**UNIVERSIDADE DE SÃO PAULO**

**FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO**

# Impressão digital metabólica do gênero *Espeletia* (Asteraceae) e sua correlação com dados filogenéticos e biogeográficos

Guillermo Federico Padilla González

Ribeirão Preto
2014

**UNIVERSIDADE DE SÃO PAULO**

**FACULDADE DE CIÊNCIAS FARMACÊUTICAS DE RIBEIRÃO PRETO**

# Impressão digital metabólica do gênero *Espeletia* (Asteraceae) e sua correlação com dados filogenéticos e biogeográficos

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

Área de concentração: Produtos Naturais e Sintéticos

**Orientado**: Guillermo Federico Padilla González
**Orientador**: Prof. Dr. Fernando Batista da Costa

Ribeirão Preto
2014

## FICHA CATALOGRÁFICA

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

Padilla-González, Guillermo Federico
Impressão digital metabólica do gênero *Espeletia* (Asteraceae) e sua correlação com dados filogenéticos e biogeográficos. Ribeirão Preto, 2014.
104 p.; 30cm.

Dissertação de Mestrado, apresentada à Faculdade de Ciências Farmacêuticas de Ribeirão Preto/USP – Área de concentração: Produtos Naturais e Sintéticos.
Orientador: Da Costa, Fernando Batista.

1. Metabolômica. 2. Quimiotaxonomia. 3. Biogeografia

Foto da capa e contracapa: Autor Guillermo Federico Padilla González

**FOLHA DE APROVAÇÃO**

Nome do aluno: Guillermo Federico Padilla González.
Título do trabalho: Impressão digital metabólica do gênero *Espeletia* (Asteraceae) e sua correlação com dados filogenéticos e biogeográficos.

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

Área de Concentração: Produtos Naturais e Sintéticos

Orientador: Prof. Dr. Fernando Batista da Costa

Aprovado em:

Banca Examinadora

Prof. Dr._____

Instituição: _____Assinatura:_____

Prof. Dr._____

Instituição: _____Assinatura:_____

Prof. Dr._____

Instituição: _____Assinatura:_____

## DEDICATÓRIA

A mi familia por el apoyo incondicional y por creer siempre en mí.

## AGRADECIMENTOS

Agradeço especialmente a meu orientador, Prof. Dr. Fernando Batista da Costa, por dar-me a oportunidade de fazer parte do seu grupo de pesquisas (AsterBioChem), pela orientação, confiança, apoio e ajuda constante.

Ao Dr. Maurício Diazgranados do *National Museum of Natural History, Smithsonian Institution*, EUA, por fornecer as amostras de *Espeletia* estudadas neste projeto e pelas enriquecedoras discussões.

Ao Herbário da Universidade de São Paulo, pela recepção do material vegetal proveniente do *Smithsonian Institution*, em especial ao curador, Prof. Dr. José Rubens Pirani.

Meus mais sinceros agradecimentos aos professores Leonardo Gobbo Neto, Norberto Peporine Lopes, Mônica Tallarico Pupo, Carlos Montanari e Luiz Alberto Beraldo de Moraes, pela paciência, ensinamentos e grandes contribuições à minha formação profissional e consequentemente ao desenvolvimento deste projeto.

Agradeço também aos funcionários do laboratório de farmacognosia, Mário, Waltinho e Angélica, pela grande ajuda facilitando a consecução de muitos materiais e pela valiosa amizade.

Gostaria de agradecer aos meus amigos do AsterBioChem: Regane Barbosa (RB), Daniela Chagas (DC), Tiago Branquinho (TB), Bruno Leite (BL), Rosana Casoti (RC), Danniela Faleiro (DF), Felipe Antunes (FA), Annylory Lima (AL), Marcelo de Souza (MS), e Gari Ccana (GC), por terem contribuído de uma forma ou outra para a conclusão deste projeto, incluindo a construção do banco de dados ChemEsp (TB), preparo de extratos (DC, RC, AL), análise no UHPLC-UV-MS (DC), análises quimiométricas (DC, TB), isolamento de padrões cromatográficos (RB, BL, DF, GC), críticas e discussões enriquecedoras (RC, FA, MS), pois sem sua valiosa contribuição este projeto não teria sido possível.

Agradeço também a Fabian Niño por ter contribuído voluntariamente no preparo de extratos, mesmo como ao Dr. Ricardo Silva, pela ajuda técnica na análise no software R.

Agradeço especialmente à Muriel Coneo por voluntariamente ter contribuído na revisão do banco de dados ChemEsp e no preparo de extratos, além do constante apoio moral.

Finalmente agradeço à Faculdade de Ciências Farmacêuticas de Ribeirão Preto-USP, local de desenvolvimento deste trabalho, ao programa de pós-graduação da FCFRP-USP e às agências de fomento CAPES, CNPq e FAPESP pelo apoio financeiro recebido direta ou indiretamente.

# RESUMO

PADILLA GONZÁLEZ, G. F. **Impressão digital metabólica do gênero *Espeletia* (Asteraceae) e sua correlação com dados filogenéticos e biogeográficos.** 2014. 104f. Dissertação (Mestrado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2014.

O gênero *Espeletia* Mutis ex Bonpl (Asteraceae, Millerieae, Espeletiinae) constitui um exemplo clássico de um táxon sujeito a rápidos processos adaptativos evolutivos. Com *ca.* 71 espécies, este gênero se diversificou em elevadas altitudes num ecossistema emergente formado após o retraimento dos glaciares no final do Plioceno e início do Pleistoceno, os páramos, um tipo de ecossistema que biogeograficamente funciona como um complexo de "ilhas" presentes principalmente no noroeste dos Andes tropicais da Venezuela, Colômbia e Equador. Devido à sua complexa história evolutiva e morfologia característica, este gênero tem sido foco de muitas pesquisas envolvendo sua biologia, ecologia, taxonomia e filogenia. No entanto, do ponto de vista fitoquímico, tem sido pouco estudado e é desconhecida a influência da geografia na química do metabolismo secundário destas espécies. Neste estudo, a impressão digital metabólica de 120 amostras do gênero *Espeletia* foi obtida por UHPLC-UV-MS e submetida a análises quimiométricas. A correlação com dados geográficos revelou uma forte influência da biogeografia no metabolismo secundário das espécies deste gênero, apresentando uma impressão digital característica de acordo com seu país de origem numa escala global e de acordo com complexo de páramos de origem numa escala regional, o qual concorda com a filogenia atual da subtribo baseada nos marcadores ITS, ETS, rp116 e AFLPs. A análise por OPLS-DA e a desreplicação de extratos permitiu identificar as principais substâncias correlacionadas com a origem geográfica das espécies, mesmo como permitir a construção de modelos de predição da origem geográfica de novos extratos baseados nos seus perfis químicos. Além disso, uma análise filogenética efetuada com dados metabolômicos revelou, embora com pouco suporte, uma correlação geográfica em que as espécies da Venezuela correspondem ao clado ancestral do gênero. Adicionalmente, a química do metabolismo secundário da subtribo Espeletiinae foi revisada e um bando de dados, o ChemEsp (Chemistry of Espeletiinae), foi construído incluindo todos os metabólitos descritos na literatura, mesmo como outras informações químicas, botânicas e geográficas. Finalmente, a desreplicação de extratos permitiu a identificação de várias substâncias não descritas previamente no gênero *Espeletia* e/ou na subtribo Espeletiinae, incluindo uma grande variedade de flavonoides e ácidos cafeoilquínicos, junto com outros metabólitos já descritos no gênero por métodos fitoquímicos clássicos. Os resultados obtidos, além de contribuírem para a geração de conhecimento sobre a química do metabolismo secundário do gênero *Espeletia*, revelaram informações sobre as relações filogenéticas e biogeográficas do gênero baseadas em marcadores químicos, empregando-se modernas metodologias analíticas e computacionais.

**Palavras chave:** Asteraceae; Biogeografia; Espeletiinae; Metabolômica; Quimiotaxonomia.

## ABSTRACT

PADILLA GONZÁLEZ, G. F. **Metabolomic fingerprint of the genus *Espeletia* (Asteraceae) and its correlation with phylogenetic and biogeographic data.** 2014. 104f. Dissertation (Master). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2014.

The genus *Espeletia* Mutis ex Bonpl (Asteraceae, Millerieae, Espeletiinae) constitutes a classic example of a taxon undergoing rapid adaptive evolutionary processes. With *ca*. 71 species, this genus diversified at high altitudes in an emerging ecosystem formed after the retreat of the glaciers in the late Pliocene and early Pleistocene, the páramos, a type of ecosystem that biogeographically acts as an “island” complex in the northwestern tropical Andes of Venezuela, Colombia and Ecuador. Due to its complex evolutionary history and characteristic morphology, this genus has been focus of several investigations regarding its biology, ecology, taxonomy and phylogeny. However, phytochemically, they have been poorly studied and it is unknown the influence of geography in their secondary metabolism. In this study, the metabolomic fingerprint of 120 samples of the genus *Espeletia* was analyzed by UHPLC-UV-MS and chemometrics. Correlations with geographical data revealed a strong influence of biogeography on their secondary metabolism, displaying a characteristic fingerprint according to their country of origin in a global scale and by páramo complex in a regional scale, which is in agreement with the current phylogeny of the subtribe based on ITS, ETS, rp116 and AFLPs markers. An OPLS-DA analysis and extract dereplication allowed the identification of the main compounds correlated with their geographical origin, which also served for building models to predict the geographical origin of unknown samples based on their chemical profile. Also, a phylogenetic analysis was performed with metabolomic data which interestingly, although with low support, also displayed a geographical structure with the Venezuelan cluster as the ancestral clade of the genus. Furthermore, the secondary metabolite chemistry of the subtribe Espeletiinae was reviewed and a data base, ChemEsp (Chemistry of Espeletiinae), was built including all secondary metabolite reports available in the literature as well as other chemical, botanical and geographical information. Finally, the extract dereplication allowed the identification of several compounds previously unreported in the genus *Espeletia* and/or the subtribe Espeletiinae, including a wide range of flavonoids and caffeoylquinic acids, along with other metabolites already reported in the genus by classic phytochemical methods. These results, besides of contributing to further knowledge about the secondary metabolite chemistry of the genus *Espeletia*, provided valuable information about the phylogenetic and biogeographic relationships in the genus based on chemical markers by means of modern analytical and computation techniques.

**Keywords:** Asteraceae; Biogeography; Chemotaxonomy; Espeletiinae; Metabolomics.

## LISTA DE FIGURAS

**Figura 1**. Extensão aproximada dos páramos (em preto) na cordilheira dos Andes (fonte: BUYTAERT et al., 2006). .....1

**Figura 2.** Ecossistema de páramo (A, B) e espécies do gênero *Espeletia* (C-G). .....3

**Figura 3.** *Screenshot* da tela do banco de dados ChemEsp no software MarvinView. .....18

**Figura 4.** Visualização de estrutura química individual em 2D com informaçoes específicas (acima) e sua estrutura química em 3D com potenciais grupos aceptores (A) e doadores de ligações de hidrogênio (abaixo), empregando-se o software MarvinView. .....19

**Figura 5.** Classes de metabólitos secundários descritas na subtribo Espeletiinae…20

**Figura 6.** Análise hierarquica de clusters (HCA) da subtribo Espeletiinae obtida com os metabólitos secundários descritos na literatura. Em azul, espécies principalmente de origem colombiana; em vermelho, espécies de origem venezuelana. Abreviações dos gêneros: Ca., *Carramboa*; Co., *Coespeletia*; E., *Espeletia*; Es., *Espeletiopsis*; L., *Libanothamnus*; R., *Ruilopezia*; T., *Tamania*; P., *Paramiflos*. .....22

**Figura 7.** Cromatograma da amostra A, extrato fresco (acima) e seco no fluxo de ar (abaixo). A seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.....24

**Figura 8.** Cromatograma da amostra B, extrato fresco (acima) e seco no fluxo de ar (abaixo). A seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.....24

**Figura 9.** Cromatograma da amostra C, extrato fresco (acima) e seco no fluxo de ar (abaixo). A seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.....24

**Figura 10.** Cromatograma da amostra D, extrato fresco (acima) e seco no fluxo de ar (abaixo). A seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.....24

**Figura 11.** Cromatograma da amostra E, extrato fresco (acima) e seco no fluxo de ar (abaixo). A seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.....25

**Figura 12.** Cromatograma da amostra A, extrato fresco (acima) e seco no *SpeedVac* (abaixo). A Seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco. ..... 25

**Figura 13.** Cromatograma da amostra E, extrato fresco (acima) e seco no *SpeedVac* (abaixo). ..... 25

**Figura 14.** Análise hierárquica de clusters (HCA) do extrato fresco de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*. ..... 26

**Figura 15. A**. *Score plot* da análise por componentes principais (PCA) de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*; **B**. Mapa de distribuição geográfica das cinco espécies acima mencionadas (Adaptado de DIAZGRANADOS, 2012). ..... 27

**Figura 16.** Análise hierárquica de clusters (HCA) do extrato seco no fluxo de ar comprimido de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*. ..... 28

**Figura 17.** Análise hierárquica de clusters (HCA) dos tricomas de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*. ..... 29

**Figura 18.** Análise hierárquica de clusters (HCA) das folhas sem tricomas de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*. ..... 30

**Figura 19.** *Score plot* da análise por OPLS-DA dos tricomas (em verde) e das folhas sem tricomas (em azul) de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*. ..... 31

**Figura 20.** Cromatogramas de íons totais (TIC) sobrepostos de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). ..... 33

**Figura 21.** *Score plot* da análise por componentes principais (PCA) de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Países de origem das espécies indicados em diferentes cores. .....34
**Figura 22.** Cromatogramas de íons totais (TIC) sobrepostos das quatro réplicas da amostra 34 analisadas por UHPLC-UV-MS (modo negativo). (Ver códigos das amostras no apêndice I). .....35
**Figura 23.** Análise hierárquica de clusters (HCA) de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Caixa A: réplicas da amostra 34; caixa B: primeiro agrupamento das espécies venezuelanas; caixa C: segundo agrupamento das espécies venezuelanas. .....36
**Figura 24.** *Score-plot* da análise por OPLS-DA de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Países de origem das espécies indicados em diferentes cores. .....39
**Figura 25.** *S-plot* da análise por OPLS-DA de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Em vermelho as variáveis correlacionadas (VIP>3,0) com o agrupamento das espécies em cada país. Números indicando o ID das principais variáveis discriminantes. .....40
**Figura 26.** *Score-plot* da análise por OPLS-DA de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo), com áreas dos picos transformadas em variáveis binárias. Países de origem das espécies indicados em diferentes cores. .....43
**Figura 27.** Dendograma da análise por OPLS-DA de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Cordilheiras de origem das espécies indicadas em diferentes cores. .....45
**Figura 28.** Dendograma da análise por OPLS-DA de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Principais maciços de páramos indicados em diferentes cores. .....46
**Figura 29.** Mapa da Colômbia, Equador e Venezuela com principais maciços de páramos considerados neste estudo ressaltados em diferentes cores (Adaptado de DIAZGRANADOS, 2012). .....46
**Figura 30.** Matriz de classificação do conjunto de espécies empregadas na validação externa do modelo por OPLS-DA. .....51
**Figura 31.** Modelo de predição geográfica por árvores de decisão .....52

**Figura 32.** *Score-plot* da análise por OPLS-DA de três híbridos do gênero *Espeletia* e suas espécies parentais analisadas por UHPLC-UV-MS (modo negativo). .....54
**Figura 33.** *Score-plot* da análise por OPLS-DA de três híbridos do gênero *Espeletia* e suas espécies parentais analisadas por UHPLC-UV-MS (modo negativo). .....55
**Figura 34.** *Score-plot* da análise por OPLS-DA de quatro variedades e uma forma do gênero *Espeletia* e suas espécies parentais analisadas por UHPLC-UV-MS (modo negativo). .....57
**Figura 35.** Estruturas químicas dos ácidos clorogênicos desreplicadas nas espécies de *Espeletia*. a. ácido protocatecuico; b. ácido quínico; c. ácido 3-*O*-(*E*)-cafeoilquínico; d. ácido 5-*O*-(*E*)-cafeoilquínico; e. ácido 3,4-di-*O*-(*E*)-cafeoilquínico; f. metil 3-*O*-cafeoil-4-*O*-feruloilquinato; g. ácido 4,5-di-*O*-(*E*)-cafeoilquínico; h. ácido 3,5-di-*O*-(*E*)-cafeoilquínico; i. ácido tricafeoilaltrárico.....63
**Figura 36.** Estruturas químicas dos flavonoides (agliconas) desreplicados nas espécies de *Espeletia*. a. quercetina; b. 3-metóxiquercetina; c. pinobanksina; d. hesperetina; e. 8,8''-metilenobisquercetina. .....65
**Figura 37.** Estruturas químicas dos flavonoides glicosilados desreplicados nas espécies de *Espeletia*. a. quercetina-3,4'-*O*-di-glucopiranosídeo; b. quercetina-3-*O*-(6''-malonil-glucosídeo)-7-*O*-glucosídeo; c. rutina; d. quercetina-3-*O*-galactosídeo; e. quercetina-3-*O*-arabinosídeo; f. Isoramnetina-3-glicosídeo; g. quercetrina. .....66
**Figura 38.** Estruturas químicas dos flavonoides esterificados com ácido caféico desreplicados nas espécies de *Espeletia*. a. castillicetina; b. quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo; c. quercetina-3-(6″-cafeoilsoforosídeo).....68
**Figura 39.** Estruturas químicas dos diterpenos desreplicados nas espécies de *Espeletia*. a. ácido *ent*-12-oxocauro-9(11),16-dien-18-oico; b. ácido *ent*-12-hidroxicauro-9(11),16-dien-18-oico; c. ácido grandiflorólico; d. ácido *ent*-16α-hidroxicauran-19-oico; e. ácido pimárico; f. ácido *ent*-15-O-isobutiroxicaur-16-en-19-oico; g. ácido *ent*-15-O-senecioxicaur-16-en-19-oico.....69
**Figura 40.** Espectros de massas no modo positivo do metil 3-*O*-cafeoil-4-*O*-feruloilquinato (acima) e do ácido 3,4-di-*O*-(*E*)-cafeoilquínico (abaixo).....70
**Figura 41.** Espectro de massas no modo positivo e possíveis fragmentos do metil 3-*O*-cafeoil-4-*O*-feruloilquinato. .....71
**Figura 42.** Espectro de massas no modo negativo e possíveis fragmentos do ácido tricafeoilaltrárico. .....72

**Figura 43.** Árvore de consenso estrito obtida a partir de dados químicos gerados na análise de 120 indivíduos do gênero *Espeletia* por UHPLC-UV-MS (modo negativo). *Outgroup*: *Senecio nivoaureus* (Asteraceae, Senecioneae). Em azul espécies de origem venezuelana e em verde espécies de origem colombiana... .....76

**Figura 44.** Árvore de *maximum likelihood* obtida a partir de dados químicos gerados na análise de 120 indivíduos do gênero *Espeletia* por UHPLC-UV-MS (modo negativo). *Outgroup*: *Senecio nivoaureus.* (Asteraceae, Senecioneae). Em azul espécies de origem venezuelana e em verde espécies de origem colombiana... .....77

## LISTA DE TABELAS

**Tabela 1.** Espécies do gênero *Espeletia* analisadas no estudo preliminar. .....9

**Tabela 2.** Principais substâncias presentes nos tricomas de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS, de acordo com o modelo por OPLS-DA.....31

**Tabela 3.** Principais substâncias presentes nas folhas sem tricomas de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS, de acordo com o modelo por OPLS-DA.....32

**Tabela 4.** Principais substâncias correlacionadas com o agrupamento das espécies do gênero *Espeletia* no cluster colombiano (VIP>3.0) e sua desreplicação.....41

**Tabela 5.** Principais substâncias correlacionadas com o agrupamento das espécies do gênero *Espeletia* no cluster venezuelano (VIP>3.0) e sua desreplicação.....42

**Tabela 6.** Principais substâncias correlacionadas com o agrupamento das espécies do gênero *Espeletia* no cluster colombiano (VIP>3.0) e sua desreplicação (matriz 2). .....43

**Tabela 7.** Principais substâncias correlacionadas com o agrupamento das espécies do gênero *Espeletia* no cluster venezuelano (VIP>1.8) e sua desreplicação (Matriz 2). .....44

**Tabela 8.** Principais substâncias correlacionadas com o agrupamento das espécies do gênero *Espeletia* nos páramos de origem e sua desreplicação. .....48

**Tabela 9.** Substâncias desreplicadas nas amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS. .....60

**Tabela 10.** Critérios estatísticos obtidos nas análises filogenéticas por máxima parcimônia (MP) e *maximum likelihood* (ML). .....74

## LISTA DE ABREVIATURAS E SIGLAS

| | |
|---|---|
| AsterDB | *Asteraceae Database* |
| CAS | *Chemical Abstract Service* |
| ChemEsp | *Chemistry of Espeletiinae* |
| IC | Índice de Consistência |
| DNP | *Dictionary of Natural Products* |
| EtOH | Etanol |
| FT | Transformada de Fourier |
| GC-MS | *Gas Chromatography coupled to Mass Spectrometry* |
| HCA | *Hierarchical Clustering Analysis* |
| HCD | *Higher energy Collisional Dissociation* |
| HESI | *Heated Electrospray* |
| IH | Índice de Homoplasia |
| HPLC | *High Performance Liquid Chromatography* |
| HPLC-MS | *High Performance Liquid Chromatography coupled to Mass Spectrometry* |
| IDH | Índice de Deficiência de Hidrogênio |
| KNN | *K-Nearest Neighbors* |
| LiNPA | *Library of Natural Products from Asteraceae* |
| *m/z* | Unidades de massa/carga |
| MeCN | Acetonitrila |
| MeOH | Metanol |

| | |
|---|---|
| ML | *Maximum Likelihood* |
| MP | Máxima Parcimônia |
| MRFA | Acetato de L-metionil-arginil-fenilalanina |
| MS/MS | Espectrometria de massas sequencial |
| OPLS-DA | *Orthogonal Partial Least Squares – Discriminant Analysis* |
| PCA | *Principal Component Analysis* |
| QSAR | *Quantitative Structure-Activity Relationship* |
| QSRR | *Quantitative Structure-Retention Relationship* |
| CR | Índice de Consistência Re-escalonada |
| RF | Rádio Frequência |
| IR | Índice de Retenção |
| *SpeedVac* | *Speed* vácuo |
| Sdf | *Structure Data Format* |
| TBR | *Tree-Bisection-and-Reconection* |
| UHPLC-UV-MS | *Ultra High Performance Liquid Chromatography coupled to Ultraviolet and Mass Spectrometry detectors* |
| UPLC | *Ultra Performance Liquid Chromatography* |
| UV | Ultravioleta |
| UVar | *Unit variance* |
| VIP | *Variable Importance in the Projection* |

# LISTA DE EQUAÇÕES

**Equação 1.** Formula de escalonamento pelo método de Pareto. .....12
**Equação 2.** Formula de escalonamento pelo método de *Unit variance.* .....13
**Equação 3.** Estimativo de *Jack-knife.* .....39

**SUMÁRIO**

**RESUMO** ..... **i**

**ABSTRACT** ..... **ii**

**LISTA DE FIGURAS** ..... **iii**

**LISTA DE TABELAS** ..... **iv**

**LISTA DE ABREVIATURAS E SIGLAS** ..... **v**

**LISTA DE EQUAÇÕES** ..... **vi**

**1 INTRODUÇÃO** ..... **1**

**2 OBJETIVOS** ..... **6**

2.1 Objetivo geral ..... 6

2.2 Objetivos específicos ..... 6

**3 MATERIAL E MÉTODOS** ..... **7**

3.1 Criação de um banco de dados químicos de Espeletiinae ..... 7

3.2 Material vegetal ..... 7

3.3 Preparo de extratos ..... 8

3.4 Impressão digital metabólica e análise estatística multivariada ..... 8

3.5 Análise de tricomas ..... 11

3.6 Análise de correlação *in silico* ..... 12

3.7 Desreplicação de extratos ..... 13

3.8 Análise filogenética ..... 14

**4 RESULTADOS E DISCUSSÃO** ..... **17**

4.1 Banco de dados químicos de Espeletiinae ..... 17

4.2 Revisão da química de Espeletiinae ..... 20

4.3 Impressão digital metabólica e análise estatística multivariada ..... 23

4.3.1 Análise de extratos frescos e secos ..... 23

4.3.2 Análise de tricomas ..... 28

4.3.3 Análise dos 120 táxons de *Espeletia* ..... 33

4.4 Análise de correlação *in silico* ..... 38

4.4.1 Correlação com dados biogeográficos .....38

4.4.2 Modelos de predição da origem geográfica.....50

4.4.3 Correlação com dados taxonômicos .....54

4.5 Desreplicação de extratos .....57

4.6 Análise filogenética.....73

**5 CONCLUSÕES .....78**

**6 REFERÊNCIAS.....79**

**APÊNDICE I.** Nomes, códigos, locais e datas de coleta das espécies estudadas neste projeto.....90

**APÊNDICE II.** Parâmetros utilizados no software MZmine. .....95

**APÊNDICE III.** Parâmetros utilizados para seleção de atributos e construção da árvore de decisão J48 no software Weka. .....98

**APÊNDICE IV.** Grupos de treino e teste para predição e validação do modelo por OPLS-DA.....99

**APÊNDICE V.** Substâncias puras injetadas no UHPLC-UV-MS. .....102

**ANEXO I.** Árvores filogenéticas da subtribo Espeletiinae. .....103

# 1 INTRODUÇÃO

No final da era do Plioceno e início do Pleistoceno, entre dois e quatro milhões de anos atrás, um novo ecossistema emergiu após o retraimento dos glaciares na região norte dos Andes tropicais na América do Sul: o páramo (HOOGHIEMSTRA; RAN, 1994; HOOGHIEMSTRA; VAN DER HAMMEN, 2004). Este ecossistema, presente em alturas de 3.500 a 4.800 m acima do nível do mar, cobre as áreas altas da cordilheira dos Andes formando uma faixa descontínua desde a cordilheira de Talamanca em Panamá e Costa Rica, passando por Colômbia, Venezuela e Equador até a depressão de Huancabamba no norte do Perú (BUYTAERT et al., 2006) (Figura 1).

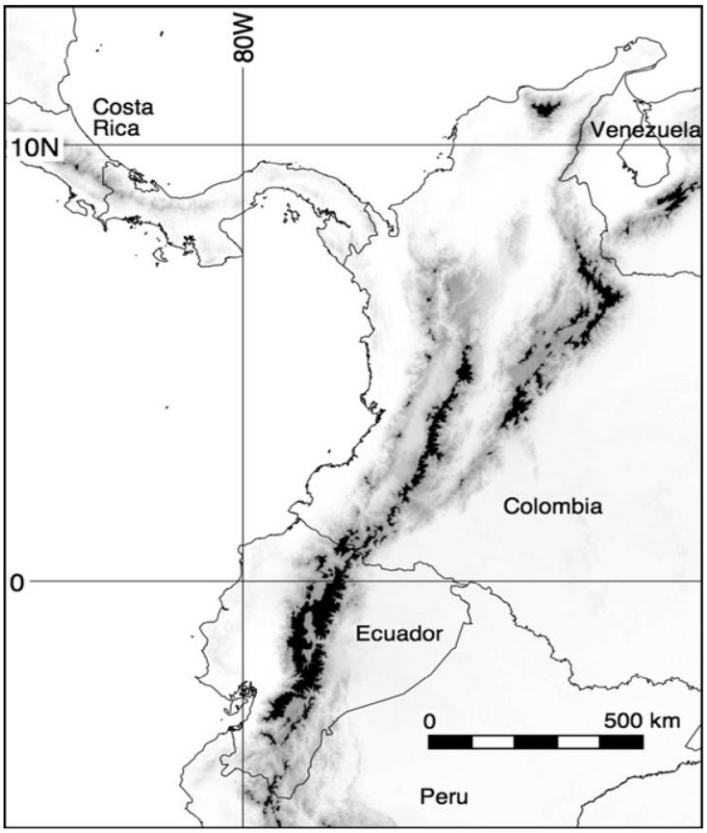



**Figura 1**. Extensão aproximada dos páramos (em preto) na cordilheira dos Andes (fonte: BUYTAERT et al., 2006).

O páramo (Figura 2A-B) é caracterizado por suas condições climáticas extremas como altos níveis de radiação ultravioleta (UV), solos pobres, muito drenados e com pH ácido, além de sazonalidade diária chegando a temperaturas inferiores a 0 °C a noite (em alturas superiores aos 4000 m) e superiores a 15 °C no

dia, enquanto a temperatura média anual não apresenta variação (MONASTERIO; SARMIENTO, 1991; MORALES et al., 2007). Estas características somadas aos eventos sucessivos de glaciações e inter-glaciações que deram origem a este ecossistema, geraram uma alta biodiversidade, sendo considerado o ecossistema de alta montanha de maior diversidade florística e riqueza de espécies endêmicas do mundo (SMITH; CLEEF, 1988; LUTEYN, 1999; RANGEL-CH, 2000).

Estando isolado por profundos vales andinos ou densos bosques altoandinos, o páramo é um tipo de ecossistema que biogeograficamente funciona como “ilha” já que, análogo às ilhas oceânicas, apresenta um alto grau de endemismo em que muitas espécies estão restritas a determinado local (MORALES et al., 2007), com pouca ou nenhuma comunicação entre espécies presentes em diferentes maciços de páramos. O isolamento geográfico e a origem recente deste ecossistema o levaram a ser considerado um sistema ideal para compreender relações evolutivas complexas e mecanismos de especiação em plantas (HOOGHIEMSTRA; VAN DER HAMMEN, 2004; CUATRECASAS, 2013).

Dentro dos principais grupos vegetais característicos dos páramos estão os *Frailejones* (Figura 2C-G), um grupo de espécies pertencentes à família Asteraceae, tribo Millerieae, subtribo Espeletiinae. Esta subtribo, com *ca.* 143 espécies agrupadas em oito gêneros (*Espeletia, Espeletiopsis, Coespeletia, Ruilopezia, Libanothamnus, Carramboa, Tamania* e *Paramiflos*) (CUATRECASAS, 1976), tem sido considerada um exemplo clássico para estudar radiações adaptativas rápidas, sendo o gênero *Espeletia* Mutis ex Bonpl o principal modelo por se tratar do maior e mais diverso gênero (*ca.* 71 spp) da subtribo (CUATRECASAS, 1976, 2013).

Devido à grande diversidade morfológica deste grupo e sua rápida evolução adaptativa, existe grande interesse no estudo da sua biologia, ecologia, sistemática e taxonomia. Portanto, inúmeros estudos têm sido realizados desde a primeira descrição de uma espécie de *Espeletia* em 1792 até a atualidade, com o intuito de estabelecer relações taxonômicas, esclarecer sua história evolutiva e determinar a influência da geografia na diversificação deste grupo (DIAZGRANADOS, 2012; CUATRECASAS, 2013). Por exemplo, um dos primeiros estudos detalhados da subtribo Espeletiinae (CUATRECASAS, 1933) reportou que diferentes áreas geográficas apresentam diferentes espécies, concluindo, posteriormente, que devido

à geografia Andina é possível que o padrão que distribuição destas espécies reflita a evolução do grupo (CUATRECASAS, 2013). Esta hipótese foi confirmada recentemente com a filogenia da subtribo baseada em diversos marcadores moleculares (DIAZGRANADOS, 2012). No entanto, embora muito estudadas do ponto de vista de sua biologia e ecologia, quimicamente têm sido pouco estudadas e ainda é desconhecida a influência da biogeografia na química do seu metabolismo secundário ou se a composição química destas espécies reflete suas relações de parentesco evolutivo.

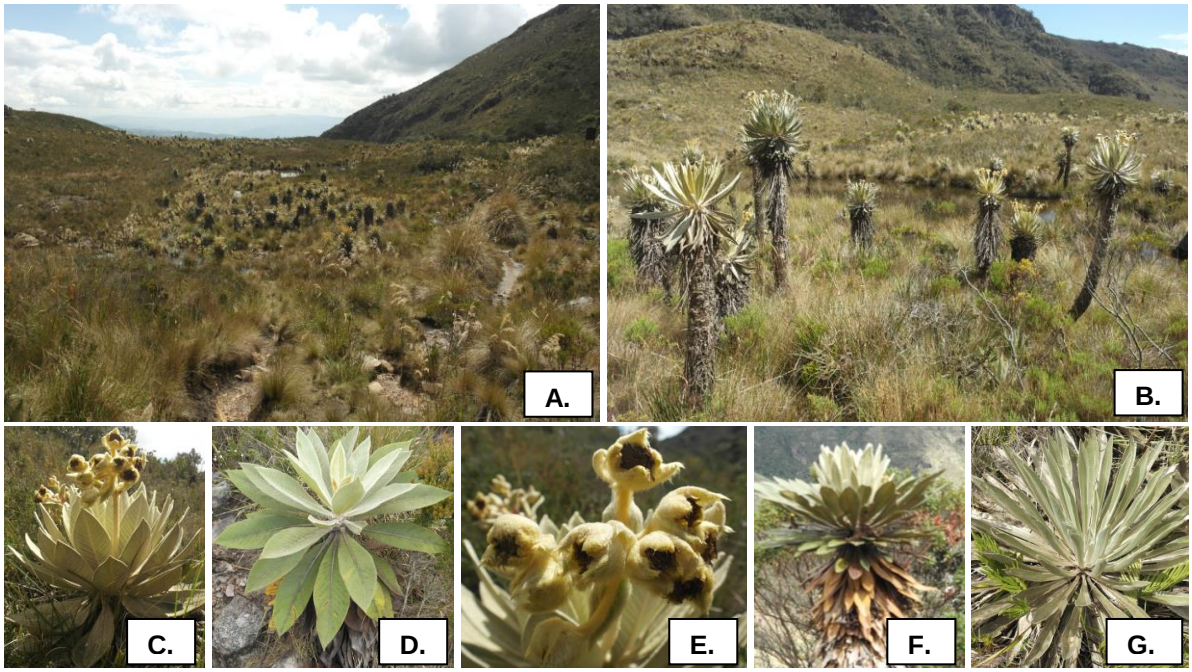



**Figura 2.** Ecossistema de páramo (A, B) e espécies do gênero *Espeletia* (C-G).

A química do metabolismo secundário do gênero *Espeletia* é caracterizada pela presença de metabólitos da classe dos terpenoides, especialmente diterpenos de tipo caurano, dos quais se destacam os ácidos *ent*-caurenóico, grandiflorênico, grandiflorólico e seus derivados, além de alguns mono e sesquiterpenos (PIOZZI et al., 1971; USUBILLAGA et al., 1973; MORALES et al., 1973; BOHLMANN; SUDING; CUATRECASAS; KING; et al., 1980; TORRENEGRA et al., 1994), sendo que estudos fitoquímicos empregando métodos clássicos resultaram na identificação de alguns dos metabólitos presentes em apenas 15 espécies: *E. argentea, E. barclayana, E. batata, E. grandiflora, E. hartwegiana, E. killipii, E. nana, E. pycnophylla, E. schultzii, E. semiglobulata, E. tenorae, E. tunjana, E. uribei*, *E. weddellii*, e *E. x aurantia* (USUBILLAGA; MORALES, 1970; PIOZZI et al., 1972;

PROANO et al., 1972; BOHLMANN; SUDING; CUATRECASAS; KING; et al., 1980; TORRENEGRA; TELLEZ, 1995; TELLEZ et al., 1998; USUBILLAGA et al., 2003; JANSEN et al., 2007; IBÁÑEZ; USUBILLAGA, 2008; PENA et al., 2012), das 71 que compõem o gênero.

O uso de marcadores químicos constitui uma abordagem interessante que pode auxiliar na compreensão da taxonomia deste gênero, considerando que estudos prévios relataram a presença de duas lactonas sesquiterpênicas de tipo melampolido (polimatina B e acetato de longipilina), o qual levou a questionar o status da subtribo Espeletiinae e propor uma possível circunscrição com a então subtribo Melampodiinae (TORRENEGRA; TELLEZ, 1995). Além disso, estudos desenvolvidos por Alvarenga et al. (2005) destacaram o valor quimiossistemático dos diterpenos do tipo caurano presentes no gênero *Espeletia*.

A fitoquímica clássica, apesar de ser muito utilizada em correlações com a biogeografia e taxonomia de um determinado grupo de espécies, apresenta desvantagens, tais como a dificuldade na identificação de metabólitos presentes em baixas concentrações, a alta quantidade de material vegetal requerida e o elevado tempo e esforço necessários no processo de isolamento e elucidação estrutural, dificultando assim a análise de um elevado número de espécies. São necessários, portanto, estudos sistemáticos empregando metodologias mais sensíveis que requeiram pequenas quantidades de material vegetal, permitam identificar um maior número de substâncias de uma forma rápida e padronizada, além de estudar várias dezenas de espécies ao mesmo tempo. Isto pode ser atingido mediante uma abordagem metabolômica, a qual é definida como o estudo e caracterização do conjunto de metabólitos presentes em determinado complexo biológico. A metabolômica apresenta três diferentes abordagens, a análise de “metabólitos alvo”, que se refere à “detecção e quantificação precisa de um conjunto único ou pequeno de substâncias alvo”; a análise do “perfil metabólico”, que permite a identificação e quantificação de um grupo de metabólitos associados a vias específicas ou pertencentes a determinada classe, e a “impressão digital metabólica ou *fingerprint* metabólico”, que é usada na comparação de metabolômas completos sem a necessidade da identificação das substâncias (FIEHN, 2002; KRASTANOV, 2010). Esta última abordagem, se aliada à estatística multivariada, pode ser empregada para fazer correlações entre a composição química e a origem geográfica ou

proximidade taxonômica, pois possibilita estabelecer relações de parentesco entre as espécies avaliadas de acordo com a similaridade das suas impressões digitais metabólicas.

Devido à importância de se realizar análises fitoquímicas mais detalhadas, que permitam fazer correlações mais precisas, bem como investigar a química do metabolismo secundário de um maior número de espécies de *Espeletia*, neste projeto foi proposto estudar o metaboloma de 120 amostras vegetais do gênero *Espeletia* (61 espécies, nove híbridos, cinco variedades, uma forma e 44 duplicatas), por *Ultra High Performance Liquid Chromatography – Ultraviolet Spectroscopy – Mass Spectrometry* (UHPLC-UV-MS) com analisador do tipo FT-Orbitrap em combinação com ferramentas de quimioinformática, estatística multivariada e métodos *in silico* de análise. Além de realizar uma abordagem moderna para detectar as impressões digitais metabólicas das espécies avaliadas, foi proposto ainda correlacionar os dados químicos obtidos com dados biogeográficos, taxonômicos e filogenéticos de *Espeletia.*

# 2 OBJETIVOS

## 2.1 Objetivo geral

Obter a impressão digital metabólica de espécies do gênero *Espeletia* (Asteraceae - Millerieae) por UHPLC-UV-MS e, com emprego de estatística multivariada e métodos *in silico*, correlacionar os dados químicos obtidos aos dados biogeográficos, filogenéticos e taxonômicos do gênero. Além disso, desreplicar os extratos vegetais usando vários bancos de dados disponíveis e mediante a criação de um banco de dados dos metabólitos secundários descritos na subtribo Espeletiinae.

## 2.2 Objetivos específicos

a) Revisar a literatura científica publicada sobre a química do metabolismo secundário de espécies da subtribo Espeletiinae e construir um banco de dados com tais informações;

b) Obter a impressão digital metabólica de espécies de *Espeletia* por UHPLC-UV-MS e correlacionar os dados químicos com dados taxonômicos, filogenéticos e biogeográficos de *Espeletia*;

c) Através das analises de correlação *in silico* e desreplicação, identificar as substâncias correlacionadas com os agrupamentos observados;

d) Desenvolver modelos de predição da origem geográfica de novas amostras de *Espeletia*;

e) Fazer uma análise filogenética do gênero *Espeletia* com base em marcadores químicos;

# 3 MATERIAL E MÉTODOS

## 3.1 Criação de um banco de dados químicos de Espeletiinae

Com a finalidade de conhecer a fitoquímica da subtribo Espeletiinae, auxiliar na desreplicação de extratos vegetais e realizar uma análise quimiotaxonômica preliminar *in silico*, foi necessário criar um banco de dados com todos os metabólitos secundários descritos na subtribo. Sua construção foi feita com base na revisão bibliográfica de todos os artigos científicos disponíveis em várias bases de dados, tais como *SciFinder Scholar, ISI Web of Knowledge*, *Google Scholar*, *ScienceDirect*, SCOPUS e Scielo. O banco de dados, denominado ChemEsp (Chemistry of Espeletiinae), faz parte do AsterDB (Asteraceae Database) do nosso grupo de pesquisas AsterBioChem (www.asterbiochem.org). O ChemEsp contem as seguintes informações: estrutura química da substância em 2D e com atribuição de estereoquímica, nomes usual e sistemático, número CAS (Chemical Abstract Service), nome da espécie e parte da planta de onde foi isolada ou descrita cada substância, país e local de coleta da planta com dados de GPS quando disponível, número de voucher e herbário onde está depositada cada exsicata, classe química e tipo de esqueleto carbônico, além da fórmula molecular e massa monoisotópica em alta resolução. A estrutura química de cada substância foi desenhada no software ChemBioDraw Ultra 12.0 (CambrigeSoft, Inglaterra), e a fórmula molecular e massa monoisotópica de cada substância foram calculadas no mesmo software. O banco de dados foi elaborado empregando o software JChem for Excel (ChemAxon Ltd., Hungria) e se encontra disponível eletronicamente em formato .sdf (Structure Data Format), sendo compatível com o software MarvinView 5.12.2 (ChemAxon Ltd., Hungria).

