UNIVERSITY OF SÃO PAULO

SCHOOL OF PHARMACEUTICAL SCIENCES OF RIBEIRÃO PRETO

Metabolomics studies of the subfamily Barnadesioideae (Asteraceae)

Estudos metabolômicos da subfamília Barnadesioideae (Asteraceae)

Gari Vidal Ccana Ccapatinta

Corrected version of the Doctoral Thesis presented to the Post-Graduate Program in Pharmaceutical Sciences on 12/06/2018. The original version is available at the School of Pharmaceutical Sciences of Ribeirao Preto/USP

Ribeirão Preto

2018

UNIVERSITY OF SÃO PAULO

SCHOOL OF PHARMACEUTICAL SCIENCES OF RIBEIRÃO PRETO

Gari Vidal Ccana Ccapatinta

Metabolomics studies of the subfamily Barnadesioideae (Asteraceae)

Estudos metabolômicos da subfamília Barnadesioideae (Asteraceae)

Doctoral Thesis presented to the Graduate Program in Pharmaceutical Sciences of the School of Pharmaceutical Sciences of Ribeirão Preto/USP for the degree of Doctor in Sciences

Area: Natural and Synthetic Products

Supervisor: Dr. Fernando Batista Da Costa

Ribeirão Preto

2018

I AUTHORIZE THE REPRODUCTION AND TOTAL OR PARTIAL DISCLOSURE OF THIS WORK, BY ANY CONVENTIONAL OR ELECTRONIC MEANS

Gari Vidal Ccana Ccapatinta

Metabolomics studies of the subfamily Barnadesioideae (Asteraceae). Ribeirão Preto, 2018.

81 p.; 30 cm.

Doctoral thesis presented to the Graduate Program of School of Pharmaceutical Sciences of Ribeirão Preto/USP for the degree of Doctor in Sciences.
Concentration Area: Natural and Synthetic Products

Supervisor: Fernando Batista Da Costa

1. Asteraceae. 2. Barnadesioideae. 3. Liquid chromatography. 4. Mass spectrometry. 5. Metabolomics. 6. Multivariate statistical analysis.

CERTIFICATE OF APPROVAL

Gari Vidal Ccana Ccapatinta

Metabolomics studies of the subfamily Barnadesioideae (Asteraceae)

Estudos metabolômicos da subfamília Barnadesioideae (Asteraceae)

Doctoral Thesis presented to the Graduate Program in Pharmaceutical Sciences of the School of Pharmaceutical Sciences of Ribeirão Preto/USP for the degree of Doctor in Sciences

Area: Natural and Synthetic Products

Approved on:

Examining Commission:

Prof. Dr. _____

Institution: _____ Signature: _____

Prof. Dr. _____

Institution: _____ Signature: _____

Prof. Dr. _____

Institution: _____ Signature: _____

Prof. Dr. _____

Institution: _____ Signature: _____

Prof. Dr. _____

Institution: _____ Signature: _____

To the always warm Mother Earth ‘Pachamama’ *of Ribeirão Preto*

ACKNOWLEDGMENTS

I would like to express mi sincere thanks to:

The São Paulo Research foundation (Fundação de Amparo à Pesquisa do Estado de São Paulo - FAPESP), grant 2014/16850-6, for the scholarship and financial support.

The Graduate Program in Pharmaceutical Sciences, School of Pharmaceutical Sciences of Ribeirão Preto, University of São Paulo, for offering an interdisciplinary and international environment of study and research.

Professor Dr. Fernando Batista Da Costa, for his mentorship, guidance, friendship and confidence along these four years of research, and specially for his constant and prospective advices to look for academic opportunities abroad.

Professors Dr. Daniel Kawano, Dr. Fernando Torres and Dr. Maikon Barros, for their support at applying for University of São Paulo. To the Federal university of Rio Grande do Sul and Professor Dr. Gilsane von Poser for my academic background in phytochemistry.

Professor Dr. João Semir and Dr. Marcelo Monge (IB-UNICAMP), as well as Professor Dr. Milton Groppo and Dr. Paola Ferreira (FFCLRP-USP) for collaboration in Barnadesioideae.

Colleagues, members and associates of the AsterBioChem research group for the introduction to metabolomics, support and guidance in laboratory and warm friendship.

Professors Dr. Jairo K. Bastos and Dr. Niege A.J.C. Furtado, technical staff, and colleagues of the Laboratory of Pharmacognosy for the pleasant working environment and friendship.

Dr. Paulo Torres and family for their friendship and support.

My parents, Elizabet and Vidal, for their patience, support and care.

My wife, Karen Lizbett, for the constant support and encouragement on this project.

Wiraqocha

Wiraqocha
"P'unchaw kachun, tuta kachun,"
nispa niq.
"Paqarichun, illarichun,"
nispa niq.
P'unchaw churikiqta
qasillaqta qespillaqta purichiq
runa ruwasqaykipta
k'anchay k'anchay kananpaq
Wiraqochaya
qasilla qespilla punchaw Inka
Runa yana michisqaykiqta
killariy k'anchariy
ama unquchispa, ama nanachispa
qasiqta qespiqta waqaychaspa.

Lord Wiracocha,
Who says
"Let there be day, let there be night!"
Who says,
"Let there be dawn, let it grow light!"
Who makes the Sun, your son,
move happy and blessed each day,
so that man whom you have made has light
My Wiracocha,
shine on your Inca people,
illuminate your servants,
whom you have shepherded,
let them live happy and blessed, preserve them in peace, free of sickness and pain.

ABSTRACT

CCANA CCAPATINTA, G. V. **Metabolomics studies of the subfamily Barnadesioideae (Asteraceae)**. 2018. 81 p. Doctoral Thesis. School of Pharmaceutical Sciences of Ribeirão Preto – University of São Paulo, Ribeirão Preto, 2018.

Metabolomics is emerging as an effective approach for the comprehensive evaluation of medicinal plants, classification of raw material, as well as chemotaxonomic studies. This work demonstrates the applicability of metabolomics, using the subfamily Barnadesioideae (Asteraceae) as a study model, for quality assessment and classification purposes of medicinal species (*Chuquiraga* genus) and a chemotaxonomy study of six Barnadesioideae genera (*Arnaldoa*, *Barnadesia*, *Chuquiraga*, *Dasyphyllum*, *Fulcaldea* and *Schlechtendalia*). First, the LC-MS metabolic profiles of Barnadesioideae demonstrated that this subfamily constitutes a chemically underinvestigated taxa with a complex diversity of phenolic compounds, phenylpropanoid derivatives, alkyl glycosides, and triterpenoid glycosides. The intergeneric relationships within Barnadesioideae genera, based on the comparison of their LC-MS metabolic profiles by exploratory and supervised analyses, displayed similarities to those of the intergeneric relationships obtained by the most recent phylogenetic study based on morphological and molecular markers. Second, the LC-MS metabolic profiles of three *Chuquiraga* species (*C. jussieui*, *C. spinosa* and *C. weberbaueri*) lead to the identification of a significant variety of phenolic compounds, phenylpropanoid derivatives, alkyl glycosides, and triterpenoid glycosides, as well as the establishment of prediction models for geographical origin and species classification, as well as the identification of discriminating metabolites by exploratory and supervised multivariate statistical analysis. Third, a classical approach was carried out by acquiring HPLC chromatographic profiles of three *Chuquiraga* species (*C. jussieui*, *C. spinosa* and *C. weberbaueri*) for profiling phenolic compounds and comparison by exploratory and supervised multivariate statistical analysis. Therefore, our results support metabolomics as a valuable tool in the quality control and classification of medicinal plants as well as in chemotaxonomy studies.

Key-words: Asteraceae, Barnadesioideae, liquid chromatography, mass spectrometry, metabolomics, multivariate statistical analysis.

RESUMO

CCANA CCAPATINTA, G. V. **Estudos metabolômicos da subfamília Barnadesioideae (Asteraceae)**. 2018. 81 f. Tese (Doutorado). Faculdade de Ciências Farmacêuticas de Ribeirão Preto – Universidade de São Paulo, Ribeirão Preto, 2018.

A metabolômica vem se tornando uma abordagem eficaz para a avaliação abrangente de plantas medicinais, classificação de matérias-primas, além de estudos quimiotaxonômicos. Este trabalho demonstra a aplicabilidade da metabolômica, utilizando a subfamília Barnadesioideae (Asteraceae) como modelo de estudo, na avaliação da qualidade e classificação de espécies medicinais (espécies de *Chuquiraga*) e no estudo quimiotaxonômico dos principais gêneros de Barnadesioideae (*Arnaldoa*, *Barnadesia*, *Chuquiraga*, *Dasyphyllum*, *Fulcaldea* e *Schlechtendalia*). Em primeiro lugar, a análise dos perfis metabólicos por LC-MS dos membros de Barnadesioideae demonstrou que esta subfamília constitui um grupo quimicamente não explorado com uma diversidade complexa de substancias fenólicas, fenilpropanoides, alquilglicósidos e glicosídeos triterpenoides. As relações intergenéricas dentro da subfamília Barnadesioideae, baseadas na comparação dos seus perfis metabólicos por análises estatísticas multivariadas, mostraram semelhanças com as relações intergenéricas propostas pelo mais recente estudo filogenético com base em marcadores morfológicos e moleculares. Em segundo lugar, a aquisição dos perfis metabólicos de espécies de *Chuquiraga* (*C. jussieui*, *C. spinosa* e *C. weberbaueri*) por analises de LC-MS, levaram à identificação de uma variedade significativa de compostos fenólicos, fenilpropanoides, alquilglicosídeos e glicosídeos triterpenoides, assim como o estabelecimento de modelos de classificação geográfica e de espécies, além da identificação de metabólitos discriminantes por meio de análises estatísticas multivariadas exploratórias e supervisionadas. Terceiro, uma abordagem clássica foi realizada através da aquisição dos perfis cromatográficos por HPLC de espécies de *Chuquiraga* para o perfilhamento de compostos fenólicos e a classificação das espécies por meio de análises estatísticas multivariadas exploratórias e supervisionadas. Logo, os resultados revelam a metabolômica como uma valiosa ferramenta auxiliar no controle de qualidade e classificação de plantas medicinais, bem como em estudos de quimiotaxonômia.

Palavras-chave: Asteraceae, Barnadesioideae, cromatografia liquida, espectrometria de massas, análise estatística multivariada, metabolômica.

LIST OF FIGURES

LIST OF TABLES.

LIST OF ABBREVIATIONS

CE-MS:	Capillary electrophoresis-mass spectrometry
FT-IR:	Fourier transform-infrared spectroscopy
GC-MS:	Gas chromatography–mass spectrometry
HCA:	Hierarchical cluster analysis
HPLC-DAD:	High-performance liquid chromatography-diode array detection
*k*NN:	*k*-nearest neighbor classification
LC-MS:	Liquid chromatography-mass spectrometry
LC-NMR:	Liquid chromatography-nuclear magnetic resonance
LC-UV:	Liquid chromatography-ultraviolet detection
MS:	Mass spectrometry
NMR:	Nuclear magnetic resonance
OPLS-DA:	Orthogonal partial least squares-discriminant analysis
PCA:	Principal components analysis
PLS-DA:	Partial least squares-discriminant analysis
TLC:	Thin-layer chromatography

INDEX

1. INTRODUCTION

1.1. The subfamily Barnadesioideae

The subfamily Barnadesioideae (Benth. & Hook. F.) K. Bremer & R.K. Jansen comprises more than 90 species distributed in nine genera entirely restricted to South America. Barnadesioideae members share a number of morphological and molecular features that support their position into a separate subfamily (JANSEN and PALMER 1987; BREMER and JANSEN 1992). The presence of axillary spines and barnadesioid trichomes (pubescences of unbranched three-celled hairs) on floral and vegetative structures constitute unique morphological characteristics within Asteraceae that distinguish Barnadesioideae from the rest of the family (CABRERA 1959; EZCURRA 1985, BREMER and JANSEN 1992; STUESSY et al. 2009). Additionally, another feature of Barnadesioideae is the lack of two DNA inversions in their chloroplast genome, which are present in all other Asteraceae (JANSEN AND PALMER 1987; KIM et al. 2005). Phylogenetically, Barnadesioideae has a well-supported position as the sister group of all other Asteraceae (FUNK et al. 2005; PANERO and FUNK 2008; GRUENSTAEUDL et al. 2009; STUESSY et al. 2009).

1.1.1. Distribution

Despite the small number of species, Barnadesioideae genera display a broad range of habits and distinct geographic distributions. The **Table 1** summarizes the current recognized taxa of Barnadesioideae and their correspondent geographic distribution.

The monotypic genera *Duseniella* K.Schum., *Huarpea* Cabrera, and *Schlechtendalia* Less. are herbaceous/subshrubby plants distributed in isolated areas of Argentina, Brazil and Uruguay (STUESSY et al. 2009). The shrubby genus *Fulcaldea* Poir. was considered monotypic until 2011, when a second species was described in the "Chapada Diamantina", Bahia, Brazil (FUNK and ROQUE 2011). The genus *Doniophyton* Wedd. includes two herbaceous species, which are found in xeric areas of Chile and Argentina (KATINAS and STUESSY 1997). The three shrubby species of *Arnaldoa* Cabrera have a narrow distribution in southern Ecuador and northern Peru and grow in more or less xerophytic habitats (STUESSY and SAGÁSTEGUI 1993; ULLOA ULLOA et al. 2002).

Table 1. Species of Barnadesioideae and their geographic distribution.

Genus	Species	Distribution*
Arnaldoa Cabrera	*A. argentea* C.Ulloa, P.Jørg. & M.O.Dillon	EC
	A. macbrideana Ferreyra	PE
	A. weberbaueri (Muschl.) Ferreyra	PE
Barnadesia Mutis ex L.f	*B. aculeata* (Benth.) I.C.Chung	EC
	B. arborea Kunth	EC, PE
	B. blakeana Ferreyra	PE
	B. caryophylla (Vell.) S.F.Blake	BO, BR, PE
	B. corymbosa (Ruiz & Pav.) D.Don	BO, PE
	B. dombeyana Less.	PE
	B. glomerata var. *glomerata* Kuntze	BO
	B. glomerata var. *mucronata* I.C.Chung	BO
	B. horrida Muschl.	BO, PE
	B. jelskii Hieron.	EC, PE
	B. lehmannii var. *lehmannii* Hieron.	EC, PE
	B. lehmannii var. *angustifolia* I.C.Chung	PE
	B. lehmannii var. *ciliata* I.C.Chung	EC
	B. lehmannii var. *villosa* (I.C.Chung) Urtubey	EC, PE
	B. macbridei Ferreyra	PE
	B. macrocephala Kuntze	BO
	B. odorata Griseb.	AR, BO
	B. parviflora Spruce ex Benth. & Hook. f.	CO, EC, PE
	B. polyacantha Wedd.	BO, EC, PE
	B. pycnophylla Muschl.	BO, PE
	B. reticulata D.Don	PE
	B. spinosa L.f.	CO, EC
	B. woodii D.J.N.Hind	BO
Chuquiraga Juss	*C. acanthophylla* Wedd.	AR, BO
	C. atacamensis Kuntze	AR, BO, CH
	C. arcuata Harling	EC
	C. aurea Skottsb.	AR
	C. avellanedae Lorentz	AR
	C. calchaquina Cabrera	AR
	C. echegarayi Hieron.	AR
	C. erinacea subsp. *erinacea* D.Don	AR
	C. erinacea subsp. *hystrix* (D.Don) C.Ezcurra	AR
	C. jussieui J.F.Gmel.	BO, CO, EC, PE
	C. kuschelii Acevedo	CH
	C. longiflora (Griseb.) Hieron.	AR, BO
	C. oblongifolia Sagást. & Sánchez Vega	PE
	C. raimondiana A.Granda	PE
	C. morenonis (Kuntze) C.Ezcurra	AR
	C. oppositifolia D.Don	AR, BO, CH
	C. parviflora (Griseb.) Hieron.	AR, BO
	C. rosulata Gaspar	AR
	C. ruscifolia D.Don	AR
	C. spinosa subsp. *spinosa* Less.	PE
	C. spinosa subsp. *australis* C.Ezcurra	AR, BO, CH
	C. spinosa subsp. *huamanpinta* C.Ezcurra	PE
	C. spinosa subsp. *rotundifolia* (Wedd.) C.Ezcurra	CH, PE
	C. straminea Sandwith	AR
	C. ulicina subsp. *ulicina* Hook.	CH
	C. ulicina subsp. *acicularis* (D.Don) C.Ezcurra	CH
	C. weberbaueri Tovar	PE
Dasyphyllum Kunth	*D. argenteum* Kunth	EC
	D. armatum (J.Kost.) Cabrera	AR, BO
	D. brasiliense var. *brasiliense* (Spreng.) Cabrera	AR, BR, PA
	D. brasiliense var. *barnadesioides* (Tovar) Cabrera	BO, PE
	D. brasiliense var. *divaricatum* (Griseb.) Cabrera	AR, BO
	D. brasiliense var. *latifolium* (Don.) Cabrera	BR
	D. brevispinum Sagást. & M.O.Dillon	PE
	D. cabrerae Sagást.	PE
	D. candolleanum (Gardner) Cabrera	BO, BR, PA
	D. colombianum (Cuatrec.) Cabrera	CO

	D. cryptocephalum (Baker) Cabrera	BR
	D. diacanthoides (Less.) Cabrera	CH, AR
	D. diamantinense Saavedra & M.Monge	BR
	D. donianum (Gardner) Cabrera	BR
	D. excelsum (D.Don) Cabrera	CH
	D. flagellare (Casar.) Cabrera	BR
	D. ferox (Wedd.) Cabrera	BO, PE
	D. floribundum (Gardner) Cabrera	BR, PR
	D. fodinarum (Gardner) Cabrera	BR
	D. horridum (Muschl.) Cabrera	PE
	D. hystrix var. *hystrix* (Wedd.) Cabrera	BO
	D. hystrix var. *peruvianum* (Wedd.) Cabrera	PE
	D. inerme (Rusby) Cabrera	AR, BO, PA
	D. infundibulare (Baker) Cabrera	BR
	D. lanosum Cabrera	BR
	D. lanceolatum (Less.) Cabrera	BR
	D. latifolium (Gardner) Cabrera	BO, BR, PA
	D. lehmannii (Hieron.) Cabrera	EC
	D. leiocephalum (Wedd.) Cabrera	BO, PE
	D. leptacanthum (Gardner) Cabrera	BR
	D. maria-lianae Zardini & Soria	PA
	D. orthacanthum (DC.) Cabrera	BR, PA
	D. popayanense (Hieron.) Cabrera	EC
	D. reticulatum var. *reticulatum* (DC.) Cabrera	BR
	D. reticulatum var. *robustum* Domke ex Cabrera	BR
	D. retinens (S.Moore) Cabrera	BR
	D. spinescens (Less.) Cabrera	BR
	D. sprengelianum var. *sprengelianum* (Gardner) Cabrera	BR
	D. sprengelianum var. *inerme* (Gardner) Cabrera	BR
	D. synacanthum (Baker) Cabrera	BR
	D. tomentosum var. *tomentosum* (Spreng.) Cabrera	AR, BO, BR
	D. tomentosum var. *multiflorum* (Baker) Cabrera	BR
	D. trichophyllum (Baker) Cabrera	BR
	D. vagans (Gardner) Cabrera	BR
	D. varians (Gardner) Cabrera	BR, PR
	D. velutinum (Baker) Cabrera	BR, BO
	D. vepreculatum (D.Don) Cabrera	VE
	D. weberbaueri (Tobar) Cabrera	EC, PE
Doniophyton Wedd.	*D. anomalum* (D.Don) Kurtz	AR, CH
	D. weddellii Katinas & Stuessy	AR, CH
Duseniella K.Schum.	*D. patagonica* (O.Hoffm.) K.Schum.	AR
Fulcaldea Poir.	*F. laurifolia* (Bonpl.) Poir.	EC, PE
	F. stuessyi Roque & V.A.Funk	BR
Huarpea Cabrera	*H. andina* Cabrera	AR
Schlechtendalia Less.	*S. luzulaefolia* Less.	AR, BR, UR

Taxonomy according to STUESSY and SAGÁSTEGUI 1993, and ULLOA ULLOA et al. 2002 for *Arnaldoa*; URTUBEY 1999 and HIND 2001 for *Barnadesia*, EZCURRA 1985, HARLING 1991, SAGÁSTEGUI AND SÁNCHEZ 1991, and GRANADA 1997 for *Chuquiraga*; CABRERA 1959 and 1997, SAGÁSTEGUI 1980, SAGÁSTEGUI and DILLON 1985, ZARDINI and SORIA 1994 and SAAVEDRA et al. 2014 for *Dasyphyllum*; KATINAS and STUESSY 1997 for *Doniophyton*; FUNK and ROQUE 2011 for *Fulcaldea*, CABRERA 1951 for *Huarpea*, and STUESSY et al. 2009 for *Duseniella* and *Schlechtendalia*. Abbreviations: AR = Argentina, BO = Bolivia, BR = Brazil, CH = Chile, CO = Colombia, EC = Ecuador, PA = Paraguay, PE = Peru, UR = Uruguay, VE = Venezuela. *Updated distribution data were consulted on TROPICOS database (www.tropicos.org).

The genera *Barnadesia* Mutis ex L.f, *Chuquiraga* Juss, and *Dasyphyllum* Kunth constitute the largest and most representative taxa of Barnadesioideae; pictures of representative species are displayed in **Figure 1**. *Barnadesia* comprises 19 species of shrubs and trees, mainly distributed in the Andes from Colombia to Argentina, and one species is

found in Brazil, mostly restricted to elevations of 1800-3400 m (URTUBEY 1999; HIND 2001). *Chuquiraga* is a genus of 22 spiny evergreen shrubs that grow along the Andes and the Patagonia at high altitude habitats; however, some species are found at sea level areas in central Chile and Argentina (EZCURRA 1985; HARLING 1991; SAGÁSTEGUI and SÁNCHEZ 1991; GRANADA 1997). *Dasyphyllum* is a genus of shrubs or trees, which comprises 41 species distributed throughout South America, with two centers of diversity, one in western South America, in Andean mountains from Venezuela to north-western Argentina, occupying arid regions such as the Puna, and the other in eastern South America, in Brazil, Bolivia, and Paraguay in Atlantic forest and savanna (CABRERA 1959, 1997; SAGÁSTEGUI 1980; SAGÁSTEGUI and DILLON 1985; ZARDINI and SORIA 1994; SAAVEDRA et al. 2014).

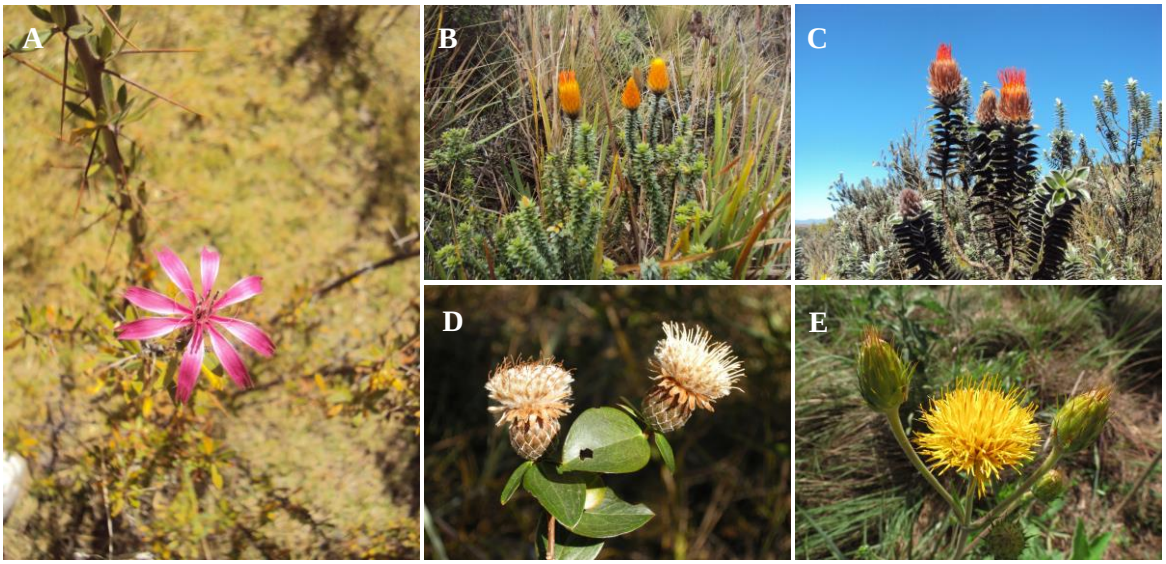


Figure 1. Pictures of representative species of Barnadesioideae: **A**) *Barnadesia horrida* (Q'orimarka, Cusco, Peru), **B**) *Chuquiraga jussieui* (Huancabamba, Piura, Peru), **C**) *Chuquiraga weberbaueri* (Celendin, Cajamarca, Peru), **D**) *Dasyphyllum sprengelianum* (Serra do Cipó, Minas Gerais, Brazil), **E**) *Schlechtendalia luzulaefolia* (Cerro do Tigre, Manoel Viana, Brazil). Photos by: G.V. Ccana-Ccapatinta, G. Shimizu and G. Heiden.

1.1.2. Chemistry

The secondary metabolite chemistry of Barnadesioideae has been sometimes described as following a simple profile (BOHM and STUESSY 1995; ZDERO et al. 1987). This possible simple chemistry profile was proposed and hypothesized as further evidence of the basal position of Barnadesioideae in the Asteraceae family (BOHM and STUESSY 1995; BOHM and STUESSY 2001; CALABRIA et al. 2007). In total, two acetophenones (**1** and **2**) (SENATORE 1996; SENATORE et al. 1999), vanillin (**3**) (HOENEISEN et al. 2000), gallic

acid (**4**) (CASTELUCCI et al. 2007), umbelliferone (**5**) (HOENEISEN et al. 2000), 13 flavonoids (kaempferol, quercetin, isorhammetin, and their 3-*O*-glycosides, **6**-**18**) (BOHM and STUESSY 1995; MENDIONDO et al. 1997, 2000; SENATORE et al. 1999; MENDIONDO and JUÁREZ 2001; JUAREZ and MENDIONDO 2002a, 2002b, 2007; LANDA et al. 2009), and 21 triterpenoids (taraxastane-, lupane-, ursane- and oleanane-type pentacyclic triterpenoids, **19**-**39**) (ZDERO et al. 1987; FLAGG et al. 1999; HOENEISEN et al. 2000; GUROVIC et al. 2010) have been described to date in Barnadesioideae. The correspondent trivial names and chemical structures are presented in **Figure 2** and **Table 2**.

Table 2. Chemical constituents reported in members of Barnadesioideae.

No	Chemical class	Trivial name
1	Phenols	*p*-Hydroxyacetophenone
2		*p*-Methoxyacetophenone
3		Vanillin
4		Gallic acid
5		Umbelliferone
6	Flavonoids	Eriodictyol
7		Kaempferol
8		Kaempferol-3-*O*-glucoside
9		Kaempferol-3-*O*-glucuronide
10		Kaempferol-3-*O*-rutinoside
11		Quercetin
12		Quercetin-3-*O*-glucoside
13		Quercetin-3-*O*-glucuronide
14		Quercetin 3-*O*-rhamnoside
15		Quercetin-3-*O*-rutinoside
16		Isorhamnetin-3-*O*-glucoside
17		Isorhamnetin-3-*O*-glucuronide
18		Isorhamnetin-3-*O*-rutinoside
19	Triterpenoids	α-Amyrin
20		α-Amyrin acetate
21		β-Amyrin
22		β-Amyrin acetate
23		Erythrodiol
24		Friedelinol
25		Taraxasterol
26		Pseudotaraxasterol
27		Faradiol
28		3β,6β-Dihydroxytaraxasta-20-ene
29		3β-Acetoxy-6β-hydroxytaraxasta-20-ene
30		6β-Dydroxytaraxasta-20-ene 3β-palmitate
31		6β-Dydroxytaraxasta-20-en-3-one
32		Lupeol
33		Lupeyl acetate
34		Calenduladiol
35		Betulin
36		Heliantriol B2
37		Lupenone
38		30-Nor-lupan-3β-ol-20-one
39		3β-Acetoxy-30-nor-lupan-20-one

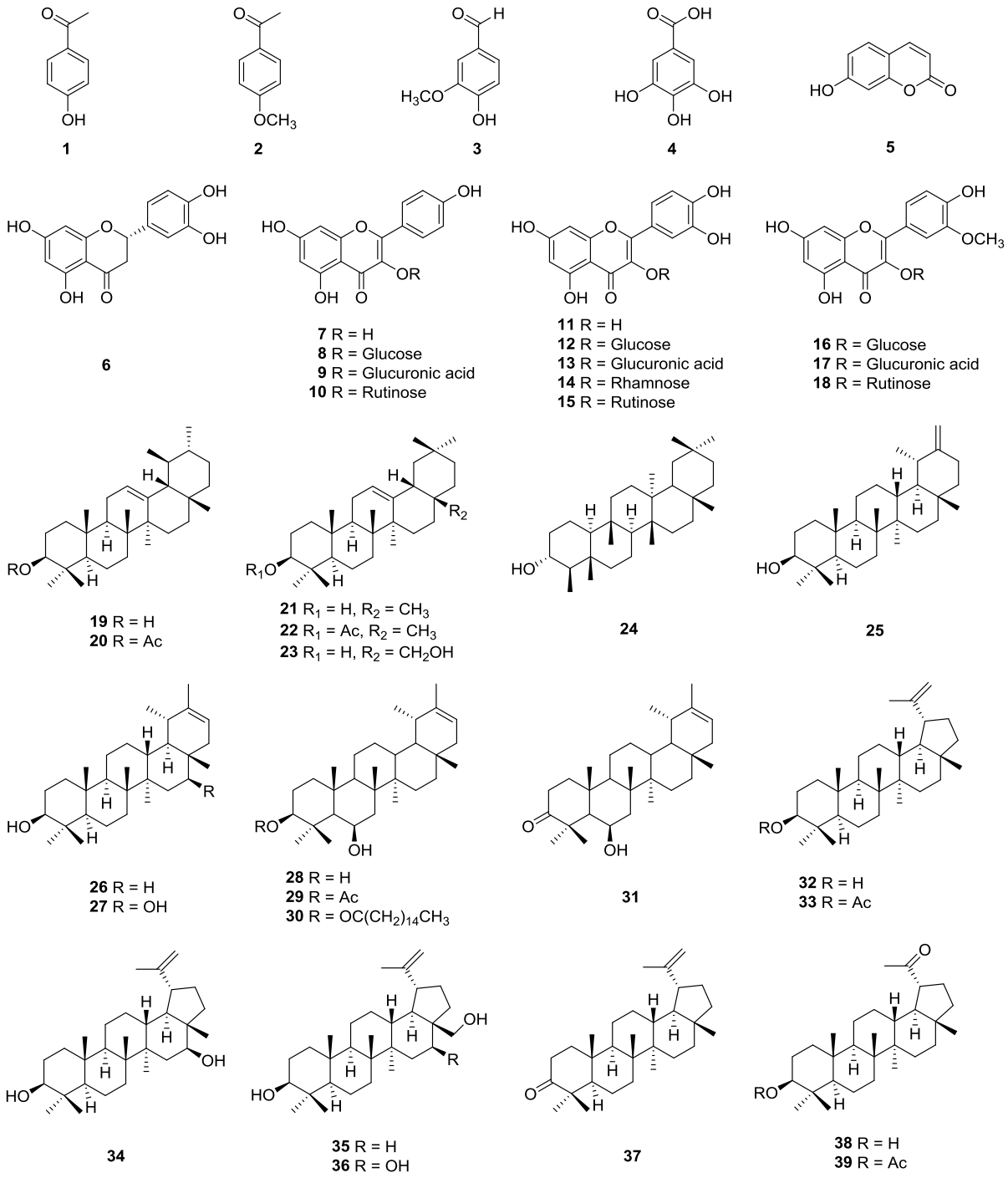


Figure 2. Chemical constituents reported in members of Barnadesioideae.

1.1.3. Medicinal uses

The infusion of the leaves of *Barnadesia arborea* Kunth, distributed among localities of Ecuador and Peru (HIND and HALL 2003), is applied externally for the relief of spasms in children (URTUBEY 1999), while the topical application of its flowers by rubbing is used in the treatment of dermatitis and influenza (TENE et al. 2007). Similarly, the infusion of the

flowers of *B. horrida* Muschl., distributed among highlands of Bolivia and Peru (HERRERA 1933), is used for the treatment of common cold, bronchopneumonia, bronchitis, cough, headache, fever, and stomach ache (YAKOVLEFF and HERRERA 1934; HERRERA 1938; ROERSCH 1994).

The infusion of leaves and thorns of *Dasyphyllum brasiliense* is used for the treatment of inflammatory diseases in São Paulo and Minas Gerais states, Brazil (CASTELUCCI et al. 2007). The cortex decoction of *D. diacanthoides* (Less.) Cabrera is used for the treatment of contusions and rheumatism in Mapuche traditional medicine in Chile (de MÖSBACH 1991).

Several species of the genus *Chuquiraga* are described as being used in the traditional medicine of Argentina, Bolivia, Chile, Colombia, Ecuador and Peru. The medicinal uses of *Chuquiraga* can be traced to times of pre-Columbian South American cultures such as the Incas (GIBERTI 1983; ROERSCH 1994; BRACK 1999; DE-LA-CRUZ et al. 2007), Aymaras (VILLAGRÁN et al. 1998, 2003), and Tehuelches (RAMÍREZ and BELOSO 2002). As a general trend, *Chuquiraga* medicinal species are used as infusions, alone or in mixture with other plants (BUSSMAN et al. 2010; BUSSMAN et al. 2015), for the treatment of respiratory, gastrointestinal, genitourinary and reproductive disorders. The common names and medicinal uses of species of the genus Chuquiraga are displayed in **Table 3**.

Chuquiraga jussieui and *C. spinosa* are frequently used in the treatment of prostatitis and prostate cancer. In this context, ARROYO-ACEVEDO et al. (2017, 2018) described for the first time the protective effect of the administration of *C. spinosa* alcoholic extract on N-methyl nitrosourea (NMU)-induced prostate cancer and gastric cancer in rats. The same extract displayed cytotoxicity in the DU-145 (prostate carcinoma) cell line with a IC_{50} of 2.98 µg/ml (ARROYO-ACEVEDO et al. 2017). Additionally, HERRERA-CALDERON et al. (2017) investigated the cytotoxicity of *C. spinosa* ethanolic extract on the MCF-7 (breast adenocarcinoma), K-562 (chronic myelogenous leukemia), HT-29 (colon adenocarcinoma), H-460 (lung large cell carcinoma), M-14 (amelanotic melanoma), HUTU-80 (duodenum adenocarcinoma), and DU-145 cell lines, obtaining IC_{50} values of 5.32-9.25 µg/ml. Interestingly, the lipophilic fractions (hexane, petrol, chloroform and ethyl acetate) obtained from the initial extract displayed IC_{50} values of 24.19-54.12 µg/ml, suggesting that the active constituents may remain in the polar fractions. The up-to-date reported biological activities of *Chuquiraga* species are summarized in **Table 4**.

Table 3. Common names and medicinal uses of species of the genus *Chuquiraga*.

Species	Country	Common name	Indications*	Reference
C. acanthophylla	AR	Espina amarilla	Cold, cough and fever. Stomachache. Urinary tract infections.	BARBARÁN 2008
C. atacamensis	AR	Hierba de san Pedro, san Pedro, kishka tola	Conjunctivitis, for which the plant is used to make a medicinal smoke. Rheumatic pain, where the plant infusion is used to wash rheumatic legs to relieve pain.	GIBERTI 1983
	BO	San Gerónimo, fundición, kutu kutu, chajllampa	Cold, cough, fever. Urinary tract infections, cystitis, prostatitis. Relief of postpartum symptoms. Not recommended in pregnant women.	ZAMORA 2008
	CH	Lengua de gallo, tastará, quebrolla, killokisca, chana chaklamba	The infusion is used as hot baths against colds. Productive and non-productive cough, fever. Genitourinary and reproductive disorders in women.	VILLAGRÁN et al. 1998, 2003
C. avellanedae	AR	Quilimbay-trayao, tratrakcha, trayau	Cough. Headache and fever, boiled leaves are chewed in a mixture with sugar.	RICHERI et al. 2013
C. erinacea	AR	Romerillo, falsa uña de gato, trifrif mamull	Stomachache and liver disease. Kidney disease. Strengthens the brain and nerves.	LADIO and LOZADA 2009
C. jussieui	CO	Chuquiragua, vela de páramo	Febrifuge, diuretic, kidney stones.	DÍAZ-PIEDRAHITA and VÉLEZ-NAUER 1993
	EC	Chuquira, chuquiragua	Liver disease, diabetes. Allergy and skin disorders. Pain of the bones, rheumatism and other inflammations. Toothache, stomachache and gastrointestinal disorders. Cold, fever, cough and respiratory disorders. Malaria, malarial fever, smallpox, internal infections. Urogenital disorders, diuretic. Relief of postpartum symptoms.	MARTÍNEZ 2006; TENE et al. 2007; ANSALONI et al. 2010
	PE	Chiquiragua (northern Peru). Inca llaulli, kentayllaulli, quishuara, kiswara, kiswara tiutumpi, qharisirviy (southern Peru)	Stomachache and liver disease. Musculoskeletal pain. Skin eruptions, inflammations. Common cold, cough, sore throat, fever, respiratory disorders. Vaginitis and vaginal infections, as external washing. Urinary tract infections, kidney disease, stones, prostatitis. Postpartum symptoms. Endoparasiticide (intestinal worms), and ectoparasiticide (lice). Rheumatic pain, an infusion is used to wash the legs.	TORRES et al. 1992; ROERSCH 1994; DE FEO 2003; VÁSQUEZ et al. 2010
C. longiflora	AR	Azafrán de la puna	The plant is added to water for personal washing.	GIBERTI 1983
C. oppositifolia	AR	Azafrán del campo	Hypoglycaemic, hypocholesterolaemic. Antifungal.	RAAD 2012
C. parviflora	BO	Chiñi michi michi	Against curse.	VANDEBROEK et al. 2003
C. spinosa	AR	Charkoma	Regulation of the menstrual cycle.	GIBERTI 1983
	BO	Huamanpinta	Kidney stones and cystitis.	CEUTERICK et al. 2011
	EC	Chuquiragua	Cold, cough and fever. Pain of the bones. Malaria.	BUSSMANN and SHARON 2006b
	PE	Huamanpinta, huancapita, huancaspita, laulinco, pucacasha, pazpapamaquin, qharisirviy, cjari sirvi	Respiratory affections. Antiblenorrhaegic and vermifuge. Conjunctivitis. Gonorrhoea. Urinary system disorders in women and men. Vaginitis and vaginal infection, the infusion of the plant is uses for external washing. Kidney and prostate inflammations. Prostate cancer. Diuretic. Sexual impotence.	BRACK 1999; MADALENO 2007; REHECHO et al. 2011
C. weberbaueri	PE	Amaro amaro	Cough, bronchitis, asthma. Liver disease. Diuretic and depurative.	BRACK 1999

*Commonly, aerial parts are used to make infusions or decoctions in water; other modes of use are detailed in the table text. Abbreviations: AR = Argentina, BO = Bolivia, CH = Chile, CO = Colombia, EC = Ecuador, PE = Peru.

