Rodriguésia 74: e01152022. 2023
http://rodriguesia.jbrj.gov.br
DOI: http://dx.doi.org/10.1590/2175-7860202374049

Review Paper

Use of mass spectrometry as a tool for the search or identification of flavonoids in Urticaceae

Fernanda Manhães Braga Gonçalves[1,3,5], *Adriana Lima de Sousa*[2] & *Rodrigo Rodrigues de Oliveira*[1,4]

Abstract

The Urticaceae family, circumscribed within the Rosales, was investigated in this study with an overview of the current literature about phytochemical studies using the Liquid Chromatography coupled to Mass Spectrometry (LC-MS) technique. The aim of this study was to review the secondary metabolites identified in the Urticaeae using LC-MS analysis. A systematic review was performed using Scifinder and ScienceDirect databases. Phenolic substances are the most abundant in the Urticaceae family, especially flavones, phenolic acids, and flavonols. We have shown that flavonoids are important chemotaxonomic markers of the chemical composition of the Urticaeae. Following chemical attributes, the *C*-glycosylated and *O*-glycosylated flavones stand out as the main skeletons. Our results revealed the chemical profile and structural variability of micromolecules from each genus of Urticaceae. This approach demonstrates a greater use of reversed-phase and liquid chromatography coupled to a mass spectrometer with a negative mode electrospray ionization (ESI) source. In addition, the mobile phase is usually composed of binary systems and eluted by gradient systems. Finally, this paper presents the identification of molecular ion patterns and fragmentation of chemical markers in Urticaceae, identified and isolated using LC-MS, which has been proven to be a valuable tool in several areas, such as phytochemistry, chemosystematics, and chemophenetics. In conclusion, this review is expected to help identify and separate phenolic compounds from the Urticaceae family.

Key words: chemical profile, fragmentation pattern, LC-MS, phenolic acid, phytochemical.

Resumo

A família Urticaceae, circunscrita em Rosales, foi investigada neste estudo com uma visão geral da literatura atual sobre estudos fitoquímicos usando a técnica de Cromatografia Líquida acoplada a Espectrometria de Massas (LC-MS). O objetivo deste estudo foi revisar os metabólitos secundários identificados em Urticaeae por meio da análise de LC-MS. Uma revisão sistemática foi realizada usando os bancos de dados Scifinder e ScienceDirect. Substâncias fenólicas são as mais abundantes na família Urticaceae, especialmente flavonas, ácidos fenólicos e flavonóis. Mostramos que os flavonoides são importantes marcadores quimiotaxonômicos da composição química da família Urticaceae. Seguindo os atributos químicos, as flavonas *C*-glicosiladas e *O*-glicosiladas se destacam como esqueletos principais. Nossos resultados revelaram o perfil químico e a variabilidade estrutural das micromoléculas de cada gênero de Urticaceae. Esta abordagem demonstra um maior uso de fase reversa e cromatografia líquida acoplada a um espectrômetro de massa com uma fonte de ionização por eletropulverização (ESI) de modo negativo. Além disso, a fase móvel é geralmente composta por sistemas binários e eluída por sistemas de gradiente. Por fim, este artigo apresenta a identificação de padrões moleculares de íons e fragmentação de marcadores químicos na família Urticaceae, identificados e isolados por LC-MS, que tem se mostrado uma ferramenta valiosa em diversas áreas como fitoquímica, quimiossistemática e quimiofenética. Em conclusão, espera-se que esta revisão ajude a identificar e separar os compostos fenólicos da família Urticaceae.

Palavras-chave: perfil químico, padrão de fragmentação, LC-MS, ácido fenólico, fitoquímica.

See supplementary material at <https://doi.org/10.6084/m9.figshare.23907960.v1>

[1] Universidade Estadual do Norte Fluminense Darcy Ribeiro, Centro de Ciência e Tecnologia, Lab. Ciências Químicas, Campos dos Goytacazes, RJ, Brasil.
[2] Instituto Federal Fluminense, IFFluminense Campus Campos Centro, Campos dos Goytacazes, RJ, Brasil. ORCID: <https://orcid.org/0000-0002-6839-9578>.
[3] ORCID: <https://orcid.org/0000-0001-5066-5738>. [4] ORCID: <https://orcid.org/0000-0002-3145-4844>.
[5] Author for correspondence: fernanda.mbraga@icloud.com

Introduction

High-resolution mass spectrometry (MS) has become a powerful highly sensitive structural tool with excellent analytical power. It can be particularly valuable and versatile in determining the structural information required for the characterization of secondary metabolites very quickly (Alvarez-Rivera *et al.* 2019).

Different ionization techniques yield molecular ions and their respective molecular fragments. The fragmentations generated provide relevant data for structural elucidation, such as molecular weight, empirical formula, detection of functional groups and substituents, stereochemical features and isotope ratios (Syage *et al.* 2008; Patel *et al.* 2010; Alvarez-Rivera *et al.* 2019).

Technological advances in recent decades have allowed Mass Spectrometry to be coupled with chromatographic methods, such as Liquid Chromatography (LC), playing an important role in the separation and identification of secondary metabolites (Fraige *et al*. 2018).

Coupling mass spectrometers in Liquid Chromatography (LC-MS) provides many advantages for the identification of organic/phenolic substances with high specificity, response time and speed of analysis. Mainly, collision-induced dissociation (CID) influenced the separation and efficiency and thus, the ability of structural elucidation in LC-MS (Seraglio *et al*. 2016; Fraige *et al*. 2018).

LC-MS is a versatile analytical tool with great potential in qualitative and quantitative studies of a wide variety of metabolites with different physicochemical properties (polarity, volatility and molecular weight) (Lanças 2009; Cajka & Fiehn 2016; Kruve 2020; Dührkop *et al.* 2021).

The technique quickly established itself as a tool for the determination of volatile micromolecules, especially phenolic micromolecules. Flavonoids are phenolic substances that are easily separated and identified by LC- MS. About glycosylated flavonoids, the approach of this technique provides detailed information on the structure of the aglycone, the type of sugar and the position of its substitutions, as well as the types of interglycosidic bonds or acyl substituents (Cuyckens & Claeys 2004).

Flavonoids are abundant in angiosperm plants, but their structural variability is unexplored in the Urticaceae family from a chemotaxonomic point of view. The Urticaceae family comprises approximately 1,200 species distributed in 54 genera (Treiber *et al.* 2016). The micromolecular profile of this family is characterized by the presence of flavonoids, phenolic acids and triterpenic acids. It should be emphasized, however, that flavonoids, especially *O*-glycosylated and *C*-glycosylated flavonoids, are the main chemical markers of this family.

Aiming to approach the use of LC-MS technique in the separation and identification of the main classes of metabolites present in the Urticaceae family, a bibliographic survey of the chemical substances identified and isolated in this family was carried out. Therefore, this review is designed to present the structural variability of the main classes identified, being differentiated by fragmentation data.

Material and Methods

The systematic review was performed on the Scifinder and ScienceDirect databases. Articles from January 2006 to April 2022 were selected using the keywords and combinations as search terms: Urticaceae, HPLC-MS (High-Performance Liquid Chromatography coupled to Mass Spectrometry), UHPLC-MS (Ultra High-Performance Liquid Chromatography coupled to Mass Spectrometry) and LC-MS. The literature searches encompassed all genera of the family, as classified by Treiber *et al*. (2016).

Results and Discussion

Analysis parameters for liquid chromatography coupled to mass spectrometry

Forty-two studies were identified using the search criteria. The characteristics of the LC-MS were shown in Table S1 (available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>) with details about the part of the plant analyzed, the column type, stationary and mobile phase chosen and the respective authors.

The liquid chromatography coupled mass spectrometry with an electrospray ionization source (ESI) is mostly used. However, two studies used Atmospheric Pressure Photon Ionization (APPI) (Pinelli *et al.* 2008; Rivera-Mondragón *et al.* 2019) and only one paper used Atmospheric Pressure Chemical Ionization (APCI) (Shrestha *et al.* 2020), with the majority of papers operating in negative mode. The LC-MS parameters are detailed in Table S1 (available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>).