## 3.2 Material vegetal

As partes aéreas de 61 espécies do gênero *Espeletia*, nove híbridos, cinco variedades, uma forma, e 33 duplicatas de espécies coletadas a partir de indivíduos presentes no mesmo páramo e 11 duplicatas de espécies coletadas em diferentes páramos (totalizando 120 amostras vegetais, Apêndice I) foram obtidas a partir de coleções biológicas presentes no *National Museum of Natural History, Smithsonian Institution* (Washington DC., EUA). As amostras foram coletadas pelo Dr. Mauricio Diazgranados em páramos da Colômbia e Venezuela entre os anos 2007 e 2010,

secas em sacolas plásticas com sílica gel, posteriormente identificadas e enviadas à FCFRP-USP pelo Dr. Diazgranados. As amostras vegetais selecionadas estão em perfeito estado de conservação.

### 3.3 Preparo de extratos

A metodologia de extração empregada neste estudo foi baseada no protocolo de extração e análise de plantas para estudos metabolômicos por cromatografia líquida acoplada a espectrometria de massas descrito por de Vos et al. (2007), com algumas modificações. Neste estudo, 15 mg de folhas secas foram pulverizadas utilizando nitrogênio líquido, capsula de porcelana e pistilo. Posteriormente, o pó foi transferido a tubos Eppendorf onde se adicionou 1,5 mL de etanol (EtOH):$H_2O$ em proporção 7:3 (v/v). Após a adição do solvente extrator, os tubos foram agitados por 10 s e colocados no banho de ultrassom (Maxiclean 1600, Unique, Brasil) por 20 min em temperatura de 25 °C. O sistema de solventes empregado foi selecionado com base em estudos prévios do nosso grupo de pesquisas (AsterBioChem), onde se determinou que a maior eficiência na extração da maior diversidade de metabólitos secundários de várias espécies da família Asteraceae foi obtida nessas condições (CASOTI et al., 2013; CHAGAS-PAULA, 2013). Ressalta-se que vários estudos metabolômicos recomendam extrações hidroalcoólicas em proporções iguais ou superiores a 1:1 (v/v) (DE VOS et al., 2007; ALVAREZ-SANCHEZ et al., 2010; KIM et al., 2011).

Após a extração, os tubos foram centrifugados por 10 min a 14.000 rpm e o sobrenadante foi transferido a outro tubo Eppendorf onde foram particionados com 0,5 mL de *n*-hexano, filtrados em filtro de PTFE de 0,2 µm, e então analisados no UHPLC-UV-MS. Todos os solventes utilizados no processo de extração foram grau HPLC e utilizou-se um filtro para cada extrato vegetal.

### 3.4 Impressão digital metabólica e análise estatística multivariada

Previamente ao início dos estudos metabolômicos das 120 amostras, realizou-se um estudo preliminar com cinco espécies de três diferentes origens geográficas (Tabela 1), com a finalidade de determinar se a secagem dos extratos e posterior ressuspenção, metodologia empregada previamente no nosso grupo de pesquisas (CASOTI et al., 2013; CHAGAS-PAULA, 2013), influenciava na impressão digital metabólica e consequentemente na correlação final entre a composição

química e a origem geográfica das espécies. A secagem dos extratos foi efetuada em fluxo de ar comprimido (amostras A-E) e no *SpeedVac* (Savant SPD 2010-220, Thermo Electron Corporation, EUA) (amostras A e E).

**Tabela 1.** Espécies do gênero *Espeletia* analisadas no estudo preliminar.

| Código | Espécie | Origem geográfica | |
|---|---|---|---|
| | | País | Estado |
| A | *E. grandiflora* | Colômbia | Cundinamarca |
| B | *E. barclayana* | Colômbia | Cundinamarca |
| C | *E. frontinoensis* | Colômbia | Antioquia |
| D | *E. jajoensis* | Venezuela | Mérida |
| E | *E. schultzii* | Venezuela | Mérida |

Este estudo preliminar foi de utilidade também na determinação das melhores condições de análise no UHPLC-UV-MS (Exactive Plus™, Thermo Fisher Scientific, Alemanha) equipado com analisador tipo *Orbitrap™* e fonte de ionização por *Heated Electrospray* (HESI), utilizando uma coluna C18 (Hypersil Gold, 50 x 2,1 mm, Thermo Fisher Scientific, Alemanha) com um tamanho de partícula de 1,9 µm e uma precoluna C18 para UPLC (Phenomenex, EUA). Empregou-se uma vazão de fase móvel de 0,3 mL/min, sendo composta por um gradiente de acetonitrila (MeCN):$H_2O$ com adição de 0,1% de ácido fórmico, obtendo-se resultados satisfatórios. O gradiente da fase móvel empregado foi: 5% de MeCN em 3 min, 5%-100% de MeCN em 30 min, seguido de 100% de MeCN em 35 min e 100%-5% de MeCN em 40 min para a limpeza e recondicionamento da coluna, respectivamente. A temperatura da coluna foi controlada em 30 °C, enquanto a temperatura do *rack* se manteve em 10 °C.

O espectrômetro de massas foi operado nos modos positivo e negativo simultaneamente numa resolução de 70,000 e faixa de varredura de 150-2000 *m/z*. Foi injetado um volume de 10 µL de cada extrato num tempo máximo de injeção de 250 ms. Os parâmetros da fonte de ionização por HESI foram: voltagem do spray de 3,6 kV no modo positivo e -3,2 kV no modo negativo, temperatura do capilar e do aquecedor de 300 ºC, fluxo de nitrogênio (sheath gas) de 30 unidades e nível de *S-lens* RF de 55.

Antes de iniciar as análises, o espectrômetro de massas foi calibrado de acordo com as especificações da *Themo Fisher Scientific* para cada polaridade

separadamente. No modo positivo empregou-se uma solução de cafeína, acetato de *L*-metionil-arginil-fenilalanina (MRFA), solução stoque *Ultramark 1621* e *n*-butilamina em MeCN: metanol (MeOH): $H_2O$ contendo 1% de ácido acético. No modo negativo se empregou uma solução de dodecil sulfato de sódio, taurocolato de sódio e a solução stoque *Ultramark 1621* em MeCN: MeOH: $H_2O$ com 1% de ácido acético.

As amostras foram analisadas randomicamente, sendo que a amostra número 34 (Apêndice I) foi injetada quatro vezes, uma no início, duas no meio e uma no fim da sequência cromatográfica, com a finalidade de conferir a reprodutibilidade das análises. A amostra "branco" (EtOH:$H_2O$, 7:3 v/v) foi injetada apenas no início e no fim da sequência, já que a análise desta amostra relativamente pura no meio dos extratos brutos pode re-acondicionar parcialmente a coluna, resultando em pequenos deslocamentos no tempo de retenção (DE VOS et al., 2007).

Os dados brutos obtidos no UHPLC-UV-MS (cromatogramas, espectros de massas e no UV) em formato .raw foram separados no modo positivo e negativo no software RecalOffline, Xcalibur (Thermo Fisher Scientific, Alemanha) e exportados ao software *open source* MZmine 2.10 (MZmine Development Team, Finlândia), onde foi efetuado o tratamento dos dados: eliminação de ruído, detecção de picos, alinhamento, eliminação de isótopos, deconvolução, etc. Após o tratamento no MZmine, os dados (área dos picos) do modo positivo e negativo foram exportados separadamente em formato .csv e transferidos ao softwares R 3.3.0 (R Foundation for Statistical Computing, Áustria), SIMCA 13.0.3.0 (Umetrics, Suécia), Weka 3.6 (University of Waikato, Nova Zelândia), STATISTICA 8 (StatSoft, Inc., EUA) e The Unscrambler X (CAMO, Finlândia) onde se realizaram análises estatísticas multivariadas supervisionadas e não supervisionadas e análises de correlação *in silico* com dados taxonômicos e geográficos.

Inicialmente foram efetuadas duas análises não supervisionadas: 1) análise por componentes principais (PCA por suas siglas em inglês), que permite reduzir o número de variáveis do conjunto de dados mediante a criação de novas variáveis chamadas "componentes principais" que descrevem a variabilidade total do sistema em forma decrescente, permitindo determinar tendências e *outliers* (BRERETON, 2003); 2) análise hierárquica de *clusters* (HCA por suas siglas em inglês) que permite agrupar objetos de acordo com sua similitude, obtendo-se como resultado

um gráfico em forma de dendrograma (BRERETON, 2003). Nessa última análise foram testados vários métodos de agrupamento e distâncias de medição, obtendo resultados satisfatórios com o método Ward's e a distância euclidiana ao quadrado. O método Ward's, diferente de outros também comumente utilizados (single, complete ou average linkage), determina a variância total entre os agrupamentos, priorizando aqueles que levam ao menor aumento na variância total do sistema (WARD, 1963). Este método tem sido utilizado em diferentes tipos de dados, incluindo variáveis contínuas e variáveis binárias, com resultados favoráveis. Por exemplo, Hands; Everitt (1987) testaram cinco técnicas de agrupamento hierárquico (single, complete, average e centroid linkage, e o método Ward's) em dados multivariados binários, encontrando os melhores resultados com o método Ward's, enquanto os outros métodos apresentaram resultados semelhantes, excetuando o método de *single linkage* que apresentou os piores resultados (HANDS; EVERITT, 1987). As análises por PCA e HCA foram efetuadas nos softwares R, SIMCA P e The Unscrambler X.

Após esta análise exploratória, empregaram-se três métodos supervisionados: 1) *Orthogonal Partial Least Squares – Discriminant Analysis* (OPLS-DA), a fim de identificar as variáveis X (picos cromatográficos) correlacionadas com a variável resposta Y (dados biogeográficos e taxonômicos das espécies), efetuado no software SIMCA P; 2) Árvores de decisão (AD), efetuado no software Weka e 3) *K-nearest neighbors* (KNN) no software STATISTICA. As análises por AD e KNN, mesmo como OPLS-DA, foram realizadas com a finalidade de construir modelos de predição geográfica de novas amostras de *Espeletia* baseados na composição química das espécies deste gênero. Todas as análises estatísticas multivariadas foram realizadas com os dados dos modos positivo e negativo separadamente.

### 3.5 Análise de tricomas

Além da análise das folhas íntegras, foi empregada outra metodologia visando determinar a composição química dos tricomas em *Espeletia*, além da sua correlação com dados geográficos. Nessa análise, os tricomas tectores presentes na superfície folhar das cinco espécies do gênero descritas na Tabela 1 foram separados manualmente com ajuda de uma pinça. Após a separação dos tricomas,

15 mg das folhas (sem tricomas) foram pulverizadas com nitrogênio líquido, capsula de porcelana e pistilo para serem logo tamisadas num tamis de 42 *mesh*. Os tricomas e as folhas sem tricomas foram extraídos e analisados seguindo os parâmetros descritos nos itens 3.3 e 3.4, respectivamente. Tal procedimento foi realizado com a finalidade de comparar a composição química dos tricomas, com a das folhas sem tricomas, para depois fazer sua correlação com dados geográficos.

### 3.6 Análise de correlação *in silico*

A correlação *in silico* com dados biogeográficos e taxonômicos foi realizada mediante uma análise por OPLS-DA no software SIMCA P considerando diferentes métodos de normalização e pré-tratamento de dados. Duas matrizes, uma com variáveis contínuas (matriz 1) e outra com variáveis binárias (matriz 2) foram empegadas nas análises. No primeiro caso, com a finalidade de determinar se existe similitude nos perfis cromatográficos das diferentes espécies de *Espeletia*, de acordo com sua origem geográfica ou proximidade taxonômica. No segundo, para determinar se existe uma similitude unicamente na presença ou ausência de determinados metabólitos nas diferentes espécies deste gênero de acordo com sua origem geográfica ou proximidade taxonômica.

Na matriz 1, em que as áreas dos picos foram consideradas como variáveis contínuas, empregou-se a normalização pelo método de Pareto, o qual é um tipo de escalonamento de dados que usa a raiz quadrada do desvio padrão como fator de escalonamento (Equação 1), assim mudanças grandes na concentração dos metabólitos são diminuídas mais que mudanças pequenas, permitindo que os metabólitos majoritários sejam menos dominantes quando comparados aos dados brutos. Este método apresenta a vantagem de manter a estrutura dos dados parcialmente intacta (VAN DEN BERG et al., 2006).

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{\sqrt{s_i}}$$

**Equação 1.** Formula de escalonamento pelo método de Pareto.

Na matriz 2, em que as áreas dos picos foram transformadas em variáveis binárias (presença/ausência), empregou-se a normalização pelo método de *Unit variance* (UVar), o qual usa o desvio padrão como fator de escalonamento (Equação 2), onde após a transformação é atribuído um desvio padrão de um a todos os metabólitos. Este método apresenta a vantagem de tornar todos os metabólitos igualmente importantes (VAN DEN BERG et al., 2006).

$$\tilde{x}_{ij} = \frac{x_{ij} - \bar{x}_i}{s_i}$$

**Equação 2.** Formula de escalonamento pelo método de *Unit variance.*

Além das vantagens mencionadas, ambos os métodos possuem desvantagens. Por exemplo, o escalonamento por Pareto é sensível a mudanças grandes na concentração dos metabólitos, enquanto o escalonamento por UVar pode aumentar consideravelmente erros de medição, onde, por exemplo, picos ruidosos terão a mesma importância que as substâncias majoritárias (VAN DEN BERG et al., 2006). No entanto, segundo van den Berg et al. (2006), todos os métodos de escalonamento apresentam suas vantagens e desvantagens, e a escolha do melhor método a ser empregado deve ser feita de forma criteriosa, de acordo com os dados, o método de análise multivariado a ser empregado posteriormente e os objetivos do estudo.

### 3.7 Desreplicação de extratos

A desreplicação dos extratos foi efetuada inicialmente no software MZmine, utilizando o abrangente banco de dados da família Asteraceae: *Library of Natural Products from Asteraceae* (LiNPA), do nosso grupo de pesquisas, o qual contem informações dos metabólitos isolados por membros do AsterBioChem, presentes no nosso banco de dados *in house* (AsterDB), além de dados da literatura. Adicionalmente, foi empregado o banco de dados específico da subtribo Espeletiinae (ChemEsp) desenvolvido neste projeto. Essa desreplicação preliminar, baseada unicamente na comparação da massa monoisotópica em alta resolução das diferentes moléculas, serviu de base para efetuar identificações mais “detalhadas” comparando os espectros de massas, espectros no UV, fórmula molecular e índice

de deficiência de hidrogênio (IDH) dos possíveis metabolitos identificados no MZmine com dados da literatura. A fórmula molecular e o IDH foram calculados no software Xcalibur e o *Dictionary of Natural Products* (DNP) e *SciFinder Scholar* foram empregados na identificação das substâncias não contidas nos bancos de dados acima mencionados.

A confirmação da possível identidade das substâncias desreplicadas foi realizada mediante a coinjeção de 21 substâncias previamente isoladas por membros do nosso grupo de pesquisas (Apêndice V), além de fazer análises por MS/MS no UHPLC-UV-MS empregando a célula de colisão (HCD) para tal propósito, com uma energia de colisão (NCE) de 35 unidades. Devido à impossibilidade de selecionar o íon percursor para fazer sua fragmentação neste equipamento, a análise foi realizada selecionando-se a faixa no tempo de retenção em que eluiu a substância de interesse. Os espectros de MS/MS obtidos foram comparados com dados da literatura e realizou-se a proposta do mecanismo de fragmentação de algumas moléculas não encontradas na literatura. Neste estudo, foi priorizada a identificação das substâncias correlacionadas com cada agrupamento.

### 3.8 Análise filogenética

Além da análise hierárquica de clusters (HCA) em que as espécies agruparam-se de acordo à similaridade na sua composição química, num enfoque netamente fenético, foi efetuada uma análise filogenética com dados químicos com a finalidade de determinar se as relações de parentesco evolutivo entre as diferentes espécies do gênero *Espeletia* podem ser desvendadas com base na sua composição química. Esta análise se baseou na premissa que os caracteres químicos e seus estados (presencia/ausência) refletem a história evolutiva deste gênero, permitindo a reconstrução das relações entre espécies ancestrais e espécies descendentes. Neste caso, a presença da mesma substância química em duas ou mais espécies seria devida a que ditos indivíduos compartilham um ancestral comum. No entanto, a presença de caracteres químicos (ou de qualquer outro tipo de caracter) em duas ou mais espécies pode ser devida também à evolução convergente.

Devido a que todos os métodos disponíveis para inferir relações filogenéticas tendem a simplificar a verdadeira complexidade dos processos evolutivos

(SCHMITT; BARKER, 2009), a busca da árvore filogenética que melhor explicasse a história evolutiva do gênero *Espeletia* com base em dados químicos foi efetuada mediante dois métodos, máxima parcimônia (MP) e *maximum likelihood* (ML), com a finalidade de avaliar a consistência dos resultados. O método de MP afirma que a melhor hipótese filogenética é aquela que requer o menor número de mudanças de caracteres. Assim, uma árvore filogenética a partir do princípio de MP tem o menor número de alterações evolutivas (SCHMITT; BARKER, 2009). Por outro lado, *likelihood* representa a probabilidade de que um evento que haja acontecido no passado gere um resultado específico. Logo, o método de ML envolve um procedimento para encontrar o valor de um ou mais parâmetros para um determinado modelo que faça que a probabilidade de observar os dados (i.e. matriz de presencia/ausência) atinja seu valor máximo (SCHMITT; BARKER, 2009). Os métodos de MP e ML foram efetuados nos softwares PAUP 4.0b 10 (SWOFFORD, 2002) e GARLI (ZWICKL, 2006), respectivamente.

A análise por MP se efetuou considerando buscas heurísticas completas com dez adições randômicas e mantendo uma árvore em cada passo. O máximo de árvores a serem salvas (MAXTREES) foi configurado em 100,000 usando a opção de aumentar automaticamente a 100, caso fosse necessário. Como algoritmo de *branch swapping* se empregou *tree-bisection-and-reconection* (TBR) e a espécie *Senecio nivoaureus* (Asteraceae) foi considerada como grupo externo. Após a seleção das árvores parcimoniosamente informativas pela busca heurística foi gerada uma única árvore de consenso estrito no mesmo software. Finalmente, foram realizadas análises por *bootstrap* e *jackknife* com a finalidade de determinar o grau de suporte dos agrupamentos observados na árvore de consenso estrito. Nessas análises se consideraram 1000 réplicas e um número máximo de árvores a serem salvas de um (MAXTREES = 1).

Na análise por ML se empregou o modelo Mkv, o qual foi desenvolvido para estabelecer filogenias a partir de dados morfológicos discretos (LEWIS, 2001) incluindo dados de presença/ausência. Nesta análise as colunas com dados constantes foram excluídas mediante a opção *standardvariable* e os parâmetros do software se mantiveram constantes, com a exceção do número de réplicas em cada busca que se configurou em 10, a fim de fazer buscas mais exaustivas. Efetuaram-se três análises independentes seguindo os mesmos parâmetros com a finalidade

de obter resultados consistentes determinados pela baixa variabilidade entre as réplicas e a ausência de melhores pontuações para novas árvores. A melhor topologia se considerou como árvore de *maximum likelihood*.

# 4 RESULTADOS E DISCUSSÃO

## 4.1 Banco de dados químicos de Espeletiinae

O banco de dados ChemEsp, desenvolvido neste período, contem a primeira revisão de literatura sobre a química do metabolismo secundário da subtribo Espeletiinae e, consequentemente, o primeiro banco de dados com tais informações. Nele, observa-se que dos 223 táxons que compõem a subtribo Espeletiinae (considerando categorias específicas e infraespecíficas), 42 espécies, duas subespécies, dois híbridos, uma variedade e uma forma (totalizando 48 táxons) foram estudadas do ponto de vista fitoquímico, sendo identificadas um total de 197 substâncias diferentes com 988 reportes. Na Figura 3, é apresentado um *screenshot* da tela do ChemEsp. O banco de dados em sua totalidade se encontra disponível eletronicamente em formato .sdf, o qual é representado em forma de uma tabela que apresenta múltiplas estruturas químicas, sendo compatível com o software MarvinView. Tal formato, além de permitir a visualização das informações contidas no banco individualmente (Figura 4) ou em forma de planilha (Figura 3), permite gerar estruturas em 3D (Figura 4), além do cálculo de diversos descritores estruturais como pKa, Log P, ponto isoelétrico, carga, polarizabilidade, estruturas ressonantes e grupos aceptores e doadores de ligações de hidrogênio, entre outras.

Este banco de dados, além de ter sido empregado para propósitos de desreplicação no nosso grupo de pesquisas, pode ser utilizado em futuros estudos quimioinformáticos envolvendo as estruturas químicas presentes nele para desenvolver estudos de *virtual screening*, *docking* molecular, relação quantitativa estrutura-atividade (QSAR) ou relação quantitativa estrutura-tempo de retenção (QSRR) em diversos sistemas cromatográficos, etc. Futuramente pretende-se disponibilizar este banco de dados para toda a comunidade científica em páginas da internet.

File Edit View Table Structure Tools Help

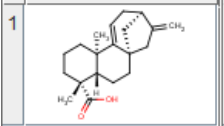
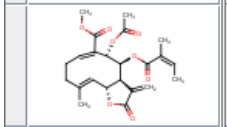
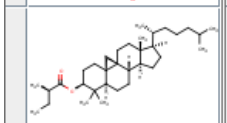
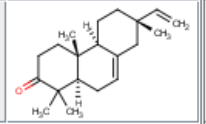
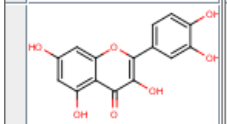
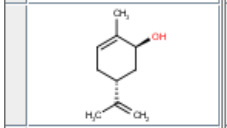
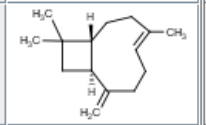
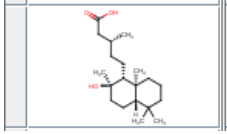
| # | structure | Species | Country of O... | Country of Co... | Locality of Collection | Voucher | Plant organ | Compound su... | Type of Skel... | Usual name | Sistematic na... | Molecular ... | CAS number | Monoisotopi... | Phytochemical Reference |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 1 |  | Espeletiopsis angustifolia (Cuatrec.) Cuatrec. | Venezuela | Venezuela | Páramo de San José de Acequias (8° 21.002 N, 71° 18.443 W) (2008-2870 m.s.n.m.) Merida | Merf Herbarium (LBR 040) | Leaves | Diterpene | Kaurane | Grandiflorenic Acid | Kaura-9(11),16- acid, (4?)- | C20H28O2 | 22338-67-6 | 300.208930 | Meccia et al., 2010 |
| |  | Espeletia killipii Cuatrec. | Colombia | Colombia | Páramo de Chingaza (3000m.s.n.m.) Cundinamarca | Herbario Nacional Colombiano 321751 | Aerial Parts | Sesquiterpen-lactone | Melampolide | Polymatin B | Cyclodeca[b]fu 5-(acetyloxy)-2, methyl ester, (? | C23H28O8 | 72023-37-1 | 432.178420 | Torrenegra & Tellez, 1995; Torrenegra et al., 1994; Torrenegra & Tellez, 1996 |
| |  | Espeletia barclayana Cuatrec. | Colombia | Colombia | Páramo El Tablazo, Subachoque, Cundinamarca | Herbario Nacional de Colombia 332528 | Leaves | Triterpene | Cycloartane | 9,19-Cyclolanostan-3-ol, (2R)-2-methyl butanoate | 9,19-Cycloland (2R)-2-methylbu (3?)- (9Cl | C35H60O2 | 207606-42-6 | 512.459330 | Tellez et al., 1998 |
| |  | Espeletia pycnophylla Cuatrec. | Colombia, Ecuador | Ecuador | Guandera Biological Station in the Eastern Cordillera (0°35' N, 77°41' W), a 11km de la ciudad de San Gabriel | Not reported | Roots | Diterpene | Isopimarane | 13B-methyl-13-vinyl-podocarb-7-en-3-one | 2(1H)-Phenanth 7-ethenyl-3,4,4 (4aR,4bS,7S,1 | C20H30O | 7715-48-2 | 300.245320 | Jansen et al., 2007 |
| |  | Espeletia barclayana Cuatrec. | Colombia | Colombia | Páramo El Tablazo, Subachoque, Cundinamarca | Herbario Nacional de Colombia 332528 | Leaves | Flavonoid | Flavonol | Quercetin | 4H-1-Benzopyra 2-(3,4-dihydrox | C15H10O7 | 117-39-5 | 302.042650 | Gutierrez et al., 1998 |
| |  | Ruilopezia atropurpurea (A. C. Sm.) Cuatrec. | Venezuela | Venezuela | Paramo El Batallon (3400 m.s.n.m.) | Merf Herbarium (RA 03, RA 05) | Leaves | Monoterpene | Monocyclic | trans-Carveol | 2-Cyclohexen-1 2-methyl-5-(1-m (1R,5S)-rel- | C10H16O | 1197-07-5 | 152.120120 | Aparicio et al., 2001 |
| |  | Libanothamnus occultus ssp. humbertii Cuatrec. | Venezuela | Venezuela | Laguna Negra (3200 m.s.n.m.) | Merf Herbarium (Usubillaga 10) | Leaves | Sesquiterpene | Bicyclic | B-Caryophyllene | Bicyclo[7.2.0]u 4,11,11-trimeth (1R,4E,9S)- | C15H24 | 87-44-5 | 204.187800 | Usubillaga et al., 2001a |
| |  | Espeletiopsis muiska (Cuatrec.) Cuatrec. | Colombia | Colombia | Via Tunja-Arcabuco 3200m.s.n.m., Boyacá | Herbario Nacional Colombiano 374424 y 374425 | Leaves | Diterpene | Labdane | (+)-8a-Hydroxy-labdan-17-oic Acid | 1-Naphthalene decahydro-2-hy (?R,1R,2R,4aS | C20H36O3 | 86189-98-2 | 324.266450 | Ramirez et al., 1998; Ramirez et al., 2000 |

**Figura 3.** *Screenshot* da tela do banco de dados ChemEsp no software MarvinView.

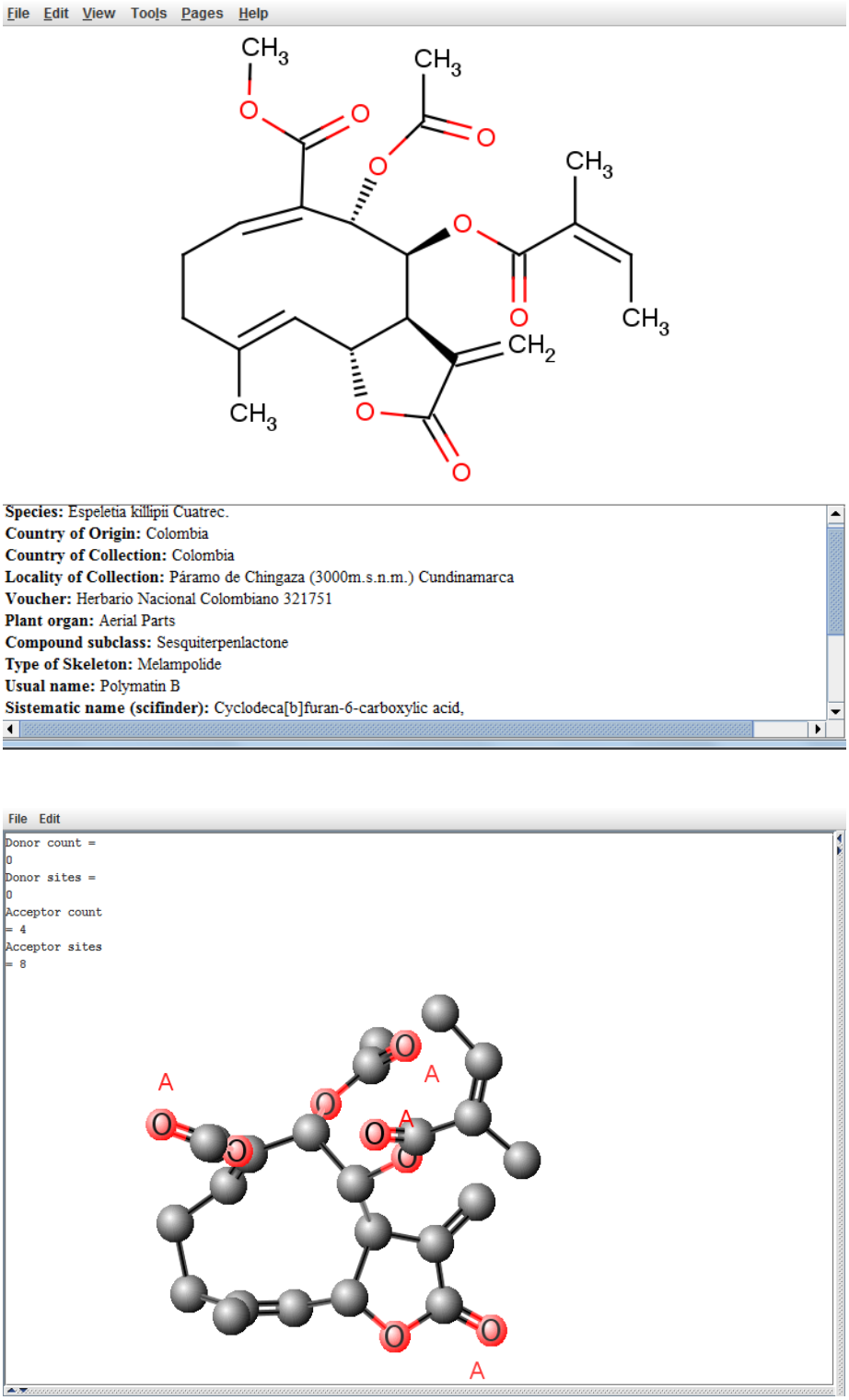



**Figura 4.** Visualização de estrutura química individual em 2D com informaçoes específicas (acima) e sua estrutura química em 3D com potenciais grupos aceptores (A) e doadores de ligações de hidrogênio (abaixo), empregando-se o software MarvinView.

## 4.2 Revisão da química de Espeletiinae

A principal classe de metabólitos secundários presentes na subtribo Espeletiinae corresponde aos terpenóides, representando 87% das substâncias identificadas (Figura 5). Dentre elas, os sesquiterpenos (32%), diterpenos (25%) e monoterpenos (25%) correspondem às subclasses de metabólitos mais comumente descritas com 63, 50 e 49 substâncias, respectivamente. Outras classes como flavonoides (2%) e lactonas sesquiterpênicas (1%) foram identificadas em menores proporções (Figura 5).

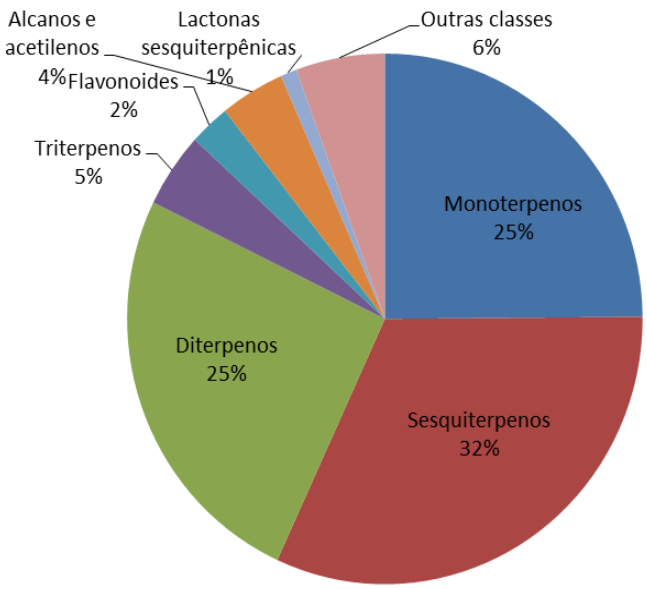



**Figura 5.** Classes de metabólitos secundários descritas na subtribo Espeletiinae.

De acordo com os dados da literatura, os diterpenos do tipo caurano representam o tipo estrutural mais comum de metabólitos na subtribo Espeletiinae. Tais substâncias, consideradas de valor quimiotaxonômico em *Espeletia* (ALVARENGA et al., 2005), têm uma ampla distribuição na subtribo, sendo descritas em 46 dos 48 táxons estudados do ponto de vista fitoquímico. As espécies *Espeletia argentea* Humb. & Bonpl. e *E. pycnophylla* Cuatrec., são as únicas onde este tipo de metabólitos não foi encontrado (PADILLA-GONZÁLEZ et al., submetido).

De acordo com Padilla-González et al. (submetido) a classificação dos gêneros *Ichthyothere* e *Smallanthus* (Asteraceae – Millerieae) como grupos irmãos da subtribo Espeletiinae com base em diversos marcadores moleculares (RAUSCHER, 2002) é suportada pela química do seu metabolismo secundário considerando a similitude na composição química entre ambos os grupos de espécies.

Uma análise por HCA efetuada com os dados da literatura no software R permitiu identificar dois agrupamentos com base na similitude da composição química das espécies (Figura 6). O primeiro (em azul) contendo espécies principalmente de origem colombiana e o segundo (em vermelho) contendo espécies todas de origem venezuelana, resultado que concorda com a filogenia atual da subtribo baseada em diferentes caracteres moleculares (DIAZGRANADOS, 2012). No entanto, é importante ressaltar que 13 espécies de origem venezuelana foram agrupadas junto com as espécies de origem colombiana no primeiro agrupamento (em azul, Figura 6), possivelmente devido aos poucos dados sobre metabólitos secundários na maioria delas (PADILLA-GONZÁLEZ et al., submetido). Tal hipótese foi confirmada no atual estudo, já que com base em dados metabolômicos, não só se atingiu um melhor agrupamento das espécies de acordo com seu país de origem, mas também a espécie *Espeletia tenorae* Aristeg, agrupada erroneamente no dendograma obtido com dados da literatura, foi agrupada corretamente com base nos dados metabolômicos.

Os resultados deste trabalho foram apresentados na *4th Brazilian Conference on Natural Products - 2013* (PADILLA-GONZÁLEZ et al., 2013) e submetidos em 2014 para publicação na modalidade de artigo de revisão.

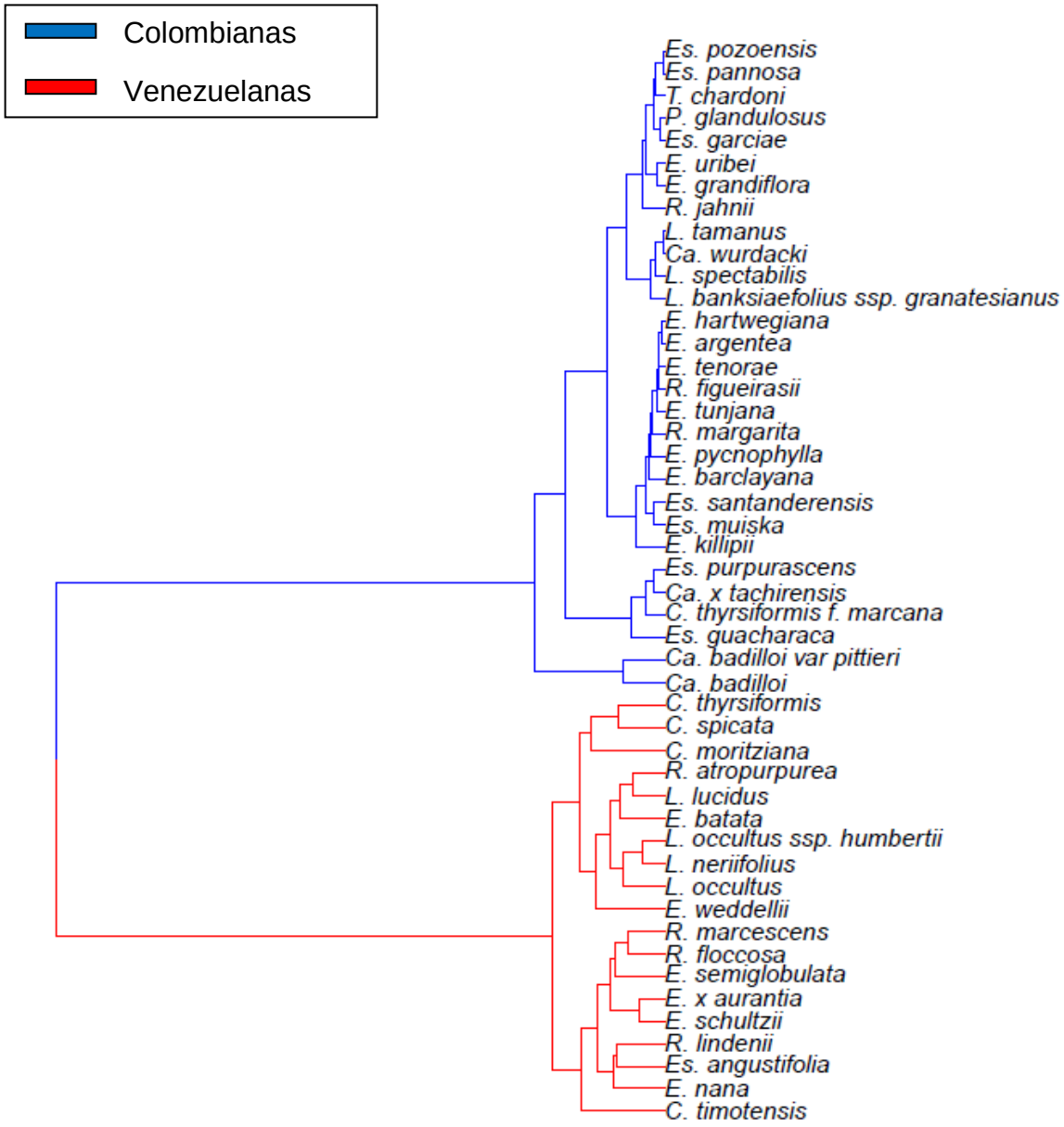



**Figura 6.** Análise hierarquica de clusters (HCA) da subtribo Espeletiinae obtida com os metabólitos secundários descritos na literatura. Em azul, espécies principalmente de origem colombiana; em vermelho, espécies de origem venezuelana. Abreviações dos gêneros: Ca., *Carramboa*; Co., *Coespeletia*; E., *Espeletia*; Es., *Espeletiopsis*; L., *Libanothamnus*; R., *Ruilopezia*; T., *Tamania*; P., *Paramiflos*.

## 4.3 Impressão digital metabólica e análise estatística multivariada

### 4.3.1 Análise de extratos frescos e secos

A análise qualitativa dos cromatogramas e espectros de massas dos extratos frescos e secos, obtidos a partir da analise por UHPLC-UV-MS das cinco espécies descritas na Tabela 1, permitiu determinar que o processo de secagem dos extratos influencia a composição química dos mesmos. Considerando que não só houve variação na intensidade de alguns picos, como também em alguns casos picos novos surgiram nos cromatogramas obtidos a partir dos extratos secos (setas vermelhas, Figuras 7-13), indicando uma possível degradação das substâncias presentes nos extratos frescos ou formação de artefatos. Por exemplo, na amostra A surgiu um pico em 9,41 min após a secagem no *SpeedVac* (seta vermelha, Figura 12). Esse pico está também presente na amostra seca no fluxo de ar comprimido (seta vermelha, Figura 7), com tempo de retenção de 8,27 min. Cabe ressaltar que após a análise detalhada do espectro de massas do pico intenso que surgiu na amostra A em aprox. 21 min, após secagem no fluxo de ar e no *SpeedVac* (Figuras 7 e 12, respectivamente), concluiu-se que tal pico corresponde ao etileno glicol por apresentar um espectro de massas característico de um polímero com perdas sucessivas de 44 *m/z.* O etileno glicol é um contaminante comum em espectrometria de massas e acredita-se que tal substância tenha vindo do tubo Eppendorff, considerando que no processo de secagem o extrato ficou em contato com este tubo por tempo prolongado e no caso do *SpeedVac* com aquecimento de 40 °C.

Além das análises qualitativas, análises estatísticas multivariadas não supervisionadas como PCA e HCA foram efetuadas no software The Unscrambler X. Nessas análises, os dados gerados pelo UHPLC-UV-MS foram tratados no software Mzmine segundo a metodologia descrita no item 3.4 e no Apêndice II.

