UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA

**FLAVONÓIDES, CUMARINAS E BENZOFURANOS COMO
MARCADORES QUIMIOTAXONÔMICOS EM ASTERACEAE**

Antônio Jedson Caldeira Brant

Dissertação de Mestrado

Prof. Dr. Vicente de Paulo Emerenciano

Orientador

São Paulo

2003

“Flavonóides, Cumarinas e Benzofuranos como Marcadores Quimiotaxonômicos em Asteraceae”

**ANTÔNIO JEDSON CALDEIRA
BRANT**

Dissertação de Mestrado submetida ao Instituto de Química da
Universidade de São Paulo como parte dos requisitos necessários à obtenção do
grau de Mestre em Química - Área: Química Orgânica.

Aprovado por:

Prof. Dr. VICENTE DE PAULO EMERENCIANO
IQ – USP
(Orientador e Presidente)

Prof. Dr. MASSAYOSHI YOSHIDA
IQ – USP

Prof. Dr. FERNANDO BATISTA DA COSTA
FCF – RP – USP

SÃO PAULO
18 DE FEVEREIRO 2003.

DEDICATÓRIAS

A minha mãe, Antônia Zumak Brant (*in memoriam*), que foi um exemplo de espírito de luta, persistência e humildade. Ela é para mim uma eterna saudade.

A Esther Trindade Ramos (*in memoriam*), como uma segunda mãe, sempre trocando comigo correspondências criativas e alegres, em prosa e verso, ao longo de muitos anos. Apelidei-a de poetisa de Ouro Fino, onde nasceu.

Aos meus Irmãos, com carinho fraternal:

Maria José Caldeira Brant

Antônio Edson Caldeira Brant (*in memoriam*)

Edna Brant Pessoa

Whitery Caldeira Brant

A Romilda Rachel Soares Silva, ex-professora de Química Geral e Inorgânica do Instituto de Ciências Exatas (ICEx) da UFMG, cujas aulas brilhantes e incentivo pessoal me motivaram a abraçar a Química como futura profissão. Além de grande mestra, Romilda era extremamente sensível a qualquer problema que um de seus alunos viesse a ter e fosse procurá-la para um conselho, diálogo. Ela sempre tinha uma palavra amiga e de ânimo. Sou-lhe muito grato e sempre seu admirador.

*"E o girassol vai girando dia afora...
O girassol é o carrossel das abelhas".
(Vinícius de Moraes)*

AGRADECIMENTOS

Ao Dr. Vicente de Paulo Emerenciano, pela acolhida em seu grupo de pesquisa, orientação para a realização deste trabalho, e pela amizade.

Ao Dr. João Pedro Simon Farah, pela ajuda e incentivo em estudos fundamentais de Estatística.

À Dra. Liliana Marzorati, pelo incentivo em dar continuidade a este trabalho futuramente.

Ao Dr. Luiz Henrique Catalani, por seu apoio todo especial em relação ao suporte parcial de meu trabalho pelo CNPq.

A todos os professores do IQ-USP, em especial Dra. Helena Maria Carvalho Ferraz, pelos ensinamentos valiosos e pela amizade.

Aos colegas de pesquisa, Marcelo José Pena Ferreira e Alessandra Rodrigues Rufino, além da amizade, pela paciência e empenho de ajudar-me no projeto diretamente na parte de digitação, desenho de estruturas químicas e diagramas, e passando-me também inúmeras dicas de uso correto de programas computacionais. Não posso deixar de mencionar também a nossa ex-estagiária, Tanja Schwabe, que me conferiu inúmeras tabelas e também desenhou algumas estruturas químicas relativas ao meu trabalho.

Ao Sr. José Rubens de Toledo Barros, cujo trabalho paciente de xerocopiar vasto acervo literário pertinente a Asteraceae foi de extrema valia para a busca de muitos dados necessários para a realização desta dissertação.

Ao Prof. Dr. Waldir Mantovani, do Departamento de Ecologia Geral do Instituto de Biociências-USP, pelos seus comentários valiosos e esclarecedores sobre diversos *clusters* obtidos ao longo de meu trabalho, que lhe apresentei para sanar algumas dúvidas de interpretação dos respectivos resultados.

Aos amigos e colegas de outros laboratórios do IQ-USP, pela atenção e amizade.

A todos os funcionários do IQ –USP, com gratidão especial a Cibeles Rosani Carlos (Secretaria de Pós-Graduação), pelo apoio e incentivo dispensado à minha pessoa.

RESUMO

Este trabalho descreve as ocorrências naturais de flavonóides, cumarinas e benzofuranos na família de plantas, Asteraceae / Compositae. Tem como objetivo principal classificar os referidos metabólitos secundários como possíveis marcadores quimiotaxonômicos da família bem como fazer previsões de suas ocorrências naquela. Para sua realização, foi montado um banco de dados de ocorrências desses produtos naturais a partir da literatura especializada. Os números de ocorrências das três classes químicas (flavonóides, 4369; cumarinas, 961; benzofuranos, 628) foram avaliados através de programas computacionais. Com estes foram feitas matrizes de correlações, regressões lineares múltiplas, e gráficos variados. Foram também correlacionados os números de oxidação (NOXs) médios de cada classe química (exceto os dos benzofuranos, devido a algumas dúvidas suscitadas de sua rota biogenética) com a finalidade de estudos de passos evolutivos de táxons. Os resultados do estudo apontam alguns esqueletos carbônicos de cumarinas e benzofuranos que podem provavelmente servir como marcadores químicos em nível de tribos e subtribos da família. Quanto aos flavonóides, já são considerados, em nível infrafamiliar, como marcadores químicos alguns tipos de esqueletos carbônicos com padrões de substituição mais específicos. Foram montadas algumas equações empíricas capazes de prever ocorrências e NOXs das classes químicas envolvidas neste trabalho para as tribos e subtribos da família, com um grau de confiabilidade de 95%. Para este estudo usou-se como base taxonômica a classificação de Asteraceae mais recente de Bremer (1994).

ABSTRACT

This work describes the natural occurrences of flavonoids, coumarins and benzofurans in the family of plants, Asteraceae / Compositae. It has as main purpose to classify the referred secondary metabolites as possible chemotaxonomic markers of the family as well as to do predictions of occurrences of the latter. For its accomplishment, a database of occurrences of those natural products was set up from the specialized literature. The numbers of occurrences of the three chemical classes (flavonoids, 4369; coumarins, 961; benzofurans, 628) were appraised with use of statistical computational programs. With aid of these were made correlation matrices of the occurrences, multiple linear regressions, numerous diverse diagrams. The average oxidation numbers (OXNs) of each chemical class (except the ones of the benzofurans, due to some doubts raised on their biogenetical route) were also correlated with the purpose of studies on evolutionary steps of the taxa. The results of the study point some carbon skeletons of coumarins and benzofurans that can probably serve as chemical markers at tribe and subtribe levels of the family. In relation to the flavonoids, these are already considered, at infrafamilial level, as chemical markers, especially some types of carbon skeletons bearing more specific substitution patterns. Were also mounted some empirical equations capable of predicting occurrences and average OXNs of the chemical classes involved in this work for the tribes and subtribes of the family, with a 95% reliability degree. For this work was also utilized as taxonomic base the most recent classification of the Asteraceae of Bremer (1994a).

SUMÁRIO

I - INTRODUÇÃO	1
1.01. Asteraceae (Compositae)	1
1.02. Dados de fósseis de plantas	3
1.03. Evolucionismo	7
1.04. Taxonomia e filogenia	8
1.05. Taxonomia de Asteraceae	9
1.06. Quimiotaxonomia	12
1.07. Nomenclatura de metabólitos secundários	14
1.08. Nome trivial e sua importância	14
1.09. Isótopos como marcadores de moléculas	15
1.10. Origem homóloga	16
1.11. Perfil químico de Asteraceae	17
II - TAXONOMIA DE ASTERACEAE	19
2.01. Taxonomia de Asteraceae segundo Bremer	19
2.02. Quimiotaxonomia de Asteraceae	24
2.03. Caracteres macromoleculares	25
III - METODOLOGIA	28
3.01. Dados de ocorrências	28
3.02. Estudo de passos oxidativos	30
3.03. Análise dos dados	32
IV - RESULTADOS E DISCUSSÃO	35
4.01. Terpenos	35
4.1.1. Monoterpenos	38
4.1.2. Sesquiterpenos	38
4.1.3. Diterpenos	39
4.1.4. Triterpenos	40
4.1.5. Rota biogenética geral dos terpenóides	40
4.1.5.1. Formação de policetídeos	40
4.1.5.2. Precursores	43
4.02. Poliacetilenos	44
4.03. Flavonóides	47

4.04. Cumarinas	53
4.05. Benzofuranos	57
4.06. Ocorrências de metabólitos secundários em tribos de Asteraceae	60
4.07. Regressão linear múltipla	71
4.08. Grupos de cumarinas classificadas por tipo de esqueleto carbônico em Asteraceae	91
4.08. Grupos de benzofuranos classificados por tipo de esqueleto carbônico em Asteraceae.....	109
V - CONCLUSÕES	147
VI - REFERÊNCIAS	150
APÊNDICE DE ESTATÍSTICA	156

SUMÁRIO DE FIGURAS

Figura 1. Girassol (<i>Helianthus annuus</i>)	1
Figura 2. Fotos de tipos de capítulo característicos de Asteraceae	2
Figura 3. Fotos de espécimes dos gêneros <i>Cotula</i> e <i>Stephanomeria</i>	3
Figura 4. Variação de massa de oxigênio em função do tempo geológico	7
Figura 5. Foto de George Bentham	10
Figura 6. Estrutura química da artemisinina	14
Figura 7. Picada do Anopheles	14
Figura 8. Diagrama de inter-relações filogenéticas de tribos de Asteraceae	20
Figura 9. Imagem 3D do citocromo c	25
Figura 10. Filogenia de <i>Encelia</i> realizada com base em caracteres fenotípicos	27
Figura 11. Ficha-modelo usada para levantamento bibliográfico	29
Figura 12. Exemplo do NOx para a 5-Metilcumarina	30
Figura 13. Estudo de passo oxidativo	31
Figura 14. Exemplos de alguns monoterpenos	38
Figura 15. Matricina, um guaianolídeo de Anthemideae (Asteraceae)	39
Figura 16. Exemplos de sesquiterpenos isolados de espécies de Asteraceae	39
Figura 17. Exemplos de diterpenos isolados de espécies de Asteraceae	40
Figura 18. Triterpenos característicos de espécies de Asteraceae	41
Figura 19. Exemplos de alguns poliacetilenos de Asteraceae	47
Figura 20. Esqueleto básico de um flavonóide	47
Figura 21. Tipos de anéis C encontrados em flavonóides	48
Figura 22. Exemplo de uma chalcona	49
Figura 23. Exemplos de flavonas	49
Figura 24. Exemplos de flavonóis	50
Figura 25. Exemplo de uma flavanona	50
Figura 26. Exemplos de antocianinas e antocianidinas	51
Figura 27. Exemplo de um isoflavonóide	51
Figura 28. Exemplo de antocianidina	52
Figura 29. Exemplos de cumarinas	53
Figura 30. Exemplos de piranocumarinas	54
Figura 31. Desidrotremetona	59
Figura 32. Ocorrências percentuais de monoterpenos em tribos de Asteraceae	64

Figura 33. Ocorrências percentuais de sesquiterpenóides em tribos de Asteraceae	65
Figura 34. Ocorrências percentuais de diterpenos em tribos de Asteraceae	66
Figura 35. Ocorrências percentuais de triterpenos em tribos de Asteraceae	66
Figura 36. Ocorrências percentuais de cumarinas em tribos de Asteraceae	67
Figura 37. Ocorrências percentuais de poliacetilenos em tribos de Asteraceae	68
Figura 38. Ocorrências percentuais de flavonóides em tribos de Asteraceae	69
Figura 39. Ocorrências percentuais de benzofuranos em tribos de Asteraceae	70
Figura 40. Diagrama de dispersão de valores previstos vs. observados para ocorrências de cumarinas em tribos de Asteraceae	75
Figura 41. Diagrama de dispersão de valores previstos vs. valores observados para ocorrências de cumarinas em subtribos de Asteraceae	83
Figura 42. Tela de computador apresentando resultados de resíduos padrão relativos a ocorrências de cumarinas para estudo de 5 outliers identificados em subtribos de Asteraceae	87
Figura 43. Histograma mostrando a distribuição de residuais brutos de ocorrências de cumarinas ao longo do eixo X versus número de observados em subtribos de Asteraceae	88
Figura 44. Histograma mostrando a distribuição de residuais padrão ao longo do eixo X versus número de observados	89
Figura 45. Diagrama de dispersão de valores previstos vs. valores observados para ocorrências de cumarinas em subtribos de Asteraceae	90
Figura 46. Diagrama tridimensional mostrando números de ocorrências de esqueletos de cumarinas por grupo em tribos de Asteraceae	101
Figura 47. Valores previstos vs. valores observados para ocorrências de benzofuranos em tribos de Asteraceae.....	105
Figura 48. Diagrama de dispersão mostrando valores previstos para ocorrências de benzofuranos em função de ocorrências de diterpenos, a partir da correlação bivariada entre as ocorrências de ambas as classes químicas.....	106
Figura 49. Diagrama de dispersão para valores previstos vs. observados de ocorrências de benzofuranos em subtribos de Asteraceae. Valores previstos obtidos a partir da regressão linear múltipla tomando OCBENZ como variável dependente	107

Figura 50. Diagrama tridimensional mostrando números de ocorrências de esqueletos de cumarinas por grupo (pode conter um ou vários tipos de esqueleto) em tribos de Asteraceae	117
Figura 51. Diagrama de dispersão para valores previstos vs. observados de ocorrências de flavonóides em tribos de Asteraceae. Valores previstos obtidos a partir da regressão linear múltipla tomando OCFLAV como variável dependente	119
Figura 52. Diagrama de dispersão para valores previstos vs. observados de ocorrências de flavonóides em subtribos de Asteraceae.....	122
Figura 53. Diagrama tridimensional de distribuição de tipos de flavonóides em tribos de Asteraceae	124
Figura 54. Diagrama de dispersão de ocorrências de flavanonas em função das ocorrências de flavonóis, a partir da correlação bivariada, em tribos de Asteraceae	126
Figura 55. Diagrama de dispersão de valores previstos versus observados de ocorrências de cumarinas em função de ocorrências de 7 tipos de produtos naturais em subtribos de Asteraceae, após remoção de 5 <i>outliers</i>	126
Figura 56. Previsão de ocorrências de monoterpenos em função de ocorrências de cumarinas em subtribos de Asteraceae	127
Figura 57. Diagrama tridimensional mostrando os valores médios de NOXs de diversas classes químicas em tribos de Asteraceae	128
Figura 58. Diagrama de dispersão de valores previstos vs. observados de NOXs de cumarinas mostrados na Tabela 37	130
Figura 59. Diagrama de dispersão de NOXs de flavonóides oxidados no anel B e cumarinas obtido a partir da regressão linear entre seus respectivos NOXs representados.....	133
Figura 60. Diagrama tridimensional mostrando NOXs de cumarinas, NOXs de flavonóides oxidados no anel B e número de observações dos referidos NOXs.	134
Figura 61. Diagrama de dispersão bivariado de valores de NOXs de flavonóides A e B em subtribos de Asteraceae, obtidos a partir de regressão linear simples....	138
Figura 62. Distribuição de luteolina e/ou apigenina entre Angiospermae	141
Figura 63. Dendrograma mostrando agrupamentos de 17 tribos de Asteraceae (ligação completa, distâncias euclidianas) baseado em números de ocorrências de 8 classes de produtos naturais nas referidas tribos	144

Figura 64. Dendrograma mostrando agrupamentos de tribos de Asteraceae
(Método de Ward – distâncias euclidianas.....145

SUMÁRIO DE TABELAS

Tabela 1. Isótopos radioativos e suas meias-vidas, utilizados em datações geológicas	4
Tabela 2. Escala absoluta do tempo geológico	6
Tabela 3. Classificações de Asteraceae segundo Bentham e Hoffmann	10
Tabela 4. Classificações principais recentes de Asteraceae	11
Tabela 5. Padrão químico de Asteraceae	18
Tabela 6. Tribos e subtribos de Asteraceae segundo a classificação de Bremer	21
Tabela 7. Matriz de diferenças de aminoácidos em citocromo c de seis espécies de plantas superiores	26
Tabela 8. Substituição de aminoácidos na posição 65 em citocromo c de filos diferentes.....	26
Tabela 9. Classificação de terpenos	37
Tabela 10. Números de ocorrências de oito classes de metabólitos secundários em 17 tribos de Asteraceae	61
Tabela 11. Números de ocorrências de oito classes químicas em 87 subtribos de Asteraceae	62
Tabela 12. Matrizes de correlações de ocorrências de oito classes de metabólitos secundários em tribos e subtribos de Asteraceae	72
Tabela 13. Matriz de correlações entre ocorrências de cumarinas, variável dependente, e ocorrências de diversas classes químicas em tribos Asteraceae	73
Tabela 14. Sumário de regressão linear múltipla para a variável dependente OCCUMA em tribos de Asteraceae.....	74
Tabela 15. Valores observados vs. previstos de ocorrências de cumarinas em tribos de Asteraceae	74
Tabela 16. Resultados de regressão linear múltipla passo-a-passo adiante de ocorrências de cumarinas em função de ocorrências de diversas classes químicas em tribos de Asteraceae	76
Tabela 17. Sumário de regressão linear múltipla passo-a-passo para trás para a váriavel dependente OCCUMA, em subtribos de Asteraceae	78
Tabela 18. Matriz de coeficientes de correlação r entre ocorrências de cumarinas e diversas classes químicas em subtribos de Asteraceae	80

Tabela 19. Sumário de regressão linear múltipla padrão de Nos de cumarinas como variável dependente em função de Nos de diversas classes químicas em subtribos de Asteraceae	81
Tabela 20. Valores previstos vs. valores observados de ocorrências de cumarinas em subtribos de Asteraceae, a partir da regressão linear múltipla.....	82
Tabela 21. Valores residuais padrão de Nos previstos vs. observados para cumarinas em subtribos de Asteraceae	85
Tabela 22. Resíduos de 5 outliers identificados em estudo de valores previstos vs. observados e respectivos pontos observados no diagrama de dispersão para ocorrências de cumarinas em subtribos de Asteraceae	88
Tabela 23. Resumo de regressão para variável dependente: ocorrências de cumarinas (OCCUMA) após remoção de 5 outliers	90
Tabela 24. Grupos de cumarinas com suas respectivas quantidades de esqueletos por tribo de Asteraceae	98
Tabela 25. Resumo de grupos de cumarinas e números totais de ocorrências de esqueletos por grupo em Asteraceae	100
Tabela 26. Esqueletos de cumarinas distribuídos em subtribos de Asteraceae..	102
Tabela 27. Resumo de regressão linear múltipla padrão tomando como variável dependente ocorrências de benzofuranos em função de diversas classes químicas, em tribos de Asteraceae.....	103
Tabela 28. Valores previstos versus valores observados para ocorrências de benzofuranos em tribos de Asteraceae.....	104
Tabela 28. Ocorrências de diterpenos (DITE) e benzofuranos (BENZ) em tribos de Asteraceae	106
Tabela 29. Sumário de regressão linear múltipla de ocorrências de benzofuranos, tomadas como variável dependente, em subtribos de Asteraceae.....	107
Tabela 30. Valores previstos e observados de ocorrências de benzofuranos em subtribos de Asteraceae. Valores previstos obtidos a partir da regressão linear múltipla tomando OCBENZ como variável dependente.....	108
Tabela 31. Grupos de benzofuranos com suas respectivas quantidades de esqueletos por tribo de Asteraceae	115
Tabela 32. Grupos de benzofuranos e números totais (T. ESQ) de ocorrências de esqueletos por grupo em Asteraceae	115

Tabela 33. Esqueletos de benzofuranos distribuídos em subtribos de Asteraceae	116
Tabela 34. Sumário de regressão linear múltipla de ocorrências de benzofuranos, tomadas como variável dependente, em tribos de Asteraceae	118
Tabela 35. Valores previstos e observados de ocorrências de flavonóides em subtribos de Asteraceae. Valores previstos obtidos a partir da regressão linear múltipla tomando OCFLAV como variável dependente	121
Tabela 36. Sumário da regressão linear múltipla, para previsão de ocorrências de flavonóides em tribos de Asteraceae	123
Tabela 37. Tipos de flavonóides encontrados em Asteraceae	124
Tabela 38. Ocorrências de flavonóis e flavanonas em tribos de Asteraceae	125
Tabela 39. Sumário de regressão linear múltipla de NOXs de cumarinas em função de NOXs de diversas classes químicas em tribos de Asteraceae	129
Tabela 40. Valores previstos vs observados de NOXs de cumarinas em tribos de Asteraceae	130
Tabela 41. Valores de NOXs médios de diversas classes químicas em subtribos de Asteraceae.....	136
Tabela 42. Sumário de regressão linear para a variável dependente OXFLAV-B, correlacionada com a variável FLAV-A, em 72 subtribos de Asteraceae	137
Tabela 43. Valores de NOXs médios de flavonóides A e B calculados em 72 subtribos de Asteraceae.....	137
Tabela 44. Parâmetros químicos para flavonóides de tribos de Asteraceae.....	139
Tabela 45. NOs de 8 classes químicas em tribos de Asteraceae	146

SUMÁRIO DE ESQUEMAS

Esquema 1. Rotas biossintéticas de produtos naturais em plantas	36
Esquema 2. Formação hipotética de mirceno e α -pineno através de unidades de isopreno, ligação cabeça-cauda	37
Esquema 3. Formação de policetídeos via condensação de Claisen	41
Esquema 4. Condensação de Claisen e mecanismos químicos envolvidos nesta reação (adição nucleofílica à carbonila)	42
Esquema 5. Condensação aldólica e mecanismos químicos envolvidos nesta reação (adição nucleofílica à carbonila)	42
Esquema 6. Formação do pirofosfato de farnesila e do esqualeno	43
Esquema 7. Rota biossintética de terpenóides a partir dos precursores pirofosfato de geranila e pirofosfato de farnesila	44
Esquema 8. Rota biogenética de poliacetilenos	45
Esquema 9. Esquema hipotético para a formação de 4,2',4',6'- tetraidroxichalcona a partir de <i>p</i> -cumaroil-CoA e malonil-CoA	48
Esquema 10. Rota biossintética de cumarinas simples	55
Esquema 11. Possíveis rotas de formação do anel furânico em cumarinas	56
Esquema 12. Rota biossintética proposta para prenil, precursor do isopreno	57
Esquema 13. Mecanismos de prenilação de enóis e fenóis	58
Esquema 14. Dois modos de ciclização de cadeias policetometilênicas	59

I - INTRODUÇÃO

1.1. Asteraceae (Compositae)

Dentro da biodiversidade de nosso planeta, as plantas sempre ocuparam um lugar de destaque no processo de desenvolvimento e continuidade de quase todo tipo de vida existente. O reino vegetal tem sido imprescindível na vida do homem desde os tempos mais remotos até os dias atuais: alimentos, abrigo, vestuário, utensílios domésticos, armas, adornos, rituais religiosos, fonte de energia (fogo), medicamentos naturais, etc.

Asteraceae (Compositae) é uma das mais importantes famílias de plantas vasculares florescentes no mundo. Até o presente, dela já foram identificados ca.1300 gêneros e 23000 espécies. Apelidada mais comumente de "a família do girassol" (Figura 1) ou "a família da margarida" (Herman *et al.*, 2000), distribui-se em quase toda a parte da Terra, envolvendo os mais variados ecossistemas. É, portanto, cosmopolita. Sua diversidade máxima está associada a áreas abertas, freqüentemente temperadas: pradarias e savanas das Américas; estepes e gramados da África, Ásia, Austrália e da Nova Zelândia. Elas ascendem muito alto nas montanhas (por exemplo, *Achillea moschata*, flores típicas dos Alpes), aventuram-se em desertos, praias, solos salgados e inóspitos, contanto que a luz lhes seja oferecida em toda a sua plenitude.

Figura 1. Girassol (*Helianthus annuus*)

Esta família de angiospermas compõe-se principalmente de plantas herbáceas perenes e anuais, muitos subarbustos, poucos arbustos ou tipos arborescentes. Podemos facilmente reconhecer o tipo "compostas" através dos capítulos, que são o motivo essencial do termo "compostas". Tipicamente, as flores são arranjadas em

uma estrutura chamada capítulo (Figura 2. Cada capítulo consiste de muitas flores individuais pequenas, chamadas flósculos, introduzidas sobre a uma estrutura central chamada disco. Alguns membros da família têm somente flósculos tubulares (tubulifloras) ou de disco (por exemplo, cárduo, *Cirsium vulgare* Bull; malmequer, *Calendula officinalis* L.); outros podem ter os flósculos de disco cercado por raios (flósculos ligulados do tipo A ou ligulifloras, por exemplo, *Cotula coronopifolia* / botão-de-latão, Fig.3-A). Num outro grupo, a subfamília Cichoriideae, por exemplo, faltam flósculos de disco; possui somente flósculos de raio (flósculos ligulados do tipo B, por exemplo, *Stephanomeria virgata*, Fig.3-B).

Figura 2. Fotos de tipos de capítulo característicos de Asteraceae

Comentário : Um capítulo de Asteraceae pode conter: 1) somente flósculos ligulados (capítulo ligulado, C 1); 2) somente flósculos tubulares (capítulo discóide, C 2); flósculos de raio e tubulares (capítulo radiado, C 3).

No capítulo das compostas, toda inflorescência se assemelha a uma só flor de ordem superior. Por isso as folhas terminais formam uma espécie de cálice (invólucro); os brotos terminais formam um receptáculo, que pode ser horizontal, convexo ou, mais raramente, côncavo; as brácteas das pequenas flores são apenas escamas, suas sépalas se tornam uma mecha plumosa (*pappus*) ou sedosa, que mais tarde coroa alguns frutos ínferos e dá a estes a faculdade de poderem transportar-se facilmente pelo ar”.

A- *Cotula coronopifolia*B- *Stephanomeria virgata*Figura 3. Fotos de espécimes dos gêneros *Cotula* e *Stephanomeria*.

As tribos e os gêneros em que os botânicos subdividem esta família são unicamente determinados pelo estudo de sua estrutura floral. Todavia descrever botanicamente Asteraceae foge ao escopo da presente dissertação, assunto que pode ser fartamente encontrado em bibliografia especializada de renomados botânicos (Bremer, 1994; Cronquist, 1981; Heywood *et al.*, 1977). A família é de grande importância econômica incluindo alimentos e plantas agroindustriais (alface, alcachofra, girassol), ornamentais (dália, crisântemo, margarida), ervas medicinais (camomila, taráxico, picão, artemísia), ervas-daninhas importantes (dente-de-leão, cardo) incluindo a espécie venenosa *Senecio jacobaea* (tasneira / erva-de-santiago), responsável por muitas mortes de animais domésticos. Poucas espécies de maior porte são classificadas como árvore (por exemplo, *Montanoa revealii*, espécie encontrada no México e em alguns países da América Latina).

1.2. **Dados de fósseis de plantas**

A idade da Terra é estimada em 4,54 bilhões de anos com base de estudos de isótopos, especialmente os isótopos de chumbo (Pb-206, Pb-207, e Pb-208 ou Pb-204, Tabela 1) de amostras terrestres e de meteoritos (Stassen, 1997). Mas o que mais intriga o homem é saber quando surgiu a vida em nosso planeta.

Tabela 1. Isótopos radioativos e suas meias-vidas, utilizados em datações geológicas (a=anos; m.a=milhão ou milhões de anos; b.a=bilhão ou bilhões de anos)

Decaimento		Meia-vida	Faixa útil	Tipo de material
Pai	Filho	$t_{1/2}$		
$^{238}\text{U}_{92}$	$^{206}\text{Pb}_{82}$	4,47 b.a	>10 milhões de anos	Rochas e minerais ígneos & às vezes, metamórficos
$^{235}\text{U}_{92}$	$^{207}\text{Pb}_{82}$	707 m.a		
$^{232}\text{Th}_{90}$	$^{208}\text{Pb}_{82}$	14 b.a		
$^{40}\text{K}_{19}$	$^{40}\text{Ar}_{18}$ & $^{40}\text{Ca}_{20}$	1, 28 b.a	>10000 anos	
			>10 milhões de anos	
$^{87}\text{Rb}_{37}$	$^{87}\text{Sr}_{38}$	48 b.a		
$^{147}\text{Sm}_{62}$	$^{143}\text{Nd}_{60}$	106 b.a		
$^{14}\text{C}_6$	$^{14}\text{N}_7$	5730 a	100 - 70000 anos	Material orgânico

Quando se fala em fósseis, temos de imaginar os tempos geológicos mais remotos, e o alvo principal é a vida surgida e desenvolvida na Terra nas mais diversas modalidades. As origens da vida não podem ser datadas precisamente, mas há evidência de que organismos como bactérias viveram na Terra há 3,5 bilhões de anos, e podem ter existido ainda mais cedo, quando a primeira crosta contínua foi formada há quase 4 bilhões de anos. Esses primeiros organismos devem ter sido bem mais simples do que os que vivem hoje. Além disso, os organismos mais adiantados devem ter antes existido lá como estruturas que não seriam chamadas "vivas", mas que são agora os componentes de coisas vivas. Hoje, todos os organismos vivos armazenam e transmitem a informação hereditária usando dois tipos de moléculas: ácido desoxirribonucléico (DNA) e ácido ribonucléico (RNA), macromoléculas que trazem um incalculável número de informações e mistérios da vida.

A escala temporal dos eventos da história da Terra, ordenados em ordem cronológica, denomina-se tempo geológico. Com base nos princípios de superposição das camadas litológicas (as mais antigas sobrepostas pelas mais jovens) e sucessão da fauna (determinadas espécies viveram em um determinado período do tempo), foi concebida uma escala de tempo relativa. Posteriormente, com o advento das técnicas de datação radiométricas, foi desenvolvida uma escala de tempo absoluta para os períodos geológicos conforme mostra a Tabela 2.

A palinologia é um ramo das ciências biológicas que estuda profundamente pólen e esporos através de microscopias óptica e eletrônica e enfatiza suas características em taxonomia de plantas. Estudos palinológicos de alguns fósseis de famílias de plantas florescentes, incluindo Asteraceae — têm sido raros tais achados —, revelam sua existência remontando do Período Cretáceo (por volta de 144 a 66 milhões de anos atrás, conforme mostra a Tabela 2. Turner (1977), por exemplo, argumentou que a família Asteraceae originou-se na metade do Cretáceo, bastante cedo para a força de tração continental promover sua distribuição no mundo, conseguida como "um resultado da tectônica de placas". Em terra, os dinossauros ainda eram dominantes, mas igualmente se extinguíram no final desse período. Mamíferos e aves ainda são insignificantes em número. Quanto à flora, as angiospermas (plantas com flores) se diversificam e adquirem bastante importância. O oxigênio em abundância na atmosfera supõe-se ter contribuído em muito no desenvolvimento das angiospermas (Gottlieb, 1989 - Figura 4). O Cretáceo foi assim o tempo em que veio junto a vida como existe agora na Terra.

O registro de fósseis para animais confirma geralmente a informação derivada de estudos filogenéticos. O número das similaridades nos aminoácidos em posições particulares em moléculas do citocromo *c* de diferentes animais torna estatisticamente improvável que poderiam ter-se originado de mais de um único tipo ancestral com uma molécula de citocromo *c* ancestral. Seguindo as diferenças em substituições do aminoácido, é possível relacionar vários grupos de animais, porque os grupos sucessivos carregam a molécula mudada de citocromo *c* após uma modificação. Nas plantas, especialmente as florescentes, não há registro extensivo de fósseis. Muito do conhecimento atual de relações e filogenia desse grupo é baseado na extrapolação de estudos em dados morfológicos. Até então, relativamente poucos citocromos *c* de plantas foram estudados, mas, nos poucos que foram investigados, há evidências, a partir de similaridades entre as seqüências de aminoácidos, de que os citocromos de plantas e animais estão relacionados (Siegler, 1981).

Éon	Era	Período	Época	Milhões de anos	Principal evento
Fanerozoico	Cenozoico	Quaternário	Recente	0,01	← Aparecimento do homem
			Pleistoceno	1,6	
		Terciário	Plioceno	5,3	← Extinção dos dinossauros
			Mioceno	23,7	
			Oligoceno	36,6	
			Eoceno	57,8	
			Paleoceno	66	
			Mesozoico	Cretáceo	
	Jurássico	208			
	Triássico	245		← Mamíferos Separação do Pangéia	
	Paleozoico	Permiano	286	← Répteis	
		Carbonífero	360		← Anfíbios
		Devoniano	408	← Plantas terrestres	
		Siluriano	438		
		Ordoviciano	505		← Peixes
		Cambriano	570		
Pré-Cambriano	Proterozóico			← Moluscos	
			2500		
	Arqueano			← Primeiro registro de vida	
				3800	
				4600	

Tabela 2. Escala absoluta do tempo geológico, onde se pode facilmente localizar o Período Cretáceo.

Linha quebrada: Variação da massa de oxigênio atmosférico no fanerozóico. Linhas sólidas: Diversidade de plantas vasculares terrestres (em número de espécies) dos quatro grupos de plantas vasculares (EP: primeiras plantas vasculares, PT: pteridófitas, GY: gimnospermas, AN: angiospermas) que sucessivamente dominaram a flora terrestre, baseada na compilação de ca. 18000 citações de espécies de plantas fósseis. Ma: milhões de anos.

Figura 4. Variação da massa de oxigênio em função do tempo geológico

1.3. Evolucionismo

A teoria do evolucionismo concebida e preconizada por Darwin (Darwin, 1859) tenta explicar a origem de muitas espécies de seres vivos com base em seleção natural e mutação fortuita. Entretanto essa teoria deixa ainda dentro de cada um de nós enormes lacunas e interrogações no âmbito filosófico, científico e religioso, às vezes conflitante para aqueles que acreditam em um Ser Supremo, Eterno e Criador, que transcende a razão e toda a imaginação humanas.

Uma abordagem crítica de Denton (Denton, 1986) sobre evolução da vida, e que contesta em muitos pontos o darwinismo, é também bastante polêmica. Tem encontrado certa resistência de muitos estudiosos do evolucionismo apoiado nas hipóteses e experiências de Darwin. Em *One Long Argument*, Mayr, 1991 (biólogo evolucionário, originador do conceito biológico da espécie) sumariza teorias de Darwin e segue a história de sua aceitação pela comunidade científica do mundo. O autor interpreta e aprofunda o pensamento de Darwin com mais clareza e enfoque moderno.

1.4. Taxonomia e filogenia

Mesmo antes de ser definida num enfoque conceitual mais abrangente e atualizado, Carl Linnaeus, sueco, médico, professor de História Natural e Botânica na Universidade de Upsala (Suécia), figura-se como o pai da taxonomia. No verão de 1745, Linnaeus publicou seus livros *Flora Suecica* e *Fauna Suecica*. Em 1753, seu livro, *Species plantarum*, foi publicado em dois volumes, nos quais já faz uma abordagem sistemática de espécies de plantas usando nomenclatura binominal, ou seja, com o nome de um gênero e um epíteto específico. Este sistema é ainda adotado presentemente. *Species plantarum* foi, portanto, o precursor da taxonomia moderna.

Taxonomia é a ciência para nomear e classificar organismos. Sistemática é o estudo científico de tipos e diversidade de organismos e as relações entre eles (Raven *et al.*, 1967). Ambos os termos ainda são tomados como sinônimos por inúmeros pesquisadores dessa área de estudos. Filogenia descreve as relações evolucionárias entre táxons (grupos de organismos).

Taxonomia foi definida por Simpson como o estudo teórico de classificação, incluindo suas bases, princípios, procedimentos e regras (Sneath, 1973). Tem como objetivo estudar morfológica e citogeneticamente caracteres de espécimes vivos, avaliar as relações de parentesco existentes entre estes e classificá-los hierarquicamente em grupos ou táxons. Os caracteres são inúmeros (capítulos de flores, corolas, frutos, inflorescências, infrutescências, raízes, pólen, anatomia de células e tecidos, número de cromossomos, etc.) e são estudados detalhadamente. Em níveis hierárquicos descendentes, temos como exemplo: divisão ou filo (mais específico para animais), classe, ordem, família, subfamília, tribo, subtribo, gênero, espécie, variedade.

Taxonomia numérica é definida como o agrupamento por métodos numéricos de unidades taxonômicas em táxons com base em seus estados de caracteres. O termo inclui o desenho de inferências filogenéticas a partir de dados de métodos estatísticos ou de outros métodos matemáticos na maior abrangência que for possível. Esses métodos requerem a conversão de informação sobre entidades taxonômicas em quantidades numéricas. O advento da taxonomia numérica, que se vale, portanto, de métodos matemáticos — matrizes, unidades taxonômicas operacionais (OTUs), distâncias euclidianas ou de Manhattan entre pares de OTUs,

etc.— para avaliar diferenças observáveis e similaridades entre grupos taxonômicos; dá , portanto, um caráter mais científico à sistemática geral dos seres vivos (Sneath, 1973). Seu emprego em sistemática vem-se ampliando cada vez mais desde a década de 50 aos tempos atuais, quando a informática robustece o seu uso, principalmente em cladística.

1.5. Taxonomia de Asteraceae

Do século XVIII ao atual, a família Asteraceae tem sido taxonomicamente estudada por renomados botânicos, em níveis hierárquicos diversificados (subfamílias, tribos , subtribos, gêneros, espécies).

Em 1816, o botânico francês Henri Cassini fez, pela primeira vez, a classificação de Asteraceae, usando genuinamente terminologia francesa e mostrando num diagrama interrelações de suas 19 tribos (Bremer, 1994).

O primeiro livro sobre Asteraceae foi publicado por Lessing em 1832. A classificação de Lessing, que foi muito diferente da de Cassini, incluía sete tribos, algumas com circunscrições bastante amplas e artificiais. Um esquema similar foi adotado por De Candolle em 1836.

Em 1873, George Bentham (Figura 5) fez uma substancial revisão de classificação tribal de Asteraceae, próxima, em grande parte, à classificação de Cassini. Bentham, 1973 classificou Asteraceae em 13 tribos e sua classificação apresenta algumas controvérsias - por exemplo, nas tribos Vernonieae, Heliantheae, Helianieae e Senecioneae – (Bremer, 1994) relativas a caracteres morfológicos que definem as tribos.

Em *Die Natürlichen Pflanzenfamilien*, Hoffmann (1890) essencialmente repetiu a classificação de Bentham com pequenas alterações, especialmente na terminologia. É ainda de grande valia nos dias atuais. Na Tabela 3, são mostradas as classificações de Bentham e Hoffmann. Por quase um século, a classificação tribal de Bentham-Hoffmann para Asteraceae se manteve, portanto, praticamente original.