As for the stationary phase, most columns were octadecyl-silica C-18 reversed-phase, with particle size ranging from 2.6 to 5 µm for HPLC and 1.7 to 1.9 µm for UHPLC. In contrast, the study developed by Pinelli *et al.* (2008) used the C-12 stationary phase specifically for the identification of anthocyanins.

The binary mobile phase solvent system was employed in most studies, which consisted of acidified water and an organic phase, also acidified, containing methanol or acetonitrile. The method of elution most noticeable was gradient elution, with different elution times and compositions. Separation of chemicals with similar structures is usually performed by polarity gradient elution, providing better resolution between peaks and shorter analysis time (Costa *et al.* 2000).

In many studies, acetic acid or formic acid have been used in the mobile phase because acidification helps reduce the ionization of phenolic substances, improving retention and separation. It also contributes to obtaining chromatograms with better sharpness of the peaks (Cuyckens & Claeys 2004). The concentration of the acids used varies in most works from 0.05 to 2% (v/v). However, in the identification of anthocyanins a high acidity (4.5% v/v) was observed, presumably due to the instability of these substances, as low pH prevents the degradation of non-acylated anthocyanin pigments (Costa *et al.* 2000).

In studies performed with LC-MS, most chemical substances were identified by interpretation of characteristic fragmentation patterns and/or comparison with existing database spectra. Furthermore, some studies performed identification by means of analytical standards and also with the aid of other techniques, such as NMR (Nuclear Magnetic Resonance).

In this context, it is worth considering that the necessary structural information provided by the Mass Spectrometry technique, combined to the separation efficiency of Liquid Chromatography, make the LC-MS technique a versatile analytical tool with great potential in qualitative and quantitative studies (Lanças 2009).

Chemical profile of Urticaceae

IOf the 54 genera of the Urticaceae family, only 11 have been studied. Of these studied genera, seven are found in Brazilian territory (*Boehmeria, Cecropia, Coussapoa, Pourouma, Urera, Urtica* and *Pilea*). In Brazil, the Urticaceae family is represented by 13 genera [Flora do Brasil 2020 (continuously updated)].

The Urticaceae family has a diversity of phenolic compounds such as flavonoids (flavonols, flavan-3-ols, flavones, and anthocyanins), phenolic acids (derivatives of cinnamic and benzoic acids), condensed tannins, coumarins, flavonolignans, and lignans.

Also, to a lesser extent, there is the presence of terpenoids (iridoids, triterpenes, triterpenoid saponins, pentacyclic triterpenoids, carotenoids, monoterpene lactone and seco-iridoid glycosides), steroids, and alkaloids (Fig. 1). Flavonoids and phenolic acids are the most abundant (52% and 21%, respectively).

Boehmeria genus.

About *Boehmeria* genus, studies were found referring to nine species (Tab. S1, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>), eight studies done with the leaves and only one with the root. *Boehmeria nivea* is the species most studied. In this genus, twelve phenolic acids, four flavonols, one sterol, one alkaloid and one monoterpene were identified (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>). The phenolic acid class (64%) was the most found, followed by flavonols (21%) (Fig. 2), which were mainly present in the leaves. To a lesser extent, a steroid (β-sitosterol), an alkaloid [(-)-cryptopleurine] and a monoterpene [(-) -loliolide] were also identified in the leaves. In a study carried out on the roots of *B. nivea*, only four phenolic acids were identified. Aqueous solvents of methanol and ethanol were used for extraction from the leaves and ethyl acetate was used for extraction from the root.

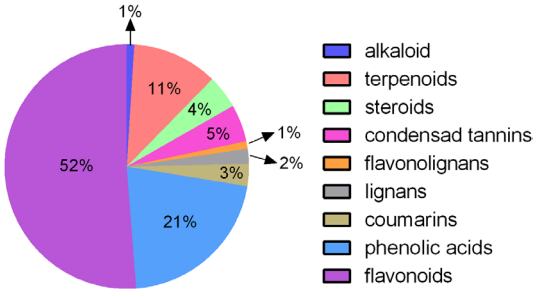


Figure 1 – Classes of secondary metabolites identified in the Urticaceae family by LC-MS.

Cecropia genus.

The genus *Cecropia*, the most studied to date, presents a diversity of types of secondary metabolites, such as flavonols, flavones, flavan-3-ols, condensed tannins, phenolic acids, flavonolignans, terpenoids (pentacyclic triterpenoids, iridoids, and triterpenoid saponins), and steroids (Tab. S2, available on supplementary material <https://doi.org/10.6084/

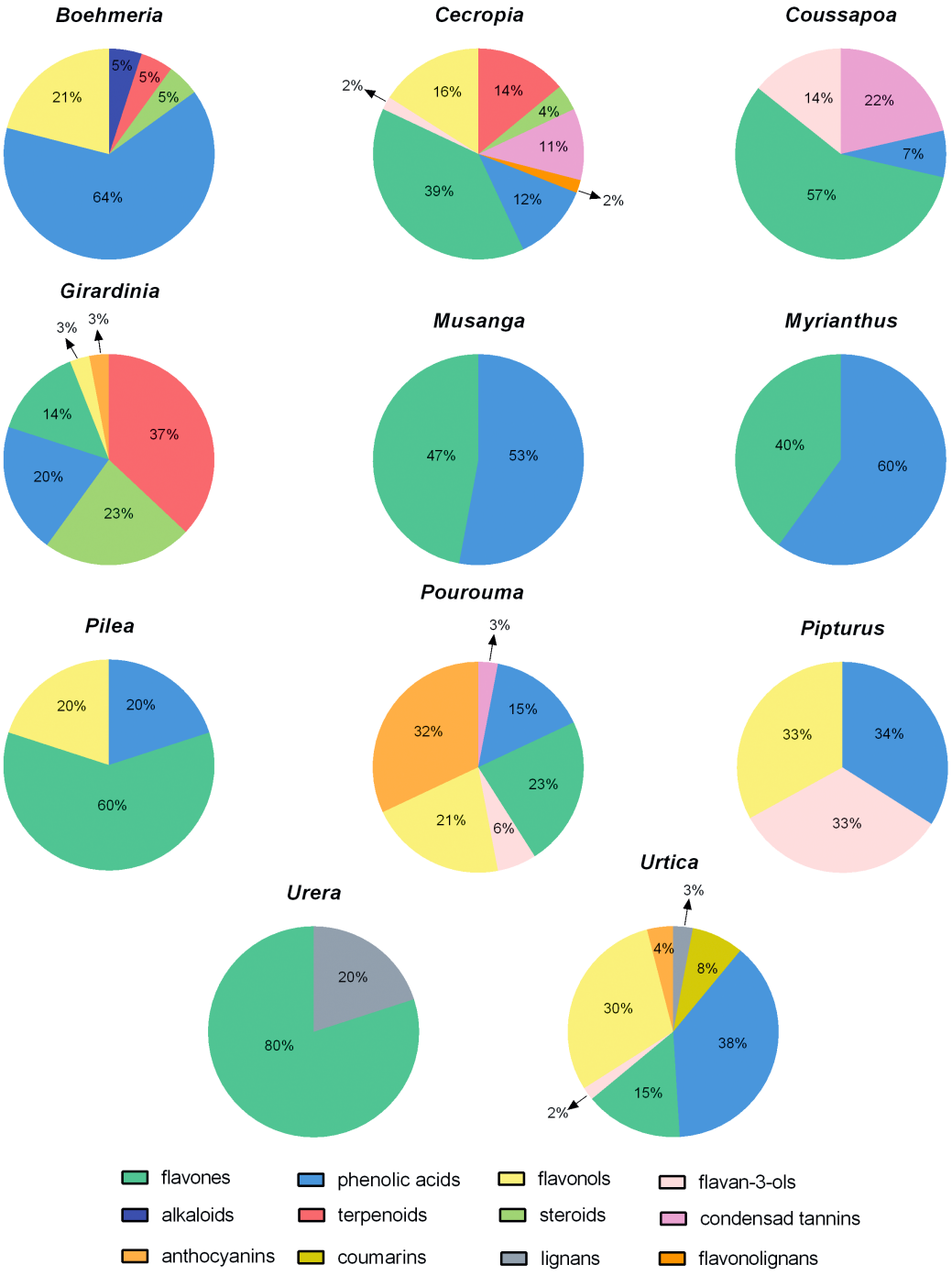


Figure 2 – Classes of secondary metabolites identified by LC-MS in genera of the Urticaceae family.

m9.figshare.23907960.v1>). All chemical metabolites were identified in the leaves of the ten species of the genus *Cecropia*. Phenolic substances are the most found in this genus (Fig. 2), flavones being the most abundant (39%). The aqueous and hydroethanolic extracts of the leaves were the most analyzed.

