Phytochemical Analysis of the methanolic extract of tigernut, tuber of *Cyperus esculentus*, by ultra-high performance liquid chromatography coupled with electrospray ionization-quadrupole-time of flight-mass spectrometry (UHPLC/ESI-Q-TOF-MS)

Caracterización del Perfil Fenólico del extracto metanólico del tubérculo de la chufa (*Cyperus esculentus*) por cromatografía líquida de alta resolución acoplada a espectrometría de masas (UHPLC/ESI-Q-TOF-MS)

Máster Interuniversitario en Descubrimiento de Fármacos

Trabajo Fin de Máster

Laura Soto Mayer

Directora: Ana Gradillas Nicolás

Cembio, Departamento de Química y Bioquímica
Facultad de Farmacia, Universidad CEU San pablo
Madrid, June 2019

CEU Universidad San Pablo

UNIVERSIDAD COMPLUTENSE MADRID

Universidad de Alcalá

Autorización de la presentación del TFM en Descubrimiento de Fármacos

Curso 2018-2019

Nombre y apellidos Directora del TFM: Ana Gradillas Nicolás

Categoría Profesional: Profesora Titular de Química Orgánica y Farmacéutica

Departamento/Unidad: Dpto. Química y Bioquímica

Centro: Facultad de Farmacia, Universidad CEU San Pablo

CERTIFICA:

Que el trabajo titulado:

"Caracterización del perfil fenólico del extracto metanólico del tubérculo de chufa (Cyperus esculentus) por Cromatografía Líquida de Alta Resolución Acoplada a Espectrometría de Masas (UHPLC/ESI-Q-TOF-MS)"

que ha realizado **Dña. Laura Soto Mayer** como Trabajo Fin de Máster para el Máster Interuniversitario en Descubrimiento de Fármacos, ha sido realizado en los laboratorios de este Departamento bajo su dirección, y que autoriza su presentación.

En Madrid, a 25 de junio de 2019

CEU Universidad San Pablo

Firma y sello

INDEX

AGRADECIMIENTOS

Este Trabajo de Fin de Master no podría haber sido desarrollado sin el aprendizaje adquirido a lo largo de las diferentes asignaturas del máster, las cuales siempre han aportado una perspectiva diferente y una visión más crítica en el descubrimiento de fármacos.

A nivel académico, quería agradecer a mi tutora, Ana Gradillas, por su dedicación y labor en la realización de este trabajo y del cual he adquirido nuevos conocimientos en el campo de la química despertando mí interés por la investigación.

Por último, y no menos importante, quería agradecer a mi familia y amigos por su apoyo y cariño incondicional.

ABSTRACT

In the present work a Spanish local variety of *Cyperus esculentus* (chufa Appellation of origin Spanish province of Valencia) was studied. The main phenolic compound profile of chufas was obtained using ultra high performance liquid chromatography coupled to a quadrupole-time-of-flight mass spectrometry (UHPLC/ESI-Q-TOF-MS). The assayed chromatographic method enables a rapid and extraordinary efficient separation, allowing easy identification of the three main phenolic compound groups present in the methanolic extracts: phenylethanoids, phenylpropanoids and flavonoids. We also investigated, for the characterization of some other compounds, the possibility of using of in-source fragmentation (ISF) with electrospray ionization (ESI) followed by product ion scan in the QTOF mass spectrometer system. This approach is based on the elucidation of the various fragmentation pathways by further dissociation of each fragment ion in the ion spectrum. This can be achieved predominately, by combining fragmentor voltage induced dissociation (in-source fragmentation IS-CID) with subsequent collision-induced dissociation (CID). 37 compounds (3 amino acids and derivatives, 6 phenolic acids and derivatives, 16 flavonoids, and 8 fatty acids) were either unambiguously indentified or tentatively characterized. One, *N*-malonyl-*L*-tryptophan, identified for the first time in tigernut.

ABBREVIATIONS

APCI: atmospheric-pressure chemical ionization

BPC: base peak chromatogram

CID: collision induced-dissociation

EFAs: essential fatty acids

EIC: extract ion chromatogram

ESI: electrospray ionization

GC-MS: gas chromatography-mass spectrometry

HPLC: high-pressure liquid chromatography

IR: infrared spectroscopy

ISF: in-source fragmentation

MS: mass spectrometry

MUFAs: monounsaturated fatty acids

NMR: nuclear magnetic resonance

n.d.: non detected

PhGs: phenylethanoid glycosides

PUFAS: polyunsaturated fatty acids

QTOF: quadrupole-time of flight

UHPLC: ultra-high pressure liquid chromatography

V: voltage

1. INTRODUCTION

Natural products (also known as secondary metabolites) and traditional medicines are of great importance since they have been recognised for many years as a source of bioactive compounds with high structural diversity and important biological activities. About half of the drugs currently in clinical use are natural products or synthetic molecules based on natural product scaffolds [1].

Drug discovery using natural products is a challenging task for designing new viable lead compounds and there is a need to describe phytochemical analysis and pharmacological investigations. The important steps in drug discovery using natural products are the extraction, isolation and characterization of bioactive compounds from plants and the isolation of active molecules by chromatographic techniques [2].

It could be said that the most intellectually demanding step toward analysis of bioactive molecules present in the plant extracts is the structural characterization. In consequence, highly sophisticated techniques such as GC-MS, HPLC, HPTLC, FTIR and NMR are powerful tools that have been developed to support the identification and characterization of bioactive compounds. These techniques are the heart and the key challenges in research of natural drug discovery. They play an important role in natural drug discovery, and without them, it would be barely possible [2].

Among all techniques, ultra-high-performance liquid chromatography coupled to a quadrupole-time-of-flight mass spectrometry (UHPLC/ESI-QTOF-MS), has become a powerful tool in the characterization of complex extracts and has been widely used. Electrospray ionization (ESI) is a soft ionization method capable of providing both protonated and deprotonated molecules. Q-TOF-MS is able to combine high sensitivity and mass accuracy and therefore makes possible to confirm the elemental composition in a rapid and efficient way. Meanwhile, the application of UHPLC can provide high resolution for the separation of complicated natural products and improve the sensitivity of a Q-TOF-MS detector [3].

The present Master's Final Dissertation (TFM) has tried to provide information about the composition of the bioactive compounds present in the methanolic extract of a Spanish

variety of tigernut "chufa" (*Cyperus esculentus* tubers), using UHPLC/ESI-QTOF-MS. An approach for the use of in-source collision-induced dissociation fragmentation (IS-CID) with ESI followed by product ion scan (MS/MS) in a Q-TOF mass spectrometer system is described. This can be achieved predominately, by combining variations in the fragmentor voltage for in-source fragmentation (ISF), with subsequent Collision-Induced Dissociation (CID) to induce further fragmentation of selected ions (MS/MS) in order to improve structural characterization.

Data presented in this study provide useful information about the phenolic profile of the methanolic extract of tigernut "chufa". To the best of our knowledge, this approach has not been performed before. The presented study allows a complete analysis of phenolic distribution and composition.

1.1. *Cyperus esculentus*: description, phytochemical profile and biological activity.

Cyperus esculentus (tigernut), which is also known as chufa, is a sedge tuber plant that is grown widely in tropical and Mediterranean regions. However, there are several types such as *esculentus* from the Mediterranean region east to India, *hermannii*, *leptostachyus*, and *macrostachyus* from the United States, and *sativus* from Asia [4]. This tuber is a perennial grass plant that belongs to the family *Cyperaceae*. Tigernut is not a nut but a tuber that is about the size of peanuts and brown colour [5] (Fig. 1). Its health and nutritious values have remained the same throughout the years, but only recently its demand has increased tremendously due to its values.

***Figure 1**. Tigernuts. Picture taken from https://organicgemini.com/pages/what-is-a-tigernut*

Cyperus esculentus has been traditionally cultivated around Valencia (Spain). The almost unique use until now has been the preparation of "horchata de chufa", a refreshing

vegetable milk. Chufa milk is a nutritive and energetic drink that can be used for all range of ages. It contains high amount of starch, carbohydrates and a small percent of protein nuts. In addition, it is rich in mineral contents (sodium, calcium, potassium, magnesium, zinc and copper), vitamins C and E contents, and essential fatty acids (EFAs) like myristic acid, oleic acid and linoleic acid. In consequence, tigernut could be considered as an interesting solution for the protein-calorie malnutrition in the developed countries. Furthermore, it can be used as a flavouring agent [6, 7].

Many health benefits have been attributed to consumption of tigernuts. Several pharmacological studies on *Cyperus sp.* indicated a myriad of biological effects such as anti-inflammatory, hepato-protective, gastro-protective, antimalarial and anti-diabetic activities [7].

Several primary and secondary metabolites have been reported from *Cyperus sp*. Including quinones, sesquiterpenes, alkaloids, saponins, essential oils and phenolic compounds. Most of them with important biological activities:

- ***Fatty Acids (EFAs)***. Consumption of EFAs plays relevant roles in many biochemical pathways resulting in cardio-protective effect, due to their considerable antiatherogenic, antithrombotic, anti-inflammatory, antiarrhythmic and hypolipidemic effect, thereby, decreasing the risk of degenerative illness such as cardiovascular diseases, cancer, osteoporosis, diabetes and others. *Cyperus esculentus* possess high content in fatty acids such as polyunsaturated (PUFAs) (linoleic acid or hydroxylinoleic acid), monounsaturated (MUFAs) (ricinoleic acid or oleic acid) and some saturated fatty acids [5].
- ***Proteins***. Protein content of tigernut is of high biological value considering the composition of essential amino acids it contains [4]. It has an antiatherogenic as well as other cardio-prevented effects and this is mainly because its arginine which is a precursor of NO that makes veins expand [8].
- ***Phenolic compounds***. Phenolic compounds are a big group of secondary metabolites found in plants that are produced under normal and stress conditions. They play an important role of adaptation of plants to their environment. Their

structure is known for having one or more aromatic rings with one or more hydroxyl groups. Phenolic compounds can be found either as free molecules or linking to other structures like sugars, proteins, lipids or other phenolic compounds. We can distinguish between non-flavonoids and flavonoids (Fig. 2) [9].

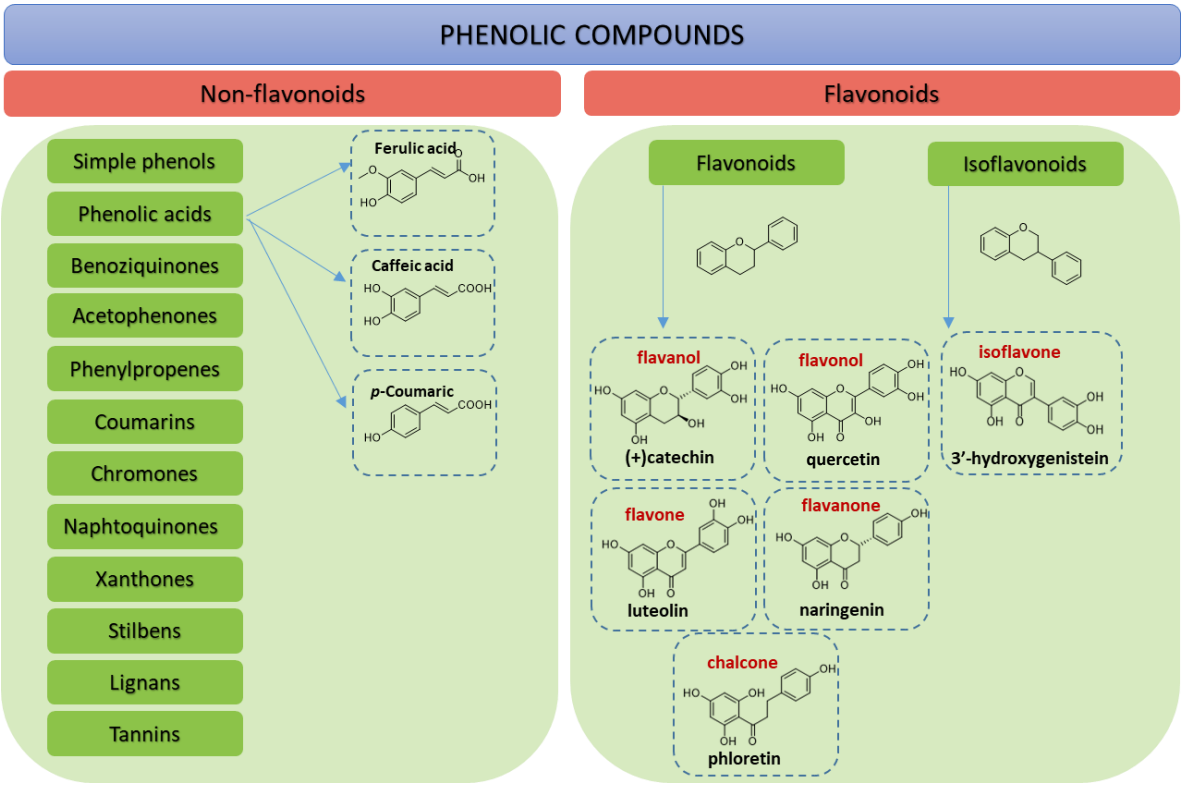


***Figure 2**. Classification of phenolic compounds.*

Phenolic compounds are some of the most numerous and widely distributed groups of bioactive compounds in the plant kingdom. They are known to have antioxidant activities which prevent formation of free radical, which are the major cause of various diseases such as cardiovascular, anti-inflammatory and cancer.

Tigernut was reported to have relatively high total phenolic content which might be responsible for its antioxidant activity but there is few data on the phenolic profile.

1.2. HPLC coupled to Mass Spectrometry. Electrospray (ESI).

In the past, the analytical methods were more focused on routine analysis. These methods were usually based on specific sample preparation protocols followed by traditional chromatographic separation. Primarily LC coupled with ultraviolet/diode array detection and fluorescence detection, was used, but detection by MS was rarely used. However, developments in the field of LC–MS technology have led to the availability of high throughput instrumentation meeting the current demands of scientists. LC–MS has become a tool of choice to deal with a number of analytical challenges related to chemical food and feed safety testing in both research and routine commercial laboratories. LC–MS-based workflows provide significantly higher selectivity and sensitivity, increased confidence in the identification of analytes, and wider analyte/matrix scope as compared with traditional methods using conventional detectors [10].

Tandem HPLC linked to a quadrupole-time-of-flight (Q-TOF) is an analyser in which ions previously formed are accelerated to a high speed by an electric field into an analyser with a long and straight drift-tube. Hence, Q-TOF separates ions based on their speed. Ions with higher *m/z* will take more time to cross the analyser. Q-TOF presents the advantage of analysing masses ranging from 50 to 20000 *m/z* with high resolution and precision. It is a powerful tool that is used in metabolomics, bioavailability and pharmacokinetic studies and an efficient instrument for the characterization of bioactive compounds present in plants. However, HPLC linked to mass spectrometry needs the use of an ionization source, where the eluent is evaporated and ionized, converting the sample molecules into a gas-phase [9].