Table 4. Biological activities reported for species of the genus *Chuquiraga.*

Plant	Extracting Solvent, standardization	Bioactivity	Results	Reference
C. atacamensis	80% Ethanol (5 g/100 ml)*, 500 µg of GAE/ml	*In vitro* COX-1 inhibition	IC_{50} = 2 µg/ml	ALBERTO et al. 2009
		In vitro COX-2 inhibition	IC_{50} = 4.7 µg/ml	
		Antioxidant, $DPPH^{\cdot}$, $ABTS^{\cdot+}$, $O^{\cdot-}_2$	IC_{50} = 3.5 - 20 µg/ml	
	80% Ethanol (5 g/100 ml)*, 500 µg of GAE/ml	*Staphylococcus aureus* strains	MIC = 80–600 µg/ml	ZAMPINI et al. 2009
		Enterococcus faecalis strains	MIC = 150–300 µg/ml	
		Escherichia coli strain	MIC = 600 µg/ml	
		Other gram-negative bacteria	MIC = 300–600 µg/ml	
	Ethanol (dry extract)	Antioxidant, $ABTS^{\cdot+}$ assay	SC_{50} = 1.5 µg/ml	ZAMPINI et al. 2010
C. erinacea	Ethanolic extract (dry extract)	*In vitro* AChE inhibitory activity	IC_{50} = 7.26 mg/ml	GUROVIC et al. 2010
C. jussieui	Water (2 g/100 ml)*	Antioxidant	IC_{50} = 64.9 mg/L	DUEÑAS et al. 2014
C. spinosa	Water (dry extract), 5.4 mg GAE/mg	Antioxidant, $DPPH^{\cdot}$, $ABTS^{\cdot+}$, $O^{\cdot-}_2$	IC_{50} = 9.6 - 30.5 µg/ml	CASADO et al. 2011
		Candina albicans	MIC = 2.5 µg on TLC plate	
		Cladosporium cucumerinum	MIC = 2.5 µg on TLC plate	
		Rhizopus stolonifer	MIC = 4.6 µg on TLC plate	
	50% Methanol (dry extract), 6.3 mg GAE/mg	Antioxidant: $DPPH^{\cdot}$, $ABTS^{\cdot+}$, $O^{\cdot-}_2$	IC_{50} = 8.5 - 21.7 µg/ml	CASADO et al. 2011
		Antiinflammatory, paw edema in rats	Maximal inhibition = 52.5%	
		Antiinflammatory, ear edema in mice	Inhibition = 88.1%	
		Candina albicans	MIC = 6.3 µg on TLC plate	
		Rhizopus stolonifer	MIC = 13.5 µg on TLC plate	
	Methanol (dry extract), 12.6 mg GAE/mg	Antioxidant, $DPPH^{\cdot}$, $ABTS^{\cdot+}$, $O^{\cdot-}_2$	IC_{50} = 10.5 – 36.5 µg/ml	CASADO et al. 2011
		Rhizopus stolonifer	MIC = 18.5 µg on TLC plate	
	Water (5g/500 ml)*	*Staphylococcus aureus* strain	13 mm, agar diffusion test	BUSSMAN et al. 2008
	96% ethanol (dry extract)	Cytotoxicity in DU-145 cell line	IC_{50} = 2.98 µg/ml	ARROYO-ACEVEDO et al. 2017
	96% ethanol (dry extract)	Cytotoxicity in MCF-7 cell line	IC_{50} = 9.25 µg/ml	HERRERA-CALDERON et al. 2017
		Cytotoxicity in K-562 cell line	IC_{50} = 7.34 µg/ml	
		Cytotoxicity in HT-29 cell line	IC_{50} = 8.52 µg/ml	
		Cytotoxicity in H-460 cell line	IC_{50} = 5.32 µg/ml	
		Cytotoxicity in M-14 cell line	IC_{50} = 8.30 µg/ml	
		Cytotoxicity in HUTU-80 cell line	IC_{50} = 6.20 µg/ml	
		Cytotoxicity in DU-145 cell line	IC_{50} = 7.09 µg/ml	
	Hexane fraction (dry extract)	Cytotoxicity in DU-145 cell line	IC_{50} = 27.03 µg/ml	
	Petroleum ether fraction (dry extract)	Cytotoxicity in DU-145 cell line	IC_{50} = 33.10 µg/ml	
	Chloroform fraction (dry extract)	Cytotoxicity in DU-145 cell line	IC_{50} = 24.19 µg/ml	
	Ethyl acetate fraction (dry extract)	Cytotoxicity in DU-145 cell line	IC_{50} = 54.12 µg/ml	
C. straminea	80% Methanol (dry extract)	Antioxidant, $DPPH^{\cdot}$, $ABTS^{\cdot+}$	SC_{50} = 14.5 - 34.9 µg/ml	MENDIONDO et al. 2011
		Staphylococcus aureus strains	MIC = 200 - 800 µg/ml	

* Plant/solvent ratio. GAE, Gallic acid equivalents

1.1.4. Toxicity

There are few data about the toxicity or side effect of species of the genus *Chuquiraga*. The aqueous extracts of *C. spinosa* and *C. weberbaueri* displayed median lethal doses (LD_{50}) >10,000 µg/mL in the brine shrimp lethality assay, whereas the ethanolic extracts displayed LC_{50} values of 1.1 and 0.25 µg/mL, respectively (BUSSMANN et al. 2011). Even though there is a report discouraging the administration of *C. atacamensis* infusions in pregnant women because it could cause miscarriage (ZAMORA 2008), additional studies are required to reveal the possible toxicity and side effect of *Chuquiraga* species and other representatives of Barnadesioideae subfamily.

1.1.5. Commercialization

The medicinal species of *Chuquiraga* are important and evident elements in medicinal plant markets of traditional cities of Ecuador and Peru but also in modern cities such as Guayaquil and Lima, and at least one species has been introduced in the international market. Differently to markets of Ecuador and Peru, where commercialization of *Chuquiraga* species is frequent, the commercialization of *Chuquiraga* species in Markets of Bolivia seems to be absent (MACÍA et al. 2005; BUSSMANN et al. 2016).

Representative pictures from commercial samples of *Chuquiraga* species are displayed in **Figure 3**. *Chuquiraga jussieui* is one of the most popular medicinal plants in Ecuador and has been noted as a plant with promising industrial potential (BUITRON 1999; MARTÍNEZ 2006; GUPTA 2006). The flowering parts of this species are also found in markets of northern Peru together with *C. weberbaueri* (BUSSMANN et al. 2007). In the markets of southern Peru, the inflorescences of *C. jussieui* are frequently commercialized separately from leaves and steams (**Figure 3A**). The aerial parts of *C. spinosa* are sold along the main cities of Peru (MADALENO 2007; CEUTERICK et al. 2011; HUAMANTUPA et al. 2011; **Figure 3B**). This species is also distributed as a dietary supplement in Europe (Huamanpinta, Esparta GmcH, www.paracelmed.com; **Figure 3C**) and North America (Huamanpinta, Alpha Omega Labs, www.alphaomegalabs.com). Products that contain *C. spinosa*, mixed with other plants, can also be found, for example, Women's Care Blend (Amazon, www.amazon.com), Prostate Care Blend and Kidney Cleanser Blend (Fito Global Inc., www.fitoglobal.com).

Figure 3. Commercial samples of *Chuquiraga* species: **A**) Flowers of *C. jussieui* (Market in Puno City, Peru); **B**) Aerial parts of *C. spinosa* (Market in La Oroya City, Peru); **C**) Capsules containing *C. spinosa* powder (Commercialized in Austria and Germany). Photos by: G.V. Ccana-Ccapatinta.

1.2. Metabolomics

In parallel to the terms "genome" (complete set of genes present in a cell or organism) and "proteome" (entire set of proteins expressed by a cell or organism), the set of metabolites synthesized by a biological system constitute the "metabolome" (OLIVER et al. 1998; FIEHN 2002). Therefore, metabolomics can be defined as the comprehensive, qualitative, and quantitative analysis of all metabolites in a biological system by high-throughput analytical strategies (GOODACRE et al. 2004; KOPKA et al. 2004; GOODACRE 2005; ROCHFORT 2005). Metabolomics is fast becoming the approach of choice across a broad range of sciences including systems biology, drug discovery, molecular and cell biology, and other medical and agricultural sciences because of the continuous analytical and computational developments conducted in this research area.

1.2.1. Metabolomics approaches

Metabolomics, in the strict sense, involves the measurement of all metabolites in a given system; however, this is not yet technically possible because of the lack of a simple automated analytical strategy that can record the metabolome in a reproducible and robust way

(GOODACRE et al. 2004; KOPKA et al. 2004). Accordingly, three technical approaches were initially described that intended to highlight the options available for monitoring the metabolome. However, the practical and conceptual boundaries between "metabolic fingerprinting" and "metabolite/metabolic profiling" could be sometimes unclear. Thus, other researchers subdivided the metabolomics analytical methodologies into only two categories, namely, "untargeted analysis/metabolite profiling" and "target analysis" (VILLAS-BÔAS et al. 2005; VINAYAVEKHIN and SAGHATELIAN 2010; ROBERTS et al. 2012). Although metabolomics is a relatively new research field and the used terminologies are still evolving (ERNST et al. 2014), the **Table 5** presents commonly used terminologies.

Table 5. Terminologies commonly used in metabolomics research.

Term	Definition
Metabolomics	Comprehensive, qualitative, and quantitative analysis of all metabolites in a biological system by high-throughput analytical strategies.
Metabolome	Complete set of small-molecule chemicals found within a biological sample.
Metabolic fingerprinting	Rapid high-throughput screening of all detectable analytes in a sample without mandatory identification.
Metabolite/metabolic profiling	The identification and quantification of a number of pre-defined metabolites, which may be associated with the same pathway or belong to the same class of compounds.
Metabolite target analysis	Metabolites are selected prior to analysis, by optimized extraction or specific separation and/or detection.
Untargeted analysis	Rapid analysis of a large number of different metabolites in which quantification is not mandatory.
Targeted analysis	Procedure that must include the identification and absolute quantification of selected metabolites.

Adapted from ERNST et al. (2014).

The analytical techniques used for metabolome data acquisition include spectroscopic (nuclear magnetic resonance, NMR) and spectrometric (mass spectrometry, MS) methods, either directly or in association with chromatography. Among them, the association of high-resolution mass spectrometry with gas or liquid chromatography (GC-MS and LC-MS) are likely the most commonly applied tools because of their high sensitivity and comprehensiveness of the acquired data (CAJKA and FIEHN 2016; FIEHN 2016; GORROCHATEGUI et al. 2016; ROCHAT 2016). Other techniques such as capillary electrophoresis tandem mass spectrometry (CE-MS), high-performance liquid chromatography with diode array detection (HPLC-DAD), infrared spectroscopy (FT-IR), among others, have also been described (ERNST et al. 2014), and are compared in **Table 6**.

Table 6. Some standard techniques for untargeted and targeted metabolomics.

	Sensitivity	Throughput	Comprehensiveness
FT-IR	Low	High	Low
NMR	Low	Low-high	Low-high
LC-NMR	Low	Low	High
LC-MS	High	High	High
GC-MS	High	High	High
CE-MS	High	Medium	High
LC-UV	Medium-high	High	Low

Modified from WECKWERTH and MORGENTHAL (2005).

1.2.2. Multivariate data analysis

Metabolomics is placed at the interface between chemistry, biology, statistics and computer science, thus requiring multidisciplinary skills. The high-dimensional nature of metabolome datasets acquired in a metabolomics study requires multivariate data analyses to turn data into knowledge (GOODACRE et al. 2004, BEISKE et al. 2015). These analyses can be classified as unsupervised and supervised multivariate methods, whose general description are displayed in **Table 7**. Popular unsupervised analyses include hierarchical cluster analysis (HCA) and principal components analysis (PCA) that are also known as exploratory methods. Supervised analyses require adequate validation by establishing a *training set*, used to build a model, a *validation set*, used to validate the model, and a *test set*, used to test the model. Popular algorithms include partial least squares discriminant analysis (PLS-DA), *k*-nearest neighbor classification (*k*NN), and neural networks.

Table 7. Description of unsupervised and supervised multivariate analyses.

Multivariate analysis	Popular algorithms	Description
Unsupervised	HCA, PCA, Kohonen neural networks	The system is shown a set of inputs and then left to cluster the metabolite data into groups. For multivariate analysis this optimization procedure is usually 'simplification' or dimensionality reduction; this means that a large body of metabolite data are summarized by a few parameters with minimal loss of information. After clustering, the ordination plots or dendrograms are interpreted
Supervised	PLS-DA, *k*NN, back-propagation neural networks	The desired responses (Y data or 'traits' or 'classes') associated with each of the inputs (X data, or 'metabolome data') are known. The goal is to find a mathematical transformation (model) that will correctly associate all or some of the inputs with the target traits. Such inductive methods allow one to discover which metabolites (inputs) are key for the separation of the traits to be predicted.

Adapted from GOODACRE et al. (2004).

1.2.3. Metabolite identification in metabolomics

In 2007, the Metabolomics Standards Initiative (MSI) Working Group on Chemical Analysis (CAWG) recommended the minimum standards in chemical analysis reports in metabolomics studies as detailed in **Table 8**. These standards recommend that authors should differentiate and report the level of identification accuracy for all reported metabolites based on a four-level system that varies from *level 1* (identified compound) through *levels 2* and *3* (annotated compounds or compound class identification) to *level 4* (unidentified or unclassified metabolites which, however, can be differentiated based on spectral data). The identification of metabolites is essential for integrating metabolomics data into other information disciplines (SUMNER et al. 2007; CREEK et al. 2014).

Table 8. Proposed minimum metadata relative to metabolite identification in metabolomics.

No	Level of identification	Details
1	*Identified compounds*	**Non-novel metabolite**: a minimum of two independent and orthogonal data relative to an authentic compound analyzed under identical experimental conditions to validate non-novel metabolite identifications (e.g. retention time/index and mass spectrum, retention time and NMR spectrum, accurate mass and tandem MS, accurate mass and isotope pattern, full ^{1}H and/or ^{13}C NMR, 2D NMR spectra). **Novel metabolite**: metabolites identified for the first time and which represent novel identifications should include sufficient evidence for full structural identification. Traditionally, it involves extraction, isolation, and purification followed by accurate mass measurement, ion mass fragmentation patterns, NMR (^{1}H, ^{13}C, 2D), and other spectral data or chemical derivatization.
2	*Putatively annotated compounds*	Without chemical reference standards, based upon physicochemical properties and/or spectral similarity with public/commercial spectral libraries. The use of literature values reported for authentic samples by other laboratories.
3	*Putatively characterized compound classes*	Based upon characteristic physicochemical properties of a chemical class of compounds, or by spectral similarity to known compounds of a chemical class.
4	*Unknown compounds*	Although unidentified or unclassified these metabolites can still be differentiated and quantified based upon spectral data.

Levels of identification following SUMNER et al. (2007).

1.2.4. Metabolomics for chemotaxonomy

According to ERDTMAN (1963), chemotaxonomy developed as "very early in the development of natural products chemistry it occurred to many botanists and chemists that it should be possible to characterize and classify plants based on their chemical constituents".

Then, chemotaxonomy is the attempt to classify and identify organisms based on differences and similarities in their biochemical compositions (ERDTMAN 1963; WINK and WATERMAN 1999). Chemotaxonomy accompanied the technical advances in phytochemical analysis from paper chromatography (PC), throughout thin layer chromatography (TLC), high-performance liquid chromatography (HPLC) to finally rely on the isolation and structure elucidation (principally by NMR and MS spectrometry) of plant constituents (WINK and WATERMAN 1999; WINK et al. 2010).

In this regard, the recent analytical improvements established by metabolomics platforms offer the possibility of scanning the metabolome of a numerous set of plants to compare them and evaluate taxonomic hypothesis (WATERMAN 2007, REYNOLDS 2007). The use of metabolomics in chemotaxonomy studies contributes to the classification of plants when uncertainty exists using classical botanical methods. For example, MESSINA et al. (2014) used LC-MS-based metabolomics to test taxonomic boundaries in the *Olearia phlogopappa* (Asteraceae) complex, confirming the limits of closely related taxa where DNA sequence data has been uninformative. Similar studies have been conducted with Brazilian members of the genus *Vernonia* and tribe *Vernonieae* (Asteraceae) (MARTUCCI et al. 2014; GALLON et al. 2018).

1.2.5. Metabolomics for quality control of medicinal plants

The quality control of botanical drugs begins with the authentication of the botanical raw material, continues through the preparation of the botanical drug extract and culminates in the botanical drug product. These products (e.g. crude plant extracts) sold as nutraceuticals or phytopharmaceuticals require that their composition is assessed with precision and kept constant. For this purpose, a usual way to standardize an extract is to quantify its active(s) principle(s). Often, however, the active(s) principle(s) are not clearly defined, and the standardization can be made on a characteristic compound of a given plant, which serves as a marker but may not be directly linked to the biological activity of the extract. Chromatographic (TLC, HPLC) and spectroscopic (NMR, IR, UV) fingerprinting have been used as tools for the quality control of medicinal plants with few phytochemical information, however, representing a challenging analytical task since these mixtures are usually composed of hundreds of different compounds (ULRICH-MERZENICH et al. 2007; WOLFENDER et al. 2010).

In this context, metabolomics enables to obtain a global idea of the compositions of a crude extract, and consequently evaluation of its quality. One advantage of applying metabolomics for quality control is that medicinal plants are evaluated based not only on a limited number of metabolites that may be (or not) pharmacologically important, but on the overall metabolome that includes known/unknow, minor/major metabolites. Then, metabolomics is now established as an approach for the comprehensive evaluation and quality control of medicinal plants, classification of raw material, definition of the degree of similarity between extracts (*phytoequivalence*), and identification of adulterations (ULRICH-MERZENICH et al. 2007; HEINRICH 2008; OKADA et al. 2010; WOLFENDER et al. 2010).

1.3. Metabolomic studies of the subfamily Barnadesioideae

Compared to other tribes or genera of, the subfamily Barnadesioideae constitutes a phytochemically underinvestigated group of Asteraceae, therefore, in this project a LC-MS approach was used to establish a metabolomics-based chemotaxonomic classification and to explore its phytochemical composition. On the other hand, species of the genus *Chuquiraga* are frequently used in traditional medicine and some products are even distributed in the international market, demanding adequate phytochemical characterization and quality control procedures that are not available in the current literature. For this purpose, two analytical setups, HPLC-DAD and LC-MS, were used to establish a species classification of medicinal species of the genus *Chuquiraga*. Therefore, the present project aimed to establish:

- A metabolomics-based chemotaxonomic classification of the subfamily Barnadesioideae.
- A metabolomics-based species classification of medicinal species of the genus *Chuquiraga*.
- A chromatographic profile-based species classification of medicinal species of the genus *Chuquiraga*.

2. OBJECTIVE

To carry out metabolomic studies and develop strategies for species classification and chemotaxonomy on members of the subfamily Barnadesioideae (Asteraceae).

3. MATERIAL AND METHODS

3.1. Plant material

3.1.1. Sampling the subfamily Barnadesioideae

Silica dried leaves of 56 Barnadesioideae taxa, corresponding to 40 species among six genera distributed along several South American localities were obtained from the SPFR Herbarium (Department of Biology, FFCLRP, USP) and UEC Herbarium (IB, UNICAMP) in collaboration with Prof. Dr. Milton Groppo Jr., Dr. Paola Ferreira, Prof. Dr. João Semir and Dr. Marcelo Monge, as well as field collections. The **Figure 4** represents the collection localities of Barnadesioideae samples along several south American countries. The identification of the species was conducted in accordance to the latest taxonomic treatments for the genera *Arnaldoa* (STUESSY and SAGÁSTEGUI 1993; ULLOA ULLOA et al. 2002), *Barnadesia* (URTUBEY 1999; HIND 2001), *Chuquiraga* (EZCURRA 1985; HARLING 1991; SAGÁSTEGUI and SÁNCHEZ 1991; GRANADA 1997), *Dasyphyllum* (CABRERA 1959 and 1997; SAGÁSTEGUI 1980; SAGÁSTEGUI and DILLON 1985; ZARDINI and SORIA 1994; SAAVEDRA et al. 2014), *Fulcaldea* (FUNK and ROQUE 2011), and *Schlechtendalia* (STUESSY et al. 2009). The corresponding scientific names, collector numbers and collection localities of Barnadesioideae samples are detailed in **Table 9**.

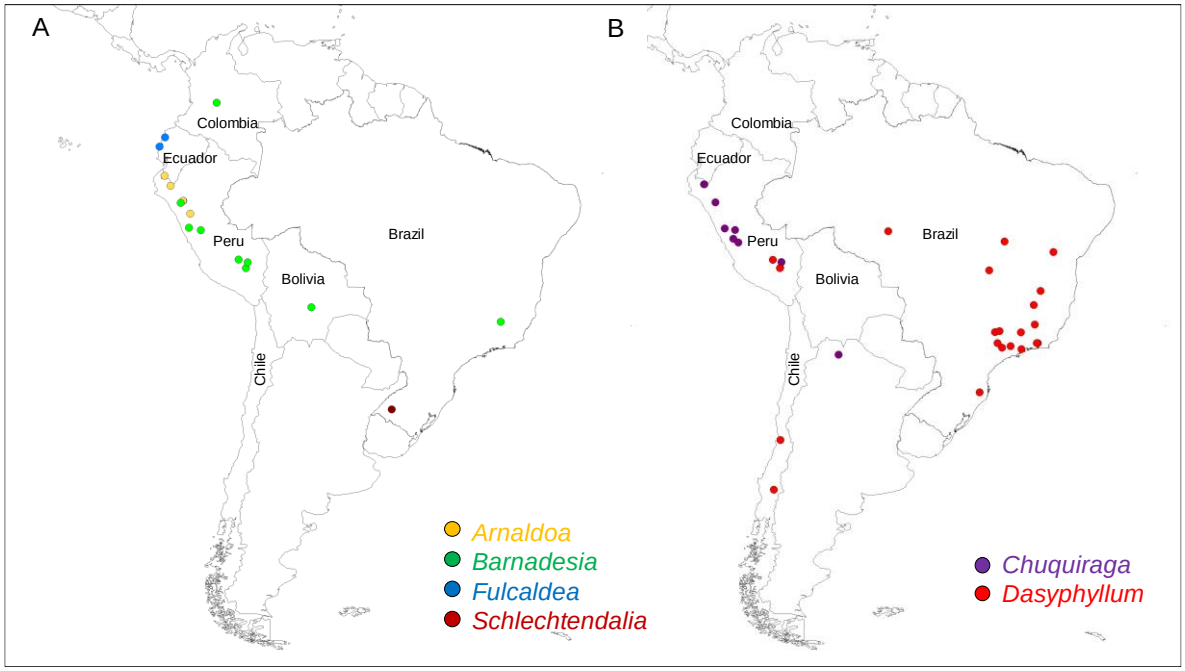


Figure 4. Collection localities of Barnadesioideae subfamily samples.

Table 9. Scientific names, collector numbers and localities of samples of Barnadesioideae.

No	Taxon	Collector and number	Locality	Date (MM-YY)
1	*A. weberbauery* 1	G.V. Ccana 44	Celendin, Cajamarca, Peru	01-2016
2	*A. weberbauery* 2	G.V. Ccana 44	Celendin, Cajamarca, Peru	01-2016
3	*A. weberbauery* 3	A. López & A. Sagástegui 8159	Pataz, La Libertad, Peru	06-1974
4	*A. weberbauery* 4	D.N. Smith & I.M. Sánchez 4323	Celendin, Cajamarca, Peru	07-1983
5	*A. macbrideana*	R. Ferreyra 13628	Huancabamba, Piura, Peru	07-1959
6	*A. argentea* 1	J.E. Madsen 8341	Pindal, Loja, Ecuador	08-2001
7	*A. argentea* 2	J.E. Madsen & C. Rosales 8175	Pindal, Loja, Ecuador	05-2001
8	*B. caryophylla*	M. Monge 2072	Santa Barbara, Minas Gerais, Brasil	08-2013
9	*B. dombeyana* 1	G.V. Ccana 45	Celendin, Cajamarca, Peru	01-2016
10	*B. dombeyana* 2	G.V. Ccana 47	Recuay, Huaraz, Peru	01-2016
11	*B. horrida* 1	G.V. Ccana 54	Urubamba, Cusco, Peru	02-2016
12	*B. horrida* 2	G.V. Ccana 57	Sicuani, Cusco, Peru	02-2016
13	*B. pycnophylla* 1	G.V. Ccana 49	Tomayquichua, Huánuco, Peru	01-2016
14	*B. pycnophylla* 2	G.V. Ccana 53	Ocongate, Cusco, Peru	01-2016
15	*B. odorata*	R. Forzza 2328	Vallegrande, Santa Cruz, Bolívia	12-2002
16	*B. spinosa* 1	M. Gutiérrez 611	Sumapaz, Bogotá, Colombia	02-2016
17	*B. spinosa* 2	M. Gutiérrez 611	Sumapaz, Bogotá, Colombia	02-2016
18	*C. jussieui* 1	G.V. Ccana 10	Huancabamba, Piura, Peru	01-2015
19	*C. jussieui* 2	G.V. Ccana 10	Huancabamba, Piura, Peru	01-2015
20	*C. longiflora* 1	C.M. Martín 736	Valle grande, Jujuy, Argentina	05-2016
21	*C. longiflora* 2	C.M. Martín 901	Valle grande, Jujuy, Argentina	04-2016
22	*C. spinosa* 1	G.V. Ccana 46	Recuay, Huaraz, Peru	01-2016
23	*C. spinosa* 2	G.V. Ccana 50	Huayllay, Cerro de Pasco, Peru	01-2016
24	*C. spinosa* 3	G.V. Ccana 51	Tarma, Junín, Peru	01-2016
25	*C. raimondiana* 1	G.V. Ccana 48	Tomayquichua, Huánuco, Peru	01-2016
26	*C. raimondiana* 2	G.V. Ccana 48	Tomayquichua, Huánuco, Peru	01-2016
27	*C. weberbaueri* 1	G.V. Ccana 43	Celendin, Cajamarca, Peru	01-2016
28	*C. weberbaueri* 2	G.V. Ccana 43	Celendin, Cajamarca, Peru	01-2016
29	*C. jussieui* 3	G.V. Ccana 52	Ocongate, Cusco, Peru	02-2016
30	*C. jussieui* 4	G.V. Ccana 59	Sicuani, Cusco, Peru	02-2016
31	*D. diacanthoides*	C.L. Silva-Luz 200	Curacautín, La Araucanía, Chile	09-2013
32	*D. excelsum*	C.L. Silva-Luz 195	Olmué, Valparaíso, Chile	09-2013
33	*D. vagans* 1	P.L. Ferreira 50	Ribeirão Preto, São Paulo, Brasil	09-2015
34	*D. candolleanum*	M.D. Moraes 756	Lavras, Minas Gerais, Brasil	07-2005
35	*D. cryptocephalum*	M.D. Moraes 131	Petrópolis, Rio de Jaeiro, Brasil	04-2006
36	*D. diamantinense*	P.L. Ferreira 9	Palmeiras, Bahia, Brasil	05-2013
37	*D. donianum*	A.V. Scatigna 126	Rio da conceição, Tocantins, Brasil	08-2012
38	*D. flagellare*	M. Monge 503	Camanducaia, Minas Gerais, Brasil	08-2008
39	*D. ferox*	G.V. Ccana 58	Sicuani, Cusco, Peru	02-2016
40	*D. fodinarum*	G.A.D. Franco 1242	Cunha, São Paulo, Brasil	08-1994
41	*D. infundibulare*	A.A. Santos 126	Niquelândia, Goiás, Brasil	07-1998
42	*D. lanceolatum*	P.L. Ferreira 49	Ribeirão Preto, São Paulo, Brasil	09-2016
43	*D. leptacanthum*	P.L. Ferreira 1	Morro do Cuca, Rio de Janeiro, Brasil	05-2013
44	*D. leiocephalum*	G.V. Ccana 55	Urubamba, Cusco, Peru	02-2016
45	*D. retinens*	J.B. Andrade 3335	Aripuanã, Mato Grosso, Brasil	05-1976
46	*D. reticulatum*	I.M. Franco 1277	Diamantina, Minas Gerais, Brasil	08-2013
47	*D. sprengelianum*	P.L. Ferreira 54	Altinópolis, São Paulo, Brasil	10-2015
48	*D. spinescens*	M. Monge 2095	Camanducaia, Minas Gerais, Brasil	08-2013
49	*D. synacanthum*	R.G. Udulutsch 370	Rio Claro, São Paulo, Brasil	06-2001
50	*D. tomentosum*	M. Monge 1592	Bom Retiro, Santa Catarina, Brasil	05 2013
51	*D. trichophyllum*	V.C. Souza 25816	Cristalia, Minas Gerais, Brasil	08-2001
52	*D. vagans* 2	M. Monge 2069	Campinas, São Paulo, Brasil	08-2013
53	*D. velutinum*	L.D. Meireles 1383	Ouro Preto, Minas Gerais, Brasil	04-2003
54	*F. laurifolia* 1	T. Delinks 399	Jaramijó, Manabí, Ecuador	10-2005
55	*F. laurifolia* 2	J. Cevallos & D.A. Neill 213	Pedernales, Manabí, Ecuador	10-2005
56	*S. luzulaefolia*	M. Monge 3140	Manoel Viana, Rio grande do Sul, Brasil	04-2014

3.1.2. Sampling *Chuquiraga* medicinal species

Commercial samples of medicinal species of *Chuquiraga* were purchased in medicinal plants markets along 31 cities between Ecuador and Peru, in January 2015, the surveyed cities are represented in **Figure 5**. The samples were composed of dried flowering aerial parts. Then, a sample portion in each city was conditioned to build a voucher for botanical identification and deposited in the SPFR Herbarium (Department of Biology, Faculty of Philosophy, Science and Letters of Ribeirão Preto, University of São Paulo, FFCLRP-USP) and UEC Herbarium (Institute of Botany, State University of Campinas, IB-UNICAMP).

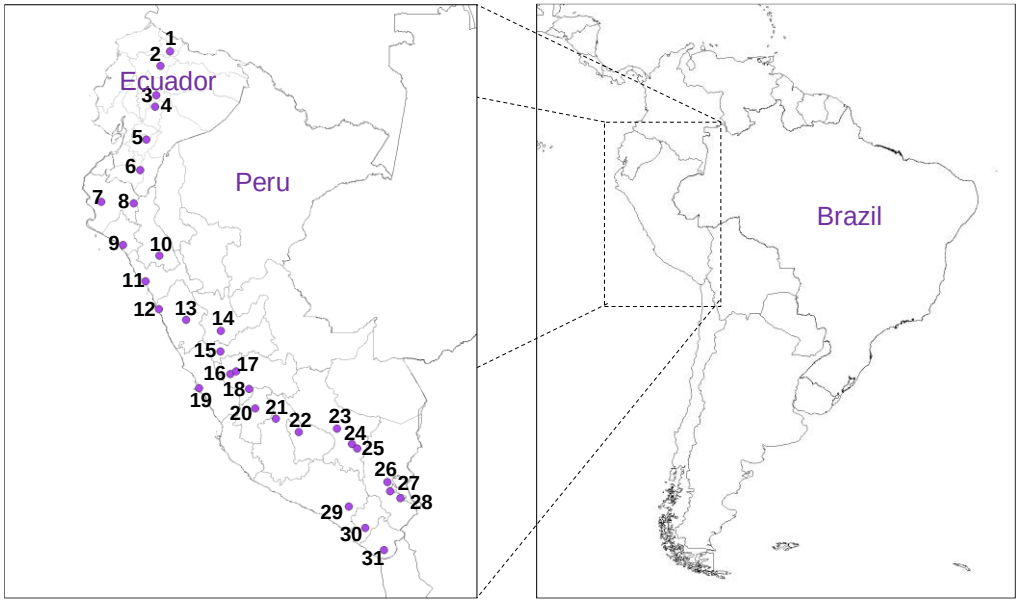


Figure 5. Commercialization localities of *Chuquiraga* species in Ecuador and Peru.

The samples were identified as *C. jussieui*, *C. weberbaueri*, and *C. spinosa*, based on the most recent taxonomic treatment (EZCURRA 1985; HARLING 1991; SAGÁSTEGUI and SÁNCHEZ 1991; GRANADA 1997), and are detailed in **Table 10**. In several cities, commercial samples of *Chuquiraga* were also distributed in small packages containing only dried leaves with the label of "Huamanpinta", some of which did not declare the scientific name of the species, while others declared to contain plant material of *C. spinosa*. Throughout the year 2015, samples of "Huamanpinta" products were also acquired over internet websites in the international market, all of which declared to contain *C. spinosa* plant material. Characteristic "Huamanpinta" commercial samples are displayed in **Figure 6**. Additionally, some samples from three species of *Chuquiraga* were purchased in January 2016.

Table 10. Localities and samples of *Chuquiraga* species acquired in Ecuador and Peru.

No.	Country	City name	Number of samples	Species/"Label"	Common names
1	Ecuador*	Ibarra	5	*C. jussieui*	Chuquiragua
2	Ecuador	Quito	7	*C. jussieui*	Chuquiragua
3	Ecuador	Ambato	5	*C. jussieui*	Chuquiragua
4	Ecuador	Riobamba	5	*C. jussieui*	Chuquiragua
5	Ecuador	Cuenca	7	*C. jussieui*	Chuquiragua
6	Ecuador	Loja	6	*C. jussieui*	Chuquiragua
7	Peru	Piura	14	*C. jussieui*	Chiquiragua
			1	*C. weberbaueri*	Amaro
			4 (Bags)	"Huamanpinta"	Huamanpinta
8	Peru	Huancabamba	7	*C. jussieui*	Chiquiragua
9	Peru	Chiclayo	5	*C. weberbaueri*	Amaro
			2 (Bags)	"Huamanpinta"	Huamanpinta
10	Peru	Cajamarca	9	*C. weberbaueri*	Amaro
			6 (Bags)	"Huamanpinta"	Huamanpinta
11	Peru	Trujillo	9	*C. weberbaueri*	Amaro
			3 (Bags)	"Huamanpinta"	Huamanpinta
12	Peru	Chimbote	5	*C. spinosa*	Huamanpinta
			5 (Bags)	"Huamanpinta"	Huamanpinta
13	Peru	Huaraz	7	*C. spinosa*	Huamanpinta
14	Peru	Huánuco	14	*C. spinosa*	Huamanpinta
15	Peru	Cerro de Pasco	15	*C. spinosa*	Huamanpinta
16	Peru	La Oroya	10	*C. spinosa*	Huamanpinta
17	Peru	Tarma	11	*C. spinosa*	Huamanpinta
18	Peru	Huancayo	23	*C. spinosa*	Huamanpinta
19	Peru	Lima	16	*C. spinosa*	Huamanpinta
			4 (Bags)	"Huamanpinta"	Huamanpinta
20	Peru	Huancavelica	11	*C. spinosa*	Huamanpinta
			2 (Bags)	"Huamanpinta"	Huamanpinta
21	Peru	Ayacucho	14	*C. spinosa*	Huamanpinta
22	Peru	Andahuaylas	6	*C. spinosa*	Intipasapra
			1 (Bags)	"Huamanpinta"	Huamanpinta
23	Peru	Cusco	-	*C. jussieui*	Quentai
			7 (Bags)	"Huamanpinta"	Huamanpinta
24	Peru	Combapata	2	*C. jussieui*	Quentai llaulli
25	Peru	Sicuani	9	*C. jussieui*	Inca llaulli
26	Peru	Juliaca	15	*C. jussieui*	Kiswara
27	Peru	Puno	4	*C. jussieui*	Kiswara
28	Peru	Ilave	7	*C. jussieui*	Kiswara
29	Peru	Arequipa	6	*C. jussieui*	Kiswara
			5 (Bags)	"Huamanpinta"	Huamanpinta
30	Peru	Moquegua	2	*C. jussieui*	Kiswara
31	Peru	Tacna	3	*C. jussieui*	Kiswara
32	Peru**	Piura	4	*C. jussieui*	Chiquiragua
33	Peru	Chiclayo	3	*C. weberbaueri*	Amaro
34	Peru	Huaraz	1	*C. spinosa*	Huamanpinta
35	Peru	Huancayo	1	*C. spinosa*	Huamanpinta
36	Austria***	-	1 (Capsules)	"Huamanpinta"	Huamanpinta
37	USA	-	1 (Powder)	"Huamanpinta"	Huamanpinta
38	USA	-	1 (Bag)	"Huamanpinta"	Huamanpinta
39	USA	-	1 (Bag)	"Huamanpinta"	Huamanpinta
40	USA	-	1 (Bag)	"Huamanpinta"	Huamanpinta
41	Peru	Lima	1 (Powder)	"Huamanpinta"	Huamanpinta
42	Peru	Lima	1 (Powder)	"Huamanpinta"	Huamanpinta
43	Peru	Lima	1 (Powder)	"Huamanpinta"	Huamanpinta

*Samples acquired in January 2015. **Samples acquired in January 2016.

*** Aboard samples (Austria, USA, Peru) acquired during 2015.

Figure 6. "Huamanpinta" samples: **A)** sample acquired in Lima, Peru; **B)** sample commercialized in Austria; **C)** sample commercialized in the USA; **D)** sample from a local Peruvian exporter.

3.2. Preparation of plant extracts

The plant samples (leaves) were powdered in an analytic mill (IKA A11 BASIC) and sieved (aperture of 0.355 mm and 42 mesh). Thirty milligrams (30 mg) of each sample were weighed in an analytical balance and transferred to Eppendorf tubes, where 3 mL of hydroalcoholic solvent (50% methanol) was added. The choice of the solvent was based on CASSADO et al. (2011) and preliminary tests. Then, the samples were submitted to ultrasonic bath for 15 min (USC 1600, 40 kHz ultrasonic frequency), and centrifuged for 5 min (New NT-810 Technique, 4000 RPM). The samples were filtered using modified cellulose syringe filters (Sartorius, 0.22 µm pore size) and conditioned for posterior chromatographic analysis.

3.4. LC-MS metabolic profile data acquisition and treatment.

Liquid chromatography analyses associated to high-resolution mass spectrometry (LC-MS) and ultraviolet detection were carried out in a setup of Thermo Scientific (USA), composed by two Accela 1250 quaternary pumps, coupled to an Accela diode array detector, and a Thermo Scientific Exactive Plus mass spectrometer equipped with Orbitrap ion trap mass analyzer synchronized by Xcalibur 2.2 software (Thermo Scientific). Water (A) and acetonitrile (B), both with 0.1% formic acid, were used as mobile phases. Chromatographic separations were conducted in a Kinetex XB-C18 (1.7 µm, 150 × 2.1 mm) column coupled to a guard column of the same characteristics, and an oven temperature of 35 ^{o}C. Four microliters of each sample were injected for analysis according to the following elution program: 2% B to

30% B in 30 min, then 100% B in 33 min, isocratic 100% until 35 min, and equilibration to 2% B until 40 min, with a mobile phase flow rate of 400 µL/min. The DAD scan wavelength was set from 190 to 600 nm.

In all analyses, the mass spectrometer operated under fast scan-to-scan polarity switching (one full positive mode scan and one full negative mode scan at a resolution setting of 70,000) in both MS and MS/MS modes with the following conditions: electrospray ionization mode (ESI), pulverization voltage of 3.6 kV in positive and 3.2 kV in negative mode, capillary temperature of 320 ºC, and a scan windows of 100 to 1500 *m/z*. Higher energy collisional dissociation (HCD) fragmentation were conducted for MS/MS in negative and positive modes. The reproductivity of the chromatographic analyses (retention times) and the accuracy of the mass spectrometer (error less than 5.0 ppm) was continually monitored by consecutive injection of a mixture of three standard compounds (5-*O*-caffeoylquinic acid, quercetin-3-*O*-rutinoside, and quercetin) before and during the analyses.

The obtained LC-UV-HRMS datasets in .RAW (Thermo Scientific) proprietary files were converted to nonproprietary .mzXML files by using the "msconvert" tool of ProteoWizard suite (version 3.0). The datasets were separated in negative and positive modes with an initial intensity cutoff at 1×10^4, including chromatographic data obtained until 30 min. These datasets were then submitted to processing tasks such as noise filtering, baseline correction, peak detection, deconvolution, deisotoping, data alignment etc., under MZmine (version 2.20) toolkit (parameters detailed in **APPENDIX**). The final dataset that included peak areas of mass patterns was transferred to a spreadsheet.

3.3. HPLC chromatographic profile data acquisition and processing

HPLC−DAD analyses were performed on a Shimadzu liquid chromatograph equipped with LC-10 Avp pumps, a SCL-10 Avp controller, a SPD-10 Avp diode array detector, and the Class VP software, version 6.2. Water (A) and acetonitrile (B), both with 1% acetic acid, were used as mobile phases. Chromatographic separations were conducted on a Zorbax SB-Phenyl (5 µm particle size, 4.6 × 250 mm) column as stationary phase. Twenty microliters of each sample were injected for analysis according to the following elution program: 4% B→30% B in 50 min (linear), then 100% B in 75 min (linear), and equilibration to 4% B until 85 min, with

a flow rate of 1.2 mL/min. The DAD scan wavelength was set from 190 to 600 nm with 0.64 s of scan rate. The reproductivity of the analyses were monitored by consecutive injection of a mixture of standard compounds and the *Chuquiraga* mixture sample before and during the analyses.

The HPLC chromatographic profiles were exported as .CSV files. The profiles at 270 nm were selected from each sample and organized in a spreadsheet and submitted to baseline correction and time shifts of peaks corrections by using the parametric time warping algorithm (PTW) implemented in the package "ptw" for R software (R version 3.30). The PTW algorithm aligns patterns by placing corresponding features at the same locations. Samples were aligned by calculating one global warping for all the set of samples and one reference chromatogram, *Chuquiraga* mixture samples.

3.5. Exploratory and supervised multivariate analyses

Initial exploratory analyses demonstrated no advantages of concatenating both dataset in negative and positive ionization modes when compared to the individual datasets. Additionally, no difference was observed between the datasets in negative and positive modes; however, the presence of a high mummer of mass patterns corresponding to fragments and adducts in the positive ionization mode dataset, favored the use of the negative ionization mode dataset for further multivariate analyses.