Coussapoa genus.

Only one study has been conducted on the genus *Coussapoa*. Flavones, phenolic acids, condensed tannins, and flavan-3-ols were identified in the methanolic extract of the leaves of the species *Coussapoa microcarpa* (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>). Phenolic substances such as flavones (57%) were predominant in the genus (Fig. 2).

Girardinia genus.

To date, only one study has been conducted on the genus *Girardinia*. In the methanolic extract made from the tips of the shoots of the species *Girardinia diversifolia*, the presence of flavonol, flavones, anthocyanins, phenolic acids, terpenoids (triterpenes, carotenoids, triterpenoid saponins, seco-iridoid glycosides), and steroids were evidenced. Of the nine genera found, the genus *Girardinia* was the one that presented a greater abundance of lipophilic substances. The most abundant secondary metabolite types were terpenoids (37%) steroids (23%), phenolic acids (20%), and flavones (14%), respectively (Fig. 2).

Musanga genus.

About genus *Musanga*, one literature record of *Musanga cecropioides* species was found. In the hydroethanolic extract of the stem bark, eight flavones and nine phenolic acids were identified (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>; Fig. 2).

Myrianthus genus.

Regarding genus *Myrianthus*, two flavones and three phenolic acids were identified from the hydroethanolic extract of the root bark from *Myrianthus arboreus* (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>; Fig. 2).

Pilea genus.

One literature record about *Pilea* genus was found. In the ethanolic extract of *Pilea microphylla*, one flavonol, one phenolic acid and three flavones were identified (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>; Fig. 2).

Pipturus genus.

Only one study of the genus *Pipturus* related to the *P. albidus* was found. In the hydromethanolic extract of the leaves, three phenolic substances (flavonol, flavan - 3 - ol and phenolic acid) were identified (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>; Fig. 2).

Pourouma genus.

About *Pourouma* genus, two studies were found with the fruits of the species *P. cecropifolia* and one study with the leaves of *P. guianensis* using methanol as extraction solvent. Four classes of phenylpropanoids (flavonols, flavones, flavan-3-ols, and anthocyanins), condensed tannins, and phenolic acids were identified, with most of these substances belonging to the anthocyanin class (32%), followed by flavones (23%) (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>; Fig. 2).

Urera genus.

A study carried out with the hydroethanolic extract of leaves of the *Urera baccifera* identified four flavones and one lignan (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>; Fig. 2).

Urtica genus.

The genus *Urtica* has been the subject of several studies, mainly concentrated on *Urtica dioica*. Secondary metabolites have been identified in different parts of species of this genus, such as flowers, stems and roots, and especially in the leaves. A large number of types of secondary metabolites have been identified in the genus *Urtica*, such as flavones, flavan-3-ols, anthocyanins, coumarins, lignans and phenolic acids, (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>). Phenolic substances such as phenolic acids (38%) and flavonols

(30%) are predominant in the genus (Fig. 2). The hydroethanolic and hydromethanolic extracts were the most analyzed.

Secondary metabolites identified by LC-MS

Flavonoids and phenolic acids were the most abundant compounds (52% and 21%, respectively). Therefore, this study presents a discussion of the fragmentations of the flavonoid structural variability and of the phenolic acid micromolecules identified by LC-MS.

Fragmentation pattern of flavones and flavonols

The main classes of flavonoids in the Urticaceae family are flavones and flavonols, attached to one or more sugar units, such as *O*- or *C*-glycosides. The flavone aglycones derived from apigenin, diosmetin and luteolin, while the flavonol aglycones derived from isorhamnetin, kaempferol, quercetin and myricetin (Fig. 3).

The flavonoids *O*-glycosides have sugar units substituents attached to an aglycone hydroxyl group, usually at positions *C*-3 or *C*-7, while in the case of *C*-glycosides flavonoids, sugars are linked to aglycone by a carbon-carbon bond at the *C*-6 or *C*-8 position (Waksmundzka-Hajnos & Sherma 2011).

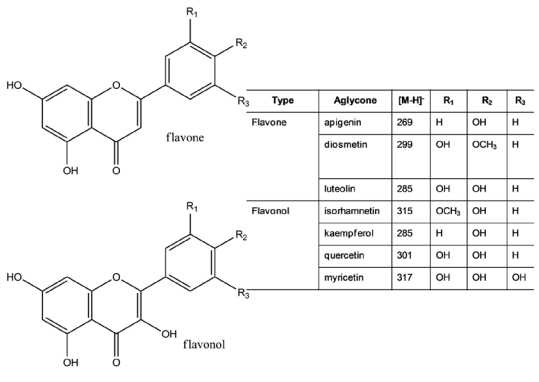
In *O*-glycoside flavonoids, fragmentation of the entire sugar unit is usually observed, because the energy required for breaking the hemiacetal C-O bonds is low (Vukics & Guttman 2010) (Fig. 4).

This type of fragmentation can be observed in data from some of the flavonoids *O*- glycosides identified in the Urticaceae family. Quercetin 3-*O*-hexoside (*m/z* 463) and luteolin-7-*O*-hexoside (*m/z* 447) generate the fragments *m/z* 301 e *m/z* 285, respectively, by the loss of a hexoside unit (162 u), as shown in Figure 5.

In the Urticaceae family, intraglycosidic cleavage commonly occurs in flavonoid *C*-glycosides, requiring high energy to break the C-C bond. Figure 6 shows the main fragmentations that occur in these cases. It is important to point out that this type of fragmentation can also be observed in *O*-glycosides, but less frequently (Vukics & Guttman 2010).

The flavone diosmetin-*C*-hexoside (*m/z* 461), identified in the *Cecropia* genus, serves as an example of this kind of fragmentation. Its most abundant fragments were generated by intraglycosidic cleavage, the base peak *m/z* 341 ($^{0,2}X^-$) and *m/z* 371 ($^{0,3}X^-$), as illustrated in Figure 7.

It is possible to differentiate the *C*-glycosylated isomers at positions *C*-6 and *C*-8, since the 6-*C* isomer has a greater abundance $^{0,3}X^-$ ion than the 8-*C* isomer (Becchi & Fraisse 1989). Two isomers present in the Urticaceae family, vitexin (apigenin 8-*C*- glucoside, *m/z* 431) and isovitexin (apigenin 6-*C*- glucoside, *m/z* 431), corroborate with Becchie & Fraisse (1989). As shown in the data collected (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>), the ion



Type	Aglycone	[M-H]⁻	R_1	R_2	R_3
Flavone	apigenin	269	H	OH	H
	diosmetin	299	OH	OCH_3	H
	luteolin	285	OH	OH	H
Flavonol	isorhamnetin	315	OCH_3	OH	H
	kaempferol	285	H	OH	H
	quercetin	301	OH	OH	H
	myricetin	317	OH	OH	OH

Figure 3 – Types of aglycone flavones and flavonols present in the Urticaceae family.