Electrospray ionization (ESI) has become one of the most important ionization techniques for the on-line coupling of liquid phase separation method with MS. It can be split into three steps: nebulization of a sample solution into electrically charged droplets, liberation of ions from droplets, and transportation of ions from the atmospheric pressure ionization source region into the vacuum and mass analyser of the mass spectrometer [11]. ESI is a soft ionization source used with peptides, proteins, carbohydrates, small oligonucleotides and lipids. ESI produces a gaseous ionized

molecules directly from a liquid solution. It operates by creating a spray of highly charged droplets in the presence of an electric field (Fig. 3).

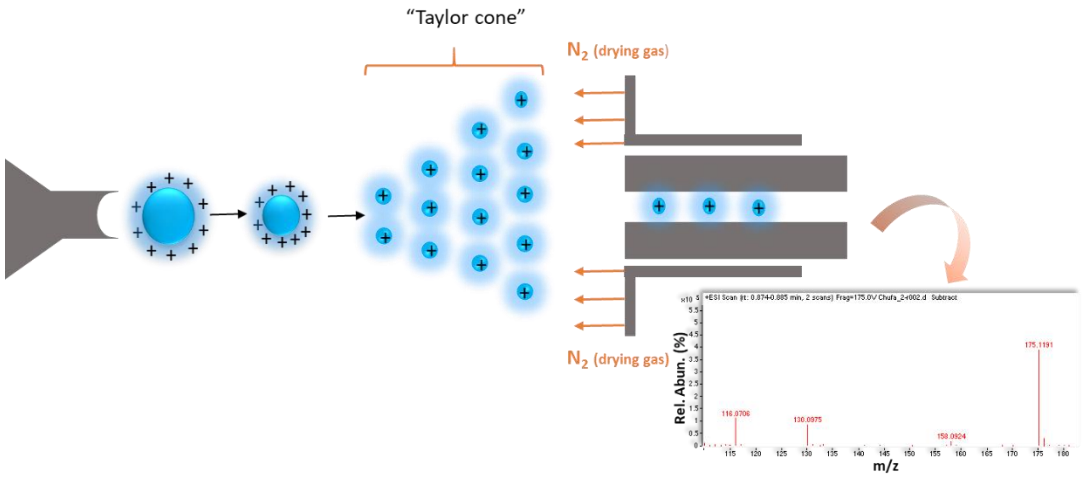


Figure 3. *Ion formation from electrospray ionization source. The ESI source uses a stream of air or nitrogen, heat, a vacuum, or a solvent sheath to facilitate desolvation of the droplets [11].*

The sample solution is sprayed from a region of strong electric field at the tip of metal nozzle maintained at a potential of anywhere from 700 to 5000 V. The needle to which the potential is applied serves to disperse the solution into a fine spray of charged droplets. Heat, dry gas or both of them are applied to the droplets at atmospheric pressure with the aim of causing the solvent to evaporate from each droplet. Due to the decreasing of the size of the droplets, the charge density on its surfaces increases. The repulsions between charges on this surface becomes so huge that exceeds the forces of surface tension, and ions are thrown out from the droplet through what is known as a "Taylor cone". ESI produces the formation of singly charged molecules as well as multiply charge species of large molecules [12].

1.3. IS-CID vs MS/MS-CID.

The MS/MS technology can be used to support molecular chromatography separations with isolation of ions formed in the initial mass analyser of the tandem MS (Fig. 4). In the first mass analyser "precursor" ions are selected. These precursors ions are fragmented by collision with an inert gas in a low-pressure collision chamber, and the chemically significant fragment ions are isolated in time (TOF) or

scanned (QQQ) with a second mass analyser. In the third chamber, there is a higher pressure and the ion selected will be fragmented [13].

However, there is a phenomenon observed with soft liquid or in-air ionization systems: in-source collision-induced dissociation (IS-CID), which can give useful information for the elucidation of molecules with a single-stage mass analyser. The generation of these ions occurs in the intermediate pressure transmission region between the atmospheric pressure source and the entrance cone of the high vacuum mass analyser, so the ions are formed without parent and there is no ion precursor as in MS/MS-CID process.[13]

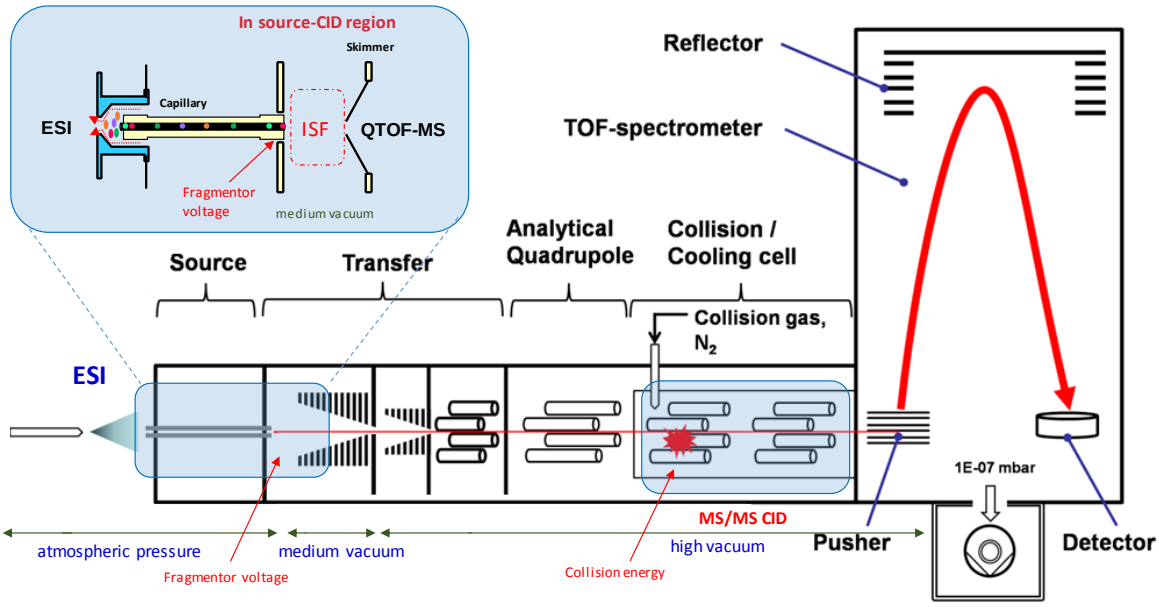


Figure 4*. In-source fragmentation in the IS-CID region and CID MS/MS in the collision cell.*

IS-CID advantages and disadvantages**.** The production of ionic products and fragments in the source provides rich information for the elucidation of compounds and makes easier the distinction between isomers and chemically similar components. Also, it is useful to generate libraries. And it can be used to quantify and fragmentation happens rapidly and high-throughput operations are feasible for simple systems [14]**.** There are issues with the reproducibility of the exact type and quantity of ions formed by the in-source fragmentation process. Moreover, the fragmentation process is more complicated and less predictable than MS/MS results.

IS-CID procedure is dependent on good chromatography resolution with little or no background contamination, otherwise it will be difficult to specify the origin of the observed IS-CID ions. The ion pattern relies on instrument and laboratory due to the instrument design and experimental control parameters (lake of a universal library for IS-CID spectra) [15]. In last instance, this is a pressure and collision energy dependent process [13].

2. OBJECTIVES

General objective: The main aim of this Master's Final Dissertation (TFM) is to investigate the chemical composition of the methanolic extracts of tigernut (*Cyperus esculentus* tubers) with special attention to the characterization of its phenolic profile by using an advanced analytical technique; UHPLC/ESI-QTOF-MS.

Specific objectives: this general objective explained above can be split into different specific objectives:

- Targeted analysis by direct comparison with a phytochemical cembio library of 48 phytochemical authentic standards.
- Quantification of the compounds that were found by targeted analysis.
- Characterization of the unknown compounds with a comparison among: retention time, isotopic profile, *m/z*, molecular formula, in-source fragmentation (ISF) and, also MS/MS experiments when it is possible.
- To give a perspective about how the tigernut extract could be useful due its bioactive compounds composition.

3. EXPERIMENTAL SECTION

3.1. Plant material and preparation of extracts.

Cyperus esculentus was purchased to the company called Món Orxata S.L. Mundo Horchata) located in Alboraia (Valencia, Spain) that is responsible for "Orxata-Coffee Shop" where "chufa" and its natural milk are promoted with appellation of origin. Tubers from *Cyperus esculentus* were milled to thine powder using a manual engine grinder. The milled sample was soaked for 1 hour in methanol. Three extractions were realised flowing the next method: 30 mg sample

were extracted using an Eppendorf tube with 300 µL of methanol in an ultrasonic water bath for 30 min at room temperature. The solution was then centrifuged at 10.000 *g* for 5 min at 4 ºC. Supernatants of the three replicates were then collected for direct analysis.

3.2 Sample and mix standards preparation. Collection of a set of 48 phenolic compounds commercially available standards (cembio phytochemical library) were all purchased from Sigma Aldrich Chemicals. All of them were of analytical grade and were used without further purifications. Methanol (mass spectrometry grade) which was purchased from Fluka and ultrapure water was obtained from Milli-Q apparatus. The standard solutions (10 µg/mL) were prepared in methanol.

3.3. Instrumentation. Samples were analysed on a 1290 Infinity series UHPLC system coupled through an electrospray ionisation source (ESI) with jet stream technology to a 6545 iFunnel qTOF-MS system (Agilent Technologies, Waldbronn, Germany). The HPLC column used was a reversed-phase column (Zorbax Eclipse XDB-C18 4.6 × 50 mm, 1.8 µm, Agilent Technology).

3.4. Chromatogram procedure. The separation of the compounds from tigernut methanolic extract was carried out with an injection volume of 2 µL to the reversed-phase column and kept at 40°C. The system was operated at 0.5 mL/min flow rate consisting of solvent A, water with formic acid (FA) at 0.1%, and solvent B, methanol, as mobile phases. Gradient started at 2% B (0-5 min), later a linear gradient from 2 to 50% B (5-10 min), then up to 95% B (10-18 min), kept at 95% B for 2 min (18-20 min), and returned to starting conditions in 1 min to finally keep the re-equilibration at 2% B until 25 min.

3.5. ESI-QTOF-MS detection. Detector was operated in full scan mode (*m/z* range from 50 to 2000) in positive and negative ESI mode at a scan rate of 1 scan/s. Accurate mass measurement was assured through an automated calibrator delivery system that continuously introduced a reference solution containing masses of *m/z* 121.0509 (purine) and *m/z* 922.0098 (HP-921) in positive ESI mode; whereas *m/z* 112.9856 (TFA) and *m/z* 922.009798 (HP-921) were introduced as a reference solution in negative ESI mode. The capillary voltage was ± 4000 V for positive and negative ionisation mode. The

source temperature was 225 °C. The nebulizer and gas flow rates were 35 psig and 11 L/min respectively. For the in-source fragmentation (ISF), fragmentor voltage was optimised with values of 100 and 175 V. Moreover, the collision energy used for the product ion (MS/MS) analysis was optimised as well by changing collision energy values (15, 20, 25, 30 eV) and was stabilised to 30 eV. The radiofrequency voltage in the octupole (OCT RF Vpp) of 750 V. For the study, MassHunter Workstation Software Data Acquisition version B.08.00 (Agilent Technologies) was used for control and acquisition of all data.

3.6. Quantification procedure. For the quantitative analysis, a cembio phytochemical library of 48 standards was used to calculate individual concentration present in the methanolic extracts. Stock standard solutions (0.01 g/L) were prepared in methanol and calibration curves were prepared for quantitative analysis of phenolic compounds in the target samples. Five concentrations (5, 2.5, 1.5, 0.25 and 0.125 mg/L) were used for the calibration, plotting peak area vs. concentration, with MassHunter Quantitative Analysis Software version B.08.00 (Agilent Technologies). The quantification of phenolic compounds was calculated by the extrapolation of the peak area values obtained for the components of every sample analyzed from the calibration curve of the standards. All the data processing was performed in MassHunter Quantitative Analysis (Agilent Technologies) Software version B.08.00. Previously, it was necessary to create a library of spectra with all the standards to be able to make a method in order to quantify the compounds which were in the mixes and in the sample. To do this library, it was used the program Library Editor.

3.7. Data processing. UHPLC-MS data processing was performed by MassHunter Qualitative Analysis (Agilent Technologies) Software version B.08.00 using Molecular Feature Extraction (MFE). The MFE algorithm creates a list of possible components that represent the full TOF mass spectral data features, which are the sum of co-eluting ions that are related by charge-state envelope, isotopic distribution and/or the presence of different adducts and dimmers. Several parameters of the algorithm were set for data extraction, applying 5000 counts as limits for the background noise. Moreover, the algorithm was

applied to find co-eluting adducts for the same possible compound, selecting +H, +Na, +K, and neutral water loss as possible adducts for positive ionization and +FA in negative ionization. Additionally, the "Generate Formula" option in the MassHunter Qualitative Analysis software was used to generate the empirical formula from accurate mass and isotopic pattern distribution to increase the confidence of compound annotation.

3.8 Annotation workflow. Annotation is defined as the process of noting each observed feature with a putative metabolite or molecular formula, but it also includes assigning every observed feature with the identity of adducts, neutral losses and in-source fragmentation. Annotation is very important in untargeted analysis as the main aim of any untargeted analysis is to identify and quantify compounds [16]. Untargeted data processing workflows involve several steps like peak deconvolution, peak detection, RT alignment, and finally feature annotation [17]. For peak identification, it is important to rely on the retention time and peak shape of the extracted ion chromatogram (EIC's) that correspond to the ions present in the average spectrum. This is because the signal of the compound of interest if often distributed over multiple entities in the mass spectrum (different isotopes, adducts, in-source fragments even dimers). (Appendix 1). Thus, a fragment is considered when it has the following characteristics: first, the same migration time than the protonated *pseudo*-molecular ion $[M-H]^-$; second increased intensity with increased in-source voltage and third, the same peak shape as the $[M-H]^-$. On the other hand, the accurate mass obtained was matched to compounds from web-based sources and compared with the MS/MS spectra available from the databases, METLIN (http://metlin.scripps.edu) and FooDB (http://foodb.ca). Ions produced by in-source fragmentation process were similar to those observed by low-energy MS/MS CID processes. The spectra were not identical but often complementary. In untargeted analysis is also important the confidence level and it is essential that the confidence of compounds assignments is transparent. There are many guidelines for reporting the minimum metadata relative to compound identification as a means to communicate the confidence of identifications. [17]. We have followed the one proposed by Schrimpe-Rutledge, et al. (Appendix 2)

4 RESULTS AND DISCUSSION

4.1. UHPLC-ESI-QTOF-MS analysis.

In this study UHPLC/ESI-Q-TOF-MS analysis has been performed effectively for characterizing various bioactive compounds present in *Cyperus esculentus*. The phenolic profile of the methanolic extract was analysed in ESI negative mode in the range of *m/z* 50-1500 Da and the resulting Base Peak Chromatogram (BPC) is shown in Fig.5. The chromatographic analysis was performed using optimized conditions that provide a satisfactory separation less than 25 min.