Figura 5. Foto de George Bentham

Tabela 3. Classificações de Asteraceae segundo Bentham e Hoffmann.

BENTHAM 1873	HOFFMANN 1890
Vernoniaceae	Vernonieae
Eupatoriaceae	Eupatorieae
Asteroideae	Astereae
Inuloideae	Inuleae
Helianthoideae	Heliantheae
Helenioideae	Helenieae
Anthemideae	Anthemideae
Senecionideae	Senecioneae
Calendulaceae	Calenduleae
Arctotideae	Arctotideae
Cynaroideae	Cynareae
Mutisiaceae	Mutisieae
Cichoriaceae	Cichorieae

Ao longo do século XX, alguns renomados botânicos têm tentado classificações mais coerentes para essa família, em sua maioria, em níveis de subfamílias e tribos conforme ilustra a Tabela 4 (Bremer *et al.*, 1992). Devem-se enfatizar também inúmeras reclassificações de gêneros e espécies da família Asteraceae.

Tabela 4. Classificações principais recentes de Asteraceae (Lactuoidae = Cichorioideae, Cichorieae = Lactuceae, Arctotideae = Arctoteae, Cardueae).

Poljakov (1967)	Heywood et al. (1977)	Wagenitz (1976)	Carlquist (1976)	Jeffrey (1978)	Robinson (1981,1983)	Thorne (1983)	Bremer (1987)	Bremer et al. (1992)	Cronquist (1955, 1977)
		<i>Cichorioideae</i>	<i>Cichorioideae</i>	<i>Lactuoidae</i>	<i>Cichorioideae</i>	<i>Cichorioideae</i>	<i>Cichorioideae</i>	<i>Barnadesioideae</i>	
Anthemideae	Eupatorieae	<i>Asteroidae</i>	Mutisieae	Lactuceae	Lactuceae	Mutisieae	Mutisieae	Barnadesieae	<i>Helianthoideae</i>
Astereae	Vernonieae		Vernonieae	Mutisieae	Vernonieae	Vernonieae	Arctoteae		Astereae
Senecioneae	Astereae	<i>Grupo 1</i>	Cardueae	Eremothamneae	Liabeae	Liabeae	Cardineae	<i>Cichorioideae</i>	Anthemideae
Inuleae	Inuleae		Arctoteae	Arctotideae	Mutisieae	Cichorieae	Echinopsiodeae		Arctotideae
Heliantheae	Heliantheae	Vernonieae	Cichorieae	Cardueae	Cardueae	Cardueae	Cardueae	Mutisieae	Inuleae
Filagineae	Tageteae	Liabeae	Eupatorieae	Vernonieae	Echinopsiodeae	Arctoteae	Lactuceae	Cardueae	Senecioneae
Arctotideae	Senecioneae	Mutisieae		Liabeae	Gundelieae		Lactuceae	Lactuceae	Calenduleae
Mutisieae	Liabeae	Cardueae	<i>Asteroidae</i>	Eupatorieae	Eremothamneae	<i>Asteroidae</i>	Vernonieae	Vernonieae	Eupatorieae
Cynareae	Anthemideae	Echinopeae			Arctoteae		Liabeae	Liabeae	Vernonieae
Vernonieae	Arctoteae	Arctotideae	Heliantheae	<i>Asteroidae</i>		Heliantheae	Arctoteae	Arctoteae	Liabeae
Cichorieae	Calenduleae		Astereae		<i>Asteroidae</i>	Tageteae	<i>Asteroidae</i>	Cardueae	Cardueae
Fitchineae	Cynareae	<i>Grupo 2</i>	Inuleae	Senecioneae		Eupatorieae		Mutisieae	Mutisieae
	Mutisieae		Anthemideae	Tageteae	Eupatorieae	Astereae	Inuleae	<i>Asteroidae</i>	Cichorieae
	Lactuceae	Eupatorieae	Senecioneae	Heliantheae	Heliantheae	Inuleae	Gnaphalieae	Inuleae	
		Heliantheae	Calenduleae	Inuleae	Calenduleae	Anthemideae	Astereae	Astereae	
		Helenieae		Anthemideae	Inuleae	Senecioneae	Eupatorieae	Anthemideae	
		Senecioneae		Ursinieae	Senecioneae	Calenduleae	Calenduleae	Senecioneae	
		Calenduleae		Calenduleae	Astereae		Senecioneae	Calenduleae	
		Astereae		Cotuleae	Anthemideae	Anthemideae	Helenieae	Helenieae	
	Inuleae			Astereae		Helenieae	Coreopsidae	Coreopsidae	
	Anthemideae					Madieae	Tageteae	Tageteae	
						Heliantheae	Heliantheae	Heliantheae	
						Tageteae	Eupatorieae	Eupatorieae	
						Coreopsidae			

Comentário: As classificações de Wagenitz (1976) a Bremer (1987) são dicotômicas. Mesmo a classificação mais recente de Bremer, p.20, mostra também uma certa dicotomia na família Asteraceae. Procurou-se manter na tabela a originalidade do conteúdo devidamente traduzido, não incluindo, portanto, naquela a nova classificação de Bremer nem modificando ordens cronológicas. NOTA: Até os anos setenta, a tribo Lactuceae era considerada distinta de toda a família Asteraceae e era classificada em sua própria subfamília, a Liguliflorae. Daí não constar em algumas classificações na Tabela 4.

1.6. Quimiotaxonomia

Metabólitos são substâncias produzidas ou necessárias em qualquer processo químico ou metabolismo inerente a um organismo vivo.

O estudo de metabólitos secundários como marcadores químicos em táxons de plantas e microorganismos vem sendo, cada vez mais, ampliado e aceito por taxonomistas de destaque. Estes podem utilizar caracteres químicos (micromoléculas) e caracteres macromoleculares na identificação de relações existentes entre seres vivos, em qualquer nível hierárquico: populações, espécies, gêneros, tribos, famílias, etc. bem como em análises evolutivas para avaliar relações filogenéticas e variação dentro de espécies (Gershenzon and Mabry, 1983; Hegnauer, 1985; Waterman and Gray, 1987; Zdero and Bohlmann, 1990). Esse processo de estudo classificatório denomina-se quimiosistemática ou quimiotaxonomia.

O *modus operandi* básico para estudos quimiotaxonômicos foi primeiramente visto no início do século vinte, quando foram desenvolvidas técnicas sorológicas para permitir uma rápida avaliação de um grande número de extratos de plantas e estudá-los comparativamente. O objetivo maior dessas técnicas era conseguir uma nova classificação filogenética (*Stammbaum*) do reino vegetal inteiro. Infelizmente, o entusiasmo dos sorotaxonomistas foi algo maior que a avaliação crítica das conclusões que alcançaram, uma vez que muitas delas contradisseram as relações que haviam sido estabelecidas a partir de uma faixa de critérios taxonômicos válidos.

Os primeiros esforços para usar dados químicos em quimiotaxonomia foram bastante dificultados pela presença de um banco de dados com que trabalhar ainda muito deficiente. Os estudos de McNair foram talvez os primeiros em que houve, de fato, um esforço de empregar caracteres químicos como contribuidores para sistemas taxonômicos fenéticos de uma forma compreensiva, usando determinados óleos e gorduras, alcalóides e óleos voláteis como caracteres adicionais em sistemas que foram de outra maneira baseados na taxonomia tradicional (Waterman *et al.*, 1987).

A quimiotaxonomia classifica três espécies de substâncias produzidas pelas plantas: metabólitos primários, aqueles que realizam funções vitais, e.g. ácido cítrico usado no ciclo de Krebs, alguns aminoácidos tais como lisina, fenilalanina; ácidos

graxos, lipídeos, etc.; metabólitos secundários, tratados na química orgânica como produtos naturais. São aqueles que, a princípio, realizam funções não-vitais como, por exemplo, alcalóides, fenólicos, glucosinolatos, terpenóides, cumarinas, etc. É bom salientar que os metabólitos secundários são produzidos a partir de metabólitos primários e têm funções bem diversificadas tais como: pigmentos (flavonóides, betalaínas), óleos essenciais (terpenóides), aleloquímicos (aflatoxinas), resinas, etc. Metabólitos semantídeos ou semânticos são moléculas que carregam informações biogenéticas — por exemplo, DNA, RNA, proteínas em geral. Na realidade, são essencialmente metabólitos primários especiais com uma classificação moderna mais técnica.

A origem de metabólitos primários e secundários em plantas envolve rotas biossintéticas complexas, onde atuam sistemas de enzimas poderosas, muitas das quais ainda não foram totalmente elucidadas. O Esquema 1 (p. 36) apresenta a rota geral biossintética de tais metabólitos.

Classificação de metabólitos secundários:

A) Baseada na estrutura química

- 1- Alifáticos de cadeia aberta : ácidos graxos, poliacetilenos, etc.
- 2- Alicíclicos: terpenóides, esteróides, alcalóides, etc.
- 3- Aromáticos / benzenoídeos: fenólicos, quinonas, etc.
- 4- Heterocíclicos: alcalóides, flavonóides, cumarinas, benzofuranos, etc.

B) Baseada em atividade biológica: morfina (1806); penicilina G (1939); prostaglandina F₂ (1963); hormônios; antibióticos; vitaminas; micotoxinas, etc.

C) Baseada na taxonomia: alcalóides do ópio, do ergot, de iboga, de Menispermaceae, etc.

D) Baseada na biossíntese: via do acetato, via do chiquimato, via do acetato-mevalonato, etc.

E) Baseada em propriedades farmacológicas: anti-hipertensivo, antitumoral, cardiotônico, analgésico, etc.

Importância dos metabólitos secundários ou produtos naturais

- Estudos de rotas biossintéticas, que podem trazer novos avanços em conhecimentos científicos e técnicos (áreas de química, física, biociências, etc.).

- Pesquisas de sínteses de novos fármacos baseados em estruturas e atividades biológicas de tais metabólitos.
- Fitoterapia e Etnobotânica (maior compreensão do uso de inúmeras plantas pelo homem no tratamento das mais diversas doenças em todo o mundo).
- Quimiotaxonomia

1.7. Nomenclatura de metabólitos secundários

A nomenclatura de produtos naturais de acordo com a IUPAC (*International Union of Pure and Applied Chemistry*) é, de modo geral, bastante complexa. Tomemos como exemplos: cumarina= 2-*H*-1-benzopiran-2-ona, relativamente simples; todavia, artemisinina (*Artemisia annua*), que é uma lactona sesquiterpênica antimalárica de interesse científico-econômico mundial, tem o nome químico (3*R*,5*aS*,6*R*,8*aS*,9*R*,12*S*,12*aR*)-octaidro-3,6,9-trimetil-3,12-epoxi-12*H*-pirano[4.3-*j*]-1,2-benzodioxepin-10(3*H*)-ona (Figura 6). No homem, pela picada do *Anopheles* (Figura 7), formas infectantes do parasita, conhecidas como esporozoítos, abandonam as glândulas salivares do inseto e invadem o organismo humano. Outra droga antimalárica ainda muito usada no Brasil e em outros países da América Latina, Ásia e África é a cloroquina (sintética).

Figura 6. Estrutura química da artemisinina

Figura 7. Picada do *Anopheles* em pele humana

1.8. Nome trivial e sua importância

Um nome trivial é o nome dado, por alguma conveniência, a um composto novo com estrutura ainda incerta. Tal nome comumente não carrega nenhuma (ou mínima) informação estrutural e é derivado geralmente da origem do material.

Nomes triviais em produtos naturais são, às vezes polêmicos, com terminologias ambíguas. Por exemplo, ostol; isolado pela primeira vez em 1909 (*Imperatoria osthutium*). Para um químico, esse nome poderia induzi-lo pensar em algum produto com função álcool, todavia esse composto é um éter de ostenol (7-hidroxi-8-prenil-cumarina); umbeliferona é simplesmente um fenol (7-hidroxi-cumarina), sem nenhuma função cetona.

A IUPAC faz algumas recomendações concernentes ao uso de nomes triviais baseados na origem biológica: *"Quando um composto for isolado de uma fonte natural e um nome trivial for necessário, o nome deve ser baseado, sempre que possível, no nome da família ou do gênero ou da espécie do material biológico do qual o composto foi isolado (Não devem ser gerados nomes baseados em uma atividade metabólica presumida da substância isolada.) Numa maior abrangência possível, o nome deve ser selecionado para refletir a distribuição conhecida ou provável do produto natural, por exemplo, uma espécie hipotética de Paradigma exemplare (família Beispieliae) poderia dar produtos com nomes baseados em beispieliae, paradigma ou exemplare. Se apropriado, a classe ou a ordem poderiam também ser usadas como a base para o nome de um composto que ocorresse em um número de famílias relacionadas"*. Um bom exemplo é angelicina, uma cumarina de *Archangelica officinalis* (Apiaceae), isolada, pela primeira vez, em 1934, com nome trivial já de acordo com a recomendação supracitada da IUPAC.

1.9. Isótopos como marcadores de moléculas

O uso de moléculas marcadas com isótopos radioativos — ^{14}C , por exemplo, na biossíntese de colesterol utilizando $\text{CH}_3^{14}\text{COOH}$ ou $^{14}\text{CH}_3\text{COOH}$ para alimentar animais, (Hendrickson *et al.*, 1970) — e naturais estáveis (^{13}C , ^{15}N , ^2H , ^{18}O , abundância natural de 1,1%, 0,36%, 0,015%, 0,204%, respectivamente) em estudos de sínteses, metabolismos, patologias (Finlay and Smith, 2002), mecanismos de reações em sínteses (Lamb *et al.*, 1996), etc., é uma técnica muito valiosa em química e ciências correlatas. Desses isótopos naturais, os três primeiros logo fazem surgir sinais na espectroscopia de ressonância magnética nuclear e podem ser analisados ou por este método ou por espectrometria de massas. ^{18}O (abundância natural de 0,204%) não dá sinais diretos, por exemplo, na RMN ^1H , mas pode ser analisado ou por espectrometria de massas ou por espectroscopia de RMN ^{13}C . A

presença de ^{18}O num átomo de carbono particular causa um deslocamento em campo alto de ressonância para esse átomo de carbono comparado quando ^{16}O está presente no mesmo (a grandeza desses deslocamentos varia de ca. 0,01-0,035 ppm para oxigênios ligados ao carbono por ligação simples e ca. 0,03-0,55 para oxigênios ligados por dupla ligação (Herbert, 1989).

No estudo de rotas biossintéticas, tais isótopos têm papel relevante na determinação de precursores ou confirmação destes propostos hipoteticamente (Shemin, 1989; Kroumova and Wagner, 1999). O acompanhamento dessas reações pode exigir ainda outras técnicas complementares tais como culturas de células, tecidos e calos de organismos (Zieg *et al.*, 1983), já amplamente utilizadas na obtenção de metabólitos secundários de plantas de interesse farmacológico, nutricional e outros.

A pesquisa de produtos naturais vem-se tornando cada dia mais eficiente e ágil, graças aos avanços no conhecimento teórico e prático do homem sobre a matéria e ao uso por ele de técnicas e dispositivos respectivos sofisticados e de alta resolução. O pesquisador hodierno de produtos naturais e afins dispõe de farto acervo literário sobre esse assunto e acesso relativamente fácil, por exemplo, a técnicas avançadas: cromatografia gasosa (CG), cromatografia líquida de alta eficiência (CLAE), eletroforese capilar (EC) (Audy *et al.*, 1990), espectrometria de massas (EM), ressonância magnética nuclear de ^1H e ^{13}C (RMN ^1H e RMN ^{13}C , respectivamente), espectroscopias no infravermelho (IV), no ultravioleta (UV) e Raman, dicroísmo circular óptico (DCO) / dispersão rotatória óptica (DRO), etc. Além disso, conta com a ajuda da informática, que lhe oferece as mais diversas opções de programas computacionais e sistemas especialistas (Lindsay *et al.*, 1980; Will *et al.*, 1996; Gastmans *et al.*, 1990; Emerenciano *et al.*, 1994; Ferreira *et al.*, 2003) como suporte para o seu trabalho em determinação de estruturas, possibilitando-lhe poupar tempo para chegar aos resultados conclusivos.

1.10. *Origem homóloga*

A determinação da origem homóloga de compostos similares em grupos taxonômicos diferentes é um dos problemas fundamentais inerentes na aplicação taxonômica de compostos secundários (Hegnauer, 1967). Dois táxons podem sintetizar ou acumular os mesmos produtos por caminhos diferentes. Como

consequência, a mera presença de um mesmo composto em dois ou mais táxons não é necessariamente uma indicação de que os mesmos possuem químicas bem relacionadas. Eles podem refletir uma similaridade evolucionária ou filética, entretanto podem também ser o resultado de processos evolucionários convergentes (Seigler, 1977). Com um conhecimento dos caminhos biossintéticos de metabólitos secundários nas plantas, deve ser possível determinar em que ponto em uma seqüência ocorreu divergência e que mudanças subseqüentes devem ocorrer (Birch, 1974). Na realidade, isso é raramente percebido; por causa de diversos fatores; diversas classes dos compostos não parecem ter exigências estruturais específicas, visto que, em outras, menos variação pode ser tolerada. Por exemplo, a maioria de substâncias fenólicas pode servir como antioxidantes, e muitos compostos lipídeos podem servir para revestimentos de superfície, conquanto sejam neles encontradas as propriedades físicas necessárias. Entretanto, atraentes químicos para polinizadores específicos e diterpenos com atividades hormonais necessitam de biossínteses específicas (Birch, 1974). Muitos produtos de planta surgem por processos simples, tal como a remoção de um grupo ativante 'por exemplo, o fosfato ou o coenzima A (CoA); ou de oxidações, reduções; ou metilações de grupos modificados facilmente (Birch, 1974). Em alguns casos, as quantidades relativas de produtos produzidos podem simplesmente refletir as taxas de duas enzimas que atuam em um precursor comum. As reações altamente prováveis, tais como a introdução de um grupo hidroxila *orto* ou *para* a um fenol existente, ocorrem freqüentemente na natureza. Estes tipos de mudanças são geralmente apenas de menor importância, considerando o significado taxonômico de compostos secundários.

1.11. *Perfil químico de Asteraceae*

Quimicamente, a família Asteraceae apresenta uma diversidade de metabólitos secundários muito ampla, despertando assim um grande interesse de seu estudo no mundo inteiro. Inúmeros deles possuem atividade biológica. A Tabela 5 resume seu perfil químico salientando as partes mais comuns da planta onde são produzidos os referidos metabólitos.

Para a realização deste trabalho, foram feitos levantamentos de ocorrências naturais das classes químicas flavonóides (4561), cumarinas (962) e benzofuranos

(628) em espécies da família Asteraceae registradas na literatura especializada. Em seguida, foram montadas bases de dados em sistemas dBASE III/ EXCEL com as referidas ocorrências.

Tabela 5. Padrão químico de Asteraceae

Classe química	Atividade biológica	Parte do vegetal
PRESENTES EM TODAS AS TRIBOS		
Frutanos do tipo inulina	---	Órgãos de armazenagem
Ácidos graxos	---	Óleos de sementes
Lactonas sesquiterpênicas	Diversas	Principalmente em folhas de gosto amargo
Álcoois de triterpenos pentacíclicos	---	Ésteres em pericarpos de frutos,
Ésteres de ácido caféico	Cinarina –diurética	Folhas
Flavonóides metilados	---	Folhas e flores
PRESENTES NA MAIORIA DAS TRIBOS		
Compostos acetilênicos	Antimicrobianos	Raízes e folhas
Óleos essenciais	Antimicrobianos	Folhas e frutos
Ciclitóis	---	Folhas
Cumarinas	---	Folhas e flores
PRESENTES EM POUCAS TRIBOS		
Poliisopreno	---	Raízes e caules
Alcalóides pirrolizidínicos	Altamente hepatotóxicos	Folhas
Ácidos triterpênicos	---	Flores
Diterpenos	---	Todos os tecidos
Glicosídeos cianogênicos	Tóxicos	Folhas e frutos
Pigmentos de antocloro	---	Flores amarelas
Cromenos	Inseticidas	Folhas e raízes
Amidas de ácido graxo	Inseticidas	Raízes

Dados de números de ocorrências (NOs) e números de oxidação (NOXs) médios das classes químicas estudadas foram também avaliados estatisticamente através de correlações em nível de tribos e subtribos, utilizando o sistema estatístico Sta_bas (Statistica for Windows, Release 4.3 B –Copyright © Inc. 1993).

Visando um estudo mais amplo de correlações e avaliação de aglomerações (*clusters*) dentro de Asteraceae, dados de ocorrências de outras classes químicas de metabólitos secundários presentes nesta família, tais como monoterpenos, sesquiterpenos, diterpenos, triterpenos, lactonas sesquiterpênicas e poliacetilenos, foram também incluídos e correlacionados juntamente com os das três classes que intitulam este trabalho.

As correlações encontradas forneceram equações capazes de prever a incidência de tais metabólitos nesses táxons bem como mostrar a existência de um

possível balanço oxidativo (a partir dos NOXs médios) entre as classes químicas, com um alto grau de confiabilidade (95%). As aglomerações (*clusters*) oriundas de correlações de compostos aromáticos mostraram uma ótima separação das tribos, bem similar à feita pelo botânico Bremer (Brant *et al.*, 2001).

Neste trabalho, foi utilizada preferencialmente a classificação sistemática mais recente de Bremer (1994b), que divide Asteraceae em 4 subfamílias (Barnadesoideae, Carduoideae, Chichorioideae e Asteroideae) e 17 tribos (Anthemideae, Archtoteae, Astereae, Barnadesieae, Calenduleae, Cardueae, Eupatorieae, Gnaphalieae, Helenieae, Heliantheae, Inuleae, Lactuceae, Liabeae, Mutisieae, Plucheeae, Senecioneae, Vernonieae). Apesar de ainda suscitar algumas controvérsias entre taxonomistas atuais, e de o próprio Bremer nela reconhecer pontos ainda polêmicos (Mutisieae, Carduoideae), essa classificação tem sido, de modo geral, bem aceita pela comunidade científica mundial estudiosa de taxonomia de Asteraceae.

II - TAXONOMIA DE ASTERACEAE

2.1. *Taxonomia de Asteraceae segundo Bremer*

Quanto à ordem Asterales — à qual pertence a família Asteraceae — Bremer e seus associados adicionaram-lhe as famílias Alseuosmiaceae, Argophyllaceae, Carpodetaceae, Phellinaceae e Rousseaceae. Eis o que diz James L. Reveal (Reveal, 1997), professor emérito da Universidade de Maryland (USA), a respeito dessa adição: “*Destas, Alseuosmiaceae, Argophyllaceae e Carpodetaceae, em minha opinião, pertencem à Cornidae na ordem Hydrangeales. Phellinaceae pertence à Aquifoliales (Rosidae), enquanto Rousseaceae está associada normalmente com Brexiales ou Celastrales (Rosidae). Eu não posso ver como sua Asterales pode distinguir-se morfologicamente de diversas de suas outras ordens, dados estes elementos diversos*”.

A classificação de Asteraceae de Bremer (Bremer, 1994a) tem como suporte o diagrama apresentado na Figura 8, e as respectivas subtribos são inseridas em tribos de acordo com a Tabela 6.

Diagrama de inter-relações filogenéticas conhecidas e desconhecidas de tribos de Asteraceae, modificadas a partir de Bremer (1994a). As quatro subfamílias reconhecidas aqui são dadas em letras maiúsculas. Os grupos VERNONIÓIDE e HELIANTÓIDE são dois outros clados bem apoiados discutidos no texto. Linhas grossas representam grupos com mais de 2000 espécies; linhas médias, grupos com 1000 a 2000 espécies (Lactuceae, Vernoniaceae, Anthemideae) e linhas finas, grupos com menos de 1000 espécies. As linhas múltiplas para Mutisieae e Helenieae indicam que estas duas tribos são conhecidas como sendo parafiléticas

Figura 8. Diagrama de inter-relações filogenéticas de tribos de Asteraceae

Tabela 6. Tribos e subtribos de Asteraceae segundo a classificação de Bremer

Subfamília	Tribos	Código	Subtribos	Código
BAR	Barnadesieae	BAR	---	---
?	Mutisieae	MUT	Mutisiinae	MUT
			Nassauviinae	NAS
CAR	Cardueae	CAR	Carduinae	CAD
			Carlininae	CAR
			Centaureinae	CEN
			Echinopsidinae	ENP
CIC	Arctoteae	ARC	Arctotinae	ARC
			Gorteriinae	GTR
CIC	Lactuceae	LAC	Crepidinae	CRE
			Cichoriinae	CCR
			Dendroseridinae	DDS
			Hieraciinae	HIE
			Hypochaeridinae	HYP
			Lactucinae	LAC
			Leontodontinae	LTD
			Microseridinae	MIC
			Scorzonerinae	SCO
			Sonchinae	SON
			Stephanomeriinae	STE
CIC	Liabeae	LIA	---	---
CIC	Vernonieae	VER	Centratherinae	CTT
			Elephantopodinae	EPP
			Lychnophorea	LYC
			Piptocarphinae	PCP)
			Rolandrinae	RLD
			Vernoniinae	VER
AST	Anthemideae	ANT	Achilleinae	ACH
			Anthemidinae	ANT
			Artemisiinae	ART
			Cancriniinae	CCR

Tabela 6. Continuação

Subfamília	Tribos	Código	Subtribos	Código
AST	Astereae	AST	Chrysantheminae	CHR
			Gonosperminae	GON
			Handeliinae	HDL
			Leucantheminae	LEU
			Matricariinae	MAT
			Tanacetinae	TAN
			Thaminophyllinae	TMP)
			Ursiniinae	URS
			Grangeinae	GRG
AST	Calenduleae	CAL	Asterinae	AST
			Solidagininae	SOL
AST	Eupatorieae	EUP	---	---
AST	Eupatorieae	EUP	Adenostenmatinae	ADT
			Ageratinae	AGE
			Alomiinae	ALO
			Ayapaninae	APN
			Critoniinae	CRI
			Disynaphiinae	DIS
			Eupatoriinae	EUP
			Fleischmanniinae	FLM
			Gyptidinae	GYP
			Hebecliniinae	HBC
			Hofmeisteriinae	HFM
			Liatrinae	LIA
			Mikaniinae	MIK
			Oxylobinae	OXY
			Praxelinae	PRA
			Trichocoroninae	TCC
			Angianthinae	ANG
			Cassiniinae	CAS
			Gnaphaliinae	GNA

Tabela 6. Continuação

Subfamília	Tribos	Código	Subtribos	Código
AST	Helenieae	HLN	Loricariinae	LRC
			Relhaniinae	RLN
			Baeriinae	BAE
			Chaenactidinae	CHA
			Flaveriinae	FLA
			Gaillardiinae	GAI
			Hymenopappinae	HYM
			Madinae	MAD
			Pectidinae	PEC
			Peritylinae	PER
AST	Heliantheae	HLT	Ambrosiinae	AMB
			Coreopsidinae	COR
			Engelmanniinae	EGM
			Galinsoginae	GLB
			Helianthinae	HLT
			Melampodiinae	MEL
			Pinilosiinae	PNS
			Rudbeckiinae	RUD
			Verbesininae	VER
			Zinninae	ZIN
AST	Inuleae	INU	---	---
AST	Plucheeae	PLU	---	---
AST	Senecioneae	SEN	Blenospermatiinae	BLS
			Senecioninae	SEN
			Tussilagininae	TUS

Código de Subfamílias: BAR – Barnadesioideae; CAR – Carduoideae;
CIC – Cichorioideae; AST – Asteroideae; ? – não-classificada

2.2 . Quimiotaxonomia de Asteraceae

Diversas classes de compostos de plantas são características desta família, notavelmente os flavonóides, as lactonas sesquiterpênicas baseadas em terpenóides, os poliacetilenos derivados de ácidos graxos e os frutanos polissacarídeos (Heywood *et al.*, 1977). Estes últimos são distintivos na bioquímica de Asteraceae, pois são polissacarídeos incomuns que se baseiam em frutose em vez de glucose, que é mais comum em outras famílias de plantas. São conhecidos também como *inulinas* e estão citados na Tabela 5.

Quimicamente, a família Asteraceae apresenta uma diversidade de metabólitos secundários muito ampla, despertando assim um grande interesse de seu estudo no mundo inteiro sob aspectos econômicos, por exemplo, óleo de girassol (*Helianthus annuus*), adoçantes naturais (*Stevia rebaudiana*, esteviosídeos – Wagner and Farnsworth., 1985), indústria de condimentos, cosméticos, perfumes; fármacos a partir de extratos de ervas medicinais de uso popular, utilizadas no combate aos mais diferentes tipos de doenças . Desses extratos têm sido isolados inúmeros metabólitos com atividade biológica / farmacológica cientificamente comprovada: antivirais (*Calendula arvensis*, saponinas triterpenoídicas, Tomasi *et al.*, 1991), antitumorais (*Helianthella quinquinervis*, o benzofurano 6-metoxi-tremetona), analgésicos e antiinflamatórios (*Cirsium subcoriaceum*, o flavonóide pectoliarina; Mariano *et al.*, 1996.), antifúngicos (*Senecio pampae*; furanoeremofilanos), antibacterianos, antimaláricos (*Artemisia annua*, a lactona sesquiterpênica artemisinina; Ngo *et al.*, 1997), antileishmaniose (*Jasonia glutinosa*, o sesquiterpeno [11R]-eudesm-4(14)-eno-5a, 11-12-triol, Villaescusa-Castillo *et al.*, 2000). Tais metabólitos servem como padrões ou protótipos para desenvolvimento de seus derivados bem como sua síntese.

São também comuns na família algumas ervas que agem como inseticidas naturais , pois contêm alcanidas especiais (*Spillanthus mauritiana*, Jondico, 1986); piretrinas e ésteres de piretrolona, substâncias existentes nas flores do *Chrysanthemum (Pyrethrum) cinerariaefolium*. Aqui vale a pena mencionar que inseticidas piretróides são compostos sintéticos que apresentam estruturas semelhantes à da piretrina: aletrina, resmetrina, decametrina, cipermetrina e fenpropanato.

No que diz respeito à sua quimiotaxonomia usando caracteres micromoleculares, flavonóides, lactonas sesquiterpênicas e poliacetilenos específicos já podem ser tomados como marcadores químicos em âmbito de família (os primeiros) e tribos (os dois últimos).

2.3. Caracteres macromoleculares

Quanto ao uso caracteres macromoleculares para estudos filogenéticos, têm sido feitos numerosos trabalhos englobando citocromo c, seqüenciamentos de DNA de núcleo de células (nDNA) e DNA de cloroplasto (cpDNA), mormente em espécies dentro de gêneros de Asteraceae e outras famílias de plantas (Doyle, 1993; *Park et al.*, 2001; *Kim et al.*, 1996). Um bom exemplo é o seqüenciamento do citocromo c (Figura 9) de várias espécies de plantas (Tabela 7).

Na figura acima, a estrutura vermelha com a bola no meio representa o grupo heme, que é importante para a função desta proteína na respiração celular. A tira azul larga representa as estruturas helicoidais da proteína. A estrutura amarela é parte de uma histidina particular que é importante para a atividade do citocromo c. Histidina é um dos cerca de vinte aminoácidos que estão presentes em todas as proteínas.

Figura 9. Imagem 3-D do citocromo c

O número de mutações necessárias ao nível de DNA para mudar tirosina para serina ou serina para fenilalanina pode ser determinado de um conhecimento do código tríplice básico. Nestes dois casos, mutações únicas separam estes aminoácidos. Através de contraste, a mudança de fenilalanina para metionina (Tabela 8) requer, no mínimo, duas mutações, uma mudança que

poderia ter ocorrido por um citocromo *c* de intermediário com leucina presente na posição 65.

Tabela 7. Matriz de diferenças de aminoácidos em citocromo *c* de seis espécies de plantas superiores (Boutler *et al.*, 1972; Harborne and Turner., 1984)

Espécie	<i>Ginkgo</i>	<i>Triticum</i>	<i>Helianthus</i>	<i>Vigna</i>	<i>Ricinus</i>	<i>Sesamum</i>
<i>Ginkgo biloba</i>	0					
<i>Triticum aestivum</i>	17	0				
<i>Helianthus annuus</i>	20	15	0			
<i>Vigna radiata</i>	20	15	12	0		
<i>Ricinus communis</i>	17	14	13	10	0	
<i>Sesamum indicum</i>	18	16	11	9	9	0

Comentário: O citocromo *c* da *Ginkgo biloba* (uma gimnosperma primitiva) é, na expectativa, mais extensamente divergente, entre 17 e 20 substituições que são diferentes. Tal análise, entretanto, não fornece a maioria do conteúdo de informação presente em cada seqüência.

Tabela 8. Substituição de aminoácidos na posição 65 em citocromo *c* de filos diferentes

Aminoácido 65	Tirosina	→	Serina	→	Fenilalanina	→	Metionina
Filo	Plantas superiores		Fungos		Insetos		Vertebrados
Mudanças mutacionais	Uma		Uma		Duas		

Comentário: A troca de fenilalamina (em insetos) para metionina (em vertebrados) requer, no mínimo, duas mutações.

Outro exemplo de uso de caráter macromolecular, nrDNA (núcleo de DNA de ribossoma) foi realizado para *Encelia* e seus parentes. A filogenia (Figura 10) é baseada em caracteres fenotípicos, e os resultados iniciais do espaçador transcrito intercistrônico de nrDNA (*ITS*) apóiam fortemente o esquema de ramificação basal. Caracteres de apoio estão relacionados logo abaixo do dendrograma.

1- O gênero *Encelia*. Cipsela com entalhe apical constricto; 2- "Clado *frutescens*". Não há tubos de resina, benzopiranos ou benzofuranos; 3- Capítulos frutificantes eretos; 4- Tricomas amplos de base multicelular; 5- "Clado *californica*". Único dímero benzopirano-benzofurano que não absorve raios UV; 6- Corolas de disco marrom, pêlos moniliformes; 7- Estigmas marrons; 8- Capitulescência paniculada, pêlos ondulados de base unicelular.

Figura 10. Filogenia de *Encelia* realizada com base em caracteres fenotípicos

OBJETIVOS

Este trabalho objetivou:

1- Complementar bancos de dados de ocorrências de flavonóides, cumarinas e benzofuranos produzidos em Asteraceae.

2- Avaliar, através de programas estatísticos, correlações de ocorrências, de estados de oxidação e passos oxidativos entre os metabólitos supracitados. Assim, é possível estabelecerem-se equações capazes de prever a incidência dessas substâncias em nível de tribos e subtribos da família e averiguar a existência de um possível balanço oxidativo entre as classes químicas correlacionadas.

3- Mostrar mais especificamente se flavonóides, cumarinas e benzofuranos podem ou não servir como marcadores quimiotaxonômicos nessa família de plantas.

III – METODOLOGIA

3.1. Dados de ocorrências

“Número de ocorrências (NO) é um parâmetro químico que fornece o grau de relevância de uma determinada categoria metabólica (X) para um táxon escolhido (A). Na prática, se para duas espécies de um táxon foram registradas cinco substâncias diferentes da mesma categoria para uma e sete para a outra, considera-se para o táxon referente à categoria micromolecular em questão $NO = 12$, mesmo que haja coincidência de substâncias nas duas espécies. Ao contar n vezes uma mesma substância encontrada em n espécies diferentes, consegue-se caracterizar a tendência de sua produção no táxon” (Gottlieb et al., 1996).

O levantamento das ocorrências naturais dos metabólitos em questão, ou seja, flavonóides, cumarinas e benzofuranos, foi feito a partir da literatura especializada englobando revisões anteriores a 1982 e subseqüentes publicações em *Chemical Abstracts*, sendo a busca realizada no *Subject Index*, através de palavras-chaves tais como nomes de categorias químicas ou de táxons. Complementando as informações extraídas do *Chemical Abstracts*, foram consultados minuciosamente as revistas e periódicos de maior peso em estudo de produtos naturais como, por exemplo, *Phytochemistry*, *Planta Medica*, *Natural Product Reports*, *Journal of Natural Products*, *Fitoterapia*, etc. até julho de 2001. O livro especializado em fitoquímica, *The Natural Coumarins: Occurrence, Chemistry and Biochemistry* (Murray, 1982), foi bastante útil no levantamento das ocorrências de cumarinas. A revisão de Proksch et al., 1983, foi muito importante como ponto de partida para o levantamento de ocorrências de benzofuranos.

Com esse levantamento, criou-se uma base de dados na qual foram enfatizados os seguintes pontos:

1- Ocorrências apenas naturais dos metabólitos em questão em espécies estudadas. Produtos derivados de tais metabólitos, oriundos de reações químicas ao longo do processo de extração e identificação, são considerados artefatos e não entram, portanto, no cômputo das ocorrências.

2- Partes da planta estudada, que podem ser toda (massiva) ou específicas: aéreas, folhas, flores, caules, raízes, frutos, tricomas, etc.

3- Local e data de coleta (se possível): Tópico de suma importância para estudo de populações, cujo metabolismo pode ser bastante influenciado pelas condições de solo e ambientais.

4- O registro do nome trivial do metabólito, que nem sempre é dado.

5- O registro do gênero e espécie da qual o metabólito de interesse foi extraído.

Fez-se uma ficha-modelo (Figura 11) para o preenchimento dos dados, conforme está exemplificada abaixo e por si só já é bastante explicativa.

Ficha N°. 353 (como exemplo)

Gênero/espécie : *Artemisia incanescens*

Referência : *Phytochemistry*, 1986, 25, 2357.

<p>Nome trivial: Escopoletina</p>	<p>N° de Esqueleto: 321</p> <p>NOX: +2</p> <p>Parte da planta: ADAP (partes aéreas secas ao ar)</p> <p>Local e data da coleta: Teruel / Espanha</p>
---	---

Obs.: Os dados das fichas foram digitados em programa dBASE, que dispõe de subprogramas capazes de colocar devidamente as espécies levantadas em suas tribos e subtribos correspondentes e de acordo com a classificação de Bremer (1994a) para a família Asteraceae.

Figura 11. Ficha-modelo usada para o levantamento bibliográfico

A seguir, os dados das fichas foram digitados nos campos específicos criados pelo sistema dBASE III. Este possui um programa que insere automaticamente nas fichas de cada ocorrência os nomes das tribos e subtribos relacionadas aos gêneros/espécies registrados, segundo a classificação de Bremer (1994b). Em montagens de tabelas, utilizou-se também o sistema EXCEL, compatível com o dBASE III e com o Sta_bas (este é específico para uso estatístico).

Os esqueletos carbônicos de cada classe química foram codificados com numerações empíricas descritas abaixo:

- 1-66, diterpenos (DITE);
- 67-129, lactonas sesquiterpênicas (LACT);
- 130-142, triterpenos (TRIT);
- 143-272, sesquiterpenos (SESQ);
- 273-320, monoterpenos (MONO);
- 321-376, cumarinas (CUMA);
- 377-404, benzofuranos (BENZ);
- 405-413, poliacetilenos (POLI);
- 414-420, flavonóides (FLAV).