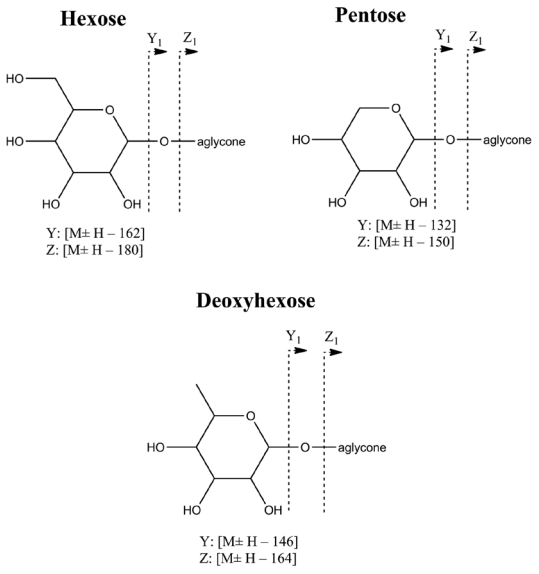


Figure 4 – Representation of the most frequent fragmentations in *O*-glycosides with their respective mass losses (adapted) (Domon & Costello 1988).

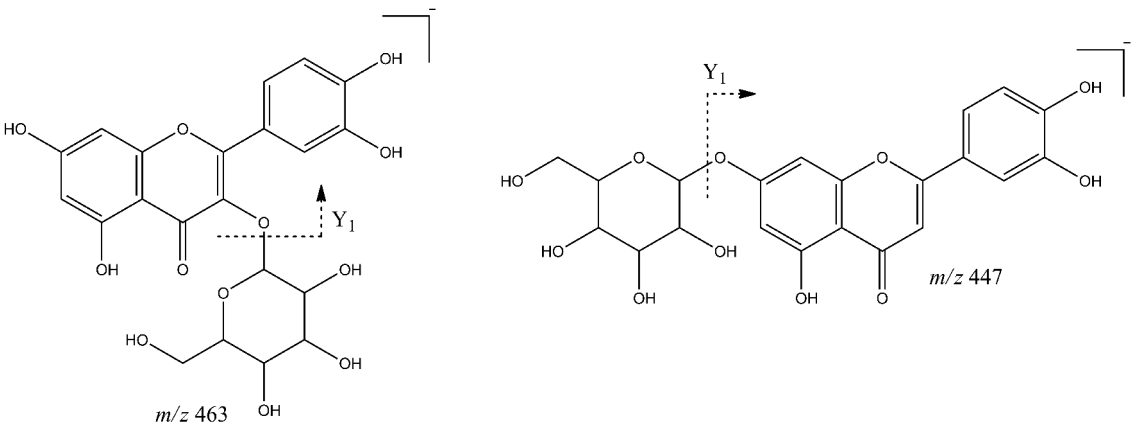
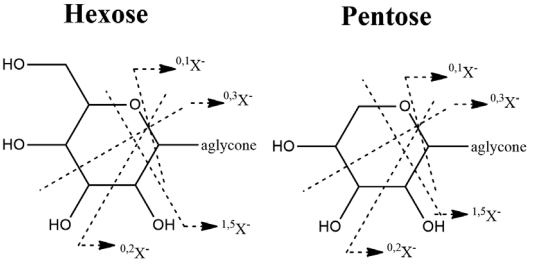


Figure 5 – Representation of the main fragmentation of quercetin 3-*O*-hexoside and luteolin-7-*O*-hexoside.

m/z 341 ($^{0,3}X^-$) has a higher intensity in isovitexin (54% - *Cecropia* and 48% *Coussapoa*) compared to vitexin (11% - *Cecropia* and 8% *Coussapoa*) (Fig. 8). Also, it is possible to differentiate the isomers orientin (luteolin-8-*C*-glucoside *m/z* 447) and isoorientin (luteolin-6-*C*-glucoside *m/z* 447) (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>). In isoorientin, the ion *m/z* 357 ($^{0,3}X^-$) is more abundant (100% - *Cecropia* and *Coussapoa*) than in orientin (34% - *Cecropia* and 40% *Coussapoa*) (Fig. 8) (Vukics & Guttman 2010).



$^{0,1}X$: [M± H - 150]
$^{0,2}X$: [M± H - 120]
$^{0,3}X$: [M± H - 90]
$^{1,5}X$: [M± H - 134]
$^{2,3}X$ - $2H_2O$: [M± H - 66]
$^{0,4}X$ - $2H_2O$: [M± H - 96]
$^{0,2}X$ - H_2O: [M± H - 138]
$^{0,2}X$ - $2H_2O$: [M± H - 156]
$^{2,3}X$ -$3H_2O$: [M± H - 84]

$^{0,1}X$: [M± H - 120]
$^{0,2}X$: [M± H - 90]
$^{0,3}X$: [M± H - 60]
$^{1,5}X$: [M± H - 104
$^{0,4}X$ - $2H_2O$: [M± H - 66]
$^{0,2}X$ - H_2O: [M± H - 108]
$^{0,2}X$ - $2H_2O$: [M± H - 126]

Figure 6 – Representation of the most frequent intraglycosidic fragmentations in *C*-glycosides with their respective mass losses (adapted) (Cuyckens & Claeys 2004; Vukics & Guttman 2010).

The flavonoid aglycone type is identified through the fragmentation spectrum, but this requires further data analysis. One way of analysis is by cleavage of the C-C bonds by the retro-Diels-Alder (RDA) mechanism of the C-ring of the aglycone. Mass losses related to water, methyl, CO_2 and CO and successive losses of these molecules are also observed and important in identifying the specific functional groups (Fig. 9) (Fabre *et al.* 2001; Benayad *et al.* 2014; Treiber *et al.* 2016; Villiers *et al.* 2016).

According to Vukins & Guttman (2010), although $^{1,3}A^+$ and $^{0,2}B^+$ ions are common for both classes, the $^{1,3}B^+$ ion was only observed in flavones, while the $^{1,3}B^+$-2H ion was detected only in flavonols. In the study developed by Ma *et al.* (1997) it was reported that the $^{0,4}B^+$ and $^{0,4}B^+$-H_2 fragments are characteristic for flavones and $^{0,2}A^+$ for flavonols (Ma *et al.* 1997). Fabre *et al.*

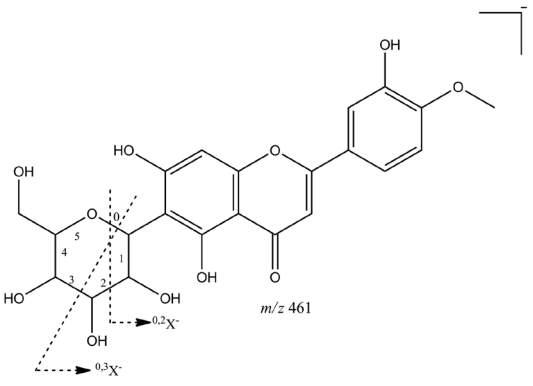


Figure 7 – Representation of diosmetin-*C*-hexoside fragments.