Hierarchical cluster analysis (HCA) and principal component analysis (PCA) were initially conducted on the datasets to determine general similarities. For these analyses, the datasets were normalized by the Pareto method. The HCA analysis was performed using the Euclidian distance and the Ward clustering method by using the "dendextend" and "circlize" packages for R software (version 3.30). The PCA analysis was performed using the NIPALS algorithm implemented in SIMCA P software (version 13.0). After the exploratory analyzes, a supervised analysis was carried out aiming to establish models for differentiation and prediction purposes. Orthogonal partial least squares discriminating analysis (OPLS-DA) was built in the SIMCA P software (version 13.0) using the Pareto normalized data. To compare OPLS-DA results, the *k*NN classifications were performed with the Euclidean distance by using the "class" and "gmodels" packages for R software (version 3.30).

Discriminating variables (metabolites) were identified after inspections of the OPLS-DA loading and score plots, considering a discriminating variable the one with a significant contribution in the loading plot, a variable importance in the projection value (VIP) greater than 1.0, and a reliable confidence interval in the coefficient plot. The selected discriminating metabolites were then represented in a heatmap constructed with the package "gplots" for R software (version 3.30).

3.6. Database construction

Chemical, spectroscopic and spectrometric data of secondary metabolites reported in members of Barnadesioideae were compiled in an electronic spreadsheet file with the support of the Marvin Suite and the JChem add-on for Excel (2016, ChemAxon Ltd., www.chemaxon.com) to build a database named AsterDB-Barnadesioiedae. This database is embedded in a larger project, the AsterDB, which contains thousands of chemical structures of secondary metabolites described in Asteraceae (www.asterbiochem.org/asterdb). The AsterDB-Barnadesioiedae.sdf file contains 294 entries of 39 2D chemical structures in .MOL file format together with their respective CAS numbers, common and scientific names, as well as their smiles, inchi,and.inchikey chemical file formats. It also contains essential spectrometric (molecular and monoisotopic mass) and spectroscopic (^{1}H- and ^{13}C-NMR) data of the secondary metabolites reported to date in Barnadesioideae (**APPENDIX**).

3.7. Metabolite identification

The process of metabolite identification was carried out considering the guidelines of the Metabolomics Standards Initiative (MSI) Working Group on Chemical Analysis (CAWG) that recommends four levels of identification (levels 1 to 4). *Level 1* identified metabolites (most accurate) were recognized by comparison of retention times (R_t), ultraviolet spectra (UV), accurate mass measurements (*m/z*), and fragmentation mass spectra (MS/MS) with those of commercial standards (Sigma-Aldrich) and isolated compounds available in the AsterDB pure compounds library of the AsterBioChem (www.asterbiochem.org) research group. For *level 2* identified metabolites, a punctual name was assigned based on spectroscopic (UV) and spectrometric data (exact mass and MS/MS spectra) and comparison with the in-house database AsterDB and other commercial databases, principally the Dictionary of Natural Products (DNP) and SciFinder Scholar. At *level 3*, it was not possible to give a punctual name, but rather

a general compound class after data analysis as described for *Level 2* identification. Finally, at *level 4* mass features with unknown identity were assigned, or for which the obtained data were not conclusive, or several hits were obtained.

The corresponding molecular formulas of the annotated metabolites were determined by the molecular finder tool of Xcalibur 2.20 (Thermo Scientific) with default settings and the ChemCalc web service (www.chemcalc.org) where elemental compositions were determined from exact monoisotopic mass data within a 5 ppm error window and the following elemental search parameters: $C_{(0-100)}$, $H_{(0-100)}$, $O_{(0-100)}$, $N_{(0-5)}$

3.8. Isolation and structure elucidation

A commercial sample of *C. spinosa* (500 g), purchased in Lima, Peru, in January 2015, was extracted with 50% methanol (3 times, 1:5 plant/solvent ratio - g/mL). This extract was dried and lyophilized to yield 110 g of a crude extract. A portion of this extract (5 g) was suspended in water (20 mL) and partitioned with *n*-hexane and dichloromethane, the organic fractions were discharged, and the aqueous fraction was filtered and subsequently subjected to column chromatography (100 g, 40 × 400 mm) using Sephadex LH-20 as stationary phase and mixtures with decreasing polarity of water-ethanol (100:0→0:100) as mobile phase, yielding 11 fractions. These fractions were then processed by semi-preparative HPLC on a C18 Shim-Pack (15 µm, 25 × 250 mm) column by using a linear gradient of water-acetonitrile (95:5→70:30 in 50 min with a flow rate of 10 mL/min), both acidified with 1% acetic acid as mobile phase, to afford the *p*-hydroxyacetophenone-4-*O*-glucoside, *p*-hydroxyacetophenone-4-*O*-(6-*O*-apiosyl)-glucoside, 5-*O*-caffeoylquinic acid, *p*-hydroxyacetophenone, quercetin-3-*O*-glucoronide, kaempherol-3-*O*-rutoside and a new caffeic acid ester derivative. To isolate further minor compounds, 100 g of the crude extract was suspended in 130 mL of water and partitioned with *n*-hexane and dichloromethane, the organic fractions were discharged and the aqueous fraction was filtered and subsequently subjected to column chromatography (250 g, 50 × 500 mm) using Sephadex LH-20 as stationary phase and mixtures with decreasing polarity of water-ethanol (100:0→0:100) as mobile phase, yielding four fractions: F1 (5 g), F2 (14 g), F3 (9 g) and F4 (5 g). Fraction F2 was resubmitted to column chromatography under the same conditions to afford four subfractions: F2A (7.7 g), F2B (2.8 g), F2C (0.7 g) and F2D (0.3 mg). Subfraction F2B was submitted to semi-preparative HPLC to afford additional knew ester derivatives of caffeic, coumaric and ferulic acids.

4. RESULTS AND DISCUSSION

4.1. Chemistry of the subfamily Barnadesioideae

The current published chemical diversity of the subfamily Barnadesioideae was limited to 39 constituents among acetophenones, flavonoids and triterpenoids (see **Table 2** and **Figure 2**). Compared to other taxa of Asteraceae, the secondary metabolite chemistry of Barnadesioideae has been sometimes described as following a simple profile. For example, regarding flavonoid chemistry, BOHM AND STUESSY (1995) declared: "no other large taxa within the Asteraceae have such a consistently simple pigment profile".

Contrarily to these reports, the LC-MS metabolome data of Barnadesioideae displayed a variety of mass features. The **Figure 7** and **Figure 8** displays typical base peak ion chromatograms (BPC) from representatives of *Arnaldoa*, *Barnadesia*, *Chuquiraga*, and *Dasyphyllum* genera. The main mass features were summarized and are listed in **Table 11** that contains retention times (R_t), UV maxima data, experimental and a calculated high-resolution mass data of 196 constituents distributed among Barnadesioideae members.

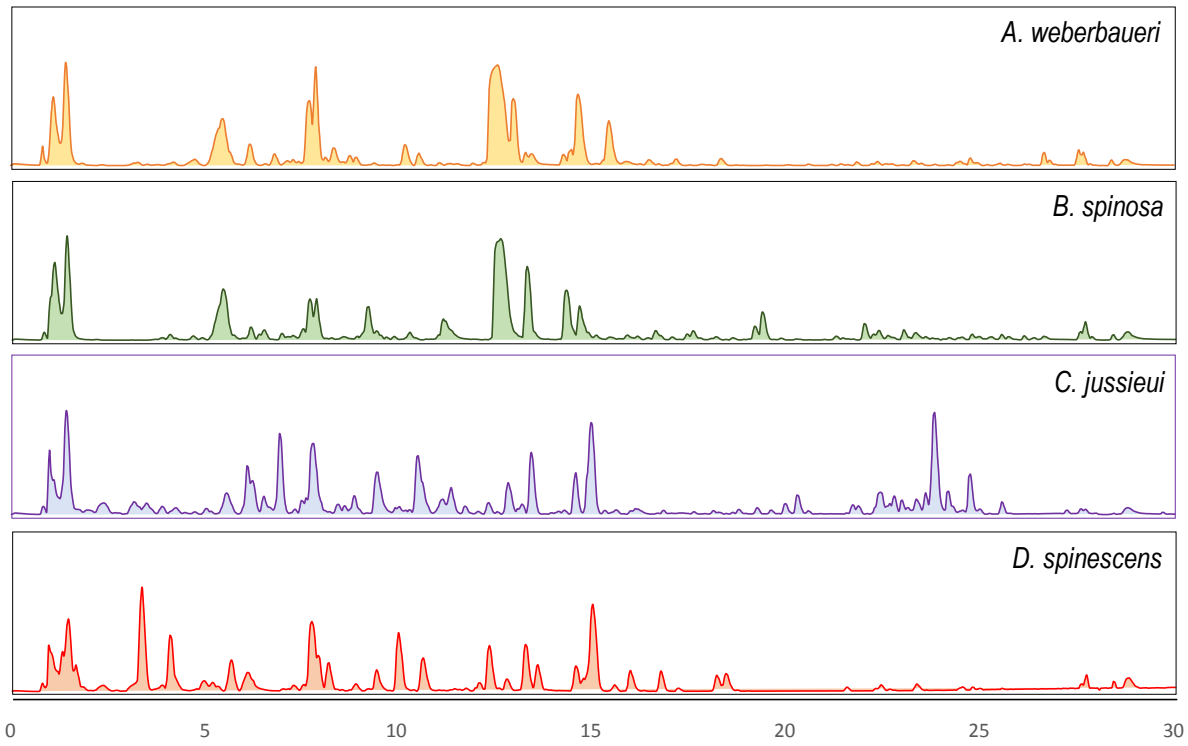


Figure 7. Base peak ion cromatogramas (BPC) in negative mode of Barnadesioideae representatives.

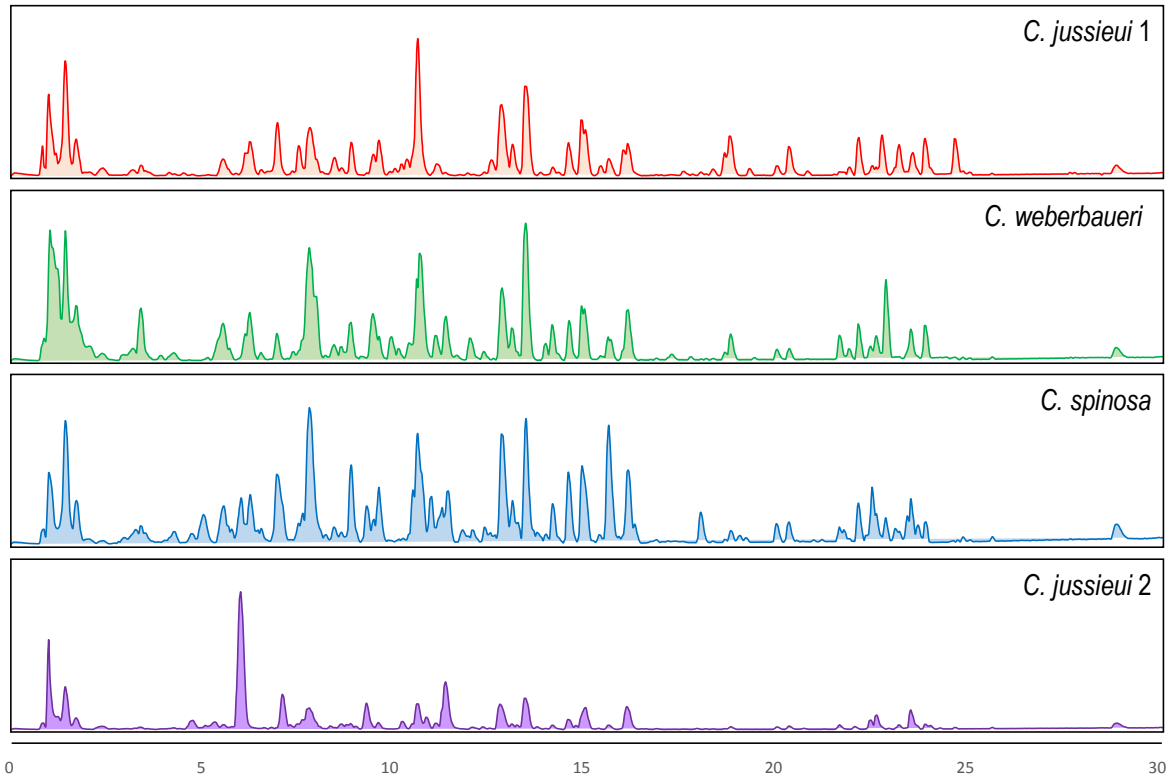


Figure 8. Base peak ion chromatograms (BPC) in negative mode of *Chuquiraga* representatives.

Twenty-four constituents were identified in *level 1* classification by direct comparison of R_t, UV and *m/z* data with those of available standard compounds: gallic acid (**40**), protocatecuic acid (**46**), 2-*O*-caffeoyltartaric acid (**54**), 3-*O*-caffeoylquinic acid (**56**), 5-*O*-(4-*O*-glucosylcaffeoyl)quinic acid (**57**), *p*-hydroxyacetophenone 4-*O*-glucoside (**58**), *p*-hydroxyacetophenone 4-*O*-(6-*O*-apiofuranosyl)-glucopyranoside (**67**), 5-*O*-caffeoylquinic acid (**74**), caffeic acid (**76**), 4-*O*-caffeoylquinic acid (**77**), *p*-hydroxyacetophenone (**86**), 2-*O*-caffeoylmalic acid (**103**), *p*-coumaric acid (**105**), 2-*O*-feruloyltartaric acid (**114**), 5-*O*-feruloylquinic acid (**115**), 2,3-*O*-dicaffeoyltartaric acid (**136**), quercetin-3-*O*-rutinoside (**138**), quercetin-3-*O*-glucuronide (**139**), quercetin-3-*O*-glucoside (**140**), kaempferol-3-*O*-rutinoside (**150**), kaempferol-3-*O*-glucuronide (**155**), 3,4-*O*-dicaffeoylquinic acid (**152**), 3,5-*O*-dicaffeoylquinic acid (**154**), and 4,5-O-dicaffeoylquinic acid (**163**). The **Figure 9** displays the chemical structure of representative constituents identified in Barnadesioideae subfamily.

Additionally, the 2-*O*-caffeoyl-4-hydroxypentanedioic acid (**75**), 2-*O*-coumaroyl-4-hydroxypentanedioic acid (**100**), 2-*O*-feruloyl-4-hydroxypentanedioic acid (**117**) and 2-*O*-caffeoyl-4-methylbutanedioic acid (**122**) correspond to new compounds isolated and identified for the first time in the genus *Chuquiraga*.

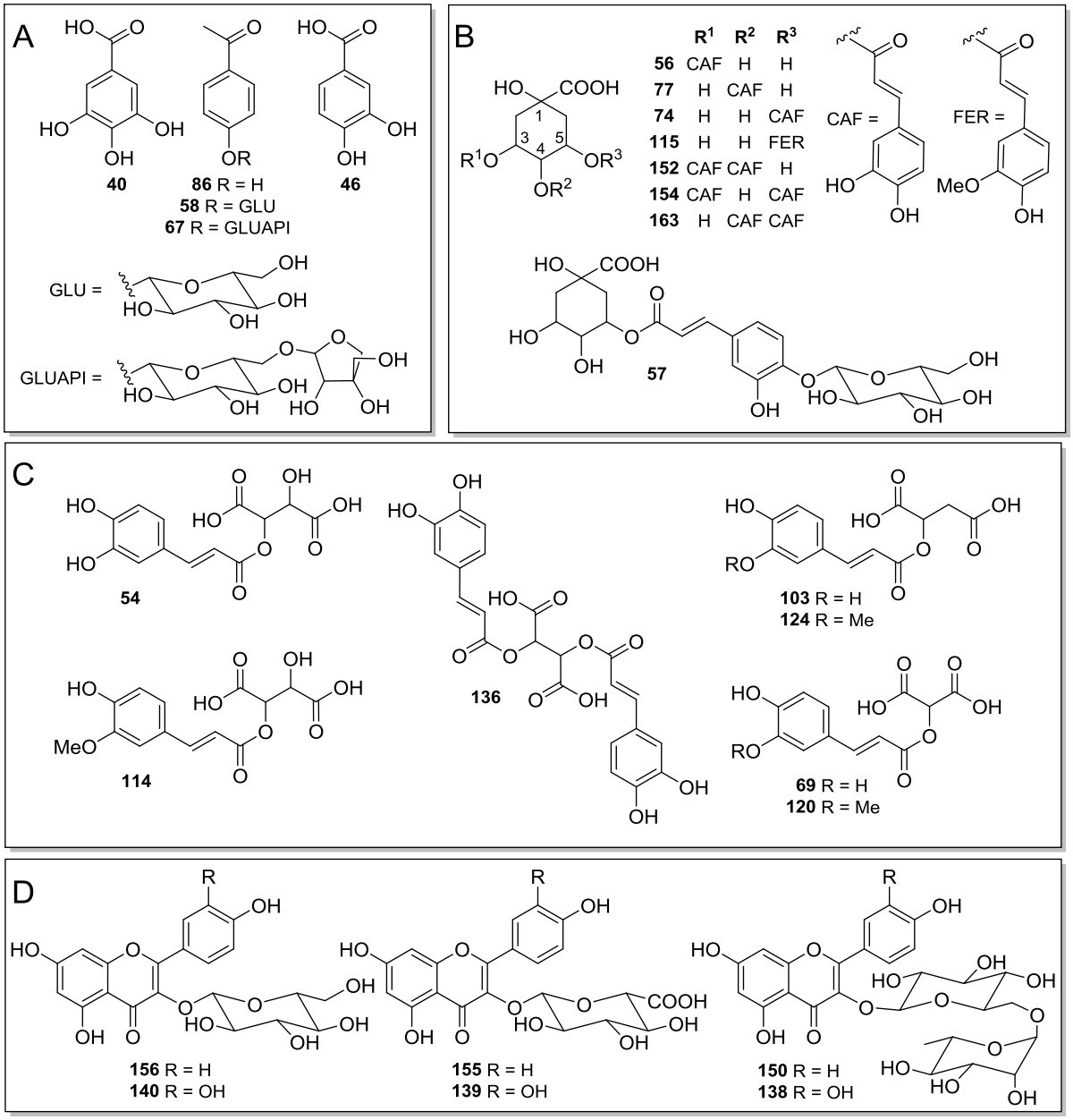


Figure 9. Representative constituents identified in Barnadesioideae subfamily. **A**) phenolic compounds and acetophenone derivatives. **B**) Caffeic and ferulic acid ester derivatives of quinic acid. **C**) Caffeic and ferulic acid ester derivatives of tartaric, malic and tartronic acids. **D**) Flavonoid glycosides of quercetin and kaempferol.

Other compounds identified in *level 3* classification, correspond to mass patterns with no UV absorption and mass values between *m*/*z* 307 and 501 (until ~11 min of the BPC chromatograms) that were identified as alkyl glycosides, principally distributed in the genus *Chuquiraga*, whereas mass patterns with no UV absorption and mass values in the range of *m*/*z* 523 to 969 were identified as triterpenoid glycosides (from 20 to 25 min of the BPC chromatograms), distributed along Barnadesioideae members

Table 11. Main metabolites in members of the subfamily Barnadesioideae.

No.	Rt (min)	UV (nm)	Molecular formula	*m/z* [M-H]$^-$			Identity	Level*
				Experimental	Calculated	Δ (ppm)		
40	2.36	l/c	$C_7H_6O_5$	169.0134	169.0137	-1.78	Gallic acid	1
41	2.41	n/d	$C_8H_{10}O_7$	217.0351	217.0348	1.38	Hydroxycitric acid lactone	2
42	3.16	n/d	$C_8H_{10}O_6$	201.0400	201.0399	0.50	Erythronolactone diacetate	2
43	3.32	n/d	$C_5H_8O_4$	131.0339	131.0344	-3.82	Methylerythronolactone	2
44	3.76	n/d	$C_{12}H_{14}O_9$	301.0564	301.0559	1.66	Glucuronolactone derivative	3
45	4.02	n/d	$C_{12}H_{12}O_{10}$	315.0356	315.0352	1.27	n/i	4
46	4.12	260, 293	$C_7H_6O_4$	153.0185	153.0188	-1.96	Protocatecuic acid	1
47	4.23	n/d	$C_{12}H_{12}O_8$	283.0461	283.0454	2.47	n/i	4
48	4.26	n/d	$C_8H_{10}O_6$	201.0399	201.0399	0.00	Erythronolactone diacetate	2
49	4.75	263	$C_{15}H_{22}O_9$	345.1193	345.1185	2.32	Phenolic glycoside	3
50	4.94	n/d	$C_{11}H_{14}O_8$	273.0615	273.0610	1.83	Pyranone glycoside	3
51	5.02	n/d	$C_{10}H_{20}O_{15}$	379.0709	379.0724	-3.96	n/i	4
52	5.21	n/d	$C_{12}H_{14}O_9$	301.0568	301.0559	2.99	Glucuronolactone derivative	3
53	5.27	n/d	$C_{19}H_{22}O_{11}$	425.1094	425.1084	2.35	n/i	4
54	5.49	249, 295 shd, 328	$C_{13}H_{12}O_9$	311.0402	311.0403	-0.32	2-*O*-Caffeoyltartaric acid	1
55	5.60	n/d	$C_{14}H_{28}O_8$	323.1713	323.1706	2.17	Alkyl glycoside	3
			$C_{14}H_{28}O_8$+CH_2O_2	369.1764	369.1760	1.08	Adduct	
56	5.63	248, 296 shd, 324	$C_{16}H_{18}O_9$	353.0873	353.0872	0.28	3-*O*-Caffeoylquinic acid	1
57	5.91	240, 292, 316	$C_{22}H_{28}O_{14}$	515.1404	515.1401	0.58	5-*O*-(4-*O*-Glucosylcaffeoyl)quinic acid	1
58	5.94	263	$C_{14}H_{18}O_7$	297.0979	297.0974	1.68	*p*-Hydroxyacetophenone 4-*O*-glucoside	1
			$C_{14}H_{18}O_7$+CH_2O_2	343.1028	343.1029	-0.29	Adduct	
59	6.01	n/d	$C_{12}H_{14}O_9$	301.0569	301.0559	3.32	Glucuronolactone derivative	3
60	6.08	n/d	$C_{14}H_{28}O_8$	323.1713	323.1706	2.17	Alkyl glycoside	3
			$C_{14}H_{28}O_8$+CH_2O_2	369.1764	369.1760	1.08	Adduct	
61	6.10	l/c	$C_{20}H_{26}O_{13}$	473.1304	473.1295	1.90	Phenolic glycoside	3
62	6.18	n/d	$C_{13}H_{24}O_9$	323.1353	323.1342	3.40	Disaccharide	3
63	6.24	n/d	$C_7H_{12}O_5$	175.0604	175.0606	-1.14	Isopropylmalic acid	2
64	6.38	n/d	$C_{19}H_{36}O_{12}$	455.2136	455.2128	1.76	Alkyl glycoside	3
			$C_{19}H_{36}O_{12}$+CH_2O_2	501.2188	501.2183	1.00	Adduct	
65	6.81	n/d	$C_{19}H_{36}O_{12}$	455.2137	455.2128	1.98	Alkyl glycoside	3
			$C_{19}H_{36}O_{12}$+CH_2O_2	501.2181	501.2183	-0.40	Adduct	
66	6.96	n/d	$C_{12}H_{22}O_8$	293.1243	293.1236	2.39	Disaccharide	3

67	6.99	215, 264	$C_{19}H_{26}O_{11}$	429.1403	429.1397	1.40	*p*-Hydroxyacetophenone 4-*O*-(6-*O*-apiofuranosyl)-glucopyranoside	1
			$C_{19}H_{26}O_{11}+CH_2O_2$	475.1461	475.1451	2.10	Adduct	
68	7.26	l/c	$C_{22}H_{28}O_{14}$	515.1415	515.1401	2.72	Glucosylcaffeoylquinic acid	2
69	7.26	249, 295 shd, 327	$C_{12}H_{10}O_8$	281.0299	281.0297	0.71	Caffeoyltartronic acid	2
70	7.31	n/d	$C_{19}H_{36}O_{12}$	455.2134	455.2128	1.32	Alkyl glycoside	3
			$C_{19}H_{36}O_{12}+CH_2O_2$	501.2189	501.2183	1.20	Adduct	
71	7.32	240, 292, 316	$C_{22}H_{28}O_{14}$	515.1411	515.1401	1.94	Glucosylcaffeoylquinic acid	3
72	7.50	n/d	$C_{19}H_{36}O_{12}$	455.2134	455.2128	1.32	Alkyl glycoside	3
			$C_{19}H_{36}O_{12}+CH_2O_2$	501.2189	501.2183	1.20	Adduct	
73	7.60	250, 298 shd, 325	$C_{13}H_{12}O_9$	311.0413	311.0403	3.22	Caffeoyltartaric acid	2
74	7.69	242, 298 shd, 324	$C_{16}H_{18}O_9$	353.0874	353.0872	0.57	5-*O*-Caffeoylquinic acid	1
75	7.78	243, 297 shd, 326	$C_{14}H_{14}O_9$	325.0571	325.0559	3.69	2-*O*-Caffeoyl-4-hydroxypentanedioic acid	1
76	7.91	244, 299 shd, 323	$C_9H_8O_4$	179.0341	179.0344	-1.68	Caffeic acid	1
77	8.22	241, 297 shd, 325	$C_{16}H_{18}O_9$	353.0875	353.0872	0.85	4-*O*-Caffeoylquinic acid	1
78	8.32	l/c	$C_{14}H_{14}O_9$	325.0570	325.0559	3.38	Caffeic acid ester derivative	3
79	8.61	n/d	$C_{19}H_{34}O_{14}$	485.1887	485.1870	3.50	Trisaccharide	3
80	8.63	n/d	$C_{18}H_{34}O_{11}$	425.2037	425.2023	3.29	Alkyl glycoside	3
			$C_{18}H_{34}O_{11}+CH_2O_2$	471.2087	471.2072	3.18	Adduct	
81	8.76	n/d	$C_{19}H_{28}O_{12}$	447.1504	447.1502	0.45	n/i	4
82	8.91	l/c	$C_{12}H_{14}O_6$	253.0716	253.0712	1.58	Caffeoylglycerol	2
83	8.93	n/d	$C_{19}H_{28}O_{12}$	401.1452	401.1447	1.25	n/i	4
84	9.05	l/c	$C_{12}H_{10}O_7$	265.0351	265.0348	1.13	Coumaroyltartronic acid	2
85	9.15	312	$C_{13}H_{12}O_8$	295.0461	295.0454	2.37	Coumaroyltataric acid	2
86	9.21	220, 275	$C_8H_8O_2$	135.0441	135.0446	-3.70	*p*-Hydroxyacetophenone	1
87	9.24	l/c	$C_{11}H_{10}O_7$	253.0354	253.0348	2.37	Benzoyltartaric acid	2
88	9.25	n/d	$C_{20}H_{36}O_{13}$	483.2085	483.2077	1.66	Alkyl glycoside	3
89	9.42	247 298 shd, 326	$C_{16}H_{16}O_8$	335.0775	335.0767	2.39	Caffeoylshikimic acid	2
90	9.43	l/c	$C_{11}H_{10}O_7$	253.0354	253.0348	2.37	Benzoyltartaric acid	2
100	9.46	l/c	$C_{14}H_{14}O_8$	309.0618	309.0610	2.59	2-*O*-Coumaroyl-4-hydroxypentanedioic acid	1
101	9.48	312	$C_{16}H_{18}O_8$	337.0927	337.0923	1.19	Coumaroylquinic acid	2
102	9.55	n/d	$C_{13}H_{24}O_8$	307.1398	307.1393	1.63	Alkyl glycoside	3
	9.64	l/c	$C_{25}H_{40}O_{14}$	563.2357	563.2340	3.02	n/i	4
103	9.85	245, 297 shd, 328	$C_{13}H_{12}O_8$	295.0461	295.0454	2.37	2-*O*-Caffeoylmalic acid	1
104	10.03	248, 297 shd, 325	$C_{16}H_{16}O_8$	335.0776	335.0767	2.69	Caffeoylshikimic acid	2
105	10.04	l/c	$C_9H_8O_3$	163.0391	163.0395	-2.45	*p*-Coumaric acid	1

106	10.11	n/d	$C_{14}H_{26}O_8$	321.1558	321.1549	2.80	Alkyl glycoside	3
107	10.16	n/d	$_{15}H_{30}O_9$	353.1822	353.1811	3.09	n/i	4
108	10.16	248, 302, 322	$C_{25}H_{24}O_{12}$	515.1199	515.1189	1.94	Caffeic acid ester derivative	3
109	10.20	268	$C_{20}H_{24}O_{11}$	439.1248	439.1240	1.82	n/i	4
110	10.25	n/d	$C_{20}H_{34}O_{10}$	433.2081	433.2073	1.85	n/i	4
111	10.28	216, 271, 335	$C_{27}H_{30}O_{15}$	593.1516	593.1506	1.69	Flavonoid glycoside	3
112	10.31	n/d	$C_9H_{18}O_{14}$	349.0601	349.0618	-4.87	n/i	4
113	10.36	l/c	$C_{12}H_{12}O_7$	267.0512	267.0505	2.62	Caffeoylglyceric acid	2
114	10.48	236, 296 shd, 327	$C_{14}H_{14}O_9$	325.0558	325.0559	-0.31	2-*O*-Feruloyltartaric acid	1
115	10.53	303 shd, 326	$C_{17}H_{20}O_9$	367.1042	367.1029	3.54	5-*O*-Feruloylquinic acid	1
116	10.59	n/d	$C_{19}H_{36}O_{11}$	439.2187	439.2179	1.82	Alkyl glycoside	3
			$C_{19}H_{36}O_{11}+CH_2O_2$	485.2242	485.2234	1.65	Adduct	
117	10.59	234, 295 shd, 326	$C_{15}H_{16}O_9$	339.0720	339.0716	1.18	2-*O*-Feruloyl-4-hydroxypentanedioic acid	1
118	10.88	247, 297 shd, 327	$C_{14}H_{14}O_8$	309.0618	309.0610	2.59	Caffeic acid ester derivative	2
119	10.90	l/c	$C_{16}H_{18}O_8$	337.0927	337.0923	1.19	Coumaroylquinic acid	2
120	10.99	245, 297 shd, 328	$C_{13}H_{12}O_8$	295.0461	295.0454	2.37	Feruloyltartronic acid	2
121	11.03	l/c	$C_{14}H_{14}O_9$	325.0563	325.0559	1.23	Feruoyltartaric acid	2
122	11.05	247, 297 shd, 327	$C_{14}H_{14}O_8$	309.0618	309.0610	2.59	2-*O*-Caffeoyl-4-methylbutanedioic acid	1
123	11.07	245, 296 shd, 327	$C_{11}H_{12}O_4$	207.0653	207.0657	-1.93	Caffeic acid ethyl ester	2
124	11.14	l/c	$C_{14}H_{14}O_8$	309.0618	309.0610	2.59	Feruloylmalic acid	2
125	11.19	244, 295 shd, 327	$C_{17}H_{20}O_9$	367.1044	367.1029	4.09	Feruloylquinic acid	2
126	11.20	l/c	$C_{14}H_{14}O_9$	325.0563	325.0559	1.23	Feruloyltartaric acid	2
127	11.32	242, 296 shd, 326	$C_{15}H_{16}O_9$	339.0721	339.0716	1.47	Ferulic acid ester derivative	1
128	11.35	l/c	$C_{14}H_{14}O_7$	293.0668	293.0661	2.39	Coumaroylmalic acid methyl ester	2
129	11.66	n/d	$C_{22}H_{38}O_{13}$	509.2238	509.2234	0.79	Monoterpene diglycoside	3
130	11.88	n/d	$C_{22}H_{32}O_{11}$	471.1872	471.1866	1.27	Acetylenic diglycoside	3
131	11.90	l/c	$C_{23}H_{20}O_{13}$	503.0836	503.0825	2.19	Flavonoid glycoside	3
132	12.30	l/c	$C_{16}H_{16}O_{10}$	367.0675	367.0665	2.72	Feruloylisocitric acid	2
133	12.33	l/c	$C_{15}H_{16}O_8$	323.0771	323.0767	1.24	Feruloylmalic acid methyl ester	2
134	12.36	n/d	$C_{22}H_{38}O_{13}$	509.2245	509.2234	2.16	Monoterpene diglycoside	3
135	12.41	l/c	$C_{21}H_{18}O_{11}$	445.0768	445.0771	-0.67	Flavonoid glycoside	3
136	12.56	246, 297 shd, 328	$C_{22}H_{18}O_{12}$	473.0726	473.0720	1.27	2,3-*O*-Dicaffeoyltartaric acid	1
137	12.79	210, 256, 353	$C_{27}H_{30}O_{16}$	609.1477	609.1455	3.61	Quercetin diglycoside	2
138	13.05	204, 256, 348	$C_{27}H_{30}O_{16}$	609.1477	609.1455	3.61	Quercetin-3-*O*-rutinoside	1
139	13.18	206, 255, 353	$C_{21}H_{18}O_{13}$	477.0678	477.0669	1.89	Quercetin-3-*O*-glucuronide	1

140	13.24	205, 256, 348	$C_{21}H_{20}O_{12}$	463.0880	463.0876	0.86	Quercetin-3-*O*-glucoside	1
141	13.26	246, 296 shd, 329	$C_{16}H_{18}O_{9}$	353.0878	353.0872	1.70	Caffeic acid ester derivative	3
142	13.48	247, 298 shd, 328	$C_{22}H_{18}O_{12}$	473.0725	473.0720	1.06	Dicaffeoyltartaric acid	2
143	13.56	l/c	$C_{28}H_{30}O_{17}$	637.1422	637.1404	2.83	Flavonoid glycoside	3
144	13.58	l/c	$C_{20}H_{20}O_{8}$	387.1087	387.1080	1.81	n/i	4
145	13.83	265, 310 shd, 346	$C_{27}H_{30}O_{15}$	593.1514	593.1506	1.35	Kaempferol diglycoside	2
146	14.08	n/d	$C_{18}H_{30}O_{10}$	405.1768	405.1760	1.97	Monoterpene diglycoside	3
147	14.26	297 shd, 329	$C_{15}H_{16}O_{8}$	323.0774	323.0767	2.17	Feruloylmalic acid methyl ester	2
148	14.30	203, 256, 347	$C_{24}H_{22}O_{15}$	549.0886	549.0880	1.09	Flavonoid glycoside	3
149	14.40	246, 296 shd, 329	$C_{22}H_{18}O_{11}$	457.0770	457.0771	-0.22	Coumaroylcaffeoyltartaric acid	2
150	14.51	265, 305 shd, 342	$C_{27}H_{30}O_{15}$	593.1514	593.1506	1.35	Kaempferol-3-*O*-rutinoside	1
151	14.55	l/c	$C_{10}H_{10}O_{4}$	193.0497	193.0501	-2.07	n/i	4
152	14.55	245, 296 shd, 328	$C_{23}H_{20}O_{12}$	515.1201	515.1189	2.33	3,4-*O*-Dicaffeoylquinic acid	1
153	14.85	254, 344	$C_{24}H_{22}O_{15}$	549.0889	549.0880	1.64	Flavonoid glycoside	3
154	14.86	247, 296 shd, 327	$C_{25}H_{24}O_{12}$	515.1182	515.1189	-1.36	3,5-*O*-Dicaffeoylquinic acid	1
155	14.92	209, 265, 343	$C_{21}H_{18}O_{12}$	461.0737	461.0720	3.69	Kaempferol-3-*O*-glucuronide	1
156	14.96	206, 256, 346	$C_{21}H_{20}O_{11}$	447.0941	447.0927	3.13	Kaempferol-3-*O*-glucoside	2
157	14.95	l/c	$C_{28}H_{32}O_{16}$	623.1624	623.1612	1.94	Flavonoid glycoside	3
158	15.36	l/c	$C_{22}H_{22}O_{12}$	477.1042	477.1033	1.89	n/i	4
159	15.54	n/d	$C_{18}H_{30}O_{10}$	405.1769	405.1760	2.22	Monoterpene diglycoside	3
160	15.55	247, 297 shd, 328	$C_{23}H_{20}O_{12}$	487.0877	487.0876	0.21	Caffeoyl-feruloyltartaric acid	2
161	15.62	l/c	$C_{22}H_{20}O_{13}$	491.0833	491.0825	1.63	Flavonoid glycoside	3
162	15.87	245, 298 shd, 329	$C_{22}H_{18}O_{12}$	473.0725	473.0720	1.06	Dicaffeoyltartaric acid	2
163	15.92	248, 295 shd, 325	$C_{25}H_{24}O_{12}$	515.1198	515.1189	1.75	4,5-*O*-Dicaffeoylquinic acid	1
164	16.04	245, 298 shd, 329	$C_{22}H_{18}O_{12}$	473.0725	473.0720	1.06	Dicaffeoyltartaric acid	2
165	16.23	n/d	$C_{27}H_{48}O_{14}$	595.2981	595.2966	2.52	Sesquiterpene glycoside	3
166	16.58	l/c	$C_{22}H_{18}O_{12}$	473.0725	473.0720	1.06	Dicaffeoyltartaric acid	2
167	16.74	208, 264, 341	$C_{21}H_{20}O_{10}$	431.0991	431.0978	3.02	Kaempferol glycoside	2
168	17.04	l/c	$C_{26}H_{26}O_{12}$	529.1353	529.1346	1.32	Caffeoyl-feruloylquinic acid	2
169	17.38	n/d	$C_{22}H_{24}O_{9}$	431.1348	431.1342	1.39	Sesquiterpene lactone	3
170	17.50	n/d	$C_{17}H_{30}O_{12}$	425.1646	425.1659	-3.06	Alkyl glycoside	3
171	17.54	245, 296 shd, 328	$C_{26}H_{26}O_{12}$	529.1351	529.1346	0.94	Caffeoyl-feruloylquinic acid	2
172	17.71	295 shd, 329	$C_{24}H_{22}O_{12}$	501.1039	501.1033	1.20	Diferuloyltartaric acid	2
173	17.95	l/c	$C_{24}H_{22}O_{12}$	501.1038	501.1033	1.00	Diferuloyltartaric acid	2
174	18.20	l/c	$C_{24}H_{22}O_{12}$	501.1038	501.1033	1.00	Diferuloyltartaric acid	2

175	18.21	296, 329	$C_{25}H_{22}O_{11}$	497.1093	497.1084	1.81	Caffeic acid ester derivative	3
176	18.29	l/c	$C_{30}H_{26}O_{14}$	609.1255	609.1244	1.81	Acylated flavonoid glycoside	3
177	18.42	296, 330	$C_{25}H_{22}O_{11}$	497.1086	497.1084	0.40	Caffeic acid ester derivative	3
178	18.49	l/c	$C_{23}H_{20}O_{12}$	487.0900	487.0876	4.93	Caffeoyl-feruloyltartaric acid	2
179	18.57	l/c	$C_{31}H_{28}O_{15}$	639.1371	639.1350	3.29	Acylated flavonoid glycoside	3
180	18.69	l/c	$C_{23}H_{20}O_{12}$	487.0879	487.0876	0.62	Caffeoyl-feruloyltartaric acid	2
181	18.73	l/c	$C_{23}H_{20}O_{12}$	487.0879	487.0876	0.62	Caffeoyl-feruloyltartaric acid	2
182	18.97	n/d	$C_{22}H_{38}O_{12}$	493.2297	493.2285	2.43	Alkyl glycoside	3
183	19.13	l/c	$C_{30}H_{26}O_{15}$	625.1213	625.1193	3.20	Acylated flavonoid glycoside	3
184	19.22	n/d	$C_{33}H_{54}O_{14}$	673.3445	673.3435	1.49	Acylglycerol diglycoside	3
185	19.33	l/c	$C_{33}H_{28}O_{18}$	711.1202	711.1197	0.70	Benzophenone glycoside	3
186	19.34	n/d	$C_{42}H_{66}O_{20}$	889.4094	889.4069	2.77	Alkyl iridoid glycoside	3
187	19.59	n/d	$C_{33}H_{54}O_{14}$	673.3447	673.3435	1.78	Acylglycerol diglycoside	3
188	19.76	266, 315	$C_{30}H_{26}O_{13}$	593.1293	593.1295	-0.34	Acylated flavonoid glycoside	3
189	19.82	n/d	$C_{21}H_{28}O_{9}$	423.1671	423.1655	3.78	n/i	3
190	19.91	266, 315	$C_{30}H_{26}O_{13}$	593.1293	593.1295	-0.34	Acylated flavonoid glycoside	3
191	20.24	266, 326	$C_{31}H_{28}O_{14}$	623.1394	623.1401	-1.12	Acylated flavonoid glycoside	3
192	20.29	l/c	$C_{30}H_{26}O_{13}$	593.1306	593.1295	1.85	Acylated flavonoid glycoside	3
193	20.44	l/c	$C_{30}H_{26}O_{13}$	593.1306	593.1295	1.85	Acylated flavonoid glycoside	3
194	20.52	n/d	$C_{33}H_{54}O_{14}$	673.3443	673.3435	1.19	Acylglycerol diglycoside	3
195	20.72	l/c	$C_{32}H_{30}O_{15}$	653.1522	653.1506	2.45	Acylated flavonoid glycoside	3
196	21.54	n/d	$C_{36}H_{56}O_{15}$	727.3549	727.3541	1.10	Triterpenoid glycoside	3
197	21.54	n/d	$C_{22}H_{30}O_{9}$	437.1823	437.1811	2.74	n/i	3
198	21.79	n/d	$C_{33}H_{50}O_{14}$	669.3138	669.3122	2.45	Triterpenoid glycoside	3
199	21.82	n/d	$C_{36}H_{56}O_{14}$	711.3608	711.3592	2.25	Triterpenoid glycoside	3
200	21.95	n/d	$C_{22}H_{30}O_{9}$	437.1823	437.1811	2.74	n/i	3
201	21.96	n/d	$C_{41}H_{62}O_{18}$	841.3871	841.3858	1.55	Triterpenoid glycoside	3
202	22.03	n/d	$C_{36}H_{56}O_{15}$	727.3555	727.3541	1.92	Triterpenoid glycoside	3
203	22.06	n/d	$C_{38}H_{58}O_{16}$	769.3649	769.3646	0.39	Triterpenoid glycoside	3
204	22.23	n/d	$C_{41}H_{64}O_{18}$	843.4051	843.4014	4.39	Triterpenoid glycoside	3
205	22.37	n/d	$C_{18}H_{32}O_{5}$	327.2182	327.2171	3.36	Fatty acid	3
206	22.50	n/d	$C_{38}H_{58}O_{16}$	769.3663	769.3646	2.21	Triterpenoid glycoside	3
207	22.55	n/d	$C_{36}H_{56}O_{14}$	711.3625	711.3592	4.64	Triterpenoid glycoside	3
208	22.64	n/d	$C_{38}H_{56}O_{16}$	767.3505	767.3490	1.95	Triterpenoid glycoside	3
209	22.66	n/d	$C_{36}H_{56}O_{14}$	711.3605	711.3592	1.83	Triterpenoid glycoside	3

210	22.70	n/d	$C_{33}H_{52}O_{13}$	655.3348	655.3329	2.90	Triterpenoid glycoside	3
211	22.75	n/d	$C_{28}H_{42}O_{10}$	537.2711	537.2699	2.23	Triterpenoid glycoside	3
212	22.92	n/d	$C_{38}H_{58}O_{16}$	769.3649	769.3646	0.39	Triterpenoid glycoside	3
213	22.94	n/d	$C_{41}H_{64}O_{18}$	843.4034	843.4014	2.37	Triterpenoid glycoside	3
214	22.96	297 shd, 328	$C_{25}H_{32}O_{12}$	523.1820	523.1815	0.96	Caffeic acid ester derivative	3
215	23.00	n/d	$C_{38}H_{58}O_{16}$	769.3636	769.3646	-1.30	Triterpenoid glycoside	3
216	23.10	n/d	$C_{40}H_{60}O_{18}$	827.3719	827.3701	2.18	Triterpenoid glycoside	3
217	23.32	n/d	$C_{18}H_{34}O_{5}$	329.2332	329.2328	1.21	Fatty acid	3
218	23.35	n/d	$C_{35}H_{54}O_{13}$	681.3503	681.3486	2.50	Triterpenoid glycoside	3
219	23.40	n/d	$C_{38}H_{58}O_{16}$	769.3660	769.3646	1.82	Triterpenoid glycoside	3
220	23.41	n/d	$C_{33}H_{52}O_{13}$	655.3341	655.3329	1.83	Triterpenoid glycoside	3
221	23.46	n/d	$C_{46}H_{74}O_{20}$	945.4697	945.4695	0.21	Triterpenoid glycoside	3
222	23.46	n/d	$C_{45}H_{68}O_{21}$	943.4188	943.4175	1.38	Triterpenoid glycoside	3
223	23.52	n/d	$C_{38}H_{56}O_{16}$	767.3502	767.3490	1.56	Triterpenoid glycoside	3
224	23.60	n/d	$C_{43}H_{66}O_{19}$	885.4141	885.4120	2.37	Triterpenoid glycoside	3
225	23.78	n/d	$C_{40}H_{60}O_{17}$	811.3768	811.3752	1.97	Triterpenoid glycoside	3
226	23.79	n/d	$C_{43}H_{66}O_{19}$	885.4135	885.4120	1.69	Triterpenoid glycoside	3
227	23.91	n/d	$C_{38}H_{60}O_{17}$	787.3762	787.3752	1.27	Triterpenoid glycoside	3
228	24.15	n/d	$C_{45}H_{68}O_{20}$	927.4225	927.4225	0.00	Triterpenoid glycoside	3
229	24.25	n/d	$C_{46}H_{74}O_{20}$	945.4703	945.4695	0.85	Triterpenoid glycoside	3
230	24.54	n/d	$C_{41}H_{66}O_{18}$	845.4186	845.4171	1.77	Triterpenoid glycoside	3
231	24.67	n/d	$C_{45}H_{68}O_{20}$	927.4250	927.4225	2.70	Triterpenoid glycoside	3
232	24.70	n/d	$C_{18}H_{28}O_{4}$	307.1917	307.1909	2.60	Fatty acid	3
232	24.77	n/d	$C_{18}H_{28}O_{4}$	307.1918	307.1909	2.93	Fatty acid	3
233	24.71	n/d	$C_{45}H_{68}O_{20}$	927.4230	927.4225	0.54	Triterpenoid glycoside	3
234	25.51	n/d	$C_{47}H_{70}O_{21}$	969.4339	969.4331	0.83	Triterpenoid glycoside	3
235	26.64	n/d	$C_{20}H_{26}O_{4}$	329.1760	329.1753	2.13	n/i	4
236	27.69	n/d	$C_{18}H_{30}O_{3}$	293.2125	293.2116	3.07	Fatty acid	3

* classification following SUMNER et al. (2007). n/d, non-detected. l/c, low concentration. n/i, non-identified. shd, shoulder peak.