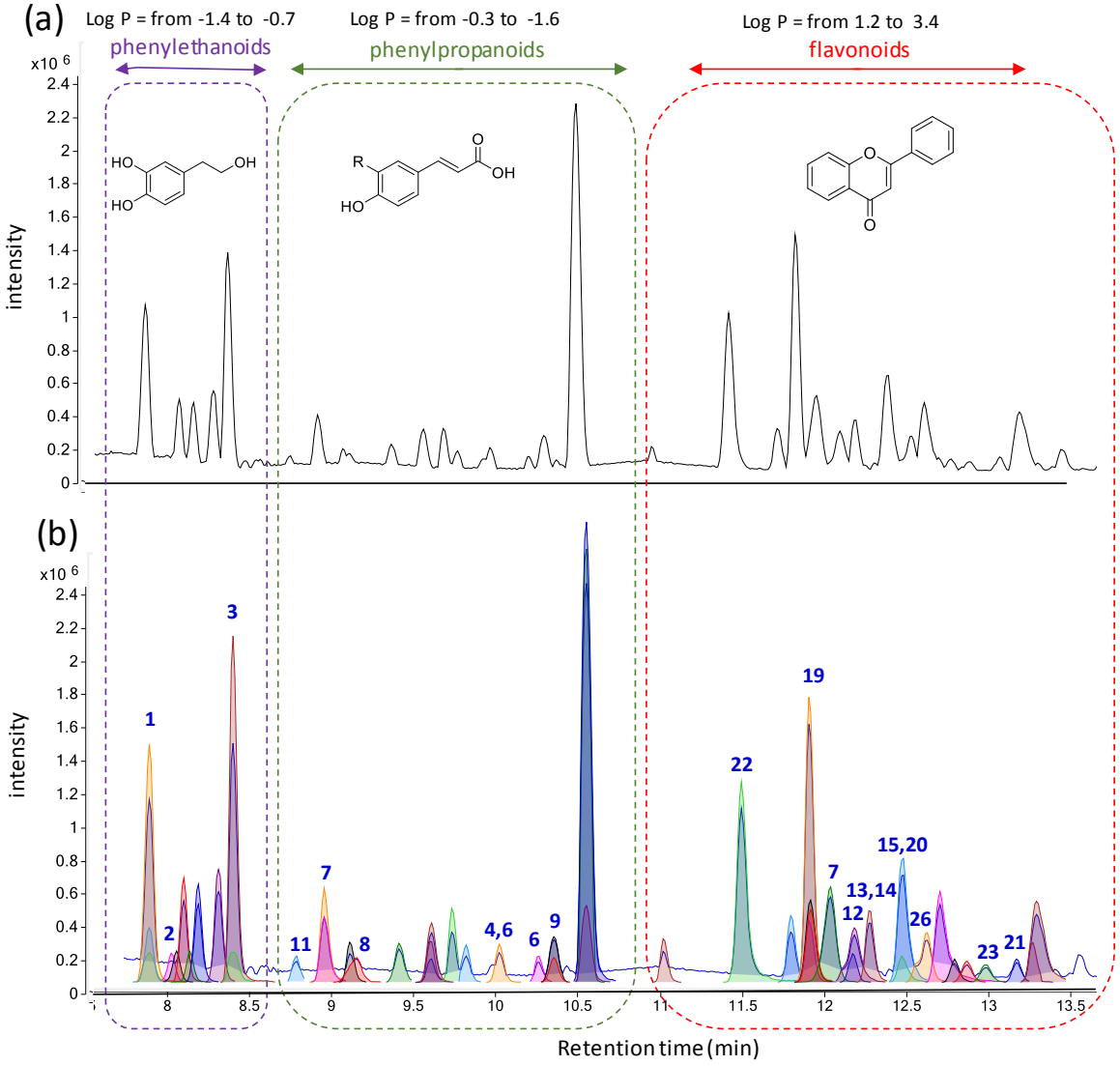


Figure 5. (a) *Representative BPC (negative ion mode) of the phenolic profile from the methanolic extract of chufa. Enlarged chromatogram from 6 to 15 min.* ***(b)*** *chromatographic deconvolution, Molecular feature Extraction (MFE).*

The use of ESI as ionization source operating in the negative mode has proved to be more efficient and sensitive than in positive for phenolic compounds and flavonoids characterization (Appendix 3).

The identification of compounds was carried out by comparing their retention times and accurate mass spectra provided by TOF-MS with those of authentic standards when available. The remaining compounds in the chromatogram, unknown compounds, for which no commercial standards available, were characterized (putative identification) by in-source fragmentation (ISF) prior to product ion (MS^2) scan. The ISF step (with variations in the fragmentor voltage from 100 to 175V) showed numerous fragments which in turn were used as precursor ions for the MS/MS step, Fig. 3. The experimental accurate mass measurement as well as the fragmentation pattern observe were matching against mass spectra databases, METLIN, FOODB and others like phenol explorer (http://phenol-explorer.eu/), and scientific bibliography.

Appendix Table 2 report all the compounds with their retention times, fragmentation ions and assigned identities. In total 26 phenolic compounds were found, the majority of them were detected and characterized from *Cyperus esculentus* and reported here for the first time.

Phenolic compounds were found to be clustered in characteristic retention time windows according to different polarity. Useful parameter (logP), which was predicted by ACD/labs was searched against free databases available as Chemspider (http://www.chemspider.com), and it was introduced to distinguish the compounds. In reversed phase liquid chromatography, the compounds with smaller logP values commonly display shorter retention times due the partition mechanism. Thus, this parameter played an important role in estimating the structures (Fig. 4).

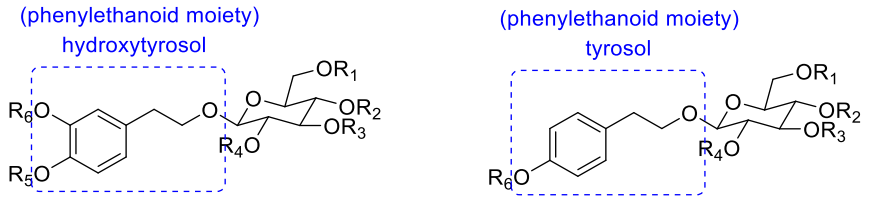
4.2 Quantitative analysis of individual phenolic compounds.

Compounds were identified by comparing their retention times and mass spectra with those of several commercial standards analyzed under the same conditions. 48 standards (CEMBIO list of phytochemicals) were analyzed. (Appendix Table 1) reports the 9 compounds found in the targeted analysis. A quantification of them was further carried out. In terms of

abundance, the compounds that were possible to be quantified and found in higher concentrations were chlorogenic acid, ferulic acid and luteolin.

4.3 Characterization of Phenylethanoid glycosides (Compounds 1-4).

PhGs are plant natural products that are widely distributed in plant kingdom and have been isolated from different families from roots, bark, leaves, stems, flowers, fruits, aerial parts as well as from callus tissue, suspension cell cultures and hairy roots cultures. Pharmaceutical research has shown that these molecules have a broad range of bioactivities (antioxidant, antibacterial and antiviral, anti-inflammatory). The core structure of PhGs is characterized by a hydroxyphenylethyl with an ethane tail (C6-C2) moiety linked with a β-glucopyranose through glycosidic bond. In most cases the basic structure of PhGs contains a cinnamic acid and hydroxyphenylethyl moiety with a β-glucopyranose through ester and glycosidic bond, respectively. Variation of PhGs occurs in both the sugar and hydroxyphenylethyl moieties. Sugars and/or aromatic acids such as cinnamic acid, caffeic acid, coumaric acid and ferulic acid can be attached to the glucose residue through both glycosidic and ester linkages (Fig. 6) [18].



R_1=H, caffeoyl, feruloyl, acetyl, acetyl, apiose, xylose, rhamnose, glucose
R_2=H, caffeoyl, feruloyl.
R_3=H, apiose, glucose, arabinose-rhamnose, rhamnose-apiose, glucose-rhamnose, apiose-rhamnose, xylose-rhamnose.
R_4=H, acetyl, apiose, glucose.
R_5,R_6=H, CH_3.

Figure 6. *Core structure of PhGs glycosides*.

Glycosylation is a common modification by which a glycan (or oligosaccharide) is covalently attached to an aglycone. Glycans are tree ensembles of monosaccharides linked via glycosidic bonds. Glycan structural analysis remains a challenging task, in part due to the vast number of topologies that they may assume even for a moderate-sized glycan. Based on the UHPLC/ESI-Q-TOF-MS/MS method, the important structural information on

the types of saccharide sequences present can also be obtained. Combined with the HPLC retention behaviour and fragmentation pathways based on literature data provide for the tentative identification. *Nomenclature*: The major diagnostic fragmentations for neutral glycans have shown that fragmentation of negative ions produces prominent C-type glycosidic cleavages and A-type cross-ring cleavages rather than the B- and Y-type glycosidic cleavages that are common to the fragmentation of positive ions (Fig. 7) [19]**.**

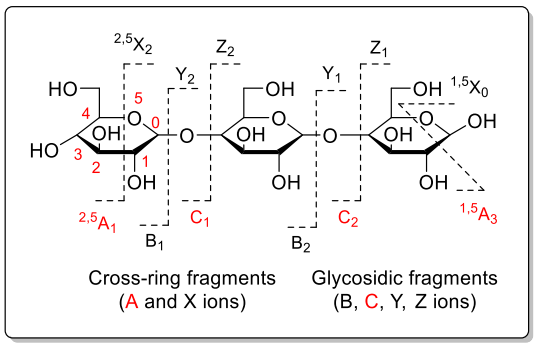


Figure 7. *Systematic nomenclature for labelling fragment ions following Domon and Costello [19]*

In this study, 4 PhGs were detected (Appendix Table 2) (MS/MS Spectra in Appendix 4). Compound **4** comes from hydroxytyrosol that was detected at *m/z* 153, with a fragment at *m/z* 123, which is due to the loss of the CH_2OH group. Compounds **1**, **2** and **3** come from tyrosol that was detected at *m/z* 131. The MS^2 analysis of compounds **1**-**4** showed the same fragments characteristic of the monosaccharide D-glucopyranose (Fig. 8a).

Peaks at *m/z* 341 (not observed) and 179 correspond to glycosidic cleavage products C_2 and C_1 formed by the loss of one or two hexose moieties respectively, whereas the peaks at *m/z* 323 and 161 are the corresponding B_2 and B_1 type ions (loss of water). Highly informative negative ions MS/MS spectra showing several diagnostic fragment ions were obtained where A-type cross-ring cleavages of monosaccharides dominate the CID-MS/MS spectra (Fig. 8b). Thus, mass spectrum of D-glucopyranose exhibits a characteristic peak due to deprotonated molecules $[M\text{-}H]^-$ at *m/z* 179 and other signals with variable intensities. It is known that monohexoses undergo retro-aldolization reactions (RDA), yielding stable α-dicarbonyl compounds (Fig.8c). Other peak signals in the low *m/z* range (i.e., *m/z* 71, 89, and 101) are likely attributable to monohexose's deprotonated by-products (Fig.8d). Other multiple cross-ring cleavages generated A type ions at *m/z* 221

($^{0,4}A_2$), 119 ($^{0,2}A_1$), 101 ($^{2,5}A_1$), and 89 ($^{0,3}A_1$), where superscripts indicate the position of the cross-ring cleavage.

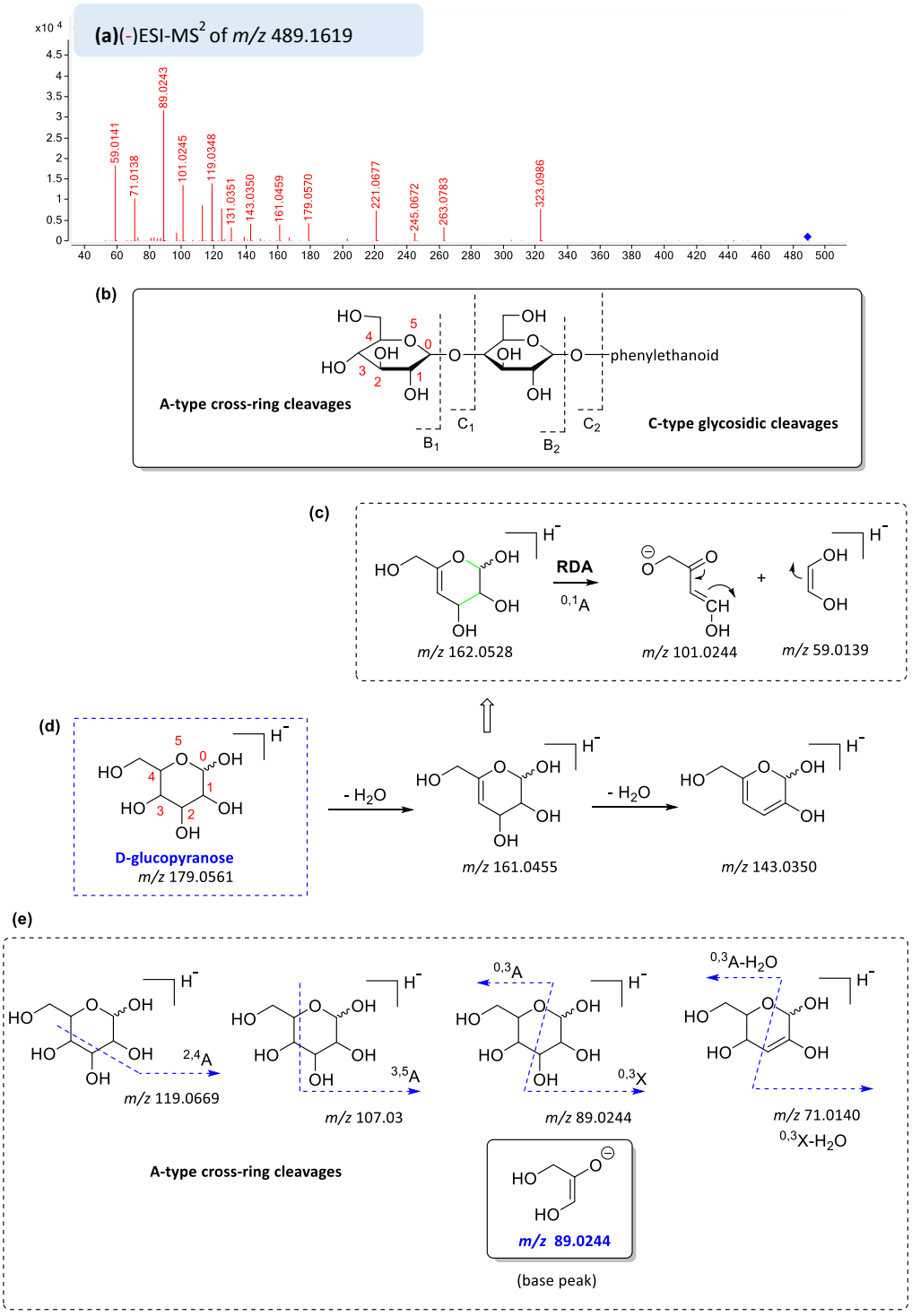


***Figure 8**. **(a)** Spectrum of m/z 489. **(b)** Nomenclature for labelling fragment ions where superscripts indicate the position of the cross-ring cleavage. **(c)** Retro-Diel-Alders (RDA). **(d)** Suggested cross-ring cleavages of deprotonated D-glucopyranose identified by MS/MS*

4.4 Characterization of free phenolic acids and derivatives (Compounds 5-10).

Phenylpropanoid compounds constitute a wide range of structural cyclic substances possessing numerous physiological functions. This family of organic compounds has an aromatic ring and a propene tail (C6-C3). Hydroxylation, methylation, and dehydrogenation of cinnamic acids results in formation of its derivatives (*p*-coumaric, caffeic, ferulic, and sinapic acid). It is often to find these compounds conjugated with organic acids, cell wall carbohydrates, or sugars [20].