Essa codificação facilita a triagem de frequências de esqueletos específicos, que podem ser utilizados como marcadores quimiotaxonômicos de um determinado táxon.

3.2. Estudo de passos oxidativos

Utilizou-se aqui o conceito de passo oxidativo (PO), que é a diferença entre o NOX do metabólito em questão e o NOX de seu precursor biogenético dividida por dois. POs são importantes em estudo de avanço evolutivo de táxons.

Como exemplo de estudo de PO, tomou-se a molécula 5-metil-cumarina (Fig.12) com o valor de estado de oxidação de cada átomo de C de sua estrutura. A soma dos estados de oxidação dos carbonos = - 4, de acordo com a regra de Cram-Hendrickson , que define como valores de estado de oxidação de carbono ligado covalentemente a C , C-C = 0; H, C-H = -1; C-heteroátomo, ou seja, C-O, C-N, C-S, C-halogênio = + 1 .

Figura 12. Exemplo de valores de oxidação de átomos de C de 5-Metil-cumarina, seguindo as regras de Cram-Hendrickson (NOX da molécula = a soma dos estados de oxidação dos átomos de C).

O NOX dos carbonos esqueléticos de 5-metil- cumarina, seguindo a regra de Cram-Hendrickson, porém levando-se em conta só os carbonos do esqueleto benzopirônico = 0; o NOX do esqueleto do precursor biogenético (ácido cinâmico, por exemplo) = -4. Estima-se, portanto, o PO conforme apresentado na Figura 13.

Figura 13. Estudo de passo oxidativo (PO).

Resultado = Um avanço evolutivo de 2 passos do táxon com relação ao acúmulo do metabólito em questão.

Nota: NOs e NOXs de mais 6 classes químicas da família foram incluídos nos testes para uma avaliação mais abrangente.

O estudo de passos oxidativos é muito importante para correlacionar quantitativamente caracteres morfológicos e micromoléculas de plantas e graus de evolução de seus táxons. O oxigênio tem aqui um papel importante conforme os dizeres do Prof. Dr. Otto Richard Gottlieb, renomado pesquisador de produtos naturais e conhecido internacionalmente na comunidade científica: (Gottlieb, 1989)

"Recentemente foram propostos métodos para a correlação quantitativa de características morfológicas e micromoleculares de plantas. Sua aplicação a grupos de plantas de posição hierárquica elevada mostra que evolução morfológica é acompanhada por oxidação progressiva de metabólitos secundários dentro de classes biossintéticas particulares. Em angiospermas, a substituição evolutiva de formas lenhosas para herbáceas é acompanhada pelo abandono gradativo do

caminho do chiquimato para a produção de metabólitos secundários e o engrandecimento concomitante dos caminhos do acetato e do mevalonato para a síntese de policetídeos e terpenóides de alto nível de oxidação. Estes resultados sugerem que evolução micromolecular é dependente da variável teor de oxigênio da atmosfera. Entretanto, o oxigênio não age diretamente na evolução micromolecular, mas propulsiona a evolução de dispositivos enzimáticos protetores, tais como eterificação seletiva, formação de base de Schiff ou redução para a regulação de meias-vidas durante a mudança oxidativa e para a orientação de caminhos biossintéticos de micromoléculas. Ambos os fenômenos são essenciais na diversificação de metabólitos secundários e por isso são requisitos para a flexibilidade da adaptação de um organismo ao ambiente”.

3.3. Análise dos dados

3.3.1. Uso de regressão linear múltipla

A finalidade geral da regressão linear múltipla é aprender mais sobre o relacionamento entre diversas variáveis independentes ou previsoras e uma variável dependente ou variável de critério.

“Quando se quer investigar se uma variável está correlacionada concomitantemente a várias outras, considera-se a primeira como variável dependente e as outras como variáveis independentes, e aplica-se aos dados a seguinte fórmula:

$$y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + \dots + b_nx_n \quad \text{em que:}$$

y = é a estimativa da variável dependente;

x = variável independente;

a = constante = intercepto múltiplo;

b = constante = coeficientes de regressão.

“A análise de regressão múltipla é trabalhosa, pois envolve a construção e multiplicação de matrizes tanto maiores quanto maior for o número de variáveis independentes analisadas. Assim, é necessário realizar tal análise em

computadores. Portanto, aqui nos preocupamos com a interpretação de resultados de análise de regressão múltipla" (Conti, 2002).

Neste trabalho, utilizou-se regressão linear múltipla, cujo objetivo principal foi encontrar equações empíricas capazes de prever a ocorrência de flavonóides, cumarinas e benzofuranos em 17 tribos e 82* subtribos de Asteraceae. A escolha deste método estatístico se deveu ao fato de procedimentos de regressão múltipla já serem muito extensamente usados em pesquisa nas ciências sociais, naturais, exatas e noutras, em que o (a) pesquisador(a) lida com muitas variáveis e deseja achar correlações entre estas. Inúmeros exemplos de aplicação de regressão múltipla estão na literatura abrangendo os mais diferentes tipos de conhecimento humano (Valeriano e Picini, 2000; Buzuk *et al.*, 2002; Barker *et al.*, 2002). Os resultados obtidos com este método são bastante confiáveis ($p < 0,05$). Dentro da estatística, outras metodologias podem ser exploradas em estudos desse enfoque, por exemplo, técnicas de PCA (*principal component analysis* / análise de componente principal) (Alvarenga *et al.*, 2001).

O banco de dados resultante — ocorrências de classes químicas e NOXs médios destas —, foi analisado no sistema computacional Statistics Standard, Sta_bas (versão de 1993 para Windows). Este sistema possui programas estatísticos diversificados, capazes de elaborar, por exemplo, matrizes de correlações, regressão linear, regressão múltipla, diagramas bi- e tridimensionais, análises canônicas, análise de fatores, análise de agrupamentos (*clusters*), etc. Neste trabalho, foi utilizado o *software* acima citado, com o qual foi feito o seguinte:

- I- Regressões lineares múltiplas de ocorrências de classes químicas em tribos de Asteraceae utilizando 17 casos com 9 variáveis em cada experimento.
- II- Regressões lineares múltiplas de ocorrências de classes químicas em subtribos de Asteraceae utilizando 82 casos com 9 variáveis em cada experimento.
- III- Correlações entre números de oxidação médios de classes químicas de Asteraceae em nível de tribos e subtribos.
- IV- Montagem de equações empíricas para previsão de ocorrências de flavonóides, cumarinas, benzofuranos, monoterpenos, sesquiterpenos, diterpenos, triterpenos, lactonas sesquiterpenênicas e poliacetilenos em tribos e subtribos de Astraceae.

- V- Formação de agrupamentos(*clusters*), visando um estudo de grau evolutivo ou parentesco entre as tribos e , possivelmente, subtribos de Asteraceae.
- VI- Correlações entre números de oxidação médios de classes químicas de Asteraceae, em nível de tribos e subtribos.

IV - RESULTADOS E DISCUSSÃO

Antes de apresentar os resultados e respectivas discussões, achou-se melhor fazer uma abordagem sucinta sobre as classes de metabólitos secundários de Asteraceae avaliadas nesta dissertação, focalizando suas definições, estruturas químicas, possíveis biogêneses / biossínteses e alguns mecanismos químicos envolvidos nestas. O Esquema 1 (Mann, 1994) apresenta, de modo geral, as principais rotas biossintéticas de metabólitos secundários em plantas. As demais rotas passam a ser mais específicas para cada tipo de classe estudada.

4.1. Terpenos

Soube-se por séculos que a co-destilação de muitos materiais de planta com vapor, uma técnica chamada destilação a vapor, produz uma mistura perfumada de líquidos conhecidos como óleos essenciais da planta. Por milhares de anos, tais extratos de planta foram usados como medicamentos, condimentos e perfumes. A investigação de óleos essenciais teve também um papel de destaque na emergência da química orgânica como uma ciência, durante o século dezanove.

Quimicamente, óleos essenciais de plantas consistem amplamente de misturas de lipídeos chamados terpenos. Os terpenos são moléculas orgânicas pequenas com uma diversidade imensa de estrutura. Milhares de terpenos diferentes são conhecidos. Alguns são hidrocarbonetos, e outros contêm oxigênio; alguns são moléculas de cadeia aberta, e outros contêm anéis.

Todos os terpenos são relacionados, não obstante suas diferenças estruturais aparentes. De acordo com a regra do isopreno proposta por Leopold Ruzicka (1887-1976, Prêmio Nobel, 1939), terpenos podem ser imaginados como oriundos da união *cabeça-cauda* de unidades de isopreno (2-metil-1,3-butadieno ou isopentenil) de cinco carbonos. O carbono 1 é chamado cabeça da unidade do isopreno, e o carbono 4 é a cauda. Por exemplo, o mirceno (Esquema 2) contém duas unidades de isopreno unidas cabeça à cauda, formando uma cadeia de oito carbonos com duas ramificações de um carbono; α -pineno (Esquema 2) contém similarmente duas unidades de isopreno montadas em uma estrutura cíclica mais complexa.

Esquema 1. Rotas biossintéticas de produtos naturais em plantas.

Esquema 2. Formação hipotética de mirceno (acima) e α -pineno (abaixo) através de unidades de isopreno, ligação cabeça-cauda.

Os terpenos são classificados de acordo com o número de unidades de isopreno que contêm. Assim, os monoterpenos são as substâncias de 10 carbonos biossintetizadas teoricamente a partir de duas unidades de isopreno; os sesquiterpenos são moléculas de 15 carbonos, ou seja, de três unidades de isopreno, e assim por diante (Tabela 9).

Tabela 9. Classificação de terpenos

Átomos de carbono	Unidades de isopreno	Classificação
10 (C ₁₀)	2 (C ₅)	Monoterpeno
15 (C ₁₅)	3 (C ₅)	Sesquiterpeno
20 (C ₂₀)	4 (C ₅)	Diterpeno
25 (C ₂₅)	5 (C ₅)	Sesterpeno
30 (C ₃₀)	6 (C ₅)	Triterpeno
40 (C ₄₀)	8 (C ₅)	Tetraterpeno

Mono- e sesquiterpenos são encontrados primariamente em plantas, porém os terpenos de massas moleculares mais elevadas ocorrem em plantas e animais, e muitos têm papéis biológicos importantes. O triterpeno lanosterol, por exemplo, é o precursor do qual todos os hormônios esteróides são feitos; o tetraterpeno β -caroteno é uma importante fonte supridora de vitamina A.

4.1.1. Monoterpenos (C_{10}):

Na família Asteraceae, estão presentes em quase todas as tribos, porém ainda não se explorou sua importância quimiotaxonômica isoladamente. Podem ter estruturas acíclicas (por exemplo, mirceno) ou cíclicas de um anel (d-limoneno) ou de duas estruturas de anel (α - e β -pineno, 3-careno). São produzidos em inúmeras famílias de plantas utilizando a rota biossintética descrita no Esquema 1. Alguns monoterpenos importantes são apresentados na Figura 14.

Figura 14. Exemplos de alguns monoterpenos

4.1.2. Sesquiterpenos / Lactonas sesquiterpênicas (C_{15})

As ocorrências de ambos os tipos de metabólitos foram somadas e representadas com a sigla SL ou SELA para os cálculos estatísticos. Presentes em quase todas as tribos, exceto em BAR e MUT (em BAR, nenhum registro de ocorrência de ambos os metabólitos; em MUT, nenhum registro de ocorrência de sesquiterpenos). Todavia as lactonas sesquiterpênicas é que são, até o presente, de maior importância quimiotaxonômica em espécies e gêneros da família. A maioria das lactonas é de gosto amargo, e várias delas são tóxicas. Estas propriedades

sugerem que lactonas sesquiterpênicas estão envolvidas contra herbívoros e parasitas de plantas. Matricina (Fig. 15) é uma dos primeiros representativos definidos quimicamente do grupo muito grande e estruturalmente diverso de lactonas sesquiterênicas de Asteraceae. Esqueletos guaionolídeos, pseudoguaianolídeos e furanoeremofilanolídeos tais como petasolídeos são abundantes em Asteraceae.

Figura 15. Matricina, um guaionolídeo de *Matricaria chamomilla* L. / Anthemideae (Asteraceae).

Outros exemplos de sesquiterpenos são apresentados na Figura 16.

Figura 16. Exemplos de sesquiterpenos isolados de espécies de Asteraceae

4.1.3. Diterpenos (C₂₀)

Podem ser bicíclicos, pertencentes ao grupo do ácido grindélico. Classificam-se como regulares (Figura 17) do tipo labdano- óxido de manool (16) e rearranjados, do tipo clerodano (17); tricíclicos de tipos pimarano (18) e abietano (19); tetracíclicos de tipos caurano (20) e estacano (21); pentacíclicos do tipo traquilobano, que são combinações de diterpenóides pertencentes a uma ou várias classes e grupos mencionados com características estruturais, tais como os sistemas furanóide (16), butenolídeo (17), lactona (17) ou fenólico (19).

Diterpenos são abundantes na família Asteraceae, especialmente nas tribos Heliantheae (gêneros *Helianthus*, *Montanoa*) com predominância de esqueletos ent-cauranos; Eupatorieae (*Brickellia*, *Stevia*) e Astereae (*Baccharis*), onde predominam

esqueletos de labdano normal, ent-labdano, clerodano normal e ent-clerodano; Inuleae (*Helichrysum*) com predominância de esqueletos de ent-labdano; Calenduleae (gêneros *Garuleum* e *Osteospermum*) com presença de esqueletos de pimarano normal.

Figura 17. Exemplos de diterpenos isolados de espécies de Asteraceae

4.1.4. Triterpenos (C₃₀)

Asteraceae é acumuladora de triterpenos. Monóis e dióis do tipo oleanol (1), ursanol (2) e lupeol (3) são os triterpenos mais característicos da família (Figura 18).

Na luz da evidência atual, a co-ocorrência dos monóis e dióis mencionados representam uma tendência metabólica na família como um todo. A síntese de ácidos triterpênicos, saponinas e triterpenóides raros, tal como a xionona* (4), pode tornar-se taxonomicamente útil no futuro, em níveis genérico e tribal.

* (*shionone* portuguesa).

4.1.5. Rota biogenética geral dos terpenóides

4.1.5.1. Formação de policetídeos

A reação de condensação de Claisen geral envolve um éster com um α -metileno que age como um componente enol gerando um carbânion e um segundo éster para agir como substrato para a substituição nucleofílica. O resultado é a formação de β -cetoésteres (Esquema 3).

Figura 18. Triterpenos característicos de espécies de Asteraceae

Esquema 3. Formação de policetídeos via condensação de Claisen.

Uma vez que a rota biogenética dos terpenóides é baseada ou na condensação de Claisen ou aldólica, são dados mais detalhadamente, a seguir, os

esquemas de ambas as reações químicas, que apresentam mecanismos muito similares (Esquema 4 e Esquema 5).

CONDENSAÇÃO DE CLAISEN

Esquema 4. Condensação de Claisen e mecanismos químicos envolvidos nesta reação (adição nucleofílica à carbonila) para formar policetídeos.

CONDENSAÇÃO ALDÓLICA

O produto da adição aldólica de um aldeído é um aldeído β -hidroxilado.

Esquema 5. Condensação aldólica e mecanismos químicos envolvidos nesta reação (adição nucleofílica à carbonila). Uma outra rota para formar policetídeos.

Comentário: O aldeído pode ser parcialmente convertido em enolato por catálise alcalina. O enolato assim formado pode atacar outra molécula de aldeído dando origem a um aldol (uma molécula com um grupo aldeído e um grupo álcool). Apresenta-se, a seguir, o mecanismo desta reação descrito no Esquema 5.

4.1.5.2. Precursores

O pirofosfato de geranila é o precursor de todos os monoterpenos. O limoneno, por exemplo, um monoterpeno encontrado em muitos óleos cítricos, origina-se do pirofosfato de geranila através de uma isomerização *cis-trans* na dupla ligação seguida pelo deslocamento nucleofílico interno do grupo pirofosfato e perda subsequente de um próton.

Uma reação adicional do pirofosfato de geranila com isopentenil resulta no pirofosfato de farnesila de 15 carbonos, o precursor de todos os sesquiterpenos. Farnesol, o álcool correspondente, é encontrado no óleo de citronela e no óleo de limão. No Esquema 6, são apresentados os mecanismos de formação do pirofosfato de farnesila e do esqualeno.

Esquema 6. Formação do pirofosfato de farnesila (A.) e do esqualeno (B).

No Esquema 7, é apresentada a rota biogenética de terpenóides em geral.

Esquema 7. Rota biogenética de terpenóides a partir dos precursores pirofosfato de geranila e pirofosfato de farnesila.

4.2. Poliacetilenos

4.2.1. Ácidos graxos saturados

Como já foi mostrado em 4.1.5.1, a formação de poli- β -cetoacil-CoA ocorre também quando para a biossíntese de ácido graxo por condensação de acetil-coenzima A com malonil-coenzima A. Este, por sua vez, é geralmente derivado por carboxilação de acetil-coenzima A.

A formação de ácidos graxos insaturados, e.g. ácido oléico, que são também metabólitos primários, pode ocorrer através de, no mínimo, duas rotas: uma aeróbica

e outra aneróbica. Essencialmente, ambas envolvem desidrogenação de um ácido graxo totalmente saturado conforme mostra o Esquema 8.

Esquema 8. Rota biogenética de poliacetilenos

Já na década de 60, um número considerável de produtos naturais acetilênicos com 9 a 18 átomos de carbono, derivados biogeneticamente através da rota do

acetato-malonato, foram isolados de plantas e fungos (Bohlmann, 1967; Jones, 1966). Dessa época até o presente, Asteraceae, dentre outras famílias de plantas (Apiaceae, Oleaceae, etc.), tem-se destacado pela produção de poliacetilenos, metabólitos de grande importância quimiotaxonômica.

Com base em evidência obtida principalmente por experimentos de incorporação de precursor, uma possível rota biogenética a partir do ácido esteárico foi proposta, conforme delineada no Esquema 8, uma adaptação ilustrativa que utiliza idéias de Nakanishi *et al.*, 1975 e Sorensen, 1977, ambos estudiosos de ácidos graxos e poliacetilenos, respectivamente.

A rota biogenética acima envolve ambos os processos de desidrogenação de ácidos menos insaturados para mais insaturados e de passos de encurtamento da cadeia, começando de ácidos de C₁₈ (1 - 4) e levando a acetilenos de C₁₆ (6), C₁₃ (7, 8, 9) e C₁₀ (5). Os passos de encurtamento da cadeia podem proceder via processos de β -oxidação. Fissão oxidativa direta no estágio de C₁₈ pode também levar a ácido de C₁₀.

Compostos acetilênicos que possuem números ímpares de átomos de carbono, e.g. C₁₇, C₁₃, C₁₁ e C₉, podem surgir através de reações envolvendo processos de α -oxidação bem como ω -oxidação. Os intermediários biossintéticos produzidos em reações do tipo acima são freqüentemente sujeitos a modificação posterior, e.g. oxidação, redução, descarboxilação, etc., levando a uma grande variedade de compostos acetilênicos de ocorrência natural. Entretanto, metabólitos acetilênicos que não são derivados de intermediários de ácidos graxos são também conhecidos, incluindo terpenóides, carotenóides e outros (Ross, 1972).

Exemplos de alguns poliacetilenos de Asteraceae são apresentados na Figura 19.

Figura 19. Exemplos de alguns poliacetilenos de Asteraceae

4.3. FLAVONÓIDES

Os flavonóides são compostos polifenólicos que possuem 15 átomos de carbono; os dois anéis benzênicos estão unidos por uma cadeia linear de 3 carbonos, os quais podem ser representados como o sistema $\text{C}_6\text{-C}_3\text{-C}_6$. (Figura 20).

Figura 20. Esqueleto básico de um flavonóide

A estrutura química dos flavonóides é baseada num esqueleto C_{15} com um anel cromano (C) suportando um segundo anel aromático B na posição 2, 3 ou 4. Em poucos casos, o anel C heterocíclico ocorre numa forma isomérica aberta ou é substituído por um anel de 5 membros (Figura 21).

Figura 21. Tipos de anéis C encontrados em flavonóides

A ponte oxigênio envolvendo o átomo de carbono central (C₂) da cadeia de 3C ocorre em um número um tanto limitado de casos, onde o heterocíclico resultante é o de tipo furano.

Os flavonóides têm sua origem a partir de duas rotas: a do chiquimato e a do acetato (mevalonato vem do acetato), conforme apresentado na Esquema 9.

ROTA BIOSSINTÉTICA DE FLAVONÓIDES E LIGNANAS

Esquema 9. Esquema hipotético para a formação de 4, 2', 4', 6'- tetraidroxichalcona a partir de *p*-cumaroil-CoA e malonil-CoA

Comentário: As lignanas têm como precursor, no esquema acima, o álcool cumarílico, que se formaria a partir do *p*-cumaroil-CoA. Porém os álcoois coniferílico e sinapílico — formados também na rota do chiquimato — podem ser seus precursores (Dewick, P. M., 1998).

Classificação de Flavonóides

Vários subgrupos de flavonóides são classificados de acordo com os padrões de substituição do anel C. O estado de oxidação do anel heterocíclico e a posição do anel B, ambos são importantes na classificação. São dados exemplos de 6 subgrupos principais:

A- Chalconas: São derivadas de três unidades de acetato e uma unidade de ácido cinâmico como é mostrado abaixo na Figura 22.

Figura 22. Exemplo de uma chalcona

B- Flavonas: Cujos exemplos são apresentados na Figura 23, geralmente ocorrem em famílias herbáceas, e.g. Lamiaceae, Apiaceae, Asteraceae.

Apigenina (aipo, salsão / *Apium graveolens*, Apiaceae)

Luteolina

(rabo-de-cavalo / cavalinha; *Equisetum arvense*, Equisetaceae)

Figura 23. Exemplos de flavonas

C- *Flavonóis*: Geralmente ocorrem em angiospermas herbáceas. Exemplos na Figura 24.

Kaempferol (sabugueiro / *Sambucus nigra*, Adoxaceae; cássia-verrugosa / *Cassia senna*, Fabaceae; urtiga-branca / *Lamium album*, Lamiaceae; bistorta / *Polygonum bistorta*, Polygonaceae)

Quercetina (arruda / *Ruta graveolens*, Rutaceae; trigo-sarraceno / *Fagopyrum esculentum*, Polygonaceae; sabugueiro / *Sambucus nigra*, Adoxaceae)

Figura 24. Exemplos de flavonóis

D- *Flavanonas*: É apresentado um exemplo abaixo:

Naringenina

Figura 25. Exemplo de uma flavanona

E- *Antocianinas*: Depois da clorofila, são o grupo mais importante e mais difundido de pigmentos solúveis em água que conferem cor às plantas (responsáveis por quase todas as cores — rosa, escarlata, vermelha, malva, violeta e azul — nas pétalas, folhas e frutos de plantas superiores). Antocianidinas são também pigmentos de plantas, constituídos da parte aglicona das antocianinas. Formam uma grande família de diferentes compostos coloridos e ocorrem em incontáveis misturas em praticamente todas as partes da maioria das plantas superiores (Hahlbrock, 1981) (Hahlbrock, K. 1981 " Flavonoids ". In *The Biochemistry of Plants*, Vol. 7, 425-456. Copyright © 1981 by Academic Press, Inc. Exemplos abaixo:

Figura 26. Exemplos de antocianina (A) e antocianidina (B)

F- Isoflavonóides

Em contraste com a maioria de outros flavonóides, isoflavonóides têm uma distribuição taxonômica um tanto limitada, ocorrendo principalmente dentro de Fabaceae. A maioria de nosso conhecimento sobre a biossíntese de isoflavonóides origina-se de estudos com isótopos radioativos, alimentando cinamatos marcados com ^{14}C .

Os isoflavonóides são incolores. Verificou-se que acetato forma o anel A e que fenilalanina, cinamato e derivados do cinamato são incorporados no anel B e nos átomos C-2, C-3 e C-4 do anel heterocíclico.

Uma vez que chalconas e flavonas são eficientes precursores de isoflavonóides, a migração arílica do anel B solicitada da posição 2 anterior ou posição *beta* para a posição 3 ou posição *alfa* do precursor fenilpropanoídico deve ocorrer depois da formação do esqueleto básico C_{15} .

Figura 27. Exemplo de um isoflavonóide

A maioria dos flavonóides (flavanonas, flavonas, flavonóis e antocianinas) possuem o anel B na posição 2 do anel heterocíclico. Nos isoflavonóides, o anel B ocupa a posição 3 (Figura 27).

G- Antocianidina

Antocianidina é uma conjugação estendida feita da aglicona de antocianinas glicosídeas. Depois da clorofila, antocianinas são o mais importante grupo de pigmentos de plantas, visíveis ao olho humano.

As antocianidinas constituem uma família grande de compostos coloridos e ocorrem em inúmeras misturas em praticamente todas as partes da maioria das plantas superiores. São de grande importância como pigmentos de frutos e, portanto, são usadas como corantes de sucos de frutas, vinhos e algumas bebidas.

Figura 28. Exemplo de antocianidina

As antocianidinas em *Hydrangea sp.* (por exemplo, hortênsia) colorem-na de vermelho em solo ácido e de azul em solo alcalino.

4.4. CUMARINAS

Cumarinas são benzo- α -pironas ou α -cromonas, e possuem, portanto, o mesmo esqueleto carbônico que os ácidos cinâmicos. Na realidade, são α -lactonas dos ácidos *o*-hidroxicinâmicos. O nome “cumarina” surgiu a partir do heterociclo oxigenado (1) da espécie de árvore nativa de Guaiana, *Coumarona odorata* = *Dipterix odorata*, isolado por Vogel em 1820. De suas sementes obteve-se cristalizada a primeira benzo- α -pirona não-substituída.

As cumarinas se acham mais difundidas nas famílias de plantas seguintes: Apiaceae, Papillonaceae (Fabaceae), Rutaceae, Lamiaceae, Asteraceae,

Solanaceae e Poaceae. Em particular, sua concentração é maior em raízes, frutos e sementes. Alguns microrganismos produzem também cumarinas. Como exemplos, temos as espécies de fungo *Aspergillus niger* e *A. flavus*, que produzem aflatoxinas, responsáveis por lesões hepáticas e morte de animais alimentados com ração contaminada por tais micotoxinas; *Streptomyces niveus* e *S. spheroides* produzem o antibiótico novobiocina, cujo esqueleto é formado por 3-amino-4,7-diidroxi-8-metil-cumarina. Atua contra bactérias gram-positivas e especialmente contra infecções por *Staphylococcus*, nas quais não se podem usar penicilina e outros antibióticos.

Algumas cumarinas e psoralenos (furanocumarinas) possuem atividade antitumoral; cumarinas naturais e sintéticas, substituídas com heterociclos em C-3, costumam apresentar atividades antimicrobiana e diversas outras atividades biológicas. Por muito tempo, a cumarina foi usada como corretor de odor e sabor em produtos farmacêuticos e na indústria alimentícia, entretanto, devido a lesões hepáticas e propriedades cancerígenas observadas em alguns experimentos com animais, seu uso está proibido em ambos os casos. Por outro lado, o dicumarol (Figura 29) [cumarina dímera, 3,3'-metilen-bis (4-hidroxycumarina) (2)] é um anticoagulante do sangue muito eficaz, usado via oral em casos de possível trombose. As cumarinas umbeliferona, esculetina e herniarina têm sido adicionadas aos preparados de proteção solar, principalmente pela sua capacidade de absorção de luz na região *uv* (280 a 315 nm), cuja radiação causa a formação de eritemas na pele.

Figura 29. Exemplos de cumarinas

Tipos de cumarinas

A- Existem as simples com uma hidroxila no C-7 e outra no C-6; e com menos freqüência em C-4, C-5 e C-8. Os glicosídeos mais comuns são os 7-O-glicosídeos.

- 1. Cumarinas condensadas
- 2.1 Diméricas (exemplo , composto 2, Fig. 29)
- 2.2. Furanocumarinas lineares e angulares

Furanocumarinas do tipo linear

Nome	R ₁	R ₂
Psoraleno	H	H
Xantotoxina	H	OMe
Bergapteno	OMe	H
Isopimpineline	OMe	OMe

Furanocumarinas do tipo angular

Nome	R ₁	R ₂
Angelicina	H	H
Isobergapteno	OMe	H
Pimpineline	OMe	OMe

B- Piranocumarinas

O anel pirano pode estar unido a posições distintas de cumarinas, por exemplo, (5,6) aloxantiletina, (6,7) xantiletina e (7,8) seselina (1-3 respectivamente, Figura 30).

Figura 30. Exemplos de piranocumarinas

C- Cumarinas complexas (serão mostradas em Tabela de Esqueletos de Cumarinas)

As cumarinas têm como precursor o ácido cinâmico, sendo, portanto, biossintetizadas através da rota do chiquimato. Os esquemas 10 e 11, de forma sucinta, sua biossíntese geral e alguns aspectos mecanísticos envolvidos nesta.

ROTA BIOSSINTÉTICA DE CUMARINAS SIMPLES (Oxigenadas na posição 7)

Esquema 10. Rota biossintética de cumarinas simples

POSSÍVEIS ROTAS DE FORMAÇÃO DO ANEL FURÂNICO EM FURANOCUMARINAS

Esquema 11. Possíveis rotas de formação do anel furânico em cumarinas

ROTA BIOSSINTÉTICA DO PRENIL

(baseada em uma das propostas de Sandmann, 2001)

Obs. : P = fosfato; PP = pirofosfato.

Esquema 12. Rota biossintética proposta para prenil, precursor do isopreno.

Mecanismo de prenilação para enóis e fenóis

O isopreno ativado, ou seja, dimetilalilpirofosfato, atua como eletrófilo sobre um de ambos os centros nucleofílicos de enóis e fenóis.

Esquema 13. Mecanismo de prenilação de enóis e fenóis (Gros, E. G. *et al.*)

4.5. Benzofuranos

ROTA BIOSSINTÉTICA

A rota biossintética, ou melhor, biogenética dos benzofuranos ainda não está totalmente elucidada. Está sendo aqui proposta, de forma simplificada, como a combinação de sistema fenólico derivado do acetato via um intermediário policetídeo — que origina a acetofenona — com uma unidade isopreno. A deidrotremetona (Figura 32 é um bom exemplo). Lin *et al.*, 1974, estudaram a biossíntese de isopentenilacetofenonas *in vivo* em *Eupatorium rugosum* utilizando os possíveis precursores de benzofuranos marcados com ^{14}C . A biossíntese da deidrodrotremetona em *Eupatorium rugosum* foi investigada alimentando precursores radioativos às plantas intactas. Os átomos do carbono de acetato-[1- ^{14}C] e de acetato-[2- ^{14}C] foram identificados na deidrotremetona pela degradação da molécula. A partir do padrão de marcação, concluiu-se que a parte da acetofenona

foi derivada do acetato através do caminho do poliacetato. Da incorporação do mevalonato, pareceu que o anel furânico e sua cadeia lateral foram formados de um composto isoprenóide . Siebertz *et al.*, 1989, usando culturas de raiz de *Eupatorium cannabinum* que acumulam os mesmos benzofuranos como os encontrados em raízes de plantas diferenciadas , pela primeira vez propuseram que tanto o anel aromático como o substituinte acetila dos benzofuranos são biogeneticamente derivados da fenilalanina.

Figura 31. Desidrotremetona

No tocante à formação do anel benzênico ou da acetofenona, mais uma vez a acondensação de Claisen ou aldólica é utilizada na ciclização das cadeias policetometilênicas, conforme mostra o Esquema 14 abaixo.

Esquema 14. Dois modos de ciclização de cadeias policetometilênicas.

4.6. Ocorrências de metabólitos secundários em tribos de Asteraceae

ABREVIações REFERENTES A OCORRÊNCIAS DE METABÓLITOS SECUNDÁRIOS E TRIBOS DE ASTERACEAE

A) *Metabólitos secundários*: MONO = monoterpenos; CUMA = cumarinas; BENZ = benzofuranos; POLI = poliacetilenos; DITE = diterpenos; TRIT = triterpenos; SELA = sesquiterpenóides = sesquiterpenos + lactonas sesquiterpênicas; FLAV = flavonóides; FLAV-A (FLAA) = flavonóides oxidados no anel A; FLAV-B (FLAB) = flavonóides oxidados no anel B.

B) *Ocorrências dos respectivos metabólitos secundários* :

OCMONO, OCCUMA (OCUMA), OCBENZ, OCPOLI, OCDITE, OCTRIT, OCSELA, OCFLAV .

Nota : Em tabelas captadas diretamente da tela do computador , as abreviações acima podem conter a letra B (= valores brutos) no final. Exemplos: OCCUMAB, OCFLAVB.

C) *Tribos*: BAR = Barnadesieae; MUT = Mutisieae; LAC = Lactuceae; VER = Vernonieae; LIA = Liabeae; ARC = Arctoteae; PLU = Plucheeae; GNA = Gnaphalieae; CAL = Calenduleae; AST = Astereae; ANT = Anthemideae; SEN= Senecioneae; HLN = Helenieae; HLT = Heliantheae; EUP = Eupatorieae.

Subjetivamente, os valores percentuais de ocorrências de cada classe química estudada nas diferentes tribos relacionadas à família Asteraceae foram tomados como: 0-0,5 (ocorrência insignificante); 0,5-5 (ocorrência muito baixa); 5-10 (ocorrência baixa); 10-20 (ocorrência razoável); 20-40 (ocorrência alta) e acima de 40 (ocorrência muito alta). Nas tabelas 10 e 11 são apresentados os números de ocorrências de 8 classes químicas por tribos e subtribos de Asteraceae, respectivamente.

Tabela 10. Números de ocorrências de 8 classes de metabólitos secundários em 17 tribos de Asteraceae

TRIBO	MONO	SELA(a)	DITE	TRIT	CUMA	POLI	FLAV	BENZ
ANT	3627	1243	15	23	230	141	852	4
ARC	1	24	21	35	1	23	0	0
AST	464	152	1419	69	179	378	454	58
BAR	7	0	0	12	0	0	139	0
CAL	0	20	93	0	3	24	1	7
CAR	24	527	6	27	24	1055	222	0
EUP	392	708	829	76	40	15	312	206
GNA	158	189	419	29	51	112	352	30
HLN	318	479	124	7	21	180	665	9
HLT	473	1636	1230	60	40	1591	861	197
INU	292	708	68	8	7	0	109	0
LAC	20	197	0	11	74	0	432	2
LIA	3	73	0	9	0	0	8	0
MUT	59	128	30	63	171	5	85	6
PLU	80	18	0	5	29	2	11	0
SEN	130	1162	71	40	25	5	18	108
VER	19	563	44	180	66	17	40	1
TOTAL	6067	7827	4369	618	961	3528	4561	628

(a) : ocorrências de sesquiterpenos + ocorrências de lactonas sesquiterpênicas.

Tabela 11. Números de ocorrências de 8 classes químicas em 87* subtribos de Asteraceae

Tribo	Subtribo	MONO	SELA	DITE	TRIT	CUMA	POLI	FLAV	BENZ
ANT	ACH	1134	157	5	9	32	97	413	0
ANT	ANT	34	47	0	0	7	23	252	0
ANT	ART	1996	693	5	3	158	0	137	4
ANT	CHR	149	36	0	0	6	0	8	0
ANT	GON	1	3	0	0	1	0	2	0
ANT	HAN	0	5	0	0	0	0	0	0
ANT	LEU	0	0	0	4	2	0	3	0
ANT	MAT	53	147	0	4	17	21	13	0
ANT	TAN	257	96	5	3	4	0	20	0
ANT	THA	1	34	0	0	0	0	2	0
ANT	URS	5	25	0	0	3	0	2	0
ARC	ARC	0	20	7	8	0	22	0	0
ARC	GRT	1	1	0	27	1	0	0	0
ARC	UNA1	0	3	14	0	0	1	0	0
AST	AST	462	70	697	57	158	287	361	44
AST	GRA	2	3	22	1	0	0	1	0
AST	SOL	0	79	700	11	21	91	92	14
BAR	UNA2	7	0	0	12	0	0	139	0
CAL	GRA	0	2	0	0	0	0	0	0
CAL	UNA3	0	18	93	0	3	24	1	7
CAR	CAD	6	174	0	14	4	407	68	0
CAR	CAR	3	2	5	0	0	92	39	0
CAR	CEN	15	351	1	12	18	313	115	0
CAR	ECH	0	0	0	1	2	243	0	0
EUP	ADE	0	1	33	0	0	1	0	0
EUP	AGE	71	113	183	13	5	2	30	21
EUP	ALO	2	41	109	7	2	0	70	13
EUP	AYA	1	10	13	0	0	1	0	29
EUP	CRI	5	38	114	7	2	0	4	48
EUP	DIS	0	23	46	2	0	0	2	4
EUP	EUP	130	136	68	11	16	2	154	27
EUP	FLE	0	5	15	2	0	0	0	0
EUP	GYP	3	108	49	18	3	0	2	21
EUP	HEB	2	20	11	1	0	0	0	1
EUP	HOF	0	0	5	0	0	0	0	8
EUP	LIA	8	49	15	1	7	5	1	23
EUP	MIK	70	105	111	8	3	4	2	0
EUP	OXY	93	56	43	6	1	0	14	11
EUP	PRA	7	3	14	0	1	0	33	0
GNA	ANG	0	49	94	0	6	4	15	0
GNA	CAS	15	90	16	1	11	46	109	0
GNA	GNA	99	40	214	9	28	59	224	15
GNA	LOR	19	4	0	18	6	3	4	3
GNA	REL	0	0	62	0	0	0	0	0
GNA	UNA4	25	6	30	1	0	0	0	12

Tabela 11. Continuação

Tribo	Subtribo	MONO	SELA	DITE	TRIT	CUMA	POLI	FLAV	BENZ
HLN	BAE	7	23	0	0	2	19	40	2
HLN	CHA	61	61	79	5	4	28	195	0
HLN	FLA	0	0	4	0	0	5	2	0
HLN	GAI	8	369	0	0	0	42	40	0
HLN	HYM	3	20	10	2	0	3	1	0
HLN	MAD	5	0	19	0	7	30	235	3
HLN	PEC	212	3	0	0	7	7	108	4
HLN	PER	7	3	0	0	0	0	44	0
HLN	UNA5	15	0	12	0	0	46	0	0
HLT	AMB	147	436	0	7	0	59	34	0
HLT	COR	13	37	0	1	4	829	190	2
HLT	ENG	0	4	0	5	0	6	0	0
HLT	GAL	0	0	0	4	0	47	0	0
HLT	HLT	37	262	514	1	20	72	501	25
HLT	MEL	99	478	99	7	4	107	57	15
HLT	RUD	10	51	2	0	1	142	10	0
HLT	UNA6	75	103	11	0	5	107	1	4
HLT	VER	68	179	597	34	6	116	68	150
HLT	ZIN	24	86	7	1	0	106	0	1
INU	UNA7	292	708	68	8	7	0	109	0
LAC	CRE	4	51	0	2	3	0	10	0
LAC	HIE	0	7	0	0	6	0	23	0
LAC	HYP	5	67	0	3	0	0	26	0
LAC	LAC	1	63	0	6	2	0	227	0
LAC	MIC	0	0	0	0	0	0	1	0
LAC	SCO	2	0	0	0	3	0	7	2
LAC	SEN	0	0	0	0	1	0	19	0
LAC	SON	3	7	0	0	51	0	56	0
LAC	STE	0	0	0	0	3	0	55	0
LAC	UNA8	5	2	0	0	5	0	8	0
LIA	LIA	3	73	0	9	0	0	8	0
MUT	MUT	58	119	30	44	107	5	80	6
MUT	NAS	1	9	0	19	64	0	5	0
PLU	UNA9	80	18	0	5	29	2	11	0
SEN	BLE	0	0	2	0	0	0	0	8
SEN	SEN	98	1039	55	37	9	1	11	40
SEN	TUS	32	123	14	3	16	4	7	60
VER	CEN	19	4	0	1	0	3	9	0
VER	ELE	0	33	0	2	0	3	0	0
VER	LYC	0	117	1	33	0	0	9	0
VER	PIP	0	8	38	2	0	0	0	0
VER	ROL	0	21	0	2	0	0	0	0
VER	VER	0	380	5	140	66	11	22	1

O uso de subtribos UNAi (i=1-9) foi devido a ocorrências de metabólitos secundários em gêneros / espécies não-determinados (*unassigned*) nas subtribos de Bremer, porém pertencentes seguramente às tribos classificadas pelo autor. De acordo com a classificação de Bremer (1994a), são 82 tribos contidas na família Asteraceae, sendo que as tribos BAR, CAL, INU, LIA, PLU não possuem subtribos. Neste caso, a estas foram dadas como subtribos respectivas UNA2, UNA3, UNA7, LIA, UNA9. Com relação às tribos ARC, GNA, HLN e HLT, foram incorporadas nestas as subtribos respectivas: UNA1, UNA4, UNA5 e UNA6.