4.2. Metabolomics-based chemotaxonomic classification of Barnadesioideae

The initial dataset of the metabolic profile of 56 samples of Barnadesioideae, contained 1915 mass features. In the present study, it was possible to evidence the chemotaxonomic relations of the subfamily Barnadesioideae by means of multivariate analyses. The exploratory multivariate analyses based on the metabolic profile data of 56 Barnadesioideae subfamily members are displayed in **Figure 10**. The hierarchical clustering analysis (HCA) dendrogram, **Figure 10A**, as well as the principal component analysis (PCA) score plot, **Figure 10B**, displayed three main clusters, where the first cluster was composed by *Dasyphyllum* genus samples, a second cluster by *Chuquiraga* genus samples, while a third cluster was composed by *Arnaldoa*, *Barnadesia*, *Fulcaldea, Schlechtendalia* genera samples.

When comparing the cluster groups displayed in the exploratory multivariate analyses based on the metabolome data of members from the subfamily Barnadesioideae with the most recent study of the phylogenetic relationships of Barnadesioideae (see **Figure 11**), it is possible to observe some important similarities. In the study of GRUENSTAEUDL et al. (2009), the genus *Dasyphyllum* is a non-monophyletic group, where *Dasyphyllum* A (subgenus *Archidasyphyllum*) and *Dasyphylum* D (subgenus *Dasyphyllum*) are sister groups of *Arnaldoa* and *Fulcaldea*. The subgenus *Archidasyphyllum* consists of two species, *D. diacanthoides* and *D. excelsum*. The segregation of these two subgenera, according to their metabolic similarities, was also verified in the present study, as evidenced in the HCA dendrogram, where it was possible to observe that all the samples of the genus *Dasyphyllum* were grouped in one main cluster, excluding the two species of the subgenus Archidasyphyllum (*D. diacanthoides* and *D. excelsum*) that were grouped distantly from the rest of species of the genus *Dasyphyllum*.

On the other hand, GRUENSTAEUDL et al. (2009) indicated that the genus *Chuquiraga* is monophyletic. In the present study, most of *Chuquiraga* samples were grouped into a main cluster, except two samples of *C. longiflora*. Additionally, GRUENSTAEUDL et al. (2009) pointed out the phylogenetic proximity of the genus *Arnaldoa, Barnadesia,* a fact that was also observed in the present study where *Arnaldoa* and *Barnadesia* genera were grouped according to their similarities on their metabolic profiles. It is worth no notice that the monotypic genus *Schlechtendalia,* which phylogenetic position is still uncertain, formed a cluster along with species of the genus *Arnaldoa, Barnadesia,* and *Fulcaldea,* according to the results provided by GRUENSTAEUDL et al. (2009).

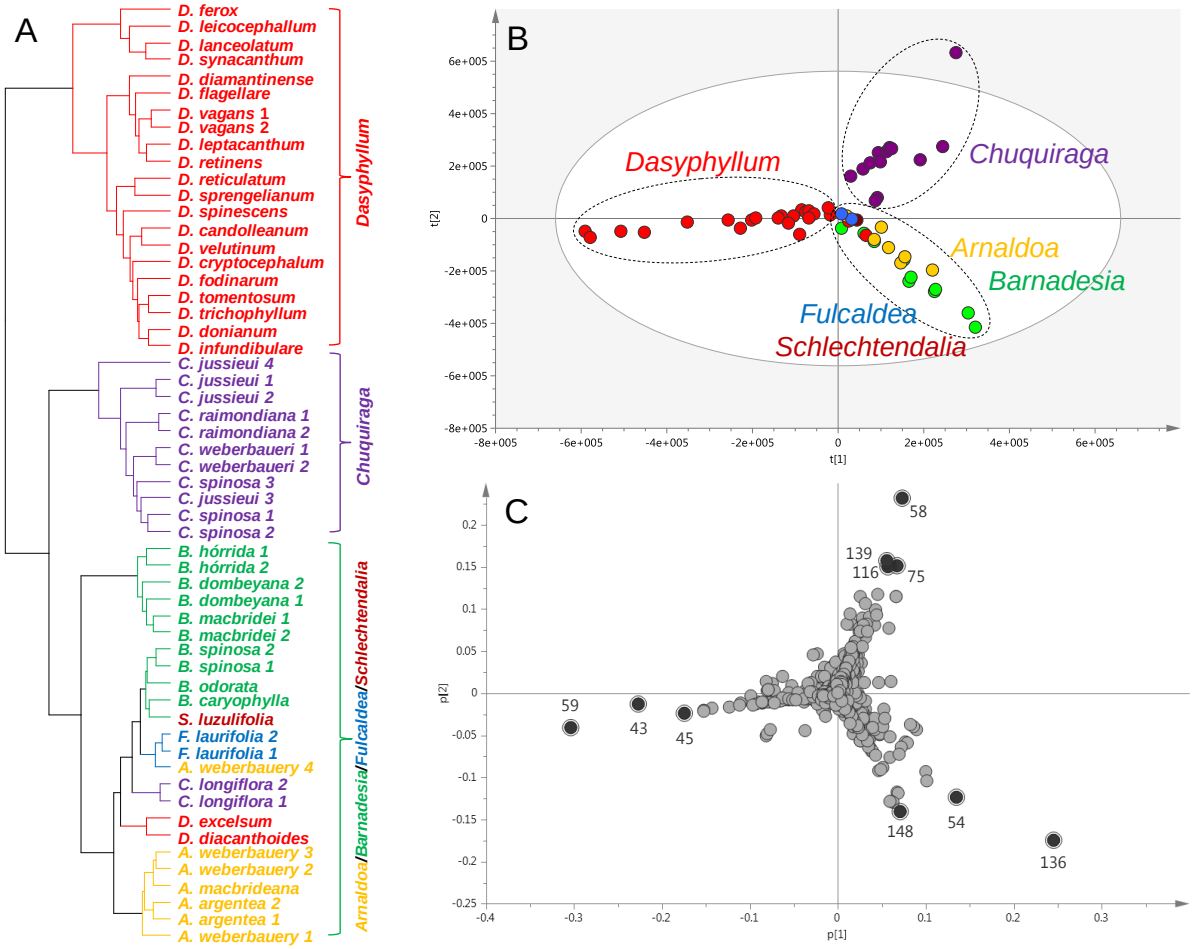


Figure 10. Exploratory multivariate analyses based on the metabolome data of 56 Barnadesioideae subfamily members. **A**) Hierarchical clustering analysis (HCA) displaying three main primary cluster groups. **B**) Principal component analysis (PCA) score plot with three main clusters. **C**) PCA loading plot exhibiting main discriminating variables.

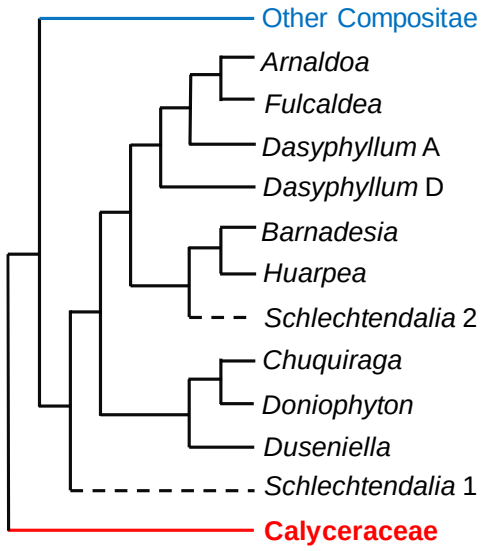


Figure 11. Simplified summary tree presenting phylogenetic relationships in Barnadesioideae based on total evidence analyses (adapted from GRUENSTAEUDL et al. 2009).

The intergeneric relationships of the subfamily Barnadesioideae based on the comparison of their metabolic profiles by exploratory multivariate analyses displayed some similarities to those intergeneric relationships proposed by GRUENSTAEUDL et al. (2009) based on morphological and molecular markers, demonstrating that metabolomics studies may be valuable auxiliary tools in taxonomy and chemotaxonomy studies.

4.3. Discriminating metabolites for Barnadesioideae

In order to determine the main discriminant variables (metabolites) for the three main clusters of Barnadesioideae subfamily, the PCA loading plot was first inspected, since this is comparable to the PCA score plot (see **Figure 10C**). From this initial evaluation on the PCA loading plot, composed by the first and second components (t[1]×t[2]), 10 metabolites were recognized as discriminating variables considering their significant contributions in the loading plot, where compounds **43**, **45** and **59** were determined as discriminating metabolites for *Dasyphyllum* genus, compounds **58**, **75**, **116** and **139** for *Chuquiraga* genus, and compounds **54**, **136** and **148** for the *Arnaldoa-Barnadesia-Fulcaldea-Schlechtendalia* cluster. The chemical structures of some discriminating metabolites are presented in **Figure 12**.

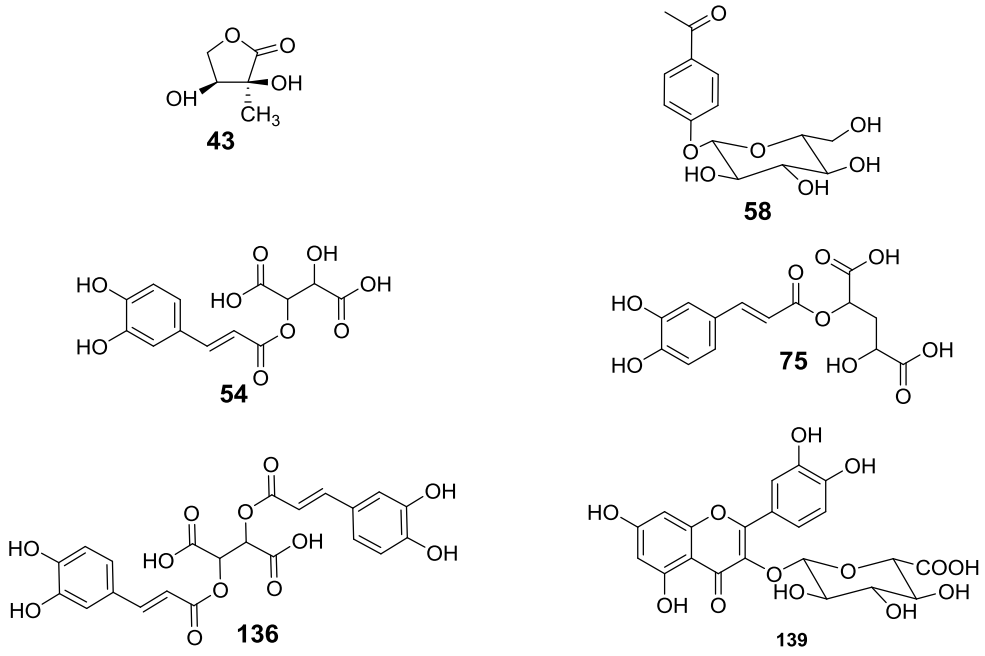


Figure 12. Chemical structures of some discriminant metabolites of Barnadesioideae clusters. *Dasyphyllum* cluster: methylerythronolactone (**43**); *Chuquiraga* cluster: *p*-hydroxyacetophenone 4-*O*-glucoside (**58**), 2-*O*-caffeoyl-4-hydroxypentanedioic acid (**75**), quercetin-3-*O*-glucuronide (**139**); *Arnaldoa-Barnadesia-Fulcaldea-Schlechtendalia* cluster: 2-*O*-caffeoyltartaric acid (**54**), 2,3-*O*-dicaffeoyltartaric acid (**136**).

To validate the proposed discriminating variables displayed in the PCA analysis, a subsequent supervised multivariate analysis (OPLS-DA) based on the metabolome data of 56 Barnadesioideae members was performed and displayed in **Figure 13**. The validity of the model was supported by adequate goodness of fit and predictability of the model (R^2 = 0.54, Q^2 = 0.45). By inspecting the OPLS-DA loading plot, it was possible to identify the same 10 variables identified as discriminating ones in the PCA loading plot. Additionally, further variables were identified as potential discriminating metabolites for Barnadesioideae subfamily by considering a variable importance in the projection value (VIP) greater than 1.0, and a reliable confidence interval in the coefficient plot (detailed in **Table 12**).

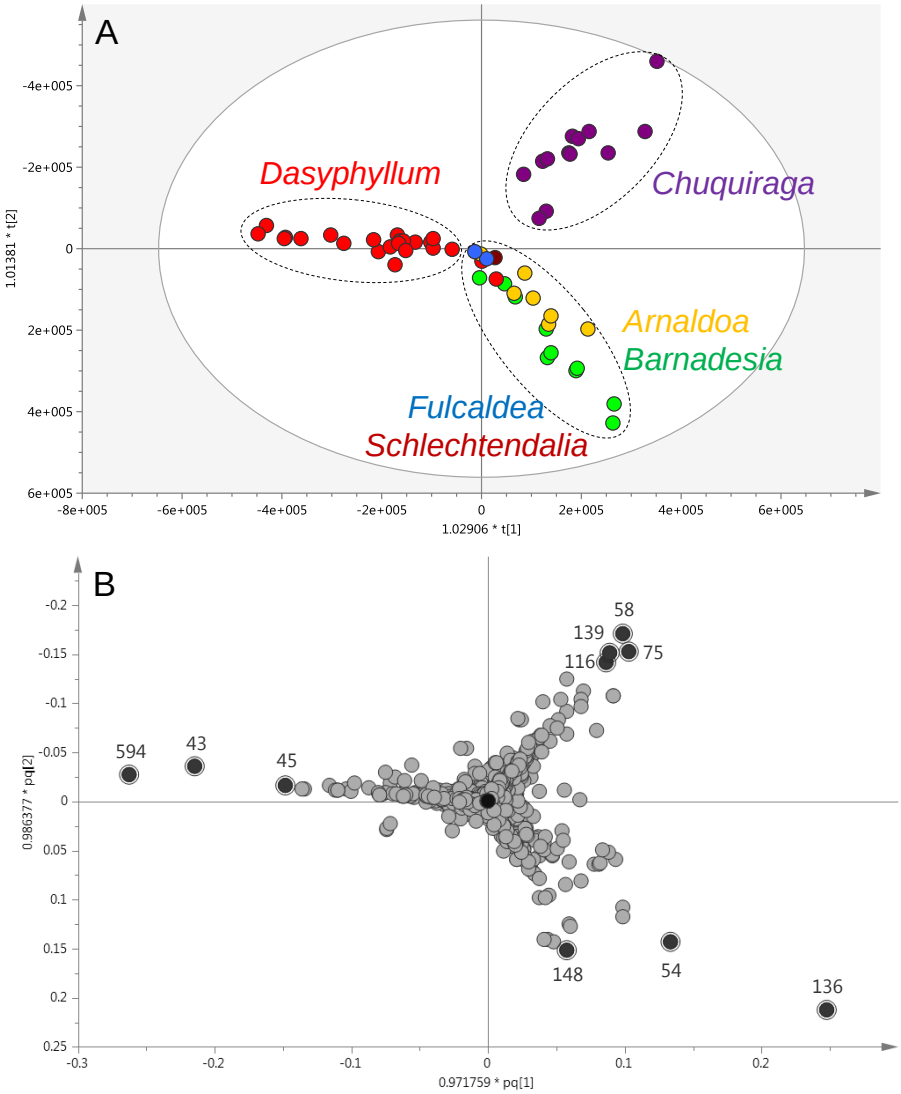


Figure 13. Supervised multivariate analysis based on the metabolome data of 56 Barnadesioideae members. **A**) OPLS-DA score plot (R^2 = 0.54, Q^2 = 0.45). **B**) OPLS-DA loading plot exhibiting main discriminating variables.

Table 12. Discriminating metabolites for Barnadesioideae.

No	Identity	Class	CoeffCS	cvSE	VIP
54	2-*O*-Caffeoyltartaric acid	*Arnaldoa**	1.57×10^{-06}	5.29×10^{-07}	5.84
136	2,3-*O*-Dicaffeoyltartaric acid	*Arnaldoa*	1.98×10^{-06}	7.54×10^{-07}	8.49
152	3,4-*O*-Dicaffeoylquinic acid	*Arnaldoa*	-4.85×10^{-07}	4.41×10^{-07}	3.40
149	Coumaroylcaffeoyltartaric acid	*Arnaldoa*	7.38×10^{-07}	2.62×10^{-07}	2.10
77	4-*O*-Caffeoylquinic acid	*Arnaldoa*	-6.63×10^{-07}	4.14×10^{-07}	4.09
148	Flavonoid glycoside	*Arnaldoa*	-8.29×10^{-07}	5.51×10^{-07}	5.27
69	Caffeoyltartronic acid	*Arnaldoa*	-5.09×10^{-07}	3.16×10^{-07}	3.25
123	Caffeic acid ethyl ester	*Arnaldoa*	-6.56×10^{-07}	2.44×10^{-07}	4.21
185	Benzophenone glycosides	*Arnaldoa*	-3.51×10^{-07}	2.89×10^{-07}	2.26
84	Coumaroyltartronic acid	*Arnaldoa*	-3.72×10^{-07}	2.94×10^{-07}	2.29
120	Feruloyltartronic acid	*Arnaldoa*	-6.46×10^{-07}	2.33×10^{-07}	4.12
229	Triterpenoid glycoside	*Arnaldoa*	-2.38×10^{-07}	9.11×10^{-08}	1.53
214	Caffeic acid ester derivative	*Arnaldoa*	-3.52×10^{-07}	2.97×10^{-07}	2.23
81	n/i	*Chuquiraga*	2.09×10^{-08}	2.01×10^{-07}	2.15
78	Caffeic acid ester derivative	*Chuquiraga*	-1.27×10^{-07}	1.29×10^{-07}	3.53
114	2-*O*-Feruloyltartaric acid	*Chuquiraga*	-4.31×10^{-07}	1.88×10^{-07}	3.31
116	Alkyl glycoside	*Chuquiraga*	-3.00×10^{-07}	2.46×10^{-07}	4.38
155	Kaempferol-3-*O*-glucuronide	*Chuquiraga*	-2.30×10^{-07}	1.44×10^{-07}	3.39
139	Quercetin-3-*O*-glucuronide	*Chuquiraga*	-3.28×10^{-07}	3.23×10^{-07}	4.75
73	Caffeoyltartaric acid	*Chuquiraga*	-2.17×10^{-07}	1.58×10^{-07}	2.33
118	Caffeic acid ester derivative	*Chuquiraga*	-1.54×10^{-07}	1.16×10^{-07}	2.19
212	Triterpenoid glycoside	*Chuquiraga*	-9.00×10^{-08}	6.26×10^{-08}	1.31
58	*p*-Hydroxyacetophenone-4-*O*-glucoside	*Chuquiraga*	-2.21×10^{-07}	4.33×10^{-07}	3.25
60	Alkyl glycoside	*Chuquiraga*	-1.40×10^{-07}	1.39×10^{-07}	2.10
117	Ferulic acid ester derivative	*Chuquiraga*	-2.62×10^{-07}	2.64×10^{-07}	3.65
75	Caffeic acid ester derivative	*Chuquiraga*	-2.28×10^{-07}	1.95×10^{-07}	4.89
65	Alkyl glycoside	*Chuquiraga*	-1.49×10^{-07}	1.44×10^{-07}	2.19
138	Quercetin-3-*O*-rutinoside	*Dasyphyllum*	1.65×10^{-06}	5.48×10^{-07}	4.82
74	5-*O*-Caffeoylquinic acid	*Dasyphyllum*	-8.62×10^{-07}	4.75×10^{-07}	4.87
156	Kaempferol-3-*O*-glucoside	*Dasyphyllum*	-3.07×10^{-07}	7.09×10^{-08}	3.27
236	Fatty acid	*Dasyphyllum*	-3.86×10^{-07}	3.42×10^{-07}	2.23
56	3-*O*-Caffeoylquinic acid	*Dasyphyllum*	-1.07×10^{-07}	1.14×10^{-07}	2.51
111	Flavonoid glycoside	*Dasyphyllum*	-9.60×10^{-08}	1.39×10^{-07}	2.05
43	Methylerythronolactone	*Dasyphyllum*	-3.01×10^{-07}	2.03×10^{-07}	5.11
59	Glucuronolactone derivative	*Dasyphyllum*	-2.53×10^{-07}	2.79×10^{-07}	5.23
50	Pyranone glycoside	*Dasyphyllum*	-1.04×10^{-07}	1.04×10^{-07}	2.16
52	Glucuronolactone derivative	*Dasyphyllum*	-1.49×10^{-07}	1.33×10^{-07}	2.94
46	n/i	*Dasyphyllum*	-1.38×10^{-07}	2.06×10^{-07}	2.85
44	Glucuronolactone derivative	*Dasyphyllum*	-9.78×10^{-08}	8.16×10^{-08}	2.06

n/i, non-identified. **Arnaldoa* cluster includes *Barnadesia*, *Fulcaldea* and *Schlechtendalia*.

The 39 discriminating metabolites summarized in **Table 12** are represented in a heatmap displayed in the **Figure 14**, where peak areas where submitted to univariate normalization, enabling a graphical representation of the relative distribution of these metabolites along members of Barnadesioideae, and supporting the occurrence of three main cluster groups in the subfamily.

Figure 14. Heatmap of discriminating metabolites in the subfamily Barnadesioideae.

4.4. Metabolomics based species classification of *Chuquiraga* species.

The initial dataset of the metabolic profile of 316 samples of *Chuquiraga* samples contained 1981 mass features. Before multivariate analyses the initial dataset was organized and aleatorily distributed (for training, validation and test sets) according to the **Table 13**.

Table 13. Sample distribution for exploratory and supervised analyses.

Sample	N	Exploratory	Supervised		
			Training set (~70%)	Validation set (~30%)	Test set
***C. jussieui* 1**	56	56	40	16	-
C. weberbaueri	24	24	16	8	-
C. spinosa	132	132	90	42	-
***C. jussieui* 2**	48	48	34	14	-
Huamanpinta*	39	-	-	-	39
Huamanpinta**	8	-	-	-	8
C. jussieui* 1**	4	-	-	-	4
C. weberbaueri**	3	-	-	-	3
C. spinosa**	2	-	-	-	2

*Peruvian samples acquired in January 2015. ** Abroad samples (Austria, USA) acquired during 2015.
***Samples acquired in January 2016.

Then, the dataset was initially submitted to hierarchical clustering analysis (HCA) to evaluate similarities among the metabolic profiles. The **Figure 15** displays the circular HCA dendrogram based on the metabolic profile of 260 *Chuquiraga* samples.

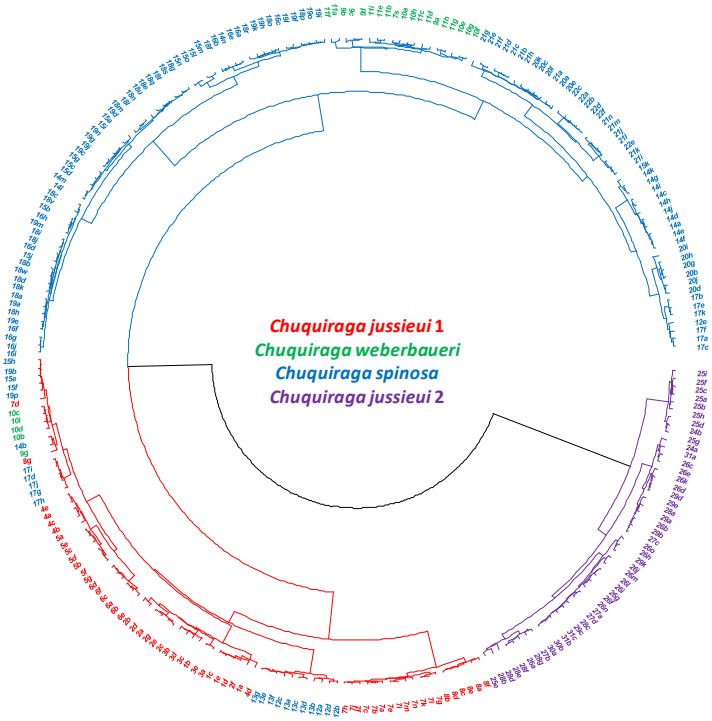


Figure 15. Circular dendrogram based on the metabolic profile of 260 *Chuquiraga* samples.

Chuquiraga jussieui occurs in two geographically separated areas (EZCURRA 1985), from highlands of Ecuador and northern Peru (*C. jussieui* 1), and from highlands of southern Peru to northern Bolivia (*C. jussieui* 2). The samples of these separated areas of *C. jussieui* were displayed in two different clusters in the circular HCA dendrogram demonstrating significant differences in the metabolic profile composition of these two geographically distant population. A third cluster in the dendrogram was composed by samples of *C. spinosa* and *C. weberbaueri*, demonstrating certain similarity in the metabolic profiles of these two species, but both were different from *C. jussieui* 1 and 2. Then, an orthogonal partial least-squares discriminant analysis (OPLS-DA) prediction model was established based on the metabolic profiles of 180 *Chuquiraga* samples as seen in **Figure 16**.

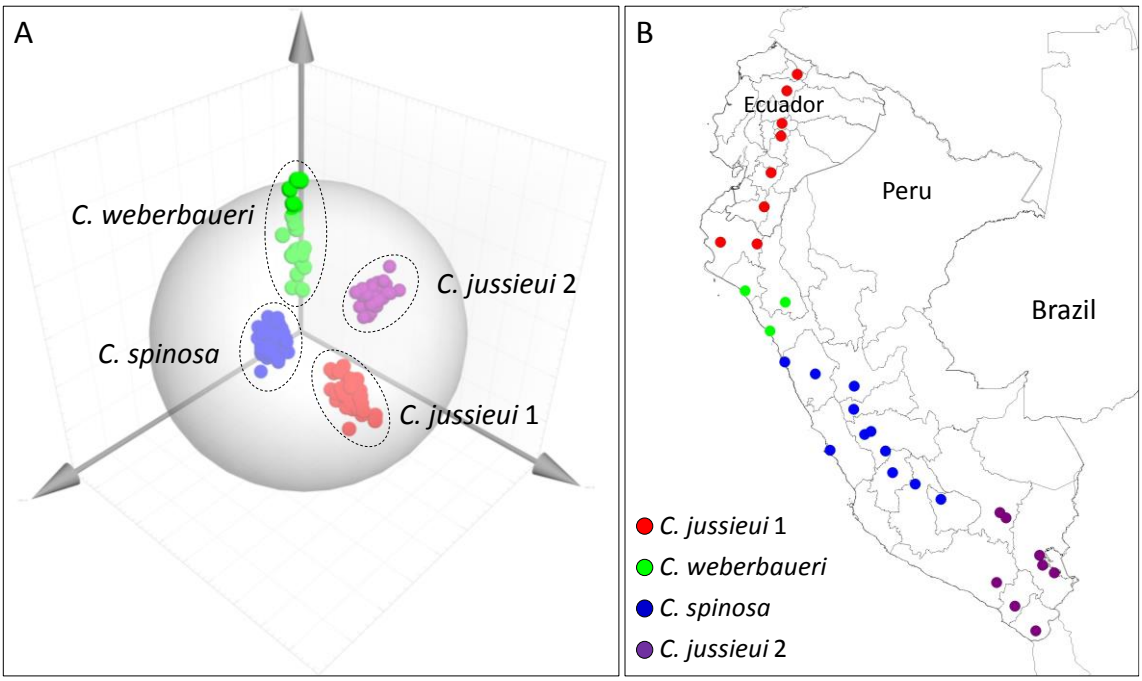


Figure 16. Orthogonal partial least-squares discriminant analysis (OPLS-DA) and distribution map of *Chuquiraga* species. **A**) 3D OPLS-DA score plot (R^2 = 0.941, Q^2 = 0.913) with four main cluster groups corresponding to samples of *C. spinosa*, *C. weberbaueri*, and two populations of *C. jussieui*. **B**) Distribution map of *Chuquiraga* samples: *C. jussieui* 1 = Ecuador and northern Peru, *C. jussieui* 2 = southern Peru), *C. weberbaueri* = northern Peru, *C. spinosa* = central Peru.

The OPLS-DA score plot, displayed in **Figure 16A**, presented adequate goodness of fit and predictability values (R^2 = 0.941, Q^2 = 0.913), displaying four main cluster groups corresponding to samples of *C. spinosa*, *C. weberbaueri*, and samples of two different geographical origin of *C. jussieui*. These cluster groups were also in agreement with their geographical distributions along Ecuador and Peru (**Figure 16B**). Afterwards, following the sample distribution established in **Table 13**, the OPLS-DA model was submitted to the external

validation step to obtain correct perfection values of 100% for *C. jussieui* (*C. jussieui* 1 and 2), *C. spinosa*, and *C. weberbaueri*, results displayed in **Table 14**, demonstrating the capacity of the model for species discrimination (*C. jussieui*, *C. spinosa*, and *C. weberbaueri*), but also for distinction of species with different geographical origin (*C. jussieui* 1 and 2).

Table 14. Classification results of validation and test sets of *Chuquiraga* by OPLS-DA and *k*NN.

Samples	**N**	***C. jussieui* 1**		***C. weberbaueri***		***C. spinosa***		***C. jussieui* 2**	
		OPLS-DA	*k*NN	OPLS-DA	*k*NN	OPLS-DA	*k*NN	OPLS-DA	*k*NN
***C. jussieui* 1**	16	16	16	-	-	-	-	-	-
C. weberbaueri	8	-	-	8	7	-	1	-	-
C. spinosa	42	-	-	-	-	42	42	-	-
***C. jussieui* 2**	14	-	-	-	-	-	-	14	14
Correct prediction	(%)	100	100	100	87.5	100	100	100	100
Huamanpinta*	39	-	-	-	-	39	32	-	7
Huamanpinta**	8	-	-	-	-	8	8	-	-
C. jussieui* 1**	4	4	4	-	-	-	-	-	-
C. weberbaueri**	3	-	-	3	3	-	-	-	-
C. spinosa**	2	-	-	-	-	2	2	-	-

*Peruvian samples acquired in 01 2015. ** Aboard samples (Austria, USA) acquired during 2015. ***Samples acquired in 01 2016.

The OPLS-DA prediction model finds applicability as a mean for assessing the identity of "Huamanpinta" samples acquired in Peru, Austria and USA. Some "Huamanpinta" samples acquired in Peru declared to contain *C. spinosa* leaves, while other samples did not declare botanical identification. "Huamanpinta" samples acquired from Austria, USA and local exporters in Peru declared *C. spinosa* as botanical source. By means of the OPLS-DA model all "Huamanpinta" samples were classified as *C. spinosa*, demonstrating its applicability for authentication and quality control of unknown samples. Additionally, to prove the applicability of the model along the time, samples of *C. jussieui*, *C. spinosa*, and *C. weberbaueri*, acquired one year later, in 2016, were also assessed and their identity were correctly predicted.

To compare the results obtained by OPLS-DA, predictions were also made using the *k*NN algorithm. In the validation set, correct predictions of 100% were obtained with $k = 30$, 20 and 10 for *C. jussieui* and *C. spinosa*, while the higher percentage of correct prediction (87.5%) for *C. weberbaueri* was obtained with $k = 10$. Therefore, in the test set a $k = 10$ was used for the identity predictions. From the 39 "Huamanpinta" samples acquired in Peru, 32 were classified as *C. spinosa*, and other seven samples as *C. jussieui* 2 (southern Peru), whereas the eight "Huamanpinta" samples acquired in Austria, USA and local exporters in Peru were

classifies as *C. spinosa*. Therefore, LC-MS metabolomics constitute a high-throughput approach for the quality assessment of *Chuquiraga* species. Supervised multivariate analyses enable geographical discrimination and species classification based on the metabolic profile of *Chuquiraga* species.

4.5. Discriminating metabolites for *Chuquiraga*

Discriminating variables (metabolites) in the OPLS-DA model were identified after inspections of the corresponding loading and score plots, considering initially a discriminating variable the one with a significant contribution in the loading plots displayed in **Figure 17**.

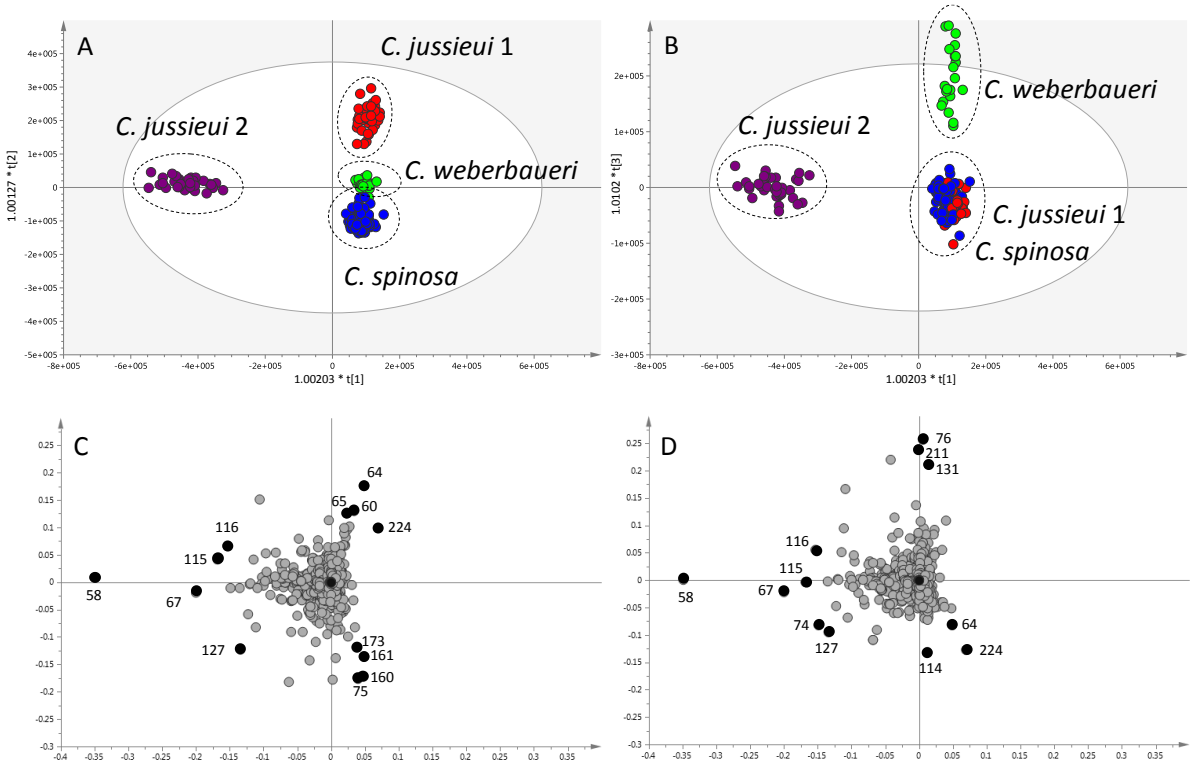


Figure 17. OPLS-DA score plots (**A** = T[1]×T[2]; **B** = T[1]×T[3],) and corresponding loading plots (**C** = T[1]×T[2]; **D** = T[1]×T[3]) based on the metabolic profile of *Chuquiraga* species. The cluster group correspond *C. spinosa*, *C. weberbaueri*, and two populations of *C. jussieui*.