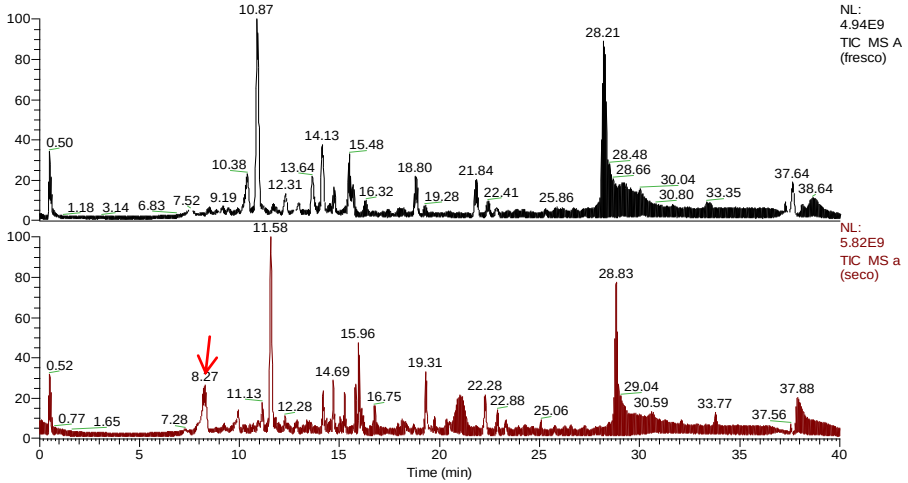



**Figura 7.** Cromatograma da amostra A, extrato fresco (acima) e seco no fluxo de ar (abaixo). A seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.

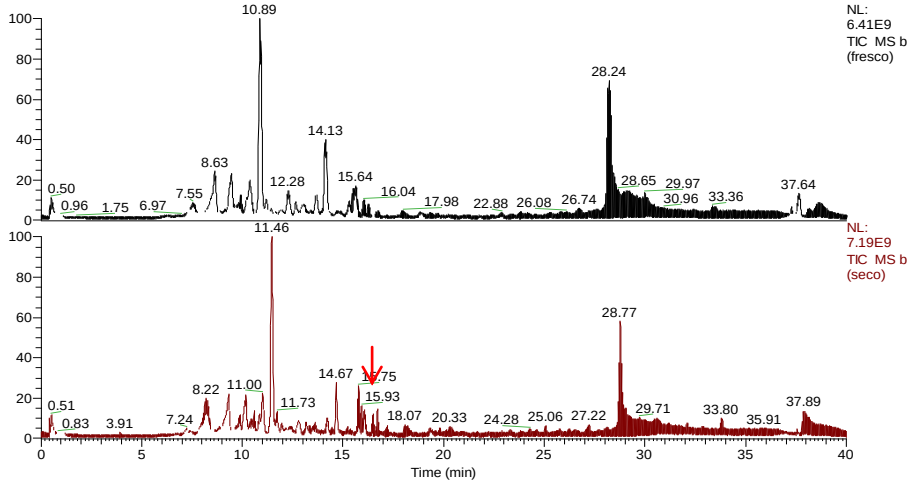



**Figura 8.** Cromatograma da amostra B, extrato fresco (acima) e seco no fluxo de ar (abaixo). A seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.

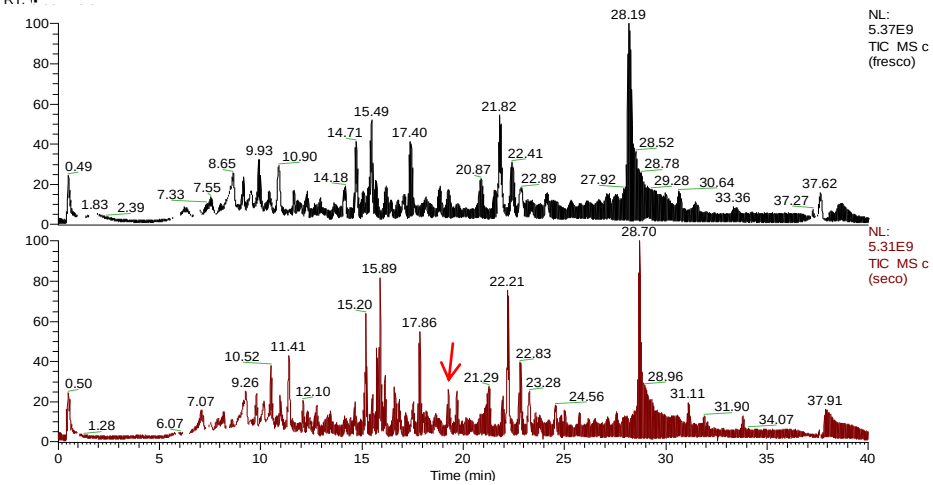



**Figura 9.** Cromatograma da amostra C, extrato fresco (acima) e seco no fluxo de ar (abaixo). A seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.

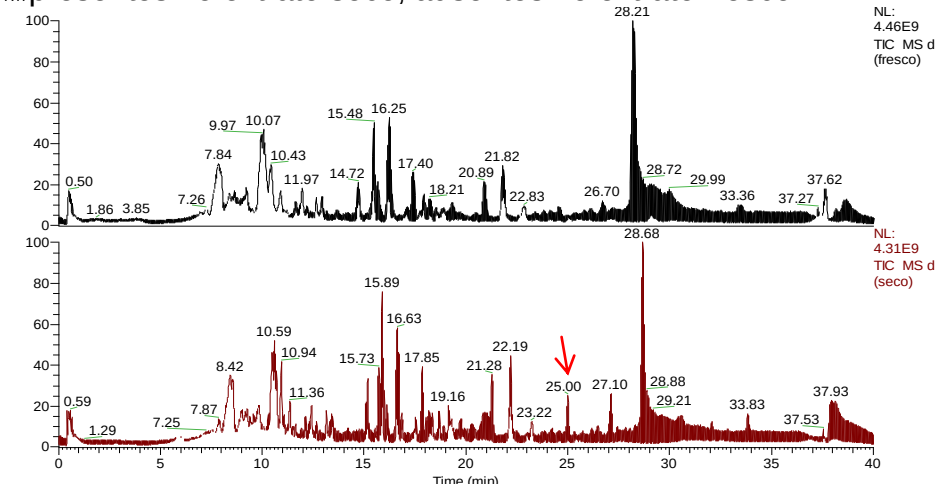



**Figura 10.** Cromatograma da amostra D, extrato fresco (acima) e seco no fluxo de ar (abaixo). A seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.

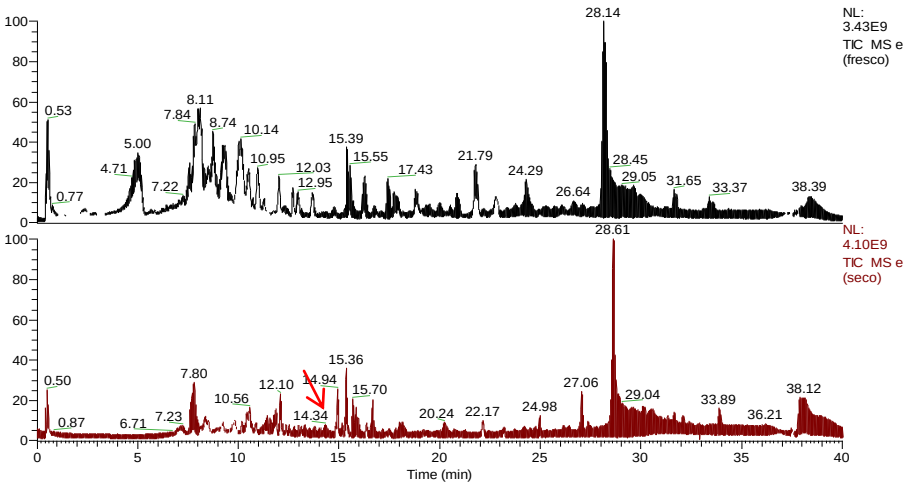



**Figura 11.** Cromatograma da amostra E, extrato fresco (acima) e seco no fluxo de ar (abaixo). A seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.

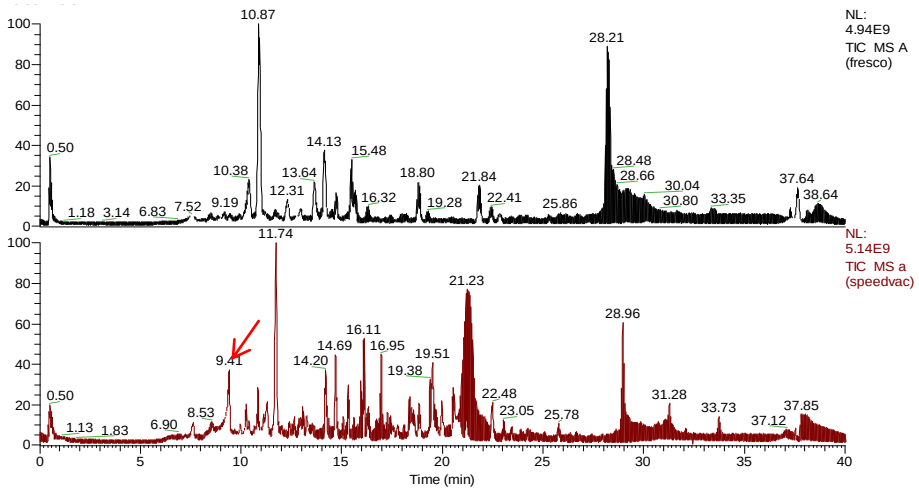



**Figura 12.** Cromatograma da amostra A, extrato fresco (acima) e seco no *SpeedVac* (abaixo). A Seta vermelha indica os picos presentes no extrato seco, ausentes no extrato fresco.

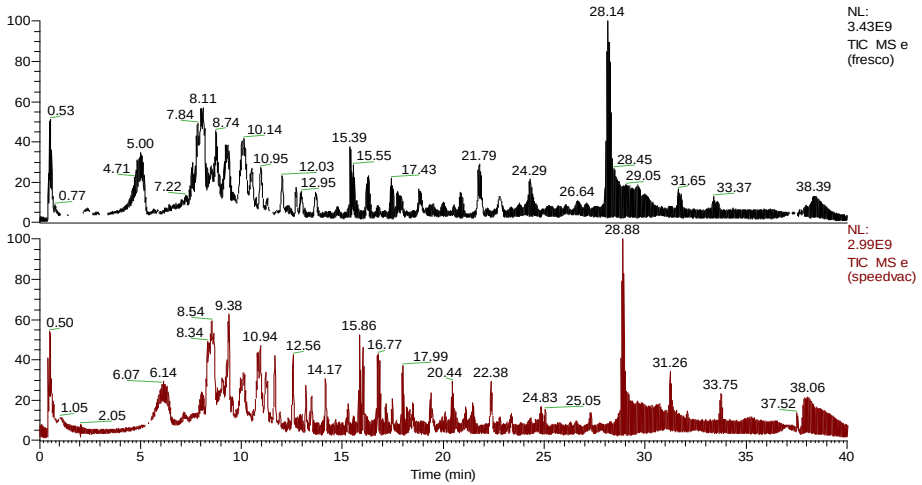



**Figura 13.** Cromatograma da amostra E, extrato fresco (acima) e seco no *SpeedVac* (abaixo).

Na análise por HCA dos extratos frescos (Figura 14) observa-se que o agrupamento das espécies com base nos seus constituintes químicos concorda com a análise por PCA (Figura 15A) e com a origem geográfica das mesmas (Figura 15B), onde as espécies *E. grandiflora* e *E. barclayana* (amostras A e B, respectivamente) de origem colombiana se agruparam juntas numa distância de agrupamento de 5, mesmo como as espécies *E. jajoensis* e *E. schultzii* (amostras D e E) de origem venezuelana (distância de agrupamento de 8). A espécie *E. frontinoensis* (amostra C) também de origem colombiana, porém proveniente de outro estado (Tabela 1), agrupou-se sozinha com a maior distância de agrupamento. Na análise por HCA a distância de agrupamento das observações corresponde ao grau de similaridade entre as mesmas. Logo, neste caso, distâncias de agrupamento menores indicam uma alta similaridade na composição química das espécies e à medida que esta distância aumenta, aumenta também o grau de divergência na composição química das mesmas.

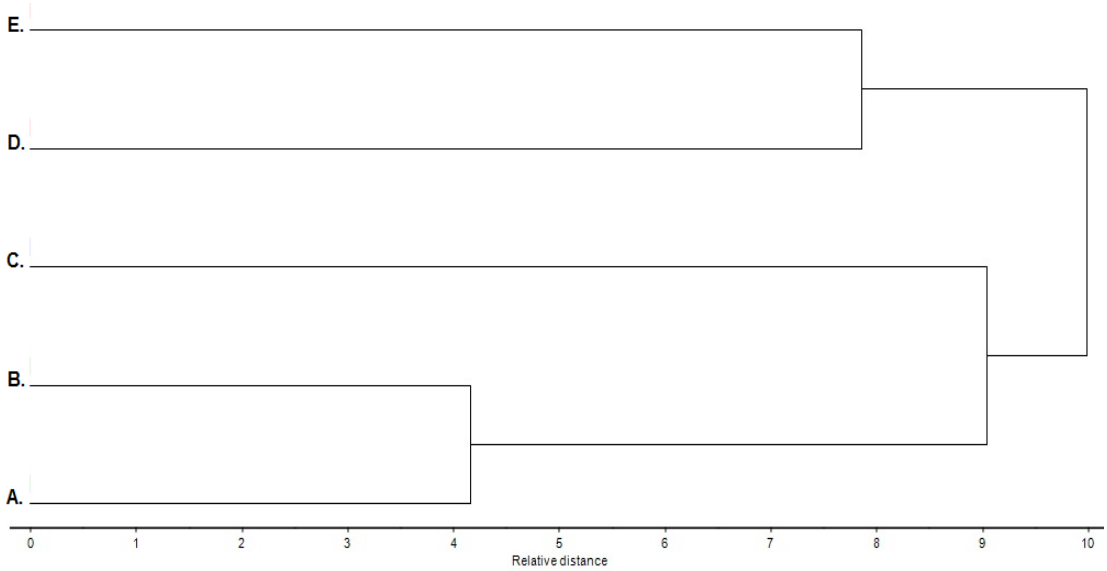



**Figura 14.** Análise hierárquica de clusters (HCA) do extrato fresco de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*.

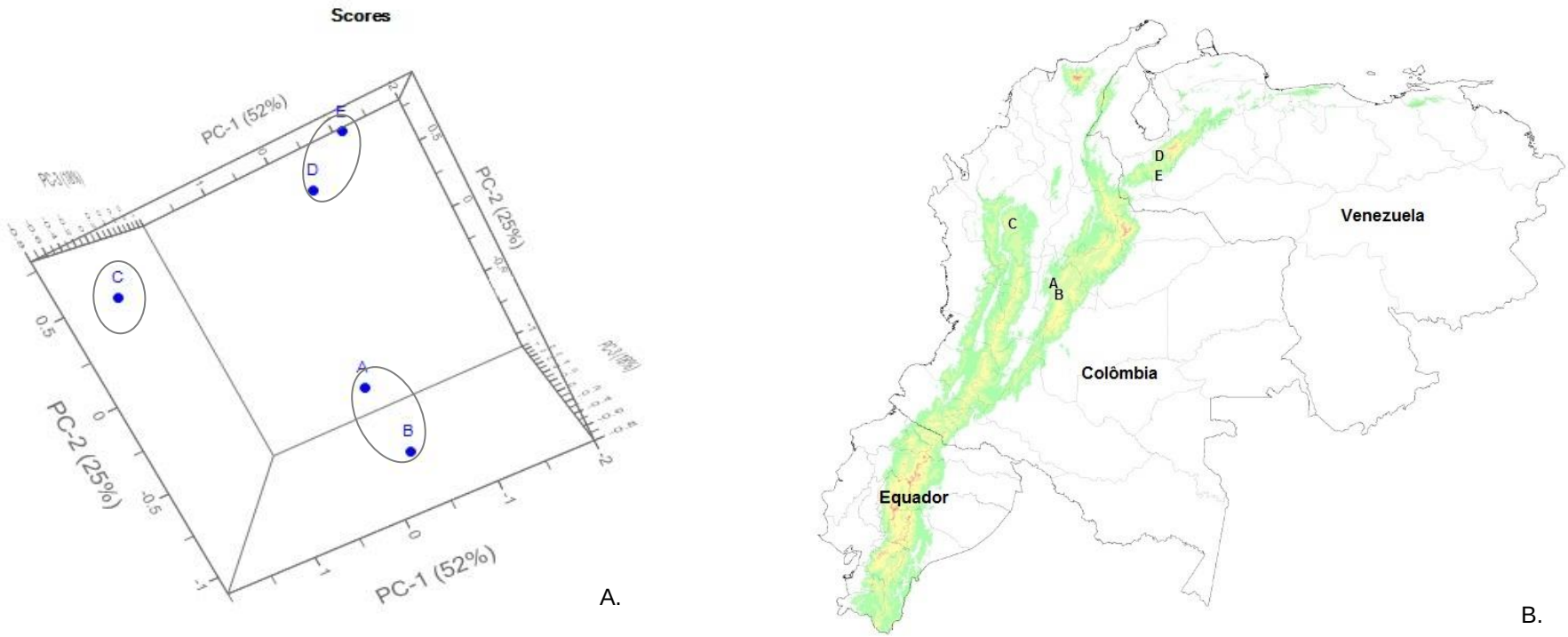



**Figura 15. A**. *Score plot* da análise por componentes principais (PCA) de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*; **B**. Mapa de distribuição geográfica das cinco espécies acima mencionadas (Adaptado de DIAZGRANADOS, 2012).

Ao comparar o resultado do HCA do extrato fresco (Figura 14) com o HCA do extrato seco no fluxo de ar (Figura 16), observa-se que houve diferenças no agrupamento das espécies, já que nesta última análise a espécie *E. frontinoensis* (amostra C), de origem colombiana, formou um cluster com *E. jajoensis* (amostra D), de origem venezuelana, enquanto a amostra E (*E. schultzii*), também venezuelana, ficou afastada dos dois clusters observados.

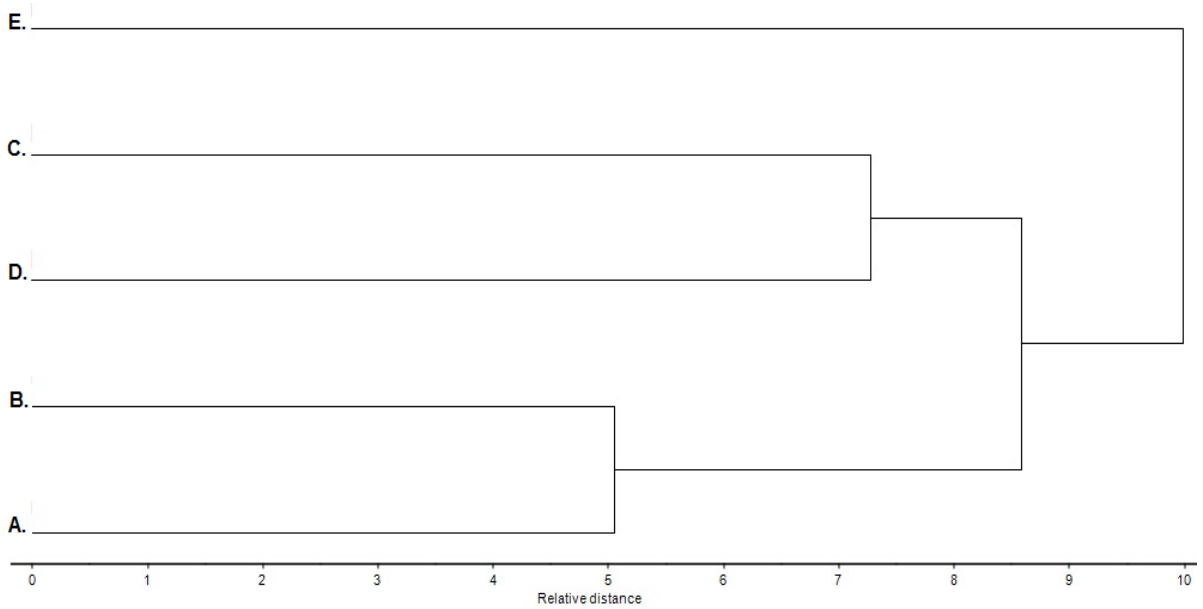



**Figura 16.** Análise hierárquica de clusters (HCA) do extrato seco no fluxo de ar comprimido de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*.

Os resultados obtidos nesse estudo permitiram concluir que o processo de secagem dos extratos altera a composição química dos mesmos de forma significativa, levando a agrupamentos diferentes. Logo, na obtenção da impressão digital metabólica das 120 amostras de *Espeletia* foram analisados somente os extratos frescos a fim de evitar uma possível mudança na sua composição química gerada pelo processo de secagem.

### 4.3.2 Análise de tricomas

A análise por HCA dos tricomas das cinco espécies já descritas (Figura 17) permitiu obter resultados semelhantes aos das análises dos extratos frescos (Figura 14), onde as espécies se agrupam de acordo com sua origem geográfica (Figura 15B). Porém, na análise dos tricomas, as espécies se agruparam numa distância

muito menor quando comparada à distância de agrupamento dos extratos frescos. Por exemplo, as espécies *E. grandiflora* e *E. barclayana* (amostras A e B, respectivamente) formaram um cluster numa distância de agrupamento de aproximadamente dois (Figura 17), enquanto as mesmas espécies se agruparam numa distância superior a quatro na análise dos seus extratos frescos (Figura 14). Igualmente as espécies *E. schultzii* e *E. jajoensis* (amostras D e E, respectivamente) se agruparam numa distância menor na análise dos seus tricomas quando comparadas às análises dos seus extratos frescos.

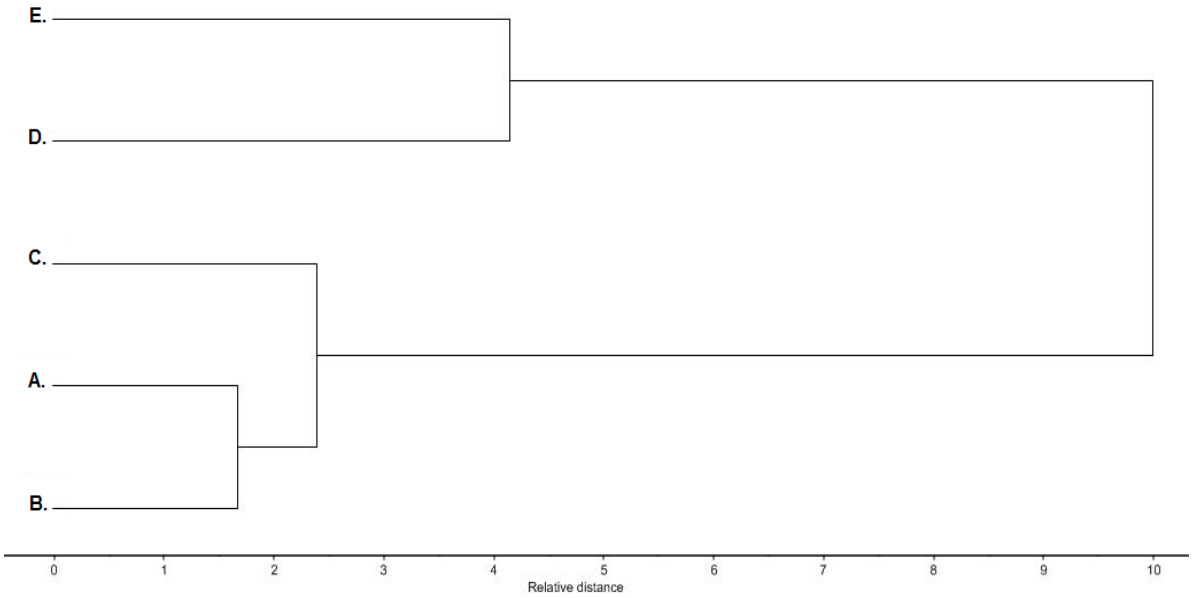



**Figura 17.** Análise hierárquica de clusters (HCA) dos tricomas de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*.

O fato das amostras não só terem mantido a correlação com sua origem geográfica na análise dos seus tricomas, mas também terem se agrupado mais proximamente do que nas análises dos extratos frescos, indica que possivelmente as substâncias responsáveis pelo agrupamento das espécies de acordo a sua origem geográfica estão presentes principalmente nos tricomas dessas espécies e não nas folhas. Este resultado é confirmado pela análise por HCA das folhas sem tricomas das mesmas espécies (Figura 18), onde se observa que ao excluir as substâncias presentes nos tricomas é perdida a correlação com a origem geográfica das mesmas. Nessa análise, observa-se também que embora as amostras A e B

tenham se agrupado juntas, a distância do seu agrupamento foi ainda maior do que na análise dos seus extratos frescos (Figura 14).

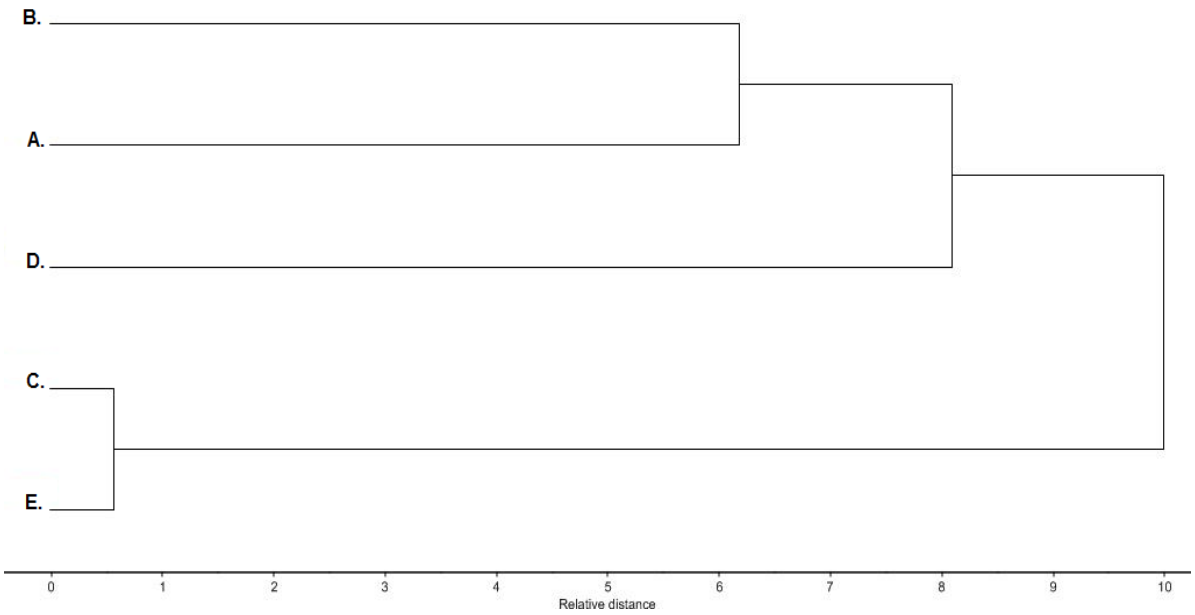



**Figura 18.** Análise hierárquica de clusters (HCA) das folhas sem tricomas de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*.

A análise por OPLS-DA dos tricomas e das folhas sem tricomas permitiu não só confirmar que existem diferenças na composição química de ambas as partes das plantas (Figura 19), mas também identificar as substâncias correlacionadas com cada agrupamento (Tabelas 2 e 3). Nos tricomas, as principais substâncias responsáveis pelo agrupamento das espécies foram diferentes flavonoides, todos eles derivados da quercetina, incluindo sua forma 3-metoxilada e 3-glicosilada, além da quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo (Tabela 2). Na análise das folhas sem tricomas (Tabela 3) observa-se que as principais substâncias correlacionadas foram diferentes ácidos clorogênicos, incluindo o ácido 5-*O*-(*E*)-cafeoilquínico e dois isômeros do ácido dicafeoilquínico (ácido 3,4-di-*O*-(*E*)-cafeoilquínico e ácido 4,5-di-*O*-(*E*)-cafeoilquínico), identificados mediante coinjeção com padrões. Além disso, vários diterpenos derivados do ácido *ent*-caurenóico, substância comum na subtribo Espeletiinae (PADILLA-GONZÁLEZ et al., submetido), foram também identificados.

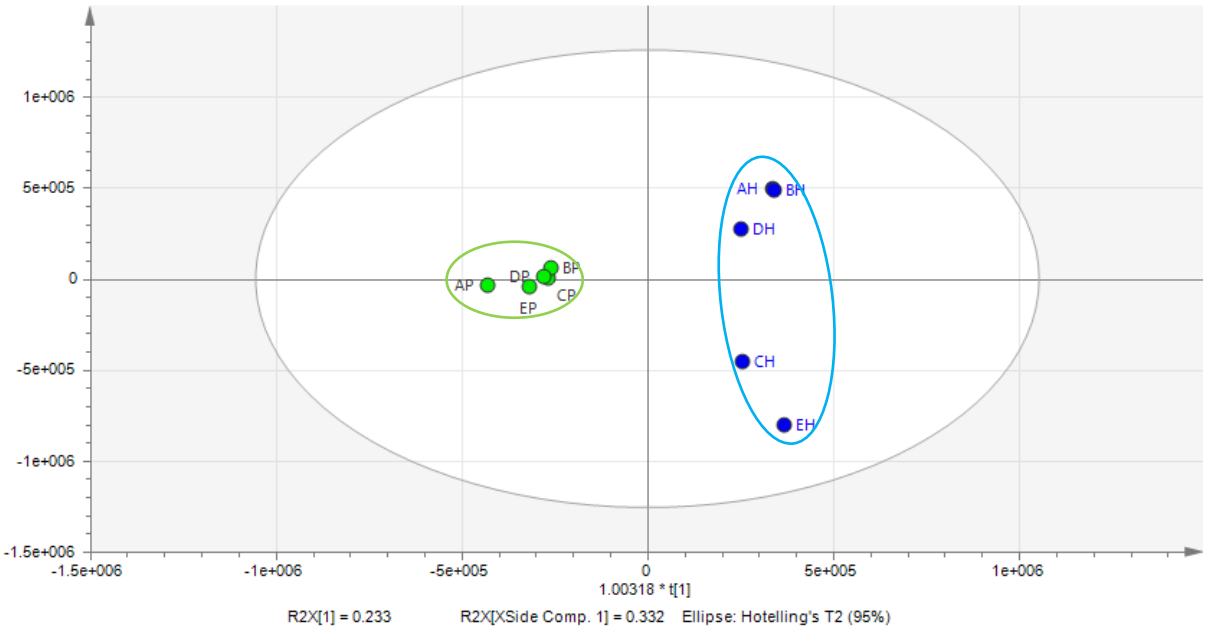



**Figura 19.** *Score plot* da análise por OPLS-DA dos tricomas (em verde) e das folhas sem tricomas (em azul) de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). (A) *E. grandiflora*, (B) *E. barclayana*, (C) *E. frontinoensis*, (D) *E. jajoensis*, (E) *E. schultzii*.

**Tabela 2.** Principais substâncias presentes nos tricomas de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS, de acordo com o modelo por OPLS-DA.

| | ID | *VIP* | *m/z* | $t_R$ (min) | Formula molecular | IDH | Classe | Substância provável |
|---|---|---|---|---|---|---|---|---|
| Positivo | 1 | 13.65 | 317.0658 | 10.94 | $C_{16}H_{12}O_7$ | 11 | Flavonoide | 3-Metóxiquercetina* |
| | 555 | 10.38 | 465.1034 | 7.95 | $C_{21}H_{20}O_{12}$ | 12 | Flavonoide | Quercetina-3-*O*-galactosídeo* |
| | 28 | 8.95 | 303.0500 | 10.41 | $C_{15}H_{10}O_7$ | 11 | Flavonoide | Quercetina* |
| | 1235 | 7.68 | 611.1401 | 10.12 | $C_{30}H_{26}O_{14}$ | 18 | Flavonoide | Flavonoide diglicosilado |
| | 4 | 7.01 | 617.0934 | 14.18 | $C_{31}H_{20}O_{14}$ | 22 | Flavonoide | 8,8''-Metileno bisquercetina |
| | 1128 | 5.78 | 627.1351 | 9.28 | $C_{30}H_{26}O_{15}$ | 18 | Flavonoide | Quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo |
| Negativo | 1 | 10.20 | 315.0510 | 10.9 | $C_{16}H_{12}O_7$ | 11 | Flavonoide | 3-Metóxiquercetina* |
| | 3 | 9.51 | 301.0360 | 10.4 | $C_{15}H_{10}O_7$ | 11 | Flavonoide | Quercetina* |
| | 76 | 8.65 | 463.0890 | 7.95 | $C_{21}H_{20}O_{12}$ | 12 | Flavonoide | Quercetina-3-*O*-galactosídeo* |
| | 229 | 6.79 | 609.1260 | 10.1 | $C_{30}H_{26}O_{14}$ | 18 | Flavonoide | Flavonoide diglicosilado |
| | 118 | 6.29 | 625.1210 | 9.27 | $C_{30}H_{26}O_{15}$ | 18 | Flavonoide | Quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo |
| | 4 | 5.16 | 615.0790 | 14.2 | $C_{31}H_{20}O_{14}$ | 22 | Flavonoide | 8,8''-Metileno bisquercetina |

* Substância confirmada por coinjeção com padrão.

A presença de flavonoides aparentemente só nos tricomas e não nas folhas de *Espeletia* sugere um possível papel adaptativo dessas substâncias como mecanismo de proteção contra os altos níveis de radiação ultravioleta característicos dos páramos, considerando que as flavonas e os flavonóis absorvem eficientemente radiação UV. Estudos prévios têm descrito que sua acumulação em células epidermais após serem expostas a altos níveis de radiação indica uma possível função como "escudos" protetores dos tecidos fotossintéticos internos (VAN TUNEN et al., 1988; SCHMELZER et al., 1988; KOES et al., 1994). No entanto, ainda são necessários estudos mais detalhados para confirmar esta hipótese em *Espeletia*.

**Tabela 3.** Principais substâncias presentes nas folhas sem tricomas de cinco espécies do gênero *Espeletia* analisadas por UHPLC-UV-MS, de acordo com o modelo por OPLS-DA.

| | ID | *VIP* | *m/z* | $t_R$ (min) | Formula molecular | IDH | Classe | Substância provável |
|---|---|---|---|---|---|---|---|---|
| Positivo | 84 | 11.42 | 517.1346 | 8.67 | $C_{25}H_{24}O_{12}$ | 14 | Ácido clorogênico | Ácido 3,4-di-*O*-(*E*)-cafeoilquínico* |
| | 5 | 10.70 | 315.1957 | 14.75 | $C_{20}H_{26}O_3$ | 8 | Diterpeno | Ácido *ent*-12-oxocauro-9(11),16-dien-18-oico |
| | 44 | 8.10 | 301.2164 | 18.84 | $C_{20}H_{28}O_2$ | 7 | N.I. | N.I. |
| | 33 | 6.52 | 333.2059 | 10.46 | $C_{20}H_{28}O_4$ | 7 | Ácido clorogênico | N.I. |
| | 168 | 5.64 | 517.1344 | 9.16 | $C_{25}H_{24}O_{12}$ | 14 | Ácido clorogênico | Ácido 4,5-di-*O*-(*E*)-cafeoilquínico* |
| Negativo | 9 | 17.95 | 515.11967 | 8.67 | $C_{25}H_{24}O_{12}$ | 14 | Ácido clorogênico | Ácido 3,4-di-*O*-(*E*)-cafeoilquínico* |
| | 52 | 13.13 | 515.11987 | 6.25 | $C_{25}H_{24}O_{12}$ | 14 | Ácido clorogênico | Isômero do ácido dicafeoilquínico |
| | 228 | 11.97 | 353.0880 | 0.85 | $C_{16}H_{18}O_9$ | 8 | Ácido clorogênico | Ácido 5-*O*-(*E*)-cafeoilquínico* |
| | 61 | 9.81 | 315.1970 | 15.50 | $C_{20}H_{28}O_3$ | 7 | Diterpeno | Ácido *ent*-12-hidroxicauro-9(11),16-dien-18-oico |
| | 55 | 9.59 | 515.11964 | 9.17 | $C_{25}H_{24}O_{12}$ | 14 | Ácido clorogênico | Ácido 4,5-di-*O*-(*E*)-cafeoilquínico* |
| | 75 | 8.67 | 313.1810 | 14.80 | $C_{20}H_{26}O_3$ | 8 | Diterpeno | Ácido *ent*-12-oxocauro-9(11),16-dien-18-oico |
| | 124 | 6.58 | 327.2180 | 12.05 | $C_{18}H_{32}O_5$ | 3 | Diterpeno | N.I. |
| | 48 | 6.21 | 347.0590 | 9.95 | $C_{13}H_{16}O_{11}$ | 6 | N.I. | N.I. |
| | 162 | 6.03 | 559.1100 | 7.62 | $C_{26}H_{24}O_{14}$ | 15 | Ácido clorogênico | N.I. |
| | 211 | 5.50 | 299.2020 | 21.80 | $C_{20}H_{28}O_2$ | 7 | N.I. | N.I. |
| | 227 | 5.35 | 331.1920 | 12.30 | $C_{20}H_{28}O_4$ | 7 | Diterpeno | Possível ácido hautriwaico |
| | 181 | 5.29 | 315.1970 | 17.06 | $C_{20}H_{28}O_3$ | 7 | Diterpeno | N.I. |

* Substância confirmada por coinjeção com padrão; N.I. não identificada.

Cabe ressaltar que embora a química do metabolismo secundário de espécies do gênero *Espeletia* tenha sido estudada durante os últimos 45 anos, este corresponde ao primeiro estudo fitoquímico do gênero, e da subtribo Espeletiinae, onde foi analisada a composição química dos tricomas separadamente das folhas.

### 4.3.3 Análise dos 120 táxons de *Espeletia*

O método cromatográfico descrito no item 3.4, empregado na obtenção da impressão digital metabólica dos 120 táxons de *Espeletia*, permitiu a obtenção de bons perfis cromatográficos tanto no modo negativo (Figura 20), quanto no modo positivo (não apresentado), considerando que os constituintes químicos de todas as espécies estão bem distribuídos por todo o cromatograma (Figura 20).

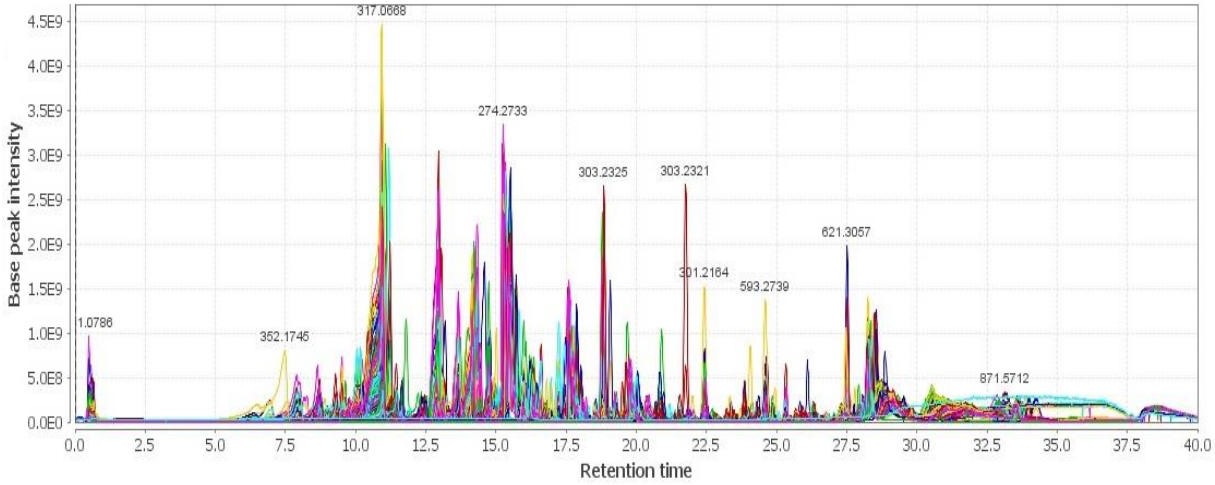



**Figura 20.** Cromatogramas de íons totais (TIC) sobrepostos de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo).

Após a obtenção dos dados cromatográficos e espectroscópicos, os mesmos foram exportados e tratados no software MZmine segundo os parâmetros descritos no Apêndice II. Posteriormente, os dados foram tratados numa planilha do Excel para subtrair os picos presentes na amostra branco, além de gerar outra matriz (matriz 2) com dados binários (presença/ausência) empregada também nas análises estatísticas efetuadas no software SIMCA P. Este procedimento foi realizado com os dados obtidos nos modos positivo e negativo separadamente.

A análise por PCA foi efetuada informando ao sistema o país de origem das espécies como variável “classe”, a fim de facilitar a visualização e interpretação dos resultados. Neste tipo de análises não supervisionadas, tal procedimento não

interfere no agrupamento das espécies, diferente das análises supervisionadas onde o país de origem das espécies foi considerado como variável resposta (variável Y).