Monoterpenos

Dentro da família, a tribo Anthemideae é a grande produtora / acumuladora de monoterpenos (ocorrência de 59,78%), seguida por Heliantheae, Astereae, Eupatorieae, Helenieae, Inuleae , que apresentam produção pequena (ocorrências 7,80; 7,65; 6,46; 5,24; 4,81 %, respectivamente). Gnaphalieae, Senecioneae, Plucheeae e Mutisieae têm produção muito pequena (ocorrências de 2,60; 2,14; 1,32; 0,97 %, respectivamente). Em Carduoideae, Lactuceae e Vernonieae, a produção é baixíssima (ocorrências de 0,40; 0,33; 0,31 %, respectivamente). Barnadesieae, Liabeae e Arctoteae produzem-nos em níveis insignificantes (0,12; 0,05; 0,02 %, respectivamente). Em Calenduleae, não se registrou sua ocorrência (Figura 32).

Figura 32. Ocorrências percentuais de monoterpenos em tribos de Asteraceae

Sesquiterpenóides (sesquiterpenos e lactonas sesquiterpênicas)

Com exceção de Barnadesieae, em todas as tribos registraram-se ocorrências de sesquiterpenóides, sendo mais pronunciadas em Heliantheae, Anthemideae e Senecioneae (ca. 21,0; 16,0 e 15,0 %, respectivamente); Eupatorieae e Inuleae apresentaram valores de ocorrências muito próximos (ca. 9,0 % em cada), o mesmo

ocorrendo com Vernonieae e Cardueae (ca. 7,0 % em cada). Helenieae apresentou valor interessante (ca. 6,0 %). Nas demais tribos, as ocorrências atingiram valores abaixo de 5 % por tribo (Figura 33).

Figura 33. Ocorrências percentuais de sesquiterpenóides em tribos de Asteraceae

Diterpenos

As tribos Astereae e Heliantheae lideram as ocorrências de diterpenos (ca.32,5 e 28,0 %, respectivamente) seguidas de Eupatorieae e Gnaphalieae (ca.19,0 e 10,0%, respectivamente). Nas demais tribos, as ocorrências dessa classe química foram abaixo de 3%, e deve-se salientar a ausência de diterpenos nas tribos Barnadesieae, Lactuceae, Liabeae e Plucheeae (Figura 34).

Triterpenos

Foram consideráveis as ocorrências de triterpenos nas tribos Vernonieae (ca. 28,0%), seguidas de Eupatorieae e Astereae (ca.11,0 % cada); Mutisieae, Heliantheae, Senecioneae e Arctoteae apresentaram ca. 10,0; 9,0; 6,0 e 5,0%, respectivamente. Nas demais tribos, as ocorrências foram abaixo de 5% por tribo, e deve-se salientar que em Calenduleae não se registrou nenhuma ocorrência da referida classe química (Figura 35).

Figura 34. Ocorrências percentuais de diterpenos em tribos de Asteraceae

Figura 35. Ocorrências percentuais de triterpenos em tribos de Asteraceae

Cumarinas

As tribos Anthemideae, Astereae e Mutisieae lideraram as ocorrências de cumarinas em Asteraceae (ca. 24,0; 19,0 e 18,0 %, respectivamente), seguidas de Lactuceae, Vernonieae e Gnaphalieae (ca.8,0; 7,0 e 5,0 %). Nas demais tribos, as ocorrências foram abaixo de 5% por tribo, e deve-se enfatizar que não se registrou nenhuma ocorrência dessa classe química em Barnadesieae e Liabeae (Figura 36).

Figura 36. Ocorrências percentuais de cumarinas em tribos de Asteraceae

Poliacetilenos

As tribos Heliantheae e Cardueae lideraram as ocorrências de poliacetilenos em Asteraceae (ca.45,0 e 29,0%, respectivamente), seguidas de Astereae (ca. 11,0 %) e Helenieae (ca. 5,0 %). Nas demais tribos, as ocorrências registradas foram abaixo de 5% por tribo, e deve-se salientar que não se registrou nenhuma ocorrência dessa classe química em Barnadesieae, Lactuceae, Liabeae e Inuleae (Figura 37).

Figura 37. Ocorrências percentuais de poliacetilenos em tribos de Asteraceae

Flavonóides

Foram registradas suas ocorrências em praticamente todas as tribos, exceto em Arctoteae. As tribos Heliantheae e Anthemideae apresentaram valores de ocorrências muito próximos (ca.19 %), seguidas de Helenieae, Astereae, Lactuceae, Gnaphalieae e Eupatorieae (ca. 15,0; 10,0; 9,5 e 7,0%, respectivamente) Nas demais tribos, as ocorrências dessa classe química foram inferiores a 5% por tribo (Figura 38).

Após estudos detalhados de quimiotaxonomia de plantas, usando flavonóides como caracteres micromoleculares naquelas, Harborne propôs uma tabela baseada principalmente nas idéias que tinha na época (início da década de 70) sobre o *status* evolucionário e em considerações biogenéticas no reino vegetal. Resumindo suas idéias e considerações, temos abaixo (Gibbs, 1974):

São caracteres primitivos: 3-Desoxiantocianidinas; Flavonóis; Chalconas; Flavanonas; Leucoantocianidinas; Diidroxichalconas; Substituição-C.

Ex.: Antocianidinas isoladas de musgos e samambaias.

São caracteres avançados: A- Resultantes de mutações por ganho incluem: O-glicosilação complexa; Hidroxilação em C-6 ou C-8; O-metilação; Oxidação de

chalconas a auronas; **B-** Resultantes de mutações por perda são: Substituição de flavonóis por flavonas; Eliminação de leucoantocianidinas; Eliminação de triidroxilação.

Figura 38. Ocorrências percentuais de flavonóides em tribos de Asteraceae

Benzofuranos

As tribos Eupatorieae, Heliantheae e Senecioneae englobaram ca. 81% das ocorrências de benzofuranos em Asteraceae (ca. 33,0; 31,0 e 17,0%, respectivamente). Astereae e Gnaphalieae apresentaram também valores de ocorrências, de certa forma, relevantes (ca. 9,0 e 5%, respectivamente). Nas demais tribos, as ocorrências foram abaixo de 1,5%, e deve-se enfatizar que não se registrou neste trabalho nenhuma ocorrência dessa classe química em Barnadesieae, Cardueae, Liabeae, Inuleae e Plucheeae (Figura 39).

Figura 39. Ocorrências percentuais de benzofuranos em tribos de Asteraceae

Resumo de ocorrências de cumarinas, flavonóides e benzofuranos em tribos de Asteraceae

CUMARINAS

As tribos que mais acumulam cumarinas são Anthemideae, Astereae e Mutisieae. Esta apresenta uma variedade de esqueletos cumarínicos muito interessante, dos mais simples (mantêm o anel benzopirônico original ou a este diretamente ligados oxigênios de funções álcool, éter ou éster) aos mais complexos (grupos metila, etila, prenila, terpenóides, farnesila, geranila, etc. podem estar diretamente ligados ao anel benzopirônico).

BENZOFURANOS

Quanto aos benzofuranos, as tribos com alta produção (número de ocorrências entre 100 e 210) desses metabólitos são, em ordem decrescente: Eupatorieae, Heliantheae e Senecioneae; Astereae e Gnaphalieae apresentaram produção razoável, 58 e 30 ocorrências, respectivamente. É impotante salientar que as tribos Arctoteae, Barnadesieae, Cardueae, Inuleae, Liabeae e Plucheeae não

apresentaram ocorrências de tais metabólitos; as demais apresentaram ocorrências muito baixas (1 a 7).

FLAVONÓIDES

Com relação aos flavonóides, cujo número total de ocorrências na família foi de 4561, já eram esperadas suas ocorrências ao longo de praticamente todas as tribos de Asteraceae. As tribos com alta produção (números de ocorrências entre 400 e 900) de tais tipos de metabólitos são, em ordem decrescente: Heliantheae, Anthemideae, Helianeae, Astereae e Lactuceae; apresentaram produções razoáveis (números de ocorrências entre 100 e 400), em ordem decrescente, as tribos Eupatorieae, Cardueae, Barnadesieae e Inuleae; baixas produções (números de ocorrências entre 40 e 100), as tribos Mutisieae (85) e Vernonieae (40). As demais apresentaram produções pouco expressivas (entre 0 e 40), devendo-se salientar aqui que não se registrou, até o presente, nenhuma ocorrência de flavonóide na tribo Arctoteae.

Obs.: Termos usados subjetiva e independentemente para produção relativa às classes de metabólitos distintas: muito baixa ou pouco expressiva; baixa; razoável; alta; muito alta.

4.7. REGRESSÃO LINEAR MÚLTIPLA

A partir do banco de dados de ocorrências das classes químicas estudadas, inúmeros resultados de correlações entre as classes químicas citadas neste trabalho foram obtidos através de estatística multivariada, entretanto somente os mais importantes, ou seja, com valores estatísticos relevantes, estão sendo apresentados nesta dissertação.

OCORRÊNCIAS E OXIDAÇÕES DE DIVERSAS CLASSES QUÍMICAS TRATADAS
ESTATISTICAMENTE, EM NÍVEL DE TRIBOS E SUBTRIBOS DE ASTERACEAE

OCORRÊNCIAS

Matrizes de correlações de ocorrências de 8 classes de
metabólitos secundários em tribos e subtribos de Asteraceae

TRIBOS

Correlações (octribu.sta)
Correlações marcadas são significantes a um
nível-p < 0,05
N=17

	OCMONO	OCSELA	OCDI TE	OCTRIT	OCCUMA	OCPOLI	OCFLAV	OCBENZ
OCMONO	1							
OCSELA	0,501709	1						
OCDI TE	0,031659	0,312903	1					
OCTRIT	-0,00119	0,225543	0,295191	1				
OCCUMA	0,571182	0,189931	0,251188	0,257163	1			
OCPOLI	0,046271	0,542234	0,522471	0,086959	0,001689	1		
OCFLAV	0,523114	0,590361	0,488506	-0,03686	0,468398	0,53317	1	
OCBENZ	0,014189	0,512544	0,719789	0,277308	-0,03153	0,444991	0,358393	1

SUBTRIBOS

N=89

	OCMONO	OCSELA	OCDI TE	OCTRIT	OCUMA	OCPOLI	OCFLAV	OCBENZ
OCMONO	1							
OCSELA	0,436231	1						
OCDI TE	0,069226	0,105414	1					
OCTRIT	0,046119	0,332153	0,24787	1				
OCUMA	0,535938	0,299946	0,333321	0,459513	1			
OCPOLI	0,031209	0,07431	0,178539	0,086397	0,115489	1		
OCFLAV	0,379446	0,181213	0,450188	0,108744	0,39111	0,318221	1	
OCBENZ	0,033202	0,196068	0,504264	0,236757	0,128585	0,073169	0,135989	1

Tabela 12. Matrizes de correlações entre ocorrências de 8 classes de
metabólitos secundários em tribos e subtribos de Asteraceae

Comentário:

As matrizes de correlações acima mostram valores de coeficientes de
correlação de Pearson (r) marcados em vermelho, considerados significativos para
a montagem de equações empíricas para previsão de ocorrências dos referidos
metabólitos em tribos e subtribos de Asteraceae.

Em nível de tribos, as ocorrências de cumarinas e flavonóides estão mais
fortemente (r > 0,60) correlacionadas com as dos monoterpenos. As correlações
entre as ocorrências de flavonóides e as de sesquiterpenos / lactonas

sesquiterpênicas, entre as de flavonóides e as de diterpenos , entre as de flavonóides e as de poliacetilenos ainda apresentam valores de r razoáveis e significativos (r = 0,50-0,60). As correlações entre as ocorrências de benzofuranos e as de monoterpenos, entre as de benzofuranos e as de diterpenos apresentam valores de r acima do razoável, r = 0,61 e r = 0,72, respectivamente.

Com relação às subtribos, a correlação entre ocorrências de cumarinas e as de monoterpenos apresentou coeficiente r = 0,63, muito próximo ao valor obtido na correlação entre as ocorrências destas classes químicas em tribos (r = 0,67). O valor de r nas correlações entre ocorrências de cumarinas e triterpenos foi 0,47; entre as de cumarinas e as de diterpenos, r = 0,33 ; entre as de cumarinas e as de sesquiterpenos / lactonas sesquiterpênicas, r = 0,30. Apesar de os valores de r estarem numa faixa abaixo do razoável, são ainda significativos. São significativas as correlações entre ocorrências de: flavonóides e monoterpenos, 0,38; flavonóides e diterpenos, 0,46; flavonóides e cumarinas, 0,39; flavonóides e poliacetilenos, 0,32, embora os valores de r estejam abaixo do razoável.

Ocorrências de cumarinas

Matriz de coeficientes de correlação de ocorrências de cumarinas em função de ocorrências de diversas classes de metabólitos secundários em tribos de Asteraceae

Variável dependente	OCMONO	OCSELA	OCDITE	OCTRIT	OCCUMA	OCPOLI	OCFLAV	OCBENZ
OCCUMA (r)	0,67	0,19	0,25	0,27	1,00	0,00	0,47	-0,03

Tabela 13. Matriz de correlações entre ocorrências de cumarinas (OCCUMA), variável dependente, e ocorrências de diversas classes químicas em tribos de Asteraceae.

Comentário: Com as ocorrências dos monoterpenos a correlação é relativamente forte, r = 0,67.

Tabela 14. Sumário de regressão linear múltipla padrão para a variável dependente OCCUMA em tribos de Asteraceae.

Sumário de regressão para a variável dependente OCCUMA						
R= 0,832959 R² = 0,69604970 R² ajustado = 0,45964392						
F(7,9) = 2,944 p< 0,06718 Err. padr. da estimativa : 51,335						
	BETA	Err. padr. de BETA	B	Err. padr. de B	t(12)	nível-p
Intercepto			20,23053	22,15217	0,913253	0,384921
OCMONO	0,732961	0,329587	0,059548	0,026777	2,223874	0,053226
OCSELA	-0,22743	0,433286	-0,03211	0,061167	-0,52489	0,612345
OCDITE	0,386907	0,388319	0,059463	0,05968	0,996364	0,345106
OCTRIT	0,337487	0,214504	0,537578	0,34168	1,573339	0,150091
OCPOLI	-0,07357	0,303949	-0,01165	0,048121	-0,24203	0,814181
OCFLAV	0,109319	0,349223	0,025906	0,082756	0,313035	0,761387
OCBENZ	-0,28114	0,413925	-0,28839	0,424591	-0,67921	0,51409

Comentário: Os valores de coeficientes de regressão BETA obtidos na Tabela 14, relacionados às ocorrências das demais classes químicas tomadas como variáveis independentes em tribos de Asteraceae, não são significantes.

Equação 1: $OC_{CUMA} = 20,23 + 0,73 OC_{MONO} + 0,39 OC_{DITE} - 0,23 OC_{SELA} + 0,34 OC_{TRIT} - 0,07 OC_{POLI} + 0,11 OC_{FLAV} - 0,28.O_{CBENZ}$

Comentário : Uma avaliação dos níneis de probabilidade (níveis-p) — exceto o dos monoterpenos, $p \sim 0,05$ — relacionados aos valores de BETA apontam a Equação 2 como de pouca utilidade na previsão de ocorrências de cumarinas.

Tabela 15. Valores observados *versus* previstos de ocorrências de cumarinas em tribos de Asteraceae. (Valores previstos obtidos a partir da regressão linear múltipla padrão tomando ocorrências de cumarinas como variável dependente em função de ocorrências de diversas classes químicas).

TRIBO	Previsto (X)	Observado (Y)
ANT	228,8367	230
ARC	39,3156	1
AST	155,0838	179
BAR	30,69917	0
CAL	22,84813	3
CAR	13,0749	24
EUP	59,4832	40
ONA	63,23868	51
HLN	47,45987	21
HLT	48,22798	40
INU	26,06541	7
LAC	31,6244	74
LJA	23,11088	0
MUT	55,69901	171
PLU	27,38805	29
SEN	-14,3481	25
VER	103,2165	66

Comentário: Os melhores valores previstos para ocorrências de cumarinas são para as tribos Anthemideae, Plucheeae, Astereae e Heliantheae (-0,5%, -5,63%, -13,4% e +20,6%, com relação aos observados, respectivamente).

A Figura 40 apresenta o diagrama de dispersão de ocorrências de cumarinas (OCCUMA) em função de ocorrências de diversas classes de metabólitos secundários (MONO, DITE, TRIT, SELA, POLI, FLAV, BENZ) em tribos de Asteraceae, a partir da regressão linear múltipla.

Figura 40. Diagrama de dispersão de valores previstos vs. observados para ocorrências de cumarinas em tribos de Asteraceae. (Valores previstos obtidos a partir da regressão linear múltipla padrão tomando ocorrências de cumarinas como variável dependente em função de ocorrências de diversas classes químicas).

Comentário: A tribo Mutisieae é facilmente identificada como um possível *outlier*. As tribos Plucheeae e Anthemideae se ajustam muito bem na reta da regressão (Figura 40), uma vez que apresentam residuais baixos (observado-previsto, valores na Tab. 15).

Resultados de regressão linear múltipla passo-a-passo adiante de ocorrências de cumarinas em função de ocorrências de diversas classes químicas em tribos de Asteraceae

Passo : 0

Variável dependente: Ocorrências de Cumarinas (OCCUMA)

R múltiplo = 0,00000000 F =
 $R^2 = 0,00000000$ df =
 Nº de casos: 17 R^2 ajustado p =
 Erro padrão da estimativa =
 Intercepto = Erro padrão t () p <
 Variável independente Valor de B

Passo : 1

Variável dependente: Ocorrências de Cumarinas (OCCUMA)

R múltiplo = 0,67 F = 12,30
 $R^2 = 0,45$ df = (1, 15)
 Nº de casos: 17 R^2 ajustado = 0,41 p = 0,00318
 Erro padrão da estimativa =
 Intercepto = 37,07 Erro padrão = 14,11 t (15) = 2,6280 p < 0,0190
 Variável independente (Valor de B)
 OCMONO (0,671)

Passo : 2

Variável dependente: Ocorrências de Cumarinas (OCCUMA)

R múltiplo = 0,74 F = 8,41
 $R^2 = 0,55$ df = (2, 14)
 Nº de casos : 17 R^2 ajustado = 0,48 p = 0,003985
 Erro padrão da estimativa = 50,31
 Intercepto = 17,56 Erro padrão = 17,48 t (14) = 1,0045 p < 0,3322
 Variável independente (Valor de B)
 OCMONO (0,690) OCTRIT (0,309)

Passo : 3

Variável dependente: Ocorrências de Cumarinas (OCCUMA)

R múltiplo = 0,79 F = 7,09
 $R^2 = 0,62$ df = (3, 13)
 Nº de casos : 17 R^2 ajustado = 0,53 p = 0,00456
 Erro padrão da estimativa = 47,71
 Intercepto = 28,82 Erro padrão = 18,02 t (13) = 1,60 p < 0,1328
 Variável independente (Valor de B)
 OCMONO (0,862) OCTRIT (0,395) OCSELA (-0,33)

Passo : 4 (Sol. final, nenhum outro F de entrada excede o limite especificado)

Variável dependente: Ocorrências de Cumarinas (OCCUMA)

R múltiplo = 0,82 F = 6,2
 $R^2 = 0,67$ df = (4, 12)
 Nº de casos : 17 R^2 ajustado = 0,57 p = 0,0061
 Erro padrão da estimativa =
 Intercepto = 26,71 Erro padrão = 17,47 t (12) = 1,53 p < 0,1522
 Variável independente (Valor de B):
 OCMONO (0,891) OCTRIT (0,341) OCSELA (-0,41) OCDITE (0,252)

Obs.: Valores de BETA em vermelho: mais significativos

Tabela 16. Resultados de regressão linear múltipla passo a passo adiante de ocorrências de cumarinas em função de ocorrências de diversas classes químicas em tribos de Asteraceae

Equação 2. $OC_{CUMA} = 26,7 + 0,89OC_{MONO} + 0,25OC_{DITE} - 0,41OC_{SELA} + 0,34OC_{TRIT}$

Comentário: Equação obtida a partir da regressão linear múltipla utilizando 4 passos adiante e como solução final o passo 4 (dados na Tabela 16) e considerando $BETA_{MONO} = 0,891121$ como valor relevante.

Passo 4: F para entrar: 1,95; Variável dependente : OCCUMA; Tolerância mínima: 0,6289; R múltiplo: 0,8209; Variáveis (E) ntradas / (R) emovidas; 1 (E) OCMONO 2 (E) OCTRIT 3 (E) OCSELA 4 (E) OCDITE

Obs. : Nenhum outro F para entrar excede o limite especificado.

Tomando os valores contidos na Tabela 15 como passo zero para iniciar a regressão[F para remover: 2,94; R: 0,6712; Variável dependente: OCCUMA; Variáveis (R) emovidas / (E) ntradas: 1 (R) OCPOLI; 2 (R) OCFLAV; 3 (R) OCBENZ; 4 (R) OCDITE; 5 (R) OCSELA; 6 (R) TRIT] para a regressão passo a passo para trás, a solução final é a matriz do passo 6 (Tabela 17).

Regressão linear múltipla passo a passo para trás, tomando como variável dependente OCCUMA em função de NOs de diversas classes químicas em tribos de Asteraceae

$R=0,83429593$
 $R^2=0,69604970$
 $R^2_{ajustado}=0,45964392$
 $F(7,9)=2,9443$
 $p<0,06718$
Err. padr. da estimativa: 51,335

Passo 0

	BETA	Err. padr. de BETA	B	Err. padr. de B	t(9)	nível-p
Intercepto			20,23053	22,16217	0,913253	0,384921
OCMONO	0,732961	0,329587	0,059548	0,026777	2,223874	0,053226
OCSELA	-0,22743	0,433286	-0,03211	0,061167	-0,52489	0,612345
OCDIITE	0,366907	0,368319	0,069463	0,06968	0,986364	0,345106
OCTRIT	0,337487	0,214504	0,537578	0,34168	1,573339	0,190091
OCPOU	-0,07357	0,303949	-0,01165	0,048121	-0,24203	0,814181
OCFLAV	0,109319	0,349223	0,025906	0,082756	0,313035	0,761387
OCBENZ	-0,28114	0,413925	-0,28839	0,424591	-0,67921	0,51409

$R=0,83310945$
 $R^2=0,69407136$
 $R^2_{ajustado}=0,51051417$
 $F(6,10)=3,7812$
 $p<0,03134$
Err. padr. da estimativa: 48,859

Passo 1

	BETA	Err. padr. de BETA	B	Err. padr. de B	t(10)	nível-p
Intercepto			21,53687	20,44827	1,053237	0,317016
OCMONO	0,776763	0,262168	0,063107	0,021299	2,962846	0,014219
OCSELA	-0,28788	0,33752	-0,04061	0,047648	-0,85234	0,413974
OCDIITE	0,348113	0,366635	0,053501	0,051737	1,034096	0,325451
OCTRIT	0,347442	0,200368	0,553435	0,319164	1,734016	0,113577
OCFLAV	0,084164	0,317318	0,019945	0,075196	0,265237	0,796214
OCBENZ	-0,24339	0,364909	-0,24966	0,374312	-0,66687	0,519881

$R=0,83181676$
 $R^2=0,69191913$
 $R^2_{ajustado}=0,55188237$
 $F(5,11)=4,9410$
 $p<0,01291$
Err. padr. da estimativa: 46,749

Passo 2

	BETA	Err. padr. de BETA	B	Err. padr. de B	t(11)	nível-p
Intercepto			23,55094	18,16608	1,296424	0,221367
OCMONO	0,807868	0,22435	0,066634	0,018227	3,600929	0,004163
OCSELA	-0,24991	0,292786	-0,03528	0,041333	-0,85365	0,411557
OCDIITE	0,400608	0,260548	0,061589	0,040043	1,537563	0,152408
OCTRIT	0,329497	0,180455	0,524851	0,287444	1,825924	0,095103
OCBENZ	-0,26961	0,336084	-0,27666	0,344745	-0,80222	0,439414

Tabela 17. Sumário de regressão linear múltipla passo a passo para trás (passo 0 - passo 6) para a variável dependente OCCUMA , em tribos de Asteraceae.

R = 0,82091097 R² = 0,67389483 R²ajustado = 0,56519310

F(4,12) = 6,1995 p< 0,00607 Err. padr. da estimativa: 46,050

Err. padr.

Err. padr.

	BETA	de BETA	B	de B	t(12)	nível-p
Intercepto			26,71132	17,4684	1,529122	0,152157
OCMONO	0,891121	0,195928	0,072398	0,015918	4,548202	0,000668
OCSELA	-0,41272	0,207874	-0,05826	0,029346	-1,98542	0,070432
OCDITE	0,251589	0,179964	0,038666	0,027658	1,397993	0,18742
OCTRIT	0,340552	0,177236	0,542461	0,282316	1,921465	0,07874

Passo 3

R = 0,78789819 R² = 0,62078356 R²ajustado = 0,53327207

F(3,13)=7.0937 p<.00456 Std.Error de estimate: 47.710

Err. padr.

Err. padr.

	BETA	de BETA	B	de B	t(13)	nível-p
Intercepto			28,90285	18,02524	1,603466	0,13284
OCMONO	0,861594	0,20181	0,069999	0,016396	4,269336	0,000914
OCSELA	-0,33139	0,206764	-0,04678	0,029189	-1,60275	0,132999
OCTRIT	0,394661	0,179194	0,62865	0,285436	2,202421	0,046291

Passo 4

R = 0,73881672 R² = 0,54585015 R²ajustado=0,48097160

F(2,14)= 8,4134 p<0,00398 Err. padr. da estimativa: 50,312

Err. padr.

Err. padr.

	BETA	de BETA	B	de B	t(14)	nível-p
Intercepto			17,56067	17,48217	1,00449	0,332189
OCMONO	0,690114	0,180447	0,056067	0,01466	3,824465	0,001859
OCTRIT	0,309392	0,180447	0,492826	0,287432	1,714584	0,108462

Passo 5

R = 0,67118195 R² = 0,45048521 R²ajustado = 0,41366089

F(1,15)= 12,297 p< 0,00318 Err. da estimativa: 53,466

Err. padr.

Err. padr.

	BETA	de BETA	B	de B	t(15)	nível-p
Intercepto			37,0689	14,1051	2,628049	0,019006
OCMONO	0,671182	0,191401	0,054529	0,01555	3,506681	0,00318

Passo 6

Obs. :Nenhum outro F para remover é menor do que o limite especificado.

Tabela 17. Continuação

Comentário: O valor do coeficiente de regressão BETA_{MONO} é significativo em todos os passos da regressão (exceto no passo 0, equivalente ao passo 7 da RLM padrão).

Equação 3: $OC_{CUMA} = 37,07 + 0,67 OC_{MONO}$ (gerada a partir da matriz do passo 6 na Tabela 17)

NOTA: Os mesmos tipos de regressão linear múltipla foram feitos para todas as classes químicas envolvidas neste trabalho, tomando cada uma como variável dependente, entretanto só serão apresentadas as principais equações empíricas pertinentes a cumarinas, benzofuranos e flavonóides.

Tabela 18. Matriz de coeficientes de correlação r entre ocorrências de cumarinas e diversas classes químicas em subtribos de Asteraceae

Matriz de coeficientes de correlação de ocorrências de cumarinas em função de ocorrências de diversas classes químicas em subtribos de Asteraceae							
Correlações (ocsu90c.sta)							
Correlações marcadas são significantes a um nível-p < 0,05							
N=89							
Variável							
dependente	OCMONO	OCSELA	OCDITE	OCTRIT	OCPOLI	OCFLAV	OCBENZ
OCCUMA (r)	0,60575	0,28475	0,32722	0,5447	0,12639	0,38598	0,25447

Comentário: Em nível de subtribos, tomando as cumarinas como variável dependente (Tabela 18), a correlação é mais fortemente influenciada pelas ocorrências dos monoterpenos (r = 0,60575). Aqui vale a pena observar que todas as classes químicas influenciam positivamente as ocorrências de cumarinas em grau menor (POLI, r = 0,12639) a médio (TRIT, r = 0,5447). Somente o valor r de OCPOLI não é significativo.

Tabela 19. Sumário de regressão linear múltipla padrão de NOs de cumarinas (OCCUMA) como variável dependente em função de NOs de diversas classes químicas em subtribos de Asteraceae.

Sumário de regressão para variável dependente: OCCUMA, em subtribos de Asteraceae						
R = 0,83069542 R ² = 0,69005488 R ² ajustado = 0,66326950						
F(7,81) = 25,762 p< 0,00000 Erro padr. da estimativa : 16,974						
	BETA	Err.padr. de BETA	B	Err.padr. de B	t(81)	nível-p
Intercepto			0,4653113	2,293866	0,20285	0,839761
OCMONO	0,633009	0,074769	0,0747201	0,008826	8,466197	9,03E-13
OCSELA	-0,16221	0,072365	-0,027847	0,012423	-2,24159	0,027724
OCDITE	0,201129	0,084711	0,044156	0,018597	2,3743	0,019948
OCTRIT	0,541563	0,070202	0,7882219	0,102176	7,71433	2,74E-11
OCPOLI	0,038882	0,065797	0,0103073	0,017442	0,590942	0,556205
OCFLAV	0,016657	0,080275	0,0051918	0,025021	0,207501	0,836139
OCBENZ	-0,07602	0,080497	-0,103042	0,109112	-0,94436	0,347792

Equação 4: $OC_{CUMA} = 0,4653 + 0,63 OC_{MONO} + 0,20 OC_{DITE} - 0,16 OC_{SELA} + 0,54 OC_{TRIT} + 0,039 OC_{POLI} + 0,017 OC_{FLAV} - 0,076.OC_{BENZ}$

Comentário: A equação gerada a partir dos valores na Tabela 19 mostra-nos os coeficientes de regressão BETA (em vermelho) que muito influenciam positivamente a previsão de ocorrências de cumarinas em subtribos de Asteraceae: BETA_{MONO} = 0,63 e BETA_{DITE} = 0,20; BETA_{TRIT}= 0,54. Os demais coeficientes influenciam em pouco a previsão.

Tabela 20. Valores previstos vs. valores observados de ocorrências de cumarinas em subtribos de Asteraceae, a partir da regressão linear múltipla.

Valores previstos versus observados para ocorrências de cumarinas em subtribos de Asteraceae

Tribo	Subtribo	Previsto	Observado	Tribo	Subtribo	Previsto	Observado
ANT	ACH	88,139	32	HLN	BAE	0,737	2
ANT	ANT	2,587	7	HLN	CAI	1,141	1
ANT	ART	134,649	158	HLN	CHA	12,322	4
ANT	CHR	11,345	6	HLN	FLA	1,369	0
ANT	GON	1,147	1	HLN	GAI	-6,665	0
ANT	HAN	1,027	0	HLN	HYM	2,899	0
ANT	LEU	3,984	2	HLN	MAD	1,997	7
ANT	MAT	4,559	17	HLN	PEC	15,96	7
ANT	TAN	20,385	4	HLN	PER	1,591	0
ANT	THA	0,441	0	HLN	UNA5	2,935	0
ANT	URS	0,942	3	HLT	AMB	7,067	0
ARC	ARC	6,77	0	HLT	COR	1,573	4
ARC	GOR	20,378	1	HLT	ENG	4,603	0
ARC	UNA1	1,871	0	HLT	GAL	3,984	0
AST	AST	105,199	158	HLT	HLT	22,948	20
AST	GRA	3,186	0	HLT	MEL	5,216	4
AST	SOL	44,276	21	HLT	RUD	0,834	1
BAR	UNA2	10,186	0	HLT	UNA6	4,176	5
CAL	GRA	1,096	0	HLT	VER	30,412	6
CAL	UNA3	4,639	3	HLT	ZIN	1,87	0
CAR	CAD	7,573	4	INU	UNA7	16,19	7
CAR	CAR	1,603	0	LAC	CRE	1,698	3
CAR	CEN	2,846	18	LAC	HIE	0,982	6
CAR	ECH	1,852	2	LAC	HYP	2,118	0
EUP	ADE	3	0	LAC	LAC	4,045	2
EUP	AGE	19,309	5	LAC	MIC	1,141	0
EUP	ALO	8,954	2	LAC	SCO	0,891	3
EUP	AYA	-4,048	0	LAC	SEN	1,141	1
EUP	CRI	2,558	2	LAC	SON	1,204	51
EUP	DIS	3,865	0	LAC	STE	1,141	3
EUP	EUP	13,978	16	LAC	UNA8	1,466	5
EUP	FLE	3,304	0	LIA	LIA	6,097	0
EUP	GYP	10,306	3	MUT	MUT	34,504	107
EUP	HEB	1,972	0	MUT	NAS	14,511	64
EUP	HOF	-0,167	0	PLU	UNA9	10,203	29
EUP	LIA	-2,398	7	SEN	BLE	-0,338	0
EUP	MIK	15,942	3	SEN	SEN	6,203	9
EUP	OXY	11,27	1	SEN	TUS	-8,313	16
EUP	PRA	2,389	1	VER	CEN	3,166	0
GNA	ANG	5,385	6	VER	ELE	1,811	0
GNA	CAS	1,825	11	VER	LYC	21,984	0
GNA	GNA	23,162	28	VER	PIP	4,547	0
GNA	LOR	14,649	6	VER	ROL	2,084	0
GNA	REL	4,676	0	VER	VER	92,059	66
GNA	UNA4	2,885	0				

Comentário: Na Tabela 20, alguns valores previstos considerados bons (margem de erro de $\pm 20\%$ do valor observado) estão marcados em vermelho; em azul, estão exemplificados valores previstos que fogem bastante dos observados.

Figura 41. Diagrama de dispersão de valores previstos vs. valores observados para ocorrências de cumarinas em subtribos de Asteraceae. (Valores previstos obtidos a partir da regressão linear múltipla, tomando NOs de cumarinas como variável dependente em função de NOs de diversas classes químicas)

Comentários: Alguns pontos estão identificados com a tribo e a subtribo respectiva abreviadas. Formação de aglomerações em torno do início da reta da regressão. Ao longo dessa reta, surgem muitos prováveis *outliers*.

NÚMEROS ERRÁTICOS (OUTLIERS)

“Os **outliers** são elementos com comportamento muito diferente dos demais. Os resíduos são grandes, em módulo. É **extremamente importante** controlar os *outliers* porque, em virtude da forma de estimação da equação, geralmente por mínimos quadrados, um erro grande modifica significativamente os somatórios, alterando os coeficientes da equação” (M.A.S. González, <http://www.unisinos.br/>, 1998-2000). Assim, um único valor muito errático pode modificar a equação. Mas eles são importantes, uma vez que podem trazer valiosas informações ou

modificações de sistemas ou projetos.. Por outro lado, os *outliers* fornecem freqüentemente a introspecção valiosa em observações particulares. Saber porque uma observação é um *outlier* é muito importante. Por exemplo, a identificação do outlier é uma parte-chave do controle de qualidade. Os *outliers* fornecem estudos de casos interessantes. Devem sempre ser identificados e discutidos; nunca devem ser ignorados, ou "varridos sob o tapete." Em toda a pesquisa científica, a divulgação plena é necessária, incluindo uma divulgação e uma discussão sobre os *outliers* (***Outliers: what to do about them?***—

http://www2.tltc.ttu.edu/westfall/images/5349/outliers_what_to_do.htm, 2003).

No caso das ocorrências de cumarinas em subtribos, o diagrama de dispersão de valores previstos *versus* observados (Figura 41) mostra a presença de vários pontos muito distantes da reta de regressão, ou seja, prováveis *outliers*. A remoção de alguns destes para um melhor ajuste dos demais pontos ao longo da reta deve ser criteriosa e a partir de estudo de resíduos ou residuais.

Tabela 21. Valores residuais padrão de NOs previstos versus observados para cumarinas em subtribos de Asteraceae (Valores previstos obtidos a partir de RLM de NOs de cumarinas como variável dependente em função de NOs de diversas classes químicas em tribos de Asteraceae).