From this initial evaluation on the OPLS-DA score plot composed by the first and second two latent variables (t[1]×t[2]), compounds **60**, **64**, **65** and **224**, were determined as discriminating metabolites for *C. jussieui* 1 (Ecuador and Northern Peru), and compounds **58**, **67**, **115**, **116** and **127** for *C. jussieui* 2 (Southern Peru). In this score plot, it was also possible to determine compounds **75**, **160**, **161** and **173** as discriminating variables for *C. spinosa*;

however, it was necessary to explore the score plot involving the first and third latent variables (t[1]×t[3]) to determinate compounds **76**, **131** and **211** as disseminating metabolites for *C. weberbaueri*.

To support this initial summary, further OPLS-DA models with only two classes were generated: *C. jussieui* 1 × [*C. jussieui* 2 + *C. spinosa* + *C. weberbaueri*]; *C. jussieui* 2 × [*C. jussieui* 1 + *C. spinosa* + *C. weberbaueri*]; *C. spinosa* × [*C. jussieui* 1 + *C. jussieui* 2 + *C. weberbaueri*]; *C. weberbaueri* × [*C. jussieui* 1 + *C. jussieui* 2 + *C. spinosa*].

These knew two-class OPLS-DA models were used to explore *S*-plots from which further variables were identified as potential discriminating metabolites considering a variable importance in the projection value (VIP) greater than 1.0, and a reliable confidence interval in the coefficient plot. This approach confirmed the compounds identified in **Figure 17**, and increased to 61 the number of potential discriminating metabolites for *Chuquiraga* species and are detailed in **Table 15**. Although some variables displayed confidence intervals greater than their correspondent coefficients, these are still included in **Table 15** because these fulfilled the other two selection criteria.

Ten main discriminating metabolites for each group are represented in a heatmap, as displayed in **Figure 18**, where peak areas where submitted to univariate normalization, enabling a graphical representation of the relative distribution of these metabolites along *Chuquiraga* species.

The samples of *C. jussieui* 1 displayed alkyl glycosides and triterpenoid glucosides as discriminating metabolites, whereas *C. jussieui* 2 displayed higher concentration of *p*-hydroxyacetophenone and *p*-hydroxyacetophenone glycosides. The concentrations of caffeic acid were higher in *C. weberbaueri* samples, while *C. spinosa* displayed higher concentrations of monoterpene diglycosides and the new caffeic acid ester derivative 2-*O*-caffeoyl-4-hydroxypentanedioic acid as distinctive metabolites.

Table 15. Discriminating metabolites for *Chuquiraga* species.

No	Identity	Class	CoeffCS	cvSE	VIP
64	Alkyl glycoside	*C. jussieui* 1	1.47×10^{-06}	2.79×10^{-07}	6.41
224	Triterpenoid glycoside	*C. jussieui* 1	2.19×10^{-07}	3.21×10^{-07}	4.76
60	Alkyl glycoside	*C. jussieui* 1	7.42×10^{-07}	2.58×10^{-07}	4.39

65	Alkyl glycoside	*C. jussieui* 1	1.13×10^{-06}	2.46×10^{-07}	4.29
231	Triterpenoid glycoside	*C. jussieui* 1	9.49×10^{-07}	3.93×10^{-07}	3.74
225	Triterpenoid glycoside	*C. jussieui* 1	1.26×10^{-06}	4.39×10^{-07}	3.06
198	Triterpenoid glycoside	*C. jussieui* 1	1.38×10^{-06}	2.42×10^{-07}	2.42
209	Triterpenoid glycoside	*C. jussieui* 1	5.51×10^{-07}	3.25×10^{-07}	2.31
210	Triterpenoid glycoside	*C. jussieui* 1	3.90×10^{-07}	2.36×10^{-07}	2.26
230	Triterpenoid glycoside	*C. jussieui* 1	8.00×10^{-07}	1.93×10^{-07}	2.07
191	Acylated flavonoid glycoside	*C. jussieui* 1	6.39×10^{-07}	1.93×10^{-07}	2.01
184	Acylglycerol diglycoside	*C. jussieui* 1	5.28×10^{-07}	1.57×10^{-07}	1.98
222	Triterpenoid glycoside	*C. jussieui* 1	5.47×10^{-07}	2.21×10^{-07}	1.87
112	n/i	*C. jussieui* 1	7.58×10^{-07}	1.27×10^{-07}	1.86
76	Caffeic acid	*C. weberbaueri*	1.65×10^{-06}	1.21×10^{-06}	8.15
211	Triterpenoid glycoside	*C. weberbaueri*	3.53×10^{-06}	2.31×10^{-06}	8.12
131	Flavonoid glycoside	*C. weberbaueri*	-4.70×10^{-10}	1.77×10^{-06}	6.78
43	Methylerythronolactone	*C. weberbaueri*	3.03×10^{-06}	9.41×10^{-07}	3.82
90	Benzoyltartaric acid	*C. weberbaueri*	3.79×10^{-06}	1.22×10^{-06}	3.68
159	Monoterpene diglycoside	*C. weberbaueri*	3.74×10^{-06}	1.80×10^{-06}	3.48
196	Triterpenoid glycoside	*C. weberbaueri*	-1.20×10^{-07}	7.56×10^{-07}	2.93
146	Monoterpene diglycoside	*C. weberbaueri*	1.97×10^{-06}	1.27×10^{-06}	2.73
113	Caffeoylglyceric acid	*C. weberbaueri*	1.74×10^{-06}	7.68×10^{-07}	2.42
107	n/i	*C. weberbaueri*	1.70×10^{-06}	1.46×10^{-06}	2.41
160	Caffeoyl-feruloyltartaric acid	*C. spinosa*	1.37×10^{-06}	5.65×10^{-07}	6.68
75	2-*O*-Caffeoyl-4-hydroxypentanedioic acid	*C. spinosa*	-3.50×10^{-07}	2.69×10^{-07}	5.71
161	Flavonoid glycoside	*C. spinosa*	2.29×10^{-06}	4.59×10^{-07}	5.63
173	Diferuloyltartaric acid	*C. spinosa*	1.73×10^{-07}	3.77×10^{-07}	4.58
117	Ferulic acid ester derivative	*C. spinosa*	2.52×10^{-07}	5.82×10^{-07}	3.73
157	Flavonoid glycoside	*C. spinosa*	1.13×10^{-06}	2.94×10^{-07}	3.46
54	2-*O*-Caffeoyltartaric acid	*C. spinosa*	-1.30×10^{-08}	3.37×10^{-07}	3.37
129	Monoterpene diglycoside	*C. spinosa*	1.32×10^{-06}	3.81×10^{-07}	3.19
182	Alkyl glycoside	*C. spinosa*	1.12×10^{-06}	2.64×10^{-07}	2.42
134	Monoterpene diglycoside	*C. spinosa*	7.69×10^{-07}	3.9×10^{-07}	2.35
100	Coumaric acid ester derivative	*C. spinosa*	3.25×10^{-07}	1.83×10^{-07}	2.13
143	Flavonoid glycoside	*C. spinosa*	3.92×10^{-07}	1.60×10^{-07}	2.02
63	Isopropylmalic acid	*C. spinosa*	8.73×10^{-08}	4.27×10^{-07}	1.96
158	n/i	*C. spinosa*	9.86×10^{-07}	3.48×10^{-07}	1.95
114	2-*O*-Feruloyltartaric acid	*C. spinosa*	8.76×10^{-07}	6.21×10^{-07}	1.92
102	Alkyl glycoside	*C. spinosa*	7.26×10^{-07}	4.72×10^{-07}	1.55
58	*p*-Hydroxyacetophenone 4-*O*-glucoside	*C. jussieui* 2	2.99×10^{-07}	7.62×10^{-07}	12.95
67	*p*-Hydroxyacetophenone 4-*O*-(6-*O*-apiosyl)-glucoside	*C. jussieui* 2	5.22×10^{-07}	3.78×10^{-07}	7.45
115	5-*O*-Feruloylquinic acid	*C. jussieui* 2	7.59×10^{-07}	1.97×10^{-07}	6.27
116	Alkyl glycoside	*C. jussieui* 2	7.86×10^{-07}	3.1×10^{-07}	5.80
74	5-*O*-Caffeoylquinic acid	*C. jussieui* 2	5.33×10^{-07}	5.12×10^{-07}	5.65
127	Ferulic acid ester derivative	*C. jussieui* 2	7.75×10-08	4.32×10^{-07}	5.13
86	*p*-Hydroxyacetophenone	*C. jussieui* 2	-3.30×10^{-07}	3.16×10^{-07}	4.38
61	Phenolic glycoside	*C. jussieui* 2	4.02×10^{-07}	3.66×10^{-07}	4.35
215	Triterpenoid glycoside	*C. jussieui* 2	5.11×10^{-07}	2.73×10^{-07}	4.29
136	2,3-*O*-Dicaffeoyltartaric acid	*C. jussieui* 2	1.25×10^{-07}	2.66×10^{-07}	4.14
88	Alkyl glycoside	*C. jussieui* 2	6.33×10^{-07}	2.81×10^{-07}	3.66
80	Alkyl glycoside	*C. jussieui* 2	6.77×10^{-07}	1.59×10^{-07}	3.61
57	5-*O*-(4-*O*-Glucosylcaffeoyl)quinic acid	*C. jussieui* 2	8.50×10^{-08}	2.00×10^{-07}	3.52
49	Phenolic glycoside	*C. jussieui* 2	-2.40×10^{-08}	4.22×10^{-07}	3.35
53	n/i	*C. jussieui* 2	5.29×10^{-07}	1.59×10^{-07}	3.32
204	Triterpenoid glycoside	*C. jussieui* 2	2.63×10^{-07}	2.15×10^{-07}	3.13
208	Triterpenoid glycoside	*C. jussieui* 2	6.53×10^{-07}	1.85×10^{-07}	3.12
68	Glucosylcaffeoylquinic acid	*C. jussieui* 2	2.65×10^{-07}	2.15×10^{-07}	2.79
208	Triterpenoid glycoside	*C. jussieui* 2	5.75×10^{-07}	1.56×10^{-07}	2.76
73	Caffeoyltartaric acid	*C. jussieui* 2	3.18×10^{-07}	2.08×10^{-07}	2.73
201	Triterpenoid glycoside	*C. jussieui* 2	6.08×10^{-07}	1.36×10^{-07}	2.67

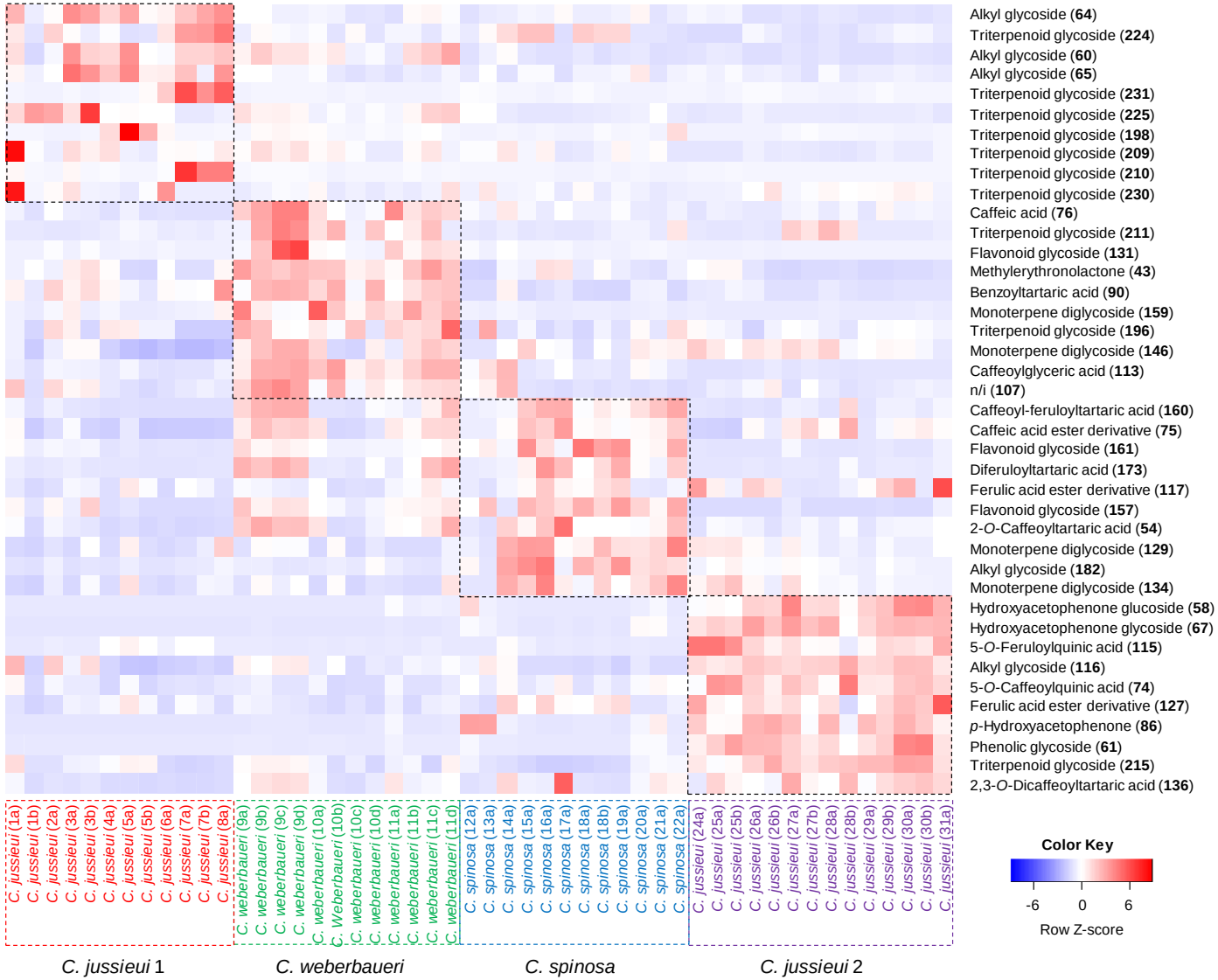


Figure 18. Heatmap of discriminating metabolites in *Chuquiraga* species.

4.6. *Chuquiraga* species classification by HPLC profiles

HPLC chromatographic fingerprinting represents a robust and versatile technique for plant drug analysis because it gives the possibility to examine metabolites without the need for derivatization in the crude extracts. When coupled to a DAD detector, HPLC provides on-line UV spectra that is particularly useful for the detection of natural products with characteristic chromophores. Therefore, HPLC chromatographic fingerprinting provides a non-sophisticated and feasible method for quality proof of herbal drugs, authentication of raw materials, exclusion of possible falsifications and adulterations, and analysis of consistency among batches of herbal medicines (WOLFENDER et al. 2010; WAGNER et al. 2011, 2015, 2016).

After analyzing the 3D plot of HPLC-DAD chromatographic profiles (retention times × wavelengths × absorbances), the traces at 270 nm were selected based on the highest absorbances observed at this wavelength and because it prevented inclusion of background noise of the gradient elution. The fingerprints of 145 samples of *Chuquiraga*, which were recorded in more than 200 h (>8 days) of continuous analysis, presented shifts in retention times, displayed in **Figure 19A**, that were corrected by using the parametric time warping algorithm, as seen in **Figure 19B**, where samples were aligned to a reference chromatogram (*Chuquiraga* mixture).

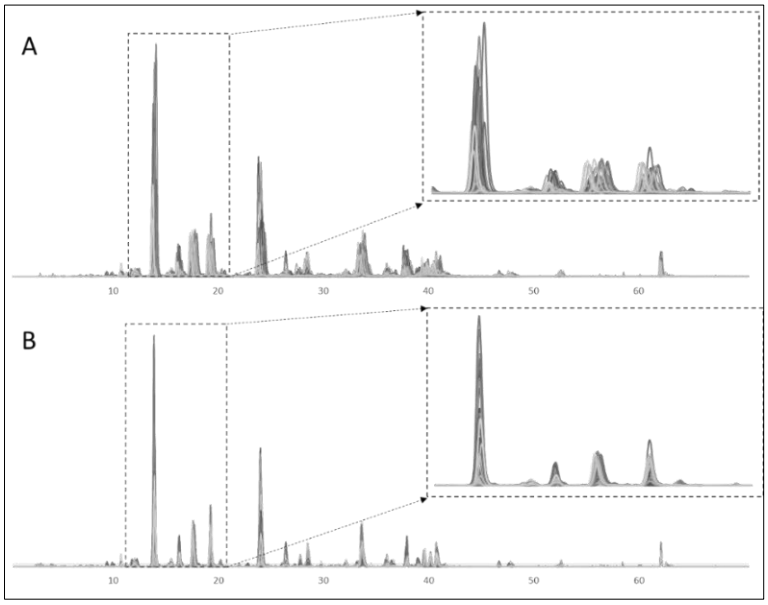


Figure 19. Alignment of HPLC chromatographic fingerprints (270 nm) of 145 samples of *Chuquiraga*. **A)** Superposition of all chromatograms before alignment. **B)** Superposition of all chromatograms after alignment by using the parametric time warping algorithm.

After the alignment, the obtained chromatographic fingerprints were exported in a spreadsheet and organized before exploratory and supervised analyses, as detailed in **Table 16**. The dataset was initially submitted to hierarchical clustering analysis (HCA) to evaluate similarities among the fingerprints (**Figure 20A**) enabling a clear differentiation between *C. jussieui* 1 samples (Ecuador and northern Peru), and *C. jussieui* 2 samples (southern Peru).

Table 16. Sample distribution for exploratory and supervised analyses.

Sample	N	Exploratory	Supervised		
			Training set (~70%)	Validation set (~30%)	Test set
C. jussieui 1	31	31	20	11	-
C. weberbaueri	20	20	13	7	-
C. spinosa	44	44	30	14	-
C. jussieui 2	32	32	24	8	-
Huamanpinta*	10	-	-	-	10
Huamanpinta**	8	-	-	-	8

*Peruvian samples acquired in January 2015. ** Aboard samples (Austria, USA) acquired during 2015.

Then, an orthogonal partial least-squares discriminant analysis (OPLS-DA) prediction model was established based on the chromatographic profiles of *Chuquiraga* species. The built model (**Figure 20B**) presented adequate fit and predictability values (R^2 = 0.842, Q^2 = 0.736).

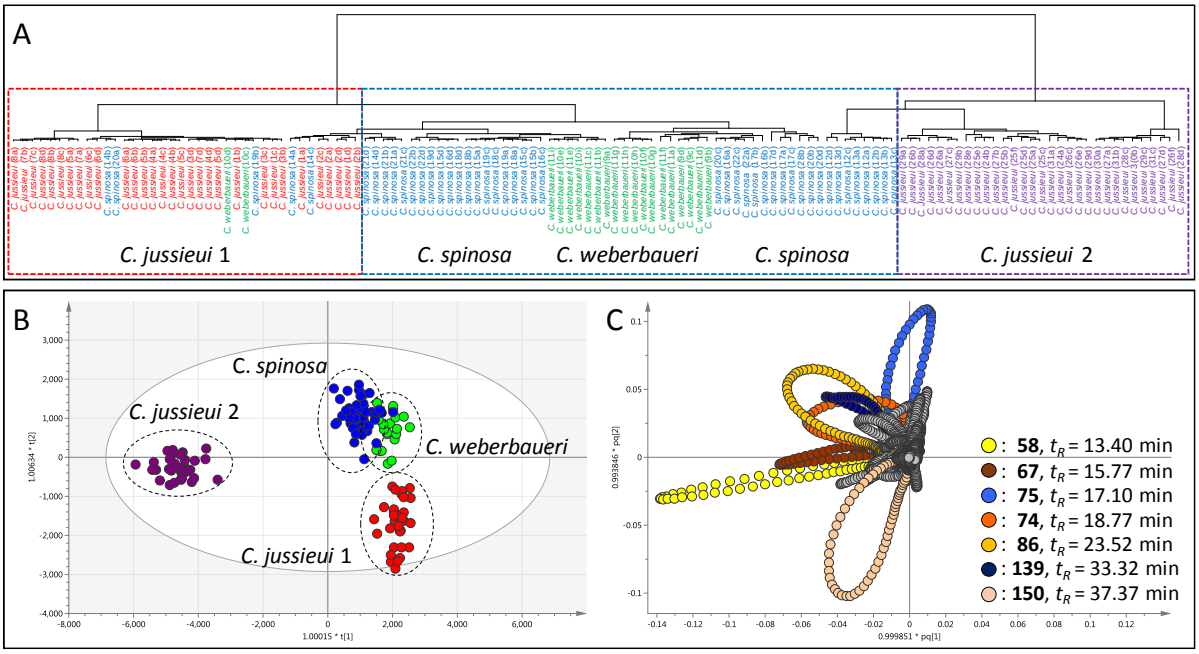


Figure 20. Exploratory and supervised multivariate analyses of commercial samples of *Chuquiraga* based on their HPLC chromatographic fingerprints at 270 nm. **A**) Hierarchical clustering analysis (HCA) dendrogram with three doted boxes corresponding to proposed primary cluster groups. **B**) Orthogonal partial least-squares discriminant analysis (OPLS-DA) score plot (R^2 = 0.842, Q^2 = 0.736) with four main cluster groups corresponding to three species of *Chuquiraga* (*C. jussieui*, *C. spinosa*, *C. weberbaueri*), and samples of separated areas of *C. jussieui* (1 = Ecuador and northern Peru, 2 = southern Peru). **C**) OPLS-DA loading plot exhibiting main variables.

The OPLS-DA model was then submitted to external validation obtaining correct perfection values of 100% for *C. jussieui* 1 and *C. jussieui* 2, 85% for *C. spinosa*, and 71% for *C. weberbaueri* (**Table 17**). Finally, the identity of "Huamanpinta" samples acquired in Peru, Austria and USA, were assessed in the OPLS-DA model and classified as *C. spinosa*, thus demonstrating the applicability of this models for authentication and quality control of unknown samples.

Table 17. Classification results of validation and test sets of *Chuquiraga* species by OPLS-DA.

Samples	N	Classified as			
		***C. jussieui* 1**	***C. weberbaueri***	***C. spinosa***	***C. jussieui* 2**
C. jussieui 1	11	11	-	-	-
C. weberbaueri	7	-	5	2	-
C. spinosa	14	1	-	12	1
C. jussieui 2	8	-	-	-	8
Correct prediction	(%)	100	71.43	85.71	100
Huamanpinta*	10	-	-	10	-
Huamanpinta*	8	-	-	8	-

*Peruvian samples acquired in January 2015. ** Aboard samples (Austria, USA) acquired during 2015.

In the OPLS-DA loading plot, **Figure 20C**, seven main chromatographic peaks were identified as main variables of the model. Their occurrence, corresponding UV spectra and chemical structures of those main constituents are displayed in **Figure 21**.

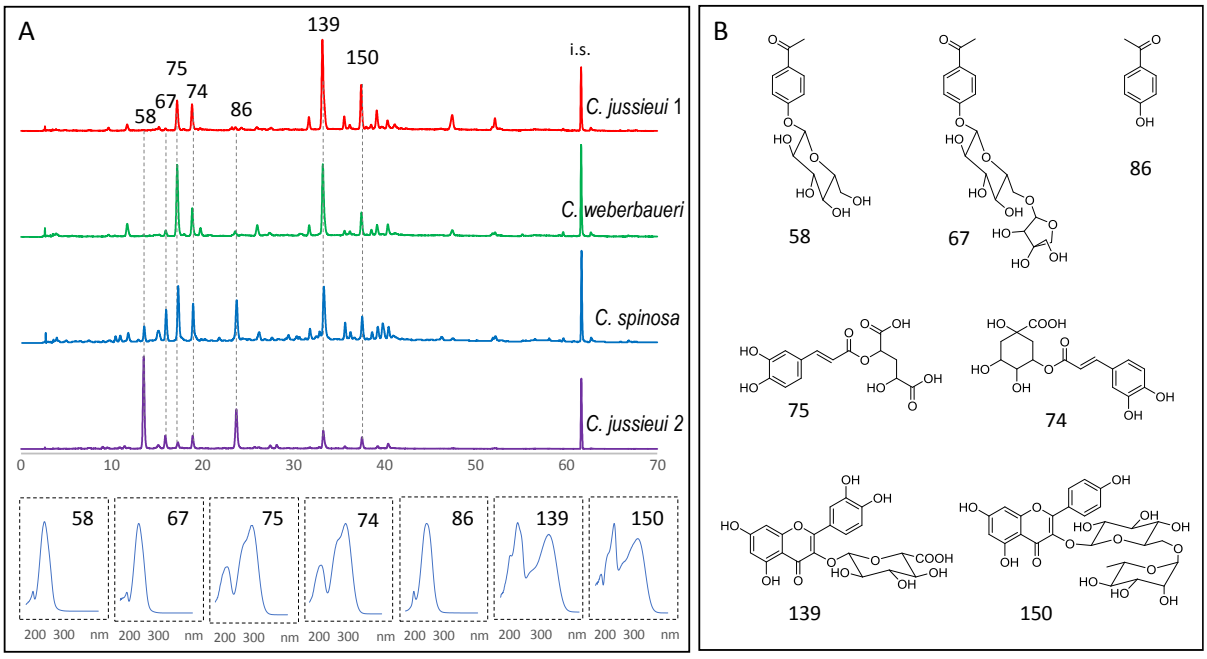


Figure 21. HPLC chromatographic profiles of *Chuquiraga* samples and chemical structures of main metabolites. **A**) *C. jussieui* 1 (sample **1a**, Ecuador), *C. weberbaueri* (sample **10a**, Peru), *C. spinosa* (sample **20a**, Peru) and *C. jussieui* 2 (sample **30a**, Peru), and online UV spectra of main metabolites. **B**) Chemical structures of main metabolites.

HPLC chromatographic profiles associated to multivariate data analysis enabled the exploratory comparison and construction of a supervised model for species discrimination of three species *Chuquiraga* (*C. jussieui*, *C. spinosa*, and *C. weberbaueri*) and samples of separated areas of *C. jussieui* (*C. jussieui* 1 and 2). The flavonoid glycosides quercetin-3-*O*-glucoronide (**139**), kaempherol-3-*O*-rutoside (**150**), the acetophenone derivatives *p*-hydroxyacetophenone-4-*O*-glucoside (**58**), *p*-hydroxyacetophenone-4-*O*-(6-*O*-apiosyl)-glucoside (**67**), *p*-hydroxyacetophenone (**86**), a known caffeic acid ester derivative 5-*O*-caffeoylquinic acid (**74**), and a new one, 2-*O*-caffeoyl-4-hydroxypentanedioic acid (**75**) were recognized as main constituents of *Chuquiraga* species.

4.7. A new caffeic acid ester derivative

Compound **75** was obtained as an amorphous yellowish powder. The high-resolution electrospray ionization mass spectra (HRESIMS) of **75** showed a *pseudo*-molecular ion peak $[M - H]^-$ at *m/z* 325.0571, consistent with the molecular formula of $C_{14}H_{14}O_9$ (calculated for $C_{14}H_{13}O_9$, 325.0559). The online UV spectra of **75** displayed absorbance maxima at 233, 295 (shd), and 324 nm, resembling UV spectra of phenylpropanoid derivatives. The ^{1}H NMR spectrum of **75** (**Table 18**) showed characteristic signals for a pair of *trans*-olefinic protons at δ_H 7.62 and 6.36 (1H, d, J = 15.9 Hz, H-7' and H-8' respectively), and an aryl moiety with characteristic coupling pattern at δ_H 6.79 (1H, d, J = 8.1 Hz, H-5'), 6.97 (1H, dd, J = 8.1, 2.0 Hz, H-6'), and 7.08 (1H, d, J = 2.0 Hz, H-2') suggesting that **75** contains a caffeic acid moiety.

Table 18. NMR spectroscopic data for compound **75**.

	Compound 75			
Position	δ_C	type	δ_H, mult (J in Hz)	HMBC
1	173.2	C		
2	70.1	CH	5.27, dd (11.0, 2.7)	C-1, C-3, C-4, C-9'
3	36.9	CH_2	2.35, m (12.6, 11.0, 3.1)	C-2, C-5, C-1
			2.18, m (13.8, 10.6, 2.9)	C-4, C-5, C-1
4	67.8	CH	4.29, dd (10.5, 3.1)	C-2, C-3, C-5
5	177.1	C		
1'	127.7	C		
2'	115.3	CH	7.08, d (2.0)	C-3', C-6', C-7'
3'	146.7	C		
4'	149.7	C		
5'	116.5	CH	6.79, d (8.1)	C-1', C-3', C-4'
6'	123.2	CH	6.97, dd (8.1, 2.0)	C-2', C-3', C-7
7'	147.8	CH	7.62, d (15.9)	C-1', C-2', C-6', C-9'
8'	114.3	CH	6.36, d (15.9)	C-1', C-7', C-9'
9'	168.4	C		

Spectra recorded in methanol-d_4 at 300 MHz for ^{1}H and 75 MHz for ^{13}C

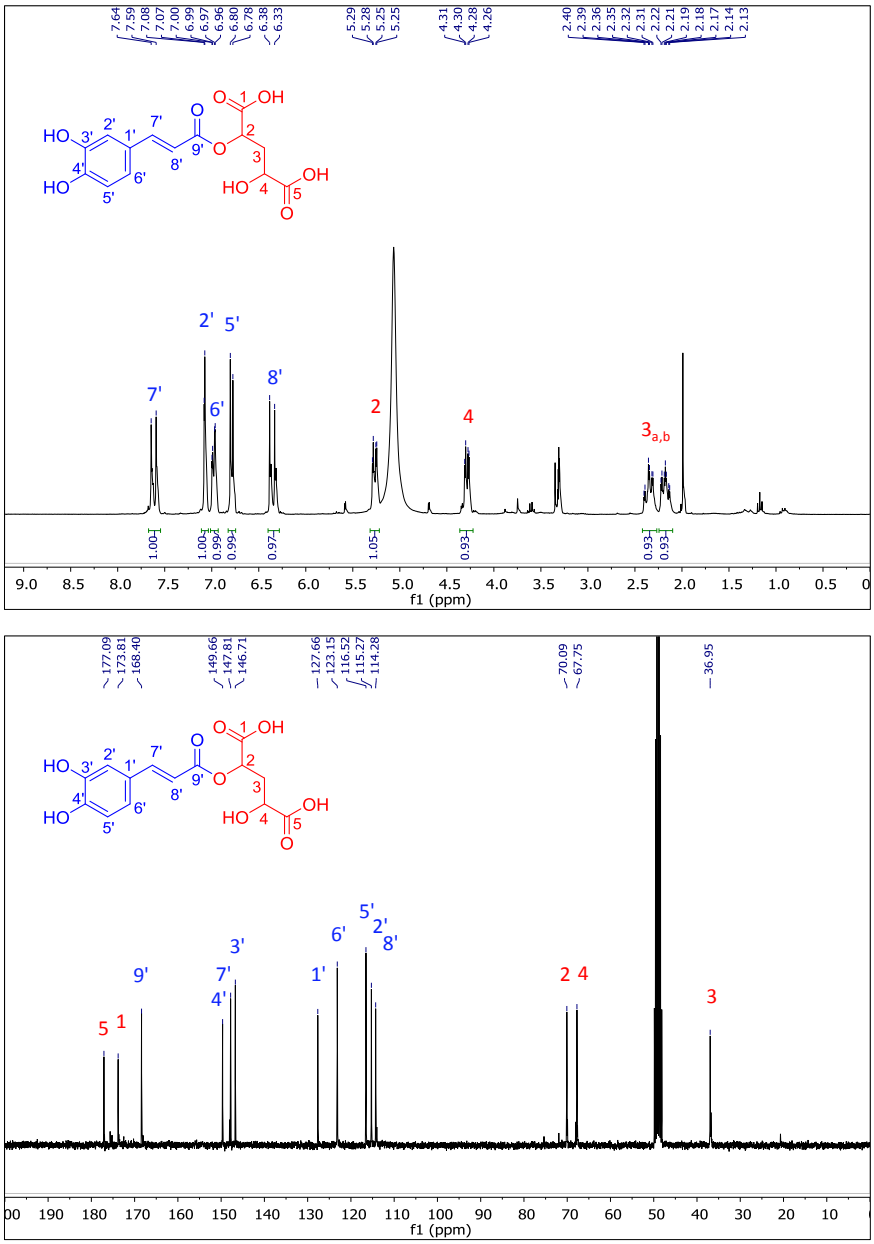


Figura 22. [1]H and [13]C NMR spectra of compounds **75** (300 MHz, methanol-d_4).

The presence of nine sp^2 carbon signals at δ_C 127.7 (C-1'), 115.3 (C-2'), 146.7 (C-3'), 149.7 (C-4'), 116.5 (C-5'), 123.2 (C-6'), 147.8 (C-7'), 114.3 (C-8') and 168.4 (C-9') in the ^{13}C NMR spectrum (**Figure 22**) confirmed the presence of a caffeic acid residue in **75**. Moreover, the ^{1}H NMR spectrum of **75** displayed signals at δ_H 5.27 (1H, dd, J =11.0, 2.7 Hz, H-2), 4.29 (1H, dd, J = 10.5, 3.1 Hz, H-4), 2.35 (1H, m, J = 12.6, 11.0, 3.1 Hz, H-3$_a$) and 2.18 (1H, m, J = 13.8, 10.6, 2.9 Hz, H-3$_b$) compatible with a 1,3-dimethine system. In addition, the signals in the ^{13}C NMR spectrum of two carbonyl units at δ_C 177.1 (C-1) and 173.8 (C-5), two aliphatic oxygen-bearing carbons at δ_C 70.1 (C-2) and 67.8 (C-3), and a methylene carbon signal at δ_C 36.9 (C-3) suggested a dihydroxypentanedioic acid moiety.

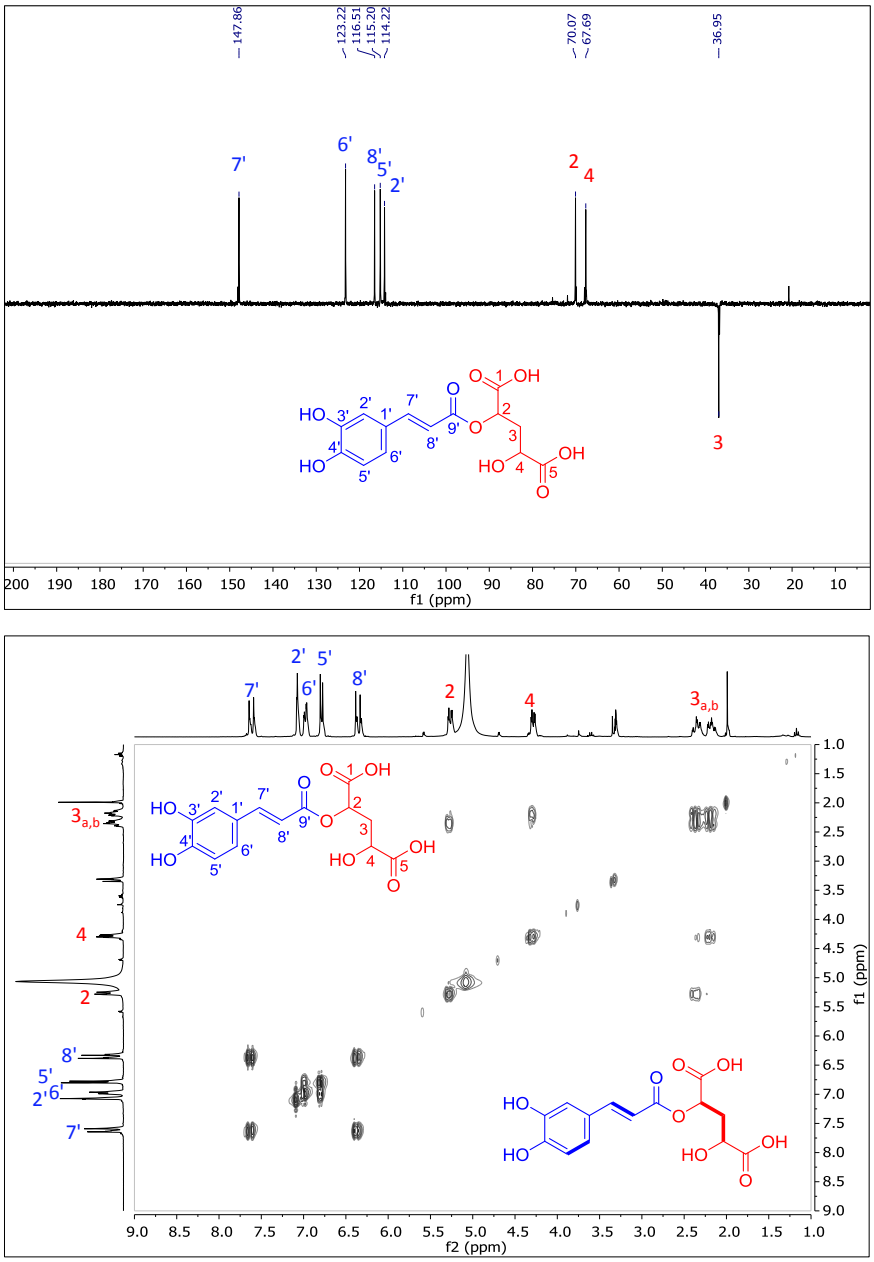


Figure 23. DEPT-135 and COSY spectra of compounds **75** (300 MHz, methanol-d_4).

The methylene and methine carbon signals assignments of the dihydroxypentanedioic acid were confirmed by the patterns displayed in the APT spectra (**Figure 23**). Furthermore, the homonuclear $^{3}J_{HH}$-coupling interaction displayed in the COSY spectra between the two protons of H-3$_{a,b}$ with H-2 or H-4 reinforced the -CH-CH2-CH- chain disposition and were in agreement with their corresponding coupling constants. Consequently, compound **75** is composed by a caffeic acid and a 2,4-dihydroxypentanedioic acid moiety.

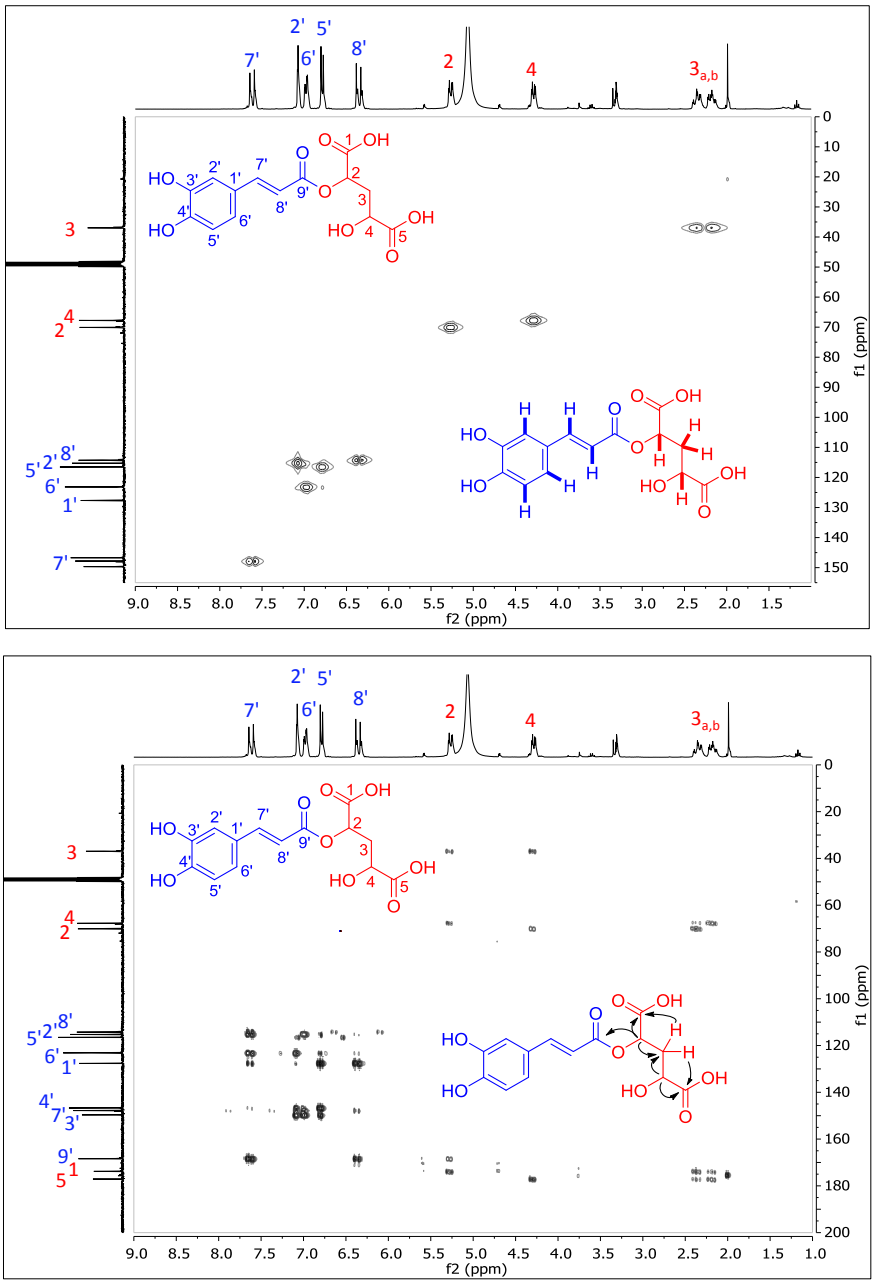


Figure 24. HSQC and HMBC spectra of compounds **75** (300 MHz, methanol-d_4).