In this study, 6 phenolic acids and derivatives were detected (Appendix Table 2). They belong to hydroxycinnamic acids (Fig 9) that have been reported to present antioxidant activity. Moreover, it has been described that these bioactive compounds possess a hypoglycaemic effect [21].The annotation was carrying out with a workflow that consisted in studying the mass spectra of each one and identifying their *m/z* and generating their molecular formula. Afterwards, a searching in spectra databases and bibliography was realised to find a matching which could suit with the parameters already defined.

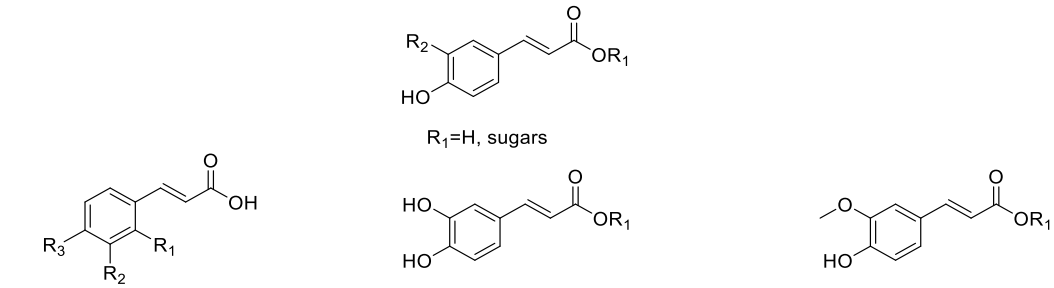


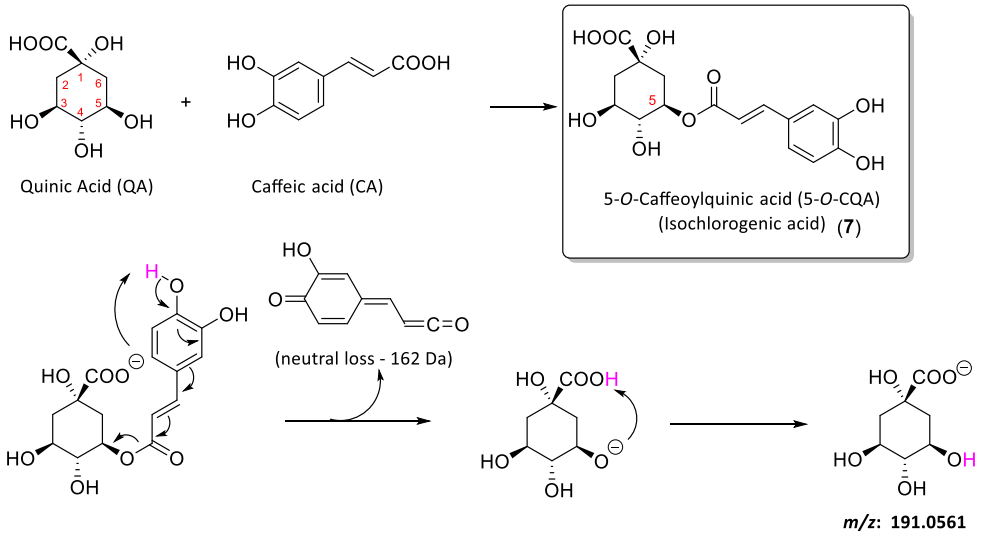
Figure 9. *General structure of phenolics acids*

Compound **5**, with its $[M-H]^-$ at *m/z* 355 in its ISF spectrum, fragmented at *m/z* 193 with a neutral loss of hexose. By comparing its molecular formula and fragmentation patterns with those reported in data bases, compound **5** was tentative identified as 1-*O*-feruloyl-β-D-glucose.

For compound **6**, with its $[M-H]^-$ at *m/z* 473, no fragmentation was detected. By comparing its molecular formula, *m/z* in data bases the best match due to *m/z*, molecular

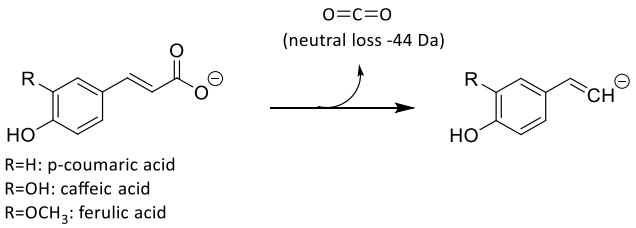
formula and log P and RT was 6-*O*-Feruloyl-β-D-glucose-2,3,4-trihydroxy-3-methyl butylglycoside.

Compound **7**, with its $[M-H]^-$ at *m/z* 353 in its ISF spectrum , fragmented at *m/z* 191 with a neutral loss of hexose (Scheme 1)(Appendix 5) and was definitely identified as 5-*O*-Caffeoylquinc acid by comparison with its standard.



Scheme 1. *Suggested fragmentation mechanism for 5-O-Caffeoylquinic acid (**7**).*

Compounds **8**, **9**, **10** suffered a neutral loss of their respective carboxylic acids (Scheme 2). These compounds were identified as caffeic (**8**), *p*-coumaric (**9**) and ferulic (**10**) acids by comparison with their standards.



Scheme 2*. Proposed fragmentation mechanism for phenolic acids.*

4.5 Characterization of Flavonoids (Compounds 11-26).

Flavonoids is the common name of a class of over 6500 molecules, which its skeleton structure is a 2-phenylbenzopyranone cyclised with oxygen (C6-C3-C6).

Flavonoids are polyphenols which have been recognised as one of the largest widespread groups of secondary metabolites in plants. The most well-known types are flavones, isoflavones, flavonols, anthocyacyanidins, flavanones, flavanols and chalcones (Fig. 10). Like almost every plant compound, flavonoids can exist either as free aglycones or as glycosidic conjugates [22]. It has been reported that flavonoids possess antioxidant activity. Anti-inflammatory properties has been found in this wide family of compounds as well [22, 23]. Inside this family, it should be highlighted the estrogenic effect that isoflavones has due to their similar structure with the human female hormone [24].

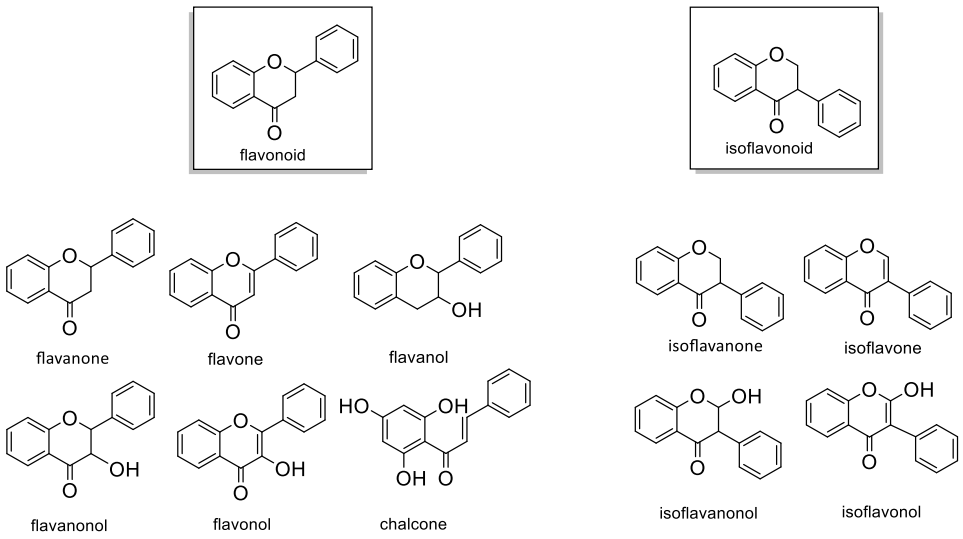


Figure 10. *Core structures of flavonoids and isoflavonoids*.

These last years, MS, especially ESI in negative mode and APCI in combination with tandem mass spectral methods has evolved as a powerful tool for elucidation and researches of these molecules [25]. *Nomenclature*: The major diagnostic fragmentations for flavonoid identification are those involving the cleavage of two C-C bonds of the C-ring giving two structurally informative fragment ions. These ions provide information on the number and type of substituents in A- and B-rings (Fig. 11) [26].

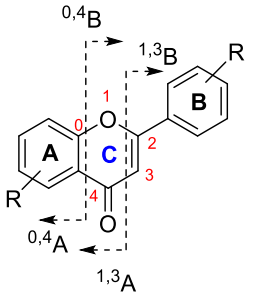


Figure 11. *Systematic flavonoid aglycone nomenclature for labelling fragment ions observed in MS/MS following Ma et al. [26]*

RetroDiels-Alder (RDA) plays an important role in the identification of flavonoids and their derivatives. The analysis of the fragmentation allows the classification of the groups attached to the structural skeleton of flavonoids and the structural variation in this class of compounds. Additionally, it is very typical to see CO elimination in flavonoids to afford a ring contraction product [27].

In this study 16 flavonoids were detected.

- ***Flavanols*** (Appendix Table 2). Compound **11**, with its $[M-H]^-$ at *m/z* 289 had an adduct $[M -H+HCOONa]^-$ at *m/z* 357, and no fragmentation was detected. This compound was identified as (+)-catechin by comparison of its chromatographic retention time and ISF pattern with its standard.
- ***Flavonols*** (Appendix Table 2). Compound **12** had its $[M-H]^-$ at *m/z* 299 and there was no ISF detected. However, this compound was identified as kaempferide by a MS^2; in the MS^2 spectrum *m/z* 299 fragmented at *m/z* 284 by a radical fragmentation. Compound **13**, with its $[M-H]^-$ at *m/z* 301, was identified as quercetin by comparison of its chromatographic retention time and ISF pattern with its standard.
- ***Flavones*** (Appendix Table 2). Compound **14**, with its $[M-H]^-$ at *m/z* 269 was identified as apigenin by its MS/MS spectrum, where *m/z* 269 fragmented at m/z 225 by a neutral loss of CO_2 and at *m/z* 245 by a neutral loss of CO (Fig. 12). Afterwards, a bibliography search was conducting and it was found in literature that tigernut has this secondary metabolite [21].

Compounds **15** and **17**, with their $[M-H]^-$ at *m/z* 299 were identified as chrysoeriol and 6-methylscutellaerin. These compounds are isomers as they shared same *m/z*, hence same

molecular formula, but they had different retention times: **15** at 12.5 min and **17** at 13.8. Compound (**15**), was definitely identified by ISF and MS^2 as chrysoeriol. In both spectra, ISF and MS/MS, *m/z* 299 fragmented at *m/z* 284 by radical fragmentation with a neutral loss of CH_3 (Appendix 6). Compound **17** at *m/z* 299, identified as 6-methylscutellarein, fragmented at *m/z* 285 by a neutral loss of CH_3. Compounds **16** and **18**, with their $[M-H]^-$ at *m/z* 285 and 253 respectively were identified as luteolin (**16**) and chrysin (**18**) by comparison of their chromatographic retention time and ISF patterns with their standards.

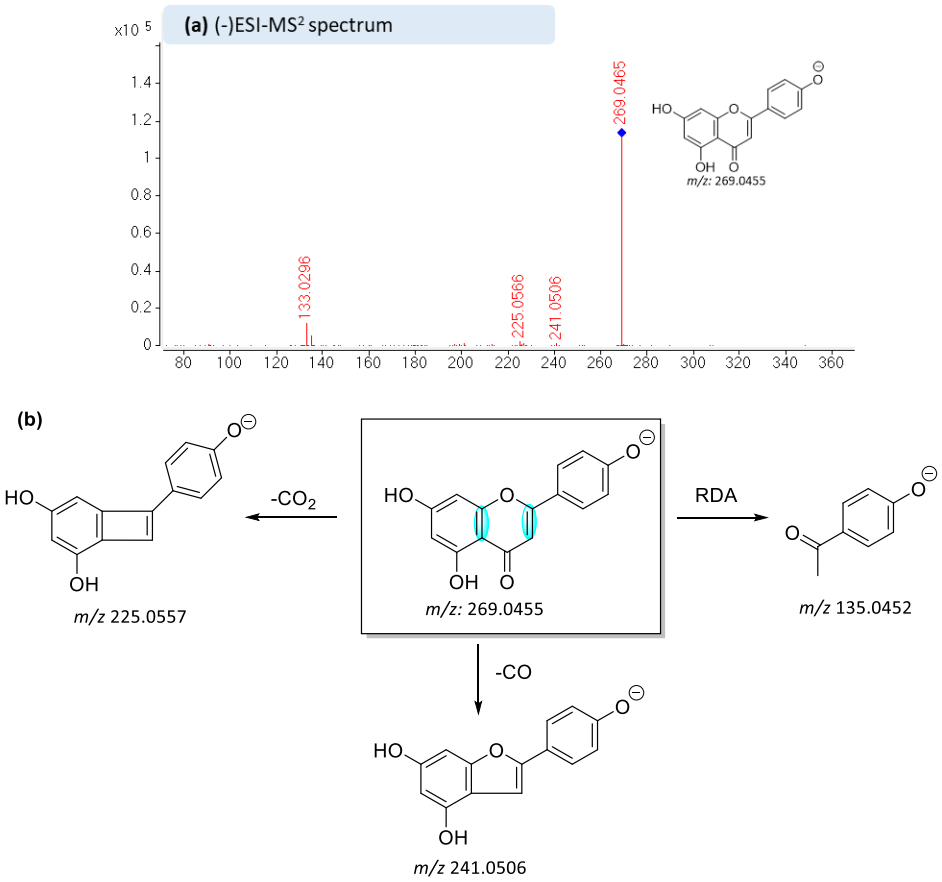


Figure 12. (a) *MS/MS spectrum of apigenin (**14**).* ***(b)*** *Proposed fragmentation mechanism of apigenin (**14**).*

- ***Flavanones*** (Appendix Table 2). Compound **19**, with its $[M-H]^-$ at *m/z* 287 and its $[2M-H]^-$ at *m/z* 575, was identified as eriodictyol by its ISF and MS^2 fragments at m/z 151 (ISF and CID-MS/MS) and at *m/z* 135 (only detected in CID-MS/MS) that are result of two pathways of RDA. (Fig. 13). Compound **20**, with its $[M-H]^-$ at *m/z* 287 was identified as naringenin by comparison of their chromatographic retention time and ISF patterns with their standards.