Tribo-Subtribo	Caso	Resíduo padrão OCCUMA (occu60c.sta)									
		Valor		Ordenado							
		observado	previsto	Resíduo padrão	Vel. Prev. padrão	Resíduo padrão	Err. padr.	Distância	Resíduo	Distância	
						dox. prev.	Milheins	deLaboz		deCook	
MUT-NUT	78.	107	37,34039	69,65961	1,046566	4,103905	4,20379542	4,40578	74,21143	0,146554	
ANT-ACH	2 *	32	91,29477	-59,2948	3,266612	-3,49268	10,2222233	30,52704	-93,0219	1,361599	
LAC-SON	74.	51	0,785284	50,21471	-0,45783	2,993334	2,33953762	0,663001	51,18713	0,021995	
AST-AST	16.	158	109,0407	49,95929	3,997346	2,894372	10,7506819	34,31225	81,75522	1,163263	
MUT-NAS	79.	64	15,29158	49,70842	0,139162	2,668992	2,74258995	1,308637	50,01413	0,028332	
VER-VER	89.	66	100,5799	-34,5799	3,649148	-2,03723	13,4695751	54,42649	-93,3993	2,392727	
HUT-VER	64.	6	39,81493	-33,8148	1,1494	-1,95216	12,5665539	47,25935	-74,8591	1,332979	
ANT-ART	4.	158	133,1934	24,80659	4,991334	1,461448	14,7773228	65,70829	102,4732	3,452921	
VER-LYC	86.	0	23,30944	-23,3094	0,469132	-1,37325	3,60766516	2,789199	-24,3493	0,010985	
ARC-GR	14.	1	21,79418	-20,7942	0,406772	-1,22506	3,35359993	2,446316	-21,6398	0,00793	
Hipóclito	1.	100	79,63918	20,36082	2,707222	1,99908	11,4977777	37,30948	36,05539	0,025499	
SEN-TUS	83.	16	-3,68097	19,69097	-0,64205	1,160062	6,37468174	11,42374	22,92439	0,03216	
FLU-UNA	80.	29	9,990513	19,03949	-0,09023	1,121666	2,25770907	0,9981	19,39239	0,002894	
AST-SOL	18.	21	37,81608	-16,8161	1,066225	-0,99081	11,4423361	39,0006	-30,6263	0,187347	
ANT-TAN	10.	4	19,69497	-15,6944	0,319945	-0,92402	2,70159219	1,240449	-16,0322	0,002846	
EUP-AGE	27.	5	18,96398	-13,9637	0,280285	-0,82265	3,01891811	1,794353	-14,4197	0,002853	
ANT-ANT	9.	17	3,769933	13,23117	-0,33605	0,779497	2,24686536	0,556681	13,46753	0,001381	
CAR-CEN	24.	18	5,137953	12,86205	-0,7287	0,757751	6,04861403	10,1857	14,73286	0,011998	
BRA-UM2	19.	0	11,16867	-11,1687	-0,03051	-0,65799	3,67171454	3,128919	-11,7189	0,002787	
EUP-MK	38.	3	14,03051	-11,0305	0,087264	-0,64995	3,10234694	1,950995	-11,4117	0,001887	
EUP-OKY	39.	1	11,42213	-10,4221	-0,02008	-0,61401	2,15519919	0,42993	-10,5929	0,000785	
GNAL-OR	44.	6	15,70416	-9,70416	0,159142	-0,57171	2,64251828	1,144041	-9,9452	0,00104	
HUN-FEC	53.	7	16,44313	-9,44313	0,166554	-0,55633	3,07462086	1,699578	-9,76348	0,001357	
GNAC-AS	42.	11	1,614661	9,38534	-0,4237	0,952925	2,70773005	1,250606	9,630407	0,001024	
EUP-GYP	34.	3	11,89017	-8,89017	-0,00123	-0,52316	2,50813246	0,932628	-9,07939	0,000781	
HUNG-AN	50.	0	-8,57182	8,571815	-0,84292	0,904997	4,47260904	5,121177	9,211371	0,002956	
EUP-UIA	37.	7	-1,16409	9,164087	-0,53906	0,490977	2,96640301	1,699903	8,421297	0,00094	
HUN-CHA	48.	4	12,05502	-8,05502	0,009964	-0,47455	3,92625546	3,719524	-8,51036	0,001691	
CAR-ON	22.	4	11,65152	-7,65152	-0,01064	-0,48078	6,69914579	12,7166	-9,05326	0,006951	
EUP-ALO	29.	2	8,827474	-6,82747	-0,12696	-0,40223	2,29394639	0,604357	-6,95335	0,00039	
ARC-ARC	13.	0	6,750003	-6,75	-0,21236	-0,39767	2,24367932	0,549087	-6,87006	0,000368	
HUT-COR	57.	4	10,51988	-6,51988	-0,05722	-0,3941	13,4949991	54,55241	-17,6757	0,086651	
UIA-UIA	77.	0	5,792197	-5,79219	-0,25178	-0,34124	2,16509414	0,442991	-5,88798	0,000245	
LAC-HE	68.	6	0,399795	5,610205	-0,47411	0,330518	2,19723034	0,469909	5,705915	0,000237	
HUT-AMB	56.	0	5,610174	-5,61017	-0,29327	-0,33052	4,612407015	5,503274	-6,05747	0,001176	
INU-UN7	66.	7	12,44234	-5,44234	0,021904	-0,32063	7,5539282	16,43968	-6,7963	0,003957	
ANT-CHR	5.	6	10,63766	-4,63766	-0,05237	-0,27322	2,45590549	0,853295	-4,73691	0,000204	
GNAG-NA	43.	28	23,51758	4,482424	0,477698	0,264076	4,16899099	4,318704	4,77012	0,000996	
HUT-ENG	58.	0	4,369977	-4,36998	-0,31095	-0,25668	2,25827575	0,599993	-4,43539	0,000151	
LAC-LAN	76.	5	0,824753	4,175247	-0,45521	0,245979	2,24544086	0,591228	4,249515	0,000137	
HUT-GAL	59.	0	4,102843	-4,10284	-0,32131	-0,2417	2,30894934	0,640135	-4,18002	0,00014	
HUN-MID	52.	7	2,89804	4,10196	-0,37099	0,241662	5,44818828	8,077245	4,573082	0,000935	
ANT-ANT	3.	7	3,242395	3,757605	-0,36571	0,221374	5,87613535	9,557501	4,289248	0,000948	
VER-AP	87.	0	3,499808	-3,49981	-0,34624	-0,20602	2,37124515	0,728622	-3,59951	0,000108	
EUP-CRI	30.	2	5,409941	-3,40994	-0,26764	-0,20071	4,45292282	5,067509	-3,65963	0,0004	
GNAL-EL	45.	0	3,202932	-3,20293	-0,39934	-0,1887	2,5822215	1,049923	-3,27991	0,000108	
GNAL-UNA	46.	0	3,042636	-3,04264	-0,36494	-0,17925	4,00222168	0,773795	-3,10492	8,39E-05	
EUP-UIS	31.	0	3,030871	-3,03087	-0,36543	-0,17855	2,26341081	0,575975	-3,09554	7,34E-05	
ANT-URS	12.	3	0,153126	2,846874	-0,49395	0,16772	2,22151586	0,518995	2,899498	6,23E-05	
LAC-LAC	70.	2	4,693551	-2,69355	-0,29699	-0,15869	5,49342251	8,229481	-3,00999	0,000411	
VER-SEM	84.	0	2,639476	-2,63948	-0,38153	-0,1555	2,2207396	0,517532	-2,69544	5,36E-05	
GNAL-NG	41.	6	3,370587	2,629413	-0,35144	0,154909	2,88028018	1,2054	2,69995	7,97E-05	
SEM-SEM	82.	9	6,393609	2,606391	-0,22703	0,153652	11,4924402	39,35158	4,812508	0,004606	
HUN-UNA5	55.	0	2,990121	-2,99012	-0,38396	-0,15259	2,35311127	0,702496	-2,64087	5,82E-05	
EUP-FLE	33.	0	2,564991	-2,56499	-0,3946	-0,15111	2,38229904	0,802199	-2,61208	5,35E-05	
LAC-SOO	72.	3	0,445011	2,554999	-0,47184	0,150524	2,24526572	0,590997	2,60049	5,13E-05	
AST-GRA	17.	0	2,290557	-2,29056	-0,39996	-0,13527	2,30929999	0,640095	-2,33936	4,39E-05	
LAC-SIE	75.	3	0,75096	2,249141	-0,49825	0,132605	2,35483367	0,704937	2,283278	4,39E-05	
HUN-HM	51.	0	2,186663	-2,18666	-0,40016	-0,12882	2,21701646	0,512466	-2,2246	3,66E-05	
EUP-AYA	29.	0	-2,14231	2,142307	-0,57832	0,126211	3,57499998	2,914196	2,241732	9,67E-05	
LAC-ORE	67.	3	0,972367	2,027633	-0,49014	0,119455	2,12905998	0,395724	2,080043	2,9E-05	
CAR-CAR	23.	0	2,005311	-2,00531	-0,40763	-0,11814	2,3732233	0,731499	-2,04529	3,55E-05	
HUT-ZIN	65.	0	1,950818	-1,95082	-0,40998	-0,11492	2,65032911	1,159996	-1,99936	4,23E-05	
EUP-DE	26.	0	1,904919	-1,90492	-0,41176	-0,11223	2,37826085	0,7398	-1,94306	3,22E-05	
CAR-EC	25.	2	3,75821	-1,75821	-0,33549	-0,10398	4,52604103	5,268033	-1,69279	0,000111	
ANT-LEU	8.	2	3,633774	-1,63377	-0,34061	-0,08625	2,25999195	0,570837	-1,66325	2,13E-05	
LAC-HYP	69.	0	1,472829	-1,47283	-0,42954	-0,08677	2,09617759	0,353294	-1,49554	1,48E-05	
VER-ROL	88.	0	1,459972	-1,45997	-0,43019	-0,08994	2,21793509	0,51373	-1,46228	1,63E-05	
HUN-BE	47.	2	0,545303	1,454697	-0,46771	0,085702	2,12514597	0,390643	1,477862	1,49E-05	

Tabela 21. Continuação

Resíduo padrão: OOCUMA(ocsu90c.sta)									
Tribo-Subtribo	Caso	Ordenado							
		Valor observado	Valor previsto	Resíduo padrão	Val. Prev. padrão	Resíduo	Err. padr. do v. prev.	Distância Mahans.	Resíduo deletado de Cook
HUN-MAD	52	7	2,89804	4,10196	-0,37089	0,241862	5,44816828	8,077245	4,573092
ANT-ANT	3	7	3,242395	3,757605	-0,36571	0,221374	5,87613635	9,557501	4,269248
VER-PIP	87	0	3,496908	-3,49691	-0,34624	-0,20602	2,37124515	0,728622	-3,58651
EUP-CRI	30	2	5,406841	-3,40684	-0,26764	-0,20071	4,45292282	5,067509	-3,68663
GNA-REL	45	0	3,202982	-3,20298	-0,38834	-0,18837	2,5829215	1,048923	-3,27891
GNA-UNA1	46	0	3,042636	-3,04264	-0,36494	-0,17925	2,40222189	0,773785	-3,10482
EUP-DIS	31	0	3,030671	-3,03067	-0,38543	-0,17855	2,26341081	0,575975	-3,08554
ANT-URS	12	3	0,153126	2,846874	-0,48866	0,16772	2,22151586	0,516886	2,886488
LAC-LAC	70	2	4,693651	-2,69365	-0,28899	-0,15889	5,49342251	8,228481	-3,00968
VER-SEM	84	0	2,639476	-2,63948	-0,38153	-0,18355	2,2207396	0,517532	-2,68544
GNA-ANG	41	6	3,370587	2,629413	-0,35144	0,154908	2,68026018	1,2054	2,69665
SEM-SEM	62	9	6,393609	2,606391	-0,22703	0,153552	11,4924402	39,35158	4,812508
HUN-UNA6	55	0	2,590121	-2,59012	-0,38856	-0,15259	2,35311127	0,702465	-2,64087
EUP-FLE	33	0	2,564861	-2,56486	-0,3846	-0,15111	2,28229904	0,802199	-2,61208
LAC-SCD	72	3	0,445011	2,554989	-0,47184	0,150524	2,24526572	0,580987	2,60049
AST-GRA	17	0	2,296057	-2,29606	-0,38866	-0,13527	2,30929689	0,640065	-2,33936
LAC-STE	75	3	0,75086	2,249141	-0,48925	0,132905	2,3548367	0,704937	2,293278
HUN-HYM	51	0	2,186663	-2,18666	-0,40016	-0,12882	2,21701646	0,512466	-2,2246
EUP-AYA	29	0	-2,14231	2,142307	-0,57832	0,126211	3,57469988	2,914198	2,241732
LAC-CRE	67	3	0,972367	2,027633	-0,48014	0,119455	2,12905698	0,395724	2,080043
CAR-CAR	23	0	2,005311	-2,00531	-0,40763	-0,1814	2,3732233	0,731489	-2,04529
HLT-ZIN	65	0	1,980618	-1,98062	-0,40388	-0,11492	2,65032911	1,156668	-1,98936
EUP-ADE	26	0	1,904919	-1,90492	-0,41176	-0,11223	2,37826085	0,7388	-1,94306
CAR-ECH	25	2	3,75821	-1,75821	-0,35549	-0,10388	4,52604108	5,268033	-1,89279
ANT-LEU	8	2	3,633774	-1,63377	-0,34061	-0,09525	2,25969195	0,570837	-1,66325
LAC-HYP	69	0	1,472629	-1,47263	-0,42564	-0,08577	2,09617798	0,353294	-1,49864
VER-ROL	88	0	1,468872	-1,46887	-0,43019	-0,08684	2,21793609	0,513373	-1,48228
HUN-BAE	47	2	0,545308	1,454692	-0,46771	0,085702	2,12514567	0,390643	1,477862
EUP-HEB	35	0	1,228712	-1,22871	-0,43869	-0,07239	2,22548053	0,523943	-1,2802
VER-ELE	85	0	1,153733	-1,15373	-0,44267	-0,06797	2,18888848	0,475971	-1,17326
HUN-PER	54	0	1,13325	-1,13325	-0,44651	-0,06676	2,25986113	0,571209	-1,1537
ARC-UNA1	15	0	1,010262	-1,01026	-0,44658	-0,06952	2,30103603	0,628432	-1,02918
HUN-FLA	49	0	0,703865	-0,70386	-0,46119	-0,04147	2,27224851	0,588218	-0,7167
EUP-FRA	40	1	1,894325	-0,89432	-0,42042	-0,04091	2,19226785	0,479155	-0,7051
HLT-UNA6	63	5	4,382725	0,617275	-0,30378	0,036386	2,71880474	1,288641	0,633805
CAL-UNA3	21	3	3,601852	-0,60185	-0,34192	-0,03546	2,60241842	1,079802	-0,61634
ANT-GON	6	1	0,466875	0,533125	-0,47094	0,031408	2,27257276	0,588668	0,542866
LAC-MIC	71	0	0,470603	-0,4706	-0,47079	-0,02772	2,28761195	0,809615	-0,47921
LAC-SEM	73	1	0,563955	0,436045	-0,48594	0,025689	2,22070551	0,517486	0,443638
CAL-GRA	20	0	0,409618	-0,40962	-0,4733	-0,02413	2,28762579	0,808635	-0,4172
ANT-THA	11	0	-0,39538	0,395376	-0,50547	0,023352	2,21069241	0,503933	0,403215
HLT-RUD	62	1	1,396195	-0,39619	-0,43269	-0,02334	2,96016388	1,887609	-0,40862
ANT-HAN	7	0	0,326077	-0,32608	-0,47673	-0,01921	2,2787416	0,597244	-0,33206
HLT-MEL	61	4	4,294023	-0,29402	-0,31344	-0,01732	5,20635319	7,290321	-0,32456
SEN-BLE	81	0	-0,27071	0,270709	-0,50129	0,015948	2,36818481	0,709795	0,276037
HLT-HLT	60	20	20,18667	-0,18667	0,340575	-0,01094	10,8323812	34,88082	-0,31324
EUP-HOF	36	0	-0,13824	0,138241	-0,49584	0,008144	2,34957938	0,697383	0,140942
EUP-EUP	32	16	16,10284	-0,10284	0,172549	-0,00806	3,82968843	3,490654	-0,10835
Mínimo		0	-8,57182	-59,2848	-0,84292	-3,49388	2,09617798	0,353294	-93,3883
Máximo		158	133,1934	69,68661	4,991334	4,103905	14,7773228	65,70829	102,4732
Média		11,910112	11,91011	-1,86507	9,96509	-1,2508	4,02902508	6,921348	-0,26609
Mediana		2	3,496908	-1,15373	-0,34624	-0,06797	2,50813246	0,932628	-1,17326

Figura 42. Tela de computador apresentando resultados de residuais padrão relativos a ocorrências de cumarinas para estudo de 5 outliers identificados em subtribos de Asteraceae. Residuais padrão que não caracterizam outliers têm valores de $\pm 2,5$.

Obs.: Os valores numéricos mantiveram-se originais, ou seja, com pontos em vez de vírgulas. Regress. = Regressão.

Tabela 22. Residuais de 5 outliers identificados em estudo de valores previstos vs. valores observados e respectivos pontos identificados no diagrama de dispersão (Figura 43) para ocorrências de cumarinas em subtribos de Asteraceae. (Valores previstos a partir da RLM, tomando NOs de cumarinas como variável dependente em função de NOs de diversas classes químicas).

Residual padrão:		OCCUMA (ocsu90.sta)				
Outliers	Valor observado	Valor previsto	Residual	V. prev. padrão	Residual padrão	Err. padr. do V. prev.
ANT-ACH	32	88,81783	-56,8178	3,452436	-3,40128	10,17359
AST-AST	158	106,633	51,36699	4,240768	3,074975	10,68048
LAC-SON	51	0,918215	50,08179	-0,43718	2,998039	2,289857
MUT-MUT	107	34,17818	72,82182	1,034598	4,359323	4,572986
MUT-NAS	64	14,20319	49,79681	0,150693	2,980979	2,770101
Mínimo	32	0,918215	-56,8178	-0,43718	-3,40128	2,289857
Máximo	158	106,633	72,82182	4,240768	4,359323	10,68048
Média	82,4	48,95008	33,44991	1,688264	2,002408	6,097404
Mediana	64	34,17818	50,08179	1,034598	2,998039	4,572986

Correlações marcadas são significantes em $p < 0,05$
N = 89

Comentário: Os residuais padrão das subtribos são ou < -2 (ACH) ou $> +2$ (AST, SON, MUT, NAS). Daí sua classificação como outliers.

Figura 43. Histograma mostrando a distribuição de residuais brutos de ocorrências de cumarinas ao longo do eixo X versus número de observados (frequências dos residuais brutos) no eixo Y, em subtribos de Asteraceae. (Valores obtidos a partir da RLM tomando NOs de cumarinas como variável dependente em função de NOs de diversas classes químicas).

Comentário : Dentro dos valores dos residuais brutos esperados , abaixo da curva ideal gaussiana traçada com linha vermelha, nota-se que ca. 50% dos observados possuem residual zero. No levantamento de ocorrências da classe química em questão, ca. 35% das subtribos de Asteraceae não apresetaram ocorrências de cumarinas; ca. 26% apresentaram números de ocorrências muito baixos (1 a 4). A maioria dos residuais brutos obtidos compreende valores entre -0,5 e +0,5. A distribuição dos residuais é, portanto, bastante coerente no histograma logo acima.

Figura 44. Histograma mostrando a distribuição de residuais padrão ao longo do eixo X versus números de observados(frequênciasdos residuais padrão) no eixo Y.

Comentário :Dispensável, uma vez que os residuais padrão são escores dos residuais brutos comentados logo abaixo da Figura 44.

Nota: *Residuais estandardizados (z-escore)*- Um residual estandardizado é um residual dividido por um tipo de desvio padrão , cuja fórmula, $z = y - \bar{y} / \text{desvio padrão}$. Aqui, o residual estandardizado = uma diferença entre um valor de alvo e um valor previsto (valor encaixado) dividido por um tipo de desvio padrão chamado erro padrão residual.

Tabela 23. Resumo de regressão para variável dependente: ocorrências de cumarinas (OCCUMA) após remoção de 5 outliers.

Resumo de regressão para a variável dependente: OCCUMA

R = 0,93843766 R² = 0,88066525 R²ajustado = 0,86967389
F(7,76) = 80,123 p < 0,00000 Err. padr. do estimativo : 7,8041

	BETA	Err. padr. de BETA	B	Err. padr. de B	t(76)	nível-p
Intercepto			-1,03452	1,082934	-0,9553	0,342456
OCMONO	0,843971	0,046794	0,082223	0,004559	18,03569	3,5E-29
OCSELA	-0,18417	0,048954	-0,02278	0,006054	-3,76208	0,000329
OCDITE	-0,00578	0,052853	-0,00107	0,009776	-0,10929	0,913263
OCTRIT	0,523695	0,045713	0,5797	0,050602	11,45617	3,17E-18
OCPOLI	0,026305	0,041486	0,005173	0,008159	0,63407	0,527939
OCFLAY	0,127106	0,046601	0,034013	0,01247	2,72756	0,00792
OCBENZ	0,050038	0,052058	0,04959	0,051592	0,961199	0,3395

Figura 45. Diagramas de dispersão de valores previstos vs. valores observados para ocorrências de cumarinas em subtribos de Asteraceae. (Valores previstos obtidos a partir de regressões lineares múltiplas, tomando NOs de cumarinas como variável dependente em função de NOs de diversas classes químicas). Diagrama acima: com os outliers (novamente Fig. 45). Diagrama abaixo: após remoção dos 5 outliers.

Comentário: A remoção de 5 *outliers* da Tabela 23 teve o intuito de exemplificar sua influência na feição do diagrama de dispersão. O que trouxe de positividade neste caso:

- a) R múltiplo da RLM passou de 0,81482546 (com os *outliers*) para 0,93843766 (sem os 5 *outliers*).
- b) o coeficiente de regressão BETA, relacionado à correlação parcial OCCUMA / OCMONO passou de 0,661934 (com os *outliers*) para 0,843971 (após a remoção dos 5 *outliers*), reforçando mais ainda a influência das ocorrências dos monoterpenos sobre as ocorrências das cumarinas, e vice-versa, em tribos e subtribos de Asteraceae.
- c) Melhora da dispersão dos pontos ao longo da curva.
- d) Melhora de valores relacionados aos demais coeficientes de regressão BETA , que são significantes para as ocorrências das cumarinas em subtribos de Asteraceae (β -SELA = -0,18417; β -TRIT = 0,523695), porém com perda significativa no valor de β -DITE; mas, por outro lado, a previsão de ocorrências de cumarinas passa a ter influência positiva das ocorrências dos flavonóides, β -FLAV = 0,127126.

4.8. Grupos de cumarinas classificadas por tipo de esqueleto carbônico em Asteraceae

Os diversos tipos de esqueleto carbônico de cumarinas presentes (alguns supostamente) em Asteraceae foram codificados com números de 321 a 376 e classificados por grupos, conforme descrito abaixo:

GI-Grupo I : Cumarinas simples, esqueleto 321

GII-Grupo II: 5-Metilcumarinas simples, esqueleto 323

GIII-Grupo III: Cumarinas simples preniladas (C-8), esqueletos 350, 357

GIV-Grupo IV: 4-Metilcumarinas simples, esqueleto 322

GV-Grupo V: 5-Etilcumarinas preniladas (C-3), esqueletos 328, 343, 351, 362.

GVI-Grupo VI: 5-Metilcumarinas preniladas / norpreniladas (C-3), esqueletos 324, 325, 334, 358, 361

GVII-Grupo VII: 5-Metilcumarinas geraniladas (C-3), esqueletos 329, 330, 335, 372

GVIII-Grupo VIII: 5-Etilcumarinas geraniladas (C-3), esqueleto 344

GIX-Grupo IX: 5-Metilcumarinas farnesiladas (C-3), esqueletos 333, 336

GX-Grupo X: 5-Metilcumarinas sesquiterpênicas (C-3), esqueletos, 346, 347, 348, 365, 366 , 367

GXI-Grupo XI: 5-Metilcumarinas monoterpênicas (C-3), esqueletos 370, 376

GXII-Grupo XII: Furanocumarinas, esqueleto 353, 354, 355, 356

GXIII-Grupo XIII: 4-Fenilcumarinas (neoflavonóides), 3-fenilcumarinas (cumestrol), esqueletos 363, 374

GXIV-Grupo XIV: 3-Fenil-5-metilcumarinas, esqueleto 359

GXV-Grupo XV: Bicumarinas , esqueletos 349, 375

Cada grupo com seus respectivos esqueletos cumarínicos (estruturas) está representado abaixo:

GI-Grupo I : cumarinas simples, esqueletos 321

Nota : $R_1, R_2, R_3, R_4, R_5 = H, OR, SR$; R = qualquer grupo

GII-Grupo II: 5-Metilcumarinas simples, esqueletos 323

Nota : $R_1, R_2, R_3, R_4, R_5 = H, OR, SR$; R = qualquer grupo

GIII-Grupo III: Cumarinas simples preniladas (C-8), esqueletos 350, 357; (C-6), esqueleto 352

GIV-Grupo IV: 4-Metilcumarinas simples, esqueleto 322

Nota : $R_1, R_2, R_3, R_4, R_5 = H, OR, SR$; R = qualquer grupo

GV-Grupo V: 5-Etilcumarinas preniladas (C-3), esqueletos 328, 343, 351, 362

GVI-Grupo VI: 5-Metilcumarinas preniladas / norpreniladas (C-3), esqueletos 324, 325, 334, 358, 361

GVII-Grupo VII: 5-Metilcumarinas geraniladas (C-3), esqueletos 326, 329, 330, 335, 369, 372

GVIII-Grupo VIII: 5-Etilcumarinas geraniladas (C-3), esqueletos 337, 344

GIX-Grupo IX: 5-Metilcumarinas farnesiladas (C-3), esqueletos 331, 333, 336

GX-Grupo X: 5-Metilcumarinas sesquiterpênicas (C-3), esqueletos, 346, 347, 348, 365, 366, 367

GXI-Grupo XI: 5-Metilcumarinas monoterpênicas (C-3), esqueletos 370, 376

GXII-Grupo XII: furanocumarinas, esqueleto 353, 354, 355, 356

GXIII-Grupo XIII: 4-Fenilcumarinas (neoflavonóides), 3-fenilcumarinas (cumestrol), esqueletos 363, 374

GXIV-Grupo XIV: 3-Fenil-5-metilcumarinas, esqueleto 359 **GXV-Grupo XV: Bicumarinas , esqueletos 349, 375**

Tabela 25. Resumo de grupos de cumarinas e números totais de ocorrências de esqueletos por grupo em Asteraceae.

GRUPOS	T. ESQ
GI	712
GII	69
GIII	4
GIV	2
GV	24
GVI	21
GVII	31
GVIII	2
GIX	21
GX	26
GXI	10
GXII	25
GXIII	9
GXIV	5
GXV	1

Comentário : A maioria das tribos de Asteraceae acumula cumarinas mais simples (esq 321), mantendo, portanto, o esqueleto benzopirônico sem aumento de átomos de carbono diretamente ligados à sua estrutura. A seguir, **os tipos de esqueleto mais predominantes são as cumarinas metiladas em C-5**, possíveis marcadores químicos para as tribos Mutisieae e Vernonieae. É importante enfatizar a química de ambas estas tribos relacionada a cumarinas: acumulam estruturas das mais simples às mais complexas. O diagrama tridimensional a seguir (Figura 46) mostra a distribuição dos grupos (tabela 24 ou Tabela 25) nas tribos de Asteraceae.

Figura 46. Diagrama tridimensional mostrando números de ocorrências de esqueletos de cumarinas por grupo (pode conter um ou vários tipos de esqueleto) em tribos de Asteraceae

Comentário: Há fortes indícios de que alguns esqueletos cumarínicos (códigos correspondentes marcados em vermelho na Tabela 26) podem servir como marcadores químicos em nível de tribo e subtribo para Vernonieae; em nível de tribo ou subtribo para Mutisieae. Ambas as tribos apresentam um quimismo muito interessante e com diversidade de esqueletos. Merecem um estudo mais aprofundado em nível macromolecular e enzimático.

Tabela 26. Esqueletos de cumarinas distribuídos em subtribos de Asteraceae

Esqueletos de cumarinas distribuídos nas subtribos de Asteraceae

Tribo	Subtribo	Código do esq.	Total por subtribo	Tribo	Subtribo	Código do esq.	Total por subtribo
ANT	ACH	321	31	LAC	CRE	321	3
ANT	ACH	322	1	LAC	HIE	321	6
ANT	ANT	321	7	LAC	SCO	321	3
ANT	ART	321	158	LAC	SON	321	51
ANT	CHR	321	6	LAC	STE	321	3
ANT	GON	321	1	LAC	UNA	321	5
ANT	MAT	321	16	LAC	LAC	353	2
ANT	TAN	321	4	MUT	MUT	321	12
ANT	URS	321	3	MUT	NAS	321	2
ANT	MAT	365	1	MUT	NAS	322	1
ARC	GOR	321	1	MUT	NAS	323	32
AST	AST	321	157	MUT	MUT	323	30
AST	SOL	321	20	MUT	MUT	326	2
AST	AST	350	1	MUT	NAS	326	6
AST	SOL	375	1	MUT	MUT	329	1
CAL	UNA	321	3	MUT	MUT	330	5
CAR	CAD	321	4	MUT	NAS	331	3
CAR	CEN	321	18	MUT	MUT	333	3
CAR	ECH	363	2	MUT	MUT	333	1
EUP	AGE	321	5	MUT	NAS	334	5
EUP	ALO	321	2	MUT	NAS	336	14
EUP	CRI	321	2	MUT	MUT	337	1
EUP	EUP	321	16	MUT	MUT	346	8
EUP	LIA	321	7	MUT	MUT	348	1
EUP	MIK	321	3	MUT	MUT	353	20
EUP	OXY	321	1	MUT	MUT	358	2
EUP	PRA	321	1	MUT	MUT	359	2
GNA	ANG	321	6	MUT	MUT	359	1
GNA	CAS	321	11	MUT	NAS	365	1
GNA	GNA	321	28	MUT	MUT	365	2
GNA	LOR	321	6	MUT	MUT	366	6
HLN	BAE	321	2	MUT	MUT	367	7
HLN	CAI	321	1	MUT	MUT	369	1
HLN	CHA	321	1	MUT	MUT	370	3
HLN	MAD	321	7	MUT	MUT	372	1
HLN	PEC	321	7	MUT	MUT	376	5
HLN	CHA	350	2	PLU	UNA	321	29
HLN	CHA	352	1	SEN	SEN	321	9
HLT	COR	321	4	SEN	TUS	321	14
HLT	HLT	321	20	SEN	TUS	370	2
HLT	MEL	321	4	VER	VER	321	3
HLT	RUD	321	1	VER	VER	323	7
HLT	UNA	321	1	VER	VER	324	8
HLT	VER	321	3	VER	VER	325	3
HLT	UNA	327	2	VER	VER	328	2
HLT	VER	327	2	VER	VER	330	2
HLT	UNA	374	2	VER	VER	335	12
HLT	VER	374	1	VER	VER	343	2
INU	UNA	321	4	VER	VER	344	1
INU	UNA	353	2	VER	VER	351	18
INU	UNA	356	1	VER	VER	358	1
				VER	VER	359	2
				VER	VER	361	2
				VER	VER	362	2
				VER	VER	372	1

Ocorrências de benzofuranos

Sumário de regressão linear múltipla para variável dependente OCBENZ (ocorrências de benzofuranos) em função de ocorrências de diversas classes químicas (MONO, DITE, SELA, TRIT, CUMA, FLAV) em tribos de Asteraceae

Tabela 27. Sumário de regressão linear múltipla padrão tomando como variável dependente ocorrências de benzofuranos em função de ocorrências de diversas classes químicas, em tribos de Asteraceae.

Sumário de regressão linear múltipla para variável dependente OCBENZ em tribos de Asteraceae

$R=0,90138630$ $R^2=0,81249726$ $R_{ajustado}=0,66666180$

$F(7,9)=5,5713$ $p<0,01025$ Err. padr. da estimativa: 39,307

	BETA	Err.padr. de BETA	B	Err.padr. de B	t(9)	nível-p
Intercepto			-5,29161	17,64225	-0,29994	0,771035
OCMONO	-0,19205	0,31581	-0,01521	0,025013	-0,60813	0,558133
OCSELA	0,719136	0,248788	0,098971	0,034239	2,890554	0,017868
OC DITE	0,74384	0,204461	0,111447	0,030634	3,638044	0,005417
OC TRIT	-0,04911	0,189533	-0,07625	0,294321	-0,25908	0,801399
OC CUMA	-0,17343	0,255344	-0,16908	0,24893	-0,67921	0,51409
OC POLI	-0,27601	0,221127	-0,0426	0,034129	-1,24821	0,243447
OC FLAV	-0,08334	0,274374	-0,01925	0,063385	-0,30375	0,768221

Valores de BETA em vermelho são significantes em um nível- $p < 0,05$; Err. padr. = Erro padrão.

Comentário : Os coeficientes de regressão BETA apontam uma forte correlação positiva entre as ocorrências de benzofuranos e as de diterpenos ($\beta = 0,74$) e sesquiterpenóides ($\beta = 0,72$). Deve-se observar também que as ocorrências dos demais metabólitos têm coeficientes de regressão negativos, apesar de não serem significantes na equação obtida (Equação 5).

São marcantes os valores de coeficientes de regressão BETA relacionados às ocorrências de diterpenos (DITE) e sesquiterpenóides (SELA), que influenciam positivamente as ocorrências de benzofuranos, tomadas como variável dependente em tribos de Asteraceae

Equação 5: $OC_{BENZ} = -5,29 - 0,08 OC_{FLAV} - 0,17 OC_{CUMA} - 0,19 OC_{MONO} + 0,74 OC_{DITE} + 0,72 OC_{SELA} - 0,05 OC_{TRIT} - 0,28 OC_{POLI}$

Comentário: A Equação 5 foi gerada a partir dos valores na Tabela 27, obtidos mediante regressão linear múltipla padrão de ocorrências de benzofuranos como

variável dependente em função de ocorrências de diversas classes químicas em tribos de Asteraceae.

Tabela 28: Valores previstos *versus* valores observados para ocorrências de benzofuranos em tribos de Asteraceae (Valores previstos obtidos através de regressão linear múltipla padrão, tomando números de ocorrências de benzofuranos como variável dependente em função de números de ocorrências de diversas classes químicas).

**Valores previstos versus observados para
ocorrências de benzofuranos em tribos de Asteraceae**

TRIBO	Previsto	Observado
ANT	1,177894	4
ARC	-4,4089	0
AST	100,4674	58
BAR	-8,98938	0
CAL	5,503507	7
CAR	-8,16488	0
EUP	132,0023	206
GNA	35,32416	30
HLN	26,54163	9
HLT	190,8169	197
INU	64,02428	0
LAC	-7,76654	2
LIA	1,047314	0
MUT	-25,7429	6
PLU	-10,3085	0
SEN	107,8112	108
VER	28,66459	1

Comentário : Os valores acima marcados em vermelho apresentam, de certa forma, uma boa coerência de previsão, uma vez que a margem de erro dos previstos atinge um máximo de apenas ca. $\pm 20\%$ dos observados, valor considerado já de boa aceitação num sistema de ocorrências de qualquer produto natural de plantas.

Figura 47. Valores previstos *versus* valores observados para ocorrências de benzofuranos em tribos de Asteraceae. (Valores previstos obtidos através de regressão linear múltipla padrão, tomando números de ocorrências de benzofuranos como variável dependente em função de números de ocorrências de diversas classes químicas).

Comentário : São marcantes os valores de coeficientes de regressão BETA relacionados às ocorrências de diterpenos (DITE) e sesquiterpenóides (SELA), que influenciam positivamente as ocorrências de benzofuranos, tomadas como variável dependente em tribos de Asteraceae.

A partir do sumário da regressão linear múltipla padrão (Tabela 28) e do gráfico apresentado na Figura 47, foi obtida uma equação empírica (Equação 5), para previsão de ocorrências de benzofuranos em tribos de Asteraceae.

Na Tabela 28, são apresentadas as ocorrências de diterpenos (DITE) e benzofuranos (BENZ) em tribos de Asteraceae e, na Figura 48, um diagrama de correlação bivariada de ocorrências de benzofuranos *versus* ocorrências de diterpenos. A partir do sumário da regressão linear múltipla (Tabela 29) e do gráfico apresentado na Figura 48, foi obtida uma equação empírica (Equação 5) para previsão de ocorrências de benzofuranos em tribos de Asteraceae.

Tabela 28. Ocorrências de diterpenos (DITE) e benzofuranos (BENZ) em tribos de Asteraceae

TRIBO	DITE	BENZ
	X	Y
ANT	15	4
ARC	21	0
AST	1419	58
BAR	0	0
CAL	93	7
CAR	6	0
EUP	829	206
GNA	419	30
HLN	124	9
HLT	1230	197
INU	68	0
LAC	0	2
LIA	0	0
MUT	30	6
PLU	0	0
SEN	71	108
VER	44	1

Comentário: É importante observar as não-ocorrências de ambas as classes químicas nas tribos Barnadesieae, Liabeae e Plucheeae. Nas tribos Arctoteae, Cardueae e Inuleae , verifica-se produção de diterpenos, porém ausência de benzofuranos; e o contrário ocorre na tribo Lactuceae. É interessante notar que o quimismo de Vernonieae relacionado a cumarinas não se observa no tocante a benzofuranos (apenas 1 ocorrência registrada).

Figura 48. Diagrama dispersão mostrando valores previstos para ocorrências de benzofuranos em função de ocorrências de diterpenos, a partir da correlação bivariada entre as ocorrências de ambas as classes químicas.

Equação 6: $OC_{BENZ} = -5,29 - 0,08 OC_{FLAV} - 0,17 OC_{CUMA} - 0,19 OC_{MONO} + 0,74 OC_{DITE} + 0,72 OC_{SELA} - 0,05 OC_{TRIT} - 0,28 OC_{POLI}$

Comentário: As ocorrências dos benzofuranos em tribos de Asteraceae são positivamente influenciadas pelas ocorrências de diterpenos e sesquiterpenóides (valores de BETA 0,74 e 0,72, respectivamente). Numa regressão bivariada tomando como variáveis as ocorrências de benzofiranos e diterpenos, a equaçõ supra passa a ser : $OC_{benz} = 9,23 + 0,108 *OC_{DITE}$, a qual apresnta um $r = 0,72$.

Ocorrências de benzofuranos em subtribos de Asteraceae

Tabela 29. Sumário de regressão linear múltipla de ocorrências de benzofuranos, tomadas como variável dependente, em subtribos de Asteraceae

Sumário de regressão para variável dependente: OCBENZ

R= 0,64490810 R²= 0,41590646 R²ajustado = 0,36542924

F(7 81) = 8,2395 p< 0,00000 Erro padrão da estimativa: 17,191

	BETA	Err. padr. de BETA	B	Err. padr. de B	t(81)	nível-p
Intercepto			2,293737	2,309709	0,993085	0,323626
OCMONO	0,113085	0,1403559	0,009848	0,012223	0,805704	0,422773
OCSELA	0,033162	0,1023091	0,0042	0,012957	0,324136	0,74667
OCDITE	0,580813	0,1014865	0,094072	0,016437	5,723061	1,7E-07
OCTRIT	0,350435	0,1208095	0,376284	0,129721	2,900725	0,004792
OCUMA	-0,14326	0,1516975	-0,10569	0,111915	-0,94436	0,347792
OCPOLI	0,021486	0,0904874	0,004202	0,017697	0,237446	0,812911
OCFLAV	-0,16546	0,1066849	-0,03805	0,024991	-1,52241	0,131801

Figura 49. Diagrama de dispersão para valores previstos vs. observados de ocorrências de benzofuranos em subtribos de Asteraceae. Valores previstos obtidos a partir da regressão linear múltipla tomando OCBENZ como variável dependente

Comentário: Alguns pontos mostram as tribos e respectivas subtribos como ilustração. A dispersão dos pontos é acentuada, e a tendência de formação de *clusters* se explica pela não-ocorrência de benzofuranos na maioria das subtribos. (Tabela 28)

Tabela 30. Valores previstos e observados de ocorrências de benzofuranos em subtribos de Asteraceae (valores previstos obtidos a partir da regressão linear múltipla tomando OCBENZ como variável dependente).