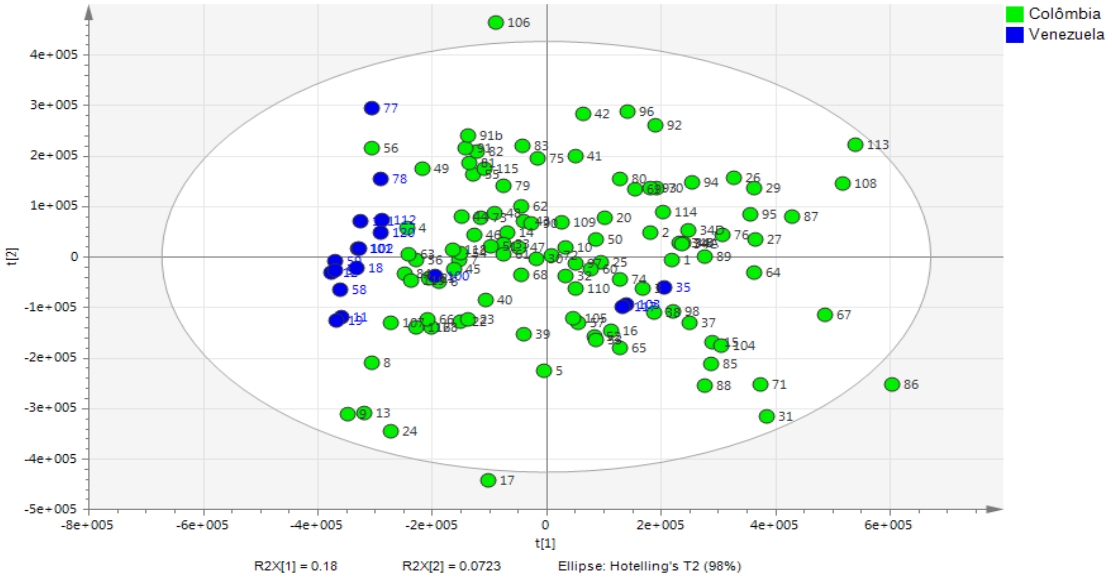



**Figura 21.** *Score plot* da análise por componentes principais (PCA) de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Países de origem das espécies indicados em diferentes cores.

Na análise por PCA dos dados obtidos no modo negativo (Figura 21), observa-se que a maioria das espécies de origem venezuelana têm uma correlação negativa com o primeiro componente principal, ficando agrupadas na parte esquerda do gráfico, enquanto as espécies colombianas estão espalhadas em todos os eixos, sugerindo uma composição química mais heterogênea. Das 120 amostras analisadas nenhuma foi identificada como *outlier*, caracterizados por ficarem longe da elipse (Hotelling´s T2) que indica a região definida com 98% de confiança da variação do modelo. Na PCA dos dados obtidos no modo positivo se obtiveram resultados semelhantes onde a maioria das espécies de origem venezuelana se agruparam juntas na parte esquerda do gráfico (dados não apresentados).

A análise por HCA (Figura 23) corrobora os resultados obtidos na PCA, além de permitir a visualização de vários dados de origem geográfica, nome científico e distância de agrupamento (informada como números nas linhas de bifurcação do dendograma) entre cada espécie e cluster. Nessa análise, as quatro réplicas da amostra 34 (caixa A, Figura 23) se agruparam juntas, sugerindo uma boa

reprodutibilidade e confiabilidade das análises cromatográficas, o que é conferido na sobreposição dos cromatogramas das mesmas amostras (Figura 22).

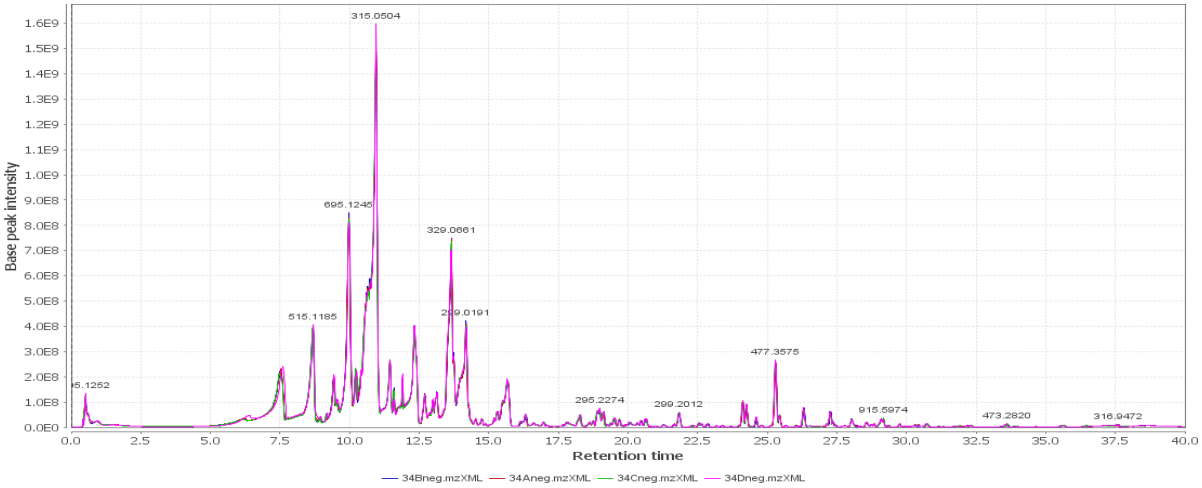



**Figura 22.** Cromatogramas de íons totais (TIC) sobrepostos das quatro réplicas da amostra 34 analisadas por UHPLC-UV-MS (modo negativo). (Ver códigos das amostras no apêndice I).

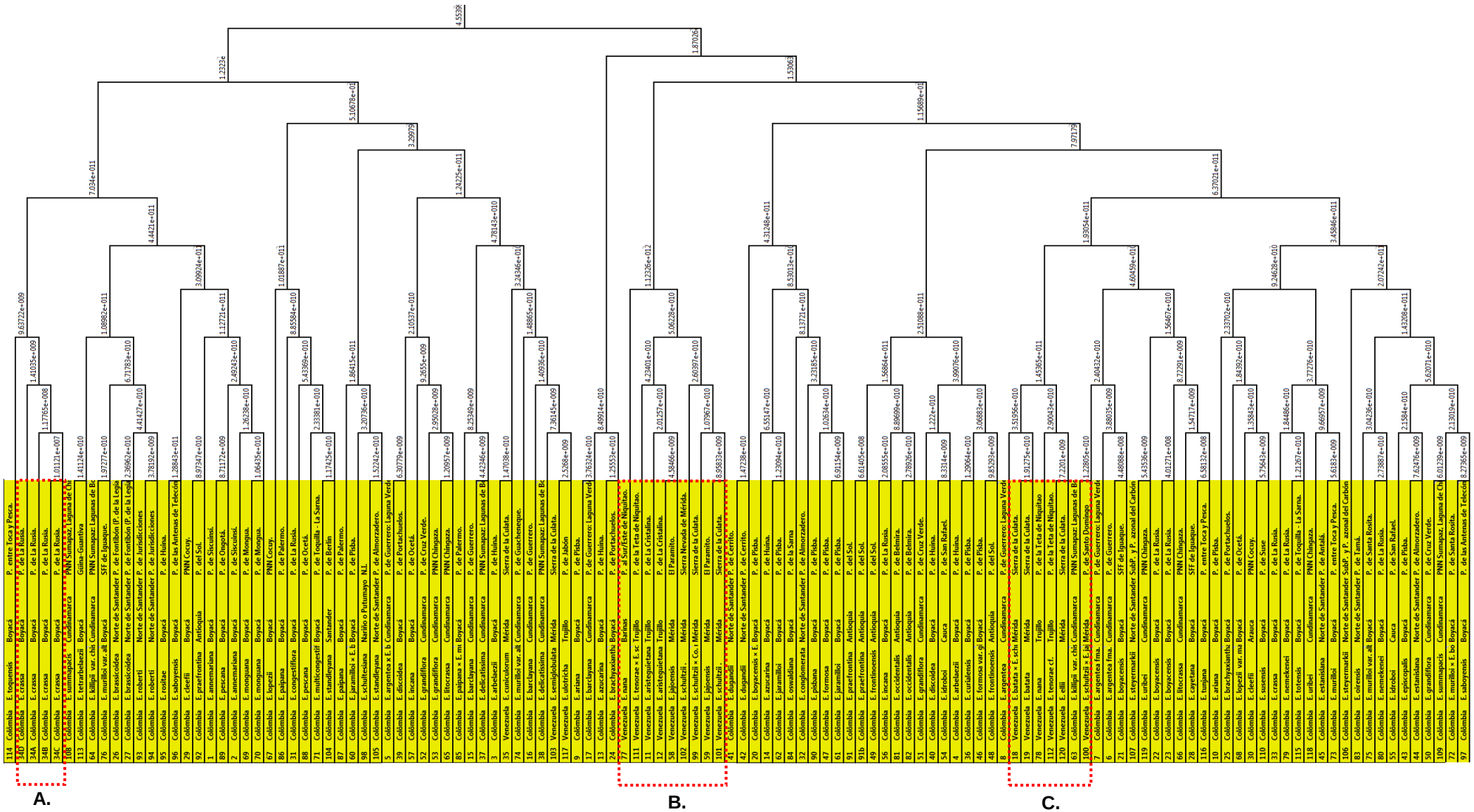



**Figura 23.** Análise hierárquica de clusters (HCA) de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Caixa A: réplicas da amostra 34; caixa B: primeiro agrupamento das espécies venezuelanas; caixa C: segundo agrupamento das espécies venezuelanas.

Na análise por HCA se observa certa correlação com a origem geográfica das espécies, onde a maioria das espécies de origem venezuelana se agruparam juntas, formando dois principais agrupamentos no meio do dendrograma (caixas B e C, Figura 23), indicando uma alta similaridade na composição química dessas espécies. No entanto, quatro espécies de origem venezuelana, *E. cuniculorum* (35), *E. semiglobulata* (103), *E. ulotricha* (117) e *E. schultzii* x *E. jajoensis* (100), se agruparam em *clusters* compostos principalmente por espécies de origem colombiana.

Além do agrupamento das espécies de acordo com seu país de origem, observa-se que houve também um agrupamento um pouco menos evidente de acordo com o estado em que cada espécie foi coletada. Por exemplo, no *cluster* venezuelano (caixa B, Figura 23), há dois principais agrupamentos correspondentes às espécies do estado de Mérida e Trujillo, enquanto a única espécie do estado de Barinas (*Espeletia nana* (77)) se encontra afastada destes dois agrupamentos (Figura 23). O agrupamento das espécies colombianas foi um pouco mais heterogêneo.

É importante mencionar que no dendograma gerado (Figura 23) as réplicas não se agruparam juntas em todos os casos, já que em alguns deles diferentes indivíduos da mesma espécie coletados na mesma época e área geográfica mostraram variação. Este resultado é análogo aos resultados obtidos em diversos estudos moleculares da subtribo Espeletiinae, onde se encontrou que a mesma espécie pode apresentar variação ou diferentes espécies podem apresentar a mesma sequência de alinhamento para alguns marcadores moleculares levando a conflito (RAUSCHER, 2002; SÁNCHEZ, 2005; DIAZGRANADOS, 2012). Por exemplo, na árvore filogenética da subtribo Espeletiinae construída com base nos marcadores ITS, ETS, e rp116 (Figura 1, Anexo I), observa-se que diferentes indivíduos da mesma espécie, como *E. boyacensis* 3654 e *E. boyacensis* 3650, mesmo como *E. pescana* 3652 e *E. pescana* 3659, apresentam variabilidade para os marcadores moleculares empregados.

Este resultado pode ser devido também a diferentes fatores externos, por exemplo, de tipo ambiental, já que a presença de pragas como insetos ou microrganismos patógenos pode induzir uma mudança na composição química das

plantas como mecanismo de defesa (WINK, 2003, 2006). Logo, se indivíduos da mesma espécie estiverem submetidos a diferentes pressões externas, podem acabar divergindo significativamente na sua composição química. Portanto, são necessários estudos mais aprofundados para corroborar se a variação química intraespecífica em *Espeletia* é devida a fatores externos ou se a química do metabolismo secundário deste gênero apresenta o mesmo padrão de variabilidade intraespecífica observado com marcadores moleculares.

## 4.4 Análise de correlação *in silico*

### 4.4.1 Correlação com dados biogeográficos

A correlação com dados biogeográficos foi realizada mediante uma análise supervisionada por OPLS-DA no software SIMCA P, com a finalidade de determinar se existem diferenças na composição química das espécies de acordo com sua origem geográfica, além de identificar as variáveis (metabólitos) correlacionadas com o agrupamento das espécies em cada cluster. Nesta análise se empregaram as duas matrizes de dados descritas anteriormente (matriz 1: variáveis contínuas e matriz 2: variáveis binárias) para a detecção em cada polaridade (modo positivo e negativo) de acordo com os objetivos descritos no item 3.6.

Na análise por OPLS-DA da matriz 1 do modo negativo (Figura 24) foi obtido um melhor agrupamento das espécies de acordo com sua origem geográfica quando comparada à análise por PCA ou HCA (Figuras 21 e 23, respectivamente), já que nenhuma espécie de origem venezuelana foi agrupada junto com as espécies colombianas e vice-versa.

Este resultado indica que realmente existem diferenças na composição química das espécies de acordo com seu país de origem, considerando que o modelo por OPLS-DA descreveu 96% da variabilidade na resposta Y (país de origem) ($R^2Y$=0,96), indicando uma excelente representação dos dados, além de conseguir predizer 80% da variabilidade na resposta Y ($Q^2$=0,80). O valor de $Q^2$, obtido por validação cruzada, é um dos critérios mais populares de validação onde o conjunto de dados é dividido em *n* partes (*folds*) iguais (sete no SIMCA P 13.0) e o sistema escolhe *n*-1 partes como conjunto de treino e uma como conjunto de teste, *n* vezes de forma independente. Logo, no SIMCA P, cada parte é usada uma vez como conjunto de teste e seis vezes no conjunto de treino. Finalmente, a média dos

sete testes corresponde ao valor de $Q^2$ (WAGNER et al., 2007; XIE et al., 2011). Resultados similares foram obtidos na análise do modo positivo, onde as espécies da Colômbia e da Venezuela formaram dois clusters independentes (dados não apresentados). Nessa análise foram atingidos valores ótimos de $R^2$ ($R^2Y$=0,89) e $Q^2$ ($Q^2Y$=0,75), confirmando os resultados obtidos na análise do modo negativo.

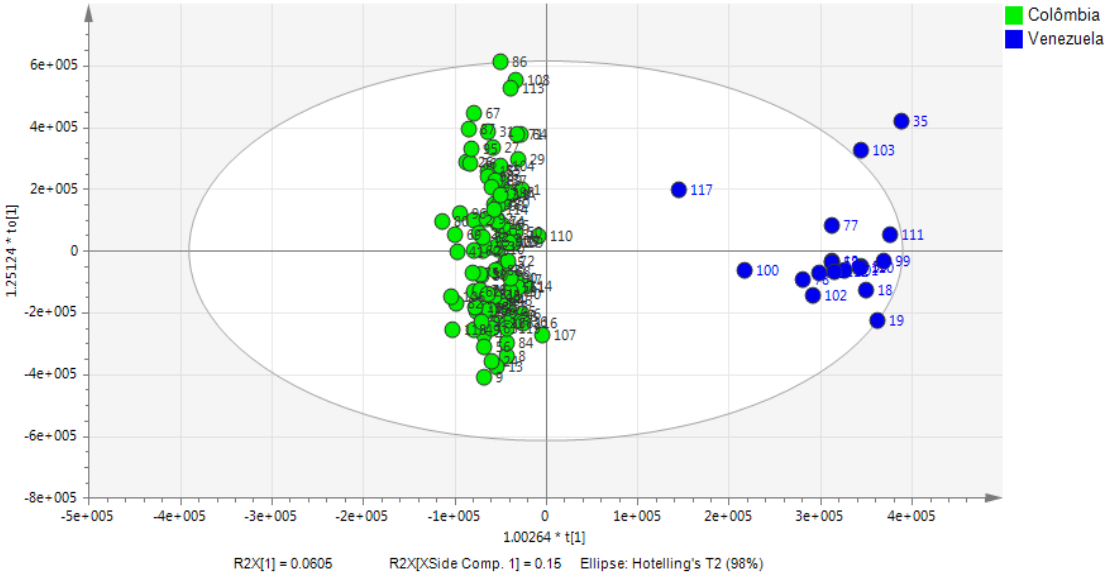



**Figura 24.** *Score-plot* da análise por OPLS-DA de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Países de origem das espécies indicados em diferentes cores.

O modelo por OPLS-DA permitiu também determinar as variáveis correlacionadas com os agrupamentos observados mediante a análise dos seus valores de VIP (Variable Importance in the Projection), os quais expressam a importância que tem cada variável tanto para explicar X quanto na correlação com Y, junto como a análise do *S-plot* (Figura 25) e os intervalos de confiança do *Loadings plot* (não apresentado) calculados pelo estimativo de *Jack-knife* (Equação 3), o qual é uma técnica de re-amostragem estatística em que cada observação é excluída uma vez da análise e a média dos cálculos obtidos é considerada tal estimativo (SEVERIANO et al., 2011).

$$\bar{\theta}_{\mathbf{Jack}} = \frac{1}{n}\sum_{i=1}^{n}(\bar{\theta}_i)$$

**Equação 3.** Estimativo de *Jack-knife.*

Esta técnica permite determinar a significância estatística das variáveis selecionadas e, junto com a análise do *S-plot* e o valor de VIP, permite ter uma confiança maior na determinação das variáveis mais importantes na separação dos grupos (SIMCA P tutorial). Neste estudo foram consideradas apenas as variáveis com valor VIP >3,0 (em vermelho, Figura 25), a fim de identificar apenas aquelas que estivessem influenciando o agrupamento em maior medida.

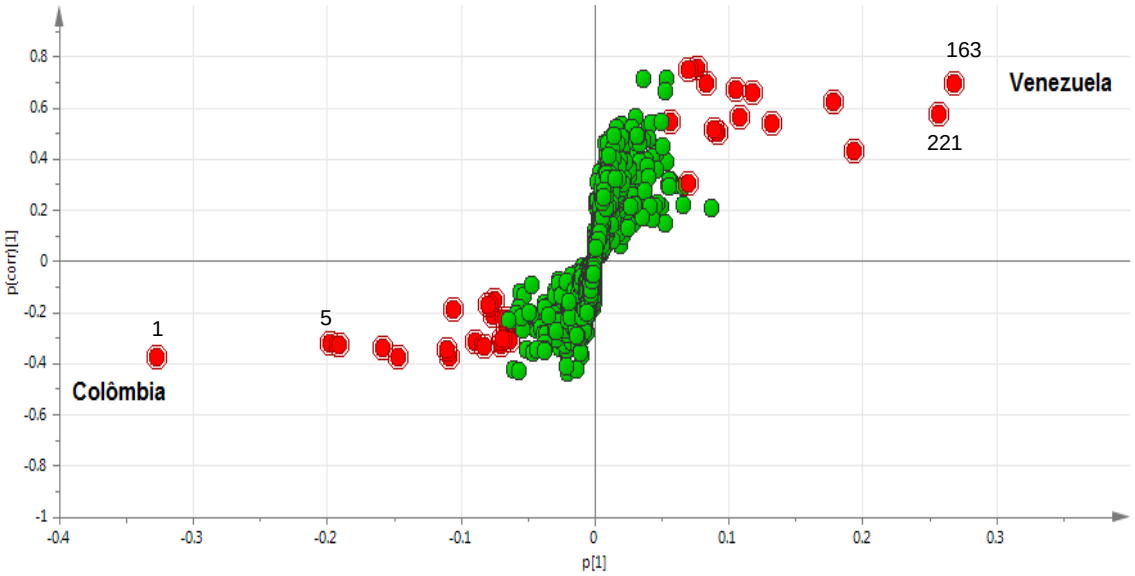



**Figura 25.** *S-plot* da análise por OPLS-DA de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Em vermelho as variáveis correlacionadas (VIP>3,0) com o agrupamento das espécies em cada país. Números indicando o ID das principais variáveis discriminantes.

Nas Tabelas 4 e 5 são apresentadas as variáveis correlacionadas com o agrupamento das espécies em cada cluster (colombiano e venezuelano), além do valor de VIP e o nome das substâncias desreplicadas, nos casos em que foi possível fazer a desreplicação das mesmas. Adicionalmente, outras informações como o valor de *m/z*, tempo de retenção, índice de deficiência de hidrogênio (IDH) e classe química com base no perfil de UV de cada substância, também são fornecidas.

Na Tabela 4, observa-se que as principais substâncias correlacionadas com o agrupamento das espécies no cluster colombiano correspondem a flavonoides metoxilados, dois quais se destacam a 3-metóxiquercetina, a hesperetina, e a 8,8''-metilenobisquercetina (Tabela 4) observadas tanto no modo positivo, quanto no modo negativo, além de outros flavonoides e ácidos clorogênicos não identificados. A 3-metóxiquercetina foi a principal substância discriminante do cluster colombiano,

de acordo com seu valor de VIP e posição no *S-plot* (ID 1, Figura 25). Este flavonoide foi identificado previamente como o componente majoritário dos tricomas de *Espeletia* (Tabela 2), confirmando os resultados obtidos nas análises anteriores, onde se descreveu que as substâncias presentes nos tricomas são as principais responsáveis do agrupamento das espécies de *Espeletia* de acordo com sua origem geográfica.

**Tabela 4.** Principais substâncias correlacionadas com o agrupamento das espécies do gênero *Espeletia* no cluster colombiano (VIP>3.0) e sua desreplicação.

| | ID | *VIP* | *m/z* | $t_R$ (min) | Formula molecular | IDH | Classe | Substância provável |
|---|---|---|---|---|---|---|---|---|
| Positivo | 1 | 21.71 | 317.0655 | 10.96 | $C_{16}H_{12}O_7$ | 11 | Flavonoide | 3-Metóxiquercetina* |
| | 5 | 12.43 | 617.0922 | 14.19 | $C_{31}H_{20}O_{14}$ | 22 | Flavonoide | 8,8''-Metilenobisquercetina |
| | 104 | 5.70 | 303.0859 | 12.96 | $C_{16}H_{14}O_6$ | 10 | Flavonoide | Hesperetina |
| | 70 | 4.55 | 917.1184 | 15.67 | - | - | Flavonoide | N.I. |
| | 12 | 4.39 | 565.1185 | 9.52 | $C_{25}H_{24}O_{15}$ | 14 | Ácido clorogênico | N.I. |
| | 19 | 4.03 | 479.1182 | 8.65 | $C_{22}H_{22}O_{12}$ | 12 | Flavonoide | Isoramnetina-3-glicosídeo |
| | 291 | 3.40 | 573.1026 | 13.11 | $C_{30}H_{20}O_{12}$ | 21 | Ácido clorogênico | N.I. |
| | 68 | 3.27 | 603.0767 | 13.71 | $C_{30}H_{20}O_{11}$ | 21 | Flavonoide | N.I. |
| Negativo | 1 | 17.21 | 315.0510 | 11.01 | $C_{16}H_{12}O_7$ | 11 | Flavonoide | 3-Metoxiquercetina* |
| | 32 | 9.66 | 695.1255 | 10.05 | $C_{33}H_{28}O_{17}$ | 20 | Ácido clorogênico | Isômero do ácido tricafeoilaltrárico |
| | 6 | 7.47 | 615.0782 | 14.20 | $C_{31}H_{20}O_{14}$ | 22 | Flavonoide | 8,8''-Metilenobisquercetina |
| | 23 | 4.70 | 915.1058 | 15.65 | - | - | Ácido clorogênico | N.I. |
| | 179 | 4.64 | 301.0717 | 12.98 | $C_{16}H_{14}O_6$ | 10 | Flavonoide | Hesperetina |
| | 71 | 3.76 | 453.0463 | 10.36 | $C_{22}H_{14}O_{11}$ | 16 | N.I. | N.I. |
| | 21 | 3.67 | 571.0884 | 13.14 | $C_{30}H_{20}O_{12}$ | 21 | N.I. | N.I. |
| | 25 | 3.54 | 603.0787 | 12.34 | $C_{30}H_{20}O_{14}$ | 21 | Flavonoide | N.I. |
| | 15 | 3.40 | 587.0834 | 13.74 | $C_{30}H_{20}O_{13}$ | 21 | Flavonoide | N.I. |
| | 1296 | 3.07 | 695.1254 | 9.58 | $C_{33}H_{28}O_{17}$ | 20 | Ácido clorogênico | Isômero do ácido tricafeoilaltrárico |

* Substância confirmada por coinjeção com padrão; N.I.: não identificada.

As principais substâncias correlacionadas com o agrupamento das espécies no cluster venezuelano (Tabela 5) foram flavonoides glicosilados como a quercetina-3-*O*-galactosídeo, cuja presença nos extratos foi confirmada por coinjeção com padrão, além da quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo e a quercetina-3-*O*-arabinosídeo. Tais substâncias apresentam uma alta correlação em ambos os modos de ionização e foram igualmente identificadas como constituintes majoritários dos tricomas de *Espeletia* (Tabela 2). Além disso, outro flavonoide presumivelmente diglicosilado (ID 195, Tabela 5) apresenta também uma alta correlação com este agrupamento.

**Tabela 5.** Principais substâncias correlacionadas com o agrupamento das espécies do gênero *Espeletia* no cluster venezuelano (VIP>3.0) e sua desreplicação.

| | ID | *VIP* | *m/z* | $t_R$ (min) | Formula molecular | IDH | Classe | Substância provável |
|---|---|---|---|---|---|---|---|---|
| Positivo | 195 | 16.06 | 611.1391 | 10.10 | $C_{30}H_{26}O_{14}$ | 18 | Flavonoide | Flavonoide diglicosilado |
| | 173 | 14.19 | 465.1027 | 7.94 | $C_{21}H_{20}O_{12}$ | 12 | Flavonoide | Quercetina-3-*O*-galactosídeo* |
| | 34 | 6.81 | 301.2159 | 16.23 | $C_{20}H_{28}O_2$ | 7 | N.I. | N.I. |
| | 3608 | 6.05 | 627.1341 | 9.31 | $C_{30}H_{26}O_{15}$ | 18 | Flavonoide | Quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo |
| | 221 | 4.40 | 435.0919 | 8.34 | $C_{20}H_{18}O_{11}$ | 12 | Flavonoide | Quercetin- 3-*O*-arabinosídeo |
| Negativo | 163 | 13.12 | 463.0882 | 8.07 | $C_{21}H_{20}O_{12}$ | 12 | Flavonoide | Quercetina-3-*O*-galactosídeo* |
| | 221 | 13.36 | 609.1252 | 10.11 | $C_{30}H_{26}O_{14}$ | 18 | Flavonoide | Flavonoide diglicosilado |
| | 169 | 9.00 | 625.1202 | 9.38 | $C_{30}H_{26}O_{15}$ | 18 | Flavonoide | Quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo |
| | 168 | 5.36 | 635.1618 | 7.86 | - | - | Flavonoide | N.I. |
| | 13 | 4.91 | 515.1193 | 6.54 | $C_{25}H_{24}O_{12}$ | 14 | Ácido clorogênico | Derivado do ácido dicafeoilquínico |
| | 225 | 4.63 | 771.1779 | 9.28 | - | - | Flavonoide | N.I. |
| | 181 | 4.23 | 433.0775 | 8.33 | $C_{20}H_{18}O_{11}$ | 12 | Flavonoide | Quercetina-3-*O*-arabinosídeo |

* Substância confirmada por coinjeção com padrão; N.I. não identificada.

Na análise por OPLS-DA da matriz 2 (Figura 26) em que as áreas dos picos foram transformadas em variáveis binárias (presença/ausência) se obtiveram resultados semelhantes aos observados na análise anterior (Figura 24), onde foram consideradas como variáveis contínuas. Nesta análise são observados também dois agrupamentos correspondentes ao país de origem das espécies, e similar à análise da matriz 1, as espécies *E. schultzii* x *E. jajoensis* (100) e *E. ulotricha* (117) foram as duas espécies de origem venezuelana mais próximas do cluster colombiano. Nesta análise se obtiveram também valores adequados de $R^2$ ($R^2Y$=0,93 modo negativo e $R^2Y$=0,96 modo positivo) e $Q^2$ ($Q^2Y$=0,65 e 0,68 no modo negativo e positivo, respectivamente), os quais resultam próximos aos obtidos na análise da matriz 1.

Apesar da similaridade entre os resultados da análise das matrizes 1 e 2, (Figuras 24 e 26, respectivamente) diferentes variáveis foram identificadas como discriminantes de cada agrupamento (Tabelas 6 e 7), indicando que a química do metabolismo secundário das espécies da Venezuela e da Colômbia difere não só quantitativamente nos metabólitos produzidos pelas espécies de acordo com seu país de origem, mas também qualitativamente na presença de determinados metabólitos únicos para cada agrupamento.

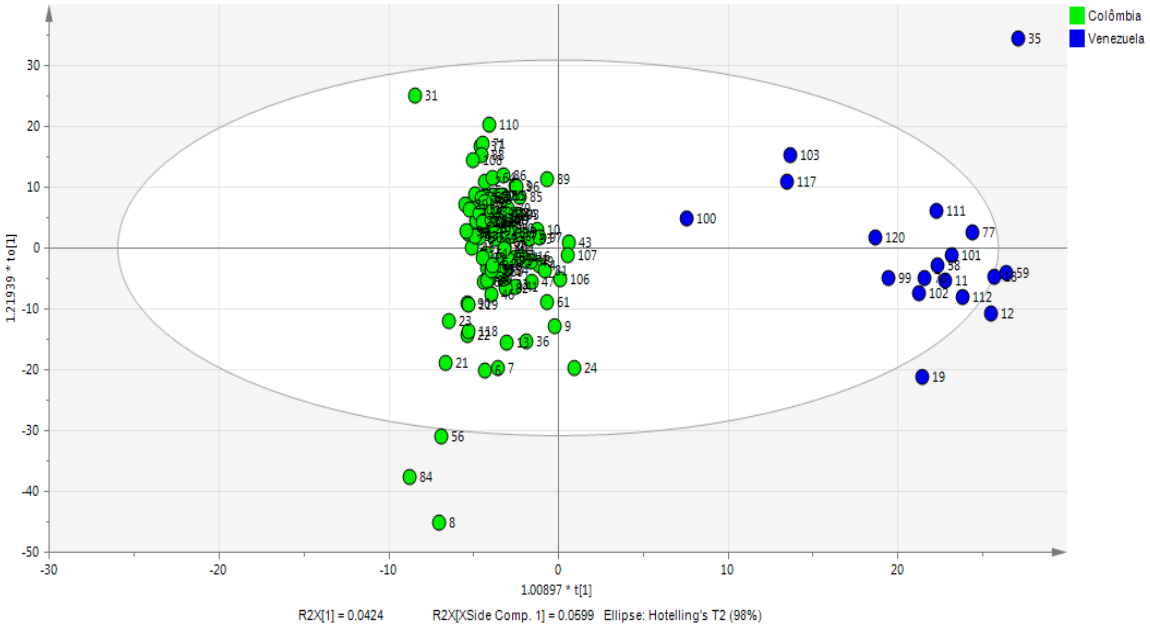



**Figura 26.** *Score-plot* da análise por OPLS-DA de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo), com áreas dos picos transformadas em variáveis binárias. Países de origem das espécies indicados em diferentes cores.

**Tabela 6.** Principais substâncias correlacionadas com o agrupamento das espécies do gênero *Espeletia* no cluster colombiano (VIP>3.0) e sua desreplicação (matriz 2).

| | ID | *VIP* | *m/z* | $t_R$ (min) | Formula molecular | IDH | Classe | Substância provável |
|---|---|---|---|---|---|---|---|---|
| Positivo | 2388 | 3.93 | 633.0873 | 12.64 | - | - | N.I. | N.I. |
| | 3123 | 3.88 | 901.1239 | 16.16 | - | - | N.I. | N.I. |
| | 1838 | 3.77 | 527.1189 | 11.34 | $C_{26}H_{22}O_{12}$ | 16 | Flavonoide | N.I. |
| | 1134 | 3.69 | 743.1812 | 12.28 | - | - | N.I. | N.I. |
| | 2135 | 3.68 | 931.1351 | 12.79 | - | - | N.I. | N.I. |
| | 1142 | 3.53 | 809.1705 | 13.54 | - | - | N.I. | N.I. |
| | 468 | 3.52 | 931.1345 | 13.50 | - | - | N.I. | N.I. |
| | 2637 | 3.28 | 931.1355 | 16.17 | - | - | N.I. | N.I. |
| | 131 | 3.19 | 865.1454 | 12.36 | | | Flavonoide | N.I. |
| | 619 | 3.08 | 537.1243 | 9.07 | $C_{24}H_{24}O_{14}$ | 13 | N.I. | N.I. |
| Negativo | 831 | 4.42 | 697.1775 | 12.29 | - | - | N.I. | N.I. |
| | 694 | 4.02 | 807.1575 | 13.53 | - | - | N.I. | N.I. |
| | 2021 | 3.95 | 633.0894 | 11.78 | - | - | N.I. | N.I. |
| | 127 | 3.61 | 769.1047 | 10.49 | - | - | Flavonoide | N.I. |
| | 1474 | 3.55 | 929.1216 | 12.79 | - | - | N.I. | N.I. |
| | 1276 | 3.08 | 1373.6608 | 15.64 | - | - | Ácido clorogênico | N.I. |
| | 529 | 3.07 | 825.0956 | 14.72 | - | - | N.I. | N.I. |

N.I.: não identificada.

**Tabela 7.** Principais substâncias correlacionadas com o agrupamento das espécies do gênero *Espeletia* no cluster venezuelano (VIP>1.8) e sua desreplicação (Matriz 2).

| | ID | VIP | *m/z* | $t_R$ (min) | Formula molecular | IDH | Classe | Substância provável |
|---|---|---|---|---|---|---|---|---|
| Positivo | 599 | 2.89 | 911.1667 | 12.03 | - | - | Ácido clorogênico | N.I. |
| | 206 | 1.80 | 627.1340 | 9.88 | $C_{30}H_{26}O_{15}$ | 18 | Ácido clorogênico | N.I. |
| Negativo | 168 | 4.62 | 635.1618 | 7.86 | - | - | Flavonoide | N.I. |
| | 1328 | 2.80 | 641.1518 | 10.77 | - | - | N.I. | N.I. |
| | 271 | 2.40 | 776.1420 | 9.00 | - | - | Ácido clorogênico | N.I. |
| | 240 | 2.32 | 359.0770 | 0.58 | $C_{18}H_{16}O_8$ | - | Flavonoide | N.I. |

N.I.: não identificada.

Ressalta-se que o flavonoide com *m/z* 635.1618 (ID 168) foi discriminante do cluster venezuelano, tanto na matriz com dados binários quanto na matriz com dados contínuos (Tabela 5), indicando que as espécies da Venezuela não só o biossintetizam em elevadas quantidades, mas também que o mesmo é exclusivo desse grupo de espécies. Isto foi conferido na matriz inicial onde se observou que efetivamente este flavonoide está ausente em todas as espécies colombianas.

Além do agrupamento das espécies de acordo com seu país de origem (Figuras 24 e 26), foram efetuadas duas análises por OPLS-DA considerando a cordilheira e os páramos (representados pelos maciços aos quais eles pertencem) de origem das mesmas como variáveis reposta. Três grupos (Venezuela, cordilheira central e ocidental e cordilheira oriental) foram considerados na correlação com a cordilheira de origem das espécies (Figura 27), enquanto cinco grupos (P. de Boyacá, P. de Cundinamarca, P. de Santander e Norte de Santander, P. da cordilheira central e ocidental e P. da Venezuela) foram considerados na correlação com o páramo de origem das mesmas (Figuras 28 e 29).

Na Figura 27 se observa que as espécies se agruparam de acordo com sua cordilheira de origem formando três agrupamentos considerando um grau de similaridade superior ao 75% como linha de corte (linha vermelha, Figura 27). Nesta análise se obtiveram valores adequados de $R^2$ ($R^2Y$=0,95) e $Q^2$ ($Q^2$=0,73), mesmo como na análise do modo positivo ($R^2Y$=0,88 e $Q^2$=0,63), na qual se obtiveram resultados semelhantes considerando a mesma linha de corte (não apresentados). Este resultado concorda com a árvore filogenética da subtribo baseada nos

marcadores AFLPs (Figura 2, Anexo I), a qual apresenta também três clados correspondentes à cordilheira de origem das espécies (Diazgranados, 2012).

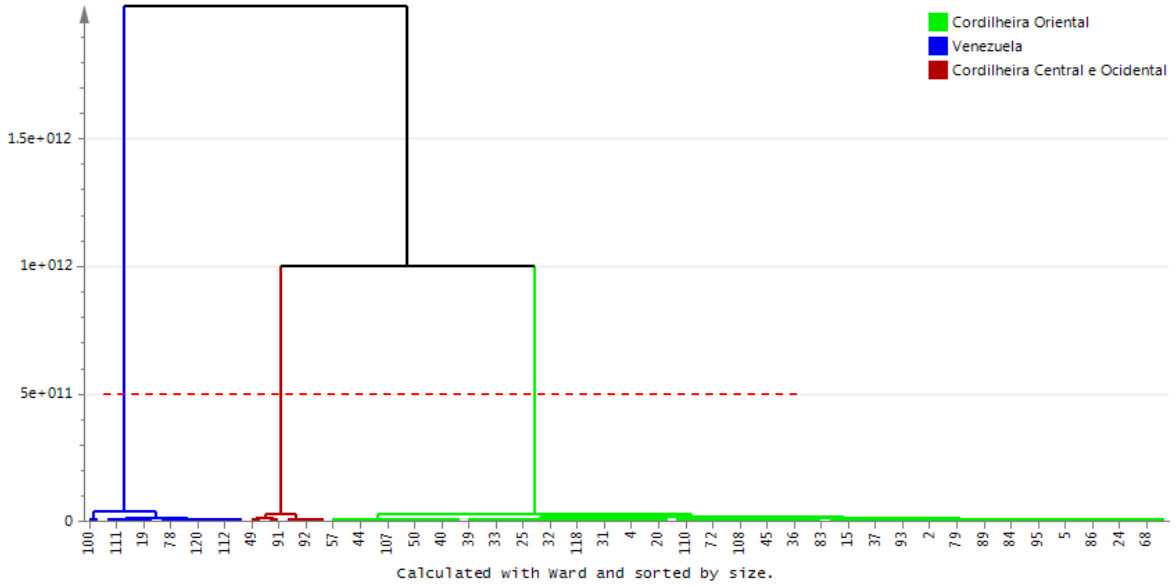



**Figura 27.** Dendograma da análise por OPLS-DA de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Cordilheiras de origem das espécies indicadas em diferentes cores.

Ao considerar os principais maciços de páramos da Colômbia como variável resposta, obtiveram-se igualmente resultados satisfatórios, confirmando que há diferenças na composição química das espécies de acordo com o páramo de origem das mesmas (Figura 28). Nesta análise, um resultado interessante foi que a proximidade observada entre os cinco agrupamentos do dendrograma (considerando um grau de similaridade superior ao 75% como linha de corte, linha vermelha, Figura 28), concorda com a proximidade geográfica dos mesmos onde, por exemplo, as espécies presentes nos páramos de Boyacá (P. de Boyacá) se agrupam junto com as espécies de Cundinamarca (P. de Cundinamarca), enquanto as espécies de Santander e Norte de Santander formam um cluster maior com os dois anteriores grupos. Esses três maciços correspondem a áreas muito próximas geograficamente (Figura 29).

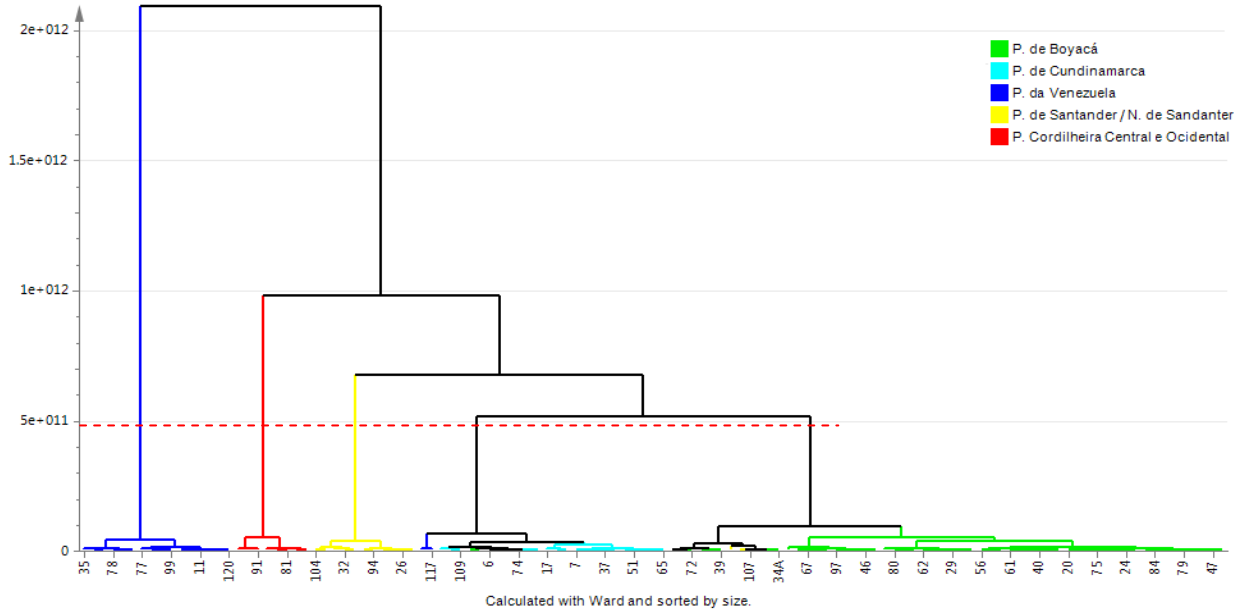



**Figura 28.** Dendograma da análise por OPLS-DA de 120 amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS (modo negativo). Principais maciços de páramos indicados em diferentes cores.