- ***Isoflavones*** (Appendix Table 2). For compound **22**, with its $[M-H]^-$ at *m/z* 285 any ISF was detected. However, it was extracted with MS/MS fragments at *m/z* 151 and 133 that are the result of a RDA (Fig. 14), being possible to identified this molecule as 3'-hydroxygenistein. Compound **23**, with its $[M-H]^-$ at 313 was identified as 2'-hydroxypseudobaptigenin by its mass measurement accuracy (*m/z*), and molecular formula. Compound **24**, with its $[M-H]^-$ at 299, was identified as barpisoflavone by its mass measurement accuracy (*m/z*), and molecular formula. Compound **25**, with its $[M-H]^-$ at 313, was also detected at m/z 298 with ISF, which was the result of a neutral loss of CH_3 and was identified as either 4',7-dihydroyx-2',5-dimethoxyisoflavone or as 4',5-dihydroxy-3',7-dimethoxyisoflavone due to the impossibility of detecting more fragments.
- ***Chalcones*** (Appendix Table 2). Compound **26**, with its $[M-H]^-$ at *m/z* 273, was identified as phloretin by comparison of its chromatographic retention time and ISF pattern with its standard.

4.6. Other compounds.

Finally, apart from the phenolic compounds described above, amino acids *L*-Arg and *L*-Trp were detected (Appendix Table 3). Amino acids showed a larger number of fragments with ESI in positive mode. *L*-arginine was reported to present a vasodilator effect which is related to cardio-protective properties [5, 8]. *L*-Tryptophan plays an important role because it is the unique precursor of peripherally and centrally produced of serotonin that is a neurotransmitter linked to states of mood [5, 28]. Moreover, a tryptophan conjugate, *N*-malonyl-*L*-tryptophan, was detected being one of the biggest peaks in the chromatogram profile, which means that it is the most abundant. Auxins are a class of plant hormones (or plant-growth regulators) that play a cardinal role in coordination of many growth processes in plant life cycles and are essential for plant body development. Indole-3-acetic acid (IAA) is the most abundant and potent native auxin and is presented naturally in conjugated forms. Relatively few IAA conjugated have been identified, these include IAA amino acid and tryptophan acid conjugates. We have identified in our samples *N(α)*-malonyl-tryptophan, an auxin precursor, which has not been reported in C*yperus* before. The compound was identified by the interpretation of its MS spectra in positive and negative

ionization mode and taking in account the fragmentation data provided in the literature and databases. (Appendix 7) [29]. In addition, it was possible to detect and identify fatty acids in ESI negative mode (Appendix Table 4).

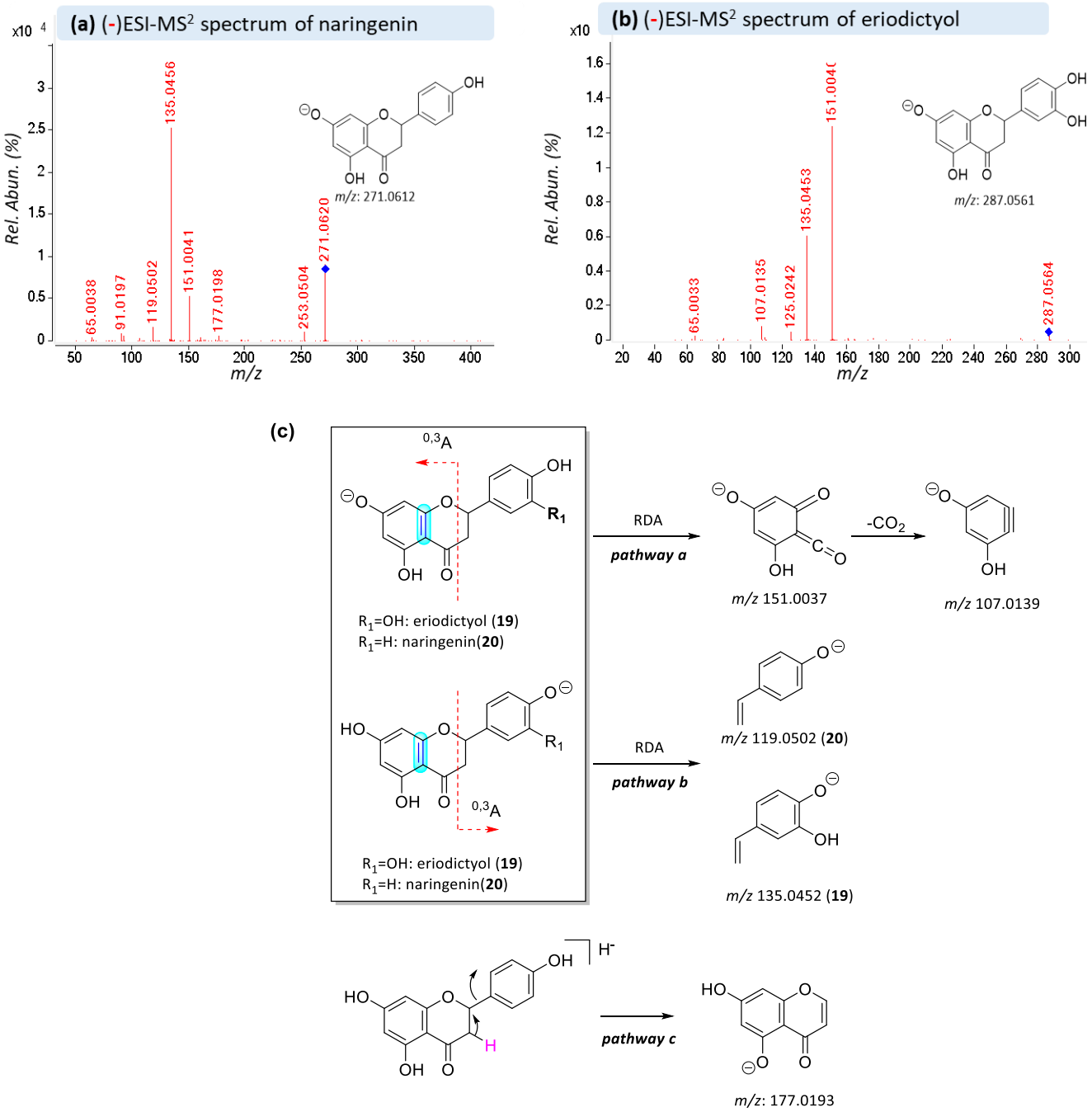


Figure 13. (a) *MS/MS spectrum of* ***20****.* ***(b)*** *MS*^{2} *spectrum of* ***19*** ***(c)*** *Proposed fragmentation pathways.*

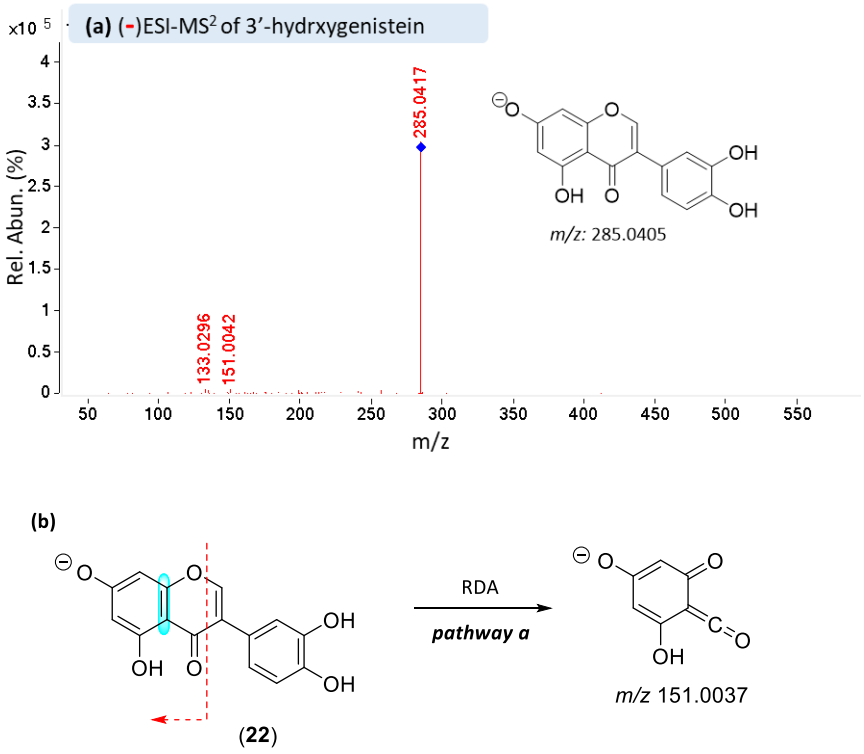


***Figure 14**. **(a)** MS/MS spectrum of 3'-hydroxygenistein (**22**). **(b)** Proposed RDA mechanism of 3'-hydroxygenistein.*

5. CONCLUSIONS

In this study, a powerful analytical method has been used to characterize 37 compounds (3 amino acids and derivatives, 6 phenolic acids and derivatives, 16 flavonoids, and 8 fatty acids) (Fig. 15a). Being 28 of them found by untargeted analysis. One, *N*-malonyl-*L*-tryptophan, identified for the first time in tigernut. Regarding phenolic compounds, they followed this trend: flavonoids compounds (62%)> phenolic acids and derivatives (23%)> phenylethanoid glycosides (15%) (Fig.15b). Also, the quantification of the nine compounds was carried out with success founding that the compounds in higher concentration were: chlorogenic acid (10)> luteolin (16)> ferulic acid (10).

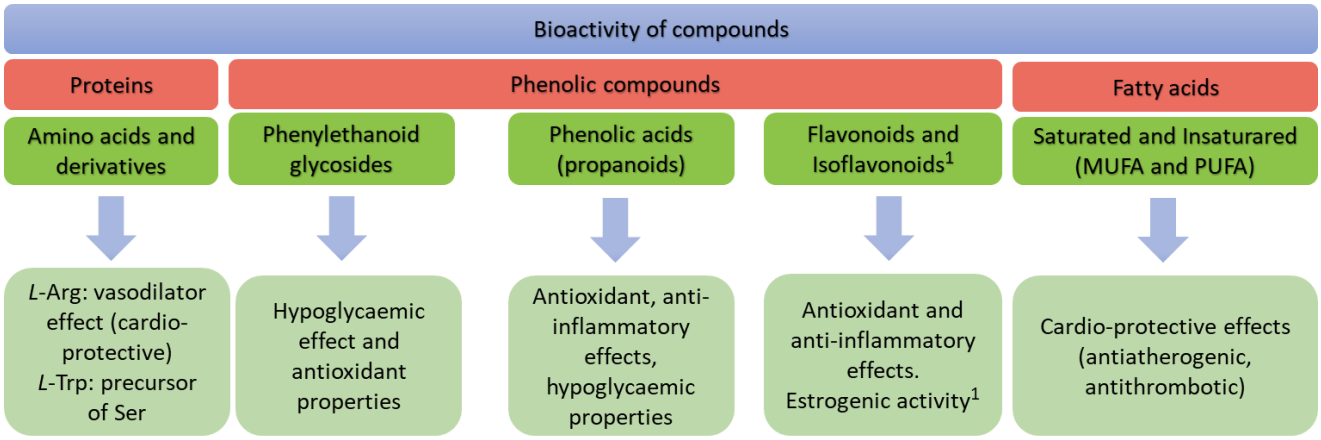


Figure 15. Bioactivity of the different compounds present in *Cyperus esculentus*.

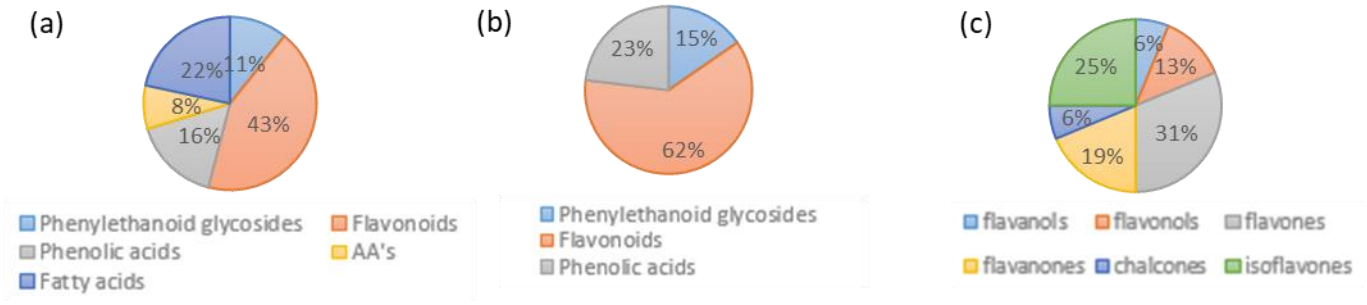


Figure 16. Percent of: ***(a)*** Total classes of compounds. ***(b)*** Phenolic compounds. ***(c)*** Flavonoid family.

All our findings contribute to a further understanding the bioactive composition of the chemical complexity of the mixture of methanol tigernut extracts.

Table 1.