Ocorrências de benzofuranos em subtribos de Asteraceae

TRIBO	SUBTRIBO	Previsto	Observado	TRIBO	SUBTRIBO	Previsto	Observado
2 ANT	ACH	-0,710472941	0	45 GNA	REL	8,126188278	0
3 ANT	ANT	-7,405138493	0	46 GNA	UNA4	5,763570309	12
4 ANT	ART	4,548604488	4	47 HLN	BAE	0,805838525	2
5 ANT	CHR	2,973757029	0	48 HLN	CHA	4,739414692	0
6 ANT	GON	2,134401798	0	49 HLN	FLA	2,614939451	0
7 ANT	HAN	2,314736605	0	50 HLN	GAI	2,576883554	0
8 ANT	LEU	3,473353386	0	51 HLN	HYM	4,07512331	0
9 ANT	MAT	2,735125303	0	52 HLN	MAD	-5,42454052	3
10 ANT	TAN	5,643339157	0	53 HLN	PEC	-0,425432861	4
11 ANT	THA	2,370287657	0	54 HLN	PER	0,701188922	0
12 ANT	URS	2,054815292	0	55 HLN	UNAS	3,763607502	0
13 ARC	ARC	8,138951302	0	56 HLT	AMB	7,160833836	0
14 ARC	GRT	12,36175537	0	57 HLT	COR	-1,214856029	2
15 ARC	UNA1	3,627543688	0	58 HLT	ENG	4,217166424	0
16 AST	AST	64,92581177	44	59 HLT	GAL	3,99636507	0
17 AST	GRA	4,733848095	0	60 HLT	HLT	31,61474228	25
18 AST	SOL	67,27748108	14	61 HLT	MEL	15,08149624	15
19 BAR	UNA2	1,589496255	0	62 HLT	RUD	2,905076504	0
20 CAL	GRA	2,302136898	0	63 HLT	UNA6	4,382633004	4
21 CAL	UNA3	10,86374855	7	64 HLT	VER	69,9357605	150
22 CAR	CAD	7,051821232	0	65 HLT	ZIN	4,371476173	1
23 CAR	CAR	1,704776764	0	66 INU	UNA7	12,66303444	0
24 CAR	CEN	3,562497854	0	67 LAC	CRE	2,602354527	0
25 CAR	ECH	3,479727983	0	68 LAC	HIE	0,813918889	0
26 EUP	ADE	5,406507969	0	69 LAC	HYP	2,763991594	0
27 EUP	AGE	23,9128952	21	70 LAC	LAC	-4,022238731	0
28 EUP	ALO	12,49875184	13	71 LAC	MIC	2,255889621	0
29 EUP	AYA	3,572719336	29	72 LAC	SCO	1,730036974	2
30 EUP	CFI	15,49717808	48	73 LAC	SEN	1,465149522	0
31 EUP	DIS	7,39411068	4	74 LAC	SON	-5,168065548	0
32 EUP	EUP	7,13925457	27	75 LAC	STE	-0,115930788	0
33 EUP	FLE	4,47838068	0	76 LAC	UNA8	1,518556714	0
34 EUP	GYP	13,76633739	21	77 LIA	LIA_	5,712049484	0
35 EUP	HEB	3,80850482	1	78 MUT	MUT	8,411931038	6
36 EUP	HOF	2,764086022	8	79 MUT	NAS	2,536496162	0
37 EUP	LIA	3,608822584	23	80 PLU	UNA9	1,563509345	0
38 EUP	MIK	16,49996567	0	81 SEN	BLE	2,481880426	8
39 EUP	OXY	9,109221458	11	82 SEN	SEN	25,35348511	40
40 EUP	PRA	2,331026554	0	83 SEN	TUS	3,630782127	60
41 GNA	ANG	10,1542511	0	84 VER	CEN	2,544109344	0
42 GNA	CAS	-0,415556878	0	85 VER	ELE	3,197507858	0
43 GNA	GNA	15,72063351	15	86 VER	LYC	14,95413399	0
44 GNA	LOR	8,497038841	3	87 VER	PIP	6,654631615	0
				88 VER	ROL	3,134502649	0
				89 VER	VER	49,27354431	1

A partir do sumário da regressão linear múltipla (Tabela 30), foi gerada a equação empírica 7 para previsão de ocorrências de benzofuranos em subtribos de Asteraceae.

$$\text{Equação 7. } OC_{\text{BENZ}} = 0,1131 - 0,17 OC_{\text{FLAV}} - 0,14 OC_{\text{CUMA}} + 0,58 OC_{\text{DITE}} + 0,33 OC_{\text{SELA}} + 0,35 OC_{\text{TRIT}} + 0,021 OC_{\text{POLI}}$$

Comentário: As influências positivas das ocorrências de diterpenos e sesquiterpenóides sobre a previsão de ocorrências de benzofuranos em subtribos de Asteraceae são mantidas como nas tribos, porém em menor intensidade (veja valores de BETA na equação supra). Diterpenos, mais uma vez, exercem maior influência na previsão de ocorrências de benzofuranos.

4.8. Grupos de benzofuranos classificadas por tipo de esqueleto carbônico em Asteraceae

GRUPOS DE BENZOFURANOS CLASSIFICADOS
POR TIPO DE ESTRUTURA (ESQUELETO CARBÔNICO)

$R_1, R_2, R_3, R_4 = \text{H, OR}$ (R, qualquer esqueleto carbônico)

A e B ocorrem naturalmente com maior frequência em Asteraceae

GI-Grupo I: benzofuranos só isopropenilados no anel furânico; esqueletos 377, 393

GII-Grupo II: benzofuranos só acetilados no anel benzênico (C-5); esqueleto 403

GIII-Grupo III: benzofuranos só acetilados no anel furânico (C-2); esqueleto 404

GIV-Grupo IV: benzofuranos simples, acetilados em C-5 e isopropenilados em C-2; esqueletos 378, 398

GV-Grupo V: benzofuranos simples, acetilados em C-6, isopropenilados em C-2; esqueleto 383

GVI-Grupo VI: benzofuranos simples, acetilados em C-7 e isopropenilados em C-2; esqueleto 401

GVII-Grupo VII: benzofuranos simples, acetilados em C-5, isopropenilados ou acetilados em C-2 e metilados no anel benzênico (C-6 ou C-7); esqueletos 379, 391

GVIII-Grupo VIII: benzofuranos acetilados em C-5, isopropenilados em C-2, prenilados em C-7; esqueleto 388

GIX-Grupo IX: benzofuranos simples, metilados no anel benzênico (C-4) e isopropenilados em C-2; esqueleto 392

GX-Grupo X: benzofuranos simples, etilados no anel benzênico (C-4) e isopropenilados em C-2; esqueleto 390

GXI-Grupo XI: benzofuranos simples, metilados nos anéis benzênico (C-6) e furânico (C-3); esqueleto 399

GXII-Grupo XII: benzofuranos metilados nos anéis benzênico (C-5) e furânico (C-3), grupo alifático ligado ao C-5; esqueletos 389, 396 (dímero de 389, C-7/C-7)

GXIII-Grupo XIII: benzofuranos acetilados em C-5 e propilados em C-2; esqueleto 384

GXIV-Grupo XIV: benzofuranos complexos com grupo isopropil adicionado ou não ao grupo acetato do C-5, prenilados ou não no grupo isopropenil em C-2 e ligados a grupos lactonizados em C-7; esqueletos 380, 381, 394, 395

GXV-Grupo XV: benzofuranos complexos; com o anel benzofurânico ligado a cromenos, grupos fenilalquilas, alicíclicos ou outros; esqueletos 385, 386, 387, 400, 402

GI-Grupo I: benzofuranos isopropenilados no anel furânico; esqueletos. 377 e 393

GII-Grupo II: benzofuranos só acetilados no anel benzênico (C-5); esqueleto 403

GIII-Grupo III: benzofuranos só acetilados no anel furânico (C-2); esqueleto 404

GIV-Grupo IV: benzofuranos simples, acetilados em C-5 e isopropenilados em C-2; esqueletos 378, 398

GV-Grupo V: benzofuranos simples, acetilados em C-6, isopropenilados em C-2; esqueleto 383

GVI-Grupo VI: benzofuranos simples, acetilados em C-7 e isopropenilados em C-2; esqueleto 401

VII-Grupo VII: benzofuranos simples, acetilados em C-5, isopropenilados ou acetilados em C-2 e metilados no anel benzênico (C-6 ou C-7); esqueletos 379, 391

GVIII-Grupo VIII: benzofuranos acetilados em C-5, isopropenilados em C-2, prenilados em C-7; esqueleto 388

GIX-Grupo IX: benzofuranos simples, metilados no anel benzênico (C-4) e isopropenilados em C-2; esqueleto 392

GX-Grupo X: benzofuranos simples, etilados no anel benzênico (C-4) e isopropenilados em C-2; esqueleto 390

GXI-Grupo XI: benzofuranos simples, metilados nos anéis benzênico (C-6) e furânico (C-3); esqueleto 399

GXII-Grupo XII: benzofuranos metilados nos anéis benzênico (C-5) e furânico (C-3), grupo alifático ligado ao C-5; esqueletos 389, 396 (dímero de 389, C₇/C₇)

XIII-Grupo XIII: benzofuranos acetilados em C-5 e propilados em C-2; esqueleto 384

GXIV-Grupo XIV: benzofuranos complexos com grupo isopropil adicionado ou não ao grupo acetato do C-5, prenilados ou não no grupo isoprenil em C-2 e ligados a grupos lactonizados em C-7; esqueletos 380, 381, 394, 395

GXV-Grupo XV: benzofuranos complexos; com o anel benzofurânico ligado a cromenos, grupos fenilalquilas, alicíclicos ou outros; esqueletos 385, 386, 387, 400, 402

Tabela 31. Grupos de benzofuranos com suas respectivas quantidades de esqueletos por tribo de Asteraceae

GRUPOS DE BENZOFURANOS COM SUAS RESPECTIVAS QUANTIDADES DE ESQUELETOS POR TRIBO DE ASTERTACEAE

Tribo	GI		GII		GIII		GIV		GV		GVI	
	Esq377	Esq393	Esq403	Esq404	Esq378	Esq390	Esq383	Esq401				
MUT	0	0	0	0	1	0	5	0				
LAC	0	0	0	0	0	0	0	0				
VER	0	0	0	0	0	0	0	0				
GNA	0	0	0	0	0	0	0	0				
CAL	0	0	0	0	7	0	0	0				
AST	0	0	0	0	37	0	14	1				
ANT	0	0	0	0	2	0	0	0				
SEN	0	9	3	2	85	2	0	0				
HLN	0	1	0	0	8	0	0	0				
HLT	0	5	0	3	123	19	0	0				
EUP	1	6	0	1	174	0	3	2				

Tribo	GVII		GVIII		GIX		GX		GXI		GXII	
	Esq379	Esq391	Esq388	Esq392	Esq390	Esq399	Esq389	Esq396				
MUT	0	0	0	0	0	0	0	0				
LAC	0	0	0	0	0	0	0	0				
VER	1	0	0	0	0	0	0	0				
GNA	1	0	0	0	0	0	0	0				
CAL	0	0	0	0	0	0	0	0				
AST	1	0	0	0	2	0	0	0				
ANT	0	0	0	0	0	0	0	0				
SEN	1	0	0	0	0	0	4	1				
HLN	0	0	0	0	0	0	0	0				
HLT	17	12	0	0	2	0	0	0				
EUP	2	0	9	1	5	2	0	0				

Tribo	GXIII		GXIV				GXV				
	Esq384	Esq380	Esq381	Esq394	Esq395	Esq385	Esq386	Esq387	Esq400	Esq402	
MUT	0	0	0	0	0	0	0	0	0	0	
LAC	0	0	0	0	0	0	0	0	0	0	1
VER	0	0	0	0	0	0	0	0	0	0	0
GNA	0	0	0	0	0	0	0	0	0	0	0
CAL	0	0	0	0	0	0	0	0	0	0	0
AST	0	0	0	0	0	4	0	1	0	0	0
ANT	0	0	0	0	0	0	0	0	0	0	0
SEN	1	0	0	0	0	0	0	0	0	0	0
HLN	0	0	0	0	0	0	0	0	0	0	0
HLT	0	0	0	0	0	15	1	0	0	0	0
EUP	0	0	0	0	0	0	0	0	0	0	1

Tabela 32. Grupos de benzofuranos e números totais (T. ESQ) de ocorrências de esqueletos por grupo em Asteraceae

Grupo	T. esq.
GI	23
GII	3
GII	6
GIV	472
GV	21
GVI	3
GVII	34
GVIII	9
GIX	1
GX	9
GXI	2
GXII	5
GXIII	1
GIV	13
GXV	26

Comentário : A maioria das tribos de Asteraceae acumula benzofuranos mais simples acetilados em C-5 e isopropenilados em C-2 (esqueletos 378 e 398), mantendo, portanto, o esqueleto benzofurânico sem aumento de átomos de carbono diretamente ligados a este e respectivos grupos em C-5 e C-2.

Tabela 33. Esqueletos de benzofuranos distribuídos em subtribos de Asteraceae

TRIBO	SUBTRIBO	CÓD.ESQ.	TOTAL	TRIBO	SUBTRIBO	CÓD. ESQ.	TOTAL
ANT	ART	378	2	HLN	BAE	378	1
ANT	ART	387	2	HLN	BAE	393	1
AST	AST	378	25	HLN	MAD	378	3
AST	AST	383	13	HLN	PEC	378	4
AST	AST	385	4	HLT	COR	378	2
AST	AST	387	1	HLT	HLT	378	13
AST	AST	401	1	HLT	HLT	391	12
AST	SOL	378	10	HLT	MEL	378	6
AST	SOL	379	1	HLT	MEL	386	1
AST	SOL	383	1	HLT	MEL	393	5
AST	SOL	390	2	HLT	MEL	398	1
CAL	UNA	378	7	HLT	MEL	404	2
EUP	AGE	378	12	HLT	UNA	378	3
EUP	AGE	383	2	HLT	UNA	398	1
EUP	AGE	390	5	HLT	VER	378	98
EUP	AGE	393	1	HLT	VER	379	17
EUP	AGE	402	1	HLT	VER	385	15
EUP	ALO	378	11	HLT	VER	390	2
EUP	ALO	401	2	HLT	VER	398	17
EUP	AYA	378	29	HLT	VER	404	1
EUP	CRI	378	36	HLT	ZIN	378	1
EUP	CRI	379	2	LAC	SCO	400	2
EUP	CRI	388	9	MUT	MUT	378	1
EUP	CRI	393	1	MUT	MUT	383	5
EUP	DIS	378	4	SEN	BLE	378	7
EUP	EUP	378	24	SEN	BLE	403	1
EUP	EUP	393	2	SEN	SEN	378	37
EUP	EUP	404	1	SEN	SEN	393	1
EUP	GYP	377	1	SEN	SEN	398	2
EUP	GYP	378	20	SEN	TUS	378	41
EUP	HEB	378	1	SEN	TUS	379	1
EUP	HOF	378	8	SEN	TUS	384	1
EUP	LIA	378	23	SEN	TUS	389	4
EUP	OXY	378	6	SEN	TUS	393	8
EUP	OXY	392	1	SEN	TUS	396	1
EUP	OXY	393	2	SEN	TUS	403	2
EUP	OXY	399	2	SEN	TUS	404	2
GNA	GNA	378	2	VER	VER	379	1
GNA	GNA	380	2				
GNA	GNA	381	4				
GNA	GNA	394	3				
GNA	GNA	394	1				
GNA	GNA	395	2				
GNA	GNA	395	1				
GNA	LOR	378	3				
GNA	UNA	378	11				
GNA	UNA	393	1				

Comentário: Alguns esqueletos benzofurânicos (códigos correspondentes marcados em vermelho na Tabela 33) podem possivelmente servir como

marcadores químicos em nível de tribo e subtribo : em EUP-CRI, esqueleto 388; em GNA-GNA, esqueleto 381 e em HLT-HLT, esqueleto 391. . Em nível de espécies , por exemplo, *Encelia* sp., benzofuranos já são considerados como marcadores químicos (Proksch and Clark, 1987). O diagrama tridimensional abaixo (Figura 50) mostra a distribuição dos grupos de esqueletos de benzofuranos (Tabela 32) em tribos de Asteraceae.

Figura 50. Diagrama tridimensional mostrando números de ocorrências de esqueletos de benzofuranos por grupo (um grupo pode conter um ou vários tipos de esqueleto) em tribos de Asteraceae

FLAVONÓIDES

Tabela 34. Sumário de regressão linear múltipla de ocorrências de benzofuranos, tomadas como variável dependente, em tribos de Asteraceae

Valores previstos e observados para ocorrências de flavonóides em tribos de Asteraceae

	Previsto	Observado
1 ANT	869,5449	852
2 ARC	69,95776	0
3 AST	567,2823	454
4 BAR	87,4389	139
5 CAL	129,6772	1
6 CAR	336,0006	222
7 EUP	271,4472	312
8 GNA	242,3441	352
9 HLN	266,8882	665
10 HLT	806,8855	861
11 INU	244,6636	109
12 LAC	145,766	432
13 LJA	100,1646	8
14 MUT	122,345	85
15 PLU	121,5056	11
16 SEN	191,5806	18
17 VER	-12,482	40

Sumário de regressão para variável dependente: OCFLAV

R =0,85209330 R²= 0,72606299 R²ajustado = 0,51300086
F(7,9)= 3,4078 p< 0,04552 Erro padrão da estimativa: 205, 66

	BETA	Err. padr. de BETA	B	Err. padr. de B	t(9)	nível-p
Intercepto			101,8375	86,33178	1,179606	0,268397
OCMONO	0,413563	0,364274	0,141786	0,124888	1,135307	0,285576
OCSELA	0,217188	0,411263	0,129385	0,245001	0,528099	0,610203
OCDITE	0,389113	0,366155	0,252359	0,237469	1,0627	0,3156
OCTRIT	-0,19081	0,220972	-1,28259	1,485336	-0,8635	0,410282
OCUMA	0,098524	0,31474	0,415764	1,328172	0,313035	0,761387
OCPOLI	0,263482	0,275846	0,17603	0,18429	0,955179	0,364438
OCBENZ	-0,12176	0,400853	-0,52706	1,735154	-0,30375	0,768221

Comentário: Na regressão linear múltipla padrão, nenhum valor de coeficiente de regressão BETA foi tido como marcante.

Figura 51. Diagrama de dispersão para valores previstos vs. observados de ocorrências de flavonóides em tribos de Asteraceae (valores previstos obtidos a partir da regressão linear múltipla tomando OCFLAV como variável dependente).

Comentário: O diagrama de dispersão dos valores previstos vs. observados das ocorrências mostra uma certa tendência de surgimento de *outliers* (por exemplo, Helianeae tem alta produção / acúmulo de flavonóides, entretanto a previsão de ocorrências destes na tribo fica abaixo de 50% do observado; o mesmo fato ocorre na tribo Lactuceae, e o inverso, na tribo Inuleae). Aglomeração ocorre entre as tribos de baixa e nenhuma incidência dos referidos metabólitos: Arctoteae, Liabeae, Calenduleae e Plucheeae. A seguir, é apresentada a equação 8 empírica, gerada do sumário da regressão linear múltipla (Tabela 34), para previsão de ocorrências de flavonóides em tribos de Asteraceae

Equação 8: $OC_{FLAV} = 101,84 + 0,099 OC_{CUMA} + 0,41 OC_{MONO} + 0,39 OC_{DITE} + 0,22 OC_{SELA} - 0,19 OC_{TRIT} + 0,26 OC_{POLI} - 0,12 OC_{BENZ}$

Comentário: Os valores de BETA relacionados às ocorrências de monoterpenos e diterpenos obtidos na regressão são muito próximos (0,41 e 0,39, respectivamente) e os que mais influenciam positivamente a previsão de ocorrências de flavonóides em nível de tribos em Asteraceae. Utilizando regressões lineares múltiplas por passo, pode-se chegar às equações 9 e 10 descritas abaixo:

Passo a passo adiante / Passo 3 como solução final:

Equação 9: $OC_{FLAV} = 98,1580 + 0,60 OC_{MONO} + 0,283 OC_{DITE} + 0,36 OC_{POLI}$

Onde $R = 0,8374$; $R^2 = 0,7013$; $F = 10,1737$; $df = (3,13)$; $p = 0,001015$

Passo a passo para trás / Passo 5 como solução final:

Equação 10: $OC_{FLAV} = 124,4420 + 0,60 OC_{MONO} + 0,51 OC_{POLI}$

Onde $R = 0,8020$; $R^2 = 0,64317$; $F = 12,6164$; $df = (2,14)$; $p = 0,000737$

As regressões passo a passo para trás apresentam uma melhor opção para estimar ocorrências de flavonóides em função das ocorrências das demais classes químicas apresentadas neste estudo, uma vez , no passo 5, o R de correlação múltipla tem valor considerado bom e os coeficientes de regressão BETA de OCMONO e OCPOLI são significantes em um nível-p muito abaixo de 0,05.

Sumário de regressão para variável dependente: OCFLAV

$R = 0,63758895$ $R^2 = 0,40651967$ $R^2_{ajustado} = 0,35523125$

$F(7,81) = 7,9261$ $p < 0,00000$ Erro padrão da estimativa: 75,358

	BETA	Err. padr. de BETA	B	Err. padr. de B	t(9)	nível-p
Intercepto			23,32402	9,851306	2,367607	0,020288
OCMONO	0,33315713	0,13713675	0,126172	0,051936	2,429379	0,0173371
OCSELA	-0,02081981	0,10316888	-0,01147	0,056823	-0,2018	0,8405765
OCDITE	0,47328526	0,10923039	0,333367	0,076938	4,332908	4,19E-05
OCTRIT	0,03042064	0,12790051	0,142054	0,597251	0,237846	0,8126011
OCUMA	0,03189507	0,15371026	0,102331	0,49316	0,207501	0,8361393
OCPOLI	0,23381721	0,08746653	0,198864	0,074391	2,673219	0,009083
OCBENZ	-0,16812226	0,11043157	-0,73114	0,480253	-1,52241	0,1318007

Tabela 35. Valores previstos e observados de ocorrências de flavonóides em subtribos de Asteraceae (valores previstos obtidos a partir da regressão linear múltipla tomando OCFLAV como variável dependente).

Valores previstos e valores observados de ocorrências de flavonóides em subtribos de Asteraceae

TRIBO	SUBTRIBO	Previsto	Observado	TRIBO	SUBTRIBO	Previsto	Observado
2 ANT	ACH	190,112106	413	45 GNA	REL	43,992794	0
3 ANT	ANT	32,3651009	252	46 GNA	UNA4	27,7788715	0
4 ANT	ART	282,552704	137	47 HLN	BAE	26,4642792	40
5 ANT	CHR	42,3247719	8	48 HLN	CHA	63,3448067	195
6 ANT	GON	23,5181255	2	49 HLN	FLA	25,6518136	2
7 ANT	HAN	23,2666893	0	50 HLN	GAI	28,4543324	40
8 ANT	LEU	24,0969028	3	51 HLN	HYM	27,6875706	1
9 ANT	MAT	34,8094521	13	52 HLN	MAD	34,7776756	235
10 ANT	TAN	57,1516228	20	53 HLN	PEC	49,2218132	108
11 ANT	THA	23,0603142	2	54 HLN	PER	24,1728249	44
12 ANT	URS	23,9751968	2	55 HLN	UNA5	38,3647537	0
13 ARC	ARC	30,9396954	0	56 HLT	AMB	49,5989609	34
14 ARC	GRT	27,3765144	0	57 HLT	COR	188,487442	190
15 ARC	UNA1	28,1556282	0	58 HLT	ENG	25,181612	0
16 AST	AST	362,338745	361	59 HLT	GAL	33,2388536	0
17 AST	GRA	31,0181007	1	60 HLT	HLT	194,567108	501
18 AST	SOL	267,347412	92	61 HLT	MEL	75,0521317	57
19 BAR	UNA2	25,9118729	139	62 HLT	RUD	53,0086937	10
20 CAL	GRA	23,3010902	0	63 HLT	UNA6	54,1383743	1
21 CAL	UNA3	54,0825081	1	64 HLT	VER	147,71196	68
22 CAR	CAD	105,421562	68	65 HLT	ZIN	48,190052	0
23 CAR	CAR	43,6419411	39	66 INU	UNA7	76,5691605	109
24 CAR	CEN	87,3161011	115	67 LAC	CRE	23,8349895	10
25 CAR	ECH	71,9947281	0	68 LAC	HIE	23,8677442	23
26 EUP	ADE	34,5125427	0	69 LAC	HYP	23,6127472	26
27 EUP	AGE	79,394722	30	70 LAC	LAC	23,7847538	227
28 EUP	ALO	51,1374321	70	71 LAC	MIC	23,3240242	1
29 EUP	AYA	6,6650176	0	72 LAC	SCO	22,4210758	7
30 EUP	CRI	27,6271782	4	73 LAC	SEN	23,4263554	19
31 EUP	DIS	35,7547112	2	74 LAC	SON	28,8411674	56
32 EUP	EUP	44,8925507	154	75 LAC	STE	23,6310177	56
33 EUP	FLE	28,5513058	0	76 LAC	UNA8	24,4436035	8
34 EUP	GYP	26,3090496	2	77 LIA	LIA	24,1439247	8
35 EUP	HEB	26,4249783	0	78 MUT	MUT	53,0856972	80
36 EUP	HOF	19,141716	0	79 MUT	NAS	32,5952187	5
37 EUP	LIA	13,8084249	1	80 PLU	UNA9	37,2869568	11
38 EUP	MIK	70,1946335	2	81 SEN	BLE	18,1416149	0
39 EUP	OXY	41,6627045	14	82 SEN	SEN	19,2398529	11
40 EUP	PRA	28,942297	33	83 SEN	TUS	-10,3914518	7
41 GNA	ANG	55,5081024	15	84 VER	CEN	26,4140625	9
42 GNA	CAS	39,9338837	109	85 VER	ELE	23,8263092	0
43 GNA	GNA	111,606544	224	86 VER	LYC	27,0036172	9
44 GNA	LOR	27,2495384	4	87 VER	PIP	36,1843491	0
				88 VER	ROL	23,3673229	0
				89 VER	VER	48,7311287	22

Figura 52. Diagrama de dispersão para valores previstos vs. observados de ocorrências de flavonóides em subtribos de Asteraceae (valores previstos obtidos a partir da regressão linear múltipla tomando OCFLAV como variável dependente).

Comentário : Similar ao que ocorre em ocorrências de flavonóides em tribos de Asteraceae, o diagrama de dispersão dos valores previstos *versus* observados das ocorrências mostra uma tendência mais acentuada de surgimento de *outliers* (por exemplo, Helianthinae — HLT-HLT — tem alta produção / acúmulo de flavonóides, entretanto a previsão de ocorrências destes na subtribo fica abaixo de 40% do observado; o inverso ocorre na subtribo Carduinae e em muitas outras). Aglomerações ocorrem entre subtribos de baixa e nenhuma incidência dos referidos metabólitos com valores observados de 0 a ca. 60 e apresentando quase sempre valores previstos elevados. Umas poucas subtribos se ajustam sobre a reta de regressão (por exemplo, AST-AST, HLT-COR, etc.).

A seguir são apresentadas as equações empíricas, geradas do sumário da regressão linear múltipla (Tabela 36), para previsão de ocorrências de flavonóides em tribos de Asteraceae.

Tabela 36. Sumário da regressão linear múltipla, para previsão de ocorrências de flavonóides em tribos de Asteraceae.

Sumário de regressão para variável dependente: OCFLAV						
R = 0,63758895 R ² = 0,40651967 R ² ajustado = 0,35523125						
F (7,81) = 7,9261 p< 0,00000 Erro padrão da estimativa: 75,358						
	BETA	Err. padr. de BETA	B	Err. padr. de B	t(9)	nível-p
Intercepto			23,32402	9,851306	2,367607	0,020288
OCMONO	0,33315713	0,13713675	0,126172	0,051936	2,429379	0,0173371
OCSELA	-0,02081981	0,10316888	-0,01147	0,056823	-0,2018	0,8405765
OCDITE	0,47328526	0,10923039	0,333367	0,076938	4,332908	4,19E-05
OCTRIT	0,03042064	0,12790051	0,142054	0,597251	0,237846	0,8126011
OCUMA	0,03189507	0,15371026	0,102331	0,49316	0,207501	0,8361393
OCPOLI	0,23381721	0,08746653	0,198864	0,074391	2,673219	0,009083
OCBENZ	-0,16812226	0,11043157	-0,73114	0,480253	-1,52241	0,1318007

Equação 10: $OC_{FLAV} = 23,3240 + 0,030OC_{CUMA} + 0,33 OC_{MONO} + 0,39 OC_{DITE} - 0,021 OC_{SELA} + 0,030 OC_{TRIT} + 0,23 OC_{POLI} - 0,17 OC_{BENZ}$

Passo a passo adiante / Passo 4 como solução final:

Equação 11: $OC_{FLAV} = 23,67025 + 0,35OC_{MONO} + 0,48 OC_{DITE} + 0,23 OC_{POLI} - 0,16 OC_{BENZ}$

Onde R = 0,6358; R² = 0,4043; F = 14,25195; df = (4,84); p = 0,000000

Passo a passo para trás / Passo 5 como solução final:

Equação 12: $OC_{FLAV} = 27,7966 + 0,35OC_{MONO} + 0,44 OC_{POLI}$

Onde R = 0,5769; R² = 0,3328; F = 21,44661; df = (2,86); p = 0,000000

Comentário:

As regressões padrão e passo a passo adiante parecem apresentar duas boas opções para estimar ocorrências de flavonóides em função das ocorrências das demais classes químicas apresentadas neste estudo, em virtude de os valores do coeficiente de correlação explicado (= R²) de ambas serem muito próximos (0,4065 e 0,4043, respectivamente).

Tabela 37. Tipos de flavonóides encontrados em Asteraceae

TRIBO	FLAVONOL	FLAVONA	FLAVANONA	CHALCONA	AURONA	DIDROFLAVONOL	DIDROCHALCONA
ANT	438	387	19	1	0	4	0
AST	316	97	21	0	0	1	0
BAR	122	0	1	0	2	0	0
CAL	1	0	0	0	0	0	0
CAR	50	161	9	0	0	0	0
EUP	149	121	26	6	0	4	0
GNA	226	64	30	18	1	2	10
HLN	552	209	61	14	16	7	0
HLT	392	181	52	138	87	1	0
INU	79	10	11	0	0	7	0
LAC	49	343	11	1	0	1	0
LIA	8	0	0	0	0	3	0
MUT	31	29	19	3	0	3	0
PLU	8	1	7	0	0	0	0
SEN	12	1	3	3	0	1	0
VER	5	22	2	0	0	1	0

Comentário: Na tribo Arctoteae (ARC), não se constatou até o presente nenhuma ocorrência de flavonóides. Em Calenduleae, só se registrou uma ocorrência de flavonol. Os flavonóis estão presentes em praticamente todas as tribos de Asteraceae, sendo mais expressivos, em ordem decrescente, em: Helenieae, Anthemideae, Heliantheae, Astereae e Gnaphalieae. Esta última apresenta uma maior diversidade de flavonóides (todos os tipos apresentados na tabela).

Figura 53. Diagrama tridimensional de distribuição de tipos de flavonóides em tribos de Asteraceae

Tabela 38. Ocorrências de flavonóis e flavanonas em tribos de Asteraceae

TRIBO	Flavonóis	Flavanonas
	X	Y
ANT	438	19
AST	316	21
BAR	122	1
CAL	1	0
CAR	50	9
EUP	149	26
GNA	226	30
HLN	552	61
HLT	392	52
INU	79	11
LAC	49	11
LIA	8	0
MUT	31	19
PLU	8	7
SEN	12	3
VER	5	2

Comentário: Foram escolhidas as duas subclasses de flavonóides nesta tabela, em virtude de apresentarem uma melhor correlação nos testes de regressão linear múltipla, tomando ocorrências de flavonóis como variável dependente, onde $R=0,89924190$; $F=6,338465$; $R^2=0,80863600$; $p=0,0007468$; valores de **BETA**= 0,303, flavonas; **0,574**, flavanonas; -0,20, chalconas; 0,209, auronas; 0,054, diidroflavonóis; 0, 273, diidrochalconas. A regressão linear múltipla por passo a passo adiante apresentou, no segundo passo, um valor significativo do coeficiente BETA das flavanonas (= **0,721**) e $R=0,8770036$.

Figura 54. Diagrama de dispersão de ocorrências de flavanonas em função das ocorrências de flavonóis, a partir da correlação bivariada, em tribos de Asteraceae.

Comentário: As ocorrências de flavonóis e as de flavanonas em tribos de Asteraceae apresentam uma correlação muito boa, coeficiente de correlação r = ca. 0,84.

Figura 55. Diagrama de dispersão de valores previstos *versus* observados de ocorrências de cumarinas em função de ocorrências de 7 tipos de produtos naturais em subtribos de Asteraceae, após remoção de 5 outliers

Figura 56. Previsão de ocorrências de monoterpenos em função de ocorrências de cumarinas em subtribos de *Asteraceae*.

Comentário: As ocorrências de cumarinas e as de monoterpenos em subtribos de *Asteraceae* estão fortemente correlacionadas ($r = \text{ca. } 0,87$). No próprio diagrama (Fig. 56) encontra-se a equação empírica para previsão de ocorrências de as.

Oxidações

Números de oxidação (NOXs) médios de diversas classes químicas em tribos de Asteraceae

Figura 57. Diagrama tridimensional mostrando os valores médios de NOXs de diversas classes químicas em tribos de Asteraceae.

Comentário: Os poliacetilenos e os sesquiterpenóides (principalmente os sesquiterpenos lactonizados) atingem índices de oxidação mais elevados em Asteraceae: os primeiros, valores máximos ca. + 12 ; os segundos, ca. + 7.

Cumarinas

Nas correlações de NOXs médios das 7 classes de metabólitos estudadas, os resultados, em nível de tribos, mostram que os NOXs das cumarinas estão forte e negativamente correlacionados com os dos flavonóides oxidados no anel B (FLAB, $r = -0,63$). São também interessantes os valores de r nas correlações com sesquiterpenos / lactonas sesquiterpênicas e poliacetilenos (0,48 e 0,51, respectivamente). Os FLABs, além de terem seus NOXs bem correlacionados com os das cumarinas, apresentam valores de r interessantes nas correlações entre os NOXs de sesquiterpenos / lactonas sesquiterpênicas e triterpenos: $-0,55$ e 0,51,

respectivamente. Quanto aos flavonóides oxidados no anel A (FLAV-A), seus NOXs só apresentam correlação razoável com os dos monoterpenos.

Tabela 39. Sumário de regressão linear múltipla de NOXs de cumarinas em função de NOXs de diversas classes químicas em tribos de Asteraceae

Sumário de regressão para variável dependente: OXCUMA

R = 0,84257200 R²=0,70992757 R²ajustado = 0,48431568

F(7,9) = 3,1467 p < 0,05647 Erro da estimativa: 0,76096

	Err. padr.		Err. padr.			
	BETA	de BETA	B	de B	t(9)	nível-p
Intercepto			1,27753	0,965354	1,32338	0,21834
OXMONO	-0,2652649	0,236514	-0,4517	0,402742	-1,12156	0,29108
OXSELA	0,3617857	0,227042	0,177773	0,111563	1,593476	0,145518
OXDITE	-0,4956048	0,224266	-0,59332	0,268481	-2,2099	0,054454
OXTRIT	-0,0706944	0,262072	-0,0878	0,325496	-0,26975	0,793434
OXPOLI	0,7612152	0,283958	0,175271	0,065382	2,680728	0,025176
OXFLAV-A	-0,2443824	0,251633	-0,42137	0,433873	-0,97119	0,356831
OXFLAV-B	-0,131829	0,327873	-0,14373	0,357474	-0,40207	0,697012

Comentário: Na RLM padrão, os NOXs de cumarinas e poliacetilenos apresentam uma forte correlação com R = 0,84 e BETA_{OXPOLI} = 0,76.

Equação empírica, gerada do sumário da regressão linear múltipla padrão (Tabela 38), para previsão de NOXs médios de cumarinas em tribos de Asteraceae:

Equação 12: OXCUMA = 1,2775 – 0,27 OXMONO – 0,50 OXDITE + 0,36 OXSELA – 0,07 OXTRIT + 0,76 OXPOLI – 0,24 OXFLAV A – 0,13 OXFLAV-B

OBS.: Valor em vermelho signficante a um p< 0,05.

Tabela 40. Valores previstos e observados de NOXs de cumarinas em tribos de Asteraceae

Tribo	Previsto	Observado
1 BAR	0,22253685	0
2 MUT	2,91014266	2,1289999
3 CAR	2,30177236	1,8329999
4 LAC	1,40748477	1,918
5 VER	3,92648482	4,7119999
6 LIA	0,29374802	0
7 ARC	2,34908843	2
8 INU	1,50366676	2,214
9 PLU	2,3368392	2,724
10 GNA	1,26288021	2,2750001
11 CAL	1,53188705	1,6670001
12 AST	1,24432921	1,232
13 ANT	1,84216428	2,0369999
14 SEN	2,3746295	1,46
15 HLN	2,02219892	1,643
16 HLT	1,90846682	2,263
17 EUP	1,54667997	0,878

Comentário: Valores previstos obtidos a partir da regressão linear múltipla padrão entre NOXs de cumarinas e de diversas classes químicas.

Figura 58. Diagrama de dispersão de valores previstos vs. observados de NOXs de cumarinas mostrados na Tabela 37

Comentário: O diagrama de dispersão acima mostra uma tendência de relação inversa (negativa) entre os NOXs de cumarinas observados vs. previstos nas tribos de Asteraceae: a maioria dos pontos cai abaixo da reta de regressão. Astereae

(AST) é a tribo que possui os NOXs observados e previstos mais ajustados à reta.; Calenduleae e Anthemideae têm os pontos correspondentes razoavelmente representados ao longo da reta.