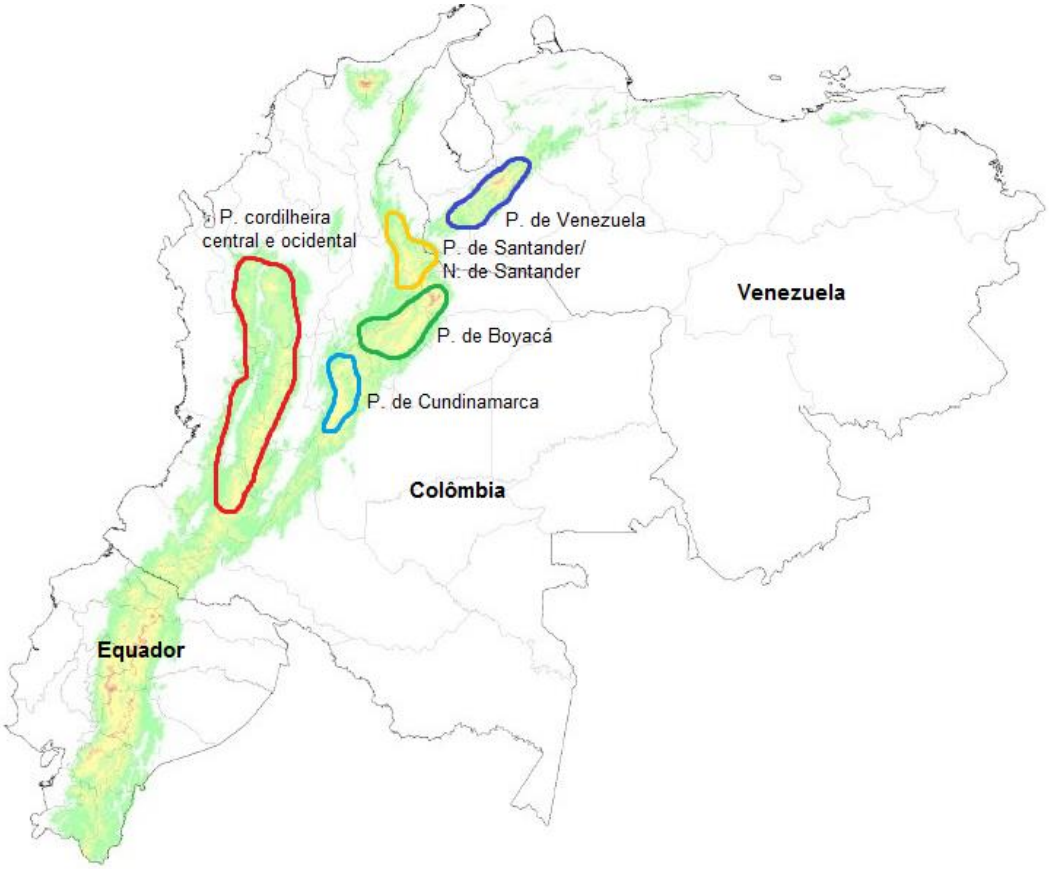



**Figura 29.** Mapa da Colômbia, Equador e Venezuela com principais maciços de páramos considerados neste estudo ressaltados em diferentes cores (Adaptado de DIAZGRANADOS, 2012).

Ressalta-se que cinco espécies, três de Boyacá (*E. boyacensis* (21), *E. incana* (57) e *E. paipana* × *E. murilloi* (85)) e duas da Venezuela (*E. schultzii* × *E. jajoensis* (100) e *E. ulotricha* (117)), foram agrupadas no cluster dos páramos de Cundinamarca, enquanto quatro espécies, três de Santander (*E. brassicoidea* (27), *E. estanislana* (45) e *E. steyermarkii* (107)) e uma de Cundinamarca (*E. uribei* (118)), se agruparam no cluster dos páramos de Boyacá. Isto sugere que espécies presentes em áreas geográficas próximas (Figura 29), mesmo pertencentes a diferentes maciços de páramos, apresentam uma composição química semelhante e, em alguns casos, o modelo gerado não consegue distingui-las, confirmando a forte influência da biogeografia da composição química das espécies. Resultado análogo àqueles obtidos com base em caracteres morfológicos (CUATRECASAS, 2013) e moleculares (DIAZGRANADOS, 2012).

Na análise dos dados gerados no modo positivo se obtiveram resultados similares aos já apresentados, onde as espécies de agruparam de acordo com o páramo de origem, considerando a mesma linha de corte (não apresentado). Este resultado concorda com a filogenia da subtribo Espeletiinae, construída com base nos marcadores moleculares ITS, ETS, rp116 e AFLPs (Figuras 1 e 2, Anexo I).

Na Tabela 8, são apresentadas as principais substâncias correlacionadas com o agrupamento das espécies nos seus páramos de origem revelando que existe certa tendência na acumulação de determinados metabólitos pelas espécies pertencentes a cada agrupamento. Por exemplo, as espécies presentes nos páramos de Santander e Norte de Santander, foram as únicas a apresentarem triterpenos como algumas das principais substâncias discriminantes deste grupo, além da presença do ácido 3,4-di-*O*-(*E*)-cafeoilquínico. Nas espécies presentes nos páramos de Boyacá, as duas principais substâncias responsáveis pelo seu agrupamento foram os flavonoides metoxilados 3-metóxiquercetina e hesperetina, enquanto as espécies dos páramos de Cundinamarca apresentam flavonoides diméricos como a 8,8''-metilenobisquercetina e o flavonoide com ID 147 (modo positivo) e 366 (modo negativo), presumivelmente também dimérico com base no seu espectro de massas. O agrupamento das espécies nos páramos das cordilheiras central e ocidental da Colômbia foi o único em que as principais substâncias discriminantes foram diterpenos, destacando-se o ácido grandiflorólico e o Ácido *ent*-12-oxocauro-9(11),16-dien-18-oico. Neste agrupamento, nenhum

flavonoide foi identificado como altamente correlacionado, embora estivessem presentes também nas amostras analisadas. Finalmente, as espécies presentes nos páramos da Venezuela se caracterizam pela presença de flavonoides glicosilados, dos quais se destaca a quercetina-3-*O*-galactosídeo como a principal substância discriminante deste agrupamento.

**Tabela 8.** Principais substâncias correlacionadas com o agrupamento das espécies do gênero *Espeletia* nos páramos de origem e sua desreplicação.

| | ID | VIP | *m/z* | $t_R$ (min) | Formula molecular | IDH | Classe | Substância provável |
|---|---|---|---|---|---|---|---|---|
| | | | | ***P. de Santander e Norte de Santander*** | | | | |
| Positivo | 134 | 5.79 | 517.1338 | 8.70 | $C_{25}H_{24}O_{12}$ | 14 | Ácido clorogênico | Ácido 3,4-di-*O*-(*E*)-cafeoilquínico* |
| | 642 | 3.28 | 479.3728 | 25.31 | $C_{29}H_{50}O_5$ | 5 | Triterpeno | N.I. |
| | 68 | 3.1 | 603.0767 | 13.71 | $C_{30}H_{18}O_{14}$ | 22 | Flavonoide | N.I. |
| | 641 | 2.7 | 507.4038 | 27.26 | - | - | Triterpeno | N.I. |
| | 516 | 2.51 | 619.0722 | 11.20 | $C_{30}H_{18}O_{15}$ | 22 | Flavonoide | N.I. |
| | 115 | 5.81 | 617.0577 | 11.11 | $C_{30}H_{18}O_{15}$ | 22 | Flavonoide | N.I. |
| Negativo | 405 | 4.3 | 477.3584 | 25.27 | $C_{29}H_{50}O_5$ | 5 | Triterpeno | N.I. |
| | 846 | 3.27 | 317.0300 | 6.37 | $C_{15}H_{10}O_8$ | 11 | Flavonoide | N.I. |
| | 845 | 3.26 | 505.3898 | 27.27 | - | - | Triterpeno | N.I. |
| | 947 | 2.72 | 491.3742 | 26.33 | $C_{30}H_{52}O_5$ | 5 | Triterpeno | N.I. |
| | 496 | 2.33 | 917.0844 | 13.21 | - | - | N.I. | N.I. |
| | | | | ***P. de Boyacá*** | | | | |
| Positivo | 1 | 12.7 | 317.0655 | 10.96 | $C_{16}H_{12}O_7$ | 11 | Flavonoide | 3-Metóxiquercetina* |
| | 1464 | 8.21 | 303.0859 | 12.96 | $C_{16}H_{14}O_6$ | 10 | Flavonoide | Hesperetina |
| | 291 | 4.45 | 573.1026 | 13.11 | $C_{30}H_{20}O_{12}$ | 21 | Flavonoide | N.I. |
| | 297 | 4.13 | 273.0757 | 11.48 | $C_{15}H_{12}O_5$ | 10 | Flavonoide | Pinobanksina |
| | 290 | 2.15 | 627.1554 | 6.96 | $C_{27}H_{30}O_{17}$ | 13 | Flavonoide | Quercetina-3,4'-*O*-di-glucopiranose |
| | 451 | 2.06 | 517.0975 | 9.99 | $C_{24}H_{20}O_{13}$ | 15 | Ácido clorogênico | N.I. |
| | 1060 | 1.59 | 789.1864 | 8.49 | $C_{36}H_{30}O_{20}$ | 19 | Flavonoide | Quercetina-3-(6″-cafeoilsoforosídeo) |
| | 501 | 1.45 | 465.1028 | 8.96 | $C_{24}H_{16}O_{10}$ | 17 | Flavonoide | Castillicetina |
| Negativo | 1 | 10.6 | 315.0510 | 11.01 | $C_{16}H_{12}O_7$ | 11 | Flavonoide | 3-Metóxiquercetina* |
| | 179 | 9.13 | 301.0717 | 12.98 | $C_{16}H_{14}O_6$ | 10 | Flavonoide | Hesperetina |
| | 455 | 3.05 | 533.0936 | 7.96 | $C_{24}H_{22}O_{14}$ | 14 | Ácido clorogênico | N.I. |
| | 177 | 2.66 | 787.1731 | 8.47 | $C_{36}H_{30}O_{20}$ | 19 | Flavonoide | Quercetina-3-(6″-cafeoilsoforosídeo) |
| | 1296 | 2.09 | 695.1254 | 9.58 | $C_{33}H_{28}O_{17}$ | 20 | Ácido clorogênico | Isómero do ácido tricafeoilaltrárico |
| | 759 | 1.23 | 711.1428 | 7.49 | $C_{30}H_{32}O_{20}$ | 15 | Flavonoide | Quercetina 3-*O*-(6''-malonil glucosídeo) 7-*O*-glucosídeo |
| | | | | ***P. de Cundinamarca*** | | | | |
| Positivo | 1465 | 12.43 | 617.0922 | 14.19 | $C_{31}H_{20}O_{14}$ | 22 | Flavonoide | 8,8''-Metilenobisquercetina |
| | 147 | 3.89 | 631.1081 | 11.26 | - | - | Flavonoide | N.I. |
| | 366 | 1.87 | 589.0976 | 13.74 | $C_{30}H_{20}O_{13}$ | 21 | Flavonoide | N.I. |
| | 623 | 1.28 | 545.1645 | 8.42 | $C_{27}H_{28}O_{12}$ | 14 | Ácido clorogênico | Metil 3-*O*-cafeoil-4-*O*-feruloilquinato |

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Negativo | 366 | 4.68 | 629.094 | 11.29 | - | - | Flavonoide | N.I. |
| | 6 | 4.13 | 615.0782 | 14.20 | $C_{31}H_{20}O_{14}$ | 22 | Flavonoide | 8,8''-Metilenobisquercetina |
| | 23 | 3.56 | 915.1058 | 15.65 | - | - | Ácido clorogênico | N.I. |
| | 15 | 2.1 | 587.0834 | 13.74 | $C_{30}H_{20}O_{13}$ | 21 | Flavonoide | N.I. |
| ***P. cordilheira central e ocidental*** | | | | | | | | |
| Positivo | 8 | 4.97 | 315.1953 | 14.77 | $C_{20}H_{26}O_3$ | 8 | Diterpeno | Ácido *ent*-12-oxocauro-9(11),16-dien-18-oico |
| | 2993 | 4.05 | 301.2161 | 14.40 | $C_{20}H_{28}O_2$ | 7 | N.I. | N.I. |
| | 51 | 3.61 | 253.1949 | 17.45 | $C_{19}H_{24}$ | 8 | N.I. | N.I. |
| | 71 | 2.93 | 271.20550 | 17.45 | $C_{19}H_{20}O$ | 7 | N.I. | N.I. |
| Negativo | 42 | 4.88 | 695.12659 | 8.22 | - | - | Ácido clorogênico | N.I. |
| | 164 | 4.61 | 315.19619 | 15.51 | $C_{20}H_{28}O_3$ | 7 | Diterpeno | Ácido *ent*-12-hidroxicauro-9(11),16-dien-18-oico |
| | 232 | 4.41 | 313.18101 | 14.77 | $C_{20}H_{26}O_3$ | 8 | Diterpeno | Ácido *ent*-12-oxocauro-9(11),16-dien-18-oico |
| | 1240 | 4.38 | 335.22271 | 14.32 | $C_{20}H_{32}O_4$ | 5 | Diterpeno | N.I. |
| | 362 | 3.97 | 317.21190 | 18.16 | $C_{20}H_{30}O_3$ | 6 | Diterpeno | Ácido grandiflorólico |
| | 12 | 3.59 | 301.21722 | 22.84 | $C_{20}H_{30}O_2$ | 6 | Diterpeno | Ácido pimárico |
| | 491 | 3.17 | 315.19630 | 17.67 | $C_{20}H_{28}O_3$ | 7 | Diterpeno | N.I. |
| | 1855 | 3.11 | 603.26630 | 12.44 | $C_{28}H_{44}O_{14}$ | 7 | N.I. | N.I. |
| | 1879 | 2.46 | 387.25416 | 23.34 | $C_{24}H_{36}O_4$ | 7 | Diterpeno | Ácido *ent*-15-O-iso butiroxicaur-16-en-19-oico |
| | 1645 | 2.41 | 399.25360 | 24.14 | $C_{25}H_{36}O_4$ | 8 | Diterpeno | Ácido *ent*-15-O-senecioxicaur-16-en-19-oico |
| ***P. da Venezuela*** | | | | | | | | |
| Positivo | 195 | 16.06 | 611.13910 | 10.10 | $C_{30}H_{26}O_{14}$ | 18 | Flavonoide | Flavonoide diglicosilado |
| | 173 | 14.19 | 465.10265 | 7.94 | $C_{21}H_{20}O_{12}$ | 12 | Flavonoide | Quercetina-3-*O*-galactosídeo* |
| | 34 | 6.81 | 301.21594 | 16.23 | $C_{20}H_{28}O_2$ | 7 | Diterpeno | N.I. |
| | 3608 | 6.05 | 627.1341 | 9.31 | $C_{30}H_{26}O_{15}$ | 18 | Flavonoide | Quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo |
| | 284 | 4.51 | 345.09634 | 15.26 | $C_{18}H_{16}O_7$ | 11 | N.I. | N.I. |
| | 221 | 4.40 | 435.0919 | 8.34 | $C_{20}H_{18}O_{11}$ | 12 | Flavonoide | Quercetina-3-*O*-arabinosídeo |
| Negativo | 163 | 7.6 | 463.08823 | 7.96 | $C_{21}H_{20}O_{12}$ | 12 | Flavonoide | Quercetina-3-*O*-galactosídeo* |
| | 221 | 7.44 | 609.12520 | 10.11 | $C_{30}H_{26}O_{14}$ | 18 | Flavonoide | Flavonoide diglicosilado |
| | 169 | 5.28 | 625.12024 | 9.38 | $C_{30}H_{26}O_{15}$ | 18 | Flavonoide | Quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo |
| | 13 | 4.75 | 515.11928 | 6.54 | $C_{25}H_{24}O_{12}$ | 14 | Ácido clorogênico | Isômero do ácido dicafeoilquínico |
| | 3 | 4.25 | 301.03525 | 10.49 | $C_{15}H_{10}O_7$ | 11 | Flavonoide | Quercetina* |
| | 168 | 2.99 | 635.16179 | 7.86 | - | - | Flavonoide | N.I. |
| | 225 | 2.57 | 771.17785 | 9.28 | - | - | Flavonoide | N.I. |
| | 181 | 2.37 | 433.07750 | 8.33 | $C_{20}H_{18}O_{11}$ | 12 | Flavonoide | Quercetina-3-*O*-arabinosídeo |

* Substância confirmada por coinjeção com padrão; N.I.: substância não identificada.

A correlação geral observada entre a composição química das espécies e sua origem geográfica concorda com a filogenia atual da subtribo Espeletiinae proposta por Diazgranados (2012). De acordo com esta filogenia, as espécies formam clados

evidentes de acordo com o país de origem numa escala global e de acordo com os complexos de páramos onde foram coletadas numa escala regional (DIAZGRANADOS, 2012). Esta característica é consequência da história evolutiva da subtribo Espeletiinae, onde espécies simpátricas (espécies que crescem na mesma área geográfica) tendem a ser mais relacionadas geneticamente do que espécies distantes, provavelmente devido à carência de polinizadores de longa distância, mesmo como a ausência de mecanismos de dispersão de sementes. Isto sugere um marcado isolamento geográfico entre espécies de diferentes páramos, que por mecanismos de especiação alopátrica (ocasionada por meio de uma barreira geográfica) acabaram divergindo da espécie parental e evoluindo em novas formas (DIAZGRANADOS, 2012).

### 4.4.2 Modelos de predição da origem geográfica

Considerando que valores de $Q^2$ superiores a 0,5 são uma condição necessária para que um modelo tenha uma boa capacidade preditiva (GOLBRAIKH; TROPSHA, 2002; BAILEY et al., 2004), foi possível elaborar modelos de predição geográfica de novas amostras unicamente de acordo com o país e a cordilheira de origem das mesmas, já que na análise em que os páramos foram considerados como variável resposta se obtiveram valores inferiores o em torno de 0,5 ($Q^2$=0,52 e 0,45 no modo negativo e positivo, respectivamente).

A construção do modelo de predição da origem geográfica foi efetuada mediante diferentes métodos estatísticos supervisionados lineares e não lineares como árvores de decisão, *random forest*, *K-nearest neighbors* e OPLS-DA. Destes, o modelo por OPLS-DA foi o que apresentou a melhor capacidade preditiva de novas amostras com 95% de acerto na classificação (Figura 30). Este modelo foi elaborado considerando 67% (n = 82) do conjunto de dados como série de treinamento e o restante 33% (n = 38) como série de validação externa. Na construção do mesmo se excluíram os principais *outliers* representados no *Hotelling's T2 column plot* (não apresentado), além de algumas replicas. O apêndice IV contem as espécies empregadas na construção do modelo.

A partir do modelo gerado, se obtiveram valores ótimos de $R^2$ ($R^2Y$=0,98) e $Q^2$ ($Q^2$=0,73) indicando não só uma boa representação dos dados, mas também um adequado poder preditivo do mesmo. Isto foi conferido na porcentagem de

observações corretamente classificadas do conjunto de espécies empregado na validação externa (Figura 30). Ressalta-se que o valor obtido a partir do índice de probabilidade de Fisher (3,4e-008) indica resultados estatisticamente significativos, considerando que valores inferiores a 0,05 sugerem uma baixa probabilidade de obter a classificação por acaso.

| | Members | Correct | Cordilheira Oriental | Venezuela | Cordilheira Central e Ocidental |
|---|---|---|---|---|---|
| Cordilheira Oriental | 29 | 100% | 29 | 0 | 0 |
| Venezuela | 6 | 66.67% | 2 | 4 | 0 |
| Cordilheira Central e Ocidental | 3 | 100% | 0 | 0 | 3 |
| No class | 0 | | 0 | 0 | 0 |
| Total | 38 | 94.74% | 31 | 4 | 3 |
| Fishers prob. | 3.4e-008 | | | | |

**Figura 30.** Matriz de classificação do conjunto de espécies empregadas na validação externa do modelo por OPLS-DA.

Das 38 espécies empregadas na validação externa, 36 (95%) foram classificadas corretamente pelo modelo, confirmando o excelente poder preditivo do mesmo. Além disso, o fato do modelo ter apresentado bons resultados tanto na validação cruzada quanto na validação externa sugere que não esteja acontecendo o fenômeno de superajuste, no qual o algoritmo perde a capacidade de predizer corretamente a classe de amostras não incluídas no conjunto de treinamento.

As espécies *E. schultzii* x *E. jajoensis* (100) e *E. ulotricha* (117) de origem venezuelana foram as duas únicas espécies classificadas incorretamente pelo modelo, indicando que ditos indivíduos apresentam uma composição química semelhante às espécies da cordilheira oriental da Colômbia, portanto o algoritmo não conseguiu predizer corretamente a origem geográfica das mesmas. Este resultado concorda com o obtido anteriormente na análise por OPLS-DA tanto da matriz com dados contínuos (Figura 23) quanto da matriz com dados binários (Figura 26), onde se observa que essas espécies são as mais próximas do cluster colombiano.

Na Figura 31, apresenta-se o modelo de predição geográfica de novas amostras de *Espeletia* pelo método de árvores de decisão (AD) implementando o algoritmo de classificação C4.5 (J48 no software Weka). O algoritmo C4.5 constrói árvores de decisão a partir de um grupo de treinamento e apresenta a vantagem de

analisar atributos contínuos e discretos simultaneamente (WITTEN et al., 2005). Este algoritmo utiliza uma estratégia de “cima” para “baixo” em que o conjunto de dados é inicialmente dividido em duas partes iguais e cada metade é subsequentemente dividida várias vezes até gerar uma árvore ideal que possa separar as classes (neste caso a origem geográfica das espécies). Portanto, os atributos (picos cromatográficos) que fornecem mais informação são selecionados com a finalidade de obter a menor árvore possível (WITTEN et al., 2005). O método de AD apresenta a vantagem de ser facilmente interpretável e uma vez o modelo e construído não é necessário o uso do software para predizer a classe de novas amostras, basta unicamente com seguir os nós da árvore, que neste caso representam o ID dos picos cromatográficos, numa estratégia de “cima” para “baixo” para conseguir predizer a classe de novas amostras com base na área dos picos cromatográficos.

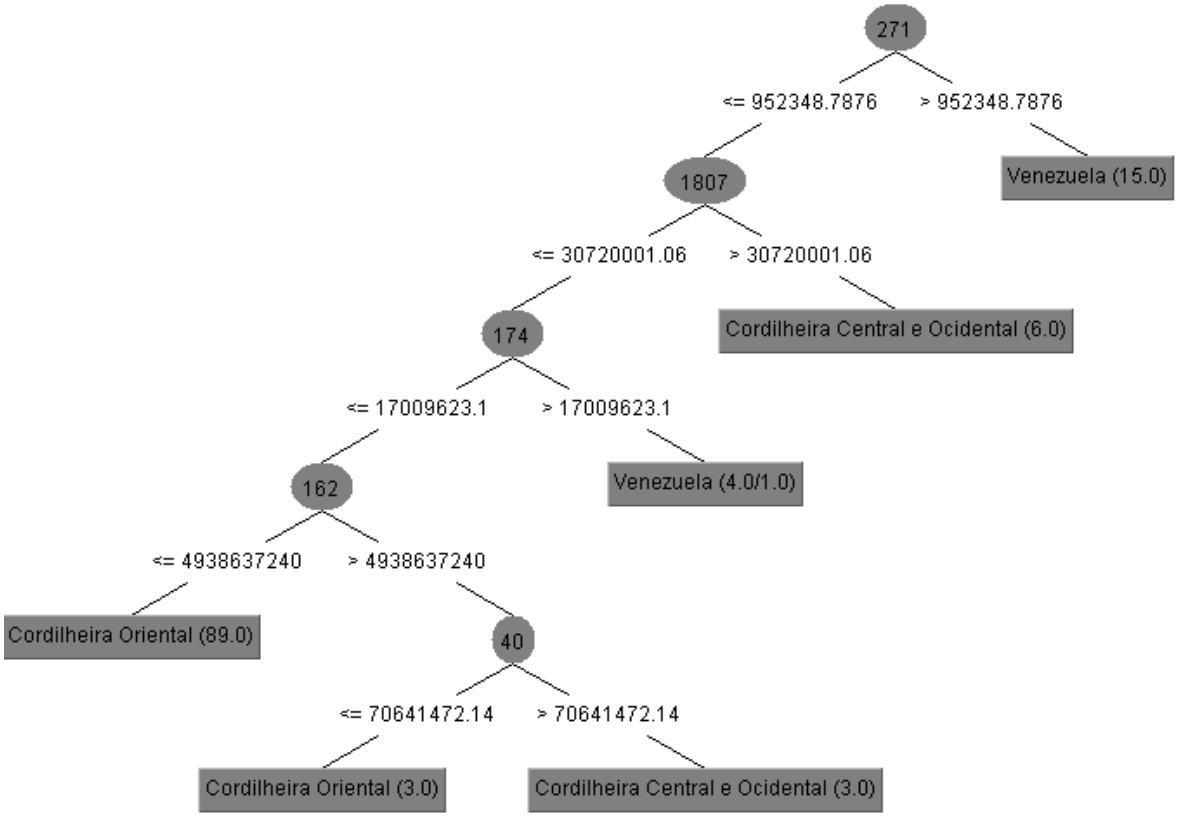



**Figura 31.** Modelo de predição geográfica por árvores de decisão

Embora o modelo por AD (Figura 31) tenha sido elaborado mediante a seleção de atributos por algoritmos genéticos, em que as variáveis pouco correlacionadas com a resposta foram descartadas, o mesmo apresentou um poder preditivo levemente inferior (90%) ao modelo por OPLS-DA (95%). Ressalta-se que

duas (ID 271 e 162, Figura 31) das cinco substâncias selecionadas como variáveis discriminantes no modelo por AD foram selecionadas também na análise por OPLS-DA (Tabelas 7 e 8), confirmando a importância de tais substâncias na separação das classes. Os parâmetros empregados na seleção de atributos e construção do modelo por AD são detalhados no Apêndice III.

Estes modelos podem ser de grande utilidade na compressão da química do metabolismo secundário deste gênero, já que se for aplicado aos 2176 espécimes de herbário do gênero *Espeletia* reportados por Diazgranados (2012), presentes em 39 diferentes herbários do mundo inteiro, seria possível comparar os padrões de distribuição atuais dos metabólitos secundários com aqueles dos exemplares de herbário mais antigos e determinar se houve ou não uma mudança nos perfis químicos das espécies deste gênero ao longo dos anos de acordo com sua origem geográfica. Isto considerando que desde a primeira descrição de uma espécie de *Espeletia* em 1792 por José Celestino Mutis, a taxonomia do gênero tem sido muito estudada e novas amostras são depositadas constantemente em diferentes herbários. Obviamente, a viabilidade deste projeto dependeria da qualidade das amostras e da metodologia empregada na sua preservação, além do fato de ter que se preparar as amostras para sua análise no UHPLC-UV-MS.

Além disso, considerando que a mais atual revisão da taxonomia da subtribo Espeletiinae foi efetuada examinando-se 4408 exemplares de herbário, dos quais unicamente 3408 foram identificados e geo-referenciados totalmente (DIAZGRANADOS, 2012), o modelo gerado poderia auxiliar na determinação taxonômica das restantes amostras pertencentes ao gênero *Espeletia*, considerando o alto grau de endemismo de suas espécies.

Finalmente, estudos de atividade biológica passados e futuros envolvendo espécies de *Espeletia* e o presente modelo, servirão para correlacionar uma determinada atividade com uma região geográfica específica, e não só predizer a atividade de novas amostras baseados na origem geográfica comum com a espécie testada, mas também aportar informações úteis para encaminhar propostas de conservação, considerando que toda a subtribo Espeletiinae se encontra em grande risco de extinção devido principalmente às alterações climáticas e à expansão agropecuária (DIAZGRANADOS, 2012).

### 4.4.3 Correlação com dados taxonômicos

A correlação com dados taxonômicos foi efetuada mediante uma análise por OPLS-DA, uma vez que este método forneceu resultados adequados na correlação com dados biogeográficos. Nesta análise se estudou a composição química de diferentes híbridos, variedades e formas presentes no gênero, junto com suas respectivas espécies parentais, com a finalidade de determinar se o uso de marcadores químicos em *Espeletia* pode ser de utilidade no estabelecimento de relações de parentesco taxonômica neste gênero.

Nas Figuras 32 e 33 são apresentadas as análises por OPLS-DA da matriz 1 (modo negativo) de seis híbridos (*E. schultzii* × *E. jajoensis, E. batata* × *E. schultzii, E. paipana* × *E. murilloi, E. murilloi* × *E. boyacensis, E. jaramilloi* × *E. boyacensis* e *E. argentea* × *E. barclayana*) com suas respectivas espécies parentais. Como vários dos híbridos estudados compartilham a mesma espécie parental (ex. *E. schultzii* × *E. jajoensis* e *E. batata* × *E. schultzii*), foram efetuadas duas análises independentes considerando três diferentes híbridos em cada uma: *E. schultzii* × *E. jajoensis*, *E. jaramilloi* × *E. boyacensis* e *E. paipana* × *E. murilloi* na primeira análise (Figura 32) e *E. batata* × *E. schultzii*, *E. murilloi* × *E. boyacensis* e *E. argentea* × *E. barclayana* na segunda (Figura 33).

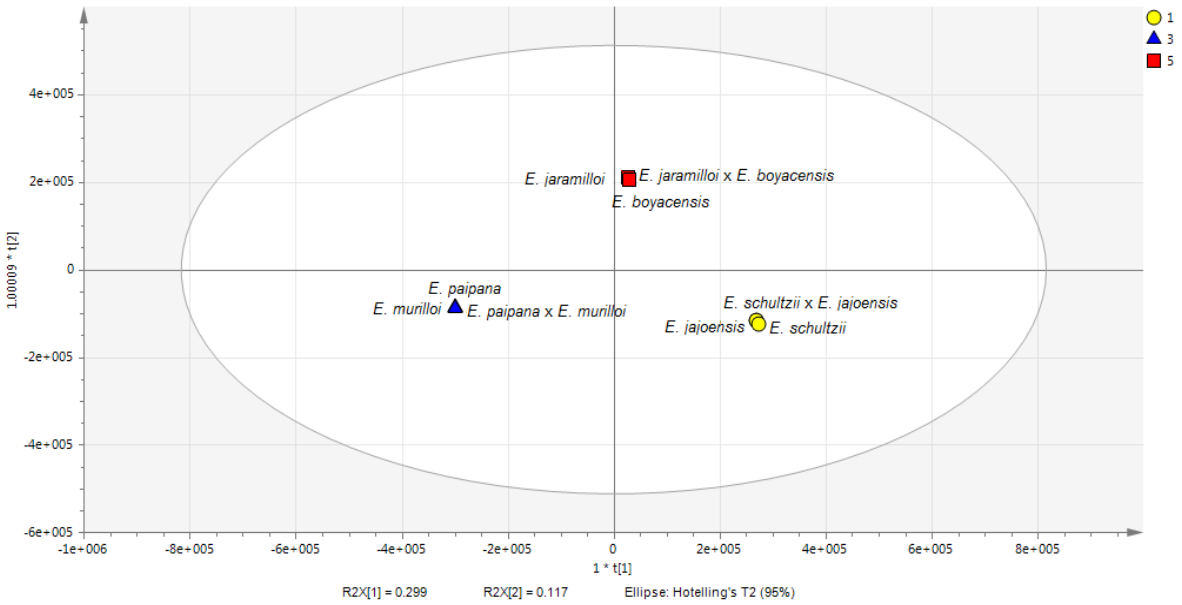



**Figura 32.** *Score-plot* da análise por OPLS-DA de três híbridos do gênero *Espeletia* e suas espécies parentais analisadas por UHPLC-UV-MS (modo negativo).

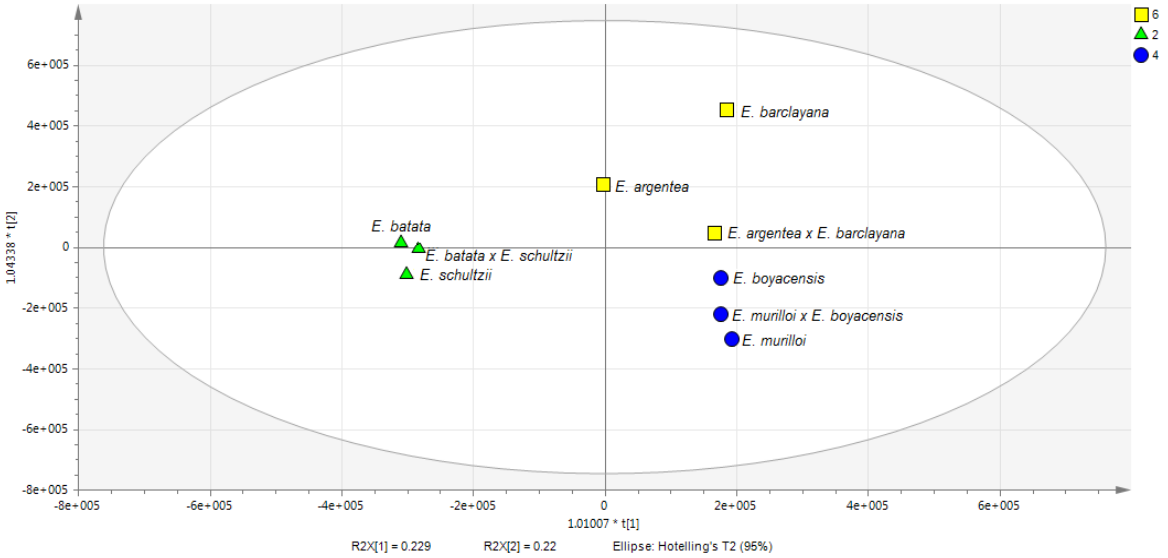



**Figura 33.** *Score-plot* da análise por OPLS-DA de três híbridos do gênero *Espeletia* e suas espécies parentais analisadas por UHPLC-UV-MS (modo negativo).

Nestas análises se obtiveram excelentes valores de $R^2$ ($R^2Y$ = 1,0) e $Q^2$ ($Q^2$ > 0,86), e da mesma forma como na análise do modo positivo ($R^2Y$ = 1,0 e $Q^2$ > 0,78) (não apresentada) se observa um excelente agrupamento dos híbridos com suas espécies parentais, indicando que existe não só similaridade na composição química das duas espécies progenitoras, o que pode ser explicado pela origem geográfica comum, mas também que os híbridos são quimicamente semelhantes às suas espécies parentais. Este resultado é interessante do ponto de vista quimiotaxonômico, pois indica que neste caso os marcadores químicos em *Espeletia* concordam com a taxonomia do gênero baseada em caracteres morfológicos. Resultados semelhantes foram obtidos na análise da matriz 2 (não apresentados).

Na subtribo Espeletiinae unicamente dois híbridos (*Carramboa* × *tachirensis* e *E.* × *aurantia*) têm sido estudados do ponto de vista fitoquímico (BOHLMANN; SUDING; CUATRECASAS; ROBINSON; et al., 1980; IBÁÑEZ; USUBILLAGA, 2008) e só um estudo (IBÁÑEZ; USUBILLAGA, 2008) comparou o perfil químico de um híbrido (*E.* × *aurantia*) com o de suas espécies parentais (*E. schultzii* e *Coespeletia moritziana*). Neste estudo, baseado na análise dos constituintes voláteis das espécies por cromatografia gasosa acoplada a espectrometria de massas (GC-MS por suas siglas em inglês), encontrou-se uma porcentagem intermediária para todos os componentes do óleo do híbrido com relação à porcentagem dos mesmos componentes nas espécies parentais (IBÁÑEZ; USUBILLAGA, 2008).

Segundo Diazgranados (2012), o fenômeno de hibridização em *Espeletia* é muito frequente quando duas ou mais espécies deste gênero compartilham a mesma área geográfica (espécies simpátricas) ou estão presentes em áreas geográficas adjacentes (espécies parapátricas). Além disso, a hibridização é também frequente num gradiente de atitude, onde o hibrido é frequentemente encontrado no meio de duas populações cujos limites de distribuição superior e inferior se sobrepõem. Nesse caso, a inflorescência dos híbridos é mais parecida com a da espécie de baixa altitude, enquanto a folha dos mesmos reflete em maior medida a folha da espécie de maior altitude (DIAZGRANADOS, 2012).

Quimicamente o fenômeno de hibridização em plantas tem sido pouco estudado e ainda há muitas dúvidas com relação à influência do mesmo no metabolismo secundário vegetal. Por tanto, este representa um campo fértil de pesquisa que pode ser especialmente abordado na subtribo Espeletiinae, considerando a importância desse fenômeno em radiações adaptativas (SEEHAUSEN, 2004), além de ser um dos principais mecanismos de especiação determinante na evolução desde grupo de plantas (CUATRECASAS, 2013).

Na análise por OPLS-DA das categorias infraespecíficas foram avaliadas quatro variedades (*E. formosa* var*. gigantea, E. lopezii* var. *major, E. murilloi* var. *alba* e *E. murilloi* var*. iguaquei*) e uma forma (*E. argentea* fma. *phaneractis*) junto com suas espécies parentais (Figura 34). Nesta análise, similar à análise dos híbridos (Figuras 32 e 33), há uma excelente correlação entre a composição química das categorias infraespecíficas e suas respectivas espécies parentais, já que em todos os casos tais indivíduos se agruparam juntos.

Este resultado sugere que existe uma alta similaridade na composição química das espécies e suas variedades ou formas, confirmando que os marcadores químicos em *Espeletia* concordam com a taxonomia do gênero baseada em caracteres morfológicos. Ressalta-se que resultados semelhantes foram obtidos na análise da matriz 2 (não apresentada), indicando que há similaridades tanto qualitativa quanto quantitativamente nos metabólitos produzidos por diferentes indivíduos pertencentes a categorias taxonômicas infraespecíficas em *Espeletia.*

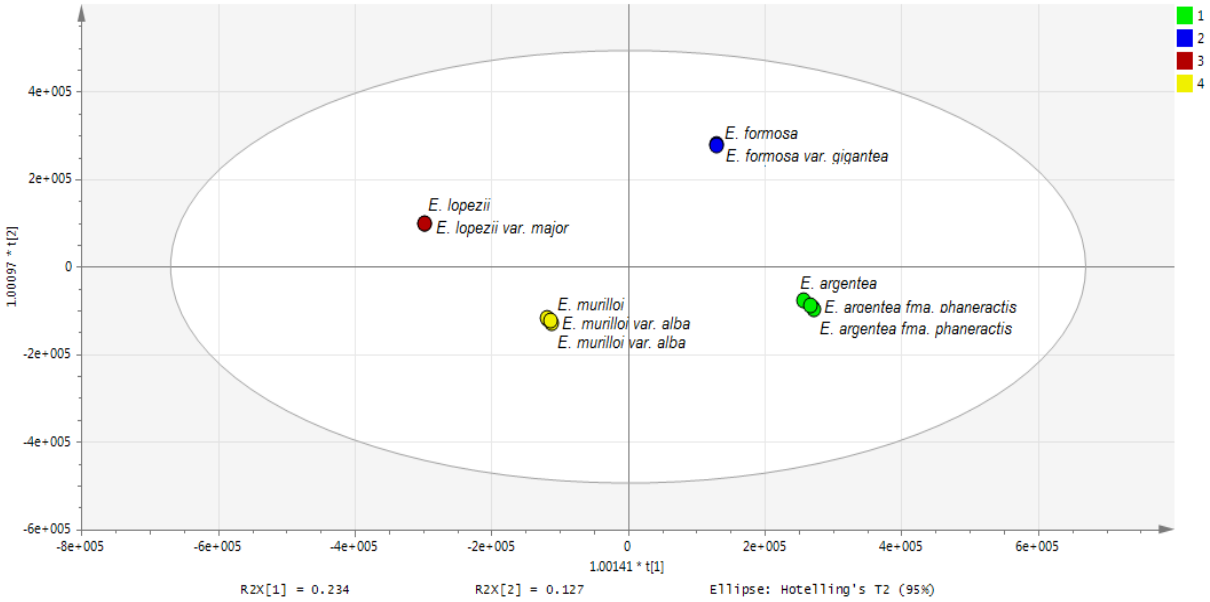



**Figura 34.** *Score-plot* da análise por OPLS-DA de quatro variedades e uma forma do gênero *Espeletia* e suas espécies parentais analisadas por UHPLC-UV-MS (modo negativo).

O fato das matrizes 1 e 2 terem apresentado resultados semelhantes nas análises de correlação tanto com dados biogeográficos, quanto com dados taxonômicos, indica que as relações de parentesco quimiotaxonômico em *Espeletia* e a correlação entre a composição química das espécies e sua origem geográfica, são devidas, possivelmente, tanto à presença de determinados metabólitos únicos para cada agrupamento, quanto ao fato das espécies pertencentes a cada grupo apresentarem uma composição química muito semelhante, inclusive nas áreas dos seus principais picos cromatográficos.

## 4.5 Desreplicação de extratos

A metodologia descrita no item 3,7 permitiu propor a identidade de 34 diferentes metabólitos (Tabela 9), os quais podem se agrupar em três grandes classes: ácidos clorogênicos, flavonoides e diterpenos. As estruturas químicas destes compostos são apresentadas nas Figuras 35 a 39.