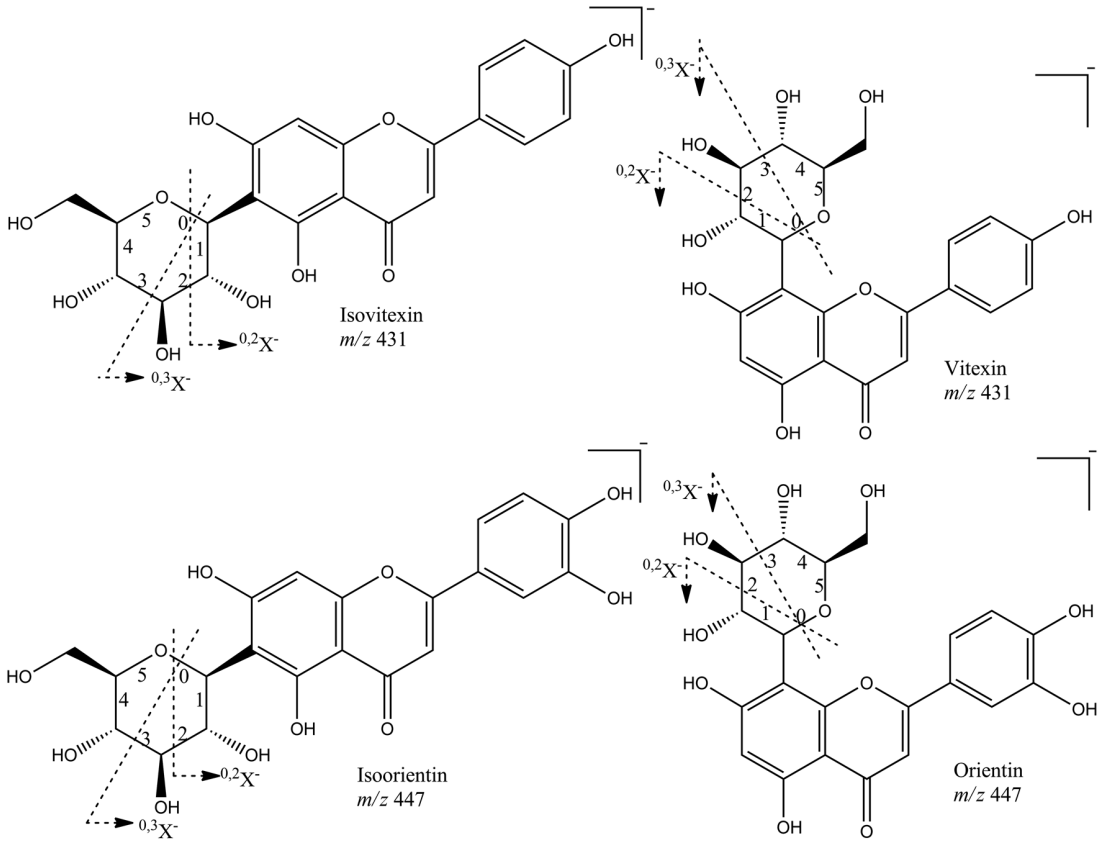


Figure 8 – Representation of the fragmentations of the flavonoids isovitexin, vitexin, isoorientin and orientin.

(2001) showed the differentiation of flavones and flavonols, distinguishing these structures by $^{1,2}B^-$ and $^{1,2}A^-$ ions for flavonols and $^{1,3}B^-$ and $^{1,3}A^-$ ions for flavones. As shown in Table S2 (available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>), the flavones luteolin (*m/z* 285) and apigenin (*m/z* 269) presented the fragmentations *m/z* 133 e *m/z* 117, respectively, which is characteristic for flavone ($^{1,3}B^-$) (Fig. 10). Quercetin (*m/z* 301) and myricetin (*m/z* 317) (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>) originated fragment $^{1,2}A^-$ (*m/z* 179) characteristic of flavonols, besides the observance of, ion 151 ($^{1,2}A^-$-CO) (Fig. 10). The flavonol kaempeferol identified in the Urticaceae family (*m/z* 285) generated only the fragment *m/z* 163 ($^{0,2}A^-$) (Fig. 10). The literature reports more fragmentations for kaempeferol, such as the ion *m/z*151 ($^{1,2}A^-$-CO) (Fabre *et al.* 2001).

Two flavan-3-ols units have been identified in the Urticaceae family, catechin (*m/z* 289) and epicatechin (*m/z* 289). These molecules exhibit characteristic fragments of this flavonoid subclass as *m/z* 245 ([M-H-44]$^-$), loss of one CH2=CHOH group and *m/z* 205 ([M-H-84]$^-$), loss of two molecules of HC≡COH, arising from heterocyclic ring fission. In addition to these ions, the presence

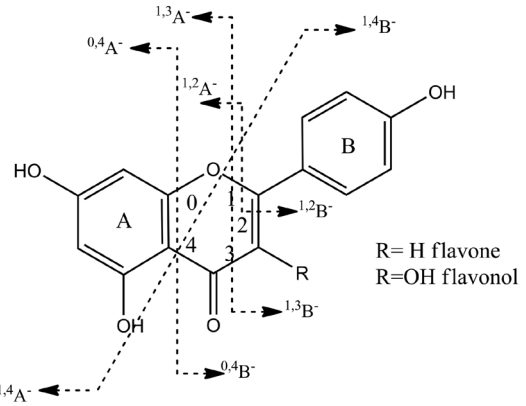


Figure 9 – Representation of retro-Diels-Alder fragmentations aglycone flavonoids followed by numbering at C-ring (adapted) (Fabre *et al.* 2001).

of the ion *m/z* 179 ([M-H-110]$^-$), characteristic of these structures, originates from the loss of a catechol unit (Fig. 11) (Cui *et al.* 2007). It was not possible to observe differences between these isomers due to fragmentation characteristics. However, it was observed in studies done with species of the Urticaceae family that the catechin was eluted first (Tanae *et al.* 2007; Lopes-Lutz *et al.* 2010; Cruz *et al.* 2013).

Fragmentation patterns of anthocyanins

The anthocyanins identified in the Urticaceae family are mainly concentrated in the fruits of the *Pourouma* genus (Tab. S2, available on supplementary material <https://doi.org/10.6084/m9.figshare.23907960.v1>), and consist of six types of aglycones such as: cyanidin (*m/z* 287), delphinidin (*m/z* 303), malvidin (*m/z* 331), pelargonidin (*m/z* 271),

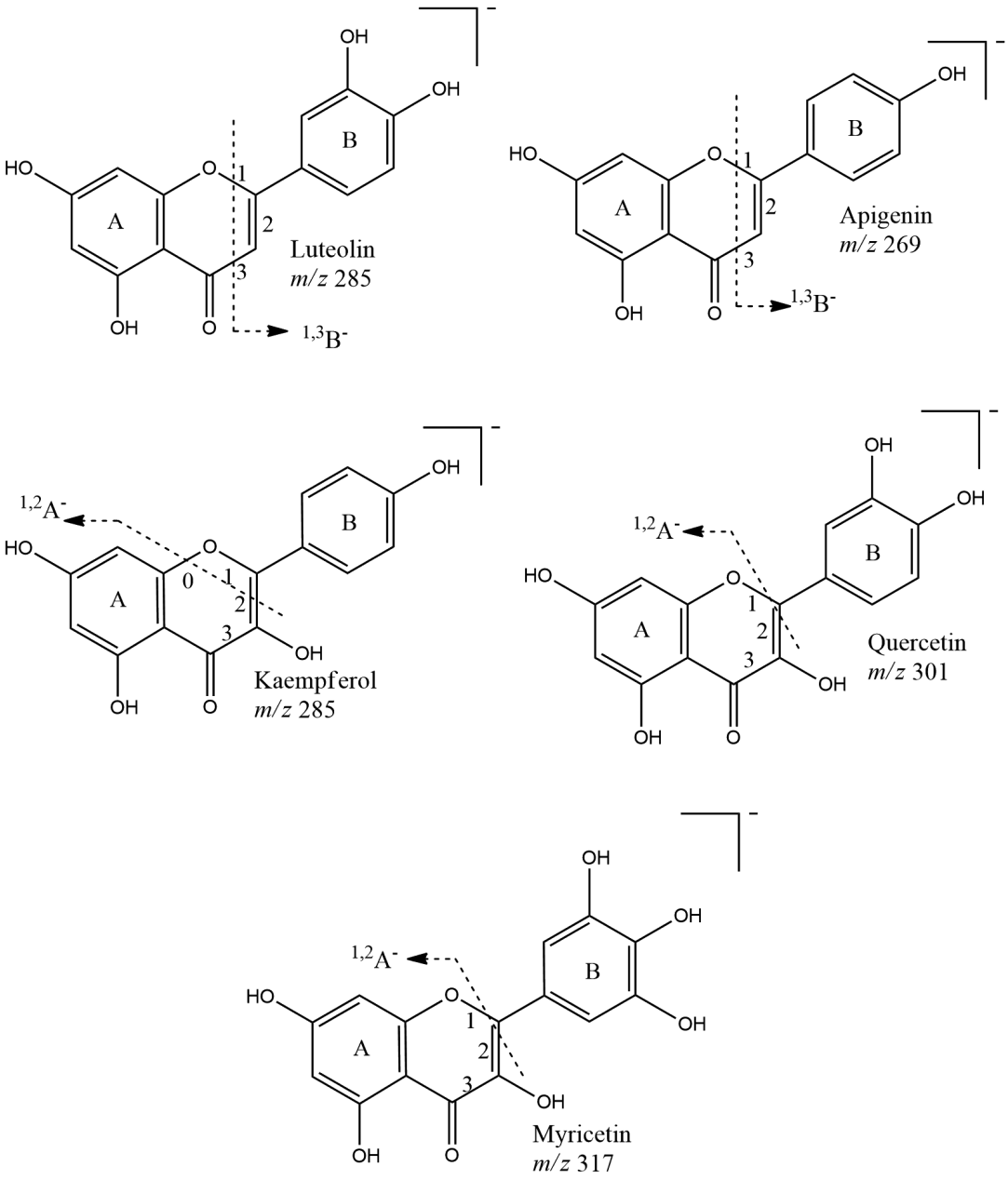


Figure 10 – Fragmentation pathways of the flavonoids luteolin, apigenin, kaempferol, quercetin and myricetin.