No.	*Rt (min)*	*Common name (Annotation Confidence)[a]*	*Monoisot. Mass*	*Molecular Formula*	*ESI negative mode[b]*		
					Selected ions	*ISF*	*CID-MS/MS*
1	7.8	PhGs	508.1816	$C_{21}H_{32}O_{14}$	$[M-H+HCOOH]^-$ = 507.1726	323.0993	323.0998, 263.0789, 299.1152, 221.0683, 179.0573, 161.0469, **89.0249**
2	8.4	PhGs	462.1737	$C_{20}H_{13}O_{12}$	$[M-H]^-$ = **461.1662**	n.d.	299.1138, 263.0782, 221.0674, 179.0573, 161.0469, 89.0244
3	10.1	PhGs	490.1707	$C_{20}H_{30}O_{13}$	$[M-H]^-$ = 489.1619	323.0993	323.0986, 263.0783, 245.0672, 221.0677, 161.0459, **89.0243**
4	7.9	PhGs	478.1691	$C_{21}H_{32}O_{14}$	$[M-H]^-$ = 477.1611	323.0993	323.0977, 221.0680, 179.0573, 161.0461, **153.0565**, 89.0248
5	8.8	1-*O*-feruloyl-β-D-glucose (level 3)	356.1110	$C_{16}H_{20}O_9$	$[M-H]^-$ = 355.1045 $[M-H+HCOOH]^-$ = **401.1086** $[M-H+HCOONa]^-$ = 423.0905	$[M-H-hex]^-$=193.0518	
6	10.2	6-*O*-Feruloyl-β-D-glucose 2,3,4-trihydroxy-3-methylbutylglycoside (level 4)	474.1740	$C_{21}H_{30}O_{12}$	$[M-H]^-$ = 473.1656 $[M-H+HCOOH]^-$ = **519.1727** $[M-H+HCOONa]^-$ = 587.1611	n.d	
7	8.9 /9.7	5-*O*-Chlorogenic acid (level 1)	354.0950	$C_{16}H_{18}O_9$	$[M-H]^-$ = **353.0878** $[M-H+HCOONa]^-$ = 421.0759 $[2M-H]^-$ = 707.1839	$[M-H-caffeoyl\ acid]^-$= 191.0579	
8	9.3	Caffeic acid (level 1)	180.0420	$C_9H_8O_4$	$[M-H]^-$ = **179.035** No adducts detected	$[M-H-CO2]^-$ = 135.0457	
9	10.3	*p*-Coumaric acid (level 2)	164.0470	$C_9H_8O_3$	$[M-H]^-$ =163.0401 No adducts detected	$[M-H-CO2]^-$ **= 119.0403**	
10	10.6	Ferulic acid (level 1)	194.0570	$C_{10}H_{10}O_4$	$[M-H]^-$ =**193.0506** $[M-H+HCOONa]^-$= 261.0399	$[M-H-CO2]^-$= 149.0613	
11	8.7	(+)-Catechin (level 1)	290.07904	$C_{15}H_{14}O_6$	$[M-H]^-$ = **289.0718** [M-H+HCOONa]- = 357.0607	n.d.	-
12	12.2	Kaempferide (level 2)	300.0634	$C_{16}H_{12}O_6$	$[M-H]^-$ = **299.0567** $[M-H+HCOONa]^-$ = 367.0438	n.d.	284.0335
13	12.4	Quercetin (level 1)	302.0427	$C_{15}H_{10}O_7$	$[M-H]^-$ = **301.0359** No adducts detected	n.d.	-
14	12.2	Apigenin (level 2)	270.0539	$C_{15}H_{10}O_5$	$[M-H]^-$ = **269.0462** $[M-H+HCOONa]^-$ = 337.0336	n.d.	241.0506, 225.0566, 135.0452, 133.0296
15	12.5	Chrysoeriol (level 2)	300.0634	$C_{16}H_{12}O_6$	$[M-H]^-$ = **299.0567** $[M-H+HCOONa]^-$ = 367.0438	284.0331	284.0331
16	12.6	Luteolin (level 1)	286.0482	$C_{15}H_{10}O_6$	$[M-H]^-$ = **285.0412** $[M-H+HCOONa]^-$ = 353.0853	n.d.	n.d.
17	13.8	6-Methylscutellarein (level 2)	300.0634	$C_{16}H_{12}O_6$	$[M-H]^-$ = **299.0567** $[M-H+HCOONa]^-$ = 367.0438	285.0770	-
18	15	Chrysin (level 1)	254.0579	$C_{15}H_{10}O_4$	$[M-H]^-$ = **253.051** $[M-H+HCOONa]^-$ = 321.0394	n.d.	-
19	11.7	Eriodictyol (level 2)	288.0634	$C_{15}H_{12}O_6$	$[M-H]^-$ = **287.0567** $[M-H+HCOONa]^-$ = 355.0440 $[2M-H]^-$ = 575.1202 $[M-H+Na]^-$ = 309.0383	151.0039	151.0453, 135.0453, 125.0242, 107.0135

Table 1 cont.

					ESI negative mode[b]		
No.	***Rt (min)***	***Common name (Annotation Confidence)***[a]	***Monois. Mass***	***Molecular Formula***	***Selected ions***	***ISF***	***CID-MS/MS***
20	12.5	Naringenin (level 1)	272.0684	$C_{15}H_{12}O_5$	$[M-H]^-$ = **271.0618** $[M-H+HCOONa]^-$ = 339.0494	n.d.	177.0198, 151.0041,119.0502, 135.0456, 107.0134
21	13.2	Homoeriodictyol (level 3)	302.0793	$C_{16}H_{14}O_6$ $C_{15}H_{12}O_4$	$[M-H]^-$ = **301.0723** $[M-H+HCOONa]^-$ = 369.0600	165.0191, 135.0454	-
22	11.5	3'-Hydroxygenistein (level 2)	286.0482	$C_{15}H_{10}O_6$	$[M-H]^-$ = **285.0415** $[M-H+HCOONa]^-$ = 353.0285	n.d.	151.0042, 133.0296
23	12.9	2'-hydroxypseudobaptigenin (level 4)	314.007	$C_{16}H_{10}O_7$	$[M-H]^-$ = **313.0392** $[M-H+HCOONa]^-$ = 381.0317	n.d.	-
24	13.3	Barpisoflavone (level 4)	300.0634	$C_{16}H_{12}O_6$	$[M-H]^-$ = **299.0567** $[M-H+HCOONa]^-$ = 367.0440	n.d.	-
25	14.7	4',7-dihydroxy-2',5-dimethoxyisoflavone (level 4)	314.0795	$C_{17}H_{14}O_7$	$[M-H]^-$ = **313.0725** No adducts detected	298,0486	-
26	12.6	Phloretin (level 1)	274.0841	$C_{15}H_{14}O_5$	$[M-H]^-$ = **301.0723** No adducts detected	n.d.	-

a) *Annotation Confidence level:* levels of confidence in reported metabolite identification (Appendix 2) [17].
b) *m/z* values for the base peak are given in bold type

6. REFERENCES

[1] Newman, D. J.; Cragg, G. M., (**2007**). Natural products as sources of new drugs over the last 25 years. *J. Nat. Prod.*, 70 (3), pp 461-77.

[2] Koparde, A. A.; Doijad, R. C.; Magdum, C. S., (**2019**). Natural Products in Drug Discovery. In Perveen, S.; Al-Taweel, A. (First Edition) *Pharmacognosy-Medicinal Plants* (pp 1-20). IntechOpen.

[3] Ganzera, M.; Sturm, S. (**2018**). Recent advances on HPLC/MS in medicinal plant analysis an update covering 2011-2016*. J. Pharm. Biomed. Anal*., 147, pp 211-233.

[4] Natividad, S.; Soler, C.; Soriano, J. M.; Mañes, J., (**2010**). Occurrence of Aflatoxins in Tigernuts and Their Beverages Commercialized in Spain. *J. Agric. Food Chem.,* 58, pp 2609-2612.

[5] Ijarotimi, O.S.; et al., (**2018**). Chemical Compositions, Functional Properties, Antioxidative Activities, and Glycaemic Indices of Raw and Fermented Tigernut Tubers (*Cyperus Esculentus Lativum*) Flour. *J. Food Biochem.,* pp. 1-14.

[6] Zommara, M.; Imaizumi, K., (**2017**). In Vitro Antioxidant of Chufa Tubers (*Cyperus esculentus L*.) Extracts in Liposome Peroxidation Systems. *J. Sus. Agric. Sci.* 43 (2), pp 69-76.

[7] Adejuyitan, J. A., (**2011**). Tigernut Processing: Its Food uses and Health Benefits. *Am. J. Food Tech.*, 6, pp 197-201.

[8] Zommara, M.; Imaizumi, K., (**2017**). Antiatherogenic Effect of Tiger Nut Tubers (Cyperus esculentus L.) Supplemented Diet in Apolipoprotein E Knockout Mice. *J. Sus. Agric. Sci*. 43 (4), pp 197-204.

[9] Rodríguez Pérez, C., (**2016**). *An Analytical Approach: From Natural Sources of Bioactive Compounds to Nutraceuticals and Functional Foods. Development of Extraction, Characterization and Bioactive Evaluation Strategies*. Thesis to aim for International Doctor in Chemistry. University of Granada, Spain

[10] Malachová, A.; et al., (**2018**). Advanced LC-MS-based methods to study the co-occurrenceand metabolization of multiple mycotoxins in cereals and cereal-based food. *Anal. Bioanal. Chem*., 410 (3), pp 801-825.

[11] Siuzdak, G. (**2004**). An Introduction to Mass Spectrometry Ionization: An Excerpt from The Expanding Role of Mass Spectrometry in Biotechnology, 2nd ed.; MCC Press: San Diego, 2005. *JALA: Journal of the Association for Laboratory Automation*, 9 (2), 50-63.

[12] Bruins, A. P., (**1998**). Mechanistic Aspects of Electrospray ionization. *J. Chromatogr. A*, 794, pp 345-357.

[13] Parcher, Jon F.; et al., (**2018**). In-source collision-induced dissociation (IS-CID): Applications, issues and structure elucidation with single-stage mass analysers. *Drug Test Anal.*, 10, pp 28-36.

[14] Crellin, K. C.; Sible, E.; Van Antwerp, J., (**2003**). Quantification and confirmation of identity of analytes in various matrices with in-source collision-induced dissociation on a single quadrupole mass spectrometer. *Int. J. Mass Spectrom*., 46 (5), pp 281-488.

[15] Venisse, N.; et al., (**2003**). A general unknown screening procedure for drugs and toxic compounds in serum using liquid chromatography-electrospray-single quadrupole mass spectrometry. *J. Anal. Toxicol*., 27 (1), pp 7-14

[16] Domingo-Almenara, Xavier; Montenegro-Burke, J. R.; Benton, H. P.; Siuzdak, G., (**2018**). Annotation: A Computational Solution for Streamlining Metabolomics Analysis. *Anal. Chem.*, 90(1), pp 489-499.

[17] Scrimpe-Rutledge, A. C.; Codreanu, S. G.; Sherrod S. D.; McLean J. A., (**2016**). Untargeted Metabolomics Strategies-Challenges and Emerging Directions. *J. Am. Soc. Mass Spectrometry*, 27, pp 1897-1905.

[18] M. Pathak; J. N. Govil (**2016**). Recent Progress in Medicinal Plants, Vol, 43: Phytotherapeutics II. (First Edition) Studium Press LLC (pp 566). (USA)

[19] Domon, B., Costello, C.E., (**1988**). A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. *Glycoconjugate J*. 5, pp 397-409.

[20] Dzhavakhiya, V. G.; Zinovyeva, S. V.; et al., (**2007**). Immune response: Phenylpropanoids and lignin. In Dyakov, Yu T., Dzhavakhiya, V. G., Korpela, T. (Elsevier) *Comprehensive and Molecular Phytopathology* (pp 265-314). Elsevier Science.

[21] Oladele, A. k.; Adobowale, J. O.; Bamidele, O. P., (**2017**). Phenolic profile and antioxidant of brown and yellow varieties of Tiger Nut (*Cyperus esculentus L*.). *Nigerian Food J.,*pp 1-9

[22] Corradini, E.; et al., (**2011**). Flavonoids: Chemical Properties and Analytical Methodologies of Identification and Quantification in Food and Plants. *Nat. Prod. Res.*, 25 (5), pp 469-495.

[23] Elshamy, A. I. and et al., (**2017**). Phenolic Constituents, Anti-inflammatory and Antidiabetic Activities of *Cyperus laevigatus* L. *Phamacogn. J.*, 9 (6), pp 828-833.

[24] Vitale, D.; et al., (**2013**). Isoflavones: estrogenic activity, biological effect and bioavailability. *J. Drug. Metab. Pharmacokinet*., 38, pp 15-25.

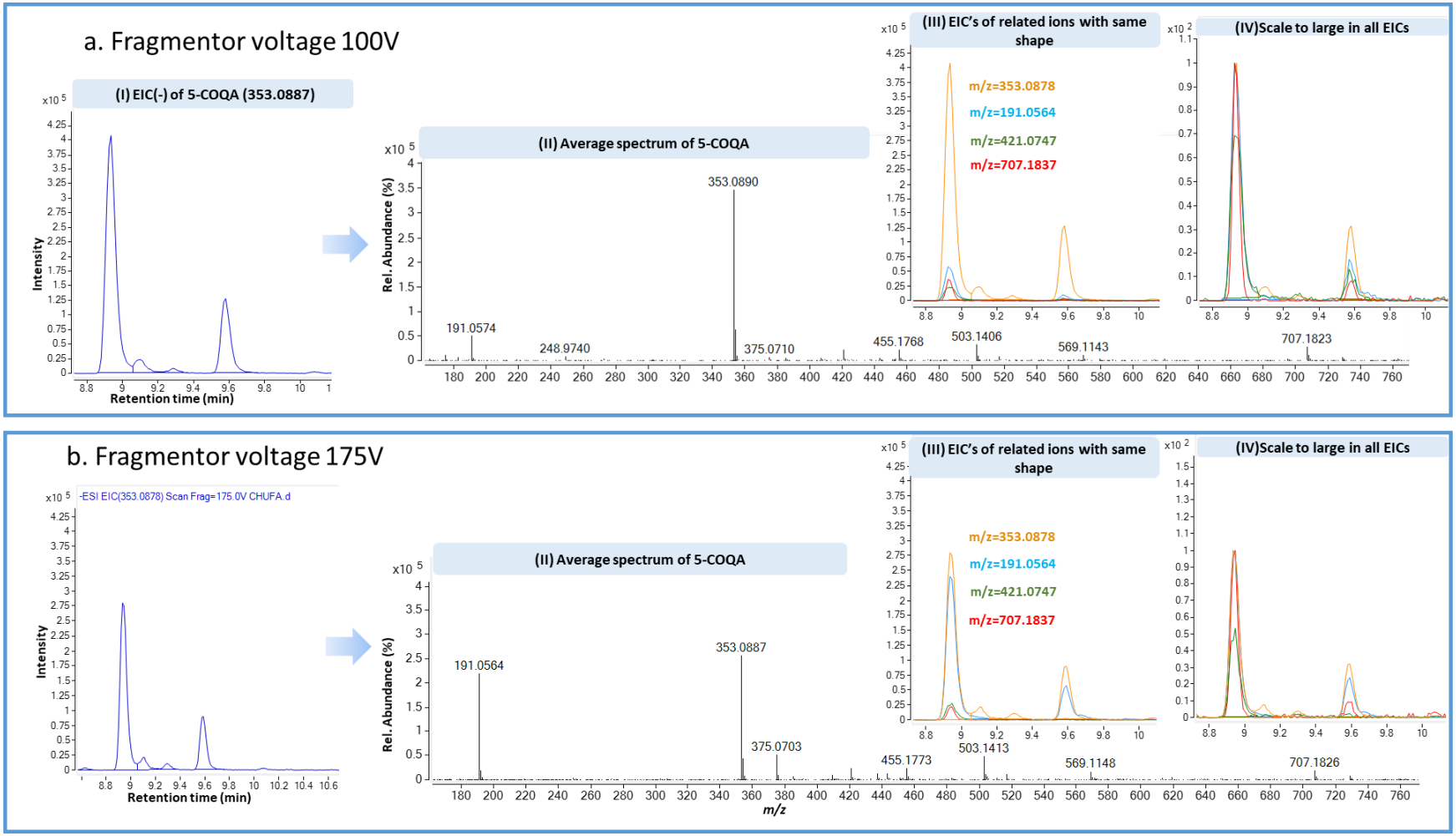
[25] Schimdt, J., (**2016**). Negative ion electrospray high-resolution tandem mass spectrometry of polyphenols. *J. Mass Spectrom.*, 51, pp 33-43.