Os ácidos clorogênicos foram identificados por comparação com substâncias previamente isoladas pelo nosso grupo de pesquisas e de acordo com características espectrais descritas por Martucci et al. (2014). Substâncias desta classe de metabólitos apresentam um espectro no UV característico, com dois picos máximos de absorção em torno de 300 e 325 nm. Além disso, perdas de 162 *m/z*

nos seus espectros de massas indica a possível perda de uma molécula neutra de ácido caféico. Os compostos 3 e 4 (Tabela 9), apresentam uma molécula desprotonada em 353 *m/z* e um íon produto em 191 *m/z* no modo de ionização negativo, indicando a possível esterificação de uma unidade de ácido caféico com uma de ácido quínico. Logo, as substâncias 3 e 4 foram identificadas como o ácido 3-*O*-(*E*)-cafeoilquínico e 5-*O*-(*E*)-cafeoilquínico, respectivamente, mediante a coinjeção com substâncias padrão, assim como os ácidos protocatecuico e quínico (Compostos 1 e 2, respectivamente).

Os ácidos 3,4, 3,5 e 4,5-di-*O*-(*E*)-cafeoilquínicos (Tabela 9) foram igualmente identificados mediante a coinjeção com substâncias padrão. Estes metabólitos apresentam espectros de massas caracterizados pela presença de uma molécula protonada em 517 *m/z* no modo de ionização positivo, seguido da perda de uma molécula de água e um pico intenso em 163 *m/z* formado após a perda de uma unidade de ácido cafeoilquínico. O composto 6 foi identificado como metil 3-*O*-cafeoil-4-*O*-feruloilquinato baseado no seu espectro no UV e massas. Embora este composto coelui com o ácido 3,4 di-*O*-(*E*)-cafeoilquínico (Figura 40), foi possível sua identificação baseada no mecanismo de fragmentação proposto na Figura 41.

É interessante a possível presença de dois isômeros do ácido tricafeoilaltrárico (compostos 10 e 11, Tabela 9) nas amostras analisadas. Esses metabólitos foram identificados pela perda sucessiva de três unidades de ácido caféico ligadas a uma unidade de ácido galactárico, o qual apresenta um íon produto característico em 209 *m/z* no modo de ionização negativo. Devido ao fato que não havia registros na literatura sobre os possíveis fragmentos desta substância nem do metil 3-*O*-cafeoil-4-*O*-feruloilquinato, o mesmo foi proposto na Figura 41.

É importante ressaltar que este estudo constitui a primeira descrição de ácidos clorogênicos no gênero *Espeletia* e na subtribo Espeletiinae, sendo que esta classe de metabólitos é característica de várias espécies da família Asteraceae (TAKENAKA et al., 2003). Além disso, a presença do ácido tricafeoilaltrárico é muito relevante do ponto de vista quimiotaxonômico, considerando que este metabólito foi descrito apenas no gênero *Smallanthus* (TAKENAKA et al., 2003; TERADA et al., 2009), o qual é considerado grupo irmão da subtribo Espeletiinae, com base em marcadores moleculares (RAUSCHER, 2002).

Os flavonoides constituem a principal classe de metabólitos identificados, tendo sido detectados nas 120 amostras analisadas. Estes podem se subdividir em três principais categorias: agliconas (subclasses – flavanonas, flavonóis, diidroflavonóis e biflavonóis), glicosilados e esterificados com unidades de ácido caféico. Como agliconas foram identificadas a pinobanksina e a quercetina, cuja presença foi confirmada pela coinjeção com as substâncias padrão. Além disso, foram identificados dois flavonoides metoxilados correspondentes à hesperetina e à 3-metóxiquercetina, sendo este último o composto majoritário das espécies colombianas e, portanto, o principal discriminante. A 3-metóxiquercetina, confirmada pela coinjeção com substância padrão, foi descrita unicamente nas espécies *E. barclayana* e *E. killipii* em estudos fitoquímicos prévios, enquanto a quercetina tem sido descrita, além de nas duas espécies já mencionadas, em *E. schultzii* (TORRENEGRA et al., 1994; TORRENEGRA; TELLEZ, 1996; GUTIERREZ et al., 1998; DE LOS RIOS et al., 1999).

Diferentes flavonoides glicosilados foram identificados com base nos seus espectros massas e pela comparação com substâncias padrão. Dentre eles, a quercetina 3-*O*-galactosídeo corresponde a um dos mais comuns, sendo discriminante do cluster venezuelano. Estes tipos de compostos se caracterizam por apresentar espectros de massas com perdas da unidade glicosídica e um íon produto correspondente à aglicona. Por exemplo, os compostos 17-22 (Tabela 9) apresentam todos um íon produto com *m/z* 303 no modo de ionização positivo, correspondente à quercetina como aglicona, enquanto o valor de *m/z* da molécula protonada desses compostos varia de acordo com o tipo de açúcar e o número destes ligados na aglicona. Dos flavonoides glicosilados, apenas a quercetina-3-*O*-galactosídeo havia sido previamente descrita no gênero *Espeletia*, estando presente nas inflorescências de *E. schultzii* (DE LOS RIOS et al., 1999).

Foram identificados também diferentes flavonoides glicosilados e esterificados com unidades de ácido caféico. Estes apresentam espectros no UV com máximos de absorção em torno de 300 e 325 nm, característicos de ácidos clorogênicos, além de apresentarem espectros de massas com perdas de unidades de açúcar e resíduos de ácido caféico, e um pico correspondente à aglicona do flavonoide. Por exemplo, a quercetina-3-(6″-cafeoilsoforosídeo) (composto 25, Tabela 9), apresenta uma molécula protonada em *m/z* 789, além de íons produto correspondentes à

perda de uma unidade de glicose seguida de uma de ácido caféico (*m/z* 465) e um pico correspondente à aglicona (303 *m/z*) no modo positivo. Ressalta-se que este tipo de metabólitos não haviam sido previamente descritos no gênero *Espeletia* ou na subtribo Espeletiinae.

Finalmente, vários diterpenos, principalmente do tipo caurano, foram identificados em todas as amostras analisadas (compostos 28-34, Tabela 9), embora em maior proporção nas espécies pertencentes aos páramos da cordilheira central e ocidental da Colômbia. Este tipo de compostos são a principal classe de metabólitos secundários presentes na subtribo Espeletiinae (PADILLA-GONZÁLEZ et al. submetido), destacando-se o ácido grandiflorólico e o ácido *ent*-12-oxocauro-9(11),16-dien-18-oico. Este resultado confirma estudos quimiotaxonômicos prévios realizados por Alvarenga et al. (2005), onde se estabeleceu que todas as espécies do gênero *Espeletia* biossintetizam tais metabólitos.

**Tabela 9.** Substâncias desreplicadas nas amostras do gênero *Espeletia* analisadas por UHPLC-UV-MS.

| No. | $t_R$ (min) | Substância provável | Ionização no modo negativo (*m/z*) | Ionização no modo positivo (*m/z*) | UV máx (nm) |
|---|---|---|---|---|---|
| 1 | 0,73 | Ácido protocatecuico* | $[M-H]^-$ 153.01817 | - | 260, 294 |
| 2 | 0,51 | Ácido quínico* | $[M-H]^-$ 191.05531 | - | - |
| 3 | 0,60 | Ácido 3-*O*-(*E*)-cafeoilquínico* | $[M-H]^-$ 353.08755, $[(M-H) - CAF]^-$ 191.05515 | $[M+H]^+$ 355.10202, $[(M+H) - AQ]^+$ 163.03868 | 300om, 324 |
| 4 | 0,81 | Ácido 5-*O*-(*E*)-cafeoilquínico* | $[M-H]^-$ 353.08676, $[(M-H) - CAF]^-$ 191.05479 | $[M+H]^+$ 355.10196, $[(M+H) - AQ]^+$ 163.03871 | 300om, 326 |
| 5 | 8,60 | Ácido 3,4-di-*O*-(*E*)-cafeoilquínico* | $[M-H]^-$ 515.11926 | $[M+H]^+$ 517.13403, $[(M+H) - H_2O]^+$ 499.12308 $[(M+H) - AQ]^+$ 163.03886 | 300, 328 |
| 6 | 8,50 | Metil 3-*O*-cafeoil-4-*O*-feruloilquinato | $[M-H]^-$ 543.15009 | $[M+H]^+$ 545.16455, $[(M+H) - H_2O]^+$ 527.15363 $[(M+H) - H2O - CAF]^+$ 351.10657, $[(M+H) - AQ]^+$ 177.05429 | 300, 326 |
| 7 | 8,70 | Ácido 3,5-di-*O*-(*E*)-cafeoilquínico* | $[M-H]^-$ 515.11880 | $[M+H]^+$ 517.13300, $[(M+H) - H_2O]^+$ 499.12271 $[(M+H) - AQ]^+$ 163.03864 | 298, 327 |
| 8 | 9,16 | Ácido 4,5-di-*O*-(*E*)-cafeoilquínico* | $[M-H]^-$ 515.11914 | $[M+H]^+$ 517.13379, $[(M+H) - H_2O]^+$ 499.12271 $[(M+H) - AQ]^+$ 163.03865 | 300, 329 |

| | | | | | |
|---|---|---|---|---|---|
| 9 | 6,54 | Isômero do ácido dicafeoilquínico | $[M-H]^-$ 515.11896 | $[M+H]^+$ 517.13330, $[(M+H) - H_2O]^+$ 499.12268, $[(M+H) - AQ]^+$ 163.03874 | 300, 322 |
| 10 | 9,50 | Isómero do ácido tricafeoilaltrárico | $[M-H]^-$ 695.12518, $[(M-H) - CAF]^-$ 533.09363, $[(M-H) - CAF - CAF]^-$ 371.06171, $[(M-H) - CAF - CAF - CAF]^-$ 209.02954 | - | 300om, 330 |
| 11 | 9,94 | Isómero do ácido tricafeoilaltrárico | $[M-H]^-$ 695.12512, $[(M-H) - CAF]^-$ 533.09344, $[(M-H) - CAF - CAF]^-$ 371.06174, $[(M-H) - CAF - CAF - CAF]^-$ 209.02959 | - | 300om, 329 |
| 12 | 11,51 | Pinobanksina | $[M-H]^-$ 271.06125 | $[M+H]^+$ 273.07565 | 229, 290 |
| 13 | 10,49 | Quercetina* | $[M-H]^-$ 301.03482 | $[M+H]^+$ 303.05057 | 255, 300om, 360 |
| 14 | 11,01 | 3-Metóxiquercetina* | $[M-H]^-$ 315.05063 | $[M+H]^+$ 317.06677 | 255, 300om, 357 |
| 15 | 12,98 | Hesperetina | $[M-H]^-$ 301.07083 | $[M+H]^+$ 303.08707 | 228, 287 |
| 16 | 14,20 | 8,8''-Metileno bisquercetina | $[M-H]^-$ 615.07758, $[(M-H) - QER+CH_3]^-$ 299.01920 | $[M+H]^+$ 617.09143 | 235, 270, 303, 350om |
| 17 | 7,96 | Quercetina-3-*O*-galactosídeo* | $[M-H]^-$ 463.08801 | $[M+H]^+$ 465.10229, $[(M+H) - glucose]^+$ 303.04935 | 255, 290om, 346 |
| 18 | 7,83 | Rutina* | $[M-H]^-$ 609.14581 | $[M+H]^+$ 611.16028, $[(M+H) - ramnose]^+$ 465.10263, $[(M+H) - ramnose - glucose]^+$ 303.04950 | 256, 268om, 300om, 347 |
| 19 | 6,96 | Quercetina-3,4'-*O*-di-glucopiranosideo | $[M-H]^-$ 625.14044, | $[M-H]^+$ 627.15448, $[(M-H]) - glucose]^+$ 465.10197, $[(M-H]) - glucose - glucose]^+$ 303.04916 | 264, 340 |
| 20 | 8,33 | Quercetina-3-*O*-arabinosideo | $[M-H]^-$ 433.07761 | $[M+H]^+$ 435.09229, $[(M+H]) - arabinose]^+$ 303.04968 | 253, 336 |
| 21 | 8,65 | Quercetrina* | $[M-H]^-$ 447.09348 | $[M+H]^+$ 449.10831 $[(M+H) - glucose]^+$ 303.04999 | 256, 300om, 345 |

| | | | | | |
|---|---|---|---|---|---|
| 22 | 10,06 | Flavonoide diglicosilado | $[M-H]^-$ 609.12463 | $[M+H]^+$ 611.13898 [(M+H) - ramnose - glucose]$^+$ 303.04941 | 258, 315, 360 |
| 23 | 7,49 | Quercetina-3-*O*-(6"-malonil, glucosídeo)-7-*O*-glucosídeo | $[M-H]^-$ 711.14093, [(M+H) - manoilglusídeo]$^-$ 463.08780, [(M+H) - manoilglucose - glucose]$^-$ 301.03522 | $[M+H]^+$ 713.15527 | 250, 296, 331 |
| 24 | 8,65 | Isoramnetina-3-glicosídeo | $[M-H]^-$ 477.10315 | $[M+H]^+$ 479.11743, [(M+H) - glicosídeo]$^+$ 317.06424 | 253, 300om, 336 |
| 25 | 8,45 | Quercetina-3-(6″-cafeoilsoforosídeo) | $[M-H]^-$ 787.17261 | $[M+H]^+$ 789.18677, [(M+H) - glucose - CAF]$^+$ 465.10205 [(M+H) - glucose - CAF - glucose]$^+$ 303.04938 | 247, 295om, 327 |
| 26 | 8,96 | Castillicetina | $[M-H]^-$ 463.08838 | $[M+H]^+$ 465.10306, [(M+H) - CAF]$^+$ 303.04941 [(M+H) - QER]$^+$ 163.03873 | 251, 268om, 300om, 332 |
| 27 | 9,52 | Quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo | $[M-H]^-$ 625.11938 | $[M+H]^+$ 627.13318, [(M+H) - galactose - CAF]$^+$ 303.04910 | 250, 300om, 332 |
| 28 | 14,77 | Ácido *ent*-12-oxocauro-9(11),16-dien-18-oico | $[M-H]^-$ 313.18039 | $[M+H]^+$ 315.19614 | 247 |
| 29 | 15,51 | Ácido *ent*-12-hidroxicauro-9(11),16-dien-18-oico | $[M-H]^-$ 315.19577 | - | 245 |
| 30 | 18,16 | Ácido grandiflorólico | $[M-H]^-$ 317.21155 | - | - |
| 31 | 18,80 | Ácido *ent*-16α-hidroxicauran-19-oico | $[M-H]^-$ 319.22787 | - | - |
| 32 | 22,84 | Ácido pimárico | $[M-H]^-$ 301.21680 | - | - |
| 33 | 23,34 | Ácido *ent*-15-*O*-iso-butiroxicaur-16-en-19-oico | $[M-H]^-$ 387.25363 | - | - |
| 34 | 24,14 | Ácido *ent*-15-*O*-senecioxicaur-16-en-19-oico | $[M-H]^-$ 399.25351 | - | - |

* Substância confirmada por coinjeção com padrão; AQ: ácido cafeoilquínico; CAF: Ácido caféico; QER: quercetina; om: ombro.

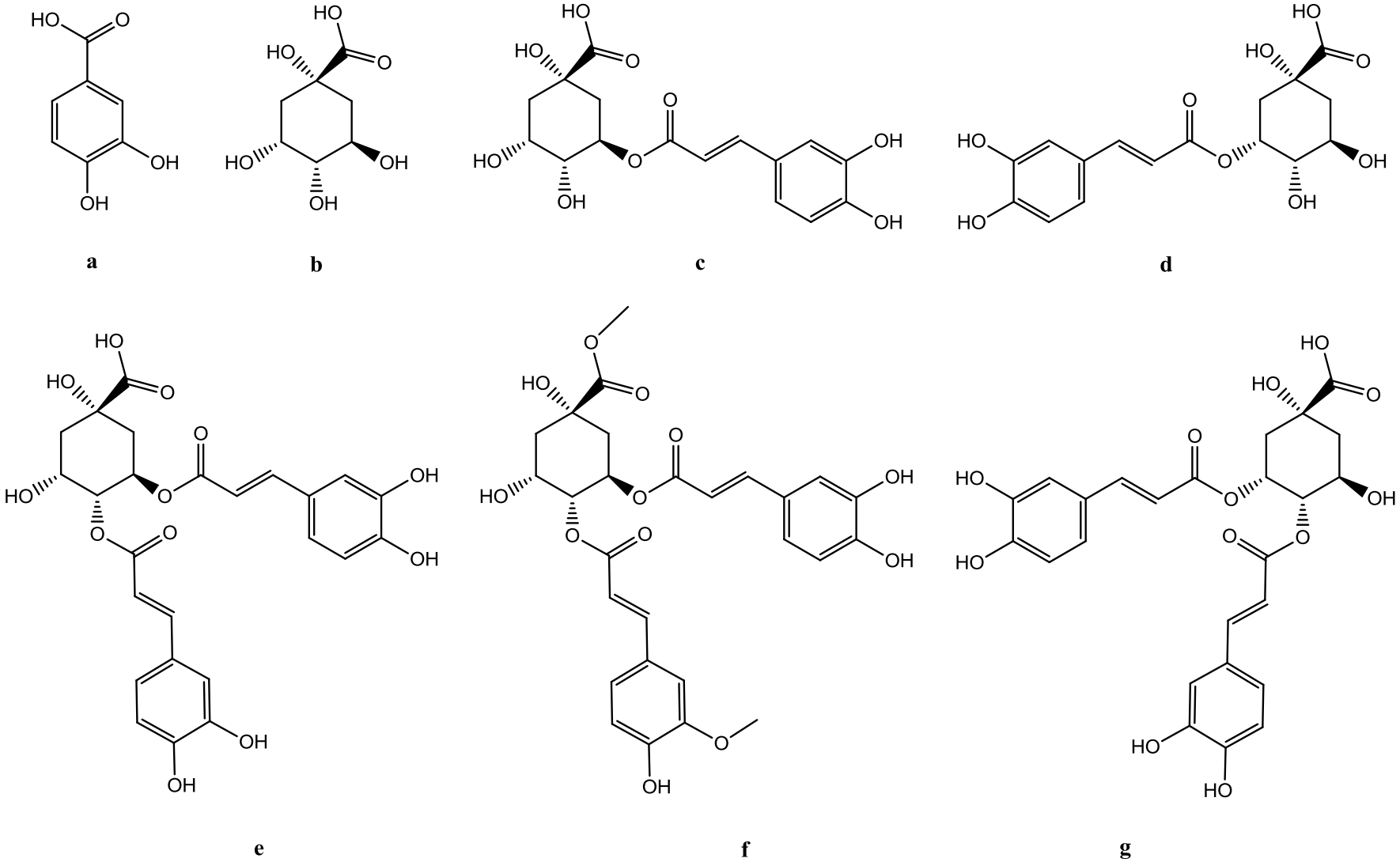
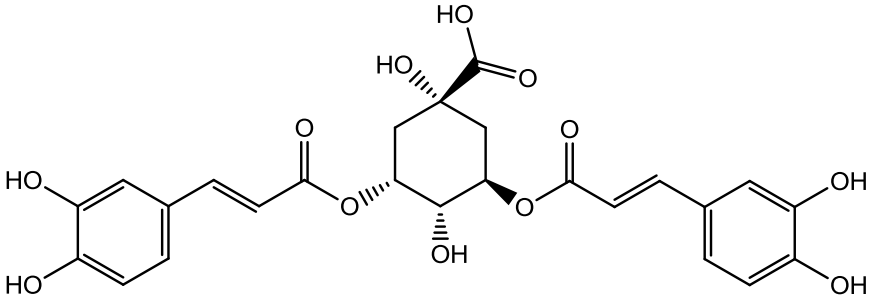
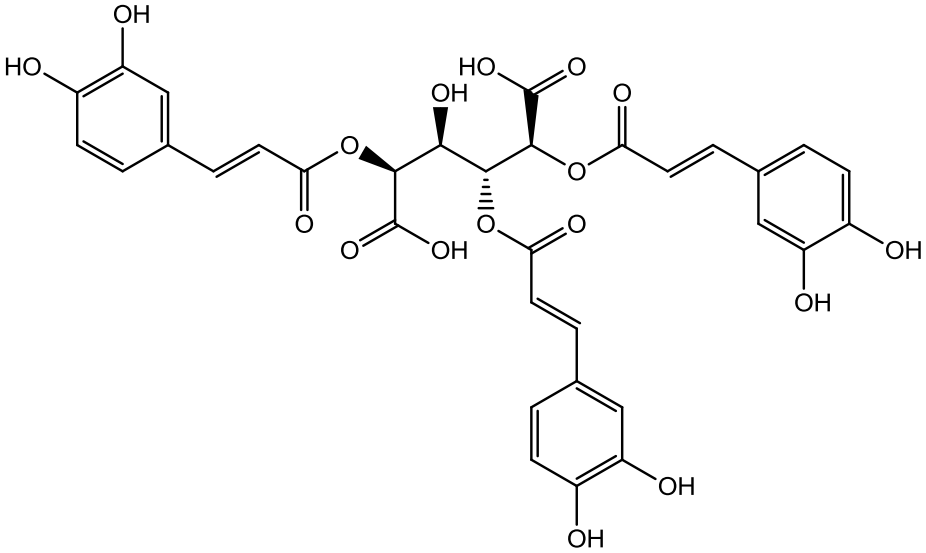



**Figura 35.** Estruturas químicas dos ácidos clorogênicos desreplicadas nas espécies de *Espeletia*. a. ácido protocatecuico; b. ácido quínico; c. ácido 3-*O*-(*E*)-cafeoilquínico; d. ácido 5-*O*-(*E*)-cafeoilquínico; e. ácido 3,4-di-*O*-(*E*)-cafeoilquínico; f. metil 3-*O*-cafeoil-4-*O*-feruloilquinato; g. ácido 4,5-di-*O*-(*E*)-cafeoilquínico; h. ácido 3,5-di-*O*-(*E*)-cafeoilquínico; i. ácido tricafeoilaltrárico.




**h**




**i**

(Continuação da Figura 35).

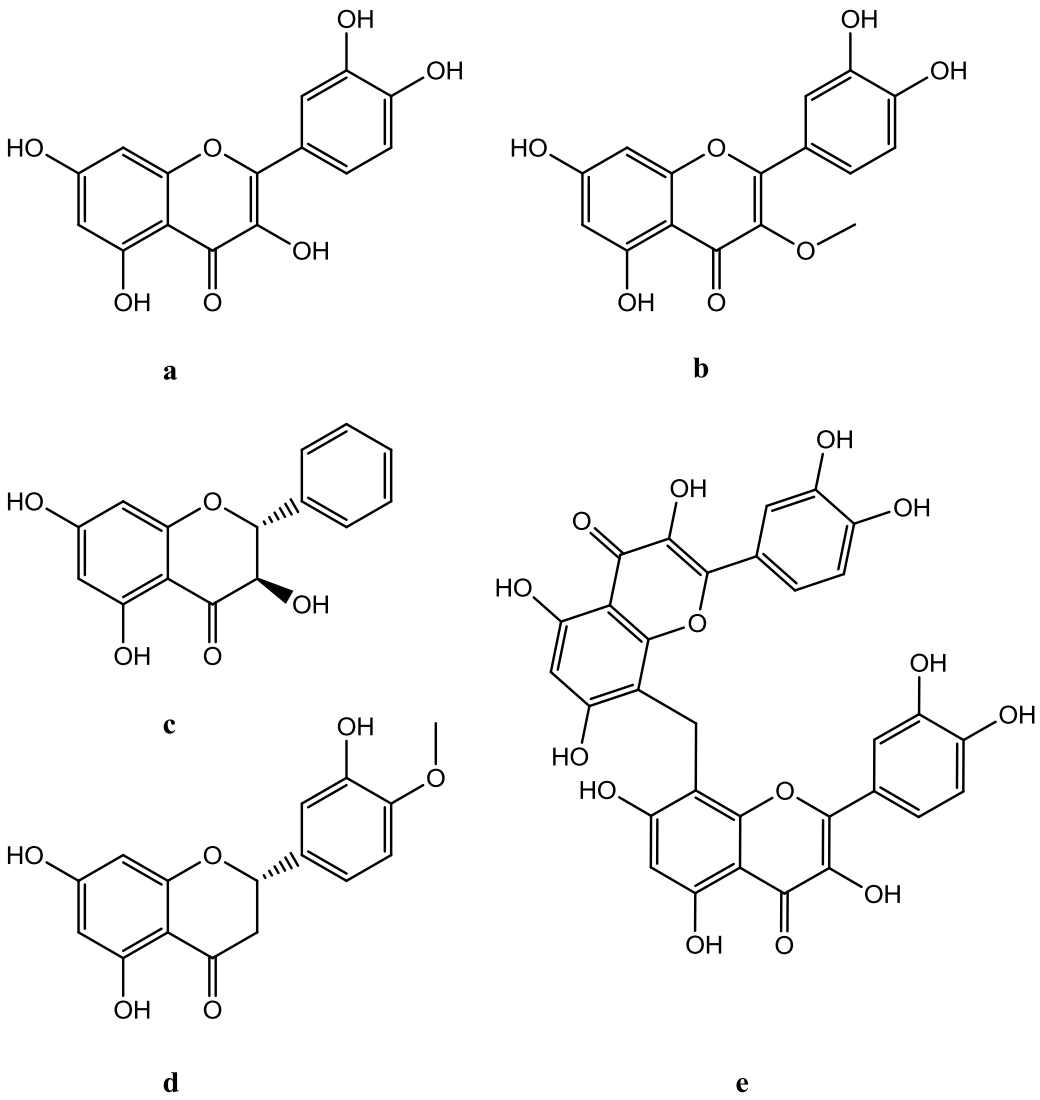



**Figura 36.** Estruturas químicas dos flavonoides (agliconas) desreplicados nas espécies de *Espeletia*. a. quercetina; b. 3-metóxiquercetina; c. pinobanksina; d. hesperetina; e. 8,8''-metilenobisquercetina.

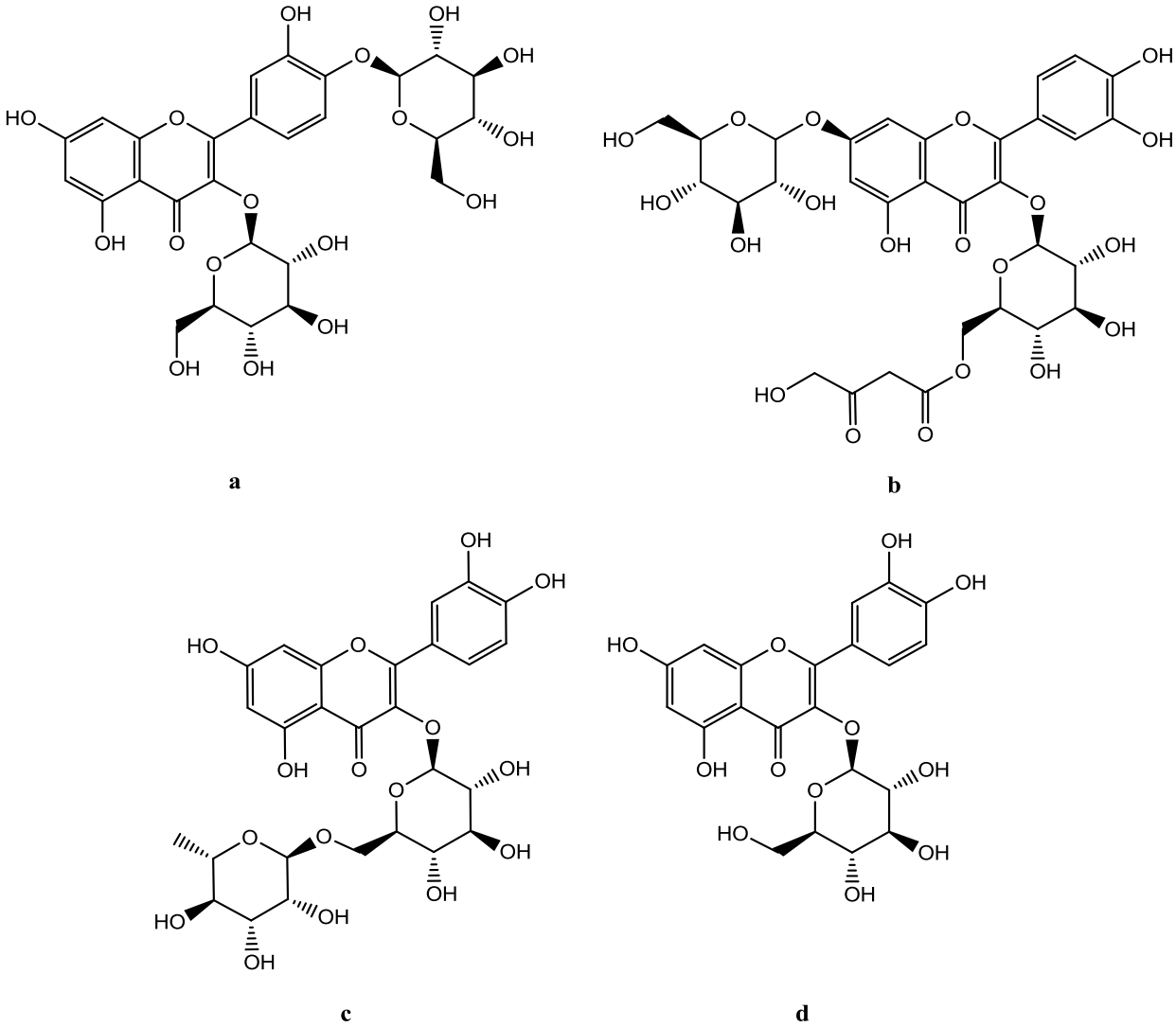
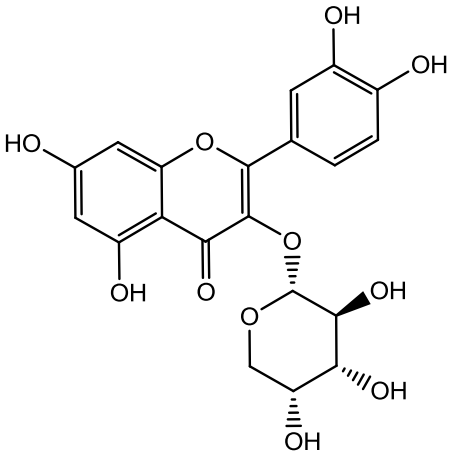
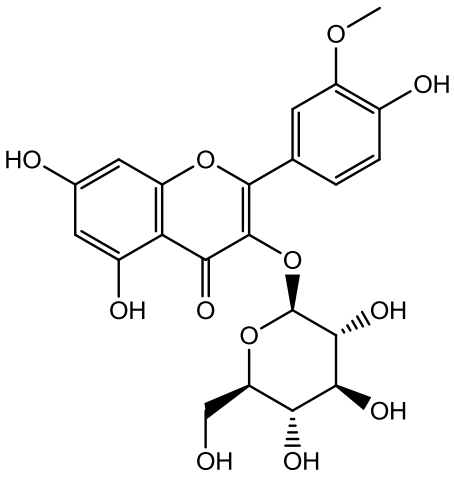


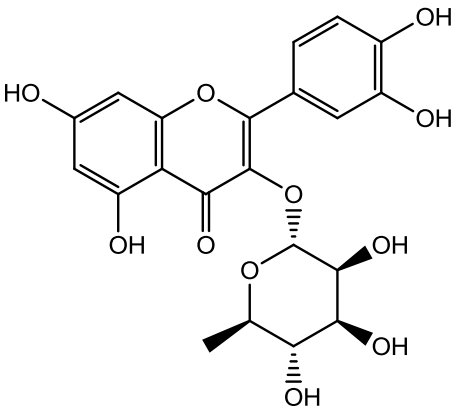

**Figura 37.** Estruturas químicas dos flavonoides glicosilados desreplicados nas espécies de *Espeletia*. a. quercetina-3,4'-*O*-di-glucopiranosídeo; b. quercetina-3-*O*-(6''-malonil-glucosídeo)-7-*O*-glucosídeo; c. rutina; d. quercetina-3-*O*-galactosídeo; e. quercetina-3-*O*-arabinosídeo; f. Isoramnetina-3-glicosídeo; g. quercetrina.




**e**

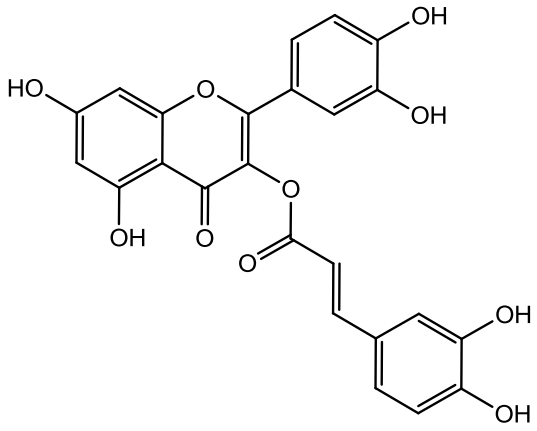



**f**

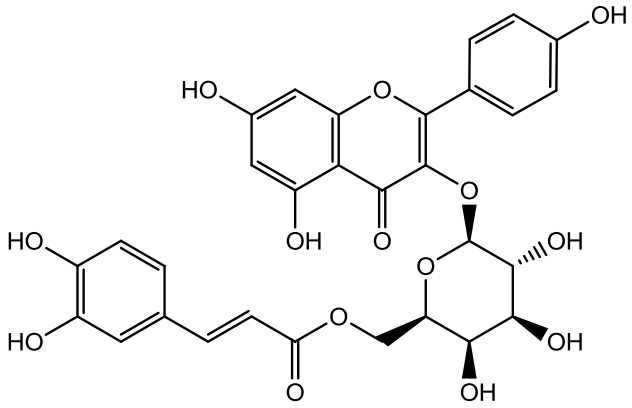



**g**

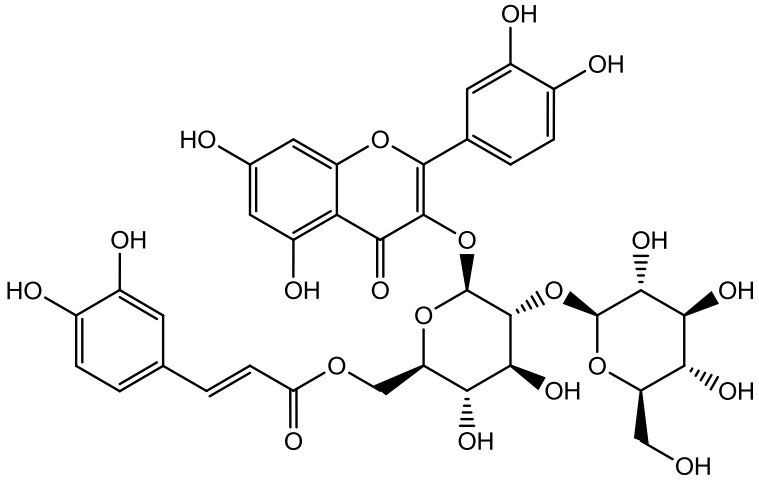
(Continuação da Figura 37).




**a**




**b**




**c**

**Figura 38.** Estruturas químicas dos flavonoides esterificados com ácido caféico desreplicados nas espécies de *Espeletia*. a. castillicetina; b. quercetina-3-*O*-(6-*O*-cafeoil)-galactosídeo; c. quercetina-3-(6″-cafeoilsoforosídeo).

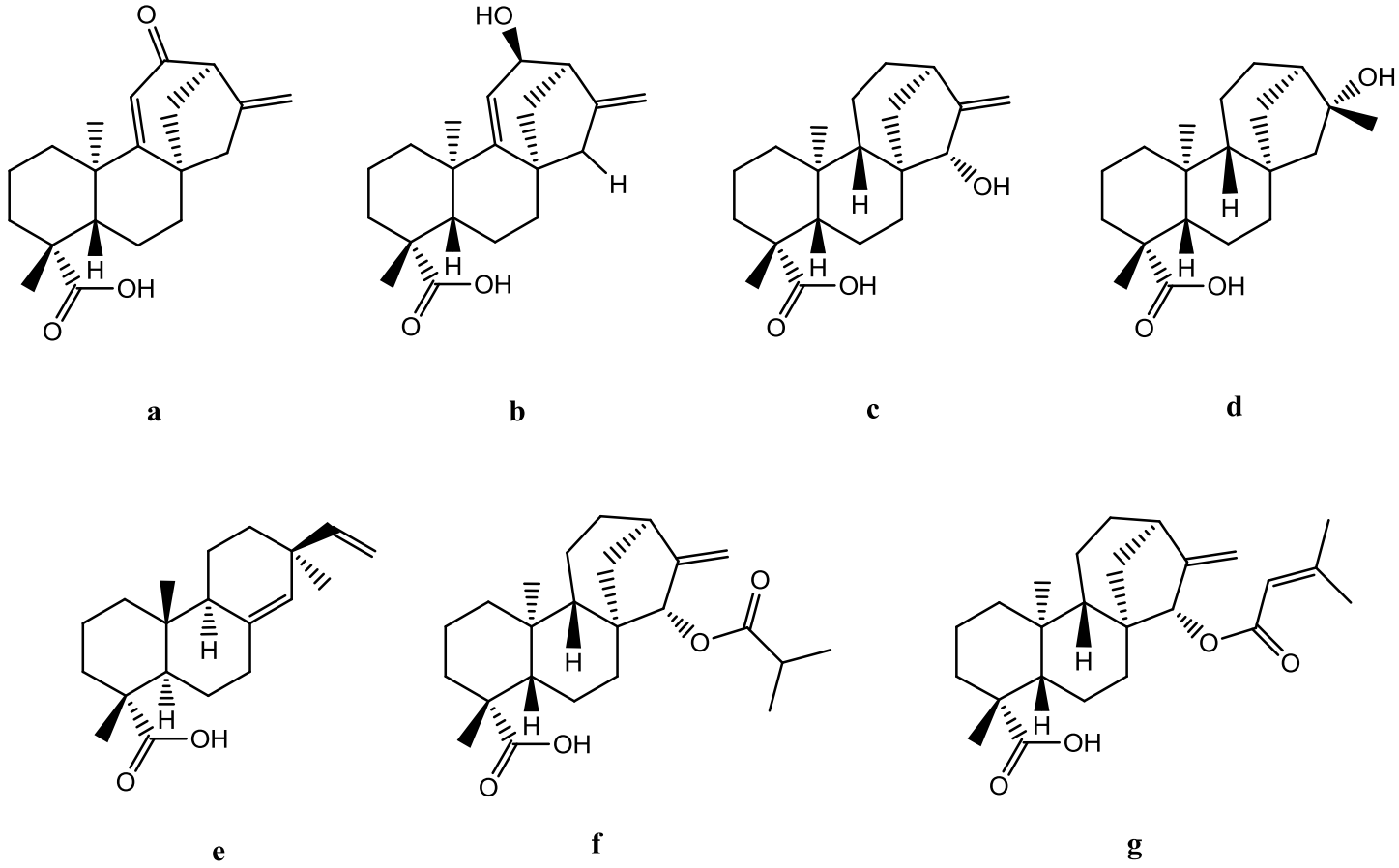



**Figura 39.** Estruturas químicas dos diterpenos desreplicados nas espécies de *Espeletia*. a. ácido *ent*-12-oxocauro-9(11),16-dien-18-oico; b. ácido *ent*-12-hidroxicauro-9(11),16-dien-18-oico; c. ácido grandiflorólico; d. ácido *ent*-16α-hidroxicauran-19-oico; e. ácido pimárico; f. ácido *ent*-15-O-isobutiroxicaur-16-en-19-oico; g. ácido *ent*-15-O-senecioxicaur-16-en-19-oico.