The heteronuclear $^{1}J_{CH}$-coupling interactions displayed in the HSQC spectra confirmed the assignments of proton-carbon single bond correlations of both caffeic and dihydroxypentanedioic acids. The long-range heteronuclear $^{2\text{-}3}J_{CH}$-coupling interactions displayed in the HSQC spectra from H-2 towards C-1, C-3, C-4 and C-9' revealed the position of the caffeic acid residue at C-2 (**Figure 24**). Therefore, compound **75** constitutes a new caffeic acid ester derivative identified as 2-*O*-caffeoyl-4-hydroxypentanedioic acid. This compound present two stereogenic centers at C-2 and C-4.

4.8. Three more new caffeic acid ester derivatives

In addition to compound **47**, further minor constituents were isolated from a sample of *C. spinosa* leading to the identification of compounds **100**, **117** and **122**, whose chemical structures are displayed in **Figure 25**.

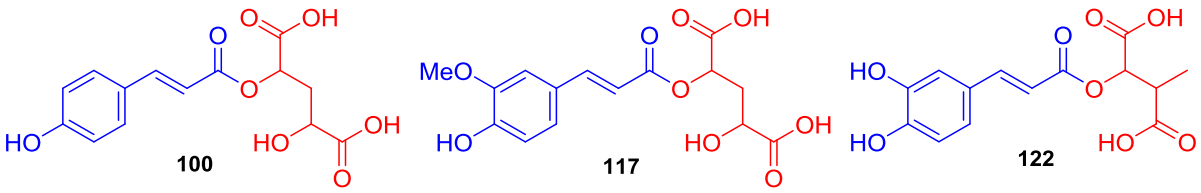


Figure 25. Chemical structures of new caffeic acid ester derivatives.

Compound **100** was obtained as an amorphous yellowish powder. The HRESIMS of **100** showed a *pseudo*-molecular ion peak $[M - H]^-$ at *m/z* 309.0618, consistent with the molecular formula of $C_{14}H_{14}O_8$ (calculated for $C_{14}H_{13}O_8$, 309.0610). The ^{1}H- and ^{13}C-NMR data of **100** (**Figure 26**, **Table 19**) exhibited signals corresponding to a 2,4-dihydroxypentanedioic acid moiety. Moreover, compound **100** displayed characteristic 1D- and 2D- NMR signal corresponding to a *p*-coumaric acid moiety attached to C-2. Therefore, compound **100** was identified as 2-*O*-coumaroyl-4-hydroxypentanedioic acid.

Table 19. NMR spectroscopic data for compounds **100**, **117** and **122**.

	100**		**117 ***		**122***	
Position	δ_C	δ_H, mult (*J* in Hz)	δ_C	δ_H, mult (*J* in Hz)	δ_C	δ_H, mult (*J* in Hz)
1	172.4		173.7		170.9	
2	69.1	5.23, d (10.9)	70.2	5.26, dd (11.1, 2.8)	72.6	5.57, d (3.5)
3	35.6	2.31, m (12.7)	37.0	2.35, m (12.5, 11.0, 3.1)	40.4	3.15, dd (7.0, 3.5)
		2.13, m (12.3)		2.17, m (12.3, 10.6, 3.0)		
4	66.3	4.23, d (9.9)	67.8	4.27, dd (10.5, 3.1)	174.4	
5	175.8		177.1			
1'	125.5		128.8		126.4	
2'	129.9	7.47, d (8.1)	114.9	7.11, d (1.8)	113.7	7.05, d (1.9)
3'	115.4	6.78, d (8.1)	151.7		144.5	
4'	159.9		148.1		148.2	
5'	115.4	6.78, d (8.1)	112.6	6.79, d (8.1)	115.2	6.78, d (8.1)
6'	129.9	7.47, d (8.1)	123.0	7.08, dd (7.8, 2.1)	121.7	6.96, dd (8.1, 2.0)
7'	146.1	7.65, d (15.8)	147.7	7.64, d (15.9)	146.1	7.57, d (16.0)
8'	112.8	6.39, d (15.9)	115.4	6.41, d (15.9)	113.1	6.30, d (16.0)
9'	166.9		168.2		167.1	
CH_3			56.4	3.89, s	10.0	1.27, d (7.2)

*Spectra recorded in methanol-d_4 at 300 MHz. ** Spectra recorded in methanol-d_4 at 500 MHz. δ_C of compounds **100** and **122** were deduced from HSQC and HMBC spectra.

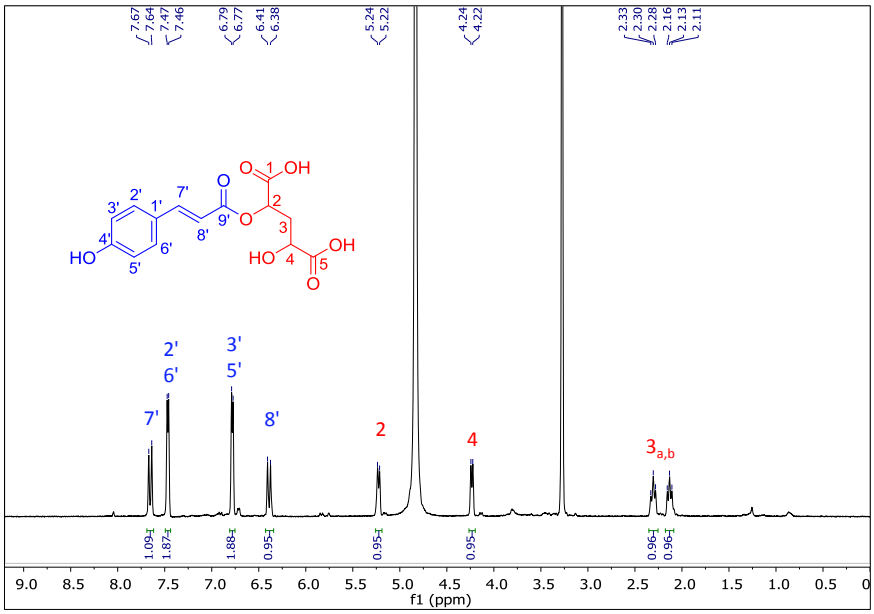


Figure 26. ^{1}H NMR spectra of compound **100** (500 MHz, methanol-d_4).

Compound **117** was obtained as an amorphous yellowish powder. The HRESIMS of **177** showed a *pseudo*-molecular ion peak $[M - H]^-$ at *m/z* 339.0720, consistent with the molecular formula of $C_{15}H_{16}O_9$ (calculated for $C_{15}H_{15}O_9$, 339.0716). The ^{1}H- and ^{13}C-NMR data of **117** (**Figure 27**, **Table 19**) resembled those data of 2-*O*-caffeoyl-4-hydroxypentanedioic acid. Differently, compound **117** displayed a characteristic singlet signal at δ_C 3.89 (δ_C 56.4) corresponding to a methoxy group attached to C-3'. Therefore, compound **117** was identified as 2-*O*-feruloyl-4-hydroxypentanedioic acid.

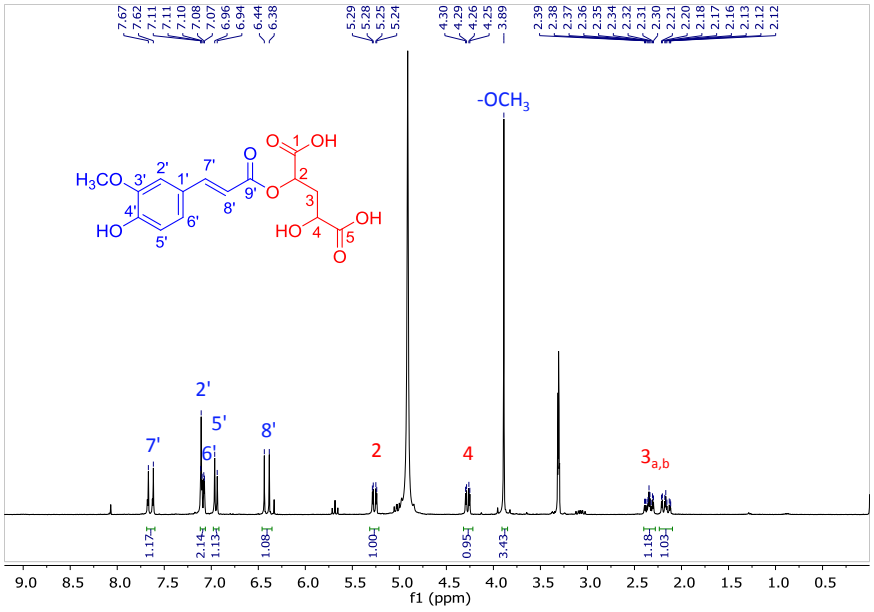


Figure 27. ^{1}H NMR spectra of compound **117** (300 MHz, methanol-d_4).

Compound **122** was obtained as an amorphous yellowish powder. The HRESIMS of **122** showed a *pseudo*-molecular ion peak $[M - H]^-$ at *m/z* 309.0618, consistent with the molecular formula of $C_{14}H_{14}O_8$ (calculated for $C_{14}H_{13}O_8$, 309.0610). The ^{1}H- and ^{13}C-NMR data of **122** (**Figure 28**, **Table 19**) exhibited signals corresponding to a caffeic acid moiety and a 2-hydroxy-3-methylbutanedioic acid moiety. The HMBC interaction of H-2 towards C-1, C-3, C-4 and C-9' confirmed the presence of the caffeoyl unit at C-2. Therefore, compound **122** was identified as 2-*O*-caffeoyl-4-methylbutanedioic acid.

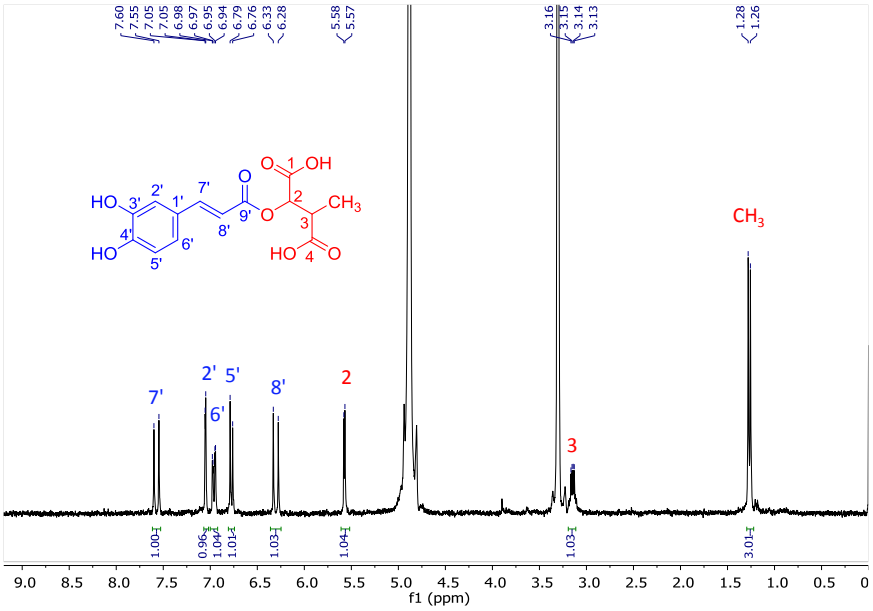


Figure 28. ^{1}H NMR spectra of compound **122** (300 MHz, methanol-d_4).

CONCLUSIONS

This work demonstrates the applicability of metabolomics for quality assessment of medicinal species of the genus *Chuquiraga* as well as a chemotaxonomy study of the subfamily Barnadesioideae.

The subfamily Barnadesioideae constitutes a chemically underinvestigated taxa with a complex diversity of phenolic compounds, alkyl glucosides, and triterpenoid glycosides. The intergeneric relationships of the subfamily Barnadesioideae based on their LC-MS metabolome data displayed similarities to those intergeneric relations proposed by the most recent phylogenetic study based on morphological and molecular markers, therefore showing metabolomics as a valuable auxiliary tool in chemotaxonomy studies.

Liquid chromatography associated to high-resolution mass spectrometry (LC-MS) and ultraviolet detection, constitutes a high-throughput platform in metabolomics studies for the quality assessment of medicinal plants. In addition to the phenolic compounds, the occurrence of phenylpropanoid derivatives of malic, tartaric and tartronic acids was recognized. Several flavonoid glycosides were also identified along with alkyl glycosides and triterpene glycosides. Exploratory and supervised multivariate analyses enabled geographical discrimination, species classification and identification of discriminating metabolites.

The HPLC chromatographic profiles of *Chuquiraga* species showed the occurrence of *p*-hydroxyacetophenone glycosides, flavonoids glycosides, and caffeic acid ester derivatives. Although the HPLC-DAD method is limited to the detection of chromophoric compounds, the multivariate analysis of HPLC chromatographic fingerprints enabled the establishment of a species classification model for *Chuquiraga* species.

REFERENCES

ABAD C, GONZÁLEZ J, CHAMORRO A (2009) El Apu Pariacaca y el Alto Cañete: Estudio de paisaje cultural. Instituto Nacional de Cultura, Lima

ALBERTO MR, ZAMPINI IC, ISLA MI (2009) Inhibition of cyclooxygenase activity by standardized hydroalcoholic extracts of four Asteraceae species from the Argentine Puna. Braz J Med Biol Res 42:787-790

ANDERSON H (1867) Notes on some of the Compositae of the Andes, and more particularly on *Chuquiraga insignis*. Trans & Proc Bot Soc Edinburgh 9:115-118

ANSALONI R, WILCHES I, LEÓN F et al (2010) Estudio preliminar sobre plantas medicinales utilizadas en algunas comunidades de las provincias de Azuay, Cañar y Loja, para afecciones del aparato gastrointestinal. Revista Tecnológica ESPOL 23:89-97

ARROYO-ACEVEDO J, HERRERA-CALDERÓN O, CHÁVEZ-ASMAT R et al (2017) Protective effect of *Chuquiraga spinosa* extract on N-methyl-N-nitrosourea (NMU) induced prostate cancer in rats. Prostate Int 5:47-52

ARROYO-ACEVEDO JL, HERRERA-CALDERON O, ROJAS-ARMAS JP et al. (2018) *Chuquiraga spinosa* Lessing: a medicinal plant for gastric cancer induced by N-methyl-N-nitrosourea (NMU). Pharmacogn J 10, 1:20-24.

BARBARÁN FR (2008) Medicinal plants of the Argentinean Puna: a common property resource and an opportunity for local people. In: Proceedings of the twelfth biennial conference of the international association for the study of commons. Indiana University, Cheltenham, 14-18 July, 2008

BEISKEN S, EIDEN M, SALEK RM (2015) Getting the right answers: understanding metabolomics challenges. Expert Rev Mol Diagn 15:197-109

BRACK A (1999) Diccionario enciclopédico de plantas útiles del Perú. Centro Bartolomé de Las Casas, Cusco

BREMER K, JANSEN RK (1992) A new subfamily of the Asteraceae. Ann Missouri Bot Gard 79:414-415

BOHM BA, STUESSY TF (1995) Flavonoid Chemistry of Barnadesioideae (Asteraceae). Syst Bot 20:22-27

BOHM BA, STUESSY TF (2001) Flavonoids of the sunfower family (Asteraceae). Springer-Verlag, Wien

BOHLMANN F, ZDERO C, SCHMEDA-HIRSCHMANN G et al (1986) Dimeric guaianolides and other constituents from *Gochnatia* species. Phytochemistry 25:1775-1178

BOUKTAIB M, ATMANI A, ROLANDO C (2002) Regio- and stereoselective synthesis of the major metabolite of quercetin, quercetin-3-*O*-β-d-glucuronide. Tetrahedron Lett 43:6263-6266

BUITRON XC (1999) Ecuador: uso y comercio de pantas medicinales, situación actual y aspectos importantes para su conservación. TRAFFIC International, Cambridge

BUSSMANN RW, SHARON D (2006a) Traditional plant use in Northern Peru: Tracking two thousand years of healing culture. J Ethnobiol Ethnomed 2:47

BUSSMANN RW, SHARON D (2006b) Traditional medicinal plant use in Loja province, Southern Ecuador. J Ethnobiol Ethnomed 2:44

BUSSMANN RW, SHARON D, VANDEBROEK I et al (2007) Health for sale: the medicinal plant markets in Trujillo and Chiclayo, Northern Peru. J Ethnobiol Ethnomed 3:37

BUSSMANN RW, SHARON D, PEREA FA et al (2008) Antibacterial activity of Northern-Peruvian medicinal plants. Arnaldoa 15:127-148

BUSSMANN RW, GLENN A (2010) Medicinal plants used in Northern Peru for reproductive problems and female health. J Ethnobiol Ethnomed 6:30

BUSSMANN RW, GLENN A, MEYER K et al (2010) Herbal mixtures in traditional medicine in Northern Peru. J Ethnobiol Ethnomed 6:10

BUSSMANN RW, MALCA G, GLENN A et al (2011) Toxicity of medicinal plants used in traditional medicine in Northern Peru. J Ethnopharmacol 137:121-140

BUSSMANN RW, SHARON D (2015) Medicinal plants of the Andes and the Amazon-The magic and medicinal flora of Northern Peru. William L. Brown Center, MBG, St. Louis.

BUSSMANN RW, PANIAGUA-ZAMBRANA N, CASTANEDA SIFUENTES RY et al (2015) Health in a pot-the ethnobotany of emolientes and emolienteros in Peru. Econ Bot 69:83-88

BUSSMANN RW, PANIAGUA ZAMBRANA NY, MOYA HUANCA LA, HART R (2016) Changing markets - Medicinal plants in the markets of La Paz and El Alto, Bolivia. J Ethnopharmacol 193:76-95

CABRERA AL (1951) *Huarpea*, nuevo género de Compuestas. Bol Soc Argent Bot 4:129-132

CABRERA AL (1959) Revisión del género *Dasyphyllum* (Compositae). Rev Mus La Plata 9(38): 21-100

CABRERA AL (1997) Nota crítica en la tribu Mutisieae (Compositae) para la flora de Paraguay. Candollea 52:216

CHAGAS-PAULA DA, OLIVEIRA TB, FALEIRO DP, OLIVEIRA RB, COSTA FB (2015) Outstanding anti-inflammatory potential of selected Asteraceae species through the potent dual inhibition of cyclooxygenase-1 and 5-lipoxygenase. Planta Med 81:1296-1307

CAJKA T, FIEHN O (2016) Toward merging untargeted and targeted methods in mass spectrometry-based metabolomics and lipidomics. Anal Chem 88:524-545

CALABRIA LM, EMERENCIANO VP, FERREIRA MJP et al (2007) A phylogenetic analysis of tribes of the Asteraceae based on phytochemical data. Nat Prod Commun 2:277-285

CALABRIA LM, EMERENCIANO VP, SCOTTI MT, MABRY TJ (2009) Secondary chemistry of Compositae. In: Funk V, Susanna A, Stuessy TF, Robinson H (ed), Compositae: Systematics, Evolution, and Biogeography of Compositae. International Association for Plant Taxonomy, Vienna

CAMAQUI AM (2007) Plantas medicinales. La experiencia de Tinguipaya. Editorial Gente Común, La Paz

CAMPOS-NAVARRO R, SCARPA GF (2013) The cultural-bound disease "empacho" in Argentina. A comprehensive botanico-historical and ethnopharmacological review. J Ethnopharmacol 148:349-60

CARULLO G, CAPPELLO AR, FRATTARUOLO L et al (2017) Quercetin and derivatives: useful tools in inflammation and pain management. Future Med Chem 9:79-93

CASADO R, LANDA A, CALVO J et al (2011) Anti-inflammatory, antioxidant and antifungal activity of *Chuquiraga spinosa*. Pharm Biol 49:620-626

CASTELUCCI S, DE PAULA ROGERIO A, AMBROSIO SR et al (2007) Anti-inflammatory activity of *Dasyphyllum brasiliensis* (Asteraceae) on acute peritonitis induced by beta-glucan from *Histoplasma capsulatum*. J Ethnopharmacol 112:192-198

CEUTERICK M, VANDEBROEK I, PIERONI A (2011) Resilience of Andean urban ethnobotanies: a comparison of medicinal plant use among Bolivian and Peruvian migrants in the United Kingdom and in their countries of origin. J Ethnopharmacol 136:27-54

CHING-WEN C, YUN-CHIEH C, YU-CHIN L, WEN-HUANG P (2017) *p*-Hydroxyacetophenone suppresses nuclear factor-κB-related inflammation in nociceptive and inflammatory animal models. J Nat Med 71(2):422-432

COLLINS J (1870) Notes on some new little-known vegetable products. Pharm J Trans 11:66-67

CREEK DJ, DUNN WB, FIEHN O, et al. (2014) Metabolite identification: are you sure? And how do your peers gauge your confidence? Metabolomics 10:350-353

DA COSTA FB, SCHORR K, ARAKAWA NS et al (2001) Infraspecific variation in the chemistry of glandular trichomes of two Brazilian *Viguiera* species (Heliantheae; Asteraceae). J Braz Chem Soc 12: 403-407

DA COSTA FB, TERFLOTH L, GASTEIGER J (2005) Sesquiterpene-lactone based classification of three Asteraceae tribes: a study based on self-organizing neural networks applied to chemosystematics. Phytochemistry 66:345-353

DE FEO V (2003) Ethnomedical field study in northern Peruvian Andes with particular reference to divination practices. J Ethnopharmacol 85:243-256

DE KRAKER JW, FRANSSEN MC, DALM MC et al (2001) Biosynthesis of germacrene A carboxylic acid in chicory roots. Demonstration of a cytochrome P450 (+)-germacrene a hydroxylase and NADP+-dependent sesquiterpenoid dehydrogenase(s) involved in sesquiterpene lactone biosynthesis. Plant Physiol 125:1930-1940

DE KRAKER JW, FRANSSEN MC, JOERINK M, et al (2002) Biosynthesis of costunolide, dihydrocostunolide, and leucodin. Demonstration of cytochrome P450-catalyzed formation of the lactone ring present in sesquiterpene lactones of chicory. Plant Physiol 129:257-268

DE-LA-CRUZ H, VILCAPOMA G, ZEVALLOS PA (2007) Ethnobotanical study of medicinal plants used by the Andean people of Canta, Lima, Peru. J Ethnopharmacol 111:284-294

DE LA TORRE L, ALARCÓN DS, KVIST LP, LECARO JS (2008) Usos medicinales de las plantas. In: de la Torre L, Navarrete H, Muriel PM et al (eds.) Enciclopedia de las plantas útiles del Ecuador. Herbario QCA & Herbario AAU, Quito & Aarhus

DE MÖSBACH EW (1991) Botánica Indígena de Chile. Editorial Andrés Bello, Santiago de Chile

DÍAZ-PIEDRAHITA S, VÉLEZ-NAUER C (1993) Revisión de las tribus Barnadesieae y Mutisieae (Asteraceae) para la flora de Colombia. Jardín Botánico José Celestino Mutis, Bogotá

DUKE JA, BOGENSCHUTZ-GODWIN MJ, OTTENSEN AR (2009) Duke's handbook of medicinal plants of Latin America. CRC Press, Boca Raton

DUENAS AA, ALCIVAR UE, OLAZABAL E, CORTES R (2014) Efecto antioxidante de la *Chuquiraga jussieui* J.F.Gmel en el ensayo de hemólisis. Medicent Electrón 18:57-64

ERDTMAN H (1963) Some aspects of chemotaxonomy. Pure Appl Chem 6:679-708

ERNST M, SILVA DB, SILVA RR, VÊNCIO RZ, LOPES NP (2014) Mass spectrometry in plant metabolomics strategies: from analytical platforms to data acquisition and processing. Nat Prod Rep 31:784-806

ESTOMBA D, LADIO A, LOZADA M (2006) Medicinal wild plant knowledge and gathering patterns in a Mapuche community from North-western Patagonia. J Ethnopharmacol 103:109-119

EZCURRA C (1985) Revisión del género *Chuquiraga* (Compositae - Mutisieae). Darwiniana 26:219-284

FERREIRA PL, GROPPO M (2018). Phylogeny and circumscription of *Dasyphyllum* (Asteraceae: Barnadesioideae) based on molecular data with the recognition of a new genus, *Archidasyphyllum*. Plos One, *submitted*

FIEHN O (2002) Metabolomics-the link between genotypes and phenotypes. Plant Mol Biol 48:155-171

FIEHN O (2016) Metabolomics by gas chromatography-mass spectrometry: combined targeted and untargeted profiling. Curr Protoc Mol Biol. 114:30.4.1-30.4.32.

FLAGG ML, VALCIC S, MONTENEGRO G et al (1999) Pentacyclic triterpenes from *Chuquiraga ulicina*. Phytochemistry 52:1345-1350

FORINO M, TENORE GC, TARTAGLIONE L et al. (2015) (1*S*,3*R*,4*S*,5*R*)-5-*O*-Caffeoylquinic acid: isolation, stereo-structure characterization and biological activity. Food Chem 178:306-310

FUNK VA, BAYER RJ, KEELEY S et al (2005) Everywhere but Antarctica: using a supertree to understand the diversity and distribution of the Compositae. Biol Skr 55:343-374

FUNK VA, ROQUE N (2011) The monotypic Andean genus *Fulcaldea* (Compositae, Barnadesioideae) gains a new species from northeastern Brazil. Taxon 60:1095-1103

GALLON ME, MONGE M, CASOTI R, DA COSTA FB, SEMIR J, GOBBO-NETO L (2018) Metabolomic analysis applied to chemosystematics and evolution of megadiverse Brazilian Vernonieae (Asteraceae). Phytochemistry 150:93-105

GÁLVEZ M, PASTOR A (1996) Estudio fitoquímico de la *Chuquiraga spinosa*. Revista de Química 10:133-134

GIBERTI GC (1983) Herbal folk medicine in northwestern Argentina: Compositae. J Ethnopharmacol 7:321-341

GIOTI K, TENTA R (2015) Bioactive natural products against prostate cancer: mechanism of action and autophagic/apoptotic molecular pathways. Planta Med 81:543-562

GOODACRE R, VAIDYANATHAN S, DUNN WB, HARRIGAN GG, KELL DB (2004) Metabolomics by numbers: acquiring and understanding global metabolite data. Trends Biotechnol 22:245-252

GOODACRE R (2005) Metabolomics - The way forward. Metabolomics 1:1-2

GORROCHATEGUI E, JAUMOT J, LACORTE S, TAULER R (2016) Data analysis strategies for targeted and untargeted LC-MS metabolomic studies: overview and workflow. Trends Anal Chem 82:425-442

GRANADA A (1997) Una nueva especie de *Chuquiraga* (Asteraceae-Mutisieae) del Perú. Kurtziana 25:151-156

GRUENSTAEUDL M, URTUBEY E, JANSEN RK et al (2009) Phylogeny of Barnadesioideae (Asteraceae) inferred from DNA sequence data and morphology. Mol Phylogenet Evol 51:572-587

GUPTA MP (2006) Medicinal plants originating in the Andean high plateau and central valleys region of Bolivia, Ecuador and Peru. UNIDO Report, United Nations

GUROVIC MS, CASTRO MJ, RICHMOND V et al (2010) Triterpenoids with acetylcholinesterase inhibition from *Chuquiraga erinacea* D. Don. subsp. *erinacea* (Asteraceae). Planta Med 76:607-610

HARLING G (1991) Compositae-Mutisieae. In: Harling G, Andersson L (ed) Flora of Ecuador, vol 42. University of Götenborg, Göteborg

HEINRICH M (2008) Ethnopharmacy and natural product research - Multidisciplinary opportunities for research in the metabolomics age. Phytochem Lett 1:1-5

HERRERA FL (1933) Plantarum Cuzcorum Herrerarianum. Estudios sobre la flora del departamento del Cuzco. Sanruarti, Lima

HERRERA FL (1938) Plantas que curan y plantas que matan de la flora del Cuzco. Revista Universitaria 75:4-76

HERRERA-CALDERON O, TINCO-JAYO JA, FRANCO-QUINO C et al (2017) Antioxidant activity and cytotoxic profile of *Chuquiraga spinosa* Lessing on human tumor cell lines: A promissory plant from Peruvian flora. Asian Pac J Trop Dis 7: 304-308

HIND DJN (2001) A New Species of *Barnadesia* (Compositae: Barnadesieae) from Bolivia. Kew Bull 56:705-710

HIND N, Hall T (2003) Plate 459. *Barnadesia arborea* Compositae. Curtis's Bot Mag 20:25-30

HOENEISEN M, ROJAS A, BITTNER M et al (2000) Constituents of *Chuquiraga atacamensis* and *C. ulicina*. Bol Soc Chil Quim 45:49-52

HUAMANTUPA I, CUBA M, URRUNAGA R et al (2011) Riqueza, uso y origen de plantas medicinales expendidas en los mercados de la ciudad del Cusco. Rev Peru Biol 18:283-291

JANSEN RK, PALMER JD (1987) A chloroplast DNA inversion marks an ancient evolutionary split in the sunflower family (Asteraceae). Proc Natl Acad Sci U S A 84:5818-5822

JADÁN MB, ORQUERA GX, MIHAI RA (2014) Establishment of an in vitro culture protocol of *Chuquiraga jussieui* J.F.Gmel. from apical and axillary buds. Rom Biotech Lett 19:9984-9991

JEON SH, CHUN W, CHOI YJ, KWON YS (2008) Cytotoxic constituents from the bark of *Salix hulteni*. Arch Pharm Res 31:978-982

JUÁREZ BE, MENDIONDO ME (2002a) Flavonoides en *Chuquiraga acanthophylla* Weddell subfamilia Barnadesioideae (Asteraceae). In: I Congreso Latinoamericano de Fitoquímica, IV Reunion de la sociedad Latinoamericana de Fitoquimica, Buenos Aires, 8-10 May 2002

JUÁREZ BE, MENDIONDO ME (2002b) Flavonoid chemistry of *Chuquiraga* (Asteraceae). Biochem Syst Ecol 30:371-373

JUÁREZ BE, MENDIONDO ME (2007) Significado quimiotaxonomico de los flavonoides presentes en *Doniophyton anomallum* (D. Don) Kurtz (Asteraceae). B Latinoam Caribe PL 6:252-253

KATINAS L, STUESSY TF (1997) Revision of *Doniophyton* (Compositae, Barnadesioideae). Plant Syst Evol 206:33-45

KIM K-J, CHOI K-S, JANSEN RK (2005) Two chloroplast DNA inversions originated simultaneously during the early evolution of the sunflower family (Asteraceae). Mol Biol Evol 22:1783-1792

LANDA A, CASADO R, CALVO MI (2009) Identification and quantification of flavonoids from *Chuquiraga spinosa* (Asteraceae). Nat Prod Commun 4:1353-1355

LADIO AH, LOZADA M (2009) Human ecology, ethnobotany and traditional practices in rural populations inhabiting the Monte region: Resilience and ecological knowledge. J Arid Environ 73:222-227

LUNDBERG J (2009) Asteraceae and relationships within Asterales. In: Funk V, Susanna A, Stuessy TF, Robinson H (ed), Compositae: Systematics, Evolution, and Biogeography of Compositae. International Association for Plant Taxonomy, Vienna

MACÍA MJ, GARCÍA E, VIDAURRE PJ (2005) An ethnobotanical survey of medicinal plants commercialized in the markets of La Paz and El Alto, Bolivia. J Ethnopharmacol 97:337-350

MADALENO IM (2007) Etno-farmacología en Iberoamérica, una alternativa a la globalización de las prácticas de cura. Cuadernos Geográficos 41:61-95

MADALENO IM (2012) Organic cultivation and use of medicinal plants in Latin America. Phcog Commn 2:34-51

MARTINS LR, PEREIRA-FILHO ER, CASS QB (2011) Chromatographic profiles of *Phyllanthus* aqueous extracts samples: a proposition of classification using chemometric models. Anal Bioanal Chem 400:469-481

MARTÍNEZ CEC (2006) Plantas medicinales de los Andes Ecuatorianos. In: Moraes MR, Øllgaard B, Kvist LP et al (ed) Botánica Económica de los Andes Centrales. Universidad Mayor de San Andrés, La Paz

MARTUCCI MEP, DE VOS RCH, CAROLLO CA, GOBBO-NETO L (2014) Metabolomics as a potential chemotaxonomical tool: application in the genus Vernonia Schreb. Plos One. 9(4):e93149

MENDIONDO ME, JUAREZ BE, SEELIGMANN P (1997) Flavonoid patterns of some Barnadesioideae (Asteraceae). Eventual chemosystematic significance. Biochem Syst Ecol 25:673-674

MENDIONDO ME, JUÁREZ BE, SEELIGMANN P (2000) Flavonoid profiles of some Argentine species of *Chuquiraga* (Asteraceae). Biochem Syst Ecol 28:283-285
MENDIONDO ME, JUÁREZ BE (2001) Flavonoids of *Doniophyton patagonicum* (Phil.) Hieron. (Asteraceae). Biochem Syst Ecol 29:437-438
MENDIONDO ME, JUÁREZ BE, ZAMPINI C et al (2011) Bioactivities of *Chuquiraga straminea* Sandwith. Nat Prod Commun 6:965-968
MESSINA A, CALLAHAN DL, WALSH NG, HOEBEE SE, GREEN PT (2014) Testing the boundaries of closely related daisy taxa using metabolomic profiling. Taxon 63:367-376
MONIGATTI M, BUSSMANN RW, WECKERLE CS (2012) Medicinal plant use in two Andean communities located at different altitudes in the Bolívar Province, Peru. J Ethnopharmacol 145:450-464
MOSTACERO J, CASTILLO F, MEJIA FR et al (2011) Plantas medicinales del Perú. Taxonomía, ecogeografía, fenología y etnobotánica. Asamblea Nacional de Rectores, Lima
NGUYEN DT, GÖPFERT JC, IKEZAWA N et al (2010) Biochemical conservation and evolution of germacrene A oxidase in Asteraceae. J Biol Chem 285:16588-16598
NGUYEN TD, FARALDOS JA, VARDAKOU M et al (2016) Discovery of germacrene A synthases in *Barnadesia spinosa*: The first committed step in sesquiterpene lactone biosynthesis in the basal member of the Asteraceae. Biochem Biophys Res Commun 479:622-627
OLIVER SG, WINSON MK, KELL DB, BAGANZ F. (1998) Systematic functional analysis of the yeast genome. Trends Biotechnol 16:373-378
OKADA T, AFENDI FM, ALTAF-UL-AMIN M, et al. (2010) Metabolomics of medicinal plants: the importance of multivariate analysis of analytical chemistry data. Curr Comput Aided Drug Des 6:179-196
PANERO JL, FUNK VA (2008) The value of sampling anomalous taxa in phylogenetic studies: major clades of the Asteraceae revealed. Mol Phylogenet Evol 47:757-782
PASSONI FD, CAROLLO C, GOBBO-NETO L et al (2008) Polifenóis na fração ativa de espinho-agulha (*Dasyphyllum brasiliense*, Asteraceae), uma planta medicinal com atividade antiinflamatória. In: 31ª Reunião Anual da Sociedade Brasileira de Química, Águas de Lindóia, 26-29 May 2008
PAULA CS, VERDAM MCS, SOUZA AM et al (2013) Prospecção fitoquímica e avaliação preliminar da atividade antibacteriana dos extratos das folhas e casca do caule de *Dasyphyllum tomentosum* (Spreng.) Cabrera. Visão Acadêmica 14:4-12
PADILLA-GONZALEZ GF, DOS SANTOS FA, DA COSTA FB (2016) Sesquiterpene lactones: more than protective plant compounds with high toxicity. Crit Rev Plant Sci 35:18-37
PIETTA PG (2000) Flavonoids as antioxidants. J Nat Prod 63:1035-1042
QUATTROCCHI U (2012) CRC world dictionary of plant names: common names, scientific names, eponyms, synonyms, and etymology. CRC Press, New York
RAAD K (2012) Medicina ancestral de los Amaychas. Amagrafik, Tucuman
RAMÍREZ C, BELOSO C (2002) Usos tradicionales de las plantas en la Meseta Patagónica. Jardín Botánico de la Patagonia Extraandina. CENPAT-CONICET-ICBG. Dirección de Impresiones Oficiales, Chubut
REHECHO S, URIARTE-PUEYO I, CALVO J et al (2011) Ethnopharmacological survey of medicinal plants in Nor-Yauyos, a part of the landscape reserve Nor-Yauyos-Cochas, Peru. J Ethnopharmacol 133:75-85
REYNOLDS T (2007) The evolution of chemosystematics. Phytochemistry 68:2887-2895
RIBEIRO D, FREITAS M, TOMÉ SM et al (2015) Flavonoids inhibit COX-1 and COX-2 enzymes and cytokine/chemokine production in human whole blood. Inflammation 38:858-870
RICHERI M, LADIO AH, BEESKOW AM (2013) Conocimiento tradicional y autosuficiencia: la herbolaria rural en la meseta central del Chubut (Argentina). B Latinoam Caribe PL 12:44-58
ROBERTS LD, SOUZA AL, GERSZTEN RE, CLISH CB (2012) Targeted metabolomics. Curr Protoc Mol Biol. Chapter 30:Unit 30.2.1-24
ROCHAT B (2016) From targeted quantification to untargeted metabolomics: Why LC-high-resolution-MS will become a key instrument in clinical lab. Trends Anal Chem 84:151-164
ROCHFORT S (2005) Metabolomics reviewed: a new "omics" platform technology for systems biology and implications for natural products research. J Nat Prod 68:1813-1820
ROERSCH C (1994) Plantas medicinales en el sur andino del Peru. Koeltz Scientific Books, Koenigstein

RONDINA RVD, BANDONI AL, COUSSIO JD (2008) Especies medicinales argentinas con potencial actividad analgésica. Dominguezia 24:47-69

SAAVEDRA MM, MONGE M, GUIMARÃES EF (2014) *Dasyphyllum diamantinense* (Asteraceae, Barnadesioideae): a new species from the Chapada Diamantina, Bahia State, Brazil. Phytotaxa 174:231-236

SAGÁSTEGUI AA (1980) Compuestas andino-peruanas nuevas para la ciencia. Bol Soc Argent Bot 19:61-68

SAGÁSTEGUI AA, DILLON MO (1985) Four new species of Asteraceae from Peru. Brittonia 37:6-13

SAGÁSTEGUI AA, SÁNCHEZ VL (1991) Una nueva especie de *Chuquiraga* (Asteraceae-Mutisieae) del norte del Perú. Arnaldoa 1:1-4

SALA A, RECIO MC, GINER RM et al (2001) New acetophenone glucosides isolated from extracts of *Helichrysum italicum* with antiinflammatory activity. J Nat Prod 64:1360-1362

SEAMANN FC (1982) Sesquiterpene lactones as taxonomic characters in the Asteraceae. Bot Rev 48:121-594

SENATORE F (1996) Composition of the essential oil of *Chuquiraga spinosa* (R. et P.) D. Don. Flav Frag J 11:215-217

SENATORE F, NUNZIATA A, D'AGOSTINO M, DE FEO V (1999) Flavonol glycosides and *p*-hydroxyacetophenone from *Chuquiraga spinosa*. Pharm Biol 37:366-368

SHAHIDI F, CHANDRASEKARA A (2010) Hydroxycinnamates and their in vitro and in vivo antioxidant activities. Phytochem Rev 9:147-170

SIURA SC, FLORES MP (2010) Etnobotánica de las plantas medicinales de las comunidades campesinas de Quero y Masma Chicche. In: Gallo MP, Galarza VG, Gabriel JM, Moris G (ed) Las plantas medicinales del Perú: etnobotánica y viabilidad comercial. Los libros de la Catarata, Madrid

SOUBEIRAN L (1868) Extrait du proces-verbal de la séance de la Société de pharmaie de Paris. Journal de Pharmacie et de Chimie 4(8):303

SPRING O (1989) Microsampling: An alternative approach using sesquiterpene lactones for systematics. Biochem Syst Ecol 17:509-517

SPRING O (2000) Chemotaxonomy based on metabolites from glandular trichomes. In: Hallahan DL, Gray JC (ed), Plant Trichomes, Advances in Botanical Research, vol 31, Academic Press, London, pp153-174

STUESSY TF, SAGÁSTEGUI AA (1993) Revisión de *Arnaldoa* (Compositae, Barnadesioideae), género endémico del Norte del Perú. Arnaldoa 1:9-21

STUESSY TF, URTUBEY E, GRUENSTAEUDL M (2009) Barnadesieae (Barnadesioideae). In: Funk V, Susanna A, Stuessy TF, Robinson H (ed), Compositae: Systematics, Evolution, and Biogeography of Compositae. International Association for Plant Taxonomy, Vienna

SUMNER LW, AMBERG A, BARRETT D, et al. (2007) Proposed minimum reporting standards for chemical analysis Chemical Analysis Working Group (CAWG) Metabolomics Standards Initiative (MSI). Metabolomics 3:211-221

TENE V, MALAGÓN O, FINZI PV et al (2007) An ethnobotanical survey of medicinal plants used in Loja and Zamora-Chinchipe, Ecuador. J Ethnopharmacol 111:63-81

TORRES H, BOREL R, BUSTAMANTE N, CENTENO MI (1992) Usos tradicionales de arbustos nativos en el sur de Puno. Publifor, Puno

ULLOA ULLOA C, JØRGENSEN PM, DILLON MO (2002) *Arnaldoa argentea* (Barnadesioideae: Asteraceae) a new species and a new generic record for Ecuador. Novon 12:415-419

ULRICH-MERZENICH G, ZEITLER H, JOBST D, PANEK D, VETTER H, WAGNER H (2007) Application of the "-Omic-" technologies in phytomedicine. Phytomedicine 14:70-82

URTUBEY E (1999) Revisión del género *Barnadesia* (Asteraceae: Barnadesioideae, Barnadesieae). Ann Missouri Bot Gard 86:57-117

VANDEBROEK I, THOMAS E, AMETRAC (2003) Plantas medicinales para la atención primaria de la salud. El conocimiento de ocho médicos tradicionales de Apillapampa (Bolivia). Industrias Graficas Serrano, Cochabamba

VÁSQUEZ LN, ESCURRA JP, AGUIRRE RT et al (2010) Plantas medicinales del Norte del Perú. Universidad Nacional Pedro Ruiz Gallo, Lambayeque

VILLAGRÁN C, CASTRO V, SÁNCHEZ G et al (1998) La tradición surandina del desierto: Etnobotánica del área del Salar de Atacama (Provincia de El Loa, Región de Antofagasta, Chile). Estudios Atacameños 16:7-105

VILLAGRÁN C, ROMO M, CASTRO V (2003) Etnobotánica del sur de los Andes de la primera región de Chile: un enlace entre las culturas altiplánicas y las de quebradas altas del Loa Superior. Chungara 35:73-124

VILLAS-BÔAS SG, MAS S, AKESSON M, SMEDSGAARD J, NIELSEN J (2005) Mass spectrometry in metabolome analysis. Mass Spectrom Rev 24:613-646

VINAYAVEKHIN N, SAGHATELIAN A (2010) Untargeted metabolomics. Curr Protoc Mol Biol. Chapter 30:Unit 30.1.1-24.