Tabela 40. Regressão seqüencial (passo a passo adiante) de oxidação de cumarinas em função de oxidação de diversas classes de metabólitos secundários em tribos de Asteraceae

Passo 1 - Sumário de regressão para variável dependente: OXCUMA						
R=0,63100600 R²=0,39816857 R² ajustado = 0,35804648						
F(1,15)=9,9239 p<0,00660 Err.. padr.da estimativa: 0,84903						
	BETA	Erro padrão de BETA	B	Erro padrão de B	T(15)	Nível - P
Intercepto			2,738866	0,356359	7,685689	1,41E-06
OXFLAB	-0,63101	0,200305	-0,687970	0,218389	-3,150230	0,006603
Passo 2 - Sumário de regressão para variável dependente: OXCUMA						
R=0,69965801 R²=0,48952133 R² ajustado = 0,41659581						
F(2,14)=6,7126 p<0,00903 Err.. padr.da estimativa: 0,80938						
	BETA	Erro padrão de BETA	B	Erro padrão de B	T(10)	Nível - P
Intercepto			1,838475	0,662568	2,774774	0,014902
OXFLAB	-0,50744	0,206293	-0,553260	0,224917	-2,459830	0,027518
OXPOLI	0,326528	0,206293	0,075184	0,047499	1,582836	0,135781
Passo 3 - Sumário de regressão para variável dependente: OXCUMA						
R=0,75025037 R²=0,56287561 R² ajustado = 0,46200075						
F(3,13)=5,5799 p<0,01104 Err.. padr.da estimativa: 0,77725						
	BETA	Erro padrão de BETA	B	Erro padrão de B	T(10)	Nível - P
Intercepto			1,888063	0,637148	2,963306	0,010987
OXFLAB	-0,49999	0,198167	-0,545130	0,216057	-2,523070	0,025456
OXPOLI	0,466657	0,219649	0,107449	0,050575	2,124558	0,053376
OXDITE	-0,30374	0,205643	-0,363620	0,246187	-1,477000	0,163486

Tabela 40. Continuação

Passo 4 - Sumário de regressão para variável dependente: OXCUMA						
R= 0,78062491 R²=0,60937524 R² ajustado = 0,47816699						
F (4,12)= 4,6800 p< 0,01654 Err.. padr.da estimativa: 0,76475						
	BETA	Erro padrão de BETA	B	Erro padrão de B	T(10)	Nível - P
Intercepto			2,235940	0,691175	3,234985	0,007153
OXFLAB	-0,40628	0,210153	-0,442960	0,229126	-1,933270	0,077142
OXPOLI	0,601836	0,243923	0,138574	0,058164	2,467317	0,029639
OXDITE	-0,39073	0,215030	-0,467770	0,257424	-1,817100	0,094243
OXFLAA	-0,25169	0,210588	-0,433980	0,363104	-1,195190	0,255101

Passo 5 - Sumário de regressão para variável dependente: OXCUMA						
R= 0,81812831 R²= 0,66933393 R² ajustado = 0,51903117						
F (5,11)= 4,4532 p< 0,01832 Err.. padr. da estimativa: 0,73490						
	BETA	Erro padrão de BETA	B	Erro padrão de B	T(10)	Nível - P
Intercepto			1,342566	0,917222	1,463730	0,171250
OXFLAB	-0,21249	0,244158	-0,231670	0,266201	-0,870280	0,402740
OXPOLI	0,65538	0,237449	0,150903	0,054673	2,760094	0,018553
OXDITE	-0,42602	0,208142	-0,510010	0,249178	-2,046760	0,065339
OXFLAA	-0,31479	0,207242	-0,542780	0,357334	-1,518960	0,156979
OXSELA	0,30061	0,212853	0,147714	0,104591	1,412302	0,185522

Passo 6 - Sumário de regressão para variável dependente: OXCUMA						
R= 0,84117811 R² = 0,70758229 R² ajustado = 0,53213167						
F (6,10)= 4,0329 p< 0,02576 Err.. padr. da estimativa: 0,72482						
	BETA	Erro padrão de BETA	B	Erro padrão de B	T(10)	Nível - P
Intercepto			1,238639	0,909197	1,362343	0,202984
OXFLAB	-0,18781	0,241775	-0,204770	0,263602	-0,776810	0,455254
OXPOLI	0,727282	0,242484	0,167458	0,055832	2,999295	0,013360
OXDITE	-0,49142	0,213105	-0,588310	0,255120	-2,306020	0,043803
OXFLAA	-0,21822	0,221157	-0,376250	0,381326	-0,986700	0,347062
OXSELA	0,35813	0,215874	0,175976	0,106075	1,658975	0,128110
OXMONO	-0,24774	0,216619	-0,421860	0,368866	-1,143680	0,279395

Comentário : Na RLM passo a passo adiante mostrada na Tabela 40, já no passo 1, a variável independente de entrada (OXFLAV-B) apresenta um valor do coeficiente BETA negativo, porém significativo (ca. -0,63).

A equação correspondente a esta correlação bivariada (Equação 13) pode ser vista conjuntamente com o diagrama de dispersão na Figura 59. OXPOLI é a segunda variável independente de entrada, porém seu coeficiente BETA só se torna significativo no passo 5 (ca. 0,66) com coeficiente de explicação $R^2 = \text{ca. } 0,67$, $F(5,11) = 4,4532$, $p < 0,01832$. No passo 6, chega-se à solução final com valores significantes dos coeficientes $BETA_{OXPOLI} = \text{ca. } 0,73$ e $BETA_{OXDITE} = \text{ca. } -0,49$, $R = 0,84$, $R^2 = 0,70758229$, $F(6,10) = 4,0329$, $p < 0,02576$. A equação obtida para previsão de NOXs de cumarinas em tribos de Asteraceae, a partir do passo 6 da RLM, é:

(OBS.: Valores em vermelho são significantes a um $p < 0,05$)

Equação 13: $OX_{CUMA} = 1,2386 - 0,23 OX_{MONO} - 0,49 OX_{DITE} + 0,36 OX_{SELA} + 0,73 OX_{POLI} - 0,22 OX_{FLAV A} - 0,19 OX_{FLAV-B}$

OXIDAÇÃO DE FLAVONÓIDES (ANEL B) VS. OXIDAÇÃO DE CUMARINAS EM TRIBOS DE ASTERACEAE

OXCUMA	OXFLAV-B	TRIBO
0.000000	2.500000	ANT
2.129000	0.000000	ARC
1.833000	0.640000	AST
1.918000	0.840000	BAR
4.712000	0.410000	CAL
0.000000	3.180000	CAR
2.000000	0.180000	EUP
2.214000	1.780000	GNA
2.724000	0.620000	HLN
2.275000	0.800000	HLT
1.667000	3.180000	INU
1.232000	2.120000	LAC
2.037000	1.430000	LIA
1.460000	0.680000	MUT
1.643000	1.640000	PLU
2.263000	1.090000	SEN
0.878000	1.550000	VER

OXIDAÇÃO DE FLAV-B vs. OXIDAÇÃO DE CUMARINAS EM TRIBOS DE ASTERACEAE

$OXCUMA = 2,7389 - 0,6880 \cdot OXFLAV-B$

Correlação: $r = -0,6310$

Flavonóides constituem um das classes mais características de compostos em plantas superiores. Muitos flavonóides são facilmente reconhecidos como pigmentos de flores na maioria das famílias de angiospermas (plantas que florescem). Porém, a ocorrência deles não é restrita a flores, mas inclui também todas as partes da planta. A estrutura química de flavonoides está baseada em um esqueleto de C₁₅ com um anel CROMANO que sustenta um segundo anel aromático B na posição 2, 3 ou 4.

Figura 59. Diagrama de dispersão de NOXs de flavonóides oxidados no anel B e NOXs de cumarinas, obtido a partir da regressão linear entre seus respectivos NOXs representados na tabela anexa ao diagrama.

Comentário: Nas correlações de NOXs médios das 7 classes de metabólitos estudadas em nível de tribos, os resultados mostram que as cumarinas estão forte e negativamente correlacionadas com os flavonóides oxidados no anel B (FLAV-B, $r=-0,63$).

Figura 60. Diagrama tridimensional mostrando NOXs de cumarinas (OXCUMA, x), NOXs de flavonóides oxidados no anel B (OXFLAV-B, y) e número de observações dos referidos NOXs (z).

Comentário : O histograma acima mostra um certo equilíbrio em número de observações de Noxs médios de 0 a +1 para as cumarinas e os flavonóides oxidados no anel B. Numa visão geral, as cumarinas e os flavonóides oxidados no anel B atingem Noxs médios em sua maioria de 0 a +4 .

Os NOXs dos FLAV-Bs, além de correlacionarem relativamente bem com os das cumarinas, apresentam valores de r interessantes nas correlações entre NOXs de sesquiterpenos / lactonas sesquiterpênicas, triterpenos e poliacetilenos: 0,48, **-0,55** (significante) e 0,51, respectivamente.

Quanto aos flavonóides oxidados no anel A (FLAV-A), seus NOXs só apresentam correlação razoável com os dos monoterpenos.

Quanto às correlações de NOXs nas 87 subtribos, os resultados não foram, de modo geral, animadores. Isto se deve ao fato de que o NOX não é um bom critério

para correlações, quando se cai de nível hierárquico (Emerenciano *et al.*, 2003). Todavia, a partir de critérios de avaliação de ocorrências dos referidos metabólitos, muitas vezes insignificantes em algumas subtribos, e com base em retas de regressão, foram descartadas 15 subtribos como *outliers*. As 72 subtribos restantes (Tabela 41), supostamente de maior representatividade de ocorrências na família, foram selecionadas para novos testes de matriz de correlações de NOXs e outros estatísticos. Na nova matriz, as cumarinas não apresentaram nenhum valor de r expressivo nas correlações de seus NOXs com os dos outros metabólitos testados. O melhor valor dentre os coeficientes de correlações está entre os flavonóides A e B ($r = 0,49$), Figura 61. Vale a pena ressaltar que FLAV-Bs apresentam correlações de NOXs baixas, porém significativas, com os dos monoterpenos e poliacetilenos ($r = 0,32$ e $r = 0,28$, respectivamente); similarmente, FLAV-Bs com os dos diterpenos ($r = 0,30$) e poliacetilenos ($r = 0,30$).

Correlações entre NOXs de flavonóides mais glicosilados e os de monoterpenos são explicáveis, pois monoterpenos de Asteraceae não se oxidaram ao ponto dos iridóides, o que ocorreu com os monoterpenos de Lamiales e ordens afins.

Os parâmetros evolutivos de avanço relacionados a EA_M e EA_G apresentando um coeficiente de correlação múltipla inversa ($R = -0,77$) bate com as propostas de Harborne.

Tabela 41. Valores de NOXs médios de diversas classes químicas em subtribos de Asteraceae

TRIBO	SUBTRIBO	OXMONO	OXSELA	OXDITE	OXTRIT	OXCUMA	OXPOLI	OXFLAV-A	OXFLAV-B
1 ANT	ACH	0,512	6,404	1	1	2,969	11,526	2,549	1,298
2 EUP	ADE	0	0	4,429	0	0	12	0	0
3 EUP	AGE	0,775	6,757	1,418	0,846	0,4	12	3,778	2,519
4 EUP	ALO	5	5,125	1,947	0,727	0	0	2,529	1,918
5 HLT	AMB	0,231	6,167	1,333	4,5	0	11,78	2,824	1,647
6 GNA	ANG	0	3,95	1,81	0	3,333	11,25	2,2	-0,867
7 ANT	ANT	0,912	9,553	0	0	2,286	13,826	3	3
8 ARC	ARC	0	6,4	1	1,125	0	11,273	0	0
9 ANT	ART	0,471	6,481	1,667	0,667	1,994	0	2,141	1,163
10 AST	AST	0,333	3,04	1,949	2,58	0,354	11,491	1,91	1,194
11 HLN	BAE	3,429	3,826	0	0	0	11,895	0,45	0,8
12 HLT	BAH	3,381	7,25	1,5	0,545	1,5	9,952	1,9	1,1
13 CAR	CAD	0,5	7,278	0	0,643	1,5	9,525	1,571	-0,119
14 CAR	CAR	0,667	3,5	2	0	0	11,228	3	1
15 GNA	CAS	0,2	3,519	1,333	4	3,091	11,087	0,828	1,736
16 CAR	CEN	1	6,945	-5	0,846	1,556	9,13	1,933	0,822
17 HLN	CHA	2,515	7,067	0,816	2	0	10,769	1,844	1,26
18 ANT	CHR	0,698	5,82	0	0	1	0	2,333	1,667
19 HLT	COR	1,5	7,605	0	6	2	10,591	0,579	0,158
20 LAC	CRE	0	6,477	0	1	2	0	1	1
21 EUP	CRI	2,2	6,147	-1,627	0,714	0	0	1	-1,5
22 EUP	DIS	0	7,522	1,8	0,5	0	0	1	1
23 HLT	ECL	0,413	8,246	2,438	1,143	5	12,376	1	3
24 VER	ELE	0	8,515	0	1	0	11,333	0	0
25 HLT	ENG	0	6	-0,161	3,4	0	8,167	0	0
26 EUP	EUP	0,869	6,952	1,538	0,818	-0,312	12	2,573	0,84
27 HLT	FIT	0	5,286	0	0	0	12	0	0
28 EUP	FLE	0	2,4	0,933	0,5	0	0	0	0
29 HLT	GAI	2,138	6,998	0	0	-2	12,131	3,231	0,5
30 HLT	GAL	0,931	4,111	-1	1	0	7,2	0	0
31 GNA	GNA	0,071	2,586	0,658	1,667	2,428	11,821	2,094	-0,853
32 ANT	GON	3	8,667	0	0	2	0	1	-3
33 ARC	GOR	0	6	0	1,037	2	0	0	0
34 CAL	GRA	1	6	4,708	1	0	0	1	3
35 EUP	GYP	2	7,889	-0,25	0,611	0	0	2,2	1,8
36 ANT	HAN	0	6,4	0	0	0	0	0	0
37 EUP	HEB	5	5,177	2,7	1	0	0	2,667	1,667
38 HLT	HEL	0,259	7,71	2,401	2,889	2,125	8,692	1,497	0,588
39 LAC	HIE	0	9,143	0	0	0	0	1	0,613
40 HLT	HLT	0	1	0	0	0	0	0	0
41 LAC	HYP	0,4	7,344	0	1	0	0	1	1,308
42 INU	INU	0	0	0	0	0	0	1	0
43 LAC	LAC	0	9,183	0	1	3	0	1,033	0,642
44 EUP	LIA	0	7,245	0	0,857	-1,429	12,6	2,5	1,5
45 GNA	LOR	5,211	-0,25	2,614	2,778	1,333	12	2,5	1
46 VER	LYC	0	8,06	-1	0,939	0	0	1,333	-0,167
47 HLT	MAD	0	0	-0,053	0	2	11,7	1,902	1,251
48 ANT	MAT	0,491	6,28	0	0	1,588	12,381	2,143	2,286
49 HLT	MEL	1,552	9,308	1,545	0	2	11,292	1,722	2,056
50 LAC	MIC	0	0	0	0	0	0	-1	-1
51 EUP	MIK	1,4	8,92	1,011	1,5	0,667	12,25	3	1
52 HLT	MIL	0,672	8,536	1,986	1	0	10,176	1,431	2,234
53 MUT	MUT	1,052	5,982	1,111	0,932	2,019	11,4	1,25	-0,194
54 MUT	NAS	4	4,683	0	0,632	2,656	0	1,571	-0,286
55 HLT	NEU	3,909	6,877	1,294	0	0	6,125	2,692	1,923
56 EUP	OXY	4,269	4,564	0,861	0,5	2	0	2,6	1,8
57 HLN	PEC	0,741	0	0	0	0,571	13	1,593	2,944
58 HLN	PER	4	5,667	0	0	0	0	2,516	1,613
59 VER	PIP	0	8,938	0	1	0	0	0	0
60 EUP	PRA	0,429	2	2,769	0	0	0	1,821	0,214
61 VER	ROL	0	7,2	0	1	0	0	0	0
62 SEN	SEN	0,429	4,61	0,234	1,73	1,889	12	1,032	0,556
63 AST	SOL	0,261	3,14	1,498	1,909	1,286	10,912	2,362	2,25
64 LAC	SON	1	7,143	0	0	2	0	1,047	0,488
65 LAC	STE	0	0	0	0	2	0	1	0,455
66 ANT	TAN	0,572	4,559	0	0,333	4	0	2,478	0,739
67 EUP	TRI	5	0	0	0	0	0	0	0
68 SEN	TUS	2,594	5,31	0,778	1,333	0,375	12	1,571	1,857
69 INU	UNA	2,012	4,425	0,393	1,703	2,932	11,667	1,613	1,978
70 ANT	URS	1,6	5,212	0	0	-1,333	0	1	-3
71 HLT	VER	0,324	8,234	1,387	0,847	7,424	11,43	1,385	0
72 HLT	ZIN	0	6,975	-0,867	0	0	10,045	0	0

Tabela 42. Sumário de regressão linear para a variável dependente OXFLAV-B, correlacionada com a variável FLAV-A, em 72 subtribos de Asteraceae

Sumário de regressão para variável dependente: OXFLAV-B						
R = 0,49065476 R² = 0,24074210 R² ajustado = 0,22989556						
F(1,70) = 22,195 p< 0,00001 Erro padrão da estimativa: 1,0434						
	Err.padr.		Err. padr			
	BETA	de BETA	B	de B	t(70)	nível-p
Intercepto			-0,04711	0,210589	-0,22373	0,8236227
OXFLAAB	0,490655	0,104147	0,564513	0,119824	4,711187	1,2116E-05

Tabela 43. Valores de NOXs médios de flavonóides A e B calculados em 72 subtribos de Asteraceae

TRIBO	SUBTRIBO	FLAV-B	FLAV-A	TRIBO	SUBTRIBO	FLAV-B	FLAV-A
1 ANT	ACH	1,298	2,549	37 EUP	HEB	1,667	2,667
2 EUP	ADE	0	0	38 HLT	HEL	0,588	1,497
3 EUP	AGE	2,519	3,778	39 LAC	HIE	0,613	1
4 EUP	ALO	1,918	2,529	40 HLT	HLT	0	0
5 HLT	AMB	1,647	2,824	41 LAC	HYP	1,308	1
6 GNA	ANG	-0,867	2,2	42 INU	INU	0	1
7 ANT	ANT	3	3	43 LAC	LAC	0,642	1,033
8 ARC	ARC	0	0	44 EUP	LIA	1,5	2,5
9 ANT	ART	1,163	2,141	45 GNA	LOR	1	2,5
10 AST	AST	1,194	1,91	46 VER	LYC	-0,167	1,333
11 HLN	BAE	0,8	0,45	47 HLT	MAD	1,251	1,902
12 HLT	BAH	1,1	1,9	48 ANT	MAT	2,286	2,143
13 CAR	CAD	-0,119	1,571	49 HLT	MEL	2,056	1,722
14 CAR	CAR	1	3	50 LAC	MIC	-1	-1
15 GNA	CAS	1,736	0,828	51 EUP	MIK	1	3
16 CAR	CEN	0,822	1,933	52 HLT	MIL	2,234	1,431
17 HLN	CHA	1,26	1,844	53 MUT	MUT	-0,194	1,25
18 ANT	CHR	1,667	2,333	54 MUT	NAS	-0,286	1,571
19 HLT	COR	0,158	0,579	55 HLT	NEU	1,923	2,692
20 LAC	CRE	1	1	56 EUP	OXY	1,8	2,6
21 EUP	CRI	-1,5	1	57 HLN	PEC	2,944	1,593
22 EUP	DIS	1	1	58 HLN	PER	1,613	2,516
23 HLT	ECL	3	1	59 VER	PIP	0	0
24 VER	ELE	0	0	60 EUP	PRA	0,214	1,821
25 HLT	ENG	0	0	61 VER	ROL	0	0
26 EUP	EUP	0,84	2,573	62 SEN	SEN	0,556	1,032
27 HLT	FIT	0	0	63 AST	SOL	2,25	2,362
28 EUP	FLE	0	0	64 LAC	SON	0,488	1,047
29 HLT	GAI	0,5	3,231	65 LAC	STE	0,455	1
30 HLT	GAL	0	0	66 ANT	TAN	0,739	2,478
31 GNA	GNA	-0,653	2,094	67 EUP	TRI	0	0
32 ANT	GON	-3	1	68 SEN	TUS	1,857	1,571
33 ARC	GOR	0	0	69 INU	UNA	1,978	1,613
34 CAL	GRA	3	1	70 ANT	URS	-3	1
35 EUP	GYP	1,8	2,2	71 HLT	VER	0	1,385
36 ANT	HAN	0	0	72 HLT	ZIN	0	0

Figura 61. Diagrama de dispersão bivariado de valores de NOXs de flavonóides A e B em subtribos de Asteraceae, obtidos a partir da regressão linear simples.

Comentário : Os pontos obtidos em torno da reta de regressão são bastante dispersos, mas , mesmo assim existe uma correlação positiva moderada ($r = 0,49$) entre os anéis A (derivado do acetato) e B (derivado do chiquimato). O ponto correspondente a ANT-MAT (subtribo Matricariinae) é o que melhor se ajusta à reta de regressão; os dois outros pontos com ajuste razoável à reta correspondem a SEN-SEN e LAC-SON, subtribos Senecioninae e Sonchinae, respectivamente. Os demais se dispersam bastante.

Comentário 2: Os resultados de estudos de NOXs conseguidos em nível de tribos e subtribos de Asteraceae neste trabalho merecem ainda um maior aprofundamento, especialmente com relação a precursores biogénéticos (no caso dos benzofuranos, é necessária uma reavaliação mais detalhada de seus precursores) . Entretanto sua utilidade pode ser avaliada em recente publicação em que participo como co-autor (Emereciano *et al.*, 2001). Nessa publicação, são mostrados padrões de substituição (metilação e glicosilação) para flavonas, flavonóis e flavanonas (nas posições 6; 6 e 8; 2',3',4'; 3',4',5' e 2',3',4',5',6', respectivamente) em tribos de Asteraceae. Padrões de oxidação da parte das rotas do acetato e do chiquimato foram também calculados separadamente para flavonóis, flavonas e flavononas e feitos estudos de índices evolutivos de avanço (EA) (Tabela 44) em tribos da família.

Em relação às subtribos, para cada uma foram também calculados parâmetros evolutivos de avanço relacionados a EA_M e EA_G (com grupos de proteção O-metil e O-glicosil, respectivamente). Das correlações entre estes parâmetros obteve-se um coeficiente de correlação múltipla inversa, $R = -0,77$. É interessante notar que este mesmo valor foi encontrado na correlação entre ambos os parâmetros químicos de evolução, em nível de tribos, trabalho publicado em 1987 (Emerenciano *et al.*, 1987). Isto significa uma corroboração da hipótese de Harborne, na qual as plantas preferem uma ou outra forma de proteção das hidroxilas.

Tabela 44. Parâmetros químicos para flavonóides de tribos de Asteraceae

TRIBO	EA_M	EA_G	EA_{AO}	EA_{OB}
ANT	0,37	0,046	2,4	1,25
ARC				
AST	0,379	0,022	2,14	1,88
BAR	0,0014	0,1893	1,34	2,32
CAL	0,2	0,2	1	3
CAR	0,168	0,142	1,8	0,46
EUP	0,506	0,026	2,6	1,37
GNA	0,21	0,071	1,52	0,46
HLN	0,196	0,068	1,91	1,46
HLT	0,14	0,113	1,36	0,91
INU	0,332	0,023	1,91	1,6
LAC	0,0038	0,214	1,06	0,64
LIA	0,092	0,05	1,5	3
MUT	0,207	0,047	1,31	-0,18
PLU	0,358	0	2	0,62
SEN	0,147	0,14	1,3	0,5
VER	0,147	0,0854	1,26	0,23

1-. EA_{OA} e EA_{OB} : índices evolutivos de avanço para táxons selecionados com respeito aos valores de oxidação dos anéis A e B dos flavonóides, respectivamente. 2-. EA_M e EA_G : índices evolutivos de avanço para táxons selecionados com respeito a O-metilflavonóides e O-glicosilflavonóides, respectivamente. 3- Para um glossário das subtribos, os acrônimos correspondentes de três letras estão na Tabela 6.

Marcadores quimiotaxonômicos em Asteraceae (resultados finais)

Flavonóides em nível de tribo e subtribo :

flavonas, **flavonóis** e **flavanonas** O-metilados ou O-glicosilados
nas posições **6**; **6** e **8**; **2'**, **3'**, **4'**; **3'**, **4'**, **5'** e **2'**, **3'**, **4'**, **5'**, **6'**, respectivamente

NOTA: Os flavonóides têm sido bastante explorados como marcadores químicos em angiospermas. Já no final da década de 70, Dahlgren (Gornall *et al.*, 1979) usou luteolina e apigenina como marcadores químicos em nível hierárquico elevado, abrangendo muitas ordens de plantas (Figura 62). Todas as famílias compreendidas dentro de Asterales (Asteraceae é uma delas) acumulam um dos ou ambos os metabólitos.

Comentário : Os tamanhos das áreas refletem o número de espécies nas ordens. As regiões pontilhadas destas representam a distribuição dos referidos metabólitos. O avanço da esquerda para a direita das ordens dá idéia de seu índice de evolução química: quanto mais à direita , mais evoluída é a ordem. É importante frisar que os desenhos dos diagramas de Dahlgren se baseiam em árvores supostamente cortadas em transversal e vistas de cima para baixo. Asterales foi realçada em amarelo para nela lembrar a presença da família Asteraceae.

Cumarinas em nível de tribo e subtribo :

Há fortes indícios de que os esqueletos **323** (MUT), **336** (MUT-NAS), **346** (MUT-MUT), **353** (MUT-MUT), **366** (MUT-MUT), **367** (MUT-MUT) e **376** (MUT-MUT) podem servir como marcadores quimiotaxonômicos em Mutisieae (o primeiro, em nível de tribo ; os demais, em nível de tribo e subtribo);

Os esqueletos **323**, **335** e **351** podem servir como marcadores quimiotaxonômicos para a subtribo Vernoniinae (Tab. 26, p. 101).

NOTA: O número de ocorrências dos referidos esqueletos é ainda baixo para uma proposta de marcadores quimiotaconômicos em nível infrafamiliar, entretanto temos de levar em conta que cumarinas não constituem uma classe química tão expressiva em ocorrências em Asteraceae como o são, por exemplo, flavonóides, poliacetilenos e sesquiterpenos lactonizados. Estas três classes químicas, em algumas circunstâncias, já são tomadas como marcadores quimiotaconômicos em nível infrafamiliar.

Benzofuranos em nível de tribo e subtribo :

Há fortes indícios de que os esqueletos **381** (GNA-GNA), **388** (EUP-CRI) e **391** (HLT-HLT) podem servir como marcadores quimiotaconômicos em nível de tribo e subtribo (Tab. 33, p. 116).

NOTA : Considerações similares às feitas para cumarinas , logo acima, servem também para os benzofuranos.

Estudo de agrupamentos (*clusters*) em tribos de Asteraceae

“A técnica classificatória multivariada da Análise de Agrupamentos é utilizada quando se deseja explorar as similaridades entre indivíduos (modo Q) ou entre variáveis (modo R) definindo-os em grupos, considerando simultaneamente, no primeiro caso, todas as variáveis medidas em cada indivíduo e, no segundo, todos os indivíduos nos quais foram feitas as mesmas mensurações. Segundo esse método, desenvolvido, inicialmente em Zoologia por taxonomistas numéricos, procura-se por agrupamentos homogêneos de itens representados por pontos num espaço n-dimensional em um número conveniente de grupos relacionando-os através de coeficientes de similaridades ou de semelhança”.(Landim, P. M. B. *Análise Estatística Multivariada de Dados Geológicos*, Aula 03 / Curso ministrado de agosto a novembro de 1999, UNESP — <http://www.rc.unesp.br/igce/sedimentar/aula03m.html>).

Dendrogramas (Gr. , δένδρον = árvore; γράμμα = letra, algo escrito ou desenhado) são diagramas similares a árvores apresentando interconexões entre objetos. Dendrogramas tais como cladogramas representam uma hierarquia de

categorias baseadas no grau de similaridade ou número de características compartilhadas especiais na taxonomia biológica. Seu uso é bastante amplo em estudos filogenéticos e pode estender-se às mais diferentes populações inseridas em todas as áreas do conhecimento humano .

A Análise de *Cluster*, conhecida também como Análise de Conglomerados, Taxonomia Numérica ou Reconhecimento de Padrões, é uma técnica estatística multivariante, cuja finalidade é dividir um conjunto de objetos em grupos (clusters, em inglês), de modo que os perfis dos objetos em um mesmo grupo sejam muito semelhantes entre si (coesão interna do grupo) e os dos objetos de *clusters* diferentes sejam distintos (separação externa do grupo).

A análise de agrupamentos ou *clusters* em dendrogramas permite identificar similaridades e dissimilaridades de populações em questão e é de interesse fundamental. *Clusters* estão amplamente exemplificados em publicações taxonômicas ou nas que fazem abordagens estatísticas de dados, cujo maior interesse é o estudo de agrupamentos. Mannina *et al.*, 2001, por exemplo, publicaram recentemente um trabalho muito interessante envolvendo a identificação geográfica de 216 amostras de azeite de oliva coletadas, ao longo de três anos, em diferentes regiões da Itália. Utilizaram a técnica de RMN ^1H de alto campo (600,13 MHz) com a finalidade investigar o seu potencial na caracterização geográfica dos azeites de oliva virgens. Após selecionarem as intensidades dos picos de RMN (ppm) significantes de 11 componentes mínimos de azeite de oliva (hexanal, *trans*-2-hexanal, dois outros aldeídos insaturados, formaldeído, três triterpenos, esqualeno, cicloartenol e β -sitosterol) avaliaram estatisticamente , através de ANOVA, LDA, TCA, etc., os dados de ressonância obtidos. A análise cuidadosa dos *clusters* do dendrograma resultante permitiu-os identificarem com 95% de probabilidade de acerto o local de produção da oliva como um fator discriminante, sem levar em consideração as culturas das quais os azeites de oliva são derivados.

Neste trabalho, fez-se uma avaliação estatística simples de agrupamentos de tribos de Asteraceae, com base em números de ocorrências das 8 classes químicas naquelas. Os dendrogramas obtidos estão logo abaixo (Figuras 62 e 63).

Figura 63. Dendrograma mostrando agrupamentos de 17 de tribos de Asteraceae (ligação completa, distâncias euclidianas) baseado em números de ocorrências de 8 classes de produtos naturais nas referidas tribos.

Comentário: No dendrograma de ligação completa, vários *clusters* podem ser facilmente identificados e assim divididos: Cluster I, contendo os grupos BAR, CAL, PLU, LIA., ART de mais forte similaridade (distâncias menores ou mais próximo de zero em ordem crescente, respectivamente. LIA e ARC correspondem muito bem à classificação de Bremer. MUT pode-se considerar um grupo à parte, porém de similaridade próxima à do *cluster I*. Na classificação de Bremer, coincidentemente Mutisieae (MUT) é uma tribo sem posição totalmente definida e está posicionada bem próximo de BAR. Cluster II, VER e INU, que são, a seguir, os grupos mais fortemente similares, formando o *cluster II*. Cluster III, HLN, LAC e GNA. Cluster IV, AST e EUP, grupos de similaridade já bem distante dos demais. Os grupos ANT (Anthemideae) e HLT (Heliantheae) são os de maior dissimilaridade e se separam nitidamente. CAR (Cardueae) pode-se considerar também um grupo distinto e , na classificação de Bremer, Cardueae é uma tribo com características de subfamília, Carduoideae, pertencente ao GRUPO VERNONIÓIDE.

Figura 64. Dendrograma mostrando agrupamentos de tribos de Asteraceae (Método de Ward-Distâncias euclidianas)

Comentário: O método de Ward praticamente dá os mesmos resultados obtidos com o de ligação completa (Fig. 63). Apenas troca as posições do *cluster* IV (EUP, AST) e do *cluster* II (VER, INU)-SEN e coloca os grupos HLT e ANT num *cluster* bem definido (aqui denominado *Cluster Ward*), porém mantendo a grande dissimilaridade relativa aos demais grupos e a “aparente similaridade” entre si. A mudança de escala pode influir nesse tipo de deslocamento. Vale a pena confrontar os números de ocorrências das classes químicas nas tribos (Tabela 45, ilustrativa) e tentar ver o que elas têm muito forte em comum ou dissimilar. No caso de Anthemideae e Heliantheae, ambas têm fortemente em comum produção excessiva de algumas classes químicas em detrimento de outras. Por exemplo, Anthemideae é a tribo que mais produz monoterpenos em Asteraceae. Com relação a esta classe química, Heliantheae produz quase oito vezes menos. Por outro lado, Heliantheae é a tribo líder em produção de poliacetilenos em Asteraceae, com número de ocorrências dessa classe química mais de dez vezes maior do que se tem registrado até então para Anthemideae. Estes dois últimos itens podem ser tomados como forte dissimilaridade entre ambas as tribos. E o que elas têm mais fortemente em comum é a produção de flavonóides, item que pode ser tomado como forte similaridade.. Mas colocando lado a lado suas variáveis de ocorrências em comum, vê-se que elas se distanciam bastante dos demais grupos em termos de similaridade, e também

entre si. É evidente que números de ocorrências de outras classes químicas comuns nessas duas tribos e nas demais poderiam ser avaliados comparativamente Nos dendrogramas aqui apresentados, os agrupamentos estão muito coerentes.

TRIBO	OCMONO	OCSELA	OCDITE	OCTRIT	OCCUMA	OCPOLI	OCFLAV	OCBENZ
MUT	59	128	30	63	171	5	85	6
BAR	7	0	0	12	0	0	139	0
CAL	0	20	93	0	3	24	1	7
PLU	80	18	0	5	29	2	11	0
LIA	3	73	0	9	0	0	8	0
ARC	1	24	21	35	1	23	0	0
VER	19	563	44	180	66	17	40	1
INU	292	708	68	8	7	0	109	0
HLN	318	479	124	7	21	180	665	9
LAC	20	197	0	11	74	0	432	2
GNA	158	189	419	29	51	112	362	30
SEN	130	1162	71	40	25	5	18	108
EUP	392	708	829	76	40	15	312	206
AST	464	152	1419	69	179	378	454	58
CAR	24	527	6	27	24	1055	222	0
HLT	473	1636	1230	60	40	1591	861	197
ANT	3627	1243	15	23	230	141	852	4

Tabela 45. NOs de 8 classes químicas em tribos de Asteraceae agrupadas (*Clusters*)

Grupo isolado

Cluster I

Cluster II

Cluster III

Cluster IV

Cluster Ward

Comentário: Os *clusters* correspondentes às Figuras 63 e 64 em forma de dendrogramas estão representados conjuntamente em listras coloridas na Tabela 45.

V - CONCLUSÕES

A abordagem quimiotaxonômica mostrada aqui permite-nos concluir que, para o entendimento das relações evolutivas dentro de Asteraceae, é necessário o uso de várias classes de metabólitos secundários produzidos na família bem como ferramentas capazes de realizar uma análise estatística multivariada de suas ocorrências e outras características interessantes.

O uso de ferramentas (*softwares*) capazes de realizar uma análise estatística multivariada dos números de ocorrências de metabólitos secundários e outras características interessantes em táxons é de grande valia e agiliza a obtenção e interpretação dos resultados.

O estudo de residuais é de suma importância para identificar, com maior probabilidade de acerto, números erráticos (*outliers*), que são considerados importantes em projetos baseados em ou inferidos de estatística. De grande importância também é a identificação de agrupamentos (*clusters*) nesses projetos.

Apesar dos grandes avanços da biologia química e molecular de plantas e animais e na alta confiabilidade de testes relacionados a macromoléculas (por exemplo, sequenciamentos de cpDNA, nDNA, citocromo c, etc.), em busca de achar relações filogenéticas entre táxons, os metabólitos secundários continuarão a ter sua importância quimiotaxonômica significativa, em paralelo à sistemática convencional.

Os flavonóides, com algumas ressalvas, são utilizados como marcadores quimiotaxonômicos em Asteraceae, mais em nível infrafamiliar. Alguns esqueletos de cumarinas (das tribos Mutisieae e Vernonieae, que são próximas filogeneticamente segundo a classificação de Bremer) e benzofuranos (das tribos Gnaphalieae, Eupatorieae e Heliantheae), com ocorrências não tão expressivas — porém representativas em termos de tribos e subtribos — possivelmente poderão servir como marcadores quimiotaxonômicos em Asteraceae.

Uma triagem mais refinada dos padrões de esqueletos de cumarinas e benzofuranos, em termos de padrões de substituição, podem revelar alguns novos marcadores quimiotaxonômicos na família. É evidente que muitas informações de natureza química e quimiotaxonômica podem-se perder, quando um grande número de esqueletos carbônicos com substituições O-R (R = metil, etil, prenil, outros) são agrupados e representados como um único tipo simples ou mais de um tipo de esqueleto. Por exemplo, GI-Grupo I de cumarinas, p.92, englobando 712 ocorrências do esqueleto 321 (Tab. 24), ou seja, ca. 75% das ocorrências desta classe química

na família; GIV-Grupo IV de benzofuranos, p. 110, englobando 437 esqueletos 378 e 21 esqueletos 398 (Tab. 31), um total de 458 esqueletos ou ca. 73% das ocorrências desta classe química na família. São, portanto, números de ocorrências muito expressivos.

O aumento de nossa base de dados com a incorporação de ocorrências de flavonóides, cumarinas e benzofuranos em Asteraceae, acompanhadas de referências e partes da planta estudadas, poderá colaborar em muito na elaboração de novos projetos de pesquisa. Por exemplo, estudos de correlações de atividades biológicas dos referidos metabólitos utilizando redes neurais.

Ocorrências de outras classes químicas tais como acetofenonas e fenilpropanóides em Asteraceae estão sendo levantadas e deverão ser colocadas em um trabalho de quimiotaxonomia em andamento. A ampliação de nossa base de dados total (ca. 29000 ocorrências, até o presente), poderá direcionar melhor o estudo de correlações das ocorrências e com resultados ainda mais confiáveis.

As matrizes de correlações e equações empíricas obtidas a partir de regressões lineares múltiplas para previsão de ocorrências de metabólitos secundários (8 classes correlacionadas) em tribos e subtribos de Asteraceae mostram-nos algumas correlações de razoáveis ($R = 0,60-0,80$ a boas ($R = 0,80-0,90$), em um nível- $p < 0,05$: cumarinas com monoterpenos; benzofuranos com diterpenos (a tribo Astereae é mais especializada em diterpenos, Fig. 34, p. 66); e flavonóides com lactonas sequiterpênicas.

O uso das equações empíricas apresentadas neste trabalho deve ser ainda criterioso, uma vez que inúmeros táxons em Asteraceae foram quimicamente pouco estudados (Mutisieae, por exemplo). Por outro lado, as equações podem dar boas pistas a pesquisadores que desejarem saber da possibilidade de ocorrências de qualquer uma das classes químicas correlacionadas neste estudo, em nível de tribo e subtribo de Asteraceae, o que lhes pode até poupar tempo de trabalho em seus projetos..