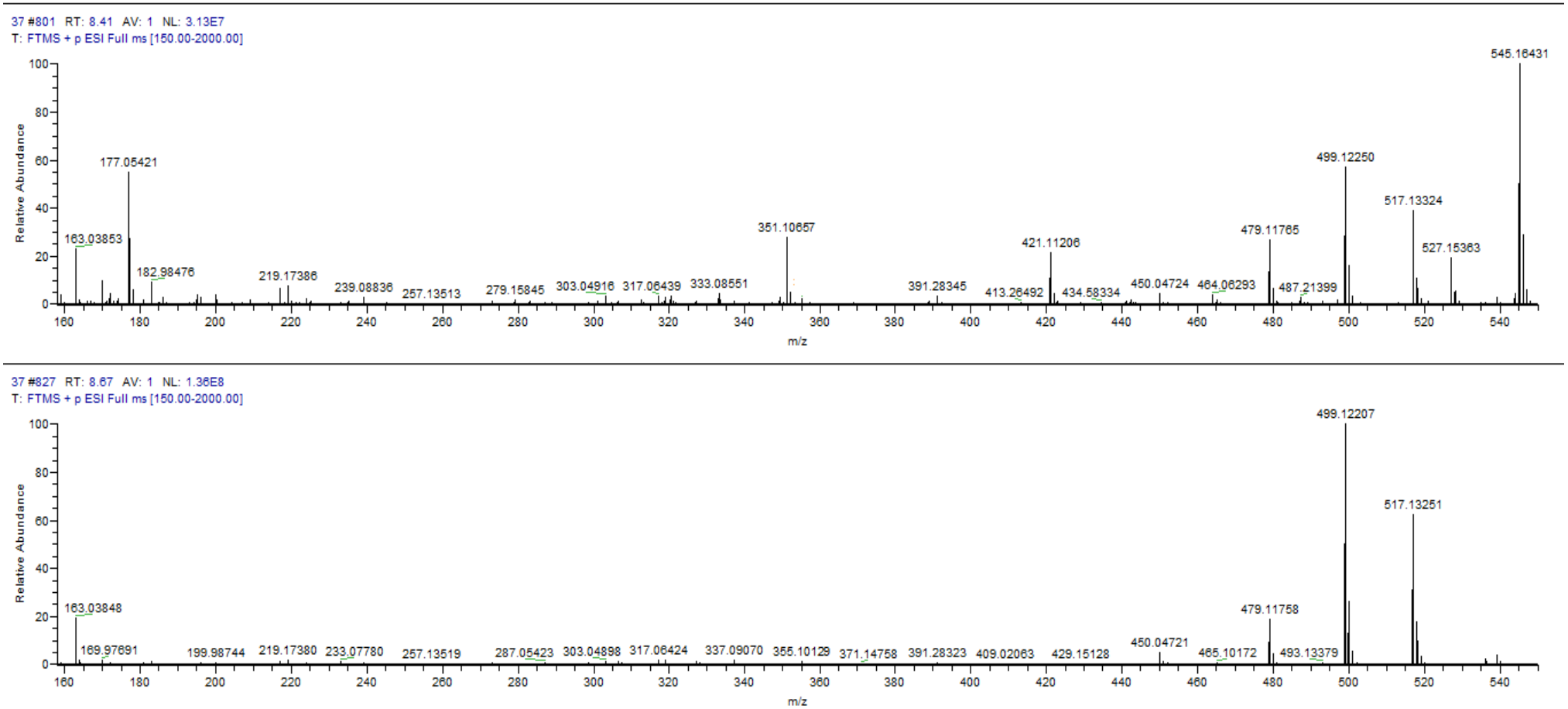



**Figura 40.** Espectros de massas no modo positivo do metil 3-*O*-cafeoil-4-*O*-feruloilquinato (acima) e do ácido 3,4-di-*O*-(*E*)-cafeoilquínico (abaixo).

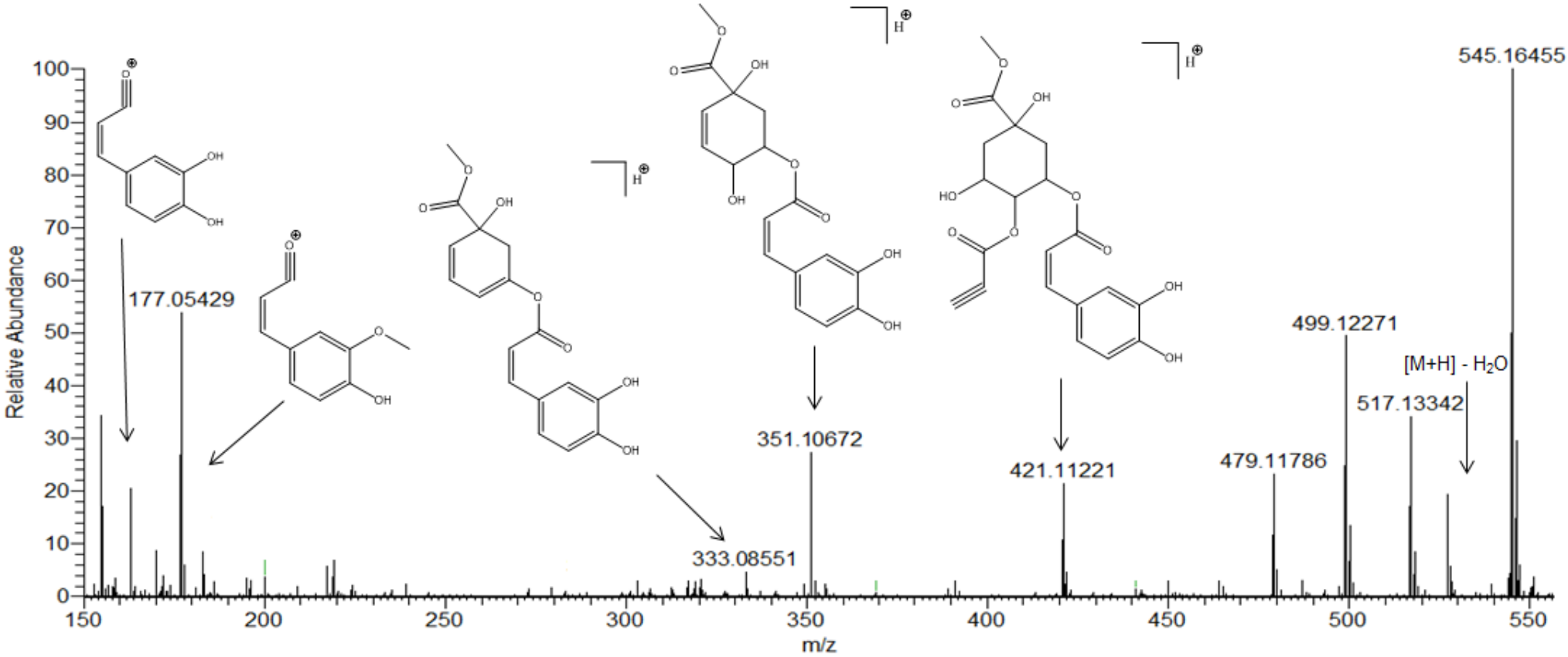



**Figura 41.** Espectro de massas no modo positivo e possíveis fragmentos do metil 3-*O*-cafeoil-4-*O*-feruloilquinato.

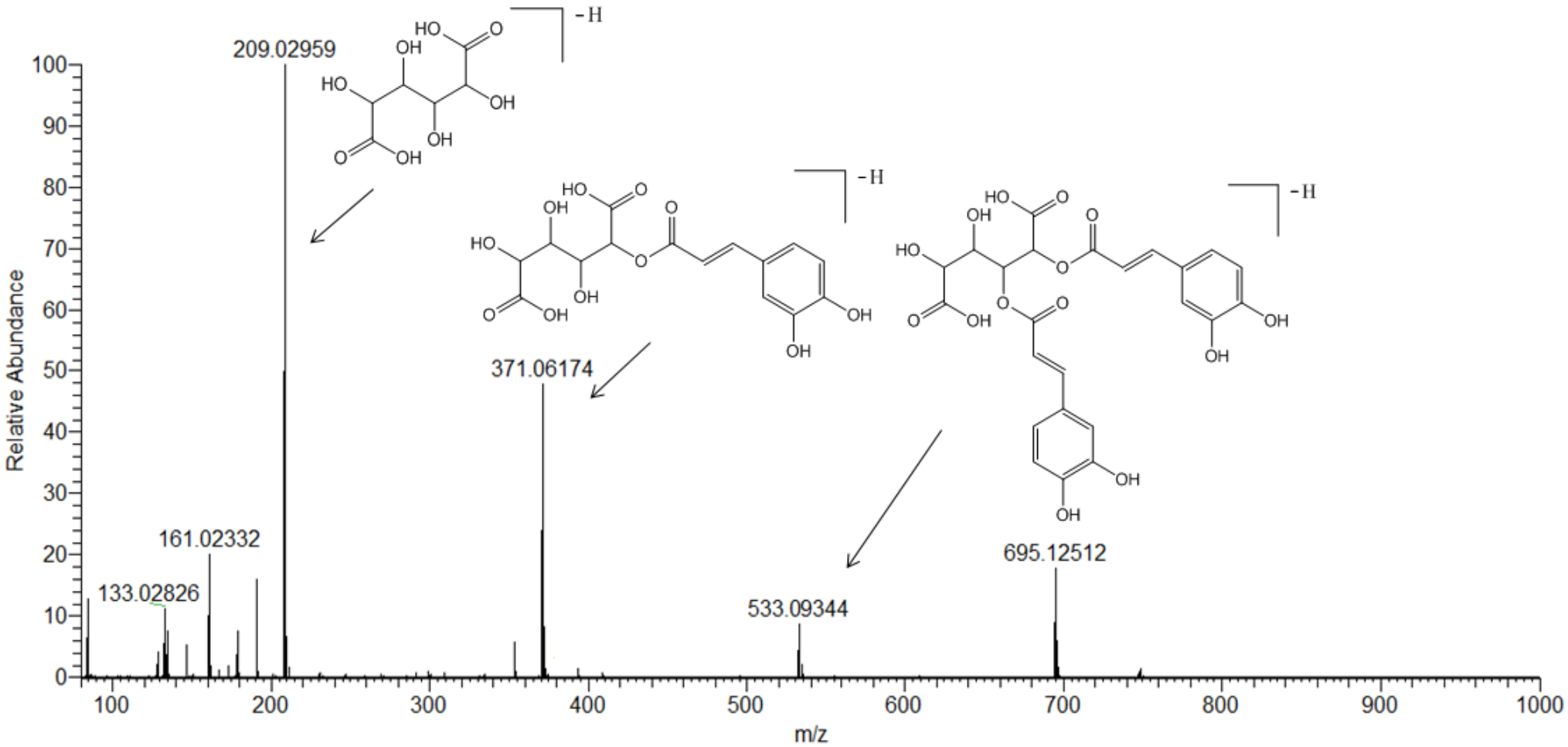



**Figura 42.** Espectro de massas no modo negativo e possíveis fragmentos do ácido tricafeoilaltrárico.

## 4.6 Análise filogenética

As análises filogenéticas por MP e ML foram efetuadas tanto com a matriz do modo negativo, quanto com a do modo positivo, separadamente e em conjunto, obtendo topologias semelhantes. Dos 2554 caracteres considerados no modo negativo, 1502 foram informativos do ponto de vista filogenético, enquanto que no modo positivo 1882 foram informativos a partir de 3661 caracteres iniciais.

Tanto na árvore de consenso estrito gerada pela análise de MP (Figura 43), quanto na árvore de ML (Figura 44), observa-se um agrupamento condizente com as análises estatísticas multivariadas (OPLS-DA), onde as espécies se agruparam de acordo com seu país de origem (Figuras 24 e 26). Porém, na análise filogenética, foi interessante o fato das espécies venezuelanas terem se agrupado na base do dendrograma (junto com *E. incana* de origem colombiana), indicando que com base em dados químicos este grupo de espécies corresponde ao clado ancestral do gênero *Espeletia.* Este resultado concorda com hipóteses prévias sobre a origem geográfica da subtribo Espeletiinae, já que segundo Cuatrecasas (1986, 2013) a cordilheira de Mérida na Venezuela corresponde possivelmente ao local de origem e diversificação primária da subtribo, considerando que nessa região ocorre a maior diversidade de Espeletiinae, com seis gêneros e 63 espécies endêmicas. Esta hipótese foi confirmada recentemente na análise filogenética da subtribo com base em diferentes marcadores moleculares (DIAZGRANADOS, 2012).

A pesar da consistência observada entre as filogenias geradas pelos métodos de MP (Figura 43) e ML (Figura 44) e sua semelhança com a filogenia baseada em marcadores moleculares (Figuras 1 e 2, Anexo I), os resultados obtidos devem ser considerados com cautela, já que ao determinar um grau de suporte superior a 50% mediante análises por *bootstrap* e *jackknife* a maioria dos nodos das árvores colapsaram formando politomias. Porém, um dos clados recuperado foi o conformado pelas espécies *E. schultzii 101, E. schultzii 102, E. aristeguietana 11, E. aristeguietana 12, E. jajoensis 58* e *E. jajoensis 59* (todas de origem venezuelana) com um valor de *bootstrap* de 67 (não apresentado). Isto indica que existe um baixo suporte aos agrupamentos observados, considerando que os métodos de *bootstrap* e *jackknife* são técnicas de re-amostragem estatística amplamente utilizadas em análises filogenéticas em que as variáveis no conjunto de dados, neste caso os

picos cromatográficos, são re-amostradas um determinado número de vezes (1000 vezes neste estudo), onde algumas variáveis são excluídas de cada análise e substituídas por outras presentes no conjunto de dados (bootstrap) (FELSENSTEIN, 1985) ou simplesmente excluídas (jackknife) (FARRIS, 1996).

O baixo suporte dos agrupamentos nas análises por *bootstrap* e *jackknife* sugere um baixo sinal filogenético na matriz inicial, o qual pode ser devido ao alto nível de homoplasias químicas (caracteres similares adquiridos em eventos evolutivos independentes) presentes no conjunto de dados (Tabela 10). Porém, segundo Sánchez (2005), é possível que incluindo um maior número de táxons na análise (por exemplo, considerando os outros gêneros da subtribo Espeletiinae e não somente ao gênero *Espeletia*) e excluindo caracteres pouco informativos, consiga-se aumentar o número de sinapomorfias químicas (caracter compartilhado por clados irmãos), diminuindo o efeito das homoplasias. Além disso, devido ao fato da espécie *Senecio niveaureus* (Asteraceae, Senecioneae), escolhida como grupo externo por ser morfológica e geograficamente próxima de *Espeletia*, apresentar uma composição química muito distante deste gênero, é possível que esteja influenciando a topologia e o suporte da árvore, por se tratar de um táxon muito distante em termos evolutivos do gênero, minimizando as diferenças interespecíficas nas espécies analisadas.

**Tabela 10.** Critérios estatísticos obtidos nas análises filogenéticas por máxima parcimônia (MP) e *maximum likelihood* (ML).

| Critério estatístico | MP | ML |
|---|---|---|
| Comprimento da árvore | 15390 | 15403 |
| Índice de consistência (IC) | 0,1316 | 0,1315 |
| Índice de homoplásia (IH) | 0,8684 | 0,8685 |
| IC após excluir caracteres não informativos | 0,1010 | 0,1009 |
| IH após excluir caracteres não informativos | 0,8990 | 0,8991 |
| Índice de retenção (IR) | 0,4244 | 0,4239 |
| Índice de consistência re-escalonado (CR) | 0,0558 | 0,0557 |

É importante ressaltar que mesmo empregando diferentes tipos de marcadores moleculares na reconstrução da história evolutiva da subtribo Espeletiinae, neste grupo de espécies há um alto índice de homoplasias (variando entre 0,54 e 0,89 de acordo com o marcador empregado), sendo um fenômeno comum em Espeletiinae (DIAZGRANADOS, 2012; SÁNCHEZ, 2005). Portanto, é

necessário escolher caracteres altamente informativos na reconstrução da história evolutiva deste grupo. Além disso, ressalta-se que este corresponde ao primeiro estudo em que se reporta o uso de dados metabolômicos na reconstrução da história evolutiva de um organismo.

Embora a distribuição de muitos metabólitos secundários nas plantas seja restrita a determinados grupos taxonômicos, premissa na qual se baseia a quimiossistemática, é comum que plantas não relacionadas filogeneticamente entre si apresentem uma composição química semelhante. Segundo Wink et al. (2010), isto pode ser devido à convergência evolutiva, transferência horizontal de genes ou à presença de microrganismos endofíticos. Estes fatores dificultam seriamente análises filogenéticas baseadas unicamente em dados químicos e têm sido a principal causa do declive na incorporação de substâncias químicas em análises filogenéticas modernas. No entanto, a pesar das limitações no uso de metabólitos secundários na reconstrução da história evolutiva de um organismo, este tipo de enfoque é valido na tentativa de entender por que as filogenias baseadas em substâncias químicas diferem, se for o caso, daquelas baseadas em dados morfológicos ou moleculares (CALABRIA et al., 2006). Além disso, em alguns casos os dados químicos têm sido úteis no esclarecimento de relações taxonômicas em diferentes categorias incluindo tribos, gêneros e espécies (ALVARENGA et al., 2001).

Segundo Larsson (2007), a quimiossistemática moderna envolve um enfoque multidisciplinar onde a biologia molecular é empregada na construção da árvore filogenética de um indivíduo, enquanto os caracteres químicos são mapeados nessa árvore e empregados na reconstrução dos caracteres ancestrais, permitindo determinar a proximidade evolutiva de espécies que produzem e acumulam uma determinada classe de metabólitos. Este enfoque apresenta a vantagem de eliminar a alta variabilidade e convergência presentes na construção de filogenias unicamente com base em dados químicos (REYNOLDS, 2007), além de aportar informações do por que ou como um determinado táxon evoluiu.

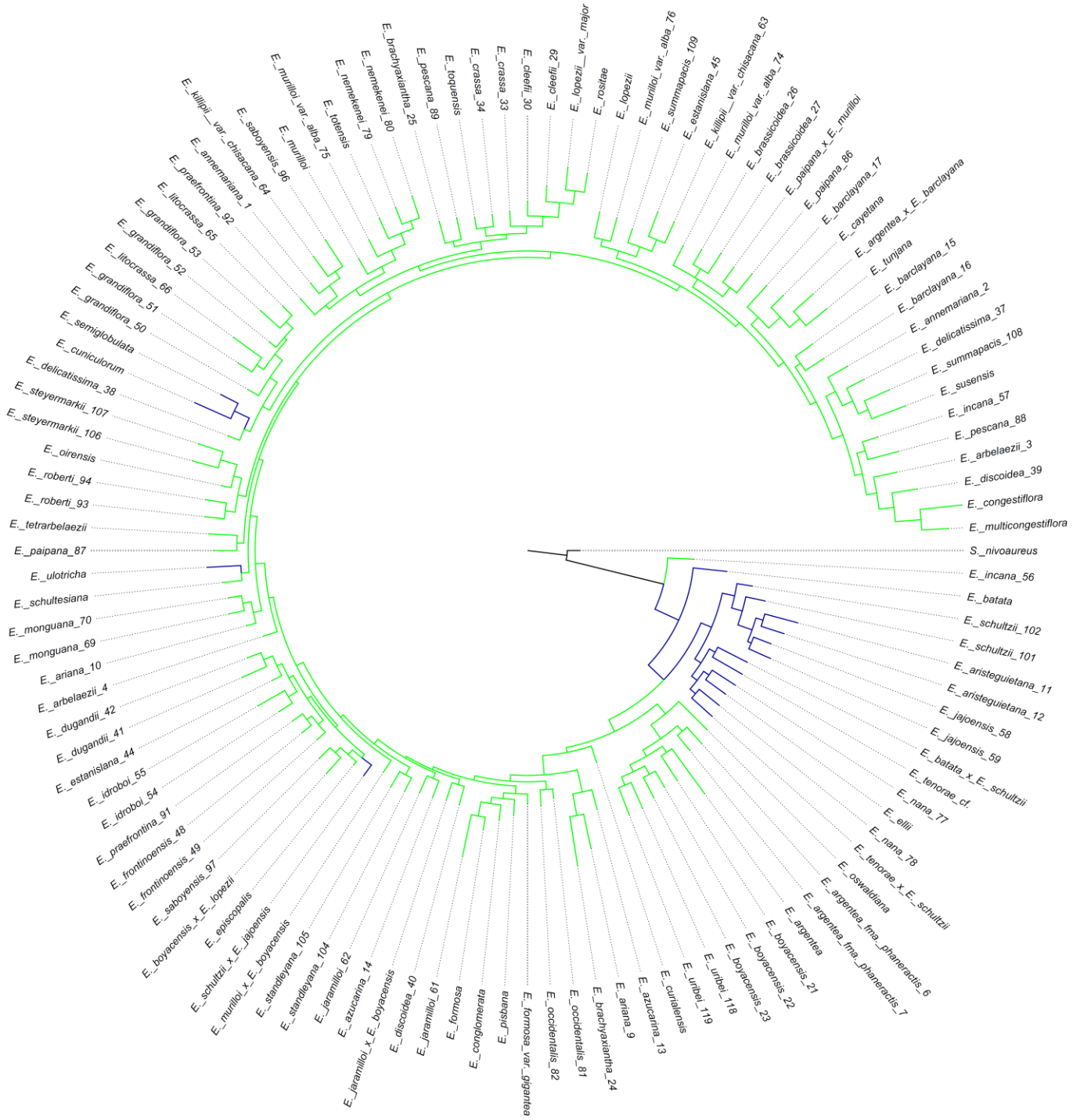



**Figura 43.** Árvore de consenso estrito obtida a partir de dados químicos gerados na análise de 120 indivíduos do gênero *Espeletia* por UHPLC-UV-MS (modo negativo). *Outgroup*: *Senecio nivoaureus* (Asteraceae, Senecioneae). Em azul espécies de origem venezuelana e em verde espécies de origem colombiana.

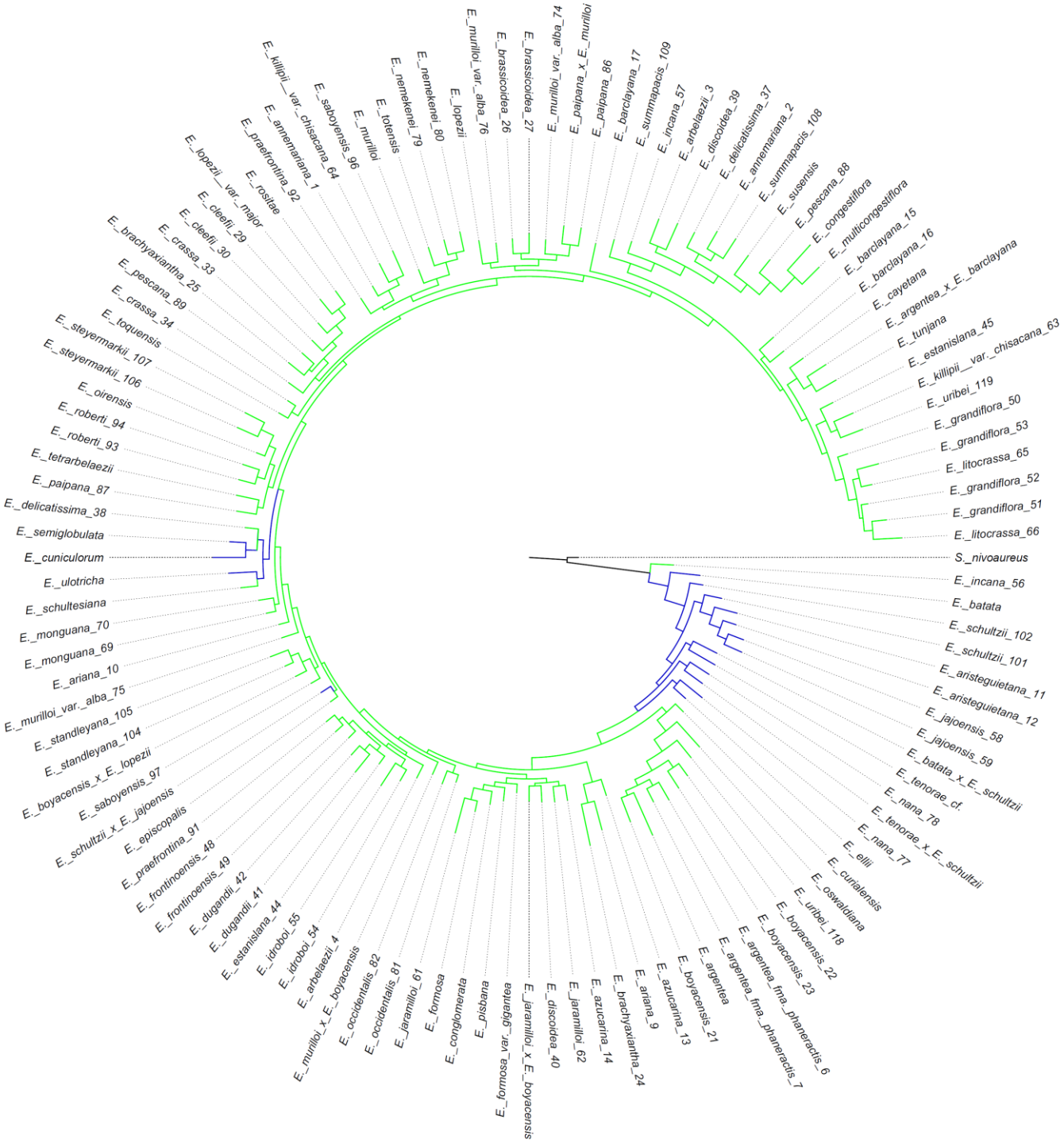



**Figura 44.** Árvore de *maximum likelihood* obtida a partir de dados químicos gerados na análise de 120 indivíduos do gênero *Espeletia* por UHPLC-UV-MS (modo negativo). *Outgroup*: *Senecio nivoaureus.* (Asteraceae, Senecioneae). Em azul espécies de origem venezuelana e em verde espécies de origem colombiana.

# 5 CONCLUSÕES

Neste trabalho, além de se apresentar a primeira revisão da química do metabolismo secundário da subtribo Espeletiinae (Asteraceae, Millerieae) e criar um banco de dados específico para este grupo (ChemEsp), apresenta-se o primeiro estudo metabolômico do gênero *Espeletia*, sendo estudadas um total de 61 espécies, nove híbridos, cinco variedades e uma forma (sem considerar réplicas), totalizando 76 táxons, dos quais 65 foram estudados pela primeira vez. A metodologia empregada permitiu desreplicar dezenas de substâncias nas espécies analisadas, muitas das quais, de valor quimiotaxonômico, não haviam sido reportadas previamente na subtribo Espeletiinae. Além disso, revelou-se uma forte influência da origem geográfica na composição química das espécies, onde as amostras se agruparam de acordo com seu país de origem numa escala global e de acordo com o maciço de páramos onde foram coletadas numa escala regional. Uma análise filogenética efetuada com dados metabolômicos revelou, embora com pouco suporte, dois principais clados correspondentes ao país de origem das espécies, em que as espécies da Venezuela correspondem ao clado ancestral do gênero. Este resultado concorda com a filogenia atual da subtribo Espeletiinae baseada em diversos marcadores moleculares. Estudos de correlação *in silico* permitiram não apenas identificar as substâncias responsáveis pelo agrupamento das espécies de acordo com sua origem geográfica, mas também a criação de modelos capazes de predizer a origem geográfica de novos extratos ainda não investigados. Isto pode ser de grande utilidade para o esclarecimento de relações taxonômicas e filogenéticas em *Espeletia,* considerando o alto grau de endemismo das espécies deste gênero.

# 6 REFERÊNCIAS

ALVARENGA, S. A. V.; FERREIRA, M. J. P.; EMERENCIANO, V. P.; CABROL-BASS, D. Chemosystematic studies of natural compounds isolated from Asteraceae: characterization of tribes by principal component analysis. **Chemometrics and Intelligent Laboratory Systems**, v. 56, p. 27–37, 2001.

ALVARENGA, S. A. V.; FERREIRA, M. J. P.; RODRIGUES, G. V.; EMERENCIANO, V. P. A general survey and some taxonomic implications of diterpenes in the Asteraceae. **Botanical Journal of the Linnean Society**, v. 147, n. 3, p. 291–308, 2005.

ALVAREZ-SANCHEZ, B.; PRIEGO-CAPOTE, F.; LUQUE DE CASTRO, M. D. Metabolomics analysis II. Preparation of biological samples prior to detection. **Trends in Analytical Chemistry**, v. 29, n. 2, p. 120–127, 2010.

BAILEY, N. J. C.; WANG, Y.; SAMPSON, J.; et al. Prediction of anti-plasmodial activity of *Artemisia annua* extracts: application of $^{1}$H NMR spectroscopy and chemometrics. **Journal of Pharmaceutical and Biomedical Analysis**, v. 35, n. 1, p. 117–26, 2004.

BOHLMANN, F.; SUDING, H.; CUATRECASAS, J.; KING, R. M.; ROBINSON, H. New diterpenes from subtribe Espeletiinae. **Phytochemistry**, v. 19, p. 267–271, 1980.

BOHLMANN, F.; SUDING, H.; CUATRECASAS, J.; ROBINSON, H.; KING, R. M. Tricyclic sesquiterpenes and further diterpenes from *Espeletiopsis* species. **Phytochemistry**, v. 19, p. 2399–2403, 1980.

BRERETON, R. G. **Chemometrics: data analysis for the laboratory and chemical plant**. Chichester, England: John Wiley & Sons, Ltd, 2003.

BUYTAERT, W.; CÉLLERI, R.; DE BIÈVRE, B.; et al. Human impact on the hydrology of the Andean páramos. **Earth-Science Reviews**, v. 79, n. 1-2, p. 53–72, 2006.

CALABRIA, L. M.; EMERENCIANO, V. P.; FERREIRA, M. J. P.; SCOTTI, M. T.; MABRY, T. J. A phylogenetic analysis of tribes of the Asteraceae based on phytochemical data. **Natural Product Communications**, v. 0, n. 0, p. 1–9, 2006.

CASOTI, R.; CHAGAS-PAULA, A. D.; OLIVEIRA, T. B.; COSTA, F. B. DA. Experimental design to optimize *Baccharis* sample extraction in metabolomic studies. 4th Brazilian Conference on Natural Products (4th BCNP) and the XXX Annual Meeting on Micromolecular Evolution, Systematics and Ecology (XXX RESEM). **Anais do Evento**. p.PS–02 16, 2013. Natal, Brazil: SBQ.

CHAGAS-PAULA, D. **Estudos metabolômicos de Asteraceae por UPLC-UV-HRFTMS, avaliação do potencial anti-inflamatório *in vitro* e suas correlações através de métodos *in silico***, 2013. Tese (Doutorado em Ciências) Departamento de Ciências Farmacêuticas, Universidade de Sao Paulo, Ribeirão Preto, SP, Brasil.

CUATRECASAS, J. Plantae Colombianae novae. **Trabajos del Museo Nacional de Ciencias Naturales Serie Botanica**, v. 26, p. 1–31, 1933.

CUATRECASAS, J. A new subtribe in the Heliantheae (Compositae): Espeletiinae. **Phytologia**, v. 35, p. 43–61, 1976.

CUATRECASAS, J. Speciation and radiation of the Espeletiinae in the Andes. In: F. Vuilleumier; M. Monasterio (Eds.); **High Altitude Tropical Biogeography**. p.267–303, 1986. New York, USA: Oxford University Press.

CUATRECASAS, J. **A systematic study of the subtribe Espeletiinae: Heliantheae, Asteraceae**. New York, USA: The New York Botanical Garden Press, 2013.

DE LOS RIOS, C.; HIDALGO, D. B.; CONTRERAS, Q.; CRESCENTE, O.; CASERTA, A. Phytochemical evaluation and antibacterial activity of *Espeletia schultzii* (Asteraceae) inflorescences. **Ciencia (Maracaibo)**, v. 7, p. 72–77, 1999.

DE VOS, R. C.; MOCO, S.; LOMMEN, A.; et al. Untargeted large-scale plant metabolomics using liquid chromatography coupled to mass spectrometry. **Nature Protocols**, v. 2, n. 4, p. 778–791, 2007.

DIAZGRANADOS, M. **Phylogenetic and biogeographic relationships of frailejones (Espeletiinae, Compositae): an ongoing radiation in the tropical andes**, 2012. Tese (Doctor of Phylosophy), Saint Louis University, St. Louis, Missouri, USA.

FARRIS, J. Parsimony jackknifing outperforms neighbor-joining. **Cladistics**, v. 12, n. 2, p. 99–124, 1996.

FELSENSTEIN, J. Confidence limits on phylogenies: an approach using the bootstrap. **Evolution**, v. 39, n. 4, p. 783, 1985.

FIEHN, O. Metabolomics – the link between genotypes and phenotypes. **Plant Molecular Biology**, v. 48, p. 155–171, 2002.

GOLBRAIKH, A.; TROPSHA, A. Beware of $q^2$! **Journal of Molecular Graphics and Modelling**, v. 20, n. 4, p. 269–276, 2002.

GUTIERREZ, S. R.; FUENTES, O.; TELLEZ, A. N.; TORRENEGRA, R. Active antibacterial principle from *Espeletia barclayana*. **Revista Latinoamericana de Química**, v. 26, p. 71–74, 1998.

HANDS, S.; EVERITT, B. A Monte carlo study of the recovery of cluster structure in binary data by hierarchical clustering techniques. **Multivariate Behavioral Research**, v. 22, n. 2, p. 235–243, 1987.

HOOGHIEMSTRA, H.; VAN DER HAMMEN, T. Quaternary Ice-Age dynamics in the Colombian Andes: developing an understanding of our legacy. **Philosophical Transactions of the Royal Society of London**, v. 359, p. 173–180, 2004.

HOOGHIEMSTRA, H.; RAN, E. T. H. Late and middle pleistocene climatic change and forest development in Colombia: pollen record Funza II (2-158 m core interval). **Palaeogeography, Palaeoclimatology, Palaeoecology**, v. 109, p. 211–246, 1994.

IBÁÑEZ, J.; USUBILLAGA, A. Estudio de la composición del aceite esencial de un frailejón híbrido entre *Espeletia schultzii* y *Coespeletia moritziana* (Espeletiinae). **Revista de la Facultad de Farmacia, Universidad de Los Andes**, v. 50, n. 1, p. 16–19, 2008.

JANSEN, B.; NIEROP, K. G. J.; TONNEIJCK, F. H.; VAN DER WIELEN, F. W. M.; VERSTRATEN, J. M. Can isoprenoids in leaves and roots of plants serve as biomarkers for past vegetation changes? A case study from the Ecuadorian Andes. **Plant Soil**, v. 291, p. 181–198, 2007.

KIM, N.; KIM, K.; CHOI, B. Y.; et al. Metabolomic approach for age discrimination of *Panax ginseng* using UPLC-QTOF MS. **Journal of Agricultural and Food Chemistry**, v. 59, n. 19, p. 10435–10441, 2011.

KOES, R. E.; QUATTROCCHIO, F.; MOL, J. N. M. The flavonoid biosynthetic pathway in plants: function and evolution. **BioEssays**, v. 16, n. 2, p. 123–132, 1994.

KRASTANOV, A. Metabolomics – the state of art. **Biotechnology & Biotechnological Equipment**, v. 24, n. 1, p. 1537–1543, 2010.

LARSSON, S. The “new” chemosystematics: phylogeny and phytochemistry. **Phytochemistry**, v. 68, p. 2903–2907, 2007.

LEWIS, P. O. A likelihood approach to estimating phylogeny from discrete morphological character data. **Systematic Biology**, v. 50, n. 6, p. 913–925, 2001.

LUTEYN, J. L. **Páramos: a checklist of plant diversity, geographical distribution and botanical literature**. New York, USA: New York Botanical Garden Press, 1999.

MARTUCCI, M. E. P.; DE VOS, R. C. H.; CAROLLO, C. A.; GOBBO-NETO, L. Metabolomics as a potential chemotaxonomical tool: application in the genus vernonia schreb. **PloS one**, v. 9, n. 4, p. e93149, 2014.

MONASTERIO, M.; SARMIENTO, L. Adaptive radiation of *Espeletia* in the cold andean tropics. **Trends in Ecology & Evolution**, v. 6, n. 12, p. 387–391, 1991.

MORALES, A. M.; USUBILLAGA, A.; BANERJEE, A. K.; NAKANO, T. Studies on the constituents of *Espeletia weddellii*. **Planta Medica**, v. 24, n. 07, p. 243–248, 1973.

MORALES, M.; OTERO, J.; HAMMEN, T. VAN DER; et al. **Atlas de páramos de Colombia**. Bogotá, D.C.: Instituto de Investigación de Recursos Biológicos Alexander von Humboldt, 2007.

PADILLA-GONZÁLEZ, G. F.; DIAZGRANADOS, M.; OLIVEIRA, T. B.; CHAGAS-PAULA, D.; COSTA, F. B. DA. Chemistry of the subtribe Espeletiinae and its correlation with phylogenetic and biogeographic data: an in silico chemotaxonomic approach. **Biochemical Systematics and Ecology**, submetido em 2014.

PENA, A.; ROJAS, L.; APARICIO, R.; et al. Chemical composition and antibacterial activity of the essential oil of *Espeletia nana*. **Natural Product Communications**, v. 7, p. 661–662, 2012.

PIOZZI, F.; PASSANNA, S.; PATERNOS, M. P.; SPRIO, V. Kauranoid diterpenes in *Espeletia grandiflora*. **Phytochemistry**, v. 10, p. 1164–1166, 1971.

PIOZZI, F.; PASSANNANTI, S.; MARINO, M. L.; SPRIO, V. Structure of grandiflorenic acid. **Canadian Journal of Chemistry**, v. 50, p. 109–112, 1972.

PROANO, O.; ARTEAGA, M.; USUBILLAGA, A. Phytochemical study of *Espeletia hartwegiana*. **Politecnica**, v. 2, p. 95–106, 1972.

RANGEL-CH, O. **La región de vida paramuna**. Bogotá, Colombia: Instituto de Ciencias Naturales - Instituto Alexander von Humboldt, 2000.

RAUSCHER, J. T. Molecular phylogenetics of the Espeletia complex (Asteraceae): Evidence from nrDNA its sequences on the closest relatives of an andean adaptive radiation. **American Journal of Botany**, v. 89, p. 1074–1084, 2002.

REYNOLDS, T. The evolution of chemosystematics. **Phytochemistry**, v. 68, p. 2887–2895, 2007.

SÁNCHEZ, A. **Filogénia molecular de los Espeletiinae, una radiación adaptativa andina**, 2005. Tese (Maestría en Biología), Departamento de Ciencias Biológicas, Universidad de los Andes, Bogotá, Colombia.

SCHMELZER, E.; JAHNEN, W.; HAHLBROCK, K. *In situ* localization of light-induced chalcone synthase mRNA, chalcone synthase, and flavonoid end products in epidermal cells of parsley leaves. **Proceedings of the National**

**Academy of Sciences of the United States of America**, v. 85, n. 9, p. 2989–93, 1988.

SCHMITT, I.; BARKER, F. K. Phylogenetic methods in natural product research. **Natural product reports**, v. 26, n. 12, p. 1585–602, 2009.

SEEHAUSEN, O. Hybridization and adaptive radiation. **Trends in ecology & evolution**, v. 19, n. 4, p. 198–207, 2004.

SEVERIANO, A.; CARRIÇO, J. A; ROBINSON, D. A.; RAMIREZ, M.; PINTO, F. R. Evaluation of jackknife and bootstrap for defining confidence intervals for pairwise agreement measures. **PloS one**, v. 6, n. 5, p. e19539, 2011.

SMITH, J. M. B.; CLEEF, A. M. Composition and origin of the world's tropical alpine floras. **Journal of Biogeography**, v. 15, p. 631–645, 1988.

SWOFFORD, D. L. PAUP*: Phylogenetic Analysis Using Parsimony (*and Other Methods), 2002. Sunderland, Massachusetts.: Sinauer Associates.

TAKENAKA, M.; YAN, X.; ONO, H.; et al. Caffeic acid derivatives in the roots of yacon (*Smallanthus sonchifolius*). **Journal of agricultural and food chemistry**, v. 51, n. 3, p. 793–6, 2003.

TELLEZ, N.; TORRENEGRA, R.; PEDROZO, J.; GRAY, A. Cycloartan-2 beta-2-methyl butanoate isolated from genus *Espeletia* (Asteraceae). **Molecules**, v. 3, p. M49–U4, 1998.

TERADA, S.; ITOH, K.; NOGUCHI, N.; ISHIDA, T. Alpha-glucosidase inhibitor for blood glucose level elevation and functional food containing tricaffeoylaldaric acid and method for producing tricaffeoylaldaric acid, 2009. (Patent) USA.

TORRENEGRA, R. D.; TELLEZ, A. A. N. Phytochemistry of *Espeletia killipii* Cuatr. and gibberellic activity of some of the isolated compounds. **Revista Latinoamericana de Química**, v. 24, p. 2–6, 1996.

TORRENEGRA, R. D.; TELLEZ, A. N. Chemotaxonomic value of melampolides in *Espeletia* species (Asteraceae). **Biochemical Systematics and Ecology**, v. 23, p. 449–450, 1995.

TORRENEGRA, R. D.; TELLEZ, A. N.; GARCIA, G. Chemical investigation of the species of the genus *Espeletia*. Part I. Chemistry of *Espeletia kilipii-Espeletia tunjana*. **Revista Colombiana de Química**, v. 23, p. 29–35, 1994.

USUBILLAGA, A.; HERNANDEZ, J. DE; PEREZ, N.; KIRIAKIDIS, M. Kauranoid diterpenes in *Espletia* species. **Phytochemistry**, v. 12, p. 2999, 1973.

USUBILLAGA, A.; MORALES, A. Derivados del kaureno en la *Espeletia tenore*. **Revista Latinoamericana de Química**, p. 128–131, 1970.

USUBILLAGA, A.; ROMERO, M.; APARICIO, R. Kaurenic acids in Espeletiinae. **Acta Horticulturae (ISHS)**, v. 597, p. 129–130, 2003.

VAN DEN BERG, R. A.; HOEFSLOOT, H. C. J.; WESTERHUIS, J. A.; SMILDE, A. K.; WERF, M. J. VAN DER. Centering, scaling, and transformations:

improving the biological information content of metabolomics data. **BMC genomics**, v. 7, p. 142, 2006.

VAN TUNEN, A. J. VAN; KOES, R. E.; SPELT, E.; et al. Cloning of the two chalcone flavanone isomerase genes from *Petunia hybrida* : coordinate, light-regulated and differential expression of flavonoid genes. **The EMBO journal**, v. 7, n. 5, p. 1257–1263, 1988.