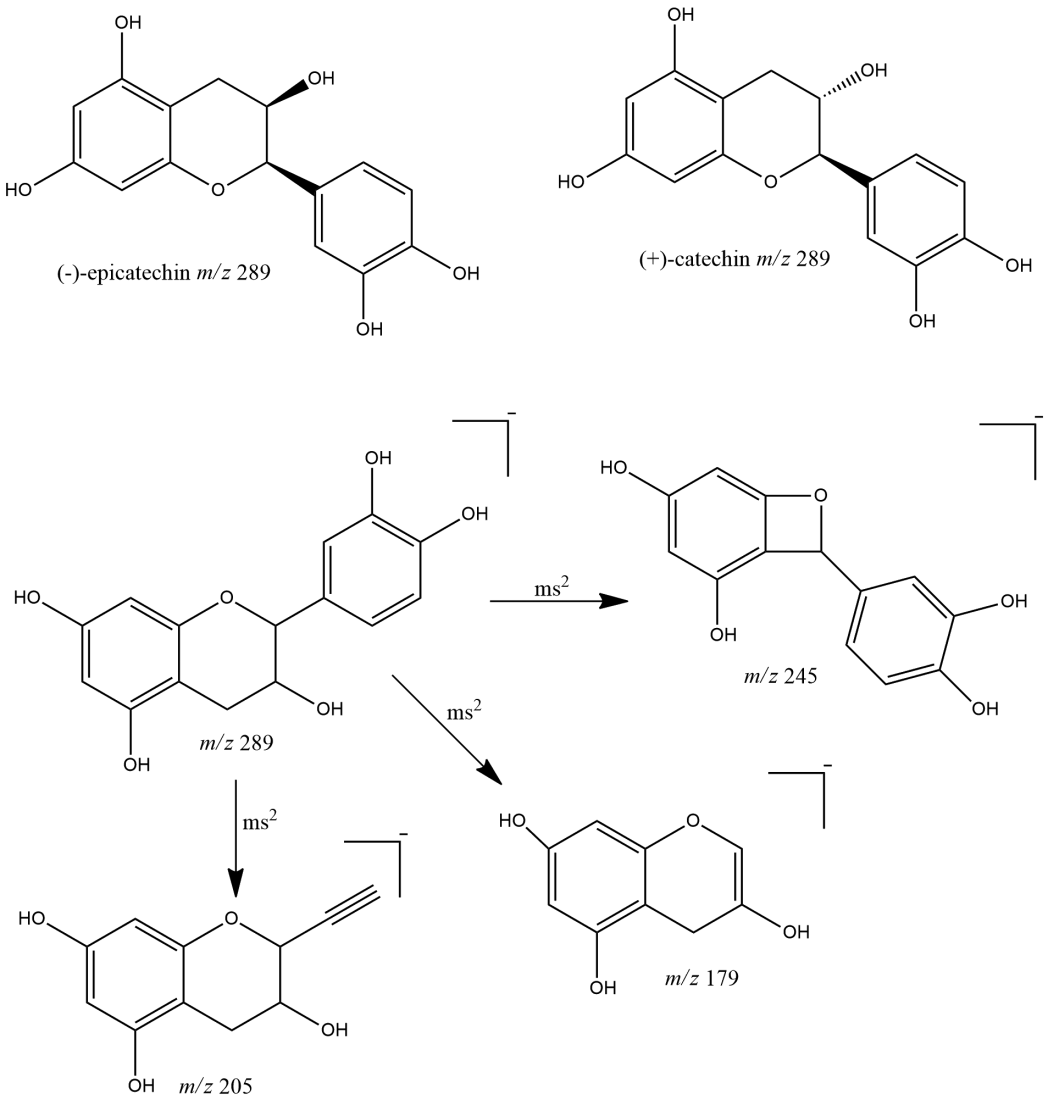
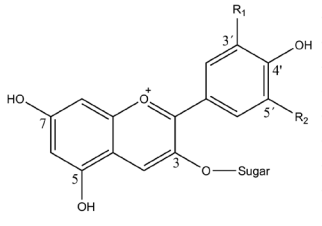


Figure 11 – Proposed fragmentation pathways for catechin and epicatechin.

petunidin (*m/z* 317) and peonidin (*m/z* 301). The main fragmentation identified in the anthocyanins corresponds to the aglycone ion, generated by sugar loss (Fig. 12).



R_1	R_2	Anthocyanidin
H	H	pelargonidin
OH	H	cyanidin
OH	OH	delphinidin
OH	OCH_3	petunidin
OCH_3	H	peonidin
OCH_3	OCH_3	malvidin

Figure 12 – Structural representation of anthocyanins.

Fragmentation patterns of phenolic acids

Phenolic acids represent the second most recorded class of micromolecules in Urticaceae in the period analyzed. Mainly the derivatives of cinnamic acid stand out. Several types of chlorogenic acids have been identified, acids formed by the esterification of quinic acid with the following trans-arylpropionic acids: caffeic, ferulic, and *p*-coumaric acids.

The 5-*O*- caffeoylquinic, 4-*O*- caffeoylquinic and 3-*O*-caffeoylquinic acids present *m/z* 353 which corresponds to the molecular ion in the negative mode of caffeoylquinic acid (Fig.

13). The differentiation between these isomers is based on the fragmentation profile by MS/MS. The isomers at the C-3 and C-5 positions are characterized by the base peak *m/z* 191 ([quinic acid–H]-), relative to quinic acid, whereas the C-3 isomer has other fragments relative to quinic acid and the corresponding cinnamic acid. The isomer 4 provides as the base peak *m/z* 173 ([quinic acid-H_2O-H]$^-$) (Clifford *et al*. 2003).

Two isomers of feruloylquinic acid (*m/z* 367) are present in the Urticaceae family (Fig. 13), the 5-*O*- feruloylquinic, which has as base peak the ion *m/z* 191, and o 4-*O*- feruloylquinic with as base peak the ion *m/z* 173 (Clifford *et al.* 2003). The *p*-coumaroylquinic acid (*m/z* 337) is fragmented at *m/z* 191 ([quinic acid-H]$^-$) and at *m/z* 173 ([quinic acid-H_2O-H]$^-$) (Fig. 14).

Arylpropionic acids have also been identified. While ferulic acid presents as base fragmentation the ion *m/z* 134 on the elimination of CO_2 and of the methyl group (by radical fragmentation), for caffeic acid the ion *m/z* 135 refers only to the loss of CO_2, resulting from decarboxylation. The illustration of these fragmentation steps is shown in Figure 15.

The *p*-coumaric acid (*m/z* 163) has as fragmentation the *m/z* 119, formed by the loss of CO_2 ([*p*-coumaric acid-CO_2-H]$^-$). On the other hand, *p*-coumaroylhexoside acid (*m/z* 325) shows the fragmentations *m/z* 163, formed by the loss of sugar and *m/z* 119 ([*p*-coumaric acid-H-CO_2]$^-$), originating from the loss of CO_2 (Fig. 16). Synapic acid (*m/z* 223) has as main fragmentation the ion *m/z* 193, resulting from the successive loss of two

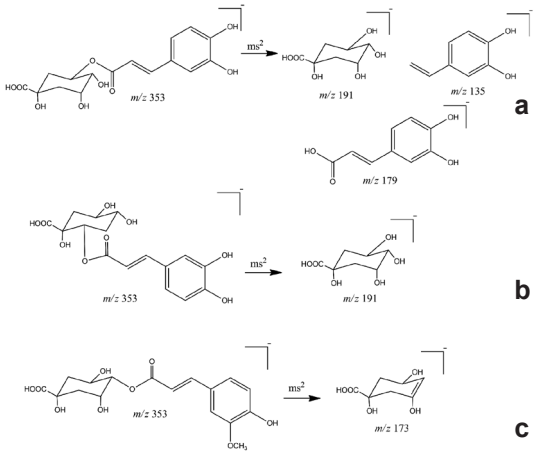


Figure 13 – a-c. Proposed fragmentation pathways for the isomers – a. 3-*O*-caffeoylquinic; b. 5-*O*-caffeoylquinic; c. 4-*O*-caffeoylquinic.