[26] Pinheiro, P. F.; Justino, G. C., (**2012**). Structural Analysis of Flavonoids and Related Compounds - A Review of Spectroscopic Applications. In Venketeshwer Rao (First Edition) *Phytochemicals - A Global Perspective of Their Role in Nutrition and Health* (pp 33-44). Rijeka, Croatia: IntechOpen

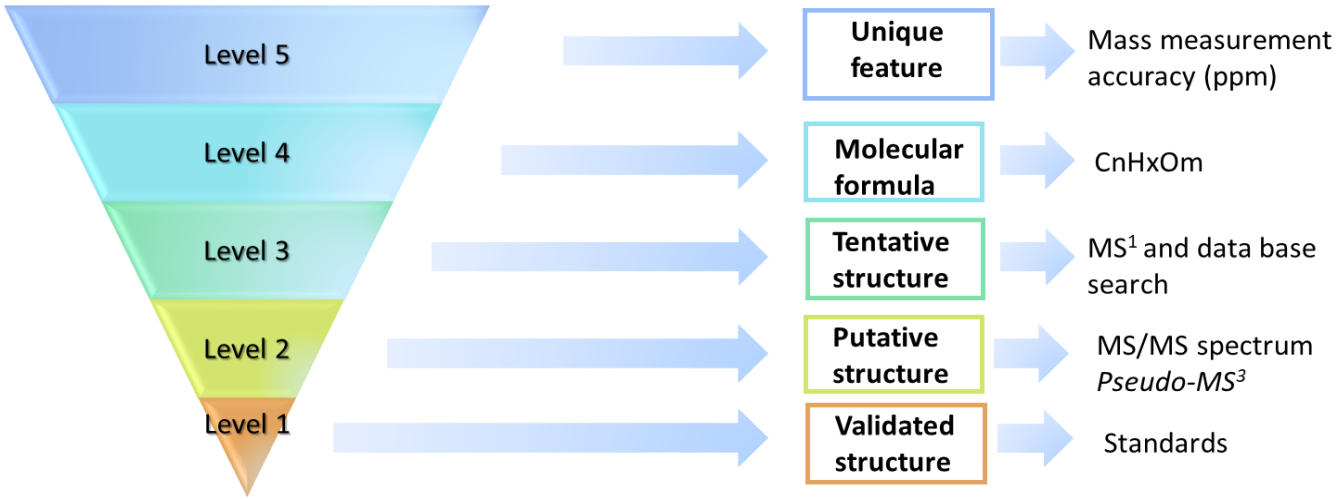
[27] Demarque, Daniel P.; et al., (**2016**). Fragmentation reactions using electrospray mass spectrometry: an important tool for the structural elucidation and characterization of synthetic and natural products. *Nat. Prod. Rep*., 33, pp 367-524.

[28] Jenkins, T. J.; Nguyen, J. C. D.; Polglaze, K. E.; Bertrand, P. B., (**2016**). Influence of Tryptophan and Serotonin on Mood and Cognition with a Possible Role of the Gut-Brain Axis*. Nutrients*, 8 (1), pp 56.

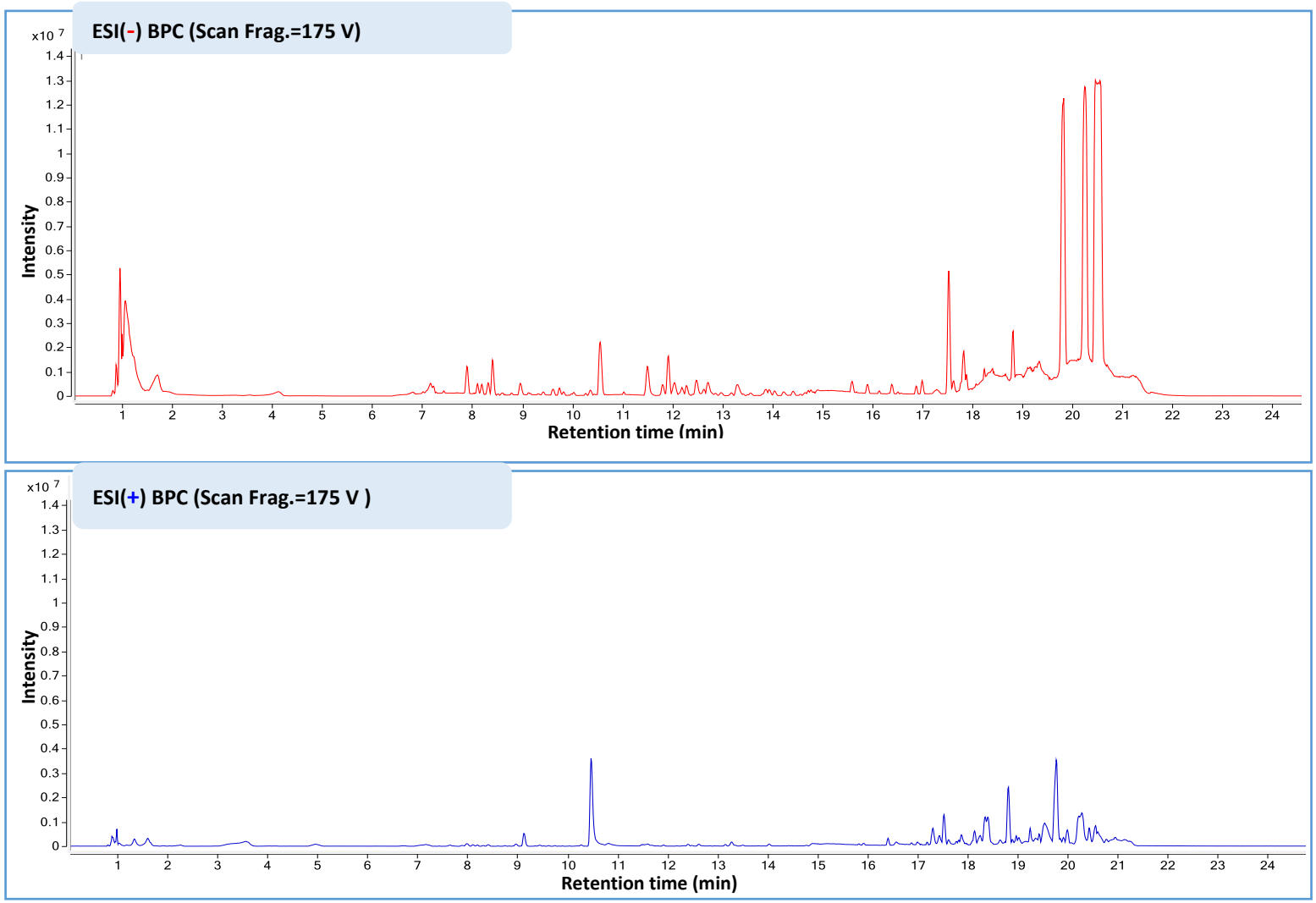
[29] Yu, P.; Hegeman, A. D.; Cohen, J. D., (**2014**). A facile means for the identification of indolic compounds from plant tissues. *Plant J.*, 6, pp 1065-75.



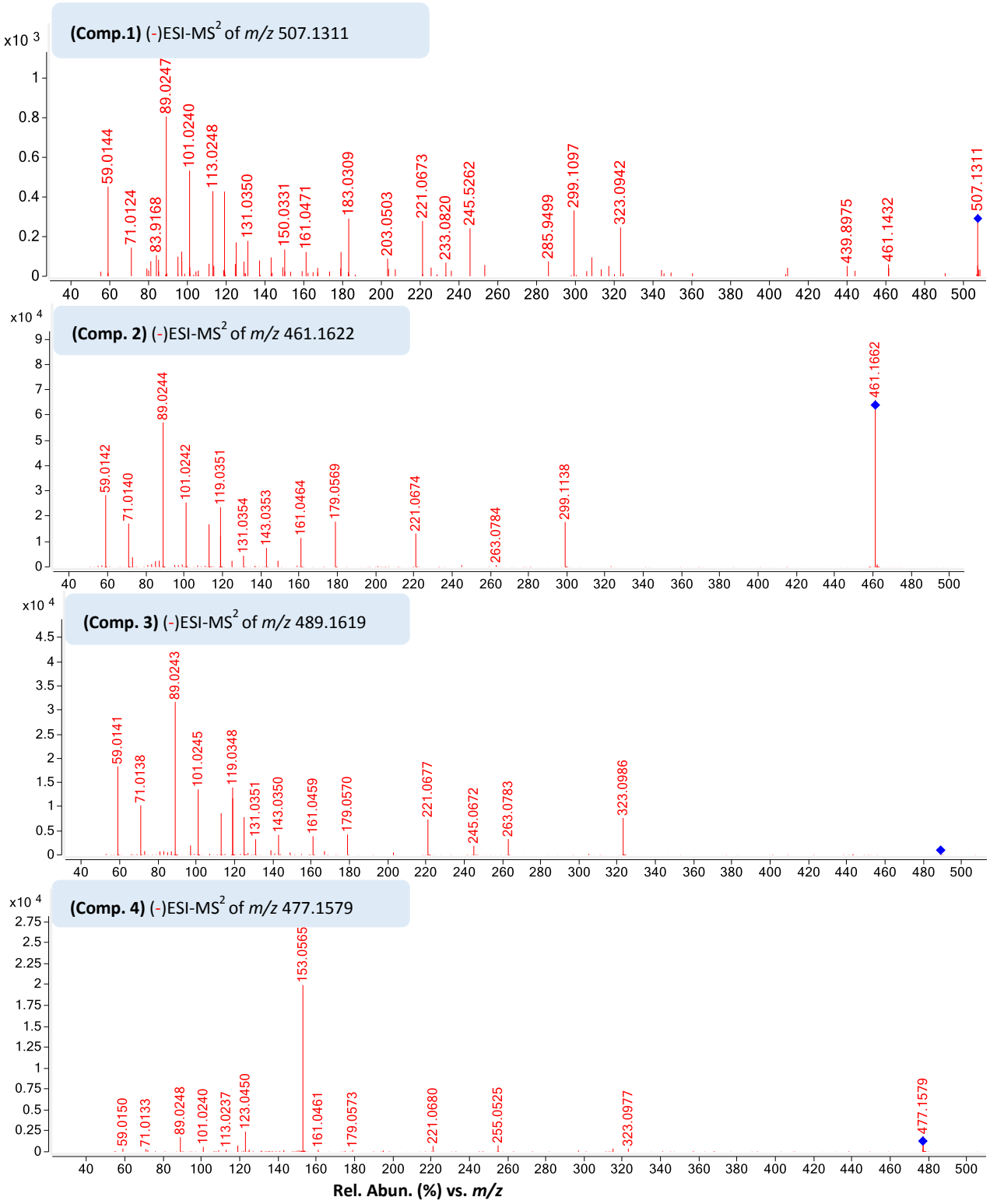
Appendix 1. Peak identification steps, where it is important to rely on retention time and the peak shape of all *m/z* that are suspictious to be fragments of the same molecule.



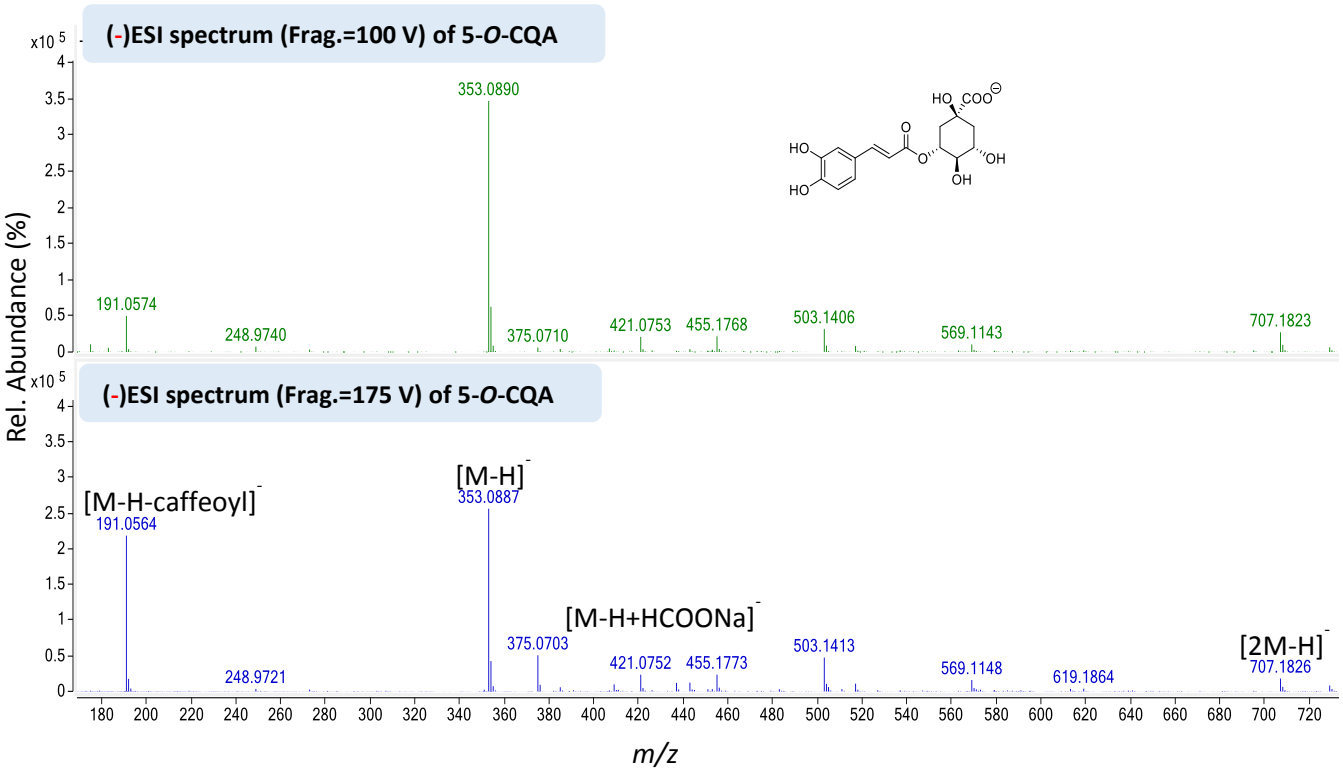
Appendix 2. *Proposed workflow for metabolite identification confidence using mass spectrometry [17]*