WAGNER, S.; SCHOLZ, K.; SIEBER, M.; KELLERT, M.; VOELKEL, W. Tools in metabonomics: an integrated validation approach for LC-MS metabolic profiling of mercapturic acids in human urine. **Analytical chemistry**, v. 79, n. 7, p. 2918–26, 2007.

WARD, J. H. Hierarchical grouping to optimize an objective function. **Journal of the American Statistical Association**, v. 58, n. 301, p. 236–244, 1963.

WINK, M. Evolution of secondary metabolites from an ecological and molecular phylogenetic perspective. **Phytochemistry**, v. 64, n. 1, p. 3–19, 2003.

WINK, M. Importance of plant secondary metabolites for protection against insects and microbial infections. In: C. Capinella; M. Ray (Eds.); **Naturally Occurring Bioactive Compounds**. p.251–268, 2006. London, UK.: Elsevier.

WINK, M.; BOTSCHEN, F.; WATERMAN, P. G.; SCH, H. Chemotaxonomy seen from a phylogenetic perspective and evolution of secondary metabolism. In: M. Wink (Ed.); **Annual Plant Reviews, Biochemistry of Plant Secondary Metabolism**. Second ed., p.364–433, 2010. Blackwell Publishing Ltd.

WITTEN, I. H.; EIBE, F.; HALL, M. A. **Data Mining: Practical Machine Learning Tools and Techniques**. Third Edit ed. San Francisco, USA: Morgan Kaufmann, 2005.

XIE, G. X.; CHEN, T. L.; QIU, Y. P.; et al. Urine metabolite profiling offers potential early diagnosis of oral cancer. **Metabolomics**, v. 8, n. 2, p. 220–231, 2011.

ZWICKL, D. J. **Genetic algorithm approaches for the phylogenetic analysis of large biological sequence datasets under the maximum likelihood criterion**, 2006. Tese (Doctor of Philosophy), School of Biological Sciences, The University of Texas at Austin, Austin, TX, USA.

**APÊNDICE I.** Nomes, códigos, locais e datas de coleta das espécies estudadas neste projeto.

| Cod. | Nome da espécie | País de coleta | Estado de coleta | Região de coleta | Data de coleta |
|---|---|---|---|---|---|
| 1 | *E. annemariana* Cuatrec. | Colômbia | Boyacá | P. de Siscuinsí. | 19/08/2009 |
| 2 | *E. annemariana* Cuatrec. | Colômbia | Boyacá | P. de Siscuinsí. | 19/08/2009 |
| 3 | *E. arbelaezii* Cuatrec. | Colômbia | Boyacá | P. de Huina. | 04/09/2009 |
| 4 | *E. arbelaezii* Cuatrec. | Colômbia | Boyacá | P. de Huina. | 05/09/2009 |
| 5 | *E. argentea* × *E. barclayana* Humb. & Bonpl. Cuatrec. | Colômbia | Cundinamarca | P. de Guerrero | 16/12/2007 |
| 6 | *E. argentea fma. phaneractis* (S. F. Blake) Cuatrec. | Colômbia | Cundinamarca | P. de Guerrero. | 10/12/2007 |
| 7 | *E. argentea fma. phaneractis* (S. F. Blake) Cuatrec. | Colômbia | Cundinamarca | P. de Guerrero | 16/12/2007 |
| 8 | *E. argentea* Humb. & Bonpl. | Colômbia | Cundinamarca | P. de Guerrero | 16/12/2007 |
| 9 | *E. ariana* S. Díaz & Rodr.-Cabeza | Colômbia | Boyacá | P. de Pisba. | 09/09/2009 |
| 10 | *E. ariana* S. Díaz & Rodr.-Cabeza | Colômbia | Boyacá | P. de Pisba. | 09/09/2009 |
| 11 | *E. aristeguietana* Cuatrec. | Venezuela | Trujillo | P. de La Cristalina. | 23/07/2010 |
| 12 | *E. aristeguietana* Cuatrec. | Venezuela | Trujillo | P. de La Cristalina. | 23/07/2010 |
| 13 | *E. azucarina* Cuatrec. | Colômbia | Boyacá | P. de Huina. | 06/09/2009 |
| 14 | *E. azucarina* Cuatrec. | Colômbia | Boyacá | P. de Huina. | 06/09/2009 |
| 15 | *E. barclayana* Cuatrec. | Colômbia | Cundinamarca | P. de Guerrero. | 10/12/2007 |
| 16 | *E. barclayana* Cuatrec. | Colômbia | Cundinamarca | P. de Guerrero. | 10/12/2007 |
| 17 | *E. barclayana* Cuatrec. | Colômbia | Cundinamarca | P. de Guerrero | 16/12/2007 |
| 18 | *E. batata* × *E. schultzii* Cuatrec.Wedd. | Venezuela | Mérida | Sierra de la Culata. | 27/10/2009 |
| 19 | *E. batata* Cuatrec. | Venezuela | Mérida | Sierra de la Culata. | 21/07/2010 |
| 20 | *E. boyacensis* × *E. lopezii* Cuatrec.Cuatrec. | Colômbia | Boyacá | P. de Pisba. | 10/09/2009 |
| 21 | *E. boyacensis* Cuatrec. | Colômbia | Boyacá | SFF de Iguaque. | 20/12/2007 |
| 22 | *E. boyacensis* Cuatrec. | Colômbia | Boyacá | P. de La Rusia. | 28/12/2007 |
| 23 | *E. boyacensis* Cuatrec. | Colômbia | Boyacá | P. de La Rusia. | 28/12/2007 |

| | | | | | |
|---|---|---|---|---|---|
| 24 | *E. brachyaxiantha* S. Díaz | Colômbia | Boyacá | P. de Portachuelos. | 30/12/2007 |
| 25 | *E. brachyaxiantha* S. Díaz | Colômbia | Boyacá | P. de Portachuelos. | 30/12/2007 |
| 26 | *E. brassicoidea* Cuatrec. | Colômbia | Norte de Santander | P. de Fontibón | 02/10/2009 |
| 27 | *E. brassicoidea* Cuatrec. | Colômbia | Norte de Santander | P. de Fontibón | 02/10/2009 |
| 28 | *E. cayetana* Cuatrec. | Colômbia | Boyacá | SFF de Iguaque. | 20/12/2007 |
| 29 | *E. cleefii* Cuatrec. | Colômbia | Boyacá | PNN Cocuy. | 07/01/2008 |
| 30 | *E. cleefii* Cuatrec. | Colômbia | Arauca | PNN Cocuy. | 10/01/2008 |
| 31 | *E. congestiflora* Cuatrec. | Colômbia | Boyacá | P. de La Rusia. | 28/12/2007 |
| 32 | *E. conglomerata* A. C. Sm. | Colômbia | Norte de Santander | P. de Almorzadero. | 04/10/2009 |
| 33 | *E. crassa* Diazgranados | Colômbia | Boyacá | P. de La Rusia. | 28/12/2007 |
| 34 | *E. crassa* Diazgranados | Colômbia | Boyacá | P. de La Rusia. | 28/12/2007 |
| 35 | *E. cuniculorum* Cuatrec. | Venezuela | Mérida | Sierra de la Culata. | 25/10/2009 |
| 36 | *E. curialensis* Cuatrec. | Colômbia | Boyacá | P. de Pisba. | 08/09/2009 |
| 37 | *E. delicatissima* Diazgranados | Colômbia | Cundinamarca | PNN Sumapaz | 18/12/2007 |
| 38 | *E. delicatissima* Diazgranados | Colômbia | Cundinamarca | PNN Sumapaz | 14/08/2009 |
| 39 | *E. discoidea* Cuatrec. | Colômbia | Boyacá | P. de Portachuelos. | 30/12/2007 |
| 40 | *E. discoidea* Cuatrec. | Colômbia | Boyacá | P. de Huina. | 06/09/2009 |
| 41 | *E. dugandii* Cuatrec. | Colômbia | Norte de Santander | P. de Cerrito. | 04/10/2009 |
| 42 | *E. dugandii* Cuatrec. | Colômbia | Norte de Santander | P. de Cerrito. | 04/10/2009 |
| 43 | *E. episcopalis* S. Díaz & Rodr.-Cabeza | Colômbia | Boyacá | P. de Pisba. | 08/09/2009 |
| 44 | *E. estanislana* Cuatrec. | Colômbia | Norte de Santander | P. de Almorzadero. | 04/10/2009 |
| 45 | *E. estanislana* Cuatrec. | Colômbia | Norte de Santander | P. de Antalá. | 12/10/2009 |
| 46 | *E. formosa var. gigantea* Diazgranados | Colômbia | Boyacá | P. de Pisba. | 08/09/2009 |
| 47 | *E. formosa* S. Díaz & Rodr.-Cabeza | Colômbia | Boyacá | P. de Pisba. | 08/09/2009 |
| 48 | *E. frontinoensis* Cuatrec. | Colômbia | Antioquia | P. del Sol. | 25/09/2009 |
| 49 | *E. frontinoensis* Cuatrec. | Colômbia | Antioquia | P. del Sol. | 25/09/2009 |

| | | | | | |
|---|---|---|---|---|---|
| 50 | *E. grandiflora* Humb. & Bonpl. | Colômbia | Cundinamarca | P. de Cruz Verde. | 15/12/2007 |
| 51 | *E. grandiflora* Humb. & Bonpl. | Colômbia | Cundinamarca | P. de Cruz Verde. | 15/12/2007 |
| 52 | *E. grandiflora* Humb. & Bonpl. | Colômbia | Cundinamarca | P. de Cruz Verde. | 15/12/2007 |
| 53 | *E. grandiflora* Humb. & Bonpl. | Colômbia | Cundinamarca | PNN Chingaza. | 22/12/2007 |
| 54 | *E. idroboi* Cuatrec. | Colômbia | Cauca | P. de San Rafael. | 10/08/2011 |
| 55 | *E. idroboi* Cuatrec. | Colômbia | Cauca | P. de San Rafael. | 10/08/2011 |
| 56 | *E. incana* Cuatrec. | Colômbia | Boyacá | P. de La Rusia. | 28/12/2007 |
| 57 | *E. incana* Cuatrec. | Colômbia | Boyacá | P. de Ocetá. | 29/12/2007 |
| 58 | *E. jajoensis* Aristeg. | Venezuela | Mérida | El Paramito. | 22/07/2010 |
| 59 | *E. jajoensis* Aristeg. | Venezuela | Mérida | El Paramito. | 22/07/2010 |
| 60 | *E. jaramilloi* × *E. boyacensis* S. Díaz Cuatrec. | Colômbia | Boyacá | P. de Pisba. | 07/09/2009 |
| 61 | *E. jaramilloi* S. Díaz | Colômbia | Boyacá | P. de Pisba. | 07/09/2009 |
| 62 | *E. jaramilloi* S. Díaz | Colômbia | Boyacá | P. de Pisba. | 08/09/2009 |
| 63 | *E. killipii Cuatrec. var. chisacana* Cuatrec. | Colômbia | Cundinamarca | PNN Sumapaz | 18/12/2007 |
| 64 | *E. killipii Cuatrec. var. chisacana* Cuatrec. | Colômbia | Cundinamarca | PNN Sumapaz | 18/12/2007 |
| 65 | *E. litocrassa* Diazgranados | Colômbia | Cundinamarca | PNN Chingaza. | 22/12/2007 |
| 66 | *E. litocrassa* Diazgranados | Colômbia | Cundinamarca | PNN Chingaza. | 22/12/2007 |
| 67 | *E. lopezii* Cuatrec. | Colômbia | Boyacá | PNN Cocuy. | 11/01/2008 |
| 68 | *E. lopezii* Cuatrec. *var. major* Cuatrec. | Colômbia | Boyacá | P. de Ocetá. | 29/12/2007 |
| 69 | *E. monguana* Diazgranados | Colômbia | Boyacá | P. de Mongua. | 22/08/2009 |
| 70 | *E. monguana* Diazgranados | Colômbia | Boyacá | P. de Mongua. | 22/08/2009 |
| 71 | *E. multicongestiflora* Diazgranados | Colômbia | Boyacá | P. de Toquilla - La Sarna. | 17/08/2009 |
| 72 | *E. murilloi* × *E. boyacensis* Cuatrec.Cuatrec. | Colômbia | Boyacá | P. de Santa Rosita. | 25/08/2009 |
| 73 | *E. murilloi* Cuatrec. | Colômbia | Boyacá | P. entre Toca y Pesca. | 15/08/2009 |
| 74 | *E. murilloi var. alba* Diazgranados | Colômbia | Cundinamarca | P. de Guacheneque. | 21/12/2007 |
| 75 | *E. murilloi var. alba* Diazgranados | Colômbia | Boyacá | P. de Santa Rosita. | 25/08/2009 |

| | | | | | |
|---|---|---|---|---|---|
| 76 | *E. murilloi var. iguaquei* Diazgranados | Colômbia | Boyacá | SFF de Iguaque. | 20/12/2007 |
| 77 | *E. nana* Cuatrec. | Venezuela | Barinas | P. al Sur/Este de Niquitao. | 28/12/2009 |
| 78 | *E. nana* Cuatrec. | Venezuela | Trujillo | P. de la Teta de Niquitao | 29/12/2009 |
| 79 | *E. nemekenei* Cuatrec. | Colômbia | Boyacá | P. de La Rusia. | 28/12/2007 |
| 80 | *E. nemekenei* Cuatrec. | Colômbia | Boyacá | P. de La Rusia. | 28/12/2007 |
| 81 | *E. occidentalis* A. C. Sm. | Colômbia | Antioquia | P. de Belmira. | 23/09/2009 |
| 82 | *E. occidentalis* A. C. Sm. | Colômbia | Antioquia | P. de Belmira. | 23/09/2009 |
| 83 | *E. oirensis* Diazgranados | Colômbia | Norte de Santander | P. de Tamá. | 15/10/2009 |
| 84 | *E. oswaldiana* S. Díaz | Colômbia | Boyacá | P. de la Sarna | 19/08/2009 |
| 85 | *E. paipana* × *E. murilloi* S. Díaz & Pedraza Cuatrec. | Colômbia | Boyacá | P. de Palermo. | 04/09/2009 |
| 86 | *E. paipana* S. Díaz & Pedraza | Colômbia | Boyacá | P. de Palermo. | 04/09/2009 |
| 87 | *E. paipana* S. Díaz & Pedraza | Colômbia | Boyacá | P. de Palermo. | 04/09/2009 |
| 88 | *E. pescana* (Diaz S.) S. Diaz | Colômbia | Boyacá | P. de Ocetá. | 29/12/2007 |
| 89 | *E. pescana* (Diaz S.) S. Diaz | Colômbia | Boyacá | P. de Ongotá. | 20/08/2009 |
| 90 | *E. pisbana* S. Díaz & Rodr.-Cabeza | Colômbia | Boyacá | P. de Pisba. | 08/09/2009 |
| 91 | *E. praefrontina* Cuatrec. | Colômbia | Antioquia | P. del Sol. | 25/09/2009 |
| 92 | *E. praefrontina* Cuatrec. | Colômbia | Antioquia | P. del Sol. | 25/09/2009 |
| 93 | *E. roberti* Cuatrec. | Colômbia | Norte de Santander | P. de Jurisdicciones | 10/10/2009 |
| 94 | *E. roberti* Cuatrec. | Colômbia | Norte de Santander | P. de Jurisdicciones | 10/10/2009 |
| 95 | *E. rositae* Cuatrec. | Colômbia | Boyacá | P. de Huina. | 06/09/2009 |
| 96 | *E. saboyensis* Diazgranados | Colômbia | Boyacá | P. de las Antenas de Telecóm | 02/09/2009 |
| 97 | *E. saboyensis* Diazgranados | Colômbia | Boyacá | P. de las Antenas de Telecóm | 02/09/2009 |
| 98 | *E. schultesiana* | Colômbia | Nariño o Putumayo | N.I. | N.I. |
| 99 | *E. schultzii* × *Co. moritziana x E. weddellii* Cuatrec. | Venezuela | Mérida | Sierra de la Culata. | 27/10/2009 |
| 100 | *E. schultzii* × *E. jajoensis Wedd.* Aristeg. | Venezuela | Mérida | P. de Santo Domingo | 29/10/2009 |
| 101 | *E. schultzii* Wedd. | Venezuela | Mérida | Sierra de la Culata. | 27/10/2009 |

| | | | | | |
|---|---|---|---|---|---|
| 102 | *E. schultzii* Wedd. | Venezuela | Mérida | Sierra Nevada de Mérida. | 28/10/2009 |
| 103 | *E. semiglobulata* Cuatrec. | Venezuela | Mérida | Sierra de la Culata. | 21/07/2010 |
| 104 | *E. standleyana* A. C. Sm. | Colômbia | Santander | P. de Berlín | 01/10/2010 |
| 105 | *E. standleyana* A. C. Sm. | Colômbia | Norte de Santander | P. de Almorzadero. | 04/10/2009 |
| 106 | *E. steyermarkii* Cuatrec. | Colômbia | Norte de Santander | SubP. y P. azonal del Carbón | 03/10/2009 |
| 107 | *E. steyermarkii* Cuatrec. | Colômbia | Norte de Santander | SubP. y P. azonal del Carbón | 03/10/2009 |
| 108 | *E. summapacis* Cuatrec. | Colômbia | Cundinamarca | PNN Sumapaz | 14/08/2009 |
| 109 | *E. summapacis* Cuatrec. | Colômbia | Cundinamarca | PNN Sumapaz | 14/08/2009 |
| 110 | *E. susensis* Diazgranados | Colômbia | Boyacá | P. de Suse | 16/08/2009 |
| 111 | *E. tenorae* × *E. schultzii* Aristeg. Wedd. | Venezuela | Trujillo | P. de la Teta de Niquitao. | 29/12/2009 |
| 112 | *E. tenorae cf.* Aristeg. | Venezuela | Trujillo | P. de la Teta de Niquitao. | 29/12/2009 |
| 113 | *E. tetrarbelaezii* Diazgranados | Colômbia | Boyacá | Güina-Guantiva | 12/01/2008 |
| 114 | *E. toquensis* Diazgranados | Colômbia | Boyacá | P. entre Toca y Pesca. | 15/08/2009 |
| 115 | *E. totensis* Diazgranados | Colômbia | Boyacá | P. de Toquilla - La Sarna. | 17/08/2009 |
| 116 | *E. tunjana* Cuatrec. | Colômbia | Boyacá | P. entre Toca y Pesca. | 15/08/2009 |
| 117 | *E. ulotricha* Cuatrec. | Venezuela | Trujillo | P. de Jabón | 23/12/2009 |
| 118 | *E. uribei* Cuatrec. | Colômbia | Cundinamarca | PNN Chingaza. | 22/12/2007 |
| 119 | *E. uribei* Cuatrec. | Colômbia | Cundinamarca | PNN Chingaza. | 22/12/2007 |
| 120 | *E. weddellii* Sch. Bip. ex Wedd. | Venezuela | Mérida | Sierra de la Culata. | 27/10/2009 |

N.I., não informado

**APÊNDICE II.** Parâmetros utilizados no software MZmine.

**Importação dos dados:** Uma vez gerados os dados das análises por UHPLC-UV-MS em formato .raw, foi feita a separação dos modos de detecção positivo e negativo no software RecalOffline – Xcalibur. Posteriormente, os dados foram transformados no formato .mzXML usando o software MassMatrix File Conversion (Case Western Reserve University) ou inseridos diretamente em formato .raw no MZmine.

**1. Raw data methods - Peak detection**

*1.1 Mass detection:*

Mass detector: Exact

Noise level: 1,0E6

*1.2 FTMS shoulder peaks filter:*

Mass resolution: 70.000

Peak model function: Lorentzian extended

*1.3 Chromatogram builder:*

Min time span (min): 0.3

Min height: 1.1E6

*m/z* tolerance: 0.002 (absolute)

*m/z* tolerance: 5 ppm (relative)

**2. Peak list methods - Peak detection**

*2.1 Chromatogram deconvolution:*

Algorithm: Local minimum search

Chromatographic theshold: 5%

Search minimum in RT range (min): 0.2

Minimum relative height: 15%

Minimum absolute height: 1.1E6

Min ration of peak top/edge: 5

Peak duration range (min): 0,3 – 10

**3. Peak list methods - Isotopes**

*3.1 Isotopic peaks grouper:*

*m/z* tolerance: 0.002 (absolute)

*m/z* tolerance: 5 ppm (relative)

Retention time tolerance: 0.2

Maximum charge: 3

Representative isotope: Most intense

**4. Peak list methods - Alignment**

*4.1 Joint aligner:*

*m/z* tolerance: 0.002 (absolute)

*m/z* tolerance: 5 ppm (relative)

Weight for m/z: 15

Retention time tolerance: 0.2 - absolute (min)

Weight for RT: 10

Isotope *m/z* tolerance: 0.002

Minimum absolute intensity: 1.1E6

Minimum score: 65%

**5. Peak list methods – Gap filling**

*5.1 Peak finder:*

intensity tolerance: 1%

*m/z* tolerance: 0.002

Retention time tolerance: 0.2 - absolute (min)

**6. Peak list methods – Identification**

*6.1 Adduct search (positive mode):*

[M+Na-H] 21.9825 *m/z*

[M+K-H] 37.9559 *m/z*

[M+Mg-2H] 21.9694 *m/z*

[M+NH3] 17.0265 *m/z*

[M+ACN+H] 41.0270 *m/z*

*m/z* tolerance: 0.002 (absolute)

*m/z* tolerance: 5 ppm (relative)

Max relative adduct peak height: 50%

**Obs:** No modo de detecção negativo são considerados os adutos [M+Cl+H] 35.9774 *m/z* e [M+ACN+H] 41.0270 *m/z*

*6.2 Complex search (positive mode):*

Ionization method: [M+H]+

Retention time tolerance:  0.2 - absolute (min)

*m/z* tolerance: 0.002 (absolute)

*m/z* tolerance: 5 ppm (relative)

Max complex peak height: 50%

**Obs:** No modo de detecção negativo é considerado o método de ionização [M-H]-

*6.3 Custom database search:*

Database file: (Banco de dados em formato .csv)

Field order: (de acordo à ordem do banco de dados)

*m/z* tolerance: 0.002

**APÊNDICE III.** Parâmetros utilizados para seleção de atributos e construção da árvore de decisão J48 no software Weka.

**1. Seleção de atributos por algoritmos genéticos.**

Attribute evaluator: CfsSubsetEval

Search Method: GeneticSearch

CrossoverProb: 0.6

MaxGenerations: 100

MutationProb: 0.033

PopulationSize: 120

ReportFrequency: 20

Seed: 1

**2. Construção da árvore de decisão J48.**

Classifier: J48

BinarySplits: False

ConfidenceFactor: 0.25

Debug: False

MinNumObj: 2

NumFolds: 3

reducedErrorPruning: False

SaveInstanceData: False

Seed: 1

subtreeRising: True

Unpruned: True

UseLaplace: False

**2.1 Test options.**

Percentage split: 66%

**APÊNDICE IV.** Grupos de treino e teste para predição e validação do modelo por OPLS-DA.

Espécies empregadas no grupo de treino:

| Cod. | Nome | Cordilheira de origem |
|---|---|---|
| 1 | *E. annemariana* | Cordilheira Oriental |
| 4 | *E. arbelaezii* | Cordilheira Oriental |
| 8 | *E. argentea* | Cordilheira Oriental |
| 10 | *E. ariana* | Cordilheira Oriental |
| 11 | *E. aristeguietana* | Venezuela |
| 14 | *E. azucarina* | Cordilheira Oriental |
| 16 | *E. barclayana* | Cordilheira Oriental |
| 18 | *E. batata* × *E. schultzii* | Venezuela |
| 21 | *E. boyacensis* | Cordilheira Oriental |
| 25 | *E. brachyaxiantha* | Cordilheira Oriental |
| 26 | *E. brassicoidea* | Cordilheira Oriental |
| 28 | *E. cayetana* | Cordilheira Oriental |
| 30 | *E. cleefii* | Cordilheira Oriental |
| 32 | *E. conglomerata* | Cordilheira Oriental |
| 33 | *E. crassa* | Cordilheira Oriental |
| 35 | *E. cuniculorum* | Venezuela |
| 36 | *E. curialensis* | Cordilheira Oriental |
| 38 | *E. delicatissima* | Cordilheira Oriental |
| 40 | *E. discoidea* | Cordilheira Oriental |
| 43 | *E. episcopalis* | Cordilheira Oriental |
| 45 | *E. estanislana* | Cordilheira Oriental |
| 46 | *E. formosa var. gigantea* | Cordilheira Oriental |
| 48 | *E. frontinoensis* | Cordilheira Central e Ocidental |
| 49 | *E. frontinoensis* | Cordilheira Central e Ocidental |
| 50 | *E. grandiflora* | Cordilheira Oriental |
| 51 | *E. grandiflora* | Cordilheira Oriental |
| 55 | *E. idroboi* | Cordilheira Central e Ocidental |
| 57 | *E. incana* | Cordilheira Oriental |
| 58 | *E. jajoensis* | Venezuela |
| 59 | *E. jajoensis* | Venezuela |
| 60 | *E. jaramilloi* × *E. boyacensis* | Cordilheira Oriental |
| 61 | *E. jaramilloi* | Cordilheira Oriental |
| 62 | *E. jaramilloi* | Cordilheira Oriental |
| 63 | *E. killipii var. chisacana* | Cordilheira Oriental |
| 64 | *E. killipii var. chisacana* | Cordilheira Oriental |
| 65 | *E. litocrassa* | Cordilheira Oriental |
| 66 | *E. litocrassa* | Cordilheira Oriental |
| 67 | *E. lopezii* | Cordilheira Oriental |
| 68 | *E. lopezii var. major* | Cordilheira Oriental |

| | | |
|---|---|---|
| 69 | *E. monguana* | Cordilheira Oriental |
| 70 | *E. monguana* | Cordilheira Oriental |
| 71 | *E. multicongestiflora* | Cordilheira Oriental |
| 72 | *E. murilloi* × *E. boyacensis* | Cordilheira Oriental |
| 73 | *E. murilloi* | Cordilheira Oriental |
| 74 | *E. murilloi var. alba* | Cordilheira Oriental |
| 75 | *E. murilloi var. alba* | Cordilheira Oriental |
| 76 | *E. murilloi var. alba* | Cordilheira Oriental |
| 79 | *E. nemekenei* | Cordilheira Oriental |
| 80 | *E. nemekenei* | Cordilheira Oriental |
| 81 | *E. occidentalis* | Cordilheira Central e Ocidental |
| 82 | *E. occidentalis* | Cordilheira Central e Ocidental |
| 83 | *E. oirensis* | Cordilheira Oriental |
| 84 | *E. oswaldiana* | Cordilheira Oriental |
| 85 | *E. paipana* × *E. murilloi* | Cordilheira Oriental |
| 86 | *E. paipana* | Cordilheira Oriental |
| 88 | *E. pescana* | Cordilheira Oriental |
| 89 | *E. pescana* | Cordilheira Oriental |
| 90 | *E. pisbana* | Cordilheira Oriental |
| 93 | *E. roberti* | Cordilheira Oriental |
| 94 | *E. roberti* | Cordilheira Oriental |
| 95 | *E. rositae* | Cordilheira Oriental |
| 96 | *E. saboyensis* | Cordilheira Oriental |
| 97 | *E. saboyensis* | Cordilheira Oriental |
| 98 | *E. schultesiana* | Cordilheira Central e Ocidental |
| 99 | *E. schultzii* × *Co. moritziana x E. ellii* | Venezuela |
| 101 | *E. schultzii* | Venezuela |
| 102 | *E. schultzii* | Venezuela |
| 103 | *E. semiglobulata* | Venezuela |
| 104 | *E. standleyana* | Cordilheira Oriental |
| 105 | *E. standleyana* | Cordilheira Oriental |
| 107 | *E. steyermarkii* | Cordilheira Oriental |
| 108 | *E. summapacis* | Cordilheira Oriental |
| 109 | *E. summapacis* | Cordilheira Oriental |
| 110 | *E. susensis* | Cordilheira Oriental |
| 111 | *E. tenorae* × *E. schultzii* | Venezuela |
| 112 | *E. tenorae cf.* | Venezuela |
| 114 | *E. toquensis* | Cordilheira Oriental |
| 115 | *E. totensis* | Cordilheira Oriental |
| 116 | *E. tunjana* | Cordilheira Oriental |
| 118 | *E. uribei* | Cordilheira Oriental |
| 119 | *E. uribei* | Cordilheira Oriental |
| 120 | *E. ellii* | Venezuela |

Espécies empregadas no grupo de teste:

| Cod. | Nome | Cordilheira de origem |
|---|---|---|
| 2 | *E. annemariana* | Cordilheira Oriental |
| 3 | *E. arbelaezii* | Cordilheira Oriental |
| 5 | *E. argentea x E. barclayana* | Cordilheira Oriental |
| 6 | *E. argentea fma. phaneractis* | Cordilheira Oriental |
| 7 | *E. argentea fma. phaneractis* | Cordilheira Oriental |
| 9 | *E. ariana* | Cordilheira Oriental |
| 12 | *E. aristeguietana* | Venezuela |
| 13 | *E. azucarina* | Cordilheira Oriental |
| 15 | *E. barclayana* | Cordilheira Oriental |
| 17 | *E. barclayana* | Cordilheira Oriental |
| 19 | *E. batata* | Venezuela |
| 20 | *E. boyacensis* × *E. lopezii* | Cordilheira Oriental |
| 22 | *E. boyacensis* | Cordilheira Oriental |
| 23 | *E. boyacensis* | Cordilheira Oriental |
| 24 | *E. brachyaxiantha* | Cordilheira Oriental |
| 27 | *E. brassicoidea* | Cordilheira Oriental |
| 29 | *E. cleefii* | Cordilheira Oriental |
| 31 | *E. congestiflora* | Cordilheira Oriental |
| 34 | *E. crassa* | Cordilheira Oriental |
| 37 | *E. delicatissima* | Cordilheira Oriental |
| 39 | *E. discoidea* | Cordilheira Oriental |
| 41 | *E. dugandii* | Cordilheira Oriental |
| 42 | *E. dugandii* | Cordilheira Oriental |
| 44 | *E. estanislana* | Cordilheira Oriental |
| 47 | *E. formosa* | Cordilheira Oriental |
| 52 | *E. grandiflora* | Cordilheira Oriental |
| 53 | *E. grandiflora* | Cordilheira Oriental |
| 54 | *E. idroboi* | Cordilheira Central e Ocidental |
| 56 | *E. incana* | Cordilheira Oriental |
| 77 | *E. nana* | Venezuela |
| 78 | *E. nana* | Venezuela |
| 87 | *E. paipana* | Cordilheira Oriental |
| 91 | *E. praefrontina* | Cordilheira Central e Ocidental |
| 92 | *E. praefrontina* | Cordilheira Central e Ocidental |
| 100 | *E. schultzii* × *E. jajoensis* | Venezuela |
| 106 | *E. steyermarkii* | Cordilheira Oriental |
| 113 | *E. tetrarbelaezii* | Cordilheira Oriental |
| 117 | *E. ulotricha* | Venezuela |

**APÊNDICE V.** Substâncias puras injetadas no UHPLC-UV-MS.

| | **Substância** | **Massa exata** | **UV** |
|---|---|---|---|
| 1 | Ácido caféico | 180.04226 | 299, 320 |
| 2 | Ácido 5-*O*-(*E*)-cafeoilquínico | 354.09508 | 299, 325 |
| 3 | Ácido p-coumárico | 164.04734 | 274, 310 |
| 4 | Ácido quínico | 192.06339 | - |
| 5 | Ácido ferúlico | 194.05791 | 298, 321 |
| 6 | Rutina | 610.15338 | 260, 299om, 350 |
| 7 | Quercetina | 302.04265 | 255, 299om, 367 |
| 8 | Ácido protocatecuico | 154.02661 | - |
| 9 | Ácido 3,5-di-*O*-(*E*)-cafeoilquínico | 516.12678 | 299, 325 |
| 10 | Ácido 4,5-di-*O*-(*E*)-cafeoilquínico | 516.12678 | 299, 325 |
| 11 | Ácido 3-*O*-(*E*)-cafeoilquínico | 354.09508 | 299, 325 |
| 12 | Ácido 3,4-di-*O*-(*E*)-cafeoilquínico | 516.12678 | 299, 325 |
| 13 | Tagitinina C | 348.15729 | 229, 255 |
| 14 | Budleína A | 374.13655 | 231, 265 |
| 16 | Quercetrina | 448.10056 | 262, 299om, 351 |
| 17 | Afizelina | 432.10565 | - |
| 18 | Ácido trans-ferúlico | 194.05791 | - |
| 20 | Quercetina 3-galactosídeo | 464.09548 | 290, 348 |
| 21 | 3-Metóxiquercetina | 316.22058 | 260, 292om, 353 |

om: ombro

**ANEXO I.** Árvores filogenéticas da subtribo Espeletiinae.

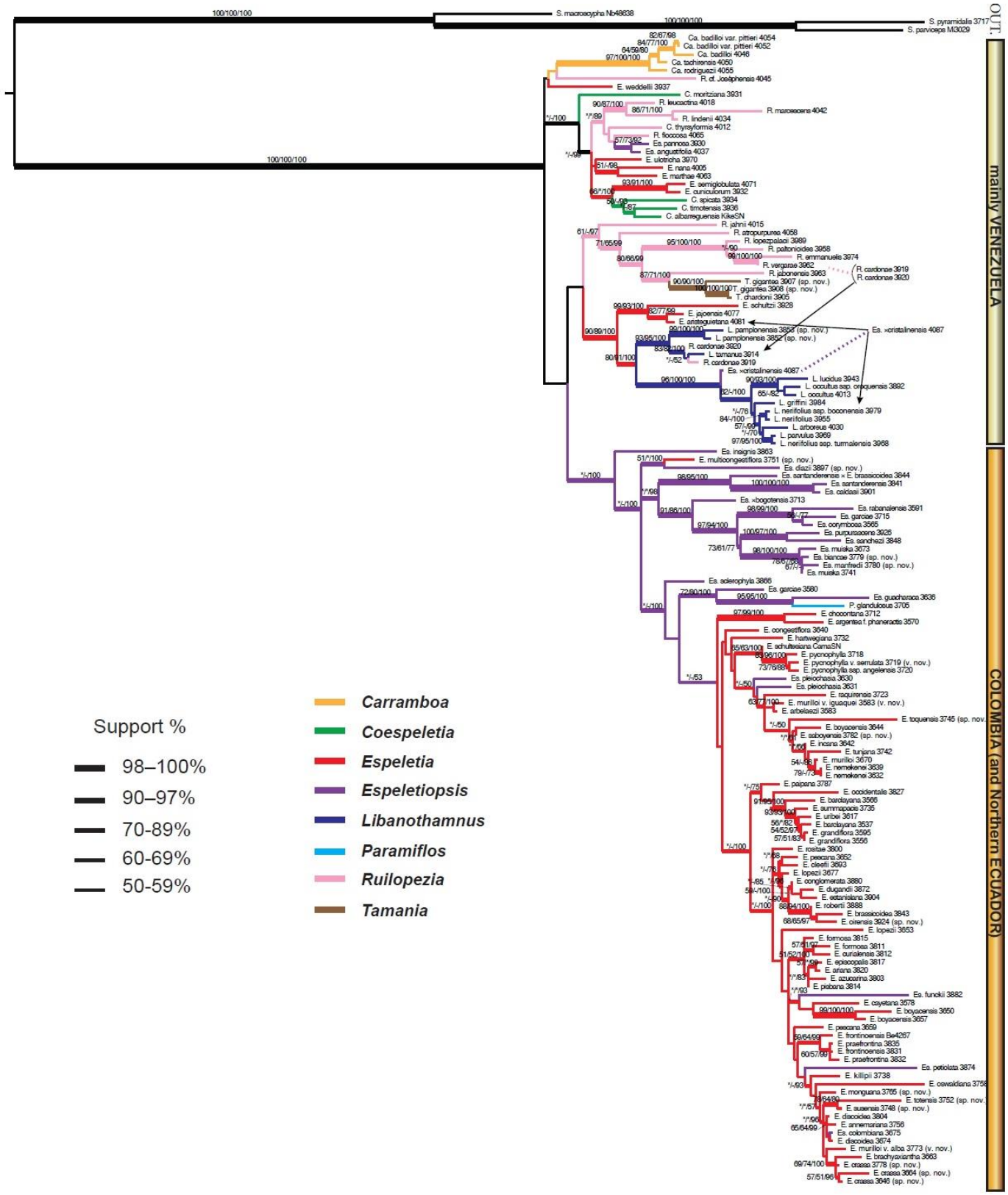



**Figura 1.** Arvore filogenética da subtribo Espeletiinae obtida com base nos marcadores moleculares ETS, ITS e rp116. Abreviações dos gêneros: S., *Smallanthus*; Ca., *Carramboa*; C., *Coespeletia*; E., *Espeletia*; Es., *Espeletiopsis*; L., *Libanothamnus*; P., *Paramiflos*; R., *Ruilopezia*; T., *Tamania.* Gêneros indicados em diferentes cores (Fonte: Diazgranados, 2012).

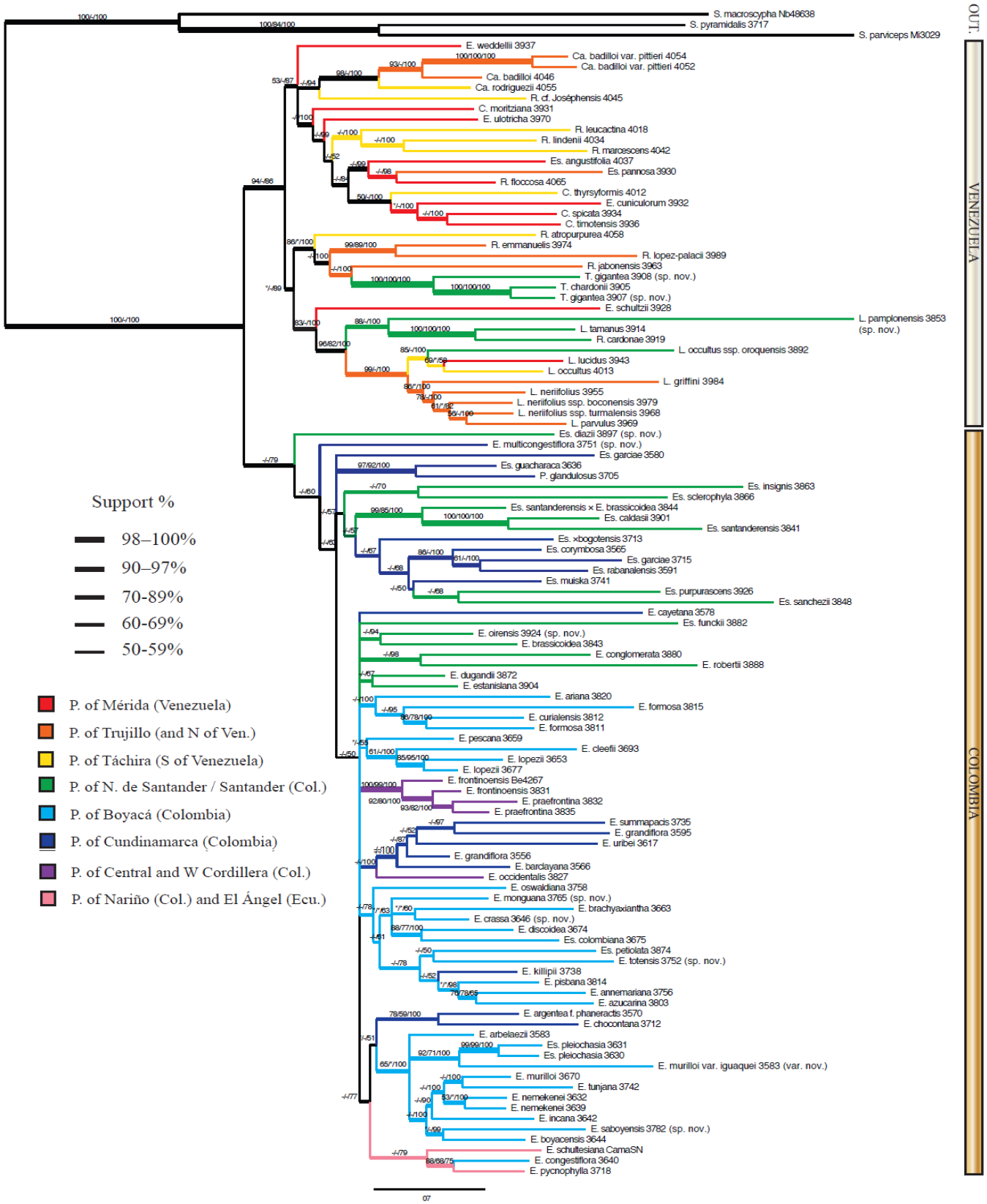



**Figura 2.** Arvore filogenética da subtribo Espeletiinae obtida com base nos marcadores moleculares ETS, ITS, rp116 e AFLPs. Abreviações dos gêneros: S., *Smallanthus*; Ca., *Carramboa*; C., *Coespeletia*; E., *Espeletia*; Es., *Espeletiopsis*; L., *Libanothamnus*; P., *Paramiflos*; R., *Ruilopezia*; T., *Tamania.* Principais *páramos* indicados em diferentes cores (Fonte: Diazgranados, 2012).