WAGNER H, BAUER R, MELCHART D, STAUDINGER A (2011) Chromatographic fingerprint analysis of herbal medicines v 1-2. Springer, Wien

WAGNER H, BAUER R, MELCHART D, STAUDINGER A (2015) Chromatographic fingerprint analysis of herbal medicines v 3. Springer, Wien

WAGNER H, BAUER R, MELCHART D, STAUDINGER A (2016) Chromatographic fingerprint analysis of herbal medicines v 4. Springer, Wien

WANG L, WANG X, KONG L (2012) Automatic authentication and distinction of *Epimedium koreanum* and *Epimedium wushanense* with HPLC fingerprint analysis assisted by pattern recognition techniques. Biochem Syst Ecol 40:138-145

WANG M, SHAO Y, LI J, ZHU N, RANGARAJAN M, LAVOIE EJ, HO CT (1999) Antioxidative phenolic glycosides from sage (*Salvia officinalis*). J Nat Prod 62:454-456

WINK M, WATERMAN P (1999) Chemotaxonomy in relation to molecular phylogeny of plants. Annual Plant Reviews 2:300-341

WINK M, BOTSCHEN F, GOSMANN C, SCHÄFER H, WATERMAN PG (2010) Chemotaxonomy seen from a phylogenetic perspective and evolution of secondary metabolism. Annual Plant Reviews 40:364-433

WOLFENDER JL, MARTI G, QUEIROZ EF (2010) Advances in techniques for profiling crude extracts and for the rapid identification of natural products: dereplication, quality control and metabolomics. Curr Org Chem 14:1808-1832

YAKOVLEFF E, HERRERA FH (1934) El mundo vegetal de los antiguos Peruanos. Rev Mus Nac/Lima 3:240-322

YANG F, SONG L, WANG H et al (2015) Quercetin in prostate cancer: Chemotherapeutic and chemopreventive effects, mechanisms and clinical application potential. Oncol Rep 33:2659-2668

ZAMORA VHC (2008) Estudio de aproximaciones etnobotánicas en áreas productoras del intersalar de Quinua Real del departamento de Potosi (Parte I). Fundacion Alitapo, Potosi

ZAMPINI IC, CUDMANI N, ISLA MI (2007) Actividad antimicrobiana de plantas medicinales argentinas sobre bacterias antibiótico-resistentes. Acta Bioquim Clin L 41:385-393

ZAMPINI IC, CUELLO S, ALBERTO MR et al (2009) Antimicrobial activity of selected plant species from "the Argentine Puna" against sensitive and multi-resistant bacteria. J Ethnopharmacol 124:499-505

ZAMPINI IC, ORDOÑEZ RM, ISLA MI (2010) Autographic assay for the rapid detection of antioxidant capacity of liquid and semi-solid pharmaceutical formulations using ABTS•+ immobilized by gel entrapment. AAPS PharmSciTech 11:1159-1163

ZARDINI EM, SORIA N (1994) A new species of *Dasyphyllum* (Asteraceae–Mutisieae) from Paraguay. Novon 4:80-82

ZDERO C, BOHLMANN F, KING RM, ROBINSON H (1986a) Further 5-methyl coumarins and other constituents from the subtribe mutisiinae. Phytochemistry 25:509-516

ZDERO C, BOHLMANN F, KING RM, ROBINSON H (1986b) α-Isocedrene derivatives, 5-methyl coumarins and other constituents from the subtribe Nassauviinae of the Compositae. Phytochemistry 25:2873-2882

ZDERO C, BOHLMANN F, KING RM (1987) Chemistry of the Barnadesiinae (Asteraceae). Phytologia 63:313-315

ZDERO C, BOHLMANN F (1990) Systematics and evolution within the Compositae, seen with the eyes of a chemist. Plant Syst Evol 171:1–14

APPENDIX 1. Steps for using the ProteoWizard toolkit

A. Install the latest version of **ProteoWizard** (http://proteowizard.sourceforge.net/). Reading and conversion of vendor formatted files only works on Windows OS. There are reports on the successful use of **ProteoWizard** in Linux OS by using "*wine emulator*"; even though, the use of a Windows virtual machine on Linux installed PCs may be more suitable for those Linux OS users. Here we are using the Windows 64-bit installer (able to convert vendor files except T2D): pwiz-setup-3.0.9798-x86_64.

Copy the destination folder where ProteoWizard 3.0.9798 is installed:

```
C:\Program Files\ProteoWizard\ProteoWizard 3.0.9798\
```

Set the path to add ProteoWizard. To change the system environment variables in Windows 7 and 10, follow the below steps:

1. From the desktop, right-click **My Computer** and click **Properties**.
2. In the **System Properties** window, click on the **Advanced** tab.
3. In the **Advanced section**, click the **Environment Variables** button.
4. Finally, in the **Environment Variables** window, highlight the **Path** variable in the **Systems Variable** section and click the **Edit** button. Add the ProteoWizard path [C:\Program Files\ProteoWizard\ProteoWizard 3.0.9798\].
5. Click OK to authenticate the Path changes.

B. In the folder that contains the original LC-MS data, create three text (.txt) documents and two folders.

Text 1, save it as **configneg**:

```
mzXML=true
zlib=true
filter="polarity negative"
filter="threshold absolute 10000 most-intense"
filter="scanTime [0.01,1800]"
```

*The threshold and scantime filters must be changes as necessary.

Text 2, save it as **configpos**:

```
mzXML=true
zlib=true
filter="polarity positive"
filter="threshold absolute 10000 most-intense"
filter="scanTime [0.01,1800]"
```

*The threshold and scantime filters must be changes as necessary.

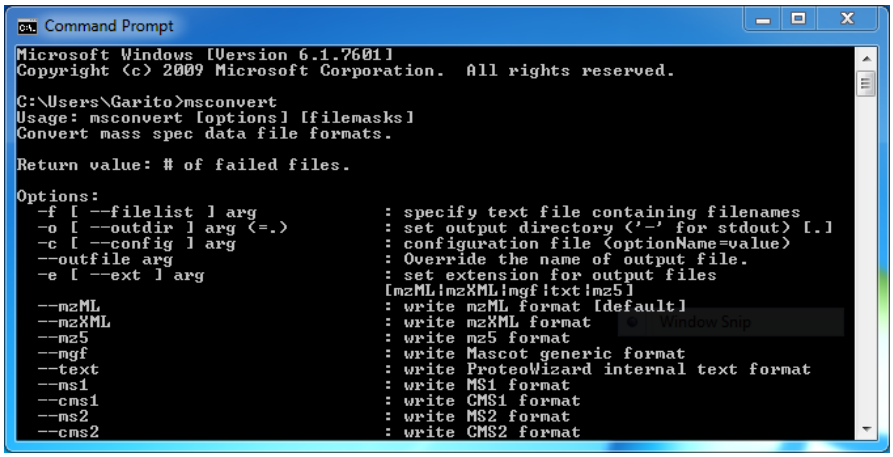
Text 3, save it as **filenames**:
Must contain a filenames list of LC-MS data, including its extension (e.g., .raw)

```
Filename_1.raw
Filename_2.raw
Filename_3.raw
Filename_4.raw
Filename_5.raw
Filename_6.raw
.....
```

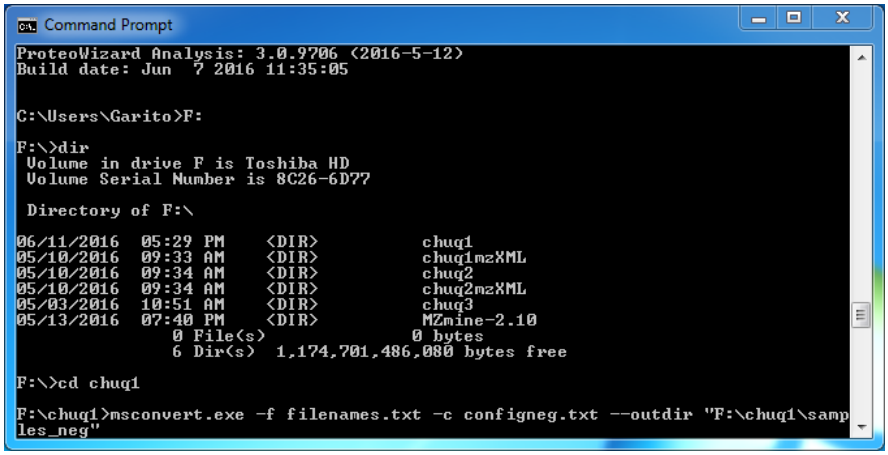
Folder 1, save it as **samples_neg**:

Folder 2, save it as **samples_pos**:

C. Open a **Command Prompt** window. Type **msconvert** to test if command line functions of ProteoWizard are enable in Command Prompt.



Change the directory to the folder that contains the LC-MS data. Then execute the msconvert.exe function as follows:




```
msconvert.exe -f filenames.txt -c configneg.txt --outdir ""

msconvert.exe -f filenames.txt -c configneg.txt --outdir
"F:\chuq1\samples_neg"
```

Repeat the function to convert the data to positive mode.

```
msconvert.exe -f filenames.txt -c configpos.txt --outdir ""

msconvert.exe -f filenames.txt -c configpos.txt --outdir
"F:\chuq1\samples_pos"
```

Examine the data.

APPENDIX 2. Parameters used in MZmine 2.20

1. Raw data import:
Raw data methods > Raw data import

2. Raw data filtering:
Raw data filtering > Scan by scan filtering
Filter: Savitzky-Golay filter
Number of datapoints: 5
Remove source file afhter filtering

3. Generation of mass lists (detected ions) for each scan using the Mass detector
Peak Picking > Mass detection
MS level 1
Set filters:
Mass detector: Exact mass
Noise level: 1.0E6

4. Mass list filtering for FTMS shoulder peaks
Peak Picking > FTMS shoulder peaks filter
Mass resolution: 70000
Peak model function: Lorentzian extended

5. Detection of chromatograms using the Chromatogram builder
Peak Picking > Chromatogram builder
Min time span (min): 0.1
Min height: 5.0E6
m/z tolerance: 0.001 m/z or 5.0 ppm

6. Deconvolution of chromatograms into individual peaks
Peak list processing > Chromatogram deconvolution
Local minimum search
Chromatographic threshold: 5.0
Search minimum in RT range (min): 0.10
Minimum relative height: 15 %
Minimum absolute height: 5.0E6
Min ratio of peak top/edge: 5
Peak duration range (min): 0.10 – 4

7. Removing of isotopes
Isotopes > Isotopic peaks grouper
m/z tolerance: 0.001 m/z or 5.0 ppm
Retention time tolerance: 0.9 min
Maximum charge: 2
Representative isotope: Most intense

8. Identification of fragments, adducts and peak complexes

Identification > Fragment search

Retention time tolerance: 0.9 min
m/z tolerance: 0.001 m/z or 5.0 ppm
Max fragment peak height: 50.0 %
Min MS2 peak height: 5.0E6

Identification > Adduct search

Retention time tolerance: 0.9 min
m/z tolerance: 0.001 m/z or 5.0 ppm
Max relative adduct peak height: 50.0%

Identification > Complex search

Ionization method: [M-H]-
Retention time tolerance: 0.9 min
m/z tolerance: 0.001 m/z or 5.0 ppm
Max complex peak height: 50.0%

9. Alignment using the Join aligner

Alignment > Join aligner

m/z tolerance: 0.001 m/z or 5.0 ppm
Weight for m/z: 15
Retention time tolerance: 0.9 min
Weight for RT: 10
Compare isotope pattern: Setup
Isotope m/z tolerance: 0.001 m/z or 5.0 ppm
Minimum absolute intensity: 5.0E6
Minimum score: 65.0%

10. Gap filling using the Peak finder

Gap filling > Same RT and m/z range gap filler

m/z tolerance: 0.001 m/z or 5.0 ppm

11. Identification using a custom database

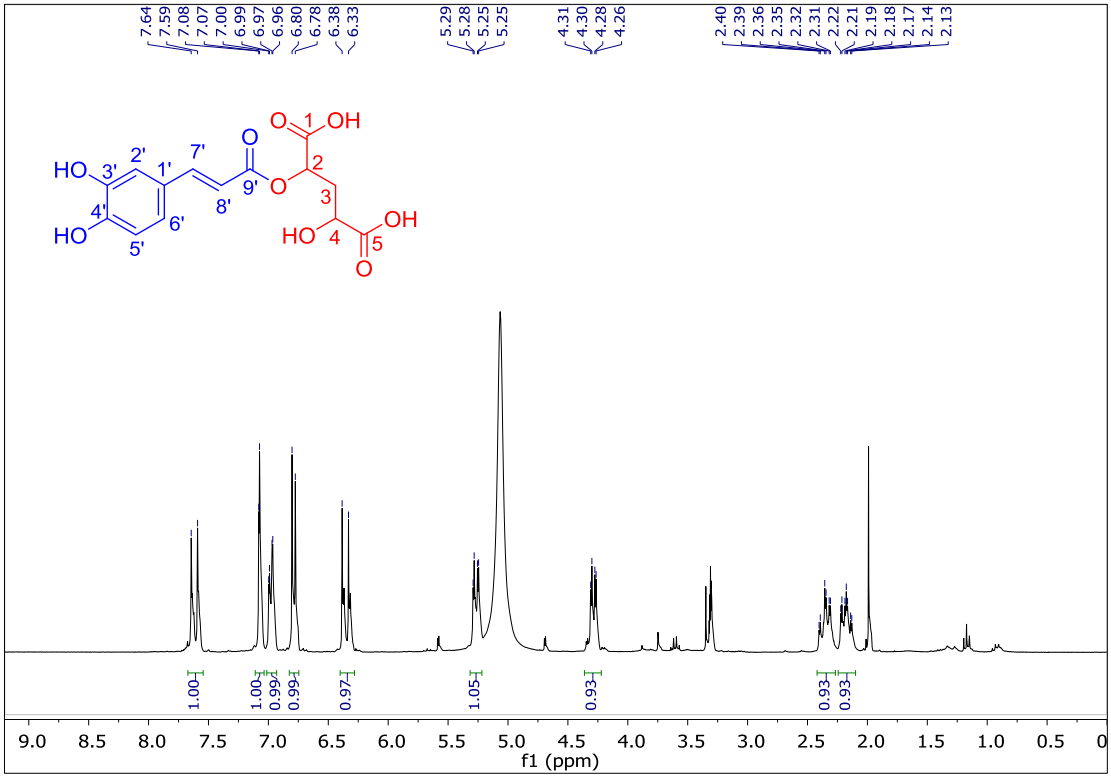
Identification > Custom database search

12. Export

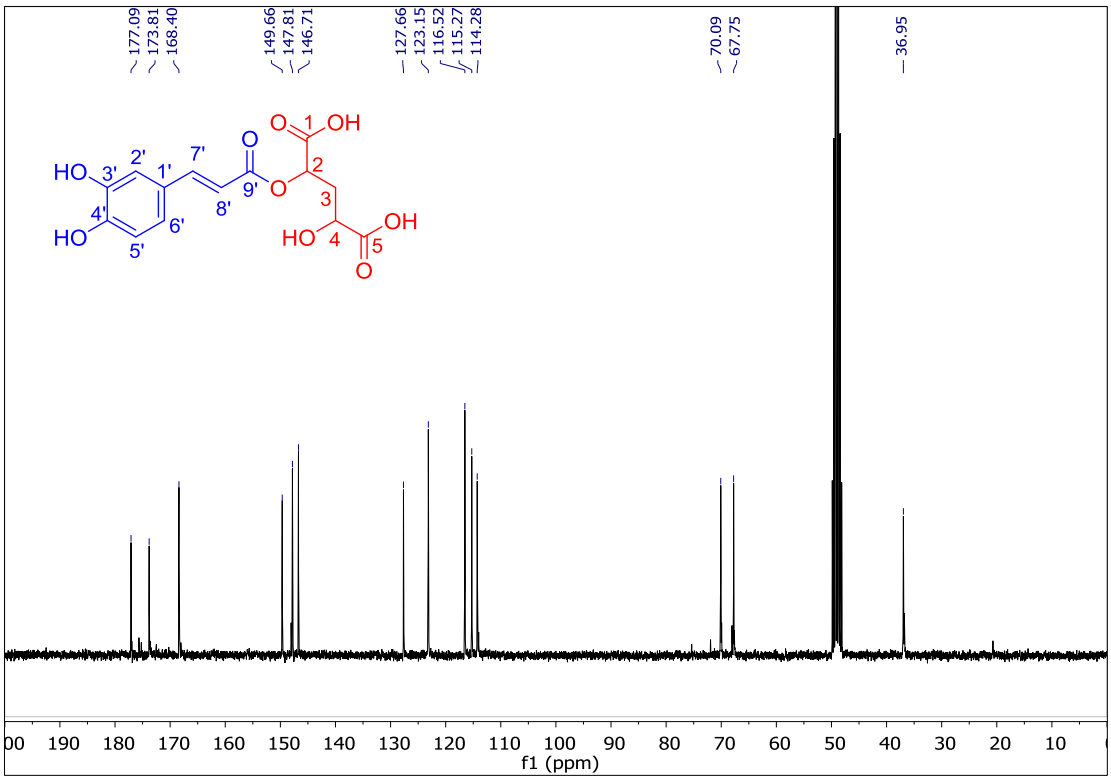
Peak list export > Export to CSV file

13. Save protect

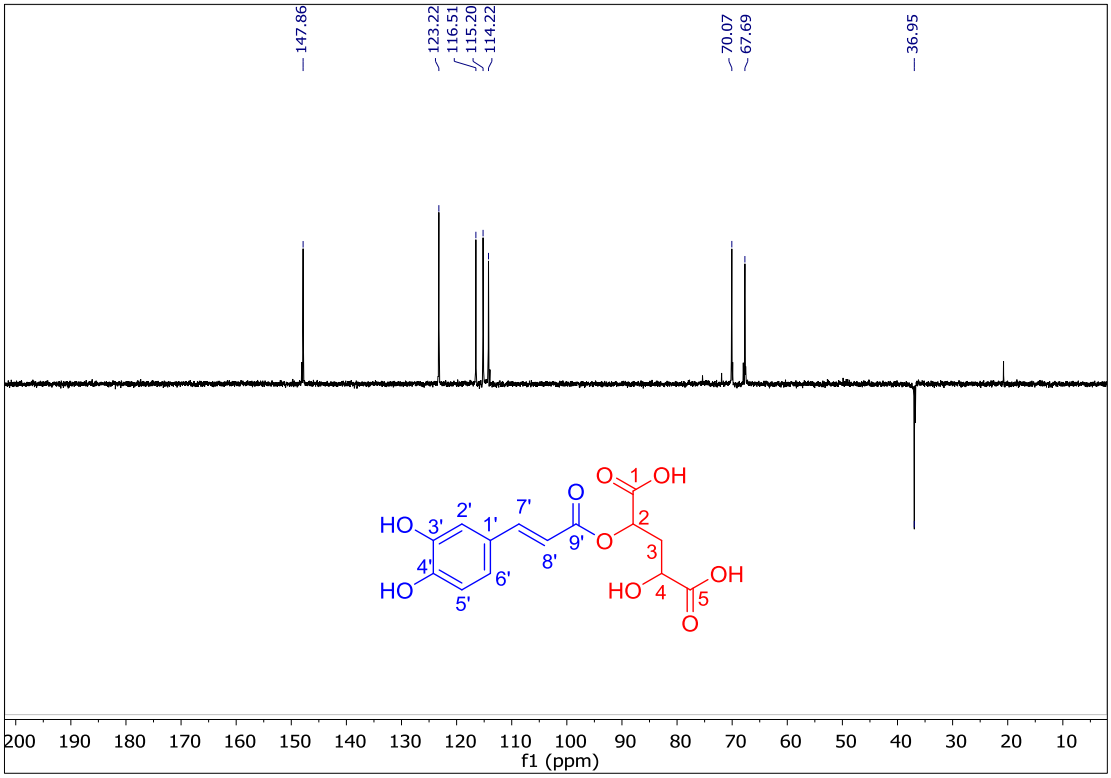
APPENDIX 3. NMR spectra of new caffeic acid ester derivatives



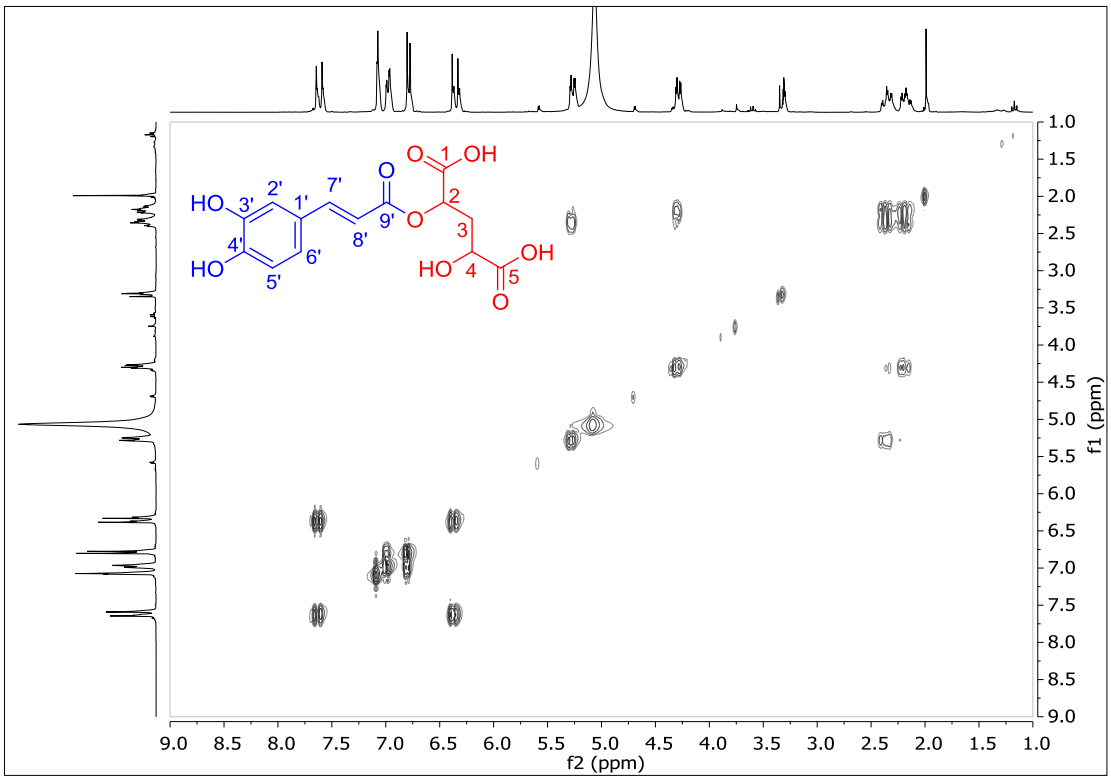
APPENDIX 3.1. ^{1}H NMR spectra of compound **75** (300 MHz, methanol-d_4).



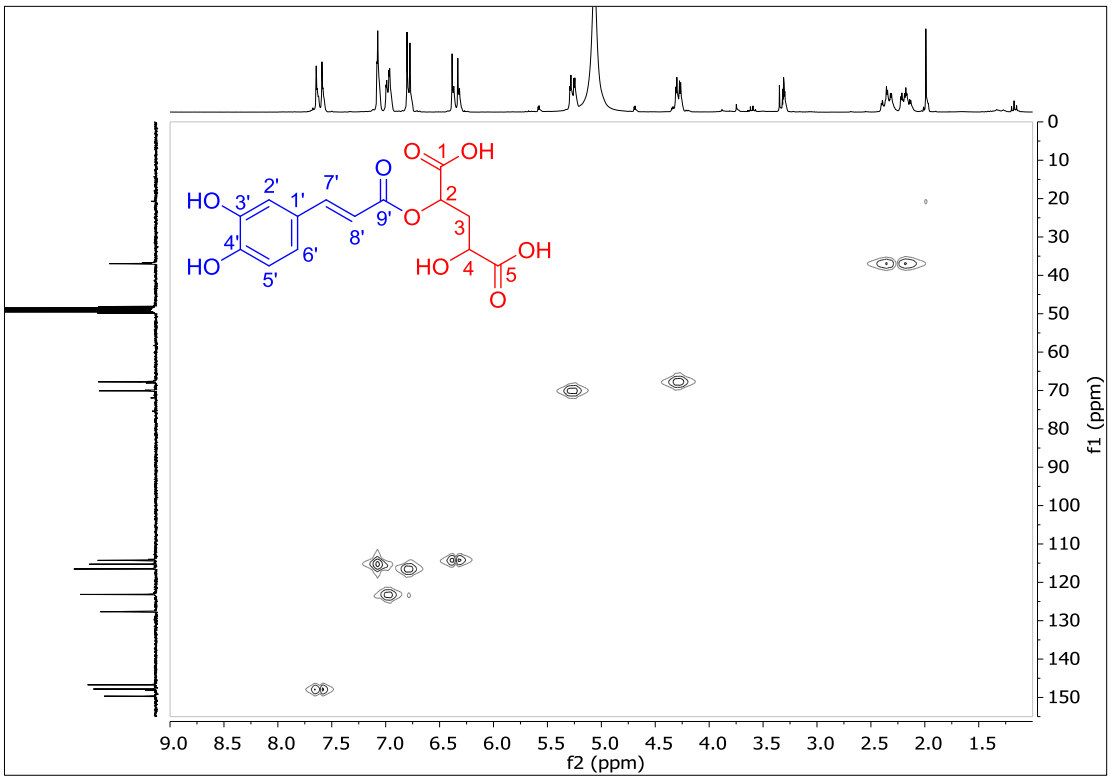
APPENDIX 3.2. ^{13}C NMR spectra of compound **75** (75 MHz, methanol-d_4).



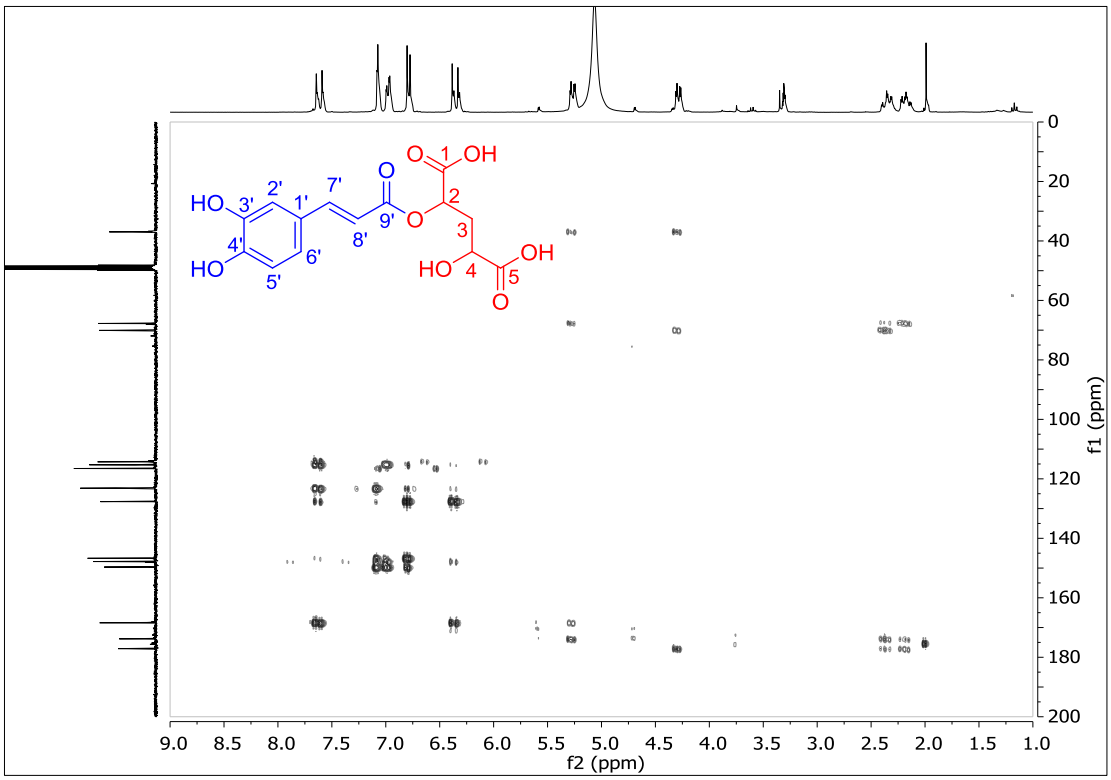
APPENDIX 3.3. DEPT-135 spectra of compound **75** (75 MHz, methanol-d_4).



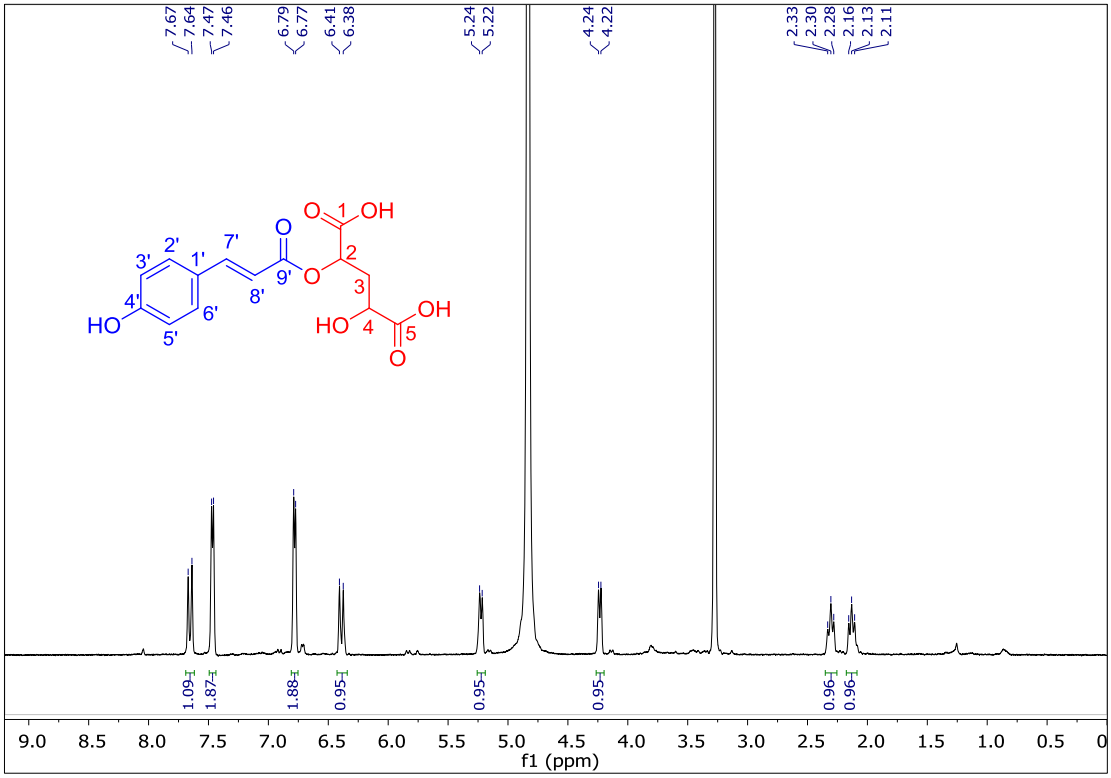
APPENDIX 3.4. COSY spectra of compound **75** (300 MHz, methanol-d_4).



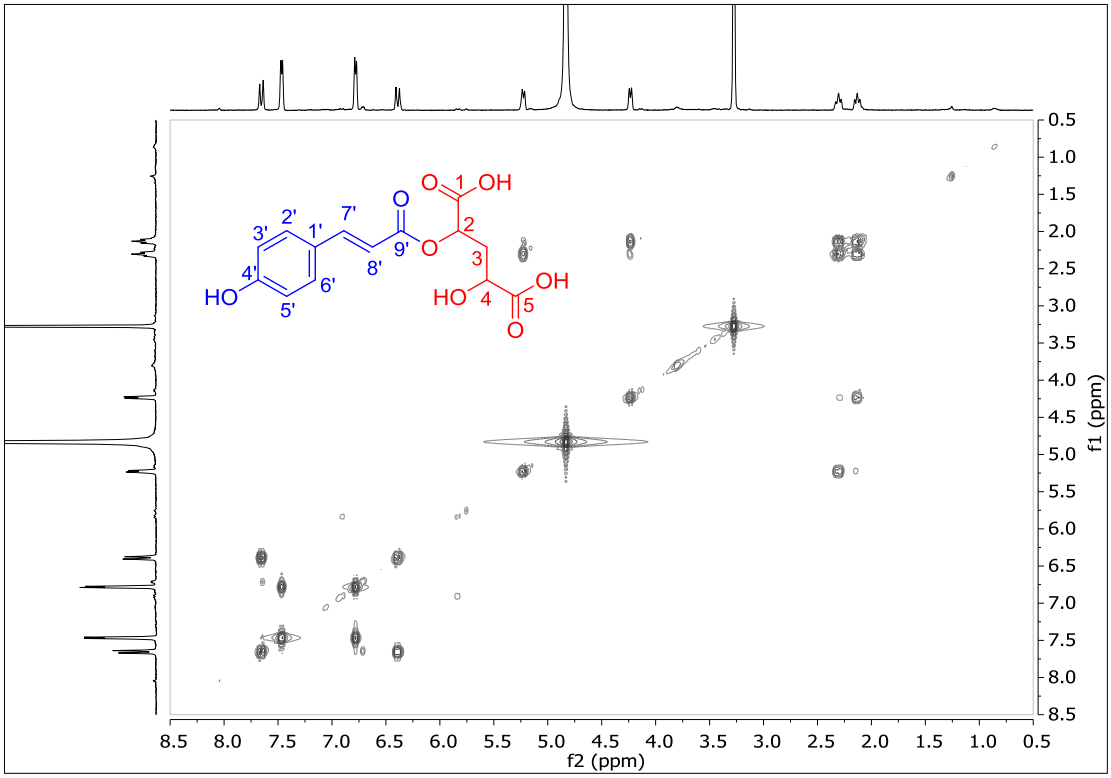
APPENDIX 3.5. HSQC spectra of compound **75** (300 MHz, methanol-d_4).



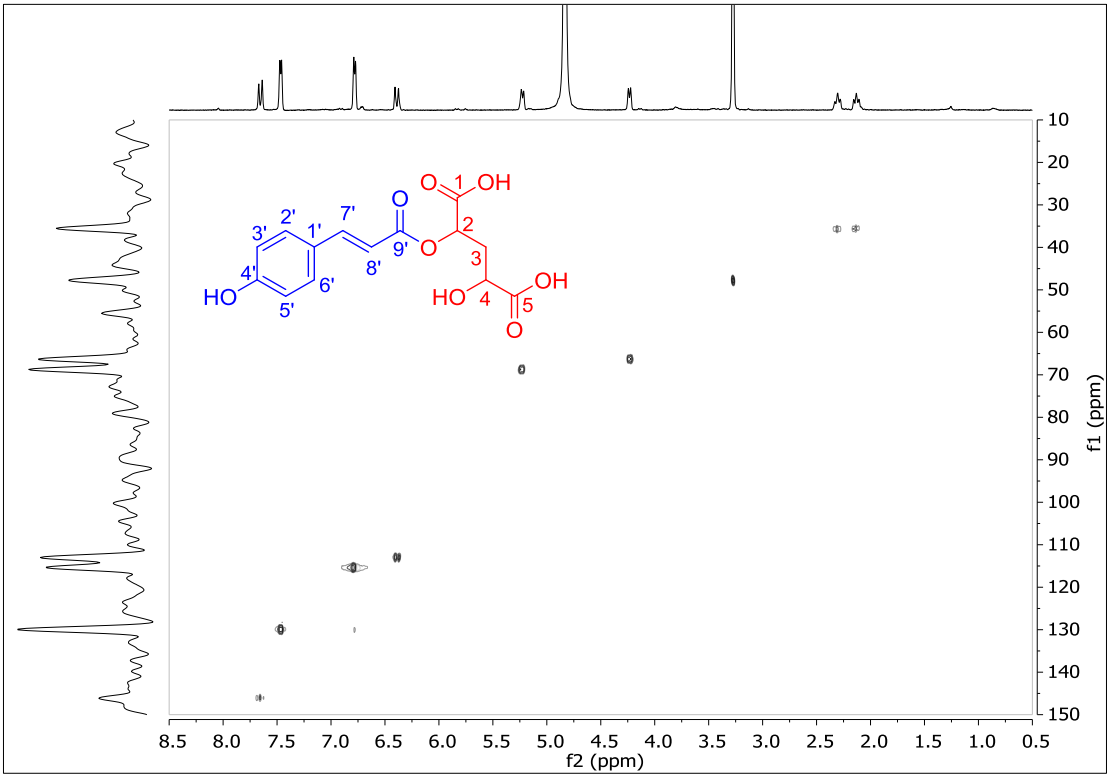
APPENDIX 3.6. HMBC spectra of compound **75** (300 MHz, methanol-d_4).



