography. A, v. 1259, p. 3–15, 2012.

Hu, Q.; Noll, R.J.; Li, H.; Makarov, A.; Hardman, M.; Cooks, G. The Orbitrap: a new mass spectrometer, v. 40, p. 430-443, 2005.

Kachlicki, P.; Einhorn, J.; Muth, D.; Kerhoas, L.; Stobiecki, M. Evaluation of glycosylation and malonylation patterns in flavonoid glycosides during LC/MS/MS metabolite profiling. Journal of Mass Spectrometry, v. 43, p. 572-586, 2008.

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

Kim, H.P.; Son, K.H.; Chang, H.W.; Kang, S.S. Anti-inflammatory plant flavonoids and cellular action mechanisms. Journal of Pharmacological Sciences, v. 96, p. 229-245, 2004.

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

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

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

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.

Magenta, M.A.G & Pirani, J.R. Novidades taxonômicas em *Aldama* (Asteraceae-Heliantheae). *Rodriguesia*, v. 65, p. 175-192, 2014.

Markham, K.R. Techniques of flavonoid identification. London: Academic press, v. 31, 1982.

Mutihac, L. & Mutihac, R. Mining in chemometrics. *Analytic Chimica ACTA*, v. 612, p. 1-18, 2008.

Ngo, L.T.; Okogun, J.I.; & Folk, W.R. 21St Century Natural Product Research and Drug Development and Traditional Medicines. *Natural product reports*, v. 30, p. 584–92, 2013.

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

Oliver, S.G.; Wilson, M.K.; Kell, D.B.; Baganz, F.; Systematic functional analysis of the yeast genome. *Focus*, v. 16, p. 373-378, 1998.

Parente, L. Pros and Cons of Selective Inhibition of Cyclooxygenase-2 versus Dual Lipoxygenase / Cyclooxygenase Inhibition : Is Two Better than One ? *The Journal of Rheumatology*, v. 28, p. 2375–2382, 2001.

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*, v. 11, p. 395–406, 2010.

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

Quinlan, J.R. Induction of decision trees. *Machine Learning*, v. 1, p. 81-106, 1986.

Quinlan, J.R. Simplifying decision trees. *International Journal Man-Machine Studies*, v. 27, p. 221-234, 1987.

Rang, H.P.; Dale, M.M.; Ritter, J.M.; Moore, P.K. Farmacologia. Elsevier Editora Ltda, Rio de Janeiro. Tradução da 5ª Edição Americana, 2004.

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

Robinson, H.; Moore, A.J. New species and new combinations in *Rhysolepis* (Heliantheae: Asteraceae). Proceedings of Bioolgical Society of Washington, v. 117, p. 423-446, 2004.

Sakakibara, H.; Honda, Y.; Nakagawa, S.; Ashida, H.; Kanazawa, K. Simultaneous determination of all polyphenols in vetables, fruits, and teas. Journal of Agricultural and Food Chemistry, v. 51, p. 571-581.

Schilling, E. E. & Panero, J. L. 2011. A revised classification of subtribe Helianthinae (Asteraceae: Heliantheae) II. Derived lineages. Botanical Journal of the Linnean Society, v. 167, p. 311–331.

Schilling, E.E. & Jansen, R.K. 2012. Restriction Fragment Analysis of Chloroplast DNA and the Systematics of *Viguiera* and Related Genera (Asteraceae: Heliantheae). American Journal of Botany, v. 76, p. 1769–1778.

Schilling, E.E. & Panero, J.L. 2002. A revised classification of subtribe Helianthinae (Asteraceae: Heliantheae). I. Basal lineages. Botanical Journal of the Linnean Society, v. 140, p. 65–76.

Schilling, E.E.; Da Costa, F.B.; Lopes, N.P.; Heise, P.J. Brazilian species of *Viguiera* (Asteraceae) exhibit low levels of ITS sequence variation. Edinburgh Journal of Botany, v. 57, p. 323–332, 2000.

Schmidt, T.J. Toxic activities of sesquiterpene lactones: structural and biological aspectes. Current Organic Chemistry, v. 3, p. 577–608, 1999.

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, v. 60, p. 733–40, 2002.

Schorr, K.; Rott, A.; Da Costa, F.B. Merfort, I. Optimisation of a human neutrophil elastase assay and investigation of the effect of sesquiterpene lactones. *Biologicals*, v. 33, 175–184, 2005.

Schwab, W. Metabolome diversity: too few genes, too many metabolites? *Phytochemistry*, v. 62, p. 837–849, 2003.

SciFinder ScholarTM, version 2014. American Chemical Society: Columbus, Ohio, EUA.

Disponível em: <<https://scifinder.cas.org>>.

Seaman, F.C. Sesquiterpene lactones as taxonomic characters in the Asteraceae. *The Botanic Review*, v. 48, p. 121-529, 1982

Sharma, O.P. Plant taxonomy. Tata McGraw-Hill Publishing Company Limited, New Delhi, 7ed, p. 95, 2007.

Siedle, B. *et al.* Quantitative structure-activity relationship of sesquiterpene lactones as inhibitors of the transcription factor NF-kappaB. *Journal of medicinal chemistry*. v. 47, p. 6042–6054, 2004.