O equilíbrio entre oxidação e número de ocorrências visto aqui corrobora a teoria redox proposta por Gottlieb e Kaplan, 1993. Entretanto, a nossa proposta de constatar um balanço oxidativo na família através dos flavonóides necessita ainda de um estudo mais amplo (por exemplo, confrontar Asteraceae com outras famílias ricas em flavonóides e aplicando os mesmos critérios utilizados naquela).

O estudo de números de oxidação médios das classes químicas torna-se relevante quando se conhece muito bem a biogênese dos metabólitos. A aplicação de índices evolutivos de avanço baseados em flavonóides coloca Asteraceae como uma família de plantas muito evoluída quimicamente (predominância da rota do acetato sobre a do chiquimato). A posição de Cardueae (CAR) longe da posição das outras tribos na coordenada OXFLAV-B (Fig. 59, p.132) é muito boa e bate com a classificação de Bremer, que considera essa tribo como uma subfamília distinta (Carduoideae). A tribo Lactuceae (LAC) está muito bem posicionada na reta de regressão e, apesar de pertencer à sufamília Cichorioideae, mostra um maior parentesco com Cardueae, corroborando a classificação de Bremer (Fig.8, p.20).

Em níveis hierárquicos mais elevados, não é tão simples classificar marcadores quimiotaxonômicos.

VI - REFERÊNCIAS

- Alvarenga, S.A.V., Ferreira, M.J.P., Emerenciano, V.P., Cabrol-Bass, D. 2001. " Chemosystematic studies of natural compounds isolated from Asteraceae: characterization of tribes by principal component analysis ". *Chem. Int. Lab. Syst.*, **56**, 27-37.
- Audy et al., 1990. " Electrophoretic Forms of Lysozyme Activity in Various Plant Species". *Phytochemistry*, **29**, 1143-1159.
- Barker, J.S., Hattotuwwagana, C.K., Deew, M.G.B. 2002. " Computational studies of sweet-tasting molecules ". *Pure Appl. Chem.*, **74**, 1207-1217.
- Bentham, G. 1873. *Compositae*. In *Genera Plantarum*. Vol.2, No.1. Eds. Bentham, G. and Hooker, J. D. London: Lovell Reeve, 163-533.
- Birch, A. J. 1974. *Chemistry in Botanical Classification*. Eds. Bendz, G., Santesson, J., New York: Academic Press, 261-270.
- Bohlmann, F. 1967. *Fortschr. Chem. Org. Naturst.*, **25**, 1.
- Brant, A.J.C., Emerenciano, V.P., Rufino, A.R., Ferreira, M.J.P., Kaplan, M.A.C. 2001. *Uso de estatística multivariada na análise de ocorrências de benzofuranos e de outras classes químicas aromáticas em Asteraceae*, Painel-Resem, 2001.
- Bremer, K., Jansen, R.K., Karis, P.O., Källersjö, M., Keeley, S.C., Kim, K.-J., Michaels, H.J., Palmer, J.D., Wallace, R.S. 1992." A review of the phylogeny and classification of the Asteraceae ". *Nord. J. Bot.* **12**: 141-148.
- Bremer, K. 1994. *Asteraceae: Cladistics & Classification*. Portland, Oregon, Timber Press.
- Bremer, K. 1994a " Major Clades and Grades of the Asteraceae ". *Compositae: Systematics. Proceedings of the International Compositae Conference*, Kew, 1994 , Vol. 1,. 1-7, D. J. N. Hind & H. J. Beentje (eds.).
- Buzuk, G.N., Lovkova, M.Y., Sokolova, S.M., Tyuketin, Yu.V. 2002. " Evaluation of the correlation between the content of alkaloids and chemical elements in Washington Lupine (*Lupinus polyphyllus* Lindl.) based on statistical analysis and mathematical simulation ". *Applied Biochemistry and Microbiology*, **38**, 286-293.
- Conti, F. 2002. *Biometria*. <http://www.ufpa.br/dicas/biome/bioreg.htm>, 8/3/2002.
- Cronquist, A. 1981. *An Integrated system of classification of flowering plants*. New York: Columbia University Press.

Darwin, C. 1859. *Origin of Species by Means of Natural Selection, or The Preservation of Favoured Races in the Struggle for Life*. First Edition.

Denton, M. 1986. *Evolution: a theory in crisis*. New York: Adler & Adler.

Dewick, P. M. 1998. *Medicinal Natural Products: A Biosynthetic Approach*. p. 119. John Wiley & Sons, England.

Doyle, J.J. 1993. " DNA, Phylogeny, and the Flowering of Plant Systematics ". *BioScience*, **43**, 380-389.

Gros, E. G., Pomilio, A. B., Seldes, A. M. y Burton, G. 1985 . Monografia: *Introducción al Estudio de los Productos Naturales* - Fac. de Ciencias Exactas e Naturales de Univ. Buenos Aires.

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

Emerenciano, V.P., Rodrigues, G.V., Macari, P.A.T., Vestri, S.A., Borges, J.H.G., Gastmans, J.P., Fromanteau, D.L.G. 1994." Applications d'intelligence artificielle dans la chimie organique. XVII. Nouveaux programmes du projet SISTEMAT ". *Spectroscopy*, **12**, 91-98.

Emerenciano, V.P., Militão, J.S.L.T., Campos, C.C., Romoff, P., Kaplan, M.A.C., Brant, A.J.C. 2001. " Flavonoids as chemotaxonomic markers for Asteraceae". *Biochem. Syst. Ecol.*, **29**, 947- 957

Emerenciano, V.P., Brant, A.J.C., Ferreira, M.J.P., Alvarenga, S.A.V. 2003. "Statistical methods in studies of plant chemosystematics. The redox theory revisited". *Biochem. Syst. Ecol.*, submetido.

Ferreira, M.J.P., Alvarenga, S.A.V., Macari, P.A.T., Rodrigues, G.V., Emerenciano, V.P. 2003. "A program for terpenoid skeleton prediction based on botanical information ". *Biochem. Syst. Ecol.*, **31**, 25-43.

Finlay, J.M., Smith, G.S. 2002. "A Critical Analysis of Neurochemical Methods for Monitoring Transmitter Dynamics in the Brain ". *Psychopharmacology - The Fourth Generation of Progress*. Eds. Bloom, F.E. and Kupfer, D.J.; Raven Press, USA

Gastmans, J.P., Furlan, M., Lopes, M.N., Borges, J.H.G., Emerenciano, V.P. 1990. " A inteligência artificial aplicada à química de produtos naturais. O programa SISTEMAT. Parte I - Bases teóricas ". *Química Nova*, **13**, 10-16.

- Gershenzon, J., Mabry, T.J. 1983. " Secondary metabolites and the higher classification of angiosperms ". *Nord. J. Bot.*, **3**, 5-34.
- Gibbs, R.D. 1974. *Chemotaxonomy of Flowering Plants*, **1**, McGill-Queen's, University Press, Montreal & London, 52.
- Gornall, R.J., Bohm, B.A., Dahlgren, R. 1979. " The distribution of flavonoids in the angiosperms ". *Bot. Notiser*, **132**, 1-30.
- Gottlieb, O.R., Kaplan, M.A.C., Borin, M.R.M.B. 1996. *Biodiversidade. Um Enfoque Químico-Biológico*, Editora UFRJ.
- Gottlieb, O.R., 1989. " The role of oxygen in phytochemical evolution towards diversity ". *Phytochemistry*, **28**, 2545-2558.
- Gottlieb, R. O. and Kaplan, M. A. C. 1993. " Phytochemical Evolution : The Redox Theory. *Phytochemistry* , **2**, 171-176.
- Harborne, J. B. and Turner, B. L. 1984. *Plant Chemosystematics*, p.398, Academic Press, London.
- Harborne, J. B. and Turner, B. L. 1984a. *Plant Chemosystematics*, 130-138, Academic Press, London. I
- Hegnauer, R. 1967. "Chemical characters in plant taxonomy. Some possibilities and limitations " . *Pure & Appl. Chem.*, **14**, 173-187.
- Hendrickson, J.B., Cram, D.J., Hammond, G.S. 1970. *Organic Chemistry*. Third ed., p.1074, New York : McGraw-Hill Book Co.
- Herbert, R.B. 1989. *The Biosynthesis of Secondary Metabolites*, Second edition, p.21-22, Chapman & Hall, London, UK.
- Herman, P.P.J., Retief, E., Koekemoer, M. & Welman, W.G. 2000. "Asteraceae (Compositae)". In: *Seed Plants of Southern Africa* (ed. O.A. Leistner). *Strelitzia* 10: 101-170. National Botanical Institute, Pretoria.
- <http://www.museums.org.za/bio/plants/asteraceae>
- Heywood, V.H., Harborne, J.B., Turner, B. L., eds. 1977. *The biology and chemistry of the Compositae*. 2 vols. London: Academic Press.
- Jondico, I.J.O. 1986. "A mosquito larvicide in *Spilanthes mauritiana* ". *Phytochemistry*, **25**, 2289-2290.
- Jones, E.R.H. 1966. *Chem. Brit.*, **2**, 6.
- Kim, S-C, Crawford, D.J., Jansen, R.K. 1996." Phylogeny of the subtribe Sonchinae (Asteraceae: Lactuceae): Additional information from the non-coding region of cpDNA ". *American Institute of Biological Sciences*, Seattle, WA.

Kroumova, A.B., Wagner, G.J. 1999. " Mechanisms for elongation in the biosynthesis of fatty acid components of epi-cuticular waxes ". *Phytochemistry*, **50**, 1341-1345.

Lamb, N., Wahab, N., Rose, P.A., Shaw, A.C., Abrams, S.R., Cutler, A.J., Smith, P.J., Gusta, L.V., Ewan, B. 1996. " Synthesis, metabolism and biological activity of a deuterated analogue of the plant hormone S-(+)-abscisic acid ". *Phytochemistry*, **48**, 23-28

Lessing, C.F. 1832. *Synopsis Generum Compositarum*. Berlin: Duncker & Humblot.

Lin, J.T., Ramstad, E. and Heinstein, P. 1974."In vivo Biosynthesis of isopentenylacetophenones in *Eupatorium rugosum* ". *Phytochemistry*, **13**, 1809-1815.

Lindsay, R.K., Buchanan, B.G., Fergenbaum, E.A., Lederberg, J. 1980. *Applications of Artificial Intelligence for Organic Chemistry : the DENDRAL project*. New York, McGraw-Hill.

Mann, J.H. 1994. *Chemical Aspects of Biosynthesis*, Oxford Press Inc., Oxford, England.

Mannina, L., Patumi, M., Proietti, N., Bassi, D., and Segre, A. L. 2001." Geographical Characterization of Italian Extra-Virgin Olive Oils Using High-Field Spectroscopy " *J. Agric. Food Chem.*, **49**, 2687-2696.

Mariano, M. V., Teresa, O. R. R., Aguilar, A., Basurto, F. and Bye, R. 1996. Poster from 22nd Symposium on Natural Products Chemistry, May 6,7, 1996, Mexico, D. F. "Analgesic and Anti-Inflammatory Activities of an Aqueous Extract and the Flavone Pectoliarin of *Cirsium subcoriaceum* " *Rev. Latinoamer. Quím.*, **24**, p. 124

Mayr, E. 1991. *One Long Argument: Charles Darwin and the Genesis of Modern Evolutionary Thought*. Harvard University Press, Cambridge Massachusetts. <http://www.talkorigins.org/faqs/darwinism.html>.

Murray, R.D.H. 1982. *The Natural Coumarins: Occurrence, Chemistry and Biochemistry*, John Wiley & Sons.

Nakanishi, K., Goto, T., Itô, S., Natori, S., Nozoe. S. 1975." Fatty acid derivatives and related compounds ". *Nat. Prod. Chem.*, **2**. Academic Press, New York, USA, 53.

- Ngo, T.H., Vertommen, J., Kinget, R. 1997. "Formulation of artemisinin tablets". *Int J Pharmaceut*, **146**, 271-274.
- Park, S.-J., Korompai, E., Francisco-Ortega, J., Santos-Guerra, A., Jansen, R.K. 2001. " Phylogenetic relationships of Tolpis (Asteraceae: Lactuceae) based on *ndhF* sequence data ". *Plt. Syst. Evol.*, **226**, 23-33.
- Petrusson, L. 1999. <http://www.nrm.se/fbo/hist/linnaeus/linnaeus.html>.
- Proksch, P., Rodriguez, E. 1983." Chromenes and Benzofurans of the Asteraceae, Their Chemistry and Biological Significance ". *Phytochemistry*, **22**, 2335.
- .Proksch, P. and Clark, C. 1987. "Systematic Implications of Chromenes and Benzofurans from *Encelia* (Asteraceae) ". *Phytochemistry*. **26** (1), 171-174.
- Raven, P.H., Holm, R.W. 1967." Systematics and the levels of organization approach ". *Systematic Zool.*, **16**, 1-5.
- Reveal, J.L. 1997. *PBIO 450 Lecture Notes on Selected Families of Angiosperms: Asteridae*. <http://www.inform.umd.edu/PBIO/pb450/aste.html>).
- Ross, R.A.M. 1972. " Natural acetylenes of special biosynthetic interest". *Phytochemistry*, **11**, 221-227.
- Sandmann, G. 2001 " Minireview: Carotenoid Biosynthesis and Biotechnological Application ". *Archives of Biochemistry and Biophysics*, **385** (1), 4-12.
- Shemin, D. 1989. "An Illustration of the Use of Isotopes: The Biosynthesis of Porphyrins". *30BioEssays*, **10**, No.1,
<http://www.bmolchem.wisc.edu/teachinglab/fall675/shemin.html>.
- Seigler, D.S. 1977. "The naturally occurring cyanogenic glycosides ". *Progr. Phytochem.*, **4**, 83-120.
- Seigler, D.S. 1981. " Secondary Metabolites and Plant Systematics ". In: Conn, E.E. (ed.), *The biochemistry of plants, A comprehensive treatise VII, Secondary plant products*. New York: Academic Press, p. 139-176.
- Siebertz, R., Proksch, P., Wray, V. and Witte, L. "Accumulation and biosynthesis of benzofurans in root cultures of *Eupatorium cannabinum* ". (1989) *Phytochemistry*, **28**, 789-793.
- Sneath, P.H.A., Sokal, R.R. 1973. *Numerical Taxonomy: The Principles and Practice of Numerical Classification*. p.3. W.H. Freeman and Company, San Francisco, USA.
- Sorensen, N. A 1977. " Polyacetylenes and conservatism of chemical characters in the Compositae ". In *The biology and Chemistry of the Compositae*

(Heywood, V.H., Harborne, J.B., Turner, B. L., ed. 1977. *The biology and chemistry of the Compositae*. Vol.I 385-410. Academic Press, London.

Stassen, C. 1997. *The age of the Earth*. <http://www.talkorigins.org/faqs/faq.age-of-earth.html>; acessado em 16/12/2002.

Tomasi, N., Stein, M.L., Pizza, C. 1991. "Structure and Antiviral Activity of Triterpenoid Saponins from *Calendula arvensis*". *Planta Medica*, **57**, 250-253.

Turner, B.L. 1977. "Fossil history and geography". In: *The biology and Chemistry of the Compositae*, vol. 1. Eds. Heywood, V.H., Harborne, J.B. and Turner, B. L. London: Academic Press, 21-39.

Valeriano, M.M., Picini, A.G. 2000. "Uso de sistema de informação geográfica para a geração de mapas de médias mensais de temperatura do Estado de São Paulo". *Revista Brasileira de Agrometeorologia*, **8**, 255-262.

Villaescusa-Castillo, L., Díaz-Lanza, A.M., Gasquet, M., Delmas, F., Ollivier, E., Bernabé, M., Faure, R., Elias, R., Balansard, G. 2000. "Antiprotozoal Activity of Sesquiterpenes from *Jasonia glutinosa*". *Pharm. Biol.*, **38**, 176-180.

Wagner, H.H., Farnsworth, N.R. 1985. "Current Status of Stevioside as a Sweetening Agent for Human Use ". Eds. Kinghorn, D., Soejarto, D.D. *Economic and Medicinal Plant Research*, **1**, 1-52.

Waterman, P.G., Gray, A.I. 1987. *Chemical Systematics*. *Nat. Prod. Rep.*, **4**, 175-203.

Will, M., Fachinger, W., Richert, J.R. 1996." Fully automated structure elucidation - A spectroscopist's dream comes true ". *J. Chem. Inf. Comput. Science*, **36**, 221-227.

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

Zieg et al., 1983. "Selection of High pyrethrin Producing Tissue Cultures". *Planta Medica*, **48**, 88-91.

APÊNDICE DE ESTATÍSTICA

No tratamento estatístico, a simbologia, abreviações e alguns verbetes utilizados do programa computacional Sta_bas referente a parâmetros e cálculos apresenta os significados seguintes:

Diagrama de dispersão - gráfico sobre o qual cada medida individual é representada por um ponto (ou outro símbolo qualquer), sendo que a posição de cada ponto é determinada pelos valores observados em um indivíduo para as duas características medidas (por exemplo, hematócrito e hemoglobina). Denominado também de *gráfico XY*

(<http://www.usp.fm/dim/regressao/index.htm>, 2003)

Nível-p: nível de significância do teste. *"Essa probabilidade, representada freqüentemente por α , é geralmente especificada antes da extração de quaisquer amostras, de modo que os resultados obtidos não influenciem as escolhas. Se, por exemplo, é escolhido um nível de significância 0,05, ou 5%, no planejamento de um teste de hipótese, há então cerca de 5 probabilidades em 100 de a hipótese ser rejeitada, quando deveria ser aceita, isto é, há uma confiança de cerca de 95% de que se tome uma decisão acertada. Nesses casos, diz-se que a hipótese é rejeitada no nível de significância 0,05, o que significa que a probabilidade de erro seria de 0,05"* (Spiegel, Murray R. *Schaum's Outline of Theory and Problems of Statistics*, 1961 by McGraw-Hill, Inc. USA).

df: graus de liberdade (*degrees of freedom*). Graus de liberdade é um conceito ligado ao número de dados disponíveis (livres) para o cálculo da estatística. Por exemplo, ao estimarmos a média populacional com a média amostral, perdemos um grau de liberdade, assim a estatística t-student terá $n-1$ graus de liberdade. No caso da Tabela de ANOVA, os graus de liberdade do grupo será igual ao número de

grupos menos 1, os graus de liberdade total será igual a $n-1$ e os graus de liberdade do resíduo, a diferença entre esses dois

Teste de hipótese: $H_0: r=0$

Estatística de teste:

$$t = r \sqrt{\frac{n-2}{1-r^2}}$$

e sob H_0 , t tem distribuição t-Student com $(n-2)$ graus de liberdade (df)

Nota: este teste verifica se o r obtido é diferente de zero. Pode ser necessário usar também uma análise de variância

S: O **desvio padrão** é a medida de dispersão mais geralmente empregada, pois leva em consideração a totalidade dos valores da variável em estudo. É um indicador de variabilidade bastante estável. O desvio padrão baseia-se nos desvios em torno da média aritmética, e a sua fórmula básica pode ser traduzida como: a raiz quadrada da média aritmética dos quadrados dos desvios e é representada por S.

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n}}$$

A fórmula acima é empregada quando tratamos de uma população de dados não-agrupados.

Obs.: Quando nosso interesse não se restringe à descrição dos dados mas, partindo da **amostra**, visamos tirar inferências válidas para a respectiva população, convém efetuar uma modificação, que consiste em usar o divisor $n - 1$ em lugar de n . A fórmula ficará então:

$$S = \sqrt{\frac{\sum (X_i - \bar{X})^2}{n - 1}}$$

Quando os dados estão agrupados (temos a presença de freqüências, f_i) a fórmula do desvio padrão ficará:

$$S = \sqrt{\frac{\sum [(x_i - \bar{x})^2 \cdot f_i]}{\sum f_i}}$$

ou

$$S = \sqrt{\frac{\sum [(x_i - \bar{x})^2 \cdot f_i]}{\sum f_i - 1}}$$

quando se trata de uma amostra:

Erro padrão da estimativa

$$\sigma_{est} = \sqrt{\frac{\sum (Y - \hat{Y})^2}{N}}$$

onde N é o número de pares de pontos (X,Y) .

S² (σ²): Varância é o desvio padrão elevado ao quadrado. “A variância é uma medida que tem pouca utilidade como estatística descritiva, porém é extremamente importante na inferência estatística e em combinações de amostras” (FACEV - Faculdade de Ciências Econômicas de Vitória. *Introdução à Estatística Econômica* - 2º ano –AULANET-15, 2001). Prof. Paulo César Ribeiro da Silva)

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n - 1}$$

Uma definição mais clara para a variância: é a soma dos quadrados dos desvios dividida pelo tamanho da amostra menos 1 (n – 1). Os estatísticos chamam o valor (n – 1) de *graus de liberdade*.

F ou Teste F: Nomeado para o grande estatístico R. A. Fisher, (F) é a relação de dois **S quadrados (S²)** (i.e. estimativas da variância de uma população, baseadas na informação em duas ou mais amostras fortuitas). Quando empregado no procedimento intitulado ANOVA, o valor obtido de F fornece um teste para a

significação estatística das diferenças observadas entre as médias de duas ou mais amostras fortuitas.

$$F = \frac{S_1^2}{S_2^2} \quad S_1^2, \text{ variância do grupo 1 ; } S_2^2, \text{ variância do grupo 2.}$$

O valor F está associado a n_1-1 e n_2-1 graus de liberdade no numerador e no denominador, respectivamente.

r: Dado um par de medidas relacionadas (X e Y) em cada um de conjunto de itens, o coeficiente de correlação (r) provê um índice do grau para o qual as medidas emparelhadas co-variam em uma moda linear. Em geral, r será positivo quando itens com valores grandes de X também tendem a ter valores grandes de Y. Correspondentemente, itens com valores pequenos de X tendem a ter valores pequenos de Y. Assim, r será negativo quando itens com valores grandes de X tendem a ter valores pequenos de Y, enquanto que itens com valores pequenos de X tendem a ter valores grandes de Y. O valor de r, também denominado *coeficiente de correlação de Pearson*, é definido pela fórmula:

$$r = \frac{\sum xy - \frac{\sum x \sum y}{n}}{\left[\sum x^2 - \frac{(\sum x)^2}{n} \right] \left[\sum y^2 - \frac{(\sum y)^2}{n} \right]^{\frac{1}{2}}}$$

O valor de r varia de -1 a +1.

R : coeficiente de correlação múltipla.

R² : coeficiente de explicação ou de determinação múltipla

Na regressão múltipla, a medida relativa de adequação do ajuste é chamada de coeficiente de determinação múltipla e é designada pelo símbolo R^2 . É a relação entre a variação explicada pela equação de regressão múltipla e a variação total da variável dependente. Assim, $R^2=0,75$ significa que 75% de variância é explicada pelo modelo. O coeficiente de determinação (R^2) é um número no intervalo [0;1], calculado conforme a fórmula a seguir:

$$R^2 = \frac{S (Y_i^h - \check{Y})^2 \text{ variação explicada}}{S (Y_i - \check{Y})^2 \text{ variação total}} = \frac{\quad}{\quad}$$

Onde $\bar{Y}$ é o valor médio dos Y_i constantes da amostra e Y_i^h é o correspondente valor estimado através da equação ($Y_i^h = b_0 + b_1X_{1i} + b_2X_{2i} + \dots + b_kX_{ki}$), para o elemento i da amostra.

R^2 ajustado ou R_a^2 : Alguns autores recomendam o uso do coeficiente de determinação "ajustado" (R_a^2), que leva em conta o número de variáveis explicatórias em relação ao número de observações. O propósito desta medida é facilitar a comparação de diversos modelos de regressão, quando há alteração no número de variáveis ou na quantidade de dados, de um modelo para outro. Tal coeficiente é determinado da seguinte forma:

$$R_a^2 = 1 - \frac{(n - 1)}{(n - k - 1)} * (1 - R^2)$$

Onde R_a^2 é o coeficiente ajustado, R^2 é o coeficiente de determinação normal, k é o número de regressores e n é o tamanho da amostra. O coeficiente de determinação R^2 é empregado como um indicador inicial da precisão das regressões para a seleção dos modelos mais ajustados. A maioria dos *softwares* calcula os dois coeficientes.

BETA ($\beta = b$) : *coeficientes de regressão* tais como são apresentados na equação $y = a + b_1x_1 + b_2x_2 + b_3x_3 + b_4x_4 + \dots + b_nx_n$. Em correlações bivariadas, $\beta = r$

Coeficiente Angular ($b_{y,x}$)- medida da variação que ocorre em uma característica quando outra característica se modifica de uma unidade. Também é chamado de coeficiente angular ($= \tan \theta$) (ver figura no item abaixo,. **Retas de Regressão**)

$$b_{y,x} = \frac{\sum (x - \bar{x})(y - \bar{y})}{\sum (x - \bar{x})^2} = \frac{\sum xy - [(\sum x)(\sum y) / n]}{\sum x^2 - [(\sum x)^2 / n]}$$

(*Disciplina de Métodos Quantitativos em Medicina* (mpt-164 / edição 2002),

[http ://www.usp.fm/dim/regressao/index.htm](http://www.usp.fm/dim/regressao/index.htm), 2003)

Intercepto (a) - Ponto de intersecção da reta com a ordenada (eixo Y). Equivale ao valor de Y quando $X=0$.

Equação de Regressão - equação que define a linha reta que descreve a associação entre duas características e que permite estimar o valor de uma medida pela outra.

$$(y - \bar{y}) = b_{y,x} (x - \bar{x})$$

B : *Intercepto da reta de regressão* de y para x ao longo das tabelas de sumário de regressão.

Retas de Regressão - funções resultantes do ajuste de uma função linear entre 2 variáveis y e x. Para obter a reta de regressão é necessário calcular o **coeficiente angular** (coeficiente de regressão) e o **intercepto** da reta com a ordenada.

(Disciplina de Métodos Quantitativos em Medicina (mpt-164 / edição 2002)

<http://www.usp.fm/dim/regressao/index.htm>)

Regressão linear múltipla passo a passo adiante

Começar escolhendo a variável independente que explica a maior variação na variável dependente (por exemplo, Tabela 15.). Escolher uma segunda variável que explica a variação mais residual, e, em seguida, recalculando coeficientes de regressão.

Continuar até nenhuma variável explicar "significativamente" variação residual.

Regressão linear múltipla passo a passo para trás

Começar com todas as variáveis no modelo, e remover a menos "significante", uma de cada vez, até ficar com só variáveis "significantes" ANOVA (Análise de Variância / *Analysis of Variance*); LDA (Análise de Discriminante Linear / *Linear Discriminant Analysis*); TCA (Análise de Dendrograma / *Tree clustering Analysis*).

eps : Em equações de função adequadas, eps representa *epsilo* (e), o erro fortuito associado com medições.

Um exemplo:

a equação no diagrama de dispersão logo abaixo.

Clusters

A escolha do que se aglomera para fundir pode ser feita em muitas maneiras diferentes. Por exemplo, pode-se escolher a aglomeração por **ligação única** baseada na distância mínima (dissimilaridade mínima) calculada entre quaisquer duas unidades dentro de dois *clusters* ou **vizinhos mais próximos**. Pode-se

escolher a aglomeração de ligação completa baseando a avaliação de dissimilaridade na maior dissimilaridade entre 2 unidades em dois *clusters* ou vizinhos mais distantes. Ligação única e ligação completa são as opções clássicas para aglomerações / agrupamentos

Distância euclidiana

Geralmente, as funções de distância, tais como as funções de Manhattan e de distância euclidiana, são usadas para determinar a similaridade. Uma função de distância dá um valor mais elevado para pares de objetos que são menos similares um do outro. Uma função da similaridade é usada às vezes preferivelmente, que dá valores mais elevados para os pares que são mais similares.

Dados dois objetos de dados p-dimensionais $i = (x_{i1}, x_{i2}, \dots, x_{ip})$ e $j = (x_{j1}, x_{j2}, \dots, x_{jp})$, as seguintes funções de distância comuns podem ser definidas:

Função de Distância Euclidiana:

$$d(i, j) = \sqrt{|x_{i1} - x_{j1}|^2 + |x_{i2} - x_{j2}|^2 + \dots + |x_{ip} - x_{jp}|^2}$$

Função de Distância de Manhattan:

$$d(i, j) = |x_{i1} - x_{j1}| + |x_{i2} - x_{j2}| + \dots + |x_{ip} - x_{jp}|$$

Distância de quarteirões / blocos da cidade ou de Manhattan. Nomeada assim porque na maioria de cidades americanas não é possível ir diretamente entre 2 pontos. Uma rota que siga a grade regular das estradas é usada

$$D = |x1 - x2| + |y1 - y2|$$

$$\text{Distância}(x, y) = \sum |x - y|$$

Distância de Cook

A distância de Cook mede a influência de cada observação nos coeficientes de regressão, incluindo o intercepto. Quanto mais elevada a distância, maior a influência no caso.

Uma versão geral para a distância (D_i) de Cook é dada pela equação

$$D_i = \frac{|\hat{Y} - \hat{Y}_{(i)}|}{k\sigma^2},$$

onde $\hat{Y}$ e $\hat{Y}_{(i)}$ são $n \times 1$ vetores encaixados avaliados com base nos dados completos e nos dados sem o *i-ésimo* caso, respectivamente, k é a dimensão de β (coeficiente de regressão) e σ^2

é a variância.

(Rancel and Sierra, 2000) [Rancel, M. M. S. and Sierra, M. A. G. "Weighting and Deletion Approaches to Regression Diagnostics: A Comparison and an Extension". 2000. The Journal of Economic and social Research, 2 (1), 85-97.

Distância de Mahalanobis. Pode-se pensar nas variáveis independentes (em uma equação da regressão) como definindo um espaço multidimensional em que cada observação pode ser traçada. Pode-se traçar também um ponto que representa as médias para todas as variáveis independentes. Esse "ponto médio" no espaço multidimensional é chamado também de **centróide**. A *distância de Mahalanobis* (D) é a distância de um objeto / caso da centróide no espaço multidimensional, definida pelas variáveis independentes correlacionadas (se as variáveis independentes forem não-correlacionadas, é o mesmo que a distância euclidiana simples). Assim, esta medida fornece uma indicação de se uma observação é ou não um *outlier* com respeito aos valores das variáveis independentes, calculada conforme a equação logo abaixo.

$$D^2 = (X - \bar{X})^T S^{-1} (X - \bar{X}),$$

onde S é a matriz de variância-covariância :

$$S = \frac{1}{n-1} (X - \bar{X})(X - \bar{X})'$$

Obs.: O uso da métrica de Mahalanobis remove diversas das limitações da métrica euclidiana:

1. Explica automaticamente o scaling dos eixos coordenados.
2. Corrige para a correlação entre as características diferentes
3. Pode fornecer limites de decisão lineares assim como curvados

Método de Ward (método da variância mínima)

Similaridade: Junta o par de *clusters* cujo aglomerador minimiza o aumento no total dentro da soma dos quadrados do erro do grupo, baseada na distância euclidiana entre centróides.

Clusters homogêneos em uma hierarquia simétrica; a centróide do *cluster* facilmente caracterizada por seu centro de gravidade:

Vantagens: bom em recuperação de estrutura de *cluster* , forma hierarquia única e exata.

Desvantagens: sensível a *outliers*, pobre em recuperação de *clusters* alongados.

CURRICULUM VITAE

I - Dados Pessoais

Antônio Jedson Caldeira Brant

Brasileiro, solteiro, químico, 56

Res.: Rua Mário Dias, 318 - Butantã - CEP: 05519-110 - São Paulo-SP

Tel/fax: (11) 3742-8409

Trab.: Instituto de Química da USP - Cid. Universitária - SP

e-mail: abrant@quim.iq.usp.br

II - Idiomas : alemão e inglês (fluência)

III - Dados Profissionais

Bacharel em Química, Universidade Federal de Minas Gerais, 1973

IV - Atividades em pesquisa, desenvolvimento, aplicações e assistência técnica :

- 1999 até o presente : Instituto de Química da Universidade de São Paulo - SP
Cursando pós-graduação e matriculado no Mestrado em Química Orgânica, área de Produtos Naturais de Plantas.
- Set, 1996 - Jan, 1999 : Linken Química Ltda; São Caetano do Sul – SP;
Cargo/função: Gerente Técnico
- Fev-Set ,1996 : Homy Química Ltda; Jardinópolis – SP; Cargo/função: Assistente Técnico
- 1977-1992 : Hoechst do Brasil Química e Farmacêutica S.A. - Dep. Resinas e Dispersões; São Paulo - SP; Cargos/funções: Técnico de Resinas; Técnico Especializado; Técnico Sênior
- 1974-1976 : Johnson&Johnson do Brasil Ind.e Com. S.A. - Dep. Adesivos e Plásticos - P&D; São José dos Campos – SP; Cargos/funções: Técnico de Especificações; Pesquisador
- Nota : Aposentado por tempo de serviço em dezembro de 1997.

V - Disciplinas cursadas em pós-graduação

- Didática de Ensino em Química (IQ-USP, 2000)
- Química Bioorgânica (IQ-USP, 2000)
- Evolução Química em Angiospermas (IQ-USP, 2000)
- Química Orgânica Avançada I (IQ-USP, 1999)
- Métodos Espectroscópicos em Química Orgânica I (IQ-USP, 1999)
- Fotoquímica Orgânica - Conceitos, Mecanismos e Aplicação Sintética (IQ-USP, 1999)
- Química Orgânica Avançada II (Stevens Institute of Technology, USA, 1993)
- Termodinâmica Química (Stevens Institute of Technology, USA, 1993)
- Fotoquímica Orgânica (IQ-USP, 1979)
- Química de Altos Polímeros (IQ-USP, 1979)
- Introdução à Físico-Química (ITA, 1976)
- Introdução à Físico-Química Orgânica (ITA, 1976)
- Química de Doenças Prastológicas (ITA, 1976)

VI - Trabalhos submetidos/publicados:

1. C.C. Campos, M. Zambom, J.S.L. Militão, P. Romoff, A.J.C. Brant, M.A.C. Kaplan, V.P. Emerenciano, "Flavonoids as chemotaxonomic markers in Asteraceae"; *Biochemical Systematics and Ecology*, **29**, 947-957 (2001).
2. F.M.M. Magri, J.S.L. Militão, M.J.P. Ferreira, A.J.C. Brant, V.P. Emerenciano, "A new program to ^{13}C NMR spectrum prediction based on tridimensional models"; *Spectroscopy*, **15**, 99-117 (2001)
3. M.J.P. Ferreira, A.J.C. Brant, G.V. Rodrigues, V.P. Emerenciano, "REGRAS: An auxiliary program for pattern recognition and substructure elucidation of monoterpenes"; *Spectroscopy*, **15**, 65-98 (2001).
4. M.J.P. Ferreira, A.J.C. Brant, G.V. Rodrigues, V.P. Emerenciano, "Automatic identification of terpenoid skeletons through ^{13}C NMR data disfunctionalization"; *Analytica Chimica Acta*, **429**, 151-170 (2001).
5. S.A.V. Alvarenga, J.P. Gastmans, G.V. Rodrigues, M.J.P. Ferreira, A.J.C. Brant, V.P. Emerenciano, "Application of Artificial Intelligence in Organic

Chemistry. Part XXII*. SISTAX– An intelligent tool in recovering information on natural products chemistry”; *Journal of Brazilian Chemical Society*, in press (2002).

6. M.J.P. Ferreira, A.J.C. Brant, A.R. Rufino, S.A.V. Alvarenga, F.M.M. Magri, V.P. Emerenciano, “Prediction of occurrences of diverse chemical classes in Asteraceae through neural networks”; *Phytochemical Analysis*, submetido (2002).
7. V.P. Emerenciano, D. Cabrol-Bass, M.J.P. Ferreira, S.A.V. Alvarenga, A.J.C. Brant, P. Romoff, “The use of statistical methods in studies of plant chemosystematics. The Redox theory revisited”; *Biochemical Systematics and Ecology*, submetido (2002).

VII - Trabalhos apresentados em congresso:

XXIII RESEM-Reunião Anual sobre Evolução, Sistemática e Ecologia Micromoleculares- Universidade de São Paulo, São Paulo-SP, 2001

1. A.R. Rufino, V.P. Emerenciano, M.J.P. Ferreira, S.A.V. Alvarenga, A.J.C. Brant, “Aplicação de redes neurais em quimiotaxonomia. Classificação de grupos vegetais”.
2. A.R. Rufino, V.P. Emerenciano, A.J.C. Brant, M.J.P. Ferreira, “Previsão de esqueletos de alcalóides aporfínicos utilizando técnicas de redes neurais e RMN ^1H ”.
3. A.R. Rufino, V.P. Emerenciano, M.J.P. Ferreira, A.J.C. Brant, “Identificação de esqueletos dos alcalóides benzilisoquinolínicos, quinolínicos, quinolizidínicos e aporfínicos através de dados de RMN ^{13}C ”.
4. A.J.C. Brant, V.P. Emerenciano, A.R. Rufino, M.J.P. Ferreira, M.A.C. Kaplan, “Uso de estatística multivariada na análise de ocorrências de benzofuranos e de outras classes químicas aromáticas em Asteraceae”
5. A.J.C. Brant, V.P. Emerenciano, A.R. Rufino, M.J.P. Ferreira, F.M.M. Magri, M.A.C. Kaplan, “Equilíbrio oxidativo entre nove classes de metabólitos secundários em tribos de Asteraceae”.

VIII - Atividades em ensino/magistério

- 1975-1981: Como atividade paralela ao trabalho nas indústrias J&J e Hoechst, fui professor na Faculdade de Engenharia Química de Lorena-FAENQUIL, onde lecionei as disciplinas: Processos Unitários Orgânicos e Físico-Química.
- 1969-1973: Professor de Química em vários colégios e curso de vestibular em Belo Horizonte-MG (Curso Pitágoras / Colégio Assunção, Fundação Cultural de Belo Horizonte, Colégio D. Cabral, Colégio S. Paulo, Colégio Comercial e Normal de Minas Gerais).
- 1966-1967: Professor primário em Teófilo Otoni-MG (Grupo Escolar Tancredo Neves).

IX - Atividades de extensão cultural e profissional

- Estágios na Hoechst AG / Frankfurt am Main / Alemanha, 1979 (7 meses); 1982 (2 semanas, atualização); 1989 (2 semanas, atualização). Enfoque: Aprimoramento técnico em aplicação de resinas e dispersões poliméricas aquosas nos mais diversos segmentos da indústria.
- Estágio na HARCO/ Inglaterra, 1989 (1 semana). Enfoque: Adesivos sensíveis à pressão; Dispersões acrílicas para tintas e congêneres.
- Alguns cursos voltados a relações humanas e *marketing*; Noções básicas de informática; Participação em cursos , seminários, congressos e palestras técnicas; Cursos de alemão e inglês em vários níveis de aprimoramento no Brasil e no exterior (Alemanha e USA, respectivamente).

São Paulo, 18 de dezembro de 2002.