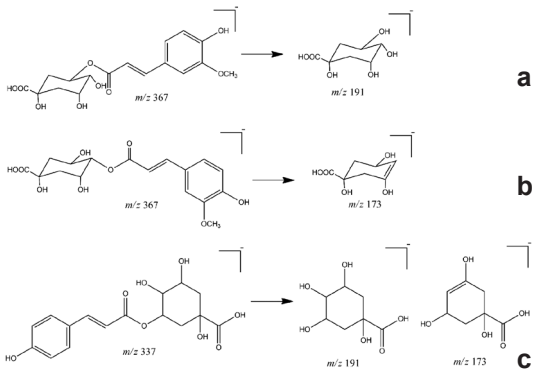


Figure 14 – Proposed fragmentation pathways for the isomers – a. 5-*O*-feruloylquinic; b. 4-*O*-feruloylquinic; c. *p*-coumaroylquinic acid.

methyl groups; whereas sinapichexoside acid (*m/z* 385) presents as fragmentation the *m/z* 223 ion ([sinapic acid-H]$^-$), formed by the loss of hexose (Fig. 16).

Arylpropionic acids, such as *p*-coumaric, caffeic and ferulic acids were also found esterified with malic acid. Acid *p*-coumaroylmalic (*m/z* 279) has the following fragmentation ions: *m/z* 163 ([p-coumaric acid-H]$^-$), *m/z* 133 ([malic acid-H]$^-$) and in *m/z* 119 ([p-coumaric acid-H-CO_2]$^-$). The caffeoylmalic acid (*m/z* 295) shows fragmentations at *m/z* 179 ([caffeic acid-H]$^-$) and at *m/z* 133, resulting from malic acid. Feruloylmalic acid (*m/z* 309) presented as pattern fragmentation *m/z* 193 (Fig. 17).

Benzoic acid derivatives have also been identified in the Urticaceae family, such as gallic, protocatechuic, gentisic acid, vanillic and syringic acids (Fig. 18). The main fragmentation of gallic (*m/z* 169), gentisic acid (*m/z* 153) and protocatechuic (*m/z* 153) acids is generated by the

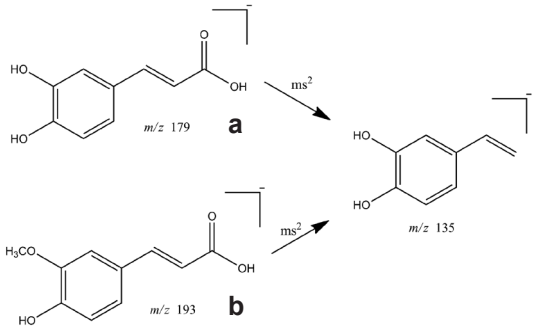


Figure 15 – a-b. Proposed fragmentation pathways for acids – a. caffeic; b. ferulic.

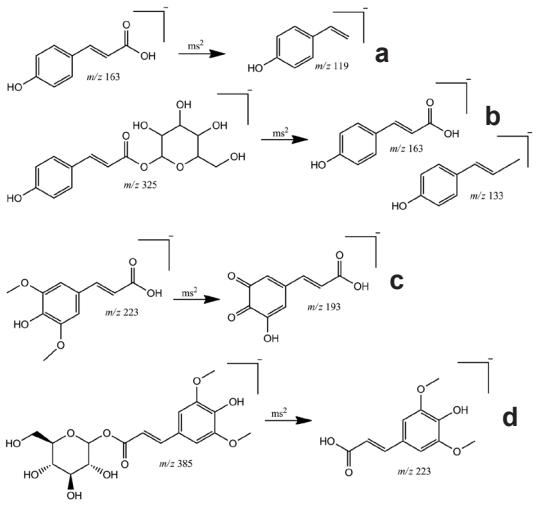


Figure 16 – a-d. Proposed fragmentation pathways – a. for *p*-coumaric acid; b. for *p*-coumaroylhexoside acid; c. for sinapic acid; d. for sinapichexoside acid.

loss of a group of CO_2 $[M-H-44]^-$ at *m/z* 125, *m/z* 109 and *m/z* 109, respectively. Syringic (*m/z* 197) and vanillic (*m/z* 167) acids generate *m/z* 182 and *m/z* 152 fragmentation, respectively, through the loss of a methyl group from the precursor ion, providing an anionic radical $[M-H-CH3]^-$. The fragmentation of vanillic acid further a loss of a CO_2 group, generating the ion *m/z* 108 (Fig. 18).

Based on the survey, it was found that of the 54 genera of the Urticaceae family, only 11 genera had phytochemical studies using the LC-MS, and seven of these genera were found in Brazil. Most

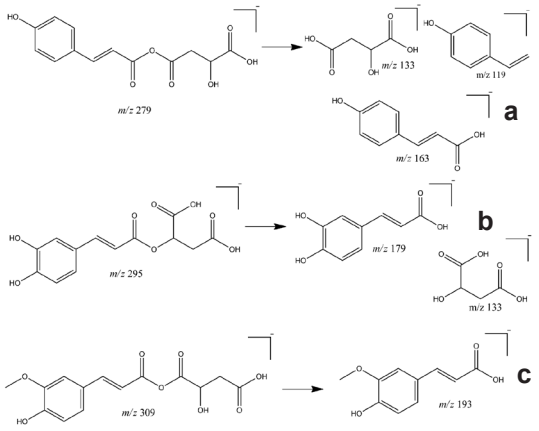


Figure 17 – a-c. Proposed fragmentation pathways – a. for *p*-coumaroylmalic acid; b. for caffeoylmalic acid; c. for feruloylmalic acid.

of the studies have focused on the *Cecropia* and *Urtica* genera. Therefore, although the family has been little explored from a phytochemical point of view, from the articles found, it could be concluded that the studies are concentrated on the leaf part of the species and that the phenolic substances are the majority in the Urticaceae family, mainly belonging to the flavone, phenolic acids, and flavonol classes. Detailed fragmentation analysis of some phenolic metabolites allows chemical identification without the use of standards and can even differentiate certain isomers. Therefore, LC-MS provides a faster and more efficient characterization of the chemical profile of plant species. Finally, with its little explored biodiversity, promising perspectives

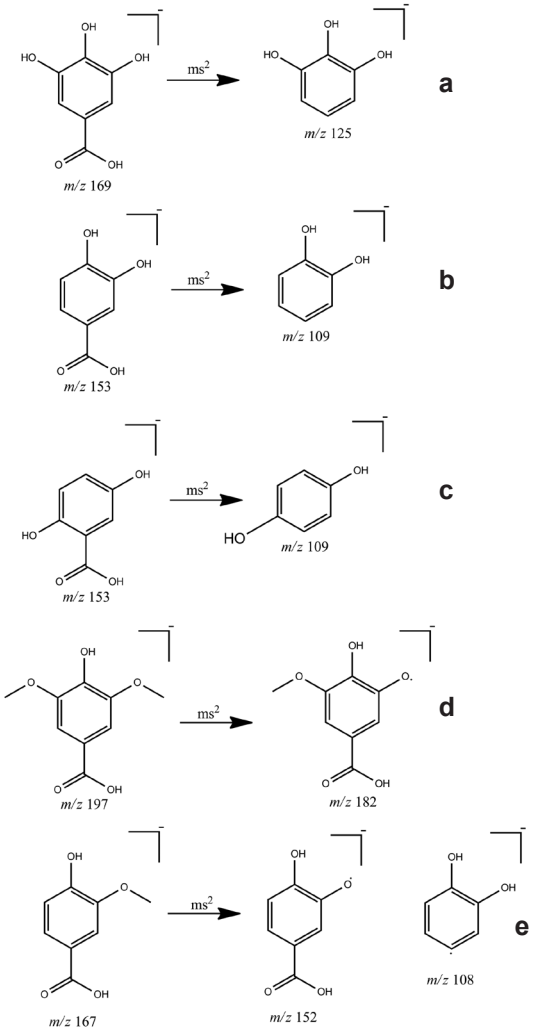


Figure 18 – a-e. Proposed fragmentation pathways – a. for gallic acid; b. for protocatechuic acid; c. for gentisic acid; d. for siryngic acid; e. for vanillic acid.

for the study of Urticaceae family species are expected in the coming years.