Sokal, R.R. Numerical Taxonomy. *Scientific American*. v. 215, p. 106-116, 1966.

Spring, O. & Schilling, E.E. The sesquiterpene lactone chemistry of *Helianthus* sect. *Atrorubentes* (Asteraceae: Heliantheae). *Biochemical Systematics and Ecology*, v. 19, p. 59–79, 1991.

Spring, O. & Schillingt, E.E. Chemosystematic Investigation of the Annual Species of *Helianthus* (Asteraceae)*. *Biochemical Systematics and Ecology*, v. 17, p. 519–528, 1989.

Spring, O. & Schillingt, E.E. The Sesquiterpene Lactone Chemistry of *Helianthus* Sect . *Ciliares* (Asteraceae: Heliantheae)*. *Biochemical Systematics and Ecology*, v. 18, p. 139–143, 1990.

Spring, O.; Zipper, R.; Reeb, S.; Vogler, B.; Da Costa, F.B. Sesquiterpene lactones and a myoinositol from glandular trichomes of *Viguiera quinqueremis* (Heliantheae; Asteraceae). *Phytochemistry*, v. 57, p. 267–72, 2001.

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

Taylor, P.G.; Dupuy Loo, O.A.; Bonilla, J.A.; Murillo, R. Anticancer activities of two sesquiterpene lactones, millerenolide and thieleanin isolated from *Viguiera sylvatica* and *Decachaeta thieleana*. *Fitoterapia*, v. 79, p. 428–32, 2008.

Tellinghuisen, J. & Bolster, C.H. Using R^2 to compare least-squares fit models: When it must fail. *Chemometrics and Intelligent Laboratory Systems*, v. 105, p. 220-222, 2011.

Tikunov, Y.M.; Laptinok, S.; Hall, R.D.; Bovy, A.; de Vos, R.C.H. MSClust: a tool for unsupervised mass spectra extraction of chromatography mass spectrometry ion-wise aligned data. *Metabolomics*, v. 8, 714-718, 2012.

Tirapelli, C.R.; Ambrosio, S.R.; Da Costa, F.B.; De Oliveira, A.M. Evidence for the mechanisms underlying the effects of pimaradienoic acid isolated from the roots of *Viguiera arenaria* on rat aorta. *Pharmacology*, v. 70, p. 31–38, 2004.

Tsimogiannis, D.; Samiotaki, M.; Panayotou, G.; Oreopoulou, V. Characterization of flavonoid subgroups and hydroxyl substitution by HPLC-MS/MS. *Molecules*, v. 12, p. 593-606, 2007.

Tuane, O.; Bombo, A.B.; Apezato-Da-Gloria, B. Anatomy of vegetative organs with an emphasis on the secretory structures of two species of *Aldama* (Asteraceae-Heliantheae). *Botany*, v. 91, p. 335-342, 2013.

Tuñón, M.J.; García-Mediavilla, M.V.; Sánchez-Campos, S.; Gonzáles-Gallego, J. Potential of flavonoids as anti-inflammatory agentes: modulation of pro-inflammatory gene expression and signal transduction pathways. *Current Drug Metabolism*, v. 10, p. 256-271, 2009.

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

Verpoorte, R. Exploration of nature's chemodiversity: the role of secondary metabolites as leads in drug development. *Drug Discovery Today*, v. 3, p. 232–238, 1998.

Vesscchi, R.; Lopes, N.P.; Gozzo, F.C.; Dörr, F.A.; Murgu, M.; Lebre, D.T.; Abreu, R.; Bustillos, O.V.; Riveros, J.M. Nomenclaturas de espectrometria de massas em língua portuguesa. *Química Nova*, v. 34, p. 1875-1887, 2011.

Villas-Bôas, S.G. & Gombert, A.K. Análise do Metaboloma. *Biotecnologia Ciência e Desenvolvimento*, v. 36, p. 58–69, 2006.

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

Vukics, V. & Guttman, A. Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. *Mass Spectrometry Reviews*, v. 29, p. 1-16, 2008.

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

Wolfgang, P. & Steinwedel, H. Ein neues massenspektrometer ohne magnetfeld. *Zeitschrift Naturforschung Teil A*, v. 8, p. 448, 1953.

Xie, Cen *et al.* Investigation of isomeric transformations of chlorogenic acid in buffers and biological matrixes by ultraperformance liquid chromatography coupled with hybrid quadrupole/ion mobility/orthogonal acceleration time-of-flight mass spectrometry. *Journal of agricultural and food chemistry*, v. 59, p. 11078-11087, 2011

Yuliana, N. D.; Khatib, A.; Verpoorte, R.; Choi, Y.H. Comprehensive extraction method integrated with NMR metabolomics: a new bioactivity screening method for plants, adenosine A1 receptor binding compounds in *Orthosiphon stamineus* Benth. *Analytical Chemistry*, v. 83, p. 6902–6906, 2011.

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

Zhou, B.; Xiao, J.F.; Tuli, L.; Ransom, H.W. LC-MS-based metabolomics. *Molecular Systems*, v. 8, p. 470-481, 2012.