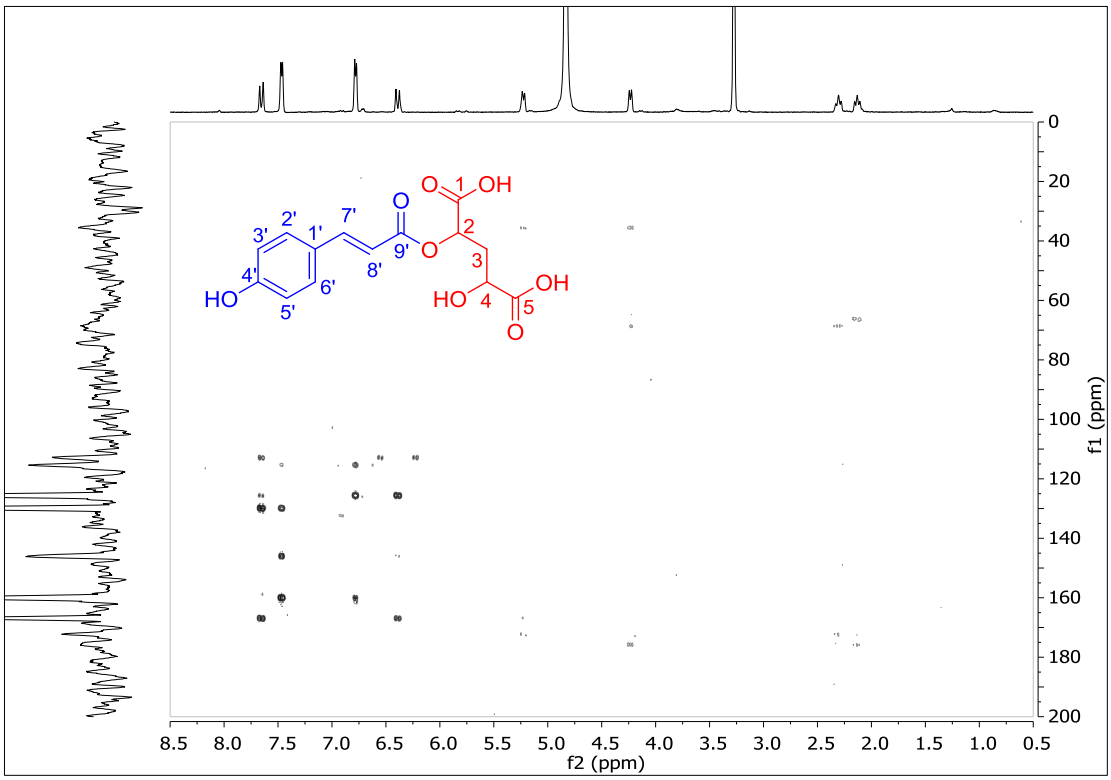
APPENDIX 3.7. ^{1}H NMR spectra of compound **100** (500 MHz, methanol-d_4).



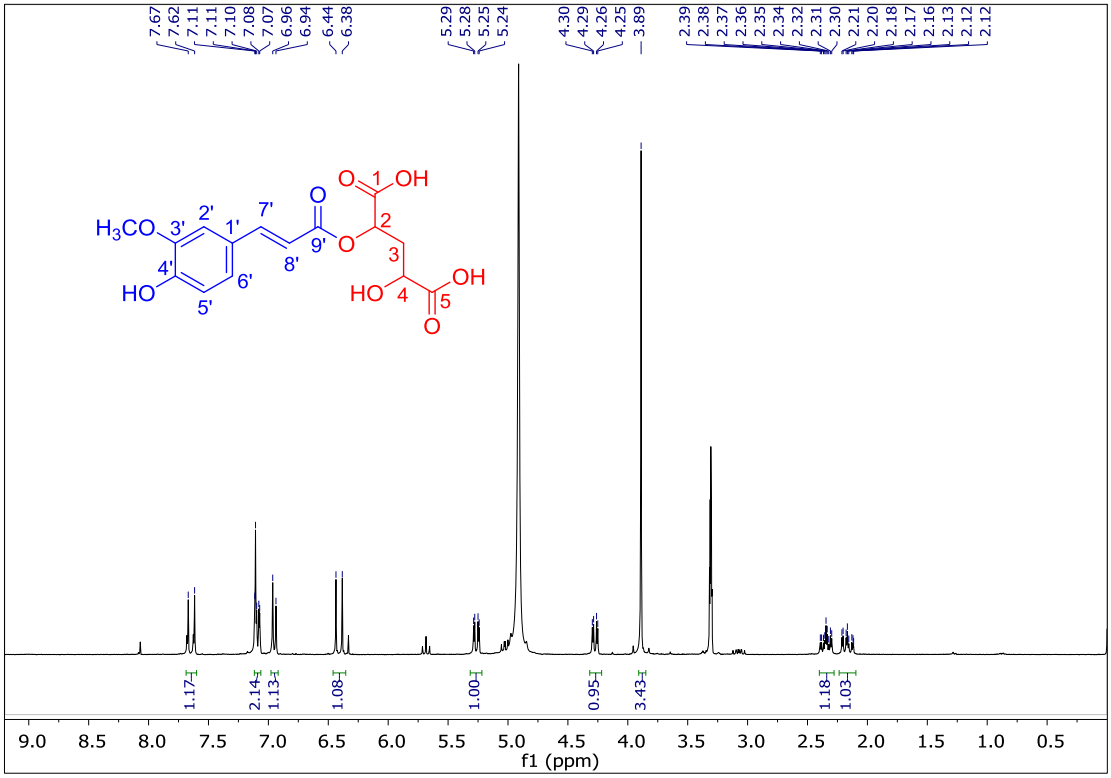
APPENDIX 3.8. COSY spectra of compound **100** (500 MHz, methanol-d_4).



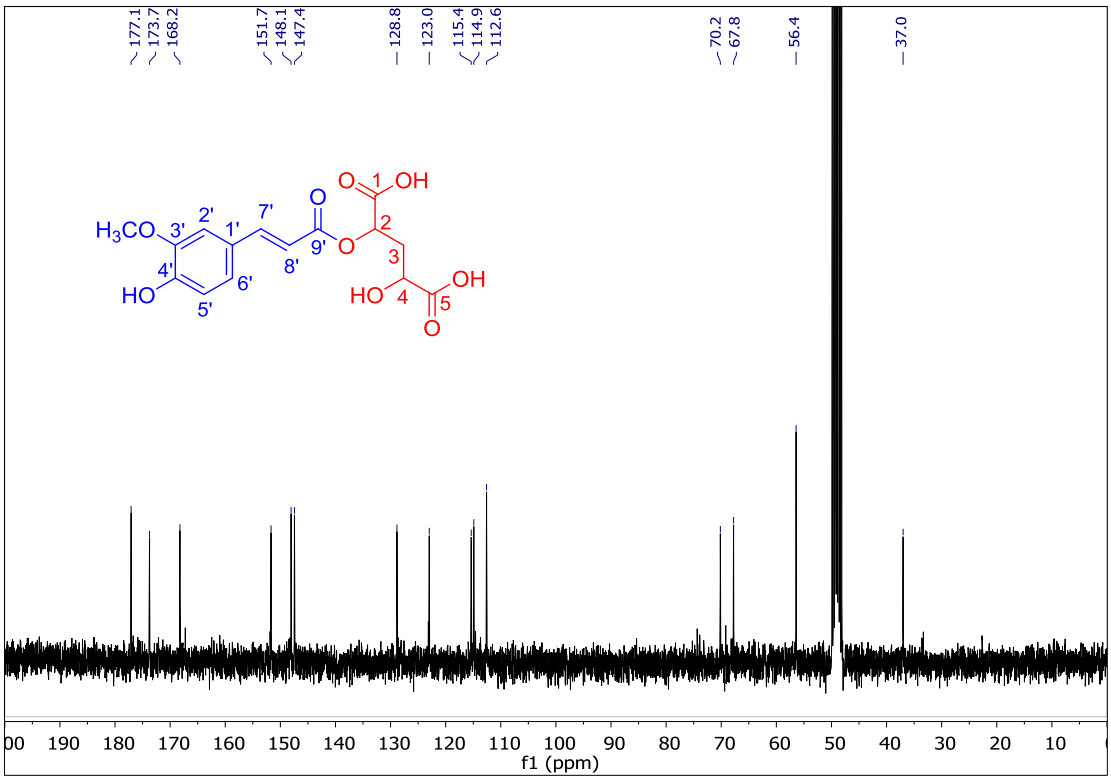
APPENDIX 3.9. HSQC spectra of compound **100** (500 MHz, methanol-d_4).



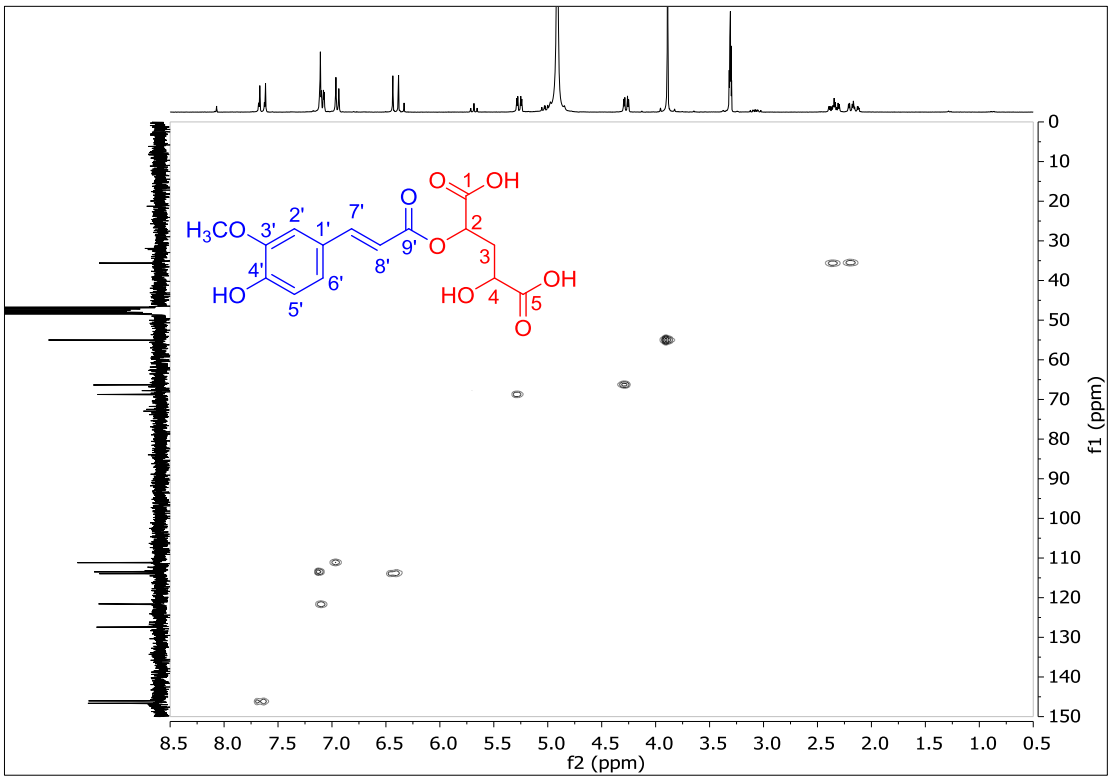
APPENDIX 3.10. HMBC spectra of compound **100** (500 MHz, methanol-d_4).



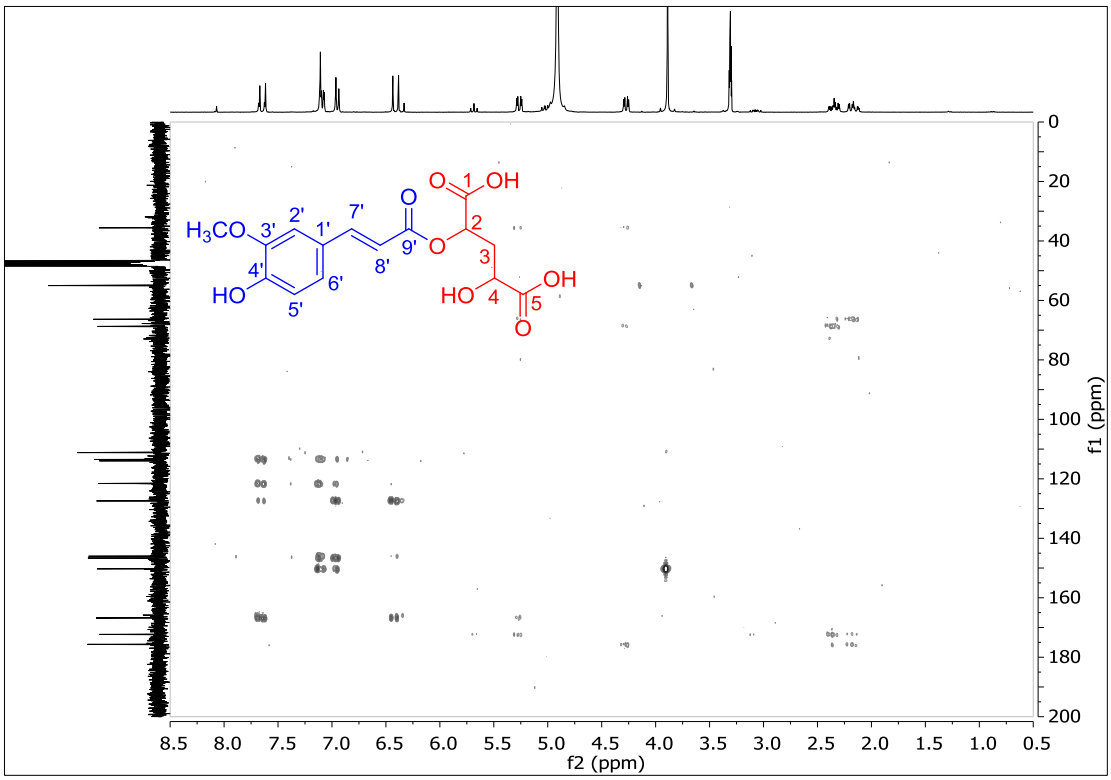
APPENDIX 3.11. ^{1}H NMR spectra of compound **117** (300 MHz, methanol-d_4).



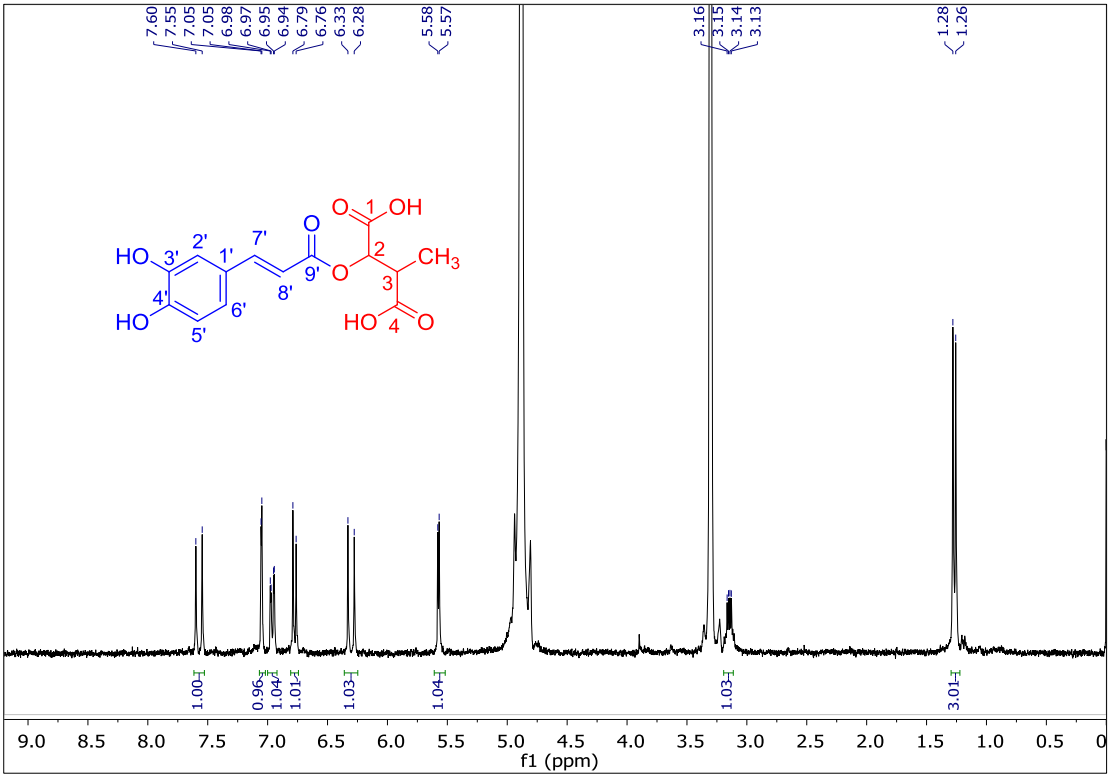
APPENDIX 3.12. ^{13}C NMR spectra of compound **117** (300 MHz, methanol-d_4).



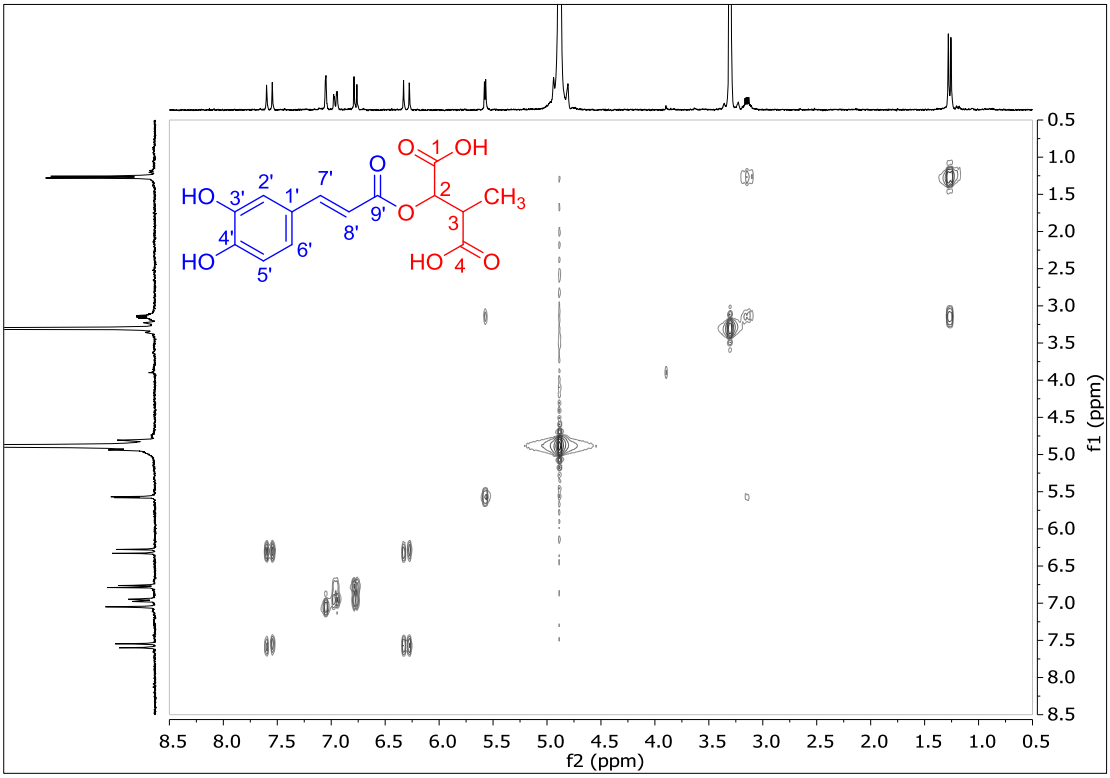
APPENDIX 3.13. HSQC spectra of compound **117** (300 MHz, methanol-d_4).



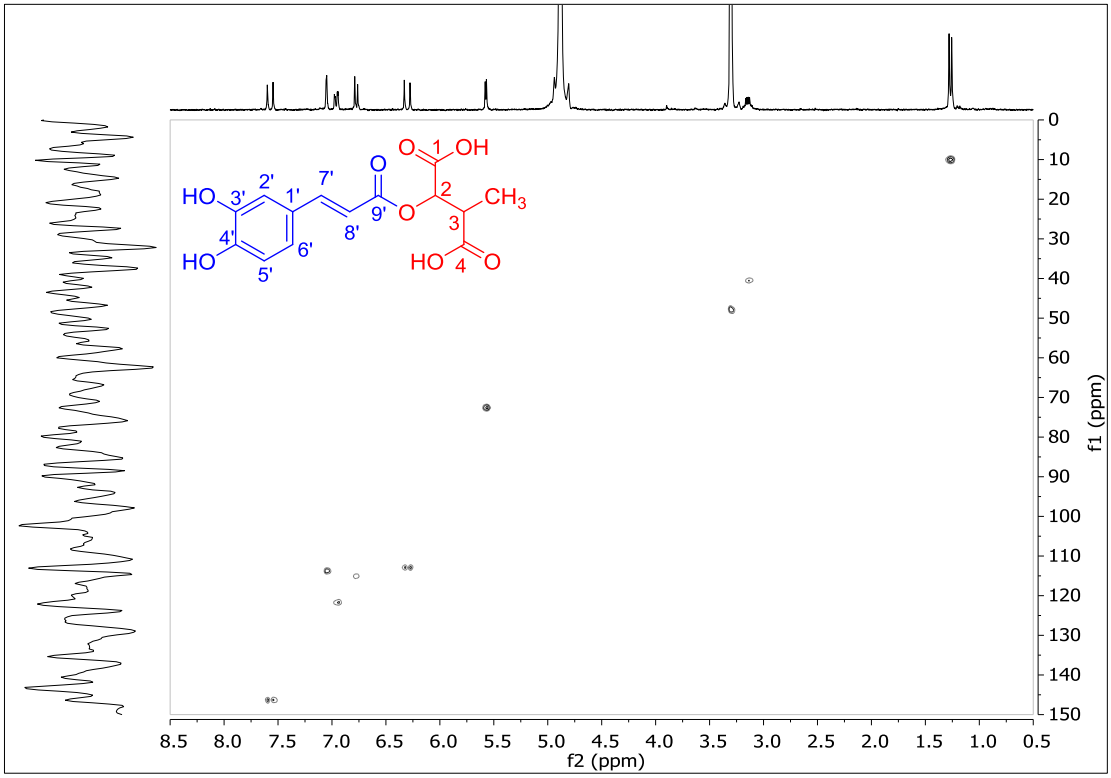
APPENDIX 3.14. HMBC spectra of compound **117** (300 MHz, methanol-d_4).



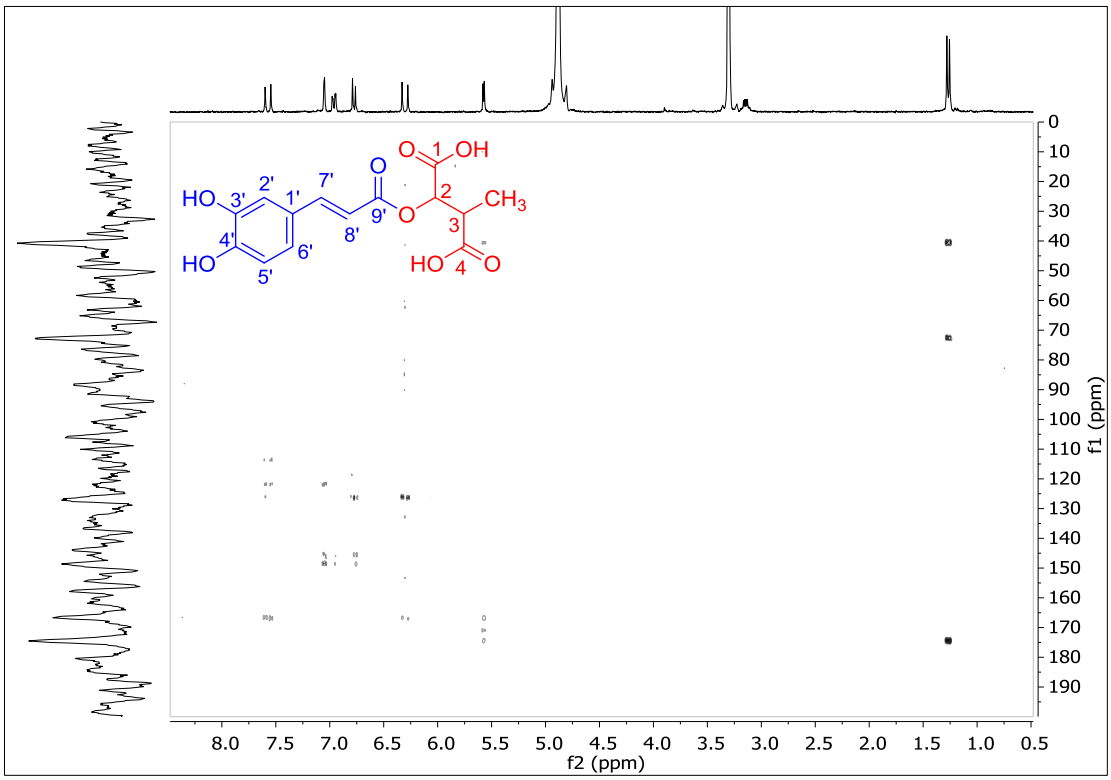
APPENDIX 3.15. ^{1}H NMR spectra of compound **122** (300 MHz, methanol-d_4).



APPENDIX 3.16. COSY spectra of compound **122** (300 MHz, methanol-d_4).



APPENDIX 3.17. HSQC spectra of compound **122** (300 MHz, methanol-d_4).

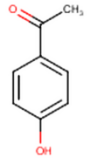
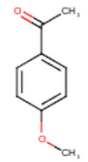
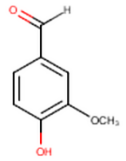
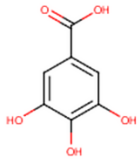
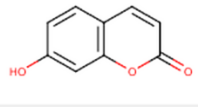


APPENDIX 3.18. HMBC spectra of compound **122** (300 MHz, methanol-d_4).

APPENDIX 4. AsterDB_Barnadesioideae database

AsterDB_Barnadesioiedae Compounds.sdf - MarvinView 16.2.15

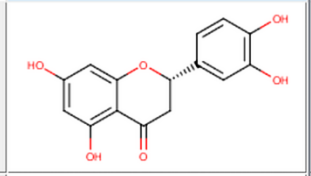
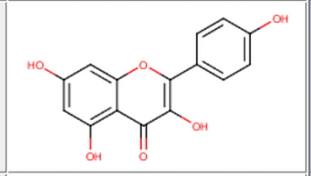
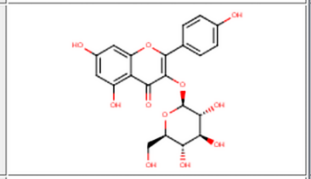
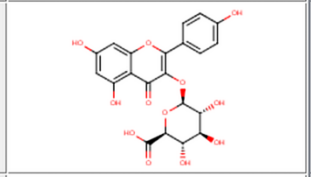
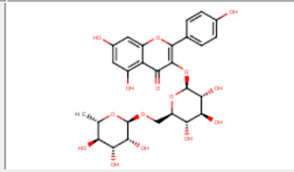
File Edit View Table Structure Tools Help

#	structure	CAS number	Compound subclass	Usual name(s)	Sistematic names (scifinder)	Smiles	InChI	InChIKey
1		99-93-4	Acetophenone	p-Hydroxyacetophenone	4'-Hydroxy-acetophenone p-Hydroxy-acetophenone 1-(4-Hydroxyphenyl)-1-ethanone 1-(4-Hydroxyphenyl)ethanone	CC(=O)C1=CC=C(O)C=C1	InChI=1S/C8H8O2/c1-6(9)7-2-4-8(10)5-3-7/h2-5,10H,1H3	InChIKey=TXFPEBPIARQU
2		100-06-1	Acetophenone	p-Methoxyacetophenone	4'-Methoxy-acetophenone 1-(4-Methoxyphenyl)ethan-1-one 1-(4-Methoxyphenyl)ethanone 1-(4-Methoxyphenyl)-ethanone 1-Acetyl-4-methoxybenzene	COC1=CC=C(C=C1)C(C)=O	InChI=1S/C9H10O2/c1-7(10)8-3-5-9(11-2)6-4-8/h3-6H,1-2H3	InChIKey=NTPLXRHDUXRF
3		121-33-5	Benzaldehyde	Vanillin	3-Methoxy-4-hydroxybenzaldehyde 4-Hydroxy-3-methoxy-benzaldehyde 4-Hydroxy-3-methoxybenzaldehyde 4-Hydroxy-5-methoxybenzaldehyde 2-Methoxy-4-formylphenol 4-Formyl-2-methoxyphenol	[H]C(=O)C1=CC(OC)=C(O)C=C1	InChI=1S/C8H8O3/c1-11-8-4-6(5-9)2-3-7(8)10/h2-5,10H,1H3	InChIKey=MWOOGOJBHIAI
4		149-91-7	Benzoic acid	Gallic acid	3,4,5-Trihydroxybenzoic acid	OC(=O)C1=CC(O)=C(O)C(O)=C1	InChI=1S/C7H6O5/c8-4-1-3(7(11)12)2-5(9)6(4)10/h1-2,8-10H,(H,11,12)	InChIKey=LNTHITQWFMAD
5		93-35-6	Coumarin	Umbelliferone	7-Hydroxy-coumarin 7-Hydroxy-2-chromenone 7-Hydroxy-2H-1-benzopyran-2-one 7-Hydroxy-2H-chromen-2-one 7-Hydroxycoumarin 7-Hydroxylcoumarin	OC1=CC2=C(C=C1)C=CC(=O)O2	InChI=1S/C9H6O3/c10-7-3-1-6-2-4-9(11)12-8(6)5-7/h1-5,10H	InChIKey=ORHBXUUXSCNI

APPENDIX 4.1. View of the AsterDB_Barnadesioideae: CAS numbers, usual and systematic names, and other chemical file formats.

AsterDB_Barnadesioiedae Compounds.sdf - MarvinView 16.2.15

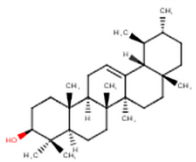
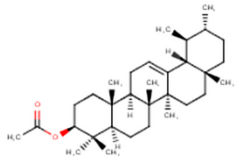
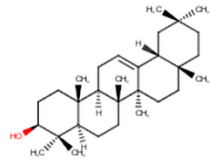
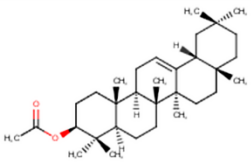
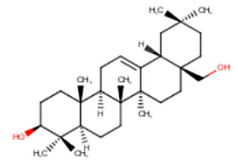
File Edit View Table Structure Tools Help

#	structure	CAS number	Compound subclass	Usual name(s)	Molecular formula	Molecular mass	Monoisotopic mass	[M-H]-	[M+H]+
6		552-58-9	Flavonoid	Eriodictyol	C15H12O6	288.255	288.063388106	287.0555630738	289.0712131382
7		520-18-3	Flavonoid	Kaempferol	C15H10O6	286.239	286.047738042	285.0399130098	287.0555630742
8		480-10-4	Flavonoid	Kaempferol-3-O-glucoside	C21H20O11	448.38	448.100561464	447.0927364318	449.1083864962
9		22688-78-4	Flavonoid	Kaempferol-3-O-glucuronide	C21H18O12	462.363	462.07982602	461.0720009878	463.0876510522
10		17650-84-9	Flavonoid	Kaempferol-3-O-rutinoside	C27H30O15	594.522	594.158470266	593.1506452338	595.1662952982

APPENDIX 4.2. View of the AsterDB_Barnadesioideae: molecular formulas, molecular and monoisotopic mass data.

AsterDB_Barnadesioiedae Compounds.sdf - MarvinView 16.2.15

File Edit View Table Structure Tools Help

#	structure	CAS number	Usual name(s)	1H NMR, published data	13C NMR, published data	Selected NMR data references (
19		638-95-9	alfa-Amyrin	(300 MHz, DMSO-d6): 3.22 (1H, dd, Jax,ax = 10.2 Hz, Jax,eq = 4.3 Hz, H-3); 5.24 (1H, m, H-12); 0.99 (3H, s, Me-23); 0.81 (3H, s, Me-24); 0.96 (3H, s, Me-25); 1.02 (3H, s, Me-26); 1.16 (3H, s, Me-27); 0.84 (3H, s, Me-28); 0.88 (3H, s, Me-29); 0.88 (3H, s, Me-30)	(15 MHz, CDCl3): 38.7 (C-1); 27.2 (C-2); 78.8 (C-3); 38.7 (C-4); 55.2 (C-5); 18.3 (C-6); 32.9 (C-7); 40.0 (C-8); 47.7 (C-9); 36.9 (C-10); 17.4 (C-11); 122.3 (C-12); 139.3 (C-13); 42.0 (C-14); 28.7 (C-15); 26.6 (C-16); 33.7 (C-17); 58.9 (C-18); 39.6 (C-19); 39.6 (C-20); 31.2 (C-21); 41.5 (C-22); 28.1 (C-23); 15.6 (C-24); 15.6 (C-25); 16.8 (C-26); 23.3 (C-27); 28.1 (C-28); 23.3 (C-29); 21.3 (C-30) (25 MHz, CDCl3): 39.0 (C-1); 27.4 (C-2); 79.0 (C-3); 38.8 (C-4); 55.3 (C-5); 18.5 (C-6); 33.1 (C-7); 40.1 (C-8); 47.9 (C-9); 37.0 (C-10); 23.4 (C-11); 124.5 (C-12); 139.6 (C-13); 42.2 (C-14); 28.7 (C-15); 28.2 (C-16); 33.8 (C-17); 59.2 (C-18); 39.8 (C-19); 39.7 (C-20);	Seo, S; Tetrahedron Lett 1975, ' Knight, S; Org Magn Reson 1974 Saeed, M; J Asian Nat Prod Res Lima, M; J Braz Chem Soc 2004
20		863-76-3	alfa-Amyrin acetate	(300 MHz, CDCl3): 4.50 (1H, dd, J = 9.7 Hz, H-3); 5.12 (1H, t, J = 3.6 Hz, H-12); 0.88 (3H, s, Me-23); 0.88 (3H, s, Me-24); 1.01 (3H, s, Me-25); 0.98 (3H, s, Me-26); 1,07 (3H, s, Me-27); 0.79 (3H, s, Me-28); 0.88 (3H, s, Me-29); 0.88 (3H, s, Me-30); 2.05 (3H, s, -OCOMe)	(15 MHz, CDCl3): 38.4 (C-1); 23.6 (C-2); 80.7 (C-3); 37.6 (C-4); 55.3 (C-5); 18.3 (C-6); 32.8 (C-7); 40.1 (C-8); 47.6 (C-9); 36.8 (C-10); 17.5 (C-11); 124.1 (C-12); 139.4 (C-13); 42.1 (C-14); 28.7 (C-15); 26.7 (C-16); 33.8 (C-17); 59.0 (C-18); 39.7 (C-19); 39.7 (C-20); 31.3 (C-21); 41.5 (C-22); 28.1 (C-23); 16.8 (C-24); 15.7 (C-25); 16.8 (C-26); 23.2 (C-27); 28.1 (C-28); 23.2 (C-29); 21.4 (C-30); 21.2 (-OOCMe); 170.4 (-OOCMe) (75 MHz, CDCl3): 38.5 (C-1); 23.4 (C-2); 80.9 (C-3); 37.7 (C-4); 55.3 (C-5); 18.3 (C-6); 32.9 (C-7); 39.7 (C-8); 47.7 (C-9); 36.8 (C-10); 22.8 (C-11); 124.2 (C-12); 139.5 (C-13); 42.1 (C-14); 28.2 (C-15); 26.7 (C-16); 33.8 (C-17); 59.0 (C-18); 39.7 (C-19); 39.7 (C-20);	Seo, S; Tetrahedron Lett 1975, ' Ali, N; BMC Complement Altern
21		559-70-6	beta-Amyrin	(300 MHz, DMSO-d6): 3.18 (1H, dd, Jax,ax = 10.0 Hz, Jax,eq = 4.6 Hz, H-3); 5.10 (1H, m, H-12); 1.01 (3H, s, Me-23); 0.81 (6H, s, Me-24); 0.96 (3H, s, Me-25); 1.02 (3H, s, Me-26); 1.09 (3H, s, Me-27); 0.80 (3H, d, J = 6.8 Hz Me-29); 0.91 (3H, d, J = 6.6 Hz; Me-30)	(15 MHz, CDCl3): 38.5 (C-1); 27.0 (C-2); 78.9 (C-3); 38.7 (C-4); 55.1 (C-5); 18.3 (C-6); 32.6 (C-7); 39.7 (C-8); 47.6 (C-9); 37.0 (C-10); 23.4 (C-11); 121.7 (C-12); 145.0 (C-13); 41.7 (C-14); 28.3 (C-15); 26.2 (C-16); 32.5 (C-17); 47.2 (C-18); 46.8 (C-19); 31.1 (C-20); 34.8 (C-21); 37.2 (C-22); 28.1 (C-23); 15.5 (C-24); 15.5 (C-25); 16.8 (C-26); 26.0 (C-27); 27.3 (C-28); 33.2 (C-29); 23.6 (C-30) (25 MHz, CDCl3): 38.7 (C-1); 27.3 (C-2); 79.0 (C-3); 38.8 (C-4); 55.3 (C-5); 18.5 (C-6); 32.8 (C-7); 38.8 (C-8); 47.7 (C-9); 37.0 (C-10); 23.6 (C-11); 121.8 (C-12); 145.1 (C-13); 41.8 (C-14); 26.2 (C-15); 27.0 (C-16); 32.5 (C-17); 47.4 (C-18); 46.9 (C-19); 31.1 (C-20);	Seo, S; Tetrahedron Lett 1975, ' Knight, S; Org Magn Reson 1974 Saeed, M; J Asian Nat Prod Res Lima, M; J Braz Chem Soc 2004
22		1616-93-9	beta-Amyrin acetate	(400 MHz, CDCl3): 4.50 (1H, dd, J = 8.8, 6.8 Hz, H-3); 5.18 (1H, dd, J = 3.2, 3.2 Hz, H-12); 0.83 (3H, s, Me-23); 0.87 (3H, s, Me-24); 0.87 (3H, s, Me-25); 0.96 (3H, s, Me-26); 0.97 (3H, s, Me-27); 1.13 (3H, s, Me-28); 0.88 (3H, s, Me-29); 0.88 (3H, s, Me-30); 2.05 (3H, s, -OCOMe) (500 MHz, CDCl3): 4.52 (1H, t like, J = 8.0 Hz, H-3); 5.20 (1H, t, J = 3.5 Hz, H-12); 0.88 (3H, s, Me-23); 0.88 (3H, s, Me-24); 0.98 (3H, s, Me-25); 0.99 (3H, s, Me-26); 1.15 (3H, s, Me-27); 0.85 (3H, s, Me-28); 0.89 (3H, s, Me-29); 0.89 (3H, s, Me-30); 2.06 (3H, s, -OCOMe)	(15 MHz, CDCl3): 38.2 (C-1); 23.6 (C-2); 80.7 (C-3); 37.6 (C-4); 55.3 (C-5); 18.3 (C-6); 32.6 (C-7); 39.7 (C-8); 47.6 (C-9); 36.8 (C-10); 23.4 (C-11); 121.5 (C-12); 144.9 (C-13); 41.7 (C-14); 28.3 (C-15); 26.2 (C-16); 32.5 (C-17); 47.2 (C-18); 46.8 (C-19); 31.1 (C-20); 34.8 (C-21); 37.1 (C-22); 28.1 (C-23); 16.8 (C-24); 15.7 (C-25); 16.8 (C-26); 26.0 (C-27); 27.0 (C-28); 33.4 (C-29); 23.6 (C-30); 21.2 (-OOCMe); 170.4 (-OOCMe) (100 MHz, CDCl3): 38.2 (C-1); 23.5 (C-2); 81.0 (C-3); 37.7 (C-4); 55.3 (C-5); 18.3 (C-6); 32.6 (C-7); 39.8 (C-8); 47.6 (C-9); 36.9 (C-10); 23.5 (C-11); 121.7 (C-12); 145.2 (C-13); 41.7 (C-14); 26.9 (C-15); 26.1 (C-16); 32.5 (C-17); 47.2 (C-18); 46.8 (C-19); 31.1 (C-20);	Seo, S; Tetrahedron Lett 1975, ' Wang, Y; Pharmazie 2003, 58, 5 Choi, S; Arch Pharm Res 2004, 2
23		545-48-2	Erythrodiol	(400 MHz, CDCl3): 3.38 (1H, dd, J=11.5, 4.6, H-3); 5.22 (1H, m, H-12); 4.17 (2H, m, H-28); 1.21, 0.99, 0.94, 0.87, 0.86, 0.80, 0.78 (each 3H, s, H-23, 24, 25, 26, 27, 29, 30)	(400 MHz, CDCl3): 38.6 (C-1); 27.2 (C-2); 79.0 (C-3); 38.7 (C-4); 55.2 (C-5); 18.3 (C-6); 32.6 (C-7); 39.8 (C-8); 47.6 (C-9); 36.9 (C-10); 23.5 (C-11); 122.4 (C-12); 144.2 (C-13); 41.7 (C-14); 25.5 (C-15); 22.0 (C-16); 36.9 (C-17); 42.3 (C-18); 46.5 (C-19); 31.0 (C-20); 34.0 (C-21); 31.0 (C-22); 28.0 (C-23); 15.5 (C-24); 15.6 (C-25); 16.7 (C-26); 25.9 (C-27); 69.7 (C-28); 33.2 (C-29); 23.6 (C-30)	Kim, D; Arch Pharm Res 2002, 2

APPENDIX 4.3. View of the AsterDB_Barnadesioideae: ^{1}H and ^{13}C NMR published data.