Acknowledgements

We thank CAPES, for the scholarship; and FAPERJ and UENF, for the financial support.

Data availability statement

In accordance with Open Science communication practices, the authors inform that there is no data sharing of this manuscript.

References

Alvarez-Rivera G, Ballesteros-Vivas D, Parada-Alfonso F, ElenaIbañez A & Cifuentes A (2019) Recent applications of high resolution mass spectrometry for the characterization of plant natural products. Trends in Analytical Chemistry 112: 87-101.

Becchi M & Fraisse D (1989) Fast atom bombardment and fast atom bombardment collision-activated dissociation/MS-analysed ion kinetic energy analysis of C-glycosidic flavonoids. Biomedical & Environmental Mass Spectrometry 18: 122-130.

Benayad Z, Gómez-Cordovés C & Es-Safi NE (2014) Characterization of Flavonoid Glycosides from Fenugreek (*Trigonella foenum-graecum*) Crude Seeds by HPLC-DAD-ESI/MS Analysis. International Journal of Molecular Sciences 15: 20668-20685.

Cajka T & Fiehn O (2016) Toward merging untargeted and targeted methods in mass spectrometry-based metabolomics and lipidomics. Analytical Chemistry 88: 524-545.

Clifford MN, Johnston KL, Knight S & Kuhnert N (2003) Hierarchical scheme for LC-MSn identification of chlorogenic acids. Journal of Agricultural and Food Chemistry 51: 2900-2911.

Costa CT, Horton D & Margolis SA (2000) Analysis of anthocyanins in foods by liquid chromatography, liquid chromatography-mass spectrometry and capillary electrophoresis. Journal Chromatography A 881: 403-410.

Cruz EM, Silva ER, Carmo Maquiaveli C, Alves ESS, Lucon Júnior JF, Reis MBG, Toledo CEM, Cruz FG & Vannier-Santos MA (2013) Leishmanicidal activity of *Cecropia pachystachya* flavonoids: arginase inhibition and altered mitochondrial DNA arrangement. Phytochemistry 89: 71-77.

Cui SY, Weng JF, Kim SJ & Lee YI (2007) Investigation of the interaction between Sodium (meta) Arsenite and Catechin *via* ESI Tandem mass spectrometry. Chemical Research in Chinese Universities 23: 524-529.

Cuyckens F & Claeys M (2004) Mass spectrometry in the structural analysis of flavonoids. Journal of Mass Spectrometry 39: 1-15.

Domon B & Costello C (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconjugate Journal 5: 397-409.

Dührkop K, Nothias LF, Fleischauer M, Reher R, Ludwig M, Hoffmann MA, Petras D, Gerwick WH, Rousu J, Dorrestein PC & Böcker S (2021) Systematic classification of unknown metabolites using high-resolution fragmentation mass spectra. Nature Biotechnology 39: 462-471.

Fabre N, Rustan I, Hoffmann E & Quetin-Leclercq J (2001) Determination of flavone, flavonol, and flavanone aglycones by negative ion liquid chromatography electrospray ion trap mass spectrometry. Journal of the American Society for Mass Spectrometry 12: 707-715.

Fraige K, Dametto AC, Zeraik ML, Freitas L, Saraiva AC, Medeiros AI, Castro-Gamboa I, Cavalheiro AJ, Silva DHS, Lopes NP & Bolzani VS (2018) Dereplication by HPLC-DAD-ESI-MS/MS and screening for biological activities of *Byrsonima* species (Malpighiaceae). Phytochemical Analysis 29: 196-204.

Kruve A (2020) Strategies for drawing quantitative conclusions from nontargeted liquid chromatography-high-resolution mass spectrometry analysis.Analytical Chemistry 92: 4691-4699.

Lanças FM (2009) A cromatografia líquida moderna e a espectrometria de massas: finalmente "compatíveis"? ScientiaChromatographica 1: 27-46.

Lopes-Lutz D, Dettmann J, Nimalaratne C & Schieber A (2010) Characterization and quantification of polyphenols in Amazon grape (*Pourouma cecropiifolia* Martius). Molecules 15: 8543-8552.

Ma YL, Li QM, Van den Heuvel H & Claeys M (1997) Characterization of flavone and flavonol aglycones by collisioninduced dissociation tandem mass spectrometry. Rapid Communications in Mass Spectrometry 11: 1357-1364.

Patel KN, Patel JK, Patel MP, Rajput GC & Patel HA (2010) Introduction to hyphenated techniques and their applications in pharmacy. Pharmaceutical Methods 1: 2-13.

Pinelli P, Ieri F, Vignolini P, Bacci L, Baronti S & Romani A (2008) Extraction and HPLC analysis of phenolic compounds in leaves, stalks, and textile fibers of *Urtica dioica* L. Journal of Agricultural and Food Chemistry 56: 9127-9132.

Rivera-Mondragón A, Bijttebier S, Tuenter E, Custers D, Ortíz OO, Pieters L, Caballero-George C, Apers S & Foubert K (2019) Phytochemical characterization and comparative studies of four *Cecropia* species collected in Panama using multivariate data analysis. Scientific Reports 9: 1-14.

Seraglio SKT, Valese AC, Daguer H, Bergamo G, Azevedo MS, Gonzaga LV, Fett R & Costa ACO (2016) Development and validation of a LC-ESI-

MS/MS method for the determination of phenolic compounds in honeydew honeys with the diluted-and-shoot approach. Food Research International 87: 60-67.

Shrestha SS, Sut S, Ferrarese I, Barbondi Marco S, Zengin G, Franco M, Pant DR, Mahomoodally MF, Ferri N, Biancorosso N, Maggi F, Acqua SD & Rajbhandary S (2020) Himalayan nettle *Girardinia diversifolia* as a candidate ingredient for pharmaceutical and nutraceutical applications-phytochemical analysisand *in vitro* bioassays. Molecules 25: 1563.

Syage JA, Short LC & Cai SS (2008) Atmospheric pressure photoionization - the second source for LC-MS?. LCGC North America 26: 286-300.

Tanae MM, Lima-Landman MT, Lima TC, Souccar C & Lapa AJ (2007) Chemical standardization of the aqueous extract of *Cecropia glaziovii* Sneth endowed with antihypertensive, bronchodilator, antiacid secretion and antidepressant-like activities. Phytomedicine 14: 309-13.

Treiber EL, Gaglioti AL, Romaniuc-Neto S, Madriñán S & Weiblen GD (2016) Phylogeny of the Cecropieae (Urticaceae) and the evolution of an ant-plant mutualism. Systematic Botany 41: 56-66.

Villiers A, Venter P & Pasch H (2016) Recent advances and trends in the líquid-chromatography - mass spectrometry analysis of flavonoids. Journal of Chromatography A 1430: 16-78.

Vukics V & Guttman A (2010) Structural characterization of flavonoid glycosides by multi-stage mass spectrometry. Mass Spectrometry Reviews 29: 1-16.

Waksmundzka-Hajnos M & Sherma J (2011) High performance liquid chromatography in phytochemical analysis. CRC Press, Boca Raton. 996p.

Area Editor: Dr. Leopoldo Baratto
Received on October 25, 2022. Accepted on March 22, 2023.