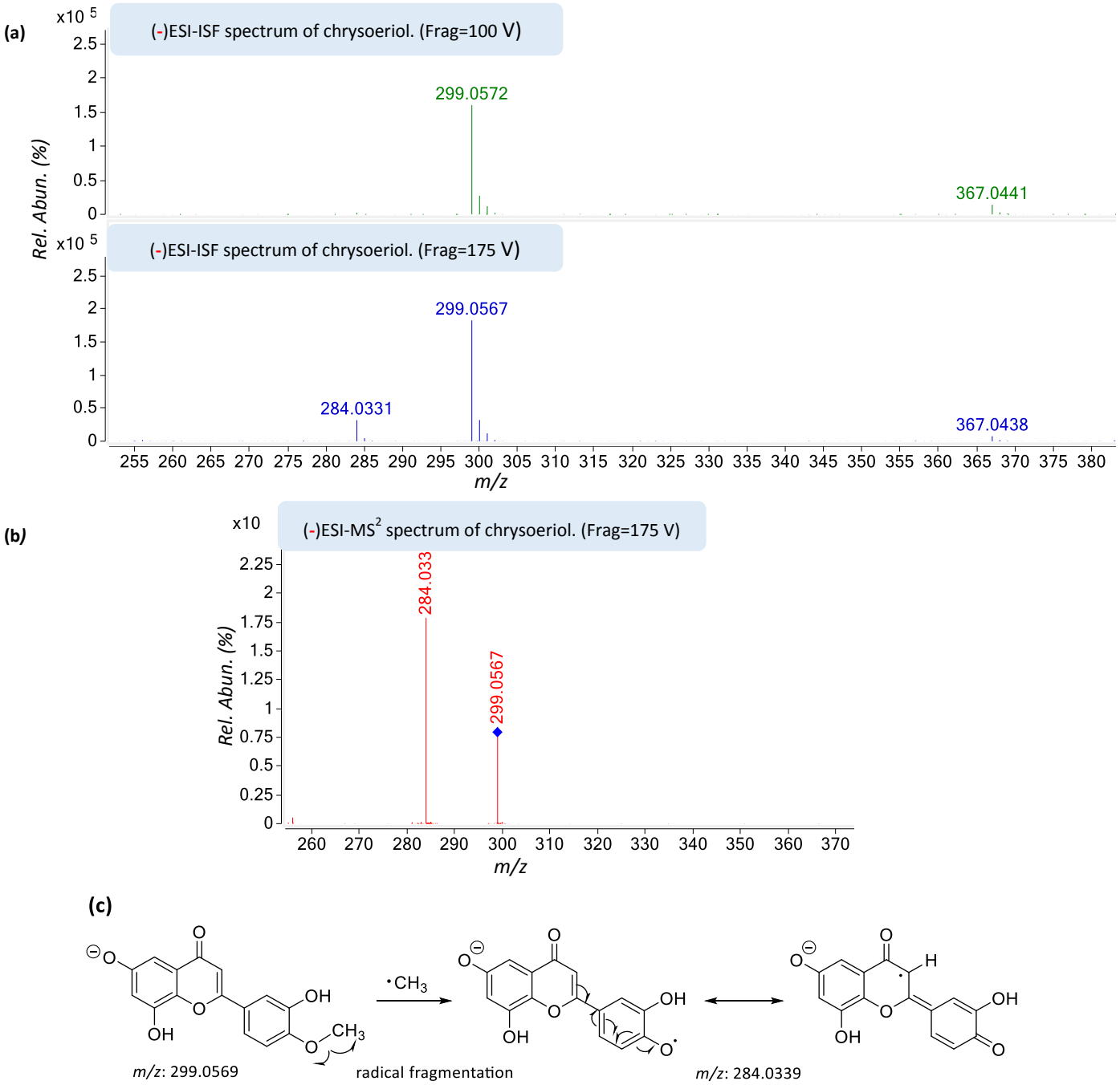
***Appendix 3*.** Typical Base Peak Chromatograms (BPC) of analysed methanolic extracts of tigernuts obtained under both negative and positive electron spray ionization (ESI^-/$ESI^{+)}$ conditions.



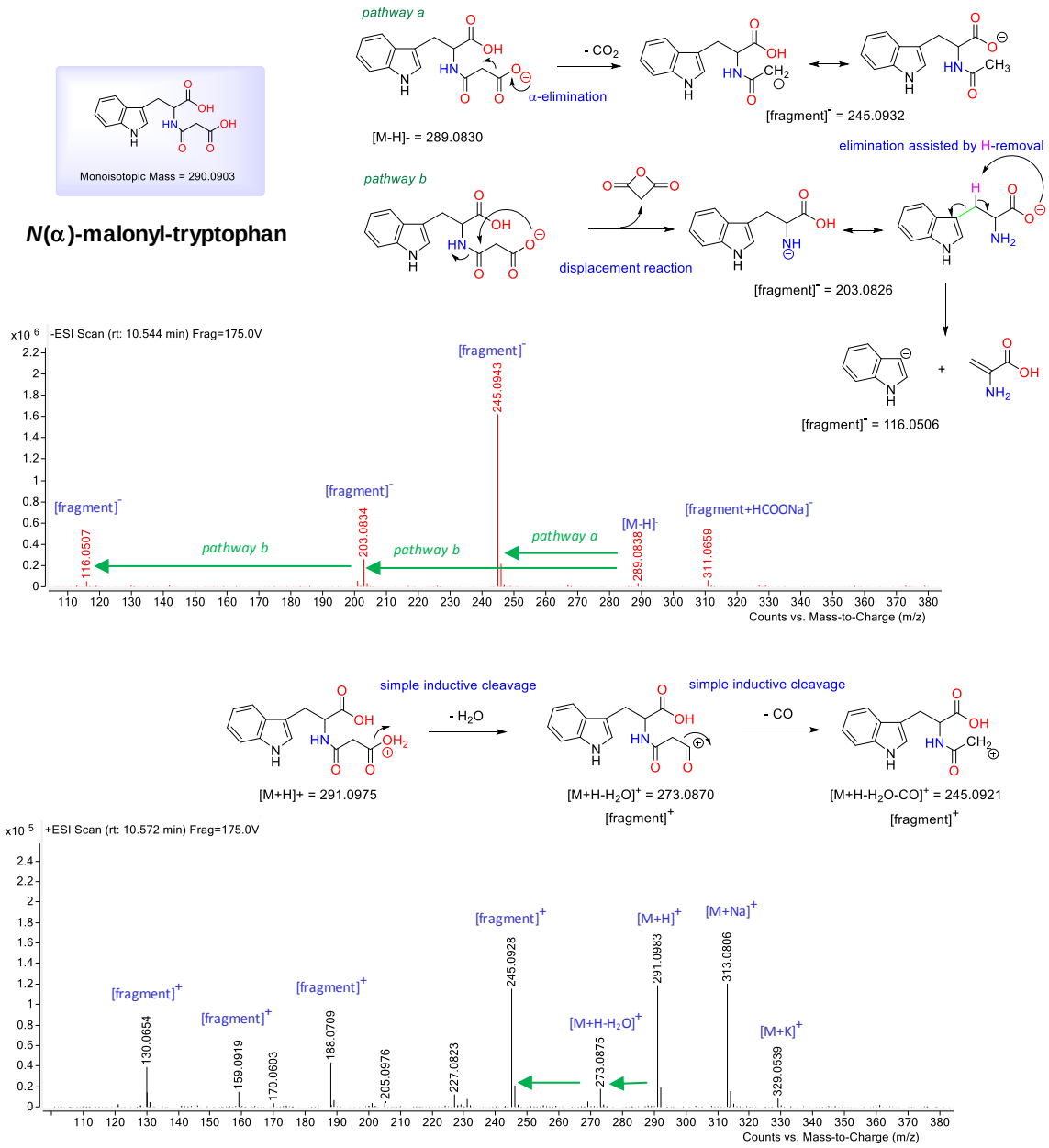
Appendix 4. MS/MS spectra of the different phenylethanoid glycosides (**1**-**4**) in which it is possible to see their fragments that share in common.



Appendix 5. In-source fragmentation spectrum of 5-O-Caffeoylquinic acid (5-*O*-CQA)



Appendix 6. **(a)** ISF spectra of chrysoeriol. **(b)** MS/MS spectrum of chrysoeriol. **(c)** Proposed fragmentation mechanism.



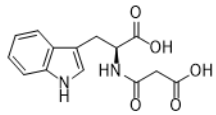
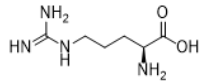
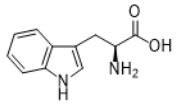
Appendix 7. Proposed fragmentation pathways of N(α)-malonyl-tryptophan.

Appendix Table 1.

Compounds	***Regression equation (y=ax+b)***	***RSD (RF)***	***LOQ***	***Concentration (ppm)***	***Concentration (µg/g)***
(+)-catechin (**11**)	y=7.21E+05x - 2.70E+03	1.21%	7.14E-08	0.43	4.27
chlorogenic acid (**7**)	y=2.59E+05x-1.09E+04	4.40%	4.38E-02	5.11	51.11
naringenin (**12**)	y=2.78E+06x+8.86E+05	22.57%	3.74E-03	0.036	0.36
Phloretin (**26**)	y=3.89E+06x+1.99E+06	42.85%	9.19E-04	0.0060	0.060
caffeic acid (**8**)	y=4.70E+05x-6.79E+04	18.97%	3.71E-02	0.31	3.05
ferulic acid (**10**)	y=1.13E+05x-1.76E+04	21.00%	7.79E-02	1.62	16.17
Chrysin (**18**)	y=5.00E+06x+7.71E+05	9.11%	4.15E-03	0.042	0.42
luteolin	y=4.07E+05x+6.16E+04	12.45%	3.30E-03	3.064	30.64
Quercetin (**13**)	y=9.44E+05x+6.04E+04	6.81%	1.72E-02	0.067	0.67

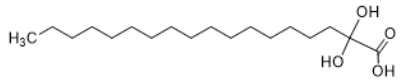
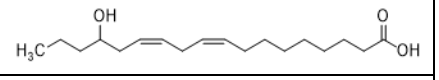
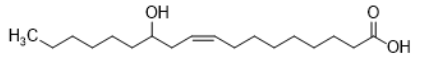
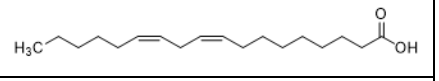
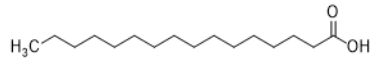
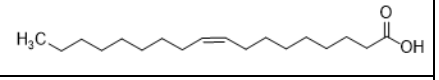
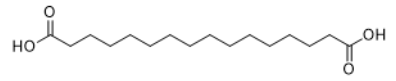
Appendix Table 1. *Quantification with standards****.***

Appendix Table 3. Amino acids and derivatives

					ESI positive mode		
No. Peak	***RT (min)***	***Common name (Annotation Confidence)***	***Monoisotopic Mass***	***Molecular Formula***	***Selected ions***	***ISF***	***Structures***
Phenylethanoid glycosides (PhGs)							
27	10.5	*N(α)*-malonyl-tryptophan (level 2)	290,0903	$C_{14}H_{14}N_2O_5$	$[M+H]^+$ = 291,0983 $[M+Na]^+$ = 313,0801	273.0873, 245.0924, 227.0811, **188.0711***, 159.0913, 130.651	
28	0.9	L-arginine (level 2)	174,1117	$C_6H_{14}N_4O_2$	$[M+H]^+$ = **175.1192*** No adducts detected	158.0924,130,0975, 116.0706	
29	7.9	L-tryptophan (level 2)	204,0899	$C_{11}H_{12}N_2O_2$	$[M+H]^+$ = 205.0974	**188.0694***,146.0589,132.0793,159.0896	

*Base peak at that *m/z*

Appendix Table 4. Fatty acids

					ESI negative mode		
No. Peak	***RT (min)***	***Common name (Annotation Confidence)***	***Monoisotopic Mass***	***Molecular Formula***	***Selected ions***	***ISF***	***Structures***
Fatty acids							
30	17.5	Dihydroxyoct adecanoic acid	316.2613	$C_{18}H_{36}O_4$	$[M-H]^-$ = **315.2556** $[2M-H]^-$ = 631.5160 $[M-H-HCOONa]^-$ = 383.2423	n.d	
31	17.8	Hydroxylinoleic acid	296.2351	$C_{18}H_{32}O_3$	$[M-H]^-$ = **295.2294** $[M-H-HCOONa]^-$ = 383.2424	n.d	
32	18.2	Ricinoleic acid	298.2507	$C_{18}H_{34}O_3$	$[M-H]^-$ = **297.2451** $[M-H-HCOONa]^-$ = 365.2319 $[2M-H]^-$ = 595.4946	n.d	
33	19.8	Linoleic acid	280.2402	$C_{18}H_{32}O_2$	$[M-H]^-$ = **279.2361** $[2M-H]^-$ = 559.4745	n.d	
34	20.1	Palmitic acid	256.2402	$C_{16}H_{32}O_2$	$[M-H]^-$ = **255.2366** $[2M-H]^-$ = 511.4754 $[M-H-HCOONa]^-$ = 323.2213	n.d	
35	20.4	Oleic acid	282.2558	$C_{18}H_{34}O_2$	$[M-H]^-$ = **281.2547** $[2M-H]^-$ =563.5046	n.d	
36	16.7 and 16.9	Octadecanedioic acid	314.2457	$C_{18}H_{34}O_4$	$[M-H]^-$ =**313.2395** $[M-H-HCOONa]^-$ = 381.2261	n.d	
37	17.5	Dihydroxystearic acid	316.26135	$C_{18}H_{34}O_5$	$[M-H]^-$ = **329.2343** $[M-H-HCOONa]^-$ = 397.2